Volume 48 Number 1 March 2024 ISSN 0350-5596

Informatica

An International Journal of Computing and Informatics

Editorial Boards

Informatica is a journal primarily covering intelligent systems in the European computer science, informatics and cognitive community; scientific and educational as well as technical, commercial and industrial. Its basic aim is to enhance communications between different European structures on the basis of equal rights and international refereeing. It publishes scientific papers accepted by at least two referees outside the author's country. In addition, it contains information about conferences, opinions, critical examinations of existing publications and news. Finally, major practical achievements and innovations in the computer and information industry are presented through commercial publications as well as through independent evaluations.

Editing and refereeing are distributed. Each editor from the Editorial Board can conduct the refereeing process by appointing two new referees or referees from the Board of Referees or Editorial Board. Referees should not be from the author's country. If new referees are appointed, their names will appear in the list of referees. Each paper bears the name of the editor who appointed the referees. Each editor can propose new members for the Editorial Board or referees. Editors and referees inactive for a longer period can be automatically replaced. Changes in the Editorial Board are confirmed by the Executive Editors.

The coordination necessary is made through the Executive Editors who examine the reviews, sort the accepted articles and maintain appropriate international distribution. The Executive Board is appointed by the Society Informatika. Informatica is partially supported by the Slovenian Ministry of Higher Education, Science and Technology.

Each author is guaranteed to receive the reviews of his article. When accepted, publication in Informatica is guaranteed in less than one year after the Executive Editors receive the corrected version of the article.

An Overview on Robot Process Automation: Advancements, Design Standards, its Application, and Limitations

Rajkumar Palaniappan
College of Engineering, Department of Mechatronics Engineering, University of Technology Bahrain, Salmabad, Kingdom of Bahrain
E-mail: r.palaniappan@utb.edu.bh

Overview Paper

Keywords: RPA, design standard, hyperautomation, cognitive, cloud, machine learning

Received: July 23, 2023

In a variety of areas, including healthcare, banking, and manufacturing, repetitive and rule-based processes are automated using robotic process automation (RPA), a fast developing technology. An overview of RPA's, its uses, limitation, and applications are given in this paper. RPA can lower costs, increase process speed, accuracy, and efficiency, and free up staff to concentrate on jobs of higher value. RPA is frequently used for tasks including data input, billing, and customer care. RPA can't, however, execute activities that call for human judgment, decision-making, or creativity, for instance. The adoption of RPA also needs a sizable initial investment and continual maintenance. This paper also touches on a few RPA-related ethical issues, like employment displacement and data privacy. While RPA has a great deal of promise to alter sectors, its deployment can only be successful if its limitations and ethical implications are carefully considered.

Povzetek: Narejen je obsežen pregled RPA, avtomatizacije robotskih procesov, kot hitro razvijajoče se tehnologije, vključno z njenimi aplikacijami, standardi, omejitvami in uporabami.

1 Introduction

Robot Process Automation (RPA) serves as a valuable instrument that empowers organizations with the ability to automate tiresome tasks typically executed by humans. Software robots within the realm of RPA exhibit astonishing skills which enable them to flawlessly emulate human actions including button clicks, data inputting as well as system navigation[1]. These tireless bots work diligently throughout all hours of the day aiming for accuracy in task completion with both speed and efficiency. Undoubtedly there lies in abundance a tremendous potential embedded within RPA; potential capable of vastly reducing costs firsthand whilst simultaneously enhancing overall efficiency along with raising operational standards by eliminating any scope for error or faults thereby Allowing relieved allocation for human resource towards activities holding relatively higher value [2]. However partial success shall accompany those who embark on such implementation endeavors unto successful execution of RPA merely via reservations shall fail upon laying strong emphasis upon meticulous planning commencing from identification associated with suitable processes up for automation followed closely thereby alongside selection relevant tools associated with processes enlisted prior plus provision put forth ensuring suitable governance in addition to appropriate oversight [3]. The rapid advancements in RPA technology have opened a world of limitless possibilities for organizations. With each progressive improvement in RPA capabilities, businesses are compelled to actively pursue and explore new use cases to harness the potential of automation. This drive to expand their automation capabilities is motivated by the highly sought-after rewards that come with adopting and implementing these innovative technological advancements.

One of the primary reasons organizations actively pursue RPA is the promise of increased efficiency and productivity. RPA technology enables the automation of repetitive and time-consuming tasks, freeing up valuable human resources to focus on more strategic and value-added activities. By automating mundane and rule-based processes, businesses can significantly reduce manual errors and increase the speed at which tasks are completed. This increased efficiency leads to cost savings and improves operational performance.

In addition to efficiency gains, RPA technology offers the potential for scalability and agility. As organizations grow and evolve, RPA can be easily scaled to accommodate increased workloads and changing business requirements [4], [5]. RPA solutions can be quickly deployed and integrated with existing systems and applications, allowing businesses to adapt to market demands and seize new opportunities more rapidly. This flexibility and agility give organizations a competitive edge in dynamic and fast-paced industries.

Furthermore, RPA can enhance accuracy and compliance within organizations. By automating processes, businesses can ensure consistent adherence to established rules and standards. RPA software can be programmed to follow predetermined workflows and perform tasks with precision, minimizing the risk of human error [6]. This level of accuracy is particularly beneficial in industries

that require strict compliance with regulations and standards, such as finance, healthcare, and legal sectors.
Moreover, RPA technology enables organizations to gain valuable insights from data. By automating data collection, processing, and analysis, businesses can extract meaningful information and make data-driven decisions more efficiently. RPA can integrate with other analytics and business intelligence tools, allowing organizations to uncover patterns, trends, and correlations that can inform strategic planning and optimize business processes.
Overall, the possibilities presented by RPA technology are virtually limitless. As organizations witness the transformative power of RPA in streamlining operations, reducing costs, improving accuracy, and enabling data-driven decision-making, they are driven to actively pursue and explore new use cases. By expanding their automation capabilities, businesses can attain the highly sought-after rewards that come with adopting and leveraging these innovative technological advancements.

2 Advancement in RPA

Robot Process Automation (RPA) is a rapidly developing discipline, and RPA technology has made several strides in recent years. The following are some significant RPA advancements:

2.1 Machine learning-based RPA:

Robotic process automation (RPA) is a sophisticated technique that uses machine learning techniques to enable intelligent automation. Traditional RPA uses pre-established rules and workflows to automate repetitive tasks, but RPA based on machine learning can gain knowledge from the past and continuously improve [7].
Machine learning algorithms are trained on big datasets to identify patterns and generate predictions in RPA that is machine learning-based. Then, these algorithms can be included into RPA systems to automate difficult activities that ordinarily call for human involvement. When processing invoices, for instance, machine learning-based RPA can be used to train the system to recognize various invoice types, extract pertinent data, and verify it against existing data [8].
Compared to standard RPA, machine learning-based RPA has a number of benefits, including better accuracy, more efficiency, and the capacity to manage unstructured data. Additionally, it is better able to manage exceptions and adapt to new circumstances. For the algorithms to be trained, machine learning-based RPA needs a lot of high-quality data, which can be difficult in particular sectors. RPA based on machine learning is still in its infancy, and there are worries about data privacy, prejudice, and the possibility that automation may replace human workers [9].
In general, machine learning-based RPA is a promising advancement in the field of automation, and it has a wide range of possible uses. However, for it to be implemented successfully, much thought must be given to its constraints and moral consequences.

2.2 Cognitive RPA

Natural language processing (NLP), machine learning (ML), and computer vision are examples of cognitive technologies that are combined with traditional robotic process automation to create cognitive RPA (Robotic Process Automation). Intuitive and adaptable automation systems that can carry out complicated and varied tasks that were previously challenging or impossible to automate are what cognitive RPA aims to build [10],[11].
To comprehend natural language inputs and extract meaning from unstructured data sources like emails, documents, and social media postings, cognitive RPA systems use NLP. In order to learn from data and increase the precision of forecasts and decision-making, machine learning algorithms are used. To recognize and analyze visual data, such as pictures and movies, computer vision is employed [12].
Cognitive RPA systems may automate a larger range of jobs and give users more individualized and knowledgeable responses by combining these technologies. For instance, a cognitive RPA system can be used to identify and prioritize customer support requests, classify, and extract data from invoices automatically, or analyze social media data to find trends and sentiment [13].
Overall, cognitive RPA has the potential to revolutionize a variety of industries by enhancing productivity, accuracy, and efficiency while lowering costs and mistakes.

2.3 Hyperautomation

Hyperautomation is a method of automation that combines several technologies in order to automate as much of a business process as possible. These technologies include Robotic Process Automation (RPA), Artificial Intelligence (AI), Machine Learning (ML), Natural Language Processing (NLP), Process Mining, and other cutting-edge technologies [14].

While hyperautomation is based on the same principles as regular automation, it employs a larger variety of tools and technology to take a more thorough approach to automation. Hyperautomation is locating and automating every routine, repetitive operation in a business process, including those that are usually done by people [15].
The creation of an end-to-end automation process that can be managed and optimized with little to no human involvement is the main objective of hyperautomation. This strategy can increase productivity, decrease manual errors, and free up resources for firms to concentrate on more difficult jobs that need for human involvement [16].
Many company activities, including customer service, finance and accounting, HR, and supply chain management, can benefit from hyperautomation. Businesses can embrace hyperautomation to streamline their operations, lower expenses, and boost productivity, which will ultimately raise their success and competitiveness in their particular marketplaces.

2.4 Cloud-based RPA

Cloud based Robotic Process Automation (RPA) systems present an opportunity for users to conveniently access RPA tools and services through the internet [17]. The utilization of cloud based RPA has the potential to bring about cost reductions and enhanced scalability by granting users a flexible and scalable platform for automating their processes. Tasks like data entry, data processing, and report generation can be efficiently automated through the implementation of cloud-based RPA solutions [18]. Typically hosted in the cloud and reachable via a web browser, cloud-based RPA systems. Users can access the automation platform from any location with an internet connection and do not need to install any software on their local computers. Scalability is one of the main advantages of cloud-based RPA. Without having to spend money on extra hardware or infrastructure, businesses can quickly scale their RPA deployment up or down based on their needs. This makes it the perfect option for companies whose customers' demands fluctuate seasonally or who need to expand quickly. The ability of cloud-based RPA to be linked with other cloud-based services and applications, including Customer relationship management (CRM), Enterprise resource planning (ERP), and other business applications, is another benefit. Without having to manually move data between separate apps, this enables enterprises to automate end-to-end business operations that include several different systems [19]. In general, firms wanting to automate their business operations can gain a lot from cloud-based RPA. It offers a scalable and adaptable solution that can boost productivity, lower errors, and free up human resources for work with higher added value.

2.5 Process mining

Process mining is a valuable tool used to analyze event logs or data from IT systems with the aim of uncovering, monitoring, and enhancing business processes [20].

By automating vital tasks such as process discovery, analysis, and optimization, process mining plays a significant role in driving operational efficiency. Moreover. It enables the identification of bottlenecks and inefficiencies within processes. Ultimately leading to improved levels of automation and effectiveness [21].

Process mining is a method for examining corporate procedures to find inefficiencies, bottlenecks, and potential areas for change. In order to see and evaluate the process flow, data must first be extracted from multiple sources, including databases, transaction logs, and other systems, using data mining algorithms. Robotic process automation, or RPA, automates routine, rule-based processes using software robots [22].

RPA may assist firms in locating opportunities for automation and putting automated solutions into place fast and effectively when used in conjunction with process mining. Process mining can be used to identify repetitive and time-consuming operations that can be automated using RPA. RPA can be used to automate these tasks and lower the risk of errors, for instance, if process mining indicates that a certain process comprises numerous data entry tasks that are prone to errors. Additionally, the process mining tool may be updated with the data gathered throughout the RPA installation, offering insights into the efficiency of the automation and pointing out potential areas for development. Organizations can gradually enhance their operations with the aid of this continuous feedback loop, leading to increased productivity and efficiency [23].

Overall, process mining and RPA are a potent duo that can aid businesses in streamlining their procedures, lowering errors, and boosting productivity. Organizations may increase the accuracy and speed of their operations while allowing staff to focus on more value-added duties by automating repetitive jobs.

3 Design standards in RPA

RPA design guidelines are essential for developing dependable, effective, and maintainable automation solutions. Here are some of the main design standards for RPA that have been suggested in various studies:

3.1 Modularity

Modularity in RPA refers to breaking down automation processes into smaller, reusable components. This approach simplifies the maintenance and updating of the automation solution over time. By dividing the tasks and functionalities into independent modules, each module can be developed and tested separately, reducing complexity. Modularity also promotes reusability, as modules can be used in different automation processes or projects, saving time and effort in development. Additionally, it enhances scalability, allowing for the integration of new modules as automation needs grow or change without affecting the entire system. Modularity in RPA fosters collaboration among developers, enabling parallel development and facilitating code reusability and version control. Overall, modularity provides a structured and efficient approach to building automation solutions, making them easier to maintain, update, and scale, while promoting collaboration and reusability among developers [24].

3.2 Error handling

In the context of RPA, it is important to design solutions that can handle mistakes graciously. This means incorporating mechanisms to detect errors and take appropriate actions to resolve them. RPA solutions should have robust error detection capabilities, allowing them to identify errors at different stages of the automation process, such as data validation, system errors, or unexpected behavior. Once an error is detected, the solution should be programmed to respond in a way that fixes the error or minimizes its impact. This may involve retrying actions, alternative approaches, data validation, or escalation to human operators for resolution. Additionally, logging and reporting mechanisms should be implemented to capture and track errors for analysis and improvement of the automation process. By designing RPA solutions to

handle mistakes graciously, organizations can ensure the reliability and resilience of their automated processes, reducing manual intervention and improving overall efficiency [25].

3.3 Security

In order to protect sensitive information and prevent unauthorized access, RPA systems should adhere to recognized security standards and best practices. This includes implementing strong access controls, such as user authentication and role-based access, to ensure that only authorized individuals can access the RPA systems and the data they handle. Encryption should be applied to data at rest and in transit to maintain confidentiality and prevent unauthorized interception or access. Regular updates and patching should be performed to address any discovered vulnerabilities, and logging and monitoring mechanisms should be in place to detect and respond to suspicious activities. Conducting regular security audits and assessments helps ensure ongoing compliance and identifies areas for improvement. By following these security measures, RPA systems can maintain the integrity and security of sensitive information, mitigating the risk of unauthorized access and data breaches [26].

3.4 Scalability

Scalability is a fundamental requirement for RPA solutions, as they need to handle growing workloads as businesses expand. Scalability in RPA refers to the ability of the solution to accommodate increased demands without sacrificing performance or efficiency. To achieve scalability, RPA solutions should be designed with flexibility and modularity, allowing for the addition of new components or replication of existing ones to handle larger workloads. Dynamic resource allocation and intelligent load balancing mechanisms are essential to optimize resource utilization. Additionally, the architecture and workflows of RPA solutions should be designed with scalability in mind, considering factors like data storage, transfer capabilities, and compatibility with different systems and platforms. By prioritizing scalability, RPA solutions can effectively meet the automation needs of growing businesses while maintaining optimal performance [27].

3.5 Documentation

Documentation is a critical aspect of well-designed RPA solutions. It involves detailing the goals, inputs, outputs, and dependencies of automation processes. Clear and detailed documentation provides a shared understanding among stakeholders, including developers, business users, and management, about the purpose and expected outcomes of the automation. It also facilitates troubleshooting and debugging by providing a comprehensive view of the data and information flows. Documenting dependencies helps identify any external systems or integrations that the RPA solution relies on, ensuring that all necessary components are in place for successful execution. Furthermore, well-documented RPA solutions serve as a reference for future enhancements, updates, and maintenance, enabling efficient collaboration and reducing dependence on specific individuals. They also support compliance and audit requirements by providing an audit trail of the automation processes. Overall, thorough documentation is essential for clarity, transparency, and maintainability of RPA solutions [24].

Following these five design guidelines will enable RPA developers to produce dependable, effective, and maintainable automation systems that will aid enterprises in achieving their automation objectives and streamlining their business procedures.

4 Applications of RPA

Robot Process Automation (RPA) is used in a wide variety of industries. Here are a few of the main RPA applications:

4.1 Finance

The finance industry is embracing the use of Robotic Process Automation (RPA) to automate crucial tasks like claims processing, invoice processing, and account reconciliation. By automating repetitive tasks that were typically performed by humans RPA can significantly reduce errors and enhance efficiency [28]. In addition. RPA aids in ensuring consistent execution of processes in accordance with regulatory requirements. Thereby helping improve compliance [29], [30].

4.2 Healthcare

RPA is now being used by the healthcare sector to automate a number of tasks, including patient scheduling, claims processing, and disease detection [31]. RPA implementation can have significant benefits, such as cost savings and improved patient outcomes, by automating tasks that were previously handled by human employees [32]. By ensuring that procedures are consistently followed in accordance with legal standards, RPA can be very helpful in improving compliance [33].

4.3 Manufacturing

In the manufacturing sector, robotic process automation (RPA) is being utilized more and more to automate crucial processes including inventory management, supply chain management, and quality control [34]. RPA has the power to significantly save expenses and increase overall efficiency by automating repetitive processes that were previously performed by people. Moreover. By routinely ensuring adherence to well defined procedures and rigorous quality standards, RPA may also dramatically improve product quality [35].

4.4 Retail

The retail sector uses RPA now for automating several processes, including order processing, inventory control, and customer support [36]. By automating processes that were previously done by people, this technology can

dramatically save costs while also improving the entire customer experience. RPA may also be very helpful in increasing compliance because it makes sure that jobs are completed consistently and in accordance with legal norms [37].

4.5 Human resources

Currently, RPA is being utilized in the human resources industry to automate a range of tasks, such as recruiting new employees, handling payroll, and administering benefits [38]. By automating repetitive operations that have historically been done by people, RPA has the potential to decrease errors and boost efficiency [39]. RPA can assist compliance initiatives by ensuring that tasks are routinely carried out in accordance with legal requirements.

5 Challenges and limitations of RPA

While Robot Process Automation (RPA) provides many advantages, there are also several difficulties and restrictions that come with using it. Here are some of the main obstacles and restrictions facing RPA:

5.1 Complexity of processes

RPA is most effective when automating repetitive and rule-based processes. However, it may have limitations when it comes to managing complex operations that require analysis and decision-making [40], [41]. In various industries, there are procedures that involve intricate workflows, data analysis, strategic planning, and subjective decision-making, which may be beyond the capabilities of RPA. These complexities can limit the use of RPA in certain sectors, as the software typically follows predefined rules and lacks the cognitive abilities to interpret unstructured data or make nuanced judgments.
While RPA may not be well-suited for handling complex processes, it can still be valuable in augmenting human work and streamlining specific aspects of these operations. By automating repetitive and well-defined subtasks within a larger complex process, RPA can free up human workers to focus on the more intricate and value-added aspects that require critical thinking and creativity. It is important for businesses to carefully evaluate their processes and determine where RPA can provide the most value based on the complexity and nature of the tasks involved. In some cases, a combination of RPA with other technologies such as AI or machine learning may be necessary to tackle the challenges posed by complex operations and achieve a more comprehensive automation strategy.

5.2 Integration with legacy systems

The utilization of antiquated software poses challenges for companies integrating RPA into their existing systems. Many businesses still rely on legacy systems that lack integration capabilities, making the integration process time-consuming and expensive [42]. Custom development work is often required to establish communication between the RPA platform and the legacy software, adding complexity and resource requirements. Additionally, extensive testing and validation are necessary to ensure smooth interaction and avoid disruptions. These factors can constrain RPA adoption as businesses assess the cost and benefits of integration and may need to prioritize system modernization efforts alongside RPA implementation. Collaboration with experienced professionals, leveraging pre-built connectors or APIs, and conducting system assessments can help streamline the integration process and overcome challenges associated with antiquated software. By carefully considering trade-offs and employing strategies to address integration complexities, companies can successfully integrate RPA into their existing systems and harness the benefits of automation.

5.3 Security concerns

RPA systems that have access to confidential data pose significant security challenges for companies. To prevent data breaches and other security issues, organizations must prioritize RPA system security [43], [44]. This involves implementing stringent access controls to ensure that only authorized personnel can interact with the RPA system and access sensitive data. Measures such as multifactor authentication, role-based access controls, and encryption of data at rest and in transit help safeguard the confidentiality and integrity of critical information. Regular security assessments and vulnerability testing should be conducted to identify and address potential weaknesses or vulnerabilities. Additionally, organizations should prioritize data privacy and compliance with relevant regulations, implementing measures to anonymize or pseudonymize data and ensuring adherence to privacy requirements. By taking a proactive approach to RPA system security, businesses can mitigate the risks associated with data breaches and protect sensitive information.

5.4 Economic concerns

The implementation and upkeep of RPA systems can be costly, which may present a challenge for smaller organizations with limited financial resources. RPA implementation involves various expenses, including software licensing, infrastructure setup, process analysis, development, and testing. Ongoing maintenance and support also add to the overall cost. For smaller organizations, these expenses can be prohibitive and act as a barrier to adopting RPA fully.

However, it is worth noting that the cost of RPA has been decreasing over time as technology becomes more accessible and competitive. Cloud-based RPA solutions, for example, provide a more cost-effective option by eliminating the need for extensive infrastructure investment. Additionally, partnering with RPA service providers or consultants can offer expertise and support without the need for significant upfront investments. These approaches can help smaller organizations

overcome the cost limitations associated with RPA and still benefit from its potential to enhance operational efficiency.

5.5 Ethical concerns

The deployment of RPA raises ethical concerns regarding its impact on jobs. Analysts predict that widespread adoption of RPA could lead to job losses, especially in sectors heavily reliant on human labor [44]. To address these concerns, businesses must carefully consider the ethical implications of RPA and develop plans to mitigate negative effects on employment. This may involve strategies such as retraining and upskilling affected employees, job rotation, and fostering transparent communication with workers. Additionally, businesses should consider broader societal impacts and invest in initiatives that support job creation and skill development, ensuring a balanced approach to automation that prioritizes the well-being of employees and society.

6 Overcoming RPA challenges and limitations

RPA (Robotic Process Automation) can pose a number of obstacles and challenges. The following advice will help you get through them:

- Select the appropriate processes for automation: Not all processes can be automated. Find the procedures that are repetitive, rule-based, time-consuming, high volume, and have a lot of room for automation. This will aid in deciding which processes should be automated initially.
- Choose the appropriate RPA tool: It's critical to pick an RPA tool that can handle the complexity of the operations you wish to automate. Pick a tool that gives solid support and training, is scalable, and is easy to use.
- • A strong business case is necessary to justify the investment in RPA. It should clearly demonstrate the return on investment and list the benefits, such as improved accuracy, cost savings, and higher productivity.
- • Include business users, IT, and management in the RPA implementation process as well as any other interested parties. The likelihood that everyone will agree and the implementation will proceed well will increase as a result.
- RPA implementation might cause a lot of organizational change. Create a change management strategy to make sure that every employee is aware of the changes and equipped to cope with them.
- Test the RPA implementation carefully to make sure it functions as expected and has no unwanted effects.
- Monitoring and optimization are necessary to make sure that RPA continues to provide the anticipated benefits. Making adjustments to the procedures or the RPA implementation itself may be necessary for this.

In summary, robotic process automation (RPA) is a technique that automates routine, rule-based, and high-volume processes using software robots. RPA technology provides a number of advantages, such as greater productivity and cost savings, but it also has certain drawbacks, such as the need for organized data and the absence of decision-making capabilities. It's critical to adhere to design principles, such as modular architecture and error handling, to ensure the efficacy of RPA installations. RPA technology can be used in a variety of sectors and processes, but success requires careful design, testing, and continuing optimization.

7 Conclusion

Process mining, cognitive automation, machine learning, and hyperautomation are a few RPA breakthroughs that are revolutionizing business processes. By increasing automation and rising productivity, this swiftly evolving technology has the potential to completely transform how company activities are carried out. RPA has the ability to boost output, reduce costs, and free up employees' time for more difficult work. Utilizing cloud technology enhances RPA's capabilities further, making it a stronger tool for enterprises overall. To realize its full potential, RPA must overcome a variety of challenges, such as those related to security, cost, and how it interacts with legacy systems, ethics, and process complexity. Businesses must keep up with the latest RPA developments and investigate how they might be used in their particular industry.

Acknowledgement

I want to express my sincere gratitude to the University of Technology Bahrain's Mechatronics Engineering department for providing the necessary library information services for carrying out this research.

References

[1] W. M. P. van der Aalst, M. Bichler, and A. Heinzl, “Robotic Process Automation,” *Business & Information Systems Engineering*, vol. 60, no. 4, pp. 269–272, 2018. https://doi.org/10.1007/s12599-018-0542-4

[2] T. Taulli, *The Robotic Process Automation Handbook*. Berkeley, CA: Apress, 2020. https://doi.org/10.1007/978-1-4842-5729-6

[3] A. Aguirre Santiago and Rodriguez, “Automation of a Business Process Using Robotic Process Automation (RPA): A Case Study,” in *Applied Computer Sciences in Engineering*, E. R. and V.-R. J. L. and F.-E. R. Figueroa-García Juan Carlos and López-Santana, Ed., Cham: Springer International Publishing, 2017, pp. 65–71. https://doi.org/10.1007/978-3-319-66963-2_7

[4] L. P. Willcocks, J. Hindle, M. Stanton, and J. Smith, “A Strategic Approach to Robotic Process Automation,” in *Maximizing Value with Automation and Digital Transformation: A Realist’s Guide*, L. P. Willcocks, J. Hindle, M.

Stanton, and J. Smith, Eds., Cham: Springer Nature Switzerland, 2024, pp. 21–29. https://doi.org/10.1007/978-3-031-46569-7_2

[5] L. P. Willcocks, J. Hindle, M. Stanton, and J. Smith, "RPA in Financial Services," in *Maximizing Value with Automation and Digital Transformation: A Realist's Guide*, L. P. Willcocks, J. Hindle, M. Stanton, and J. Smith, Eds., Cham: Springer Nature Switzerland, 2024, pp. 37–41. https://doi.org/10.1007/978-3-031-46569-7_4

[6] W. Wang *et al.*, "Cellular nucleus image-based smarter microscope system for single cell analysis," *Biosens Bioelectron*, vol. 250, p. 116052, 2024. https://doi.org/10.1016/j.bios.2024.116052

[7] R. S. Bavaresco *et al.*, "Machine learning-based automation of accounting services: An exploratory case study," *International Journal of Accounting Information Systems*, vol. 49, p. 100618, 2023. https://doi.org/10.1016/j.accinf.2023.100618

[8] T. F. and M. S. and K. S. Tyagi Amit Kumar and Fernandez, "Intelligent Automation Systems at the Core of Industry 4.0," in *Intelligent Systems Design and Applications*, V. and G. N. and S. P. and K. A. and M. A. Abraham Ajith and Piuri, Ed., Cham: Springer International Publishing, 2021, pp. 1–18. https://doi.org/10.1007/978-3-030-71187-0_1

[9] W. and M. J. M. and B. M. Lestari Nur Indah and Hussain, "A Survey of Trendy Financial Sector Applications of Machine and Deep Learning," in *Application of Big Data, Blockchain, and Internet of Things for Education Informatization*, F. Jan Mian Ahmad and Khan, Ed., Cham: Springer Nature Switzerland, 2023, pp. 619–633. https://doi.org/10.1007/978-3-031-23944-1_68

[10] P. Martins, F. Sá, F. Morgado, and C. Cunha, "Using machine learning for cognitive Robotic Process Automation (RPA)," in *2020 15th Iberian Conference on Information Systems and Technologies (CISTI)*, 2020, pp. 1–6. https://doi.org/10.23919/CISTI49556.2020.9140440

[11] F. Karim, "Cloud Computing-Based M-Government," *Informatica*, vol. 46, no. 5, Mar. 2022. https://doi.org/10.31449/inf.v46i5.3879

[12] C. Engel, P. Ebel, and J. M. Leimeister, "Cognitive automation," *Electronic Markets*, vol. 32, no. 1, pp. 339–350, 2022. https://doi.org/10.1007/s12525-021-00519-7

[13] A. Masood Adnan and Hashmi, "Cognitive Robotics Process Automation: Automate This!," in *Cognitive Computing Recipes: Artificial Intelligence Solutions Using Microsoft Cognitive Services and TensorFlow*, Berkeley, CA: Apress, 2019, pp. 225–287. https://doi.org/10.1007/978-1-4842-4106-6_5

[14] A. Haleem, M. Javaid, R. P. Singh, S. Rab, and R. Suman, "Hyperautomation for the enhancement of automation in industries," *Sensors International*, vol. 2, p. 100124, 2021. https://doi.org/10.1016/j.sintl.2021.100124

[15] S. Madakam, R. M. Holmukhe, and R. K. Revulagadda, "The Next Generation Intelligent Automation: Hyperautomation," *Journal of Information Systems and Technology Management*, vol. 19, Mar. 2022. https://doi.org/ 10.4301/S1807-1775202219009

[16] S. and W. J. Liermann Volker and Li, "Hyperautomation (Automated Decision-Making as Part of RPA)," in *The Digital Journey of Banking and Insurance, Volume II: Digitalization and Machine Learning*, C. Liermann Volker and Stegmann, Ed., Cham: Springer International Publishing, 2021, pp. 277–293. https://doi.org/10.1007/978-3-030-78829-2_16

[17] S. Karn and R. Kotecha, "RPA-based Implementation of IoT," *SSRN Electronic Journal*, 2021. https://doi.org/10.2139/ssrn.3868873

[18] O. Tembhurne, S. Milmile, G. R. Pathak, A. O. Thakare, and A. Thakare, "An Orchestrator:A Cloud-Based Shared-Memory Multi-User Architecture for Robotic Process Automation," *International Journal of Open Source Software and Processes*, vol. 13, no. 1, pp. 1–17, Sep. 2022. https://doi.org/10.4018/ijossp.308792

[19] A. Maalla, "Development Prospect and Application Feasibility Analysis of Robotic Process Automation," in *2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, 2019, pp. 2714–2717. https://doi.org/10.1109/IAEAC47372.2019.8997983

[20] A. Rautenburger Lars and Liebl, "Process Mining," in *The Digital Journey of Banking and Insurance, Volume II: Digitalization and Machine Learning*, C. Liermann Volker and Stegmann, Ed., Cham: Springer International Publishing, 2021, pp. 259–275. https://doi.org/10.1007/978-3-030-78829-2_15

[21] V. Leno, A. Polyvyanyy, M. Dumas, M. La Rosa, and F. M. Maggi, "Robotic Process Mining: Vision and Challenges," *Business & Information Systems Engineering*, vol. 63, no. 3, pp. 301–314, 2021. https://doi.org/10.1007/s12599-020-00641-4

[22] A. H. M. and K. W. and L. S. J. J. and R. M. and W. M. T. Egger Andreas and ter Hofstede, "Bot Log Mining: Using Logs from Robotic Process Automation for Process Mining," in *Conceptual Modeling*, U. and K. G. and L. S. W. and M. H. C. Dobbie Gillian and Frank, Ed., Cham: Springer International Publishing, 2020, pp. 51–61. https://doi.org/10.1007/978-3-030-62522-1_4

[23] J. Rinderle-Ma Stefanie and Mangler, "Process Automation and Process Mining in

Manufacturing," in *Business Process Management*, M. T. and V. L. A. and R. M. Polyvyanyy Artem and Wynn, Ed., Cham: Springer International Publishing, 2021, pp. 3–14. https://doi.org/10.1007/978-3-030-85469-0_1

[24] L.-V. Herm, C. Janiesch, A. Helm, F. Imgrund, A. Hofmann, and A. Winkelmann, "A framework for implementing robotic process automation projects," *Information Systems and e-Business Management*, vol. 21, no. 1, pp. 1–35, Mar. 2023. https://doi.org/10.1007/s10257-022-00553-8

[25] S.-H. Kim, "Development of Evaluation Criteria for Robotic Process Automation (RPA) Solution Selection," *Electronics (Basel)*, vol. 12, no. 4, 2023. https://doi.org/10.3390/electronics12040986

[26] Dahlia Fernandez and Aini Aman, "The Challenges of Implementing Robotic Process Automation in Global Business Services," *International Journal of Business and Society*, vol. 22, no. 3, pp. 1269–1282, Dec. 2021. https://doi.org/10.33736/ijbs.4301.2021

[27] L. A. Cooper, D. K. Holderness, T. L. Sorensen, and D. A. Wood, "Robotic Process Automation in Public Accounting," *Accounting Horizons*, vol. 33, no. 4, pp. 15–35, Dec. 2019. https://doi.org/10.2308/acch-52466

[28] D. B. D. H. Chris Lamberton, "Impact of Robotics, RPA and AI on the Insurance Industry: Challenges and Opportunities," *Journal of Financial Perspectives*, vol. 4, no. 1, pp. 1–13, 2017. https://ssrn.com/abstract=3079495

[29] F. C. M. Ortiz and C. J. Costa, "RPA in Finance: supporting portfolio management : Applying a software robot in a portfolio optimization problem," in *2020 15th Iberian Conference on Information Systems and Technologies (CISTI)*, 2020, pp. 1–6. https://doi.org/10.23919/CISTI49556.2020.9141155

[30] Z. Liang and Y. Liang, "A Study of Identification of Corporate Financial Fraud Using Neural Network Algorithms in an Information-based Environment," *Informatica*, vol. 47, no. 9, Dec. 2023. https://doi.org/10.31449/inf.v47i9.5220

[31] S. Benedict, "IoT-Enabled Remote Monitoring Techniques for Healthcare Applications -- An Overview," *Informatica*, vol. 46, no. 2, Jun. 2022. https://doi.org/10.31449/inf.v46i2.3912

[32] Ö. Doğuç, "Robotic Process Automation (RPA) Applications in COVID-19," in *Management Strategies to Survive in a Competitive Environment: How to Improve Company Performance* , S. Dincer Hasan and Yüksel, Ed., Cham: Springer International Publishing, 2021, pp. 233–247. https://doi.org/10.1007/978-3-030-72288-3_16

[33] B. A J, A. N, and I. S, "Robotic Process Automation (RPA): A software bot for healthcare sector," in *2023 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE)*, 2023, pp. 685–689. https://doi.org/10.1109/IITCEE57236.2023.10090996

[34] R. Kavitha, "Hyperautomation-Beyond RPA: : Leveraging Automation to Transform the Manufacturing Industries," in *2023 International Conference on Computer Communication and Informatics (ICCCI)*, 2023, pp. 1–5. https://doi.org/10.1109/ICCCI56745.2023.10128636

[35] F. A. Lievano-Martínez, J. D. Fernández-Ledesma, D. Burgos, J. W. Branch-Bedoya, and J. A. Jimenez-Builes, "Intelligent Process Automation: An Application in Manufacturing Industry," *Sustainability*, vol. 14, no. 14, 2022. https://doi.org/10.3390/su14148804

[36] N. Lazareva, K. Karasevskis, A. Girjatovcs, and O. Kuznecova, "Business Process Automation in Retail," in *2022 63rd International Scientific Conference on Information Technology and Management Science of Riga Technical University (ITMS)*, 2022, pp. 1–5. https://doi.org/10.1109/ITMS56974.2022.9937096

[37] S. Dey and A. Das, "Robotic process automation: assessment of the technology for transformation of business processes," *International Journal of Business Process Integration and Management*, vol. 9, no. 3, p. 220, 2019. https://doi.org/10.1504/IJBPIM.2019.100927

[38] A. Najjar, B. Amro, and M. Macedo, "An Intelligent Decision Support System For Recruitment: Resumes Screening And Applicants Ranking," *Informatica*, vol. 45, no. 4, Dec. 2021. https://doi.org/10.31449/inf.v45i4.3356

[39] S. A. Mohamed, M. A. Mahmoud, M. N. Mahdi, and S. A. Mostafa, "Improving Efficiency and Effectiveness of Robotic Process Automation in Human Resource Management," *Sustainability*, vol. 14, no. 7, 2022. https://doi.org/10.3390/su14073920

[40] L. P. Willcocks, J. Hindle, M. Stanton, and J. Smith, "Robotic Process Automation: Just Add Imagination," in *Maximizing Value with Automation and Digital Transformation: A Realist's Guide*, L. P. Willcocks, J. Hindle, M. Stanton, and J. Smith, Eds., Cham: Springer Nature Switzerland, 2024, pp. 31–35. https://doi.org/10.1007/978-3-031-46569-7_3

[41] N. Nalgozhina and R. Uskenbayeva, "Automating hybrid business processes with RPA: optimizing warehouse management," *Procedia Comput Sci*, vol. 231, pp. 391–396, 2024. https://doi.org/10.1016/j.procs.2023.12.223

[42] A. Puthuruthy and B. Marath, "Leveraging Artificial Intelligence for Developing Future Intelligent ERP Systems," *International Journal of Intelligent Systems and Applications in*

Engineering, vol. 12, no. 8s, pp. 623–629, Dec. 2023. https://ijisae.org/index.php/IJISAE/article/view/4235

[43] M. Eulerich, N. Waddoups, M. Wagener, and D. A. Wood, "The Dark Side of Robotic Process Automation (RPA): Understanding Risks and Challenges with RPA," *Accounting Horizons*, pp. 1–10, Sep. 2023. https://doi.org/10.2308/HORIZONS-2022-019

[44] A. Y. A. B. Ahmad, "Ethical implications of artificial intelligence in accounting: A framework for responsible ai adoption in multinational corporations in Jordan," *International Journal of Data and Network Science*, vol. 8, no. 1, pp. 401–414, 2024. https://doi.org/10.5267/j.ijdns.2023.9.014

Application of Agent-Based Modelling in Learning Process

Natasha Stojkovikj, Limonka Koceva Lazarova, Aleksandra Stojanova, Marija Miteva, Biljana Zlatanovska and Mirjana Kocaleva
Faculty of computer science, Goce Delcev University, Stip, Republic of North Macedonia
Email: natasa.maksimova@ugd.edu.mk, limonka.lazarova@ugd.edu.mk, aleksandra.stojanova@ugd.edu.mk, marija.miteva@ugd.edu.mk, biljana.zlatanovska@ugd.edu.mk, mirjana.kocaleva@ugd.edu.mk

Keywords: agent-based modelling, simulation, education

Received: March 23, 2022

With advances in information and communication technologies and rapid computing and technological progress, modelling, and simulation of real problems, has become the most important teaching and learning method in educational process. Representing and explaining processes through simulations can enable students to easier understand these processes and discover the essential properties of a system. In many situations, in learning different subjects it is not possible to experiment with real objects to find the right solutions, therefore modelling and simulation can be used to build models that represent the real systems. Agent-based modelling (ABM) is a powerful simulation modelling technique, that can be easily incorporated in learning and teaching processes. Agent based modelling (ABM) is a relatively new method compared to system dynamics and discrete event modelling. In ABM a system is modelled as a collection of autonomous decision-making entities called agents, that can interact among each other's. In this paper, the agent-based modelling simulation is considered as a tool in educational process for learning and teaching different subjects. Anylogic software is used for some simulation examples of agent-based modelling that can be used in educational process.

Povzetek: Programska oprema Anylogic se uporablja za nekatere simulacijske primere modeliranja na podlagi agentov, ki se lahko uporabljajo v izobraževalnem procesu.

1 Introduction

Simulation is an imitation of the operations of real-world processes or a system over time. The behaviour of the system over time is studied by developing a simulation model. It is common for models to be represented as a set of assumptions about the system itself. These assumptions are expressed through mathematical logical or symbolic relations between entities that are objects of interest in the system. Simulations can be used at the design's stage before the system is built but also on existing systems to determine whether potential changes will have an impact on system performance. Therefore, simulations can be used either as a tool to predict how changes will affect an existing system or as a tool to predict the performance of a new system under a different set of conditions. Sometimes the evolving model can be solved mathematically. Then, the solution can be obtained by using differential equations, probability theory, algebraic models, or other mathematical techniques. This solution usually consists of one or more numeric parameters called the system performance measure. However, most real systems cannot be solved mathematically because they are too complex. In this system, numerically based simulations can be used to imitate system's behaviour over time. The simulation data are collected through system monitoring. The data generated during the simulation is used to evaluate the performance of the system [1-4]. For simulation, there are three main methods: Discrete Event Simulation, System Dynamics, and Agent Based Simulation.

The Discrete Event Simulation (DES) models a process as a series of discrete events. Each event occurs at a particular point in time and represent a change of state in the system. Discrete event simulations are entity driven. The entities represent customers arriving in the system for servicing [5,6].

System dynamics (SD) are used to understand the nonlinear behaviour of complex systems over time. In SD three main objects are considered; stocks, flows and delays. Stocks are basic stores of objects; flows define the movement of items between different stocks in the system and delays are the delay between the system measuring something and then acting upon that measurement [5,6].

Agent based simulation (ABS) is a relatively new method compared to system dynamics and discrete event modelling. With agent-based modelling, the entities known as agents must be identified and their behaviour defined. The agents may be people, cells, households, vehicles etc [5-10].

In this paper, agent-based modelling and simulations and their application in the educational process is considered. It is shown how the agent-based simulation methods can help for easily and better understanding the basics of some processes, that otherwise are difficult to imagine and understand.

Also, it is shown that this type of simulation can be used in education process, especially in learning math's

subjects for math students and students of computer science also for learning and teaching science for medical students, or subjects connected to business and organization sciences and students learning that problematics.

In the paper are described models which are implemented in AnyLogic Simulation and Modelling Software. We choose this simulation software because of is its availability and its simplicity for use because it is free simulation software originally intended for educational process.

First, we describe epidemiological model SEIR-D and its usefulness for learning and understanding to the medical students, math students and students of computer science. The next model considered, is a model of Hospital Emergency Department. This model is important for the business students, medical students, and math students. And at the end, we consider example from a real world, that also can be used for easily understanding of different subjects used for math students, computer science students and business students. Precisely, we describe model of market that starts selling a new consumer product.

With considering of these simple models, the students can easily understand the basic concepts of subjects connected to this simulation and can visually and dynamically present something that can be difficult to present or imagine otherwise.

2 Agent based modelling

There are many definitions from different experts about what agents in systems are, but all agree that agents are a software component that is autonomous and aims to act as a human agent (collect data, process data, and interpret data) [6].

Agent - based modelling (ABM) is one of the newer approaches in computer simulation. This type of simulation is mainly used to model complex system, and it is based on autonomous agents and their interactions.

Agent-based models (ABMs) are computational structures where system-level behaviour can be obtained by the behaviour of individual agents. ABM basically contains three elements: agents, an environment, and rules governing each agent's behaviour and its local interactions with other agents and with the environment.

Agents have their own characteristics, rules of making decision, ability to interact with other agents in the system and environment based on which they can change and adjust their behaviour [6, 7].

This method of modelling must identify active entities, or agents and their behaviour must be defined. Agents can be people, households, vehicles, equipment, products, or companies, more precisely anything, that is related to the system [7].

In recent years, ABM has been used in various branches of science, and the largest application they found in the social sciences. It is often used to simulate the phenomenon in economic and technical sciences. Until the appearance of these models, modelling of phenomena in society were most often reduced to a simplified presentation of social phenomena, and very often they were only verbal models.

In ABM models, which models social phenomena and processes; agents represent people, and through their mutual communication and rules of conduct are modelled social processes and social communication. The main assumption is that people and their social skills can be realistically modelled.

Agent-based modelling and simulations provide more realistic models and lead to new possibilities in modelling and simulations. Agents used in these simulation models originate from the fields of robotics and artificial intelligence. Today, ABM agents are not more related to the design and understanding of artificial intelligence. The basic application is in modelling of human social behaviour, social phenomena and individual decision making.

With the obvious benefits that agent-based modelling and simulations bring, the number of simulation models of different social behaviours are increased. Today, it can be done many micro simulations that could not be done before a few years.

We are using ABM for modelling different models that can help students to easily learn and teachers to better teach different natural, social and math subjects.

3 Agent based modelling in education process

Knowledge application in realistic situations has been shown to be verry important in the process in developing complex skills. Students can acquire high level of expertise in complex real problem-solving tasks if they have enough previous knowledge and enough practice. Practice can be obtained with facing real problems which correspond to a professional field. In educational programs, the opportunity to engage in real-life problem solving is very limited. These limitations make practice in real-life situations often inaccessible especially for novice learners. Therefore, simulations can often be used in education settings. In STEM (science, technology, engineering, and mathematics) education, modelling and simulation can be used to facilitate a deeper understanding of concepts and relationships between objects and problems, easily problem solving, and decision making. Agent based modelling and simulation can be used in medical education, where simulations are used to enhance diagnostics competence, technical skills for future doctors and nurses. Agent based simulation can also be used in other fields, such as teaching education, engineering, and management, also can be used by the students of economic, biology, political science [11 -18].

Some of the most used simulation software in education are: EcoBeaker, SimBio, NetLogo, MIMOSE, AgentScript, Swarm, JAS-mine and Anylogic.

NetLogo is a multi-agent programmable modelling environment for simulating natural and social phenomena.

It is especially well suited for modelling systems that are developing over the time. NetLogo allows sophisticated modelling and allows the experienced programmers to add their own Java extensions. This software is used by many hundreds of thousands of students, teachers, and researchers from whole world [19].

MIMOSE consists of a model description language and an experimental frame for simulation of the described models. The main purpose of MIMOSE simulation software was the development of a modelling language that considers special demands of modelling in social science, especially the description of nonlinear quantitative and qualitative relations, stochastic influences, birth and death processes, and micro and multilevel models [20].

EcoBeaker is an ecological simulation program. This program is designed primarily for education goal but can be used and for research models. EcoBeaker gives a two-dimensional computer world into which agents are placed and their behaviours are designed [21].

Swarm is a simulation software package for multi-agent simulation of complex systems developed at the Santa Fe Institute. It is made to be a useful tool for researchers and students in many disciplines. The basic architecture of Swarm is the simulation of collections of concurrently interacting agents: with this architecture, a large variety of agent-based models can be implemented [22].

SimBio is a simulation software for teaching biology. The software can be used for biological systems such as cardiac cells, epithelial cells, and pancreatic β cells. With this software can be simulated experiments in evolution, cell biology, genetics, and neurobiology. SimBio is written in Java, uses XML and can solve ordinary differential equations [23].

AgentScript is a minimalist agent- based modelling framework. This tool is based on NetLogo agents' semantics. Its goal is to promote an agent-oriented programming model in a deployable CoffeeScript /JavaScript implementation [24].

JAS-mine is a Java-based computational platform that features tools to support the development of large-scale, data-driven, discrete-event simulations. JAS-mine is specifically designed for both agent-based and microsimulation modelling, anticipating a convergence between the two approaches [25].

AnyLogic is a multimode simulation modelling tool developed by AnyLogic (formerly XJ Technologies). Supports simulation methodologies based on agents, discrete events, and system dynamics. AnyLogic is a cross-platform simulation software running on Windows, macOS and Linux.

AnyLogic is used to simulate markets and competition, healthcare, manufacturing, supply chain and logistics, retail, business processes, social and ecosystem dynamics, defense, asset management, pedestrian dynamics, and road traffic.

AnyLogic models can be based on any of the three methods in simulation modelling: discrete events, system dynamics or agent-based systems [26].

The comparison of different simulation software used in education, mentioned before, and their main features are given in Table 1 .

Table 1: Comparation of different simulation software used in education

	Operating system	**Programming language**	**User support and License**	**Model development effort**	**Models' scalability level**	**Subjects covered**	**User friendliness**
EcoBeaker	Windows and Mac	No programming skills required	CD with tutorial/ Proprietary, not free for use	Simple, easy	Small scale	Ecology, conservation biology, and evolution	high
SimBio	Windows and Mac.	Java	Tutorials, Interactive Chapters, Workbook Labs, Frequently Asked Questions / General Public Licence, free for use	Moderate	Medium scale	Ecology, Evolution, Env Science, Cell Biology. Genetics, Conservation, Biology, Physiology	Medium to high
NetLogo	Cross-platform: JVM, (difficult to install on Windows)	NetLogo	Documentation; FAQ; selected references; tutorials; third party extensions; defect list; mailing lists /General Public License, free for use	Simple, easy to moderate	Medium to high scale	Different natural and social sciences	medium

MIMOSE	Linux, Windows (difficult to install on Windows)	Java	Tutorial for installation and use/ Open sourced, free for use	Moderate	Small scale	Social science	poor
AgentScript	All OS with Browsers	Javascript, NetLogo	Tutorials, Example Models /Open source free for use, GPLv3 license	Simple, easy	Small scale	Primary for social sciences but usable for natural sciences too.	medium
Swarm	Cross-platform	Java; Objective-C	Wiki; tutorials; examples; documentation; FAQ; selected publications; mailing lists/ General Public License, free for use	Hard, Complex	Extreme scale	Primary for social sciences .	poor
JAS-mine	Cross-platform: JVM	Java	Tutorials, presentations, videos/ Eclipse plugin, free for use	Simple, easy to moderate	Medium scale	Social and natural sciences, primary social, discrete-event simulations, including agent-based and microsimulation models	medium
Anylogic	Linux, macOS, Windows	Java	Demos; training; online community; ask a question; online help; tutorials; consulting services/ Free Personal Learning Edition available	Moderate	High scale	Different natural and social sciences, discrete events, system dynamics or agent-based systems	medium to high

In continuation in this paper, we give examples of the use of agent-based modelling and simulations in different areas to facilitate the study and understanding of certain problematics. We used AnyLogic as a software for implementing these examples.

We chose AnyLogic because as mentioned above and as given in Table 1, AnyLogic compared to other tools has the best features in terms of ease of use, free to use, use in multiple areas, adaptability, utility in natural and social sciences and user friendliness. And considering all these features, we decided to use AnyLogic for agent-based modelling in education. We have selected three models that can be used by students of social and natural sciences, or more specifically, medical students, biology students, mathematics and computer science students, economics and business students.

4 Examples of agent-based modelling in anylogic used for education

A. Epidemiological models

Epidemics of infectious diseases are triggering interest in predicting epidemic dynamics. Agent-based simulations can be used for education process for the medical students involved in public health and epidemiology. For this goal, universities and research centres are using simulations as teaching tools for these students.

Simulation of spread on infectious disease is playing a central role in controlling spread of infection and making prediction that can help monitoring of epidemic [27]. Some of important epidemiologic models are SI, SIS, SIR, SIRS, and SEIR, SEIR -D model without vital dynamics and with vital dynamics.

Here is given an example, where SEIR-D model without vital dynamics is explained.

In the SEIR-D model, the total population of N individuals are divided in 5 categories: susceptible (S), exposed (E), infected (I), recovered (R), and death (D).

- Susceptible – the started population people who are not infected by the virus.
- Exposed - people who are infected but who can’t infect others
- Infectious - people who are infected and who can infect others
- Recovered – people who have recovered from the virus.
- Death – people who death as consequence from infectious disease.

This model relies on the assumption of a totally susceptible population at time t0 as a starting point of the pandemic.

The goal of considering of SEIR-D is to explain the variation of S(t), E(t), I(t), R(t), D(t). This model can help medical students in public health and epidemiology, for easy understanding of a spread of the any infectious disease.

The SEIRS-D model in Anylogic simulation software is represented in Fig 1.

In Anylogic, stocks are used to represent real-world processes (material, knowledge, people, money, etc) and it define static part of the system. Flows define their rate of change - how stock values change in the time, and it define the dynamics of the system.

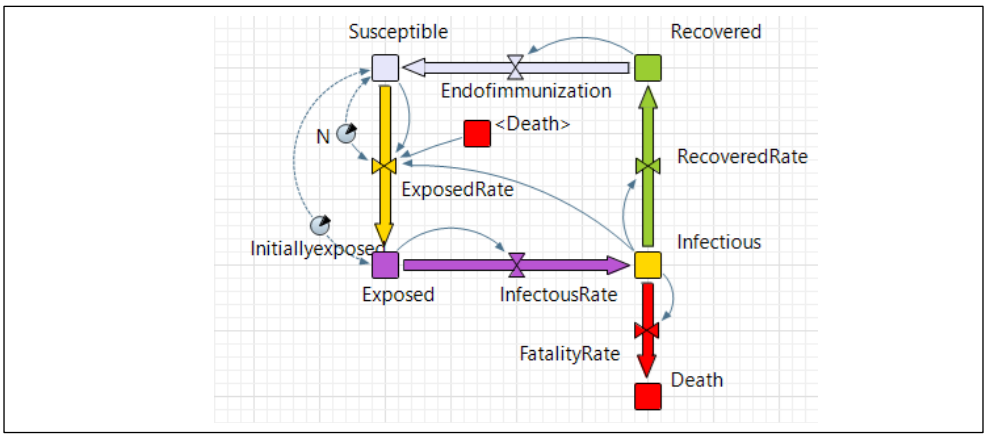


Figure 1: SEIR-D model in AnyLogic

AnyLogic automatically generates a stock's formula according to the user's stock-and-flow diagram. AnyLogic automatically created these formulas when the flow is added. This process can be easily done by students or teachers to visually represent real situation of spreading the disease.

Next step is defining the parameters and dependencies. Seven Parameters are defined: Total Population, Infectivity, ContactRate, AverageIncubationTime, AverageIllnessDuration, AverageImmunizationperiod, FalalityRate, with their default values (As shown in Fig. 2).

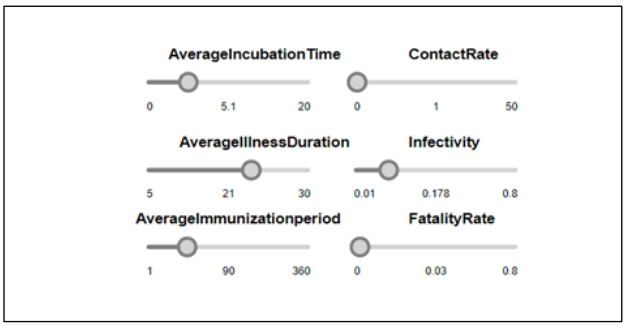


Figure 2: Parameters in the model

- Total Population=2000000
- Infectivity=0.01
- ContactRateInfectious=1
- AverageIncubationTime= 5.1 days
- AverageIllnessDuration= 21 days
- AverageImmunizationperiod = 90 days
- FatalityRate = 0.03

Initially, the number of infected persons is 1. I(0) = 1, and the number of the exposed persons E (0) = 20. The output of the model run, is given in Fig 3.

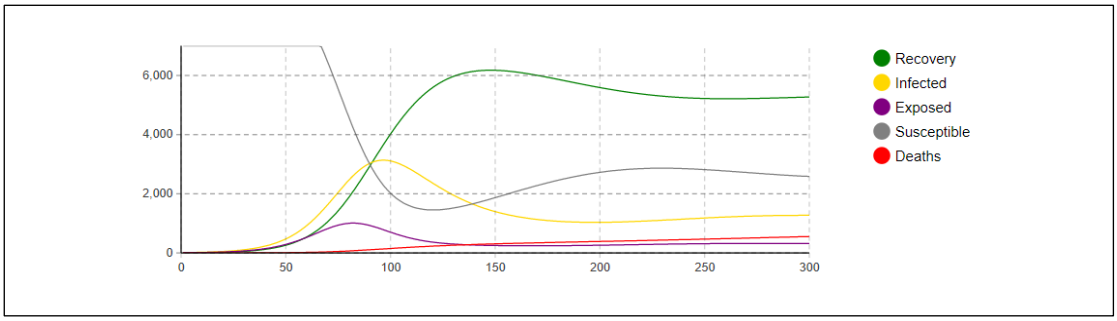


Figure 3: Output from SEIR-D model

The medical student, using this model, has powerful tools for prediction of the spread of the infectious disease. This model can be modified by students to track some epidemic spread (for example COVID-19 pandemic) [28]. After discussing death rates, prevention and treatment options and genetic and age-related variation in host susceptibility, the students can decide to focus on transmission into their model. Through discussion with the professors, they can realize how the transmission of infection disease can occur. This exercise with extending a model to reflect specific biological assumptions helps students understand the iterative process by which models are developed. Also, students can understand the utility of simpler models to understanding key features of the system's behaviour [29].

On the other hand, this model can be important for math and computer science students, because the model is given by the system of the following differential equations:

$$\frac{dS(t)}{dt} = \xi R(t) - \frac{\beta}{N - D} I(t)S(t)$$

$$\frac{dE(t)}{dt} = \frac{\beta}{N - D(t)} I(t)S(t) - \alpha E(t)$$

$$\frac{dI(t)}{dt} = \alpha E(t) - \gamma I(t) - \Delta I(t)$$

$$\frac{dR(t)}{dt} = \gamma I(t) - \xi R(t)$$

$$\frac{dD(t)}{dt} = \Delta I(t)$$

This model can be used as a good example of how differential equations can be implemented in epidemiological models.

Advantages of using agent - based simulation in epidemiology are in the fact that the mathematical representation of processes enables transparency and accuracy regarding the epidemiological assumptions. This allows students with their professors to test understanding of the epidemiology disease by comparing model results and results obtained from observation. Also, mathematical models can help predicting outcomes and adjustment of measures for stopping the spread of infections, as well as taking new appropriate measures.

B. Hospital emergency department simulation

This model is important and can be applied in process of education for the business students, and for a math and computer science's students.

For business students, the model can be used for well organizing of the healthcare systems. For the math and computer science's students is good example for hybrid model that integrates methods of discrete event simulation and agent-based simulation.

Overcrowding in the Emergency Department (ED) is one of the most important issues in healthcare systems. This situation leads to an increase in length of stay, a decrease in the quality of care and the burnout of nursing staff.

Two major causes of this congestion are identified, the first one is unjustified Emergency Department visits and the second one a lack of downstream beds. An unjustified emergency visit concerns a patient who have no health problem or a non-emergency health problem. This situation creates a work overload for the medical staff. The lack of downstream beds increases the length of stay in the Emergency Department because patients must wait for a bed in a relevant medical unit. Sometimes patients are admitted to a medical unit that is not adapted to their pathology to decrease the ED congestion. This situation is problematic because it reduces the quality of care.

First the patients come to the emergency department of the hospital, in the department they are checked whether they are emergency cases or not. In case of an emergency and in relation to the condition of the patient with an emergency, some mandatory medical tests are performed, such as different X-rays of certain parts of the body or other diagnostics tests.

For the emergencies, there must always be beds available in the hospital and after the medical tests are performed, it is decided whether to keep the patient and to determine his diagnosis or just to determine the diagnosis and patient can leave the hospital.

If the case is not urgent, the patient's vital signs such as pulse, temperature, blood pressure and respiratory rate are checked. After checking patient's vital signs, his treatment is determined. Because these patients are not urgent from high degree, additional medical tests may not be needed, therefore they can only be diagnosed if necessary and discharged from the hospital, but still, some can leave without the need for a diagnosis.

For the successful development of this simulation, a Discrete event simulation model and agent-based simulation model in Any Logic program is used. In classic discrete event tools, the entities are passive and can only have attributes that affect the way they are handled. In AnyLogic multimethod simulation software, entities and resources can be modelled as agents with individual behaviour and state changes.

In this simulation triangular distribution is used because the exact rate of patient arrival is not known, therefore, a minimum, a most probable, and a maximum value for a triangular distribution are set.

The model in Anylogic is presented in Fig 4.

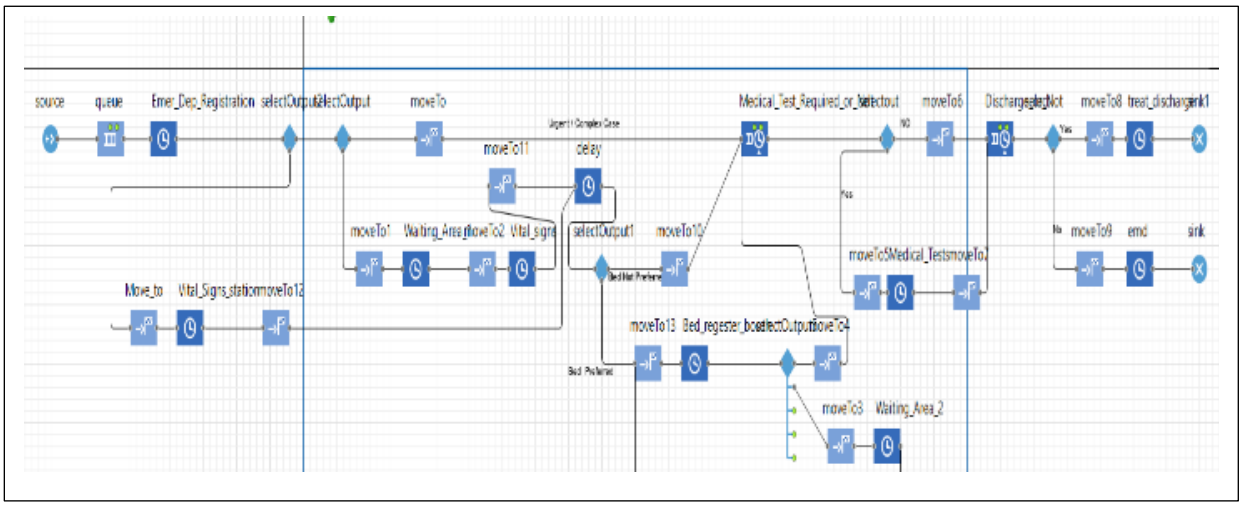


Figure 4: Emergency department model in anylogic

The result of simulation is given in Fig 5.

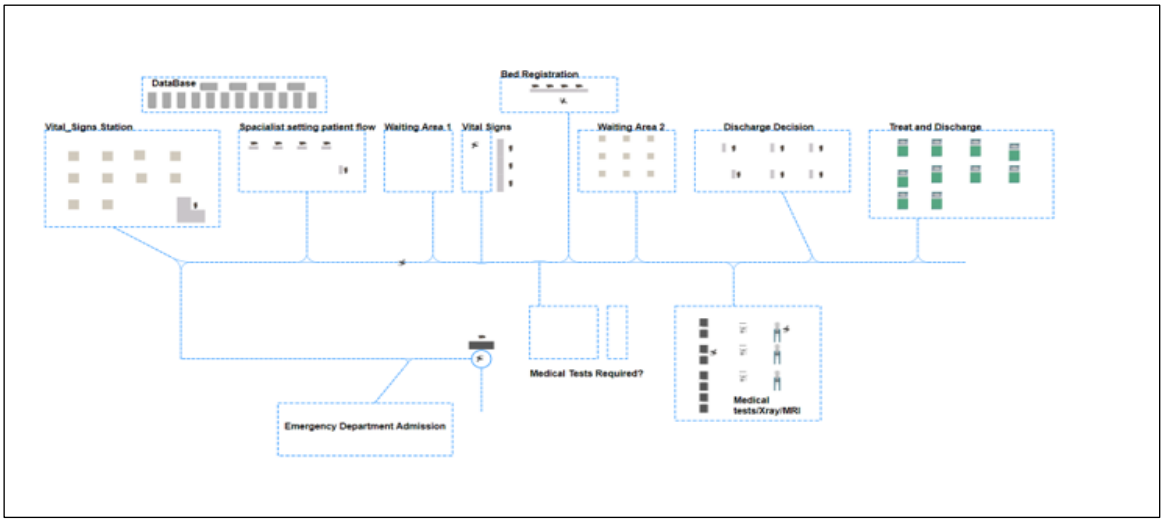


Figure 5: Output from simulation

This simulation can be helpful for the students of business and other organization sciences, managing and logistics, also and computer science students. The model can be helpful in the process of improving organization on the Hospital Emergency Department, to obtain optimal number of rooms, beds, and some other things for each sub-department of the ED, as well as have an estimated price for the ED, to optimally serve patients entering the ED with a known arrival rate.

This simulation can be useful because students can experiment on the simulation model instead of real Hospital Emergency Department.

The students can modify appropriate parameters and estimate output results from these parameters. Therefore, using this model students can easily manage with real problems like this one.

C. Market models

An agent-based model of a costumer cinema is considered for this example. In this model each costumer is an agent. The model includes 5000 people who have not seen the movie in the cinema, but a combination of advertising will eventually lead them to purchase the ticket to watch it. Also, advertising's influence on consumer demand is considered, by allowing a specific percentage of them to become interested in purchasing the ticket during a given day. Advertising effectiveness = 0.1 determines the percentage of potential users that become ready to buy the product during a given day. In Fig 6 is presented diagram of Cinema model presented in AnyLogic.

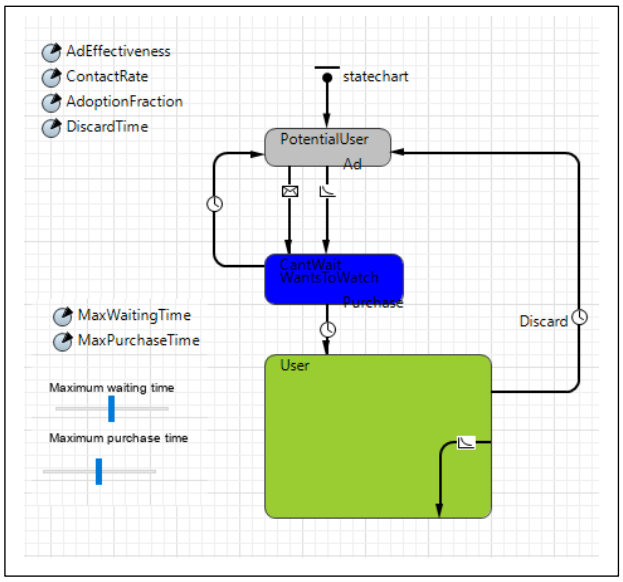


Figure 6: Cinema model

The parameters that are used represent several functions. The first parameter AdEffectiveness defines the percentage of potential users who become ready to buy the ticked and watch the movie during a given day. The second ContactRate represent how many contacts a person has per day with other PotentialUsers. The third AdoptionFraction is used to show us how much the ContactRate (the contact between two PotentialUsers) has affection. The last parameter, DiscardTime, represents how much time will the User wait to become PotentialUser again.

There are two more parameters to test the impatience of the customers. MaxWaitingTime, which is the maximum time a user will wait for the product (in this case, seven days), and MaxDeliveryTime, which is the maximum time for delivery a product (in this case, 20 days).

When the program is run, the 5000 population that are previously selected are obtained. Mostly there are gray Potential Users because the patience is very low and the max waiting time in this case is 7 days. The yellowGreen which are the Users are less and when they are done with watching the movie, they cannot go back for another 6 months.

This Cinema model simulates how 5000 people will react if they all are PotentialUsers and waiting to purchase one ticked for the one movie in the Cinema. From this model it can be concluded that 5000 people is a lot for just one selling counter and the waiting line is too long, which means that the customers will have high impatience and most of them won't wait, eventually quit, and go back to PotentialUsers. Therefore, if the purpose of the model is to sell tickets to 5000 people there must be more than one selling counter, therefore the waiting line won't be too long.

The output from simulation in given in Fig 7.

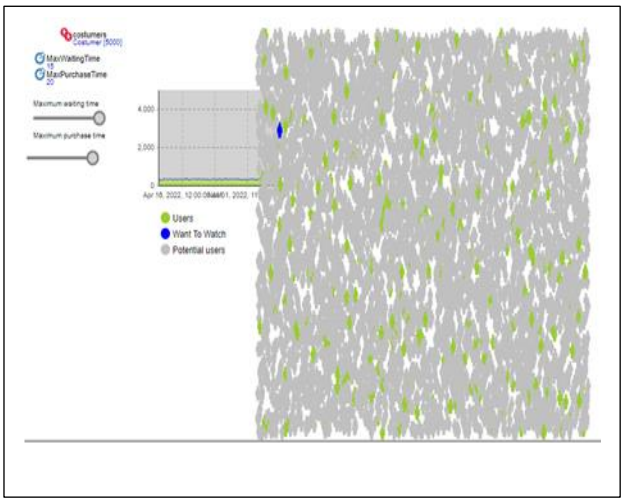


Figure 7: Output from Cinema model

This model can be good example for computer science students, economics, and business students. With making different adjustments of the parameters, students can watch changes of the behaviour of the model and can easily understand how appropriate changes reflected in consumer behaviours and whole system dynamics and can improve customer satisfaction. This model can help students to make market predictions. Students can easily apply and extend this obtained knowledge to real problems like this.

5 Conclusion

Rapid advances in computing power and increasing use of ICT in all aspects of life have made agent-based modelling and simulation (ABMS) feasible and appealing tool for easily studding teaching and understanding different subjects. Simulation-based learning can offer learning with approximation of practice, overcoming limitations of learning in real-life situations. Performing modelling and simulation activities in educational environments can be an effective tool for learning complex and dynamic systems. Students using simulation can be more motivated for learning, gaining new skills, easily understanding subjects, gaining intuition, and making generalization. The opportunity to alter and adjust real life aspects and situations, in a way that facilitates learning and practicing makes simulation an effective educational tool. Simulation-based learning can start early in study programs because it can be effective for beginners and advanced learners too. Simulation models could be used as a tool in education system, from primary and secondary school and for higher education in learning and teaching subjects in undergraduate curriculums.

Agent-based modelling and simulation (ABMS) is a powerful technique in simulating and exploring phenomena that includes a large set of active components represented by agents.

Also, agent-based models offer an extensible way to model different systems consisting of autonomous and interacting agents which perform their actions and adapt their behaviours.

In this paper are given some agent-based simulations examples that can be used by math and computer science students, medical students and business and management students, for easily understanding of learned material and gaining skills for facing with real problems.

References

[1] A. Maria. "Introduction to modeling and simulation." Proceedings of the 29th conference on Winter simulation. 1997. https://doi.org/10.1145/268437.268440

[2] J. Banks, J. S. Carson, and B. L. Nelson, "Discrete-Event System Simulation", Second Edition, Prentice Hall, 1996.

[3] J. M. Aughenbaugh, C. Paredis. "The Role and Limitations of Modeling in Systems Design", Proceedings of IMECE2004, ASME International Mechanical Engineering Congress and RD&D Expo November 13-19, 2004, Anaheim, California USA. https://doi.org/10.1115/imece2004-59813

[4] G. S. Fishman, "Discrete-Event Simulation: Modeling, Programming, and Analysis", Springer, Series in Operations Research, 2001. https://doi.org/10.1007/978-1-4757-3552-9

[5] R. Maidstone, "Discrete Event Simulation, System Dynamics and Agent Based Simulation: Discussion and Comparison", The University of Manchester, 2012.

[6] A. Borshchev, I. Grigoryev, "The Big Book of Simulation Modeling MultiMethod Modeling with AnyLogic 8", 2020.

[7] F. Klügl, A.L. C. Bazzan, "Agent-Based Modeling and Simulation", Ai Magazine, 33(3), pp. 29-40, 2012. https://doi.org/10.1609/aimag.v33i3.2425

[8] Tkaczyk Rafal, Maria Ganzha, and Marcin Paprzycki. "AgentPlanner-agent-based timetabling system". Informatica 40. no. 1, 2016.

[9] Zia Kashif, Dinesh Kumar Saini, Arshad Muhammad, and Umar Farooq. "Agent-Based Simulation of Socially-Inspired Model of Resistance against Unpopular Norms." Informatica 43, no. 2, 2019. https://doi.org/10.31449/inf.v43i2.1888

[10] Djezzar Nedjma, Iñaki Fernández Pérez, Noureddinne Djedi, and Yves Duthen. "A computational multiagent model of bioluminescent bacteria for the emergence of self-sustainable and self-maintaining artificial wireless networks." Informatica 43, no. 3, 2019: pp. 395-408. https://doi.org/10.31449/inf.v43i3.2381

[11] O. Chernikova, N. Heitzmann, M. Stadler, D. Holzberger, T. Seidel, and F. Fischer. "Simulation-based learning in higher education: a meta-analysis." Review of Educational Research 90, no. 4, 2020: pp.499-541. https://doi.org/10.3102/0034654320933544

[12] C.J. Brigas, "Modeling and simulation in an educational context: Teaching and learning sciences." Research in Social Sciences and Technology 4, no. 2, 2019 pp: 1-12. https://doi.org/10.46303/ressat.04.02.1

[13] C. Zambon, Antoni., Jana R. Saito, William H. Yonenaga, and Feginaldo S. Figueiredo. "The Introduction of Simulation as Teaching and Learning Tool." In ICSTM. 2000.

[14] H Stančić, S. Seljan, A. Cetinić, and D. Sanković. "Simulation models in education." 2007, pp: 469-481.

[15] J. P. Kincaid, R. Hamilton, R. W. Tarr, and H. Sangani. "Simulation in education and training." In Applied system simulation, pp. 437-456. Springer, Boston, MA, 2003. https://doi.org/10.1007/978-1-4419-9218-5_19

[16] K. J Murphy, S. Ciuti, and A. Kane. "An introduction to agent-based models as an accessible surrogate to field-based research and teaching." Ecology and evolution 10, no. 22, 2020, pp: 12482-12498. https://doi.org/10.1002/ece3.6848

[17] E. Bonabeau, "Agent-based modeling: Methods and techniques for simulating human systems." Proceedings of the national academy of sciences 99, no. suppl 3, 2002, pp: 7280-7287. https://doi.org/10.1073/pnas.082080899

[18] K.Rakic, M. Rosic, I. Boljat, "A Survey of Agent-Based Modelling and Simulation Tools for Educational Purpose", Technical Gazette 27, 3, pp. 1014-1020, 2020. https://doi.org/10.17559/tv-20190517110455

[19] https://ccl.northwestern.edu/netlogo/

[20] R.Hegselman, U.Mueller. G. Troitzsch, "Modelling and Simulation in the Social Sciences from the Philosophy of Science Point of View", SpringerLink, 1996. https://doi.org/10.1007/978-94-015-8686-3

[21] E. Meir, EcoBeaker 2: Teaching Ecology and Conservation Through Computer Experiments, Proceedings, University of Washington, USA, EdMedia+Innovative Learning, Association for the Advancement of Computing in Education (AACE), Waynesville, NC, 1999.

[22] https://cress.soc.surrey.ac.uk/s4ss/links.html

[23] https://simbio.com/

[24] https://swmath.org/software/30587

[25] https://www.microsimulation.ac.uk/jas-mine/

[26] A. Borshchev, "MultiMethod Modeling: AnyLogic Chapter", In book: Discrete-Event Simulation and System Dynamics for Management Decision Making, April 2014. https://doi.org/10.1002/9781118762745.ch12

[27] M, Ljubenovska, L. Koceva Lazarova, N Stojkovikj, A. Stojanova, and M. Miteva. "Mathematical modeling of COVID-19 virus." CIIT, 2021: 66-69.

[28] L.Koceva Lazarova, N.Stojkovikj, A.Stojanova, M.Miteva. "Application of differential equations in

epidemiological model", BJAMI, 4(2), pp. 91-102, 2021.

[29] E. N. Bodine, R. M. Panoff, E. O. Voit, A. E. Weisstein, "Agent-based Modeling and Simulation in Mathematics and Biology Education", Bulletin of Mathematical Biology volume 82, Article number: 101, 2020. https://doi.org/10.1007/s11538-020-00778-z

A Novel Fuzzy Modified RAFSI Method and its Applications in Multi-Criteria Decision-Making Problems

Garima Bisht*, A. K. Pal
Department of Mathematics, Statistics and Computer Science, Govind Ballabh Pant University of Agriculture and Technology, Pantnagar, 263145, Uttarakhand, India
E-mail: garimabisht98@gmail.com, ak.pal@gbpuat-cbsh.ac.in
*Corresponding author

Keywords: multi-criteria decision-making, rank reversal, RAFSI, triangular fuzzy numbers

Received: April 30, 2022

In real-life decision-making problems, the constraints may change from time to time. Change in certain decision elements can lead to the introduction of new alternatives or the removal of old alternatives to the existing decision, resulting in rank reversal. Rank reversal is the most significant problem that can't be ignored in multi-criteria decision-making (MCDM) methods. Ranking of alternatives through functional mapping of criterion subintervals into a single interval (RAFSI) method effectively removes the problem of rank reversal, but there are some limitations like standardized decision matrix is obtained by the assumption of supreme value as at least six times improved than the anti-supreme value, which is not always true. This paper aims to address those limitations by giving a modified form of the RAFSI (MRAFSI) method. As real-life problems are associated with uncertainty in the form of linguistic terms, a fuzzified form of the MRAFSI method has been given using triangular fuzzy numbers (TFNs) to deal with uncertainty. The effectiveness of the presented method is illustrated using a real-time case study to rank five stocks under the National Stock Exchange (NSE) for the year 2021 and is compared with other MCDM methods for validation. The supplier selection problem has been taken as an example to show the application of the Fuzzy Modified RAFSI (FMRAFSI) method.

Povzetek: Študija predstavlja Fuzzy Modified RAFSI (FMRAFSI) metodo za reševanje problemov večkriterijskega odločanja (MCDM), ki obvladuje negotovost z uporabo trikotnih mehkih števil in zmanjšuje problem obratnega razvrščanja.

1 Introduction

MCDM methods proved as a very important tool in solving most real-world problems. But one of the foremost significant problems that can't be ignored in most of the MCDM methods is rank reversal, the matter of unpredicted modification within the ranking of alternatives with the addition of the latest alternative or removal of an old alternative. MCDM methods are also prone to rank reversal when a problem is decomposed into multiple smaller problems keeping the standard weight and alternative scores unaltered [1]. The key explanation for rank reversal is the use of normalization, which changes with the addition or deletion of alternatives. This distorts the initial data and violates the 'Principle of Independence from Irrelevant Alternatives (PIIA). This is often true for any normalization [2]. Since differences in dimensional units of attributes can only be eliminated by normalization in most of the MADM approaches it becomes a vital part.

During the utilization of the Analytic Hierarchy Process (AHP), the matter of rank reversal was initially observed by Belton and Gear [3]. The identical was also noticed by Triantaphyllou and Mann [4] in AHP during the substitution of the worst alternative with an anti-ideal alternative. Saaty and Varga [5] presented that the matter of rank reversal can happen because of the occurrence of almost identical copies within the set of alternatives. They also opined that the addition of a new alternative can practically modify the previous preference order. Fedrizzi et al., [6] presented that the possibility of rank reversal rests on the distribution of criteria weights i.e., the entropy of the weight distribution. They established that the projected possibility of rank reversal rises with the weight's entropy. Further many authors noticed this problem in several MCDM methods because of the mutual correlations between the relevant and irrelevant alternatives, as a consequence of normalization [7]. Wang and Elhag [8] presented a technique to evade rank reversal in AHP by preserving the local significance of alternatives with the introduction of a new alternative. Mufazzal and Muzakkir [9] proposed a proximity index to minimize the rank reversal in MCDM problems. Sałabun et al., [10] developed a new MCDM method called the Characteristic Objects Method (COMET). They established that it's better than AHP concerning rank reversal.

De Farias Aires and Ferreira [11] introduced an approach targeting the identification of rank reversal during the normalization process in the TOPSIS method. Yang and Wu [12] introduced a novel R-VIKOR-based method to address rank reversal problems. Majumdar et al., [13] investigated a novel form of rank reversal specifically within the Analytic Hierarchy Process (AHP), identifying the aggregation method and criteria weight

normalization as pivotal factors contributing to its occurrence. Similarly, Liu and Ma [14] delved into the causes of rank reversal within the ELECTRE II method, offering insights into its evaluation. Additionally, Tiwari and Kumar [15] presented a robust rank reversal technique for cloud service selection using the TOPSIS method with a Gaussian distribution. Yang et al., [16] an adapted approach to minimize rank reversal occurrences within the classic TOPSIS method. However, within the previous couple of years, a huge number of advanced MADM methods gave effective outcomes for resolving real-world problems [17]. But a maximum of those methods are not able to effectively remove the matter of rank reversal.

There are abundant applications of MCDM methods in real-life problems. Some of the applications consist of construction method selection for green building projects, portfolio selection, business and marketing, supplier selection, healthcare management, wastewater management, transportation problems, site selection for solar thermoelectric power plants, infectious waste disposal, industry development, flood detection criteria, social media analysis, supply chain network design, etc. In such cases, if rank reversal exists, and that too of higher order, a non-optimal alternative gets selected, thus resulting in a big concession.

Zizovic et al., [18] developed a new method referred to as Ranking of alternatives through functional mapping of criterion subintervals into a single interval (RAFSI), and its fuzzified form has been used for solving the selection problem in health organizations for COVID-19 virus pandemic [19], and for choosing a group of construction machines for enabling mobility [20]. Although this method successfully removes the problem of rank reversal, some modifications may be done to this method to make it better for solving real-life problems. This paper aims to work on the modifications that can be made to the RAFSI method. Also, since real-life problems are associated with uncertainty in the form of linguistic terms, the fuzzified form of the MRAFSI method has been given using triangular fuzzy numbers (TFNs) to deal with uncertainty persisting in the real world. To show the applicability of the presented method it has been applied to two important decision-making problems namely indices selection and supplier selection problems. For validation comprehensive analysis has been done with other well-known MCDM methods.

The rest of the paper is organized as follows. Section 2 discusses the RAFSI method and its shortcomings. Section 3 presents the mathematical formulation of the modified RAFSI method with the real case study as an application along with the comparative analysis. Section 4 presents the fuzzification of the MRAFSI method with application and comparison with the traditional fuzzy MCDM methods. Section 5 discusses the theoretical basis of the proposed approach and compares it with existing approaches for rank reversal, followed by sensitivity analysis in Section 6. At last section 7 concludes the paper.

1.1 Related work

Extensive research has been conducted in the field of rank reversal, resulting in a vast body of literature. To gain insights into this domain, we conducted a comprehensive review of relevant studies and categorized them based on the approach employed, the method utilized, and the limitations identified. The classification of these studies is presented in Table 1, offering a systematic overview of the diverse research framework surrounding the rank reversal problem.

Table 1: Literature review on rank reversal approaches

Year	Author	Method	Limitations
2023	Saluja et al. [21]	Proximity indexed value (PIV)	Struggles with a substantial prevalence of rank reversal.
2023	Tu and Wu [22]	AHP	Intransitive preference and the prioritization methods cause rank reversals in single pairwise comparison matrices.
2023	Dehshiri and Firoozabadi [23]	Wins in league (WIL)	Sensitive to small changes, limited handling of uncertainty.
2022	Yang et al. [16]	IE-TOPSIS	Relies on supplementary data, potentially unable to eliminate rank reversal.
2021	Tiwari and Kumar [15]	G-TOPSIS	Reliance on Gaussian distribution assumptions, subjective user priority influence.
2021	Kizielewicz et al. [24]	Characteristic Objects method (COMET)	Potential sensitivity to minor variations in input data, uncertainties in handling fuzzy data representations, and a lack of robustness in maintaining consistent rankings.
2020	Stevic et al. [17]	MARCOS	Complex implementation, limited generalizability,

			sensitive to parameter changes.
2020	Zizovic et al. [18]	RAFSI	Subjective criterion interval setting, reliance on an arbitrary superiority threshold, and the potential for identical rankings among different alternatives due to its assumptions on criteria types.

2 RAFSI method

In this section, the RAFSI method given by Zizovic et al., [18] is discussed. Given the initial decision matrix with weights of criteria estimated by any of the known methods, the RAFSI method has the subsequent stages.

1) The DM describes ideal (a_{Ij}) and anti-ideal (a_{Nj}) values for individual criteria.
2) Mapping of elements of the decision matrix into criteria intervals.
   - $C_j \in [a_{Nj}, a_{Ij}]$, where C_j belongs to *max* type criteria.
   - $C_j \in [a_{Ij}, a_{Nj}]$, where C_j belongs to *min* type criteria.

   Mapping of subintervals into criteria interval $[n_1, n_{2k}]$ by the formula-

   $$f_s(x) = \frac{n_1 - n_{2k}}{a_{Ij} - a_{Nj}} x + \frac{a_{Ij} n_1 - a_{Nj} n_{2k}}{a_{Ij} - a_{N_j}}$$

   It is supposed that the optimal value is six times improved than the non-optimal value i.e., $n_1 = 1$ and $n_{2k} = 6$. In this way, a standardized decision matrix is obtained.
   - for *max* type criteria if $a_{xj} > a_{Ij}$, then $f(a_{xj}) = f(a_{Ij})$
   - for *min* type criteria if $a_{xj} < a_{Ij}$, then $f(a_{xj}) = f(a_{Ij})$
3) Next, calculate arithmetic and harmonic mean of n_1, n_{2k}.

   $$A = \frac{(n_1 + n_{2k})}{2}, H = \frac{2}{\frac{1}{n_1} + \frac{1}{n_{2k}}}$$

4) Find a normalized decision matrix
   - for max type criteria $\hat{s}_{ij} = \frac{s_{ij}}{2A}$
   - for min type criteria $\hat{s}_{ij} = \frac{H}{2s_{ij}}$
5) Calculate criteria functions of alternatives $V(A_i)$.

   $V(A_i) = \omega_1 \hat{s}_{i1} + \omega_2 \hat{s}_{i2} + \ldots\ldots + \omega_n \hat{s}_{in}$

   Finally, alternatives are ranked in descending order of $V(A_i)$.

2.1 Limitations of RAFSI method

This section discusses the limitations of the existing RAFSI method.

1) In this method the DM's set the interval for each criterion by assumption without the use of any standard formula.
2) In this method for forming a standardized decision matrix, it is supposed that the optimal value is at least six times better than the non-optimal value, but it is not always true.
3) This method assumes that
   - for max type criteria if $a_{xj} > a_{Ij}$, then $f(a_{xj}) = f(a_{Ij})$
   - for min type criteria if $a_{xj} < a_{Ij}$, then $f(a_{xj}) = f(a_{Ij})$

   but this may lead to the same ranking of two different alternatives.

The following example illustrates it more efficiently.

Example: Consider the initial decision matrix given below and let the criteria sub-intervals be defined as-

$C_1 \in [2, 10]$, $C_2 \in [4, 8]$, $C_3 \in [0, 5]$

$$A = \begin{array}{c} \\ A1 \\ A2 \\ A3 \\ A4 \\ \\ \end{array} \begin{array}{c} \begin{array}{ccc} C1 & C2 & C3 \end{array} \\ \begin{bmatrix} 12 & 6 & 1 \\ 10 & 6 & 1 \\ 5 & 7 & 4 \\ 8 & 5 & 3 \\ max & max & min \end{bmatrix} \end{array}$$

thus, according to RAFSI method $f(12) = f(10)$ for alternative A_1, and other values being same for alternatives A_1 and A_2 we get same rank for alternative A_1 and A_2. But as it can be seen since criteria C_1 is of the maximum type so A_1 must be at a higher rank than A_2.

3 Modified RAFSI (MRAFSI) method

In this section, we have tried to overcome the shortcomings of the RAFSI method. The flow chart of the MRAFSI method is shown in Figure 1.

Let the initial decision matrix consists of m alternative A_1, A_2, …. A_m and n criteria C_1, C_2, …… C_n. Find the weights of criteria by any one of the known methods considering the relative importance between criteria such that $\sum_{i=1}^{n} w_i = 1$. The initial decision matrix is shown as follows.

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

The MRAFSI has the following steps-

Step.1. Find intervals for each criterion using the mean (μ) and standard deviation (σ) of the values of criteria for different alternatives as given in the decision matrix.

$$[\ \mu\text{-}2\times\sigma,\ \mu\text{+}2\times\sigma\ ] = [\ n_1,\ n_2]$$

Step.2. Find the normalized decision matrix S = $[s_{ij}]_{m*n}$ by the use of the following formula-

$$s_{ij} = \frac{1}{1+e^{-x}} \quad (1)$$

here,

$x = \frac{a_{ij}-n_1}{n_2-n_1}$ for beneficial criteria

$x = \frac{n_2-a_{ij}}{n_2-n_1}$ for non-beneficial criteria

Step.3. Calculate the criteria functions of alternative $V(A_i) = w_1 s_{i1} + w_2 s_{i2} + \ldots\ldots + w_n s_{in}$ (2)

where $w_1, w_2, \ldots\ldots w_n$ represents the weight of criteria.

Finally, rank the alternatives in descending order of $V(A_i)$.

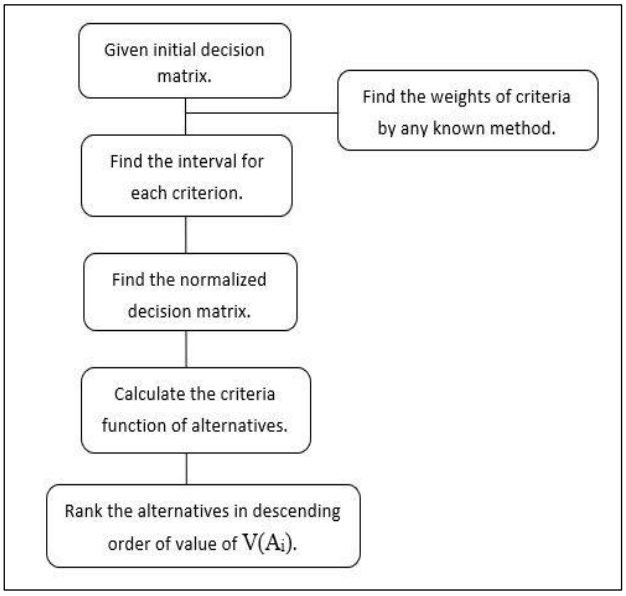


Figure 1: Block diagram of the MRAFSI method

3.1 Applications of MRAFSI multi-criteria model

This section presents the application of the MRAFSI methodology for the stock selection problem. A real case example of NSE (National Stock Exchange) is shown for selecting the best indices out of the given four indices Hindustan unilever (A_1), Asian paints (A_2), Tata consultancy services (A_3), Reliance industries (A_4) with four criteria Return on equity (ROE) (C_1), Earning per share (EPS) (C_2), Face value (C_3), P/E ratio(C_4) of year 2021 downloaded from www.ratestar.in. The weights of each criterion are given by w_i = (0.104445,0.13603,0.645511,0.114014) found by the entropy method. The decision matrix is demonstrated below.

$$\begin{array}{c} \\ A1 \\ A2 \\ A3 \\ A4 \\ \\ \end{array} \begin{bmatrix} C1 & C2 & C3 & C4 \\ 28.63 & 37.34 & 1 & 56.10 \\ 27.71 & 31.82 & 1 & 90.83 \\ 38.55 & 102.11 & 1 & 34.83 \\ 9.27 & 98.51 & 10 & 27.87 \\ max & max & max & min \end{bmatrix}$$

Applying the steps of MRAFSI method-

Step.1. Find the criteria subintervals using the mean and standard deviation of each column.
$C_1 \in [1.62,50.45]$; $C_2 \in [-8.6,143.53]$;
$C_3 \in [-5.75,12.25]$; $C_4 \in [-4.17,108.98]$;

Step.2. Find the normalized decision matrix by applying eq.1.

$$f_{A_1}(C_1) = \frac{1}{1+e^{\frac{-(28.63-1.62)}{(50.45-1.62)}}} = 0.634839$$

similarly solving other values, the normalized decision matrix can be obtained and as shown below:

$$\begin{array}{c} \\ A1 \\ A2 \\ A3 \\ A4 \\ \\ \end{array} \begin{bmatrix} C1 & C2 & C3 & C4 \\ 0.6348 & 0.5749 & 0.59267 & 0.6148 \\ 0.6305 & 0.5661 & 0.59267 & 0.5400 \\ 0.6805 & 0.6743 & 0.59267 & 0.6582 \\ 0.5391 & 0.6691 & 0.70578 & 0.7191 \\ max & max & max & min \end{bmatrix}$$

Step.3. Using eq. 2. find the criteria functions $V(A_i)$ of alternatives and rank them in descending order of $V(A_i)$ as shown in Table 2 and Figure 2.

Table 2: Final ranking of alternatives

Alternatives	$V(A_i)$	Rank
Hindustan unilever	0.597184	3
Asian Paints	0.586997	4
Tata consultancy services	0.620423	2
Reliance industries	0.679521	1

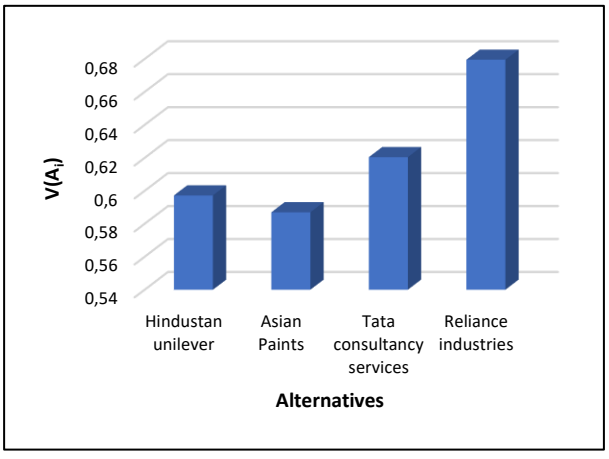


Figure 2: Ranking of stocks

Based on the above results, we found that Reliance industries is the best stock to invest in.

3.2 Rank reversal problem

The four alternatives are ranked according to MRAFSI method, now we need to check rank if we remove one alternative from them. Let us remove the alternative Hindustan unilever from the given alternatives. We find that the on removing the alternative of rank 3rd all the alternatives, after that alternative shift one rank up, without causing any rank reversal. Thus, it is observed that MRAFSI method gives effective results in dynamic environment as shown in Table 3.

Table 3: Ranking after removing one alternative

Alternatives	**V(A_i)**	**Rank**
Asian Paints	0.586997	3
Tata consultancy services	0.620423	2
Reliance industries	0.679521	1

Now let us add another alternative tata steel to the given four alternatives and check the rank. The new decision matrix formed is given below.

$$\begin{array}{c} \\ A1 \\ A2 \\ A3 \\ A4 \\ A5 \end{array}\begin{bmatrix} C1 & C2 & C3 & C4 \\ 28.63 & 37.34 & 1 & 56.10 \\ 27.71 & 31.82 & 1 & 90.83 \\ 38.55 & 102.11 & 1 & 34.83 \\ 9.27 & 98.51 & 10 & 27.87 \\ 10.87 & 317.21 & 10 & 4.3 \\ max & max & max & min \end{bmatrix}$$

After applying the steps of the MRAFSI method we found the rank of alternatives as shown below in Table 4.

Table 4: Ranking after adding one alternative

Alternatives	**V(A_i)**	**Rank**
Hindustan Unilever	0.591028	4
Asian Paints	0.575865	5
Tata consultancy services	0.613342	3
Reliance industries	0.6429	2
Tata steel	0.681706	1

The added alternative stood first in the ranking order, so all the alternatives moved single place down in the order. Thus, the MRAFSI method is resistant to rank reversal problems on adding and removing new alternatives.

3.3 Comparative analysis

For validation, the results obtained by MRAFSI method is compared with other known traditional MCDM methods. The same weights and initial decision matrix are taken in all other methods for comparison of the performance. Table 5 shows the ranking of alternatives using different methods.

Table 5: Ranking obtained by different methods

Method	**Ranking**	**Best alternative**	**Worst alternative**
MRAFSI	A4>A3>A1>A2	A4	A2
TOPSIS	A4>A3>A1>A2	A4	A2
COPRAS	A4>A3>A1>A2	A4	A2
MAUT	A4>A3>A1>A2	A4	A2

It is clear from the above table that there is no conflict in the ranking order of best and worst alternatives by all methods. Hence, this validates the MRAFSI method.

4 Fuzzy MRAFSI method

In this section, we present the fuzzified form of the MRAFSI method. This helps in handling the uncertainty persisting in real-life problems. Fuzzification is performed by applying triangular fuzzy numbers A= (a_1, a_2, a_3), where a_1 presents the smallest likely value, a_2 presents the most probable value and a_3 presents the largest possible value of any fuzzy event. Triangular fuzzy numbers (TFNs), being a specialized case of generalized fuzzy numbers, offer a competent way to present ambiguous information and linguistic preferences. The easy properties of TFNs captivated our attention to design the fuzzy RAFSI method to process the ambiguous information in the form of TFNs.
The fuzzy MRAFSI has the following stages-

Step.1. Formation of the fuzzy initial decision matrix. This matrix is formed by evaluating m alternatives (A_1, A_2, …. A_m) on *n* criteria C_1, C_2, …… C_n. The decision matrix is shown below.

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

where $a_{ij} = (a_{ij}^l, a_{ij}^s, a_{ij}^u)$ denotes the triangular fuzzy number.

Step.2. Find the criteria interval, by finding the mean and standard deviation for each element of TFNs. After finding the ideal and anti-ideal value in form of TFN we have the fuzzy criteria interval.

$C_j \in [n_1, n_2] \;\; j = 1,2,3 \ldots n$

where n_1 and n_2 are TFN's.

Step.3. Convert the initial decision matrix into normalized matrix S = $[s_{ij}]_{m*n}$ by applying the formula

$$s_{ij} = \frac{(1,1,1)}{(1,1,1)+e^{-x}} \tag{3}$$

here,

$x = \frac{a_{ij}-n_1}{n_2-n_1}$ for beneficial criteria

$x = \frac{n_2-a_{ij}}{n_2-n_1}$ for non-beneficial criteria

a_{ij}, n_1, n_2 are all TFN's.
For solving equation (3) use the operations of triangular fuzzy numbers.

Step.4. Calculate the fuzzy criteria functions of alternatives $V(A_i)$ by applying the expression:

$$V(A_i)=w_1 s_{i1} + w_2 s_{i2} + \ldots\ldots + w_n s_{in} \quad (4)$$

where w_j represents the weights of criteria, which an be found by applying any of the known methods of weight determination. Here weight determination is not taken into consideration, they are assumed to be already known.

Step.5. Defuzzification of the fuzzy criteria functions of alternatives $V(A_i)$ is done by applying the expression:

$$V(A_i) = \frac{[V(A_i)^l + 4*V(A_i)^s + V(A_i)^u]}{6} \quad (5)$$

Now rank the alternatives in the descending order of value of $V^*(A_i)$.

4.1 Applications of Fuzzy MRAFSI multi-criteria model

This section presents application of Fuzzy MRAFSI method for the supplier selection problem. An automobile company desires to select raw material suppliers. Three suppliers (S_1, S_2, S_3) are to be selected based on five criteria:

1. Quality supplied item (C_1)
2. Cost of supplied item (C_2)
3. Delivery time of supplied item (C_3)
4. Technology of supplied item (C_4)
5. Flexibility of supplied item (C_5)

The linguistic variables for weights are shown in Table 6.

Table 6: Linguistic variables for weights

Linguistic Variables	Ratings
Very Low (VL)	(0,0.1,0.2)
Low (L)	(0.1,0.3,0.5)
Medium (M)	(0.3,0.5,0.7)
High (H)	(0.6,0.8,0.9)
Very High (VH)	(0.8,0.9,1.0)

Weights of the criteria are given as:

w_1 = (0.83,0.97,1)
w_2 = (0.63,0.83,0.97)
w_3 = (0.77,0.93,1)
w_4 = (0.57,0.77,0.93)
w_5 = (0.5,0.7,0.9)

Applying the steps of fuzzy MRAFSI method to the given problem.

Step.1. Form the Fuzzy decision matrix using linguistic variables for rating shown in Table 7.

Table 7: Linguistic variables for rating

Linguistic Variables	Ratings
Very Poor (VP)	(0,1,2)
Poor (P)	(1,3,5)
Medium (M)	(3,5,7)
Good (G)	(6,8,9)
Very Good (VG)	(8,9,10)

The fuzzy decision matrix is shown below in Table 8 for the given problem.

Table 8: Fuzzy decision matrix

	C_1	C_2	C_3	C_4	C_5
S_1	(8.33,9.67,10)	(7.67,9.33,10)	(7.67,9.33,10)	(7,9,10)	(7,9,10)
S_2	(5.67,7.6,9.3)	(3.67,5.67,7.6)	(3.67,5.67,7.6)	(3.67,5.67,7.6)	(4.33,6.33,8.3)
S_3	(7,8.67,9.67)	(4.33,6.33,8.3)	(4.33,6.33,8)	(5.67,7.67,9.3)	(1.67,3.67,5.6)
	max	min	min	max	max

Step.2. Find the criteria interval by taking the mean and standard deviation of each element of TFN's in the criteria column as shown in Table 9.

Table 9: Interval for first criteria

	8.33	9.67	10
	5.67	7.67	9.33
	7	8.67	9.67
Mean(μ)	7	8.67	9.67
S. D (σ)	1.08	0.82	0.27
μ-2* σ	4.84	7.03	9.13
μ+2* σ	9.16	10.31	10.21

Thus, the interval for C_1 becomes:
$C_1 \in [(4.84,7.03,9.13), (9.16,10.31,10.21)]$
Similarly, we find intervals for all other criteria:
$C_2 \in [(1.72,3.92,6.7), (8.72,10.3,10.63)]$
$C_3 \in [(1.72,3.92,6.5), (8.72,10.3,10.62)]$
$C_4 \in [(2.7,4.7,7.04), (8.18,10.18,10.95)]$
$C_5 \in [(0,1.98,4.43), (8.68,10.68,11.56)]$

Step.3. Find the normalized matrix by applying equation (3).

$$f_{A_1}(C_2) = \frac{(1,1,1)}{(1,1,1)+e^{\frac{-((8.72,10.3,10.63)-(7.67,9.33,10))}{((8.72,10.3,10.63)-(1.72,3.92,6.7))}}}$$

$$= \frac{(1,1,1)}{(1,1,1)+e^{-(-0.63,0.15,0.146)}} = \frac{(1,1,1)}{(2.88,1.86,1.23)}$$

$$= (0.35,0.54,0.81)$$

Similarly solving other values, we get the normalized matrix as shown in Table 10.

Table 10: Normalized decision matrix

	C_1	C_2	C_3	C_4	C_5
S_1	(0,0.69,1)	(0.35,0.54,0.81)	(0.36,0.54,0.79)	(0.49,0.69,0.98)	(0.55,0.69,0.84)
S_2	(0,0.55,1)	(0.53,0.67,0.97)	(0.53,0.67,0.96)	(0.05,0.54,0.7)	(0.49,0.62,0.73)
S_3	(0,0.62,1)	(0.51,0.65,0.96)	(0.52,0.65,0.94)	(0.23,0.63,0.93)	(0.34,0.55,0.6)
	max	min	min	max	max

Step.4. Using eq. (4) calculate the final fuzzy criteria functions of alternatives $V(A_i)$.

Step.5. Final ranking of alternatives is done after defuzzification of fuzzy criteria functions of alternatives $V^*(A_i)$, as shown in Table 11 and Figure 3.

Table 11: Ranking of alternatives

Alternative	$V(A_i)$	$V^*(A_i)$	Ranking
S_1	(1.05,2.63,4.24)	2.635	1
S_2	(1.01,2.57,4.21)	2.585	3
S_3	(1.03,2.62,4.28)	2.630	2

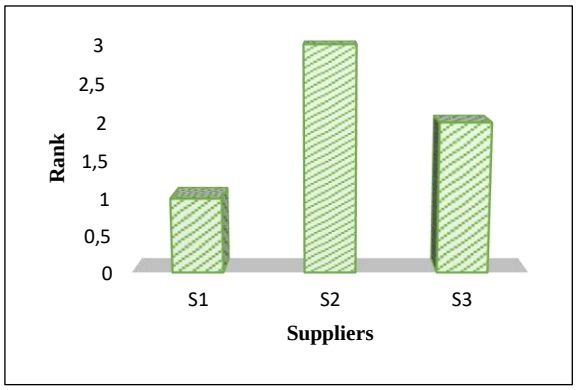


Figure 3: Ranking of suppliers

Based on the above results, we found that supplier 1 is the best alternative.

4.2 Comparative analysis

For validation, the results obtained by the FMRAFSI method is compared with the well-known Fuzzy TOPSIS and Fuzzy VIKOR method. The same weights and initial decision matrix are taken for comparison of the performance. Table 12 shows the ranking of alternatives using different methods.

Table 12: Comparison of ranking order

Method	Ranking	Best alternative	Worst alternative
FMRAFSI	A1>A3>A2	A1	A2
FTOPSIS	A1>A3>A2	A1	A2
FVIKOR	A1>A3>A2	A1	A2
FCOPRAS	A1>A3>A2	A1	A2
FELECTRE	A1>A3>A2	A1	A2
FPROMETHE	A1>A2>A3	A1	A3

It is clear from the above table that there is no conflict in the ranking order of best alternatives by different methods. Hence, this validates the FMRAFSI method.

5 Discussions

5.1 Theoretical basis

The rationale behind the mathematical formulation of mean and standard deviation in the modified RAFSI method is explained below:

Simplicity: This method offers a straightforward and easy-to-understand approach to estimate the mean and standard deviation of TFNs. By breaking down the TFN into its three values (lower, middle, upper), it simplifies the calculation process.

Transparency: It provides a transparent representation of the TFN's uncertainty. By using arithmetic operations (e.g., mean calculation, standard deviation computation) on individual terms, it offers an intuitive way to understand how these terms contribute to the overall statistics of the TFN.

Computational efficiency: Compared to some more complex methods like Monte Carlo simulation or PDF-based approaches, this method is computationally efficient. It avoids the need for extensive simulations or intricate mathematical formulations, making it suitable for quick estimations.

Applicability: This method might be particularly useful in scenarios where simplicity and a quick estimation of the mean and standard deviation are required. It can serve as a preliminary or initial estimation method, especially when dealing with a large number of TFNs in decision-making or uncertainty analysis contexts.

5.2 Comparative analysis

This section conducts a comparative analysis between the proposed approach and other methodologies for addressing rank reversal, as outlined in Table 1. It aims to elucidate the advantages inherent in the proposed approach when compared with existing methods.

1. **Stability against rank reversals:** Unlike methods such as Proximity Indexed Value (PIV), AHP, Wins in league (WIL), IE-TOPSIS, G-TOPSIS, and others prone to rank reversals, the Modified RAFSI method is designed to potentially mitigate the prevalence of rank reversals. It aims to produce more stable and consistent rankings, enhancing the reliability of decision-making processes.
2. **Enhanced handling of uncertainty:** Compared to methods like the Characteristic Objects method (COMET), which struggle with uncertainties and fuzzy data representations, Modified RAFSI offers improved handling of uncertainty. It provides a more robust means of dealing with fuzzy data representations, resulting in more reliable and consistent rankings even in uncertain scenarios.

3. **Reduced sensitivity to small changes:** In contrast to methods sensitive to small changes, such as Wins in league (WIL) and others, Modified RAFSI demonstrates lower sensitivity to minor fluctuations or variations in input data. This characteristic leads to more stable and robust rankings, less likely to be affected by insignificant changes.
4. **Objective ranking:** Similar to G-TOPSIS, RAFSI minimizes subjective bias. It aims to provide a more objective approach, enhancing the credibility and reliability of the rankings by minimizing the influence of subjective user assumptions.
5. **Simplicity and Generalizability:** Unlike complex methods like MARCOS, Modified RAFSI offers a more straightforward implementation while maintaining robustness and applicability across diverse decision-making scenarios. Its simplicity does not compromise its effectiveness in producing meaningful and reliable rankings.
6. **Reduced reliance on supplementary data:** RAFSI's design aims to reduce dependency on supplementary data, similar to how it is with IE-TOPSIS. This characteristic contributes to its practicality and efficiency, allowing it to generate rankings without relying heavily on additional information.

6 Sensitivity analysis

Decision-making is a multifaceted process susceptible to various potential errors. Therefore, a comprehensive analysis before model adoption becomes imperative. This typically involves conducting a sensitivity analysis, which can be executed through diverse approaches such as altering weight coefficients of criteria, changing measurement units expressing alternative values, comparing with alternate methodologies, etc. [25]. Most authors commonly perform sensitivity analyses focusing on adjustments in weight coefficients of criteria [26-27], as is the case in this paper as well. The primary objective of this sensitivity analysis is to gauge the impact of the most influential criterion on the ranking performance of the proposed model [28]. For the sensitivity analysis involving changes in weight coefficients, five distinct scenarios are developed. The basis for the change in weight coefficients makes the change in the weight coefficient of the best criterion C3. The changes in the weight coefficients of this criterion are made in interval $w_3 \in [0, 0.5]$.

The proportion set in this way always provides the condition where $\sum_{i=1}^{4} w_i = 1$. The values of the weight coefficients in all scenarios are shown in Figure 4.

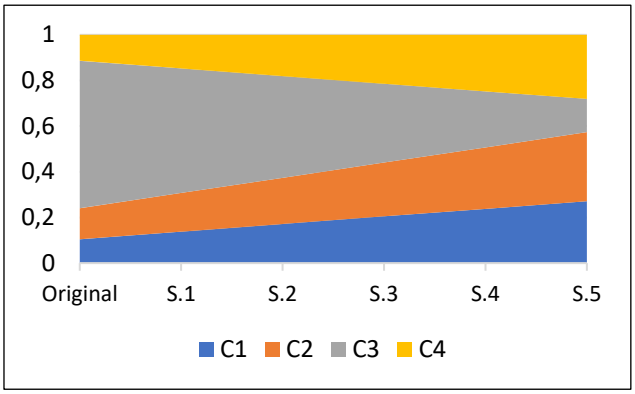


Figure 4: Weights under different scenarios

To further verify the stability of the proposed approach to attribute weights obtained by different methods, we use the objective weights obtained by critic and standard deviation method in place of weights obtained by entropy weights in the example. The weights obtained by different methods are shown in Table 13.

Table 13: The weight vector by different methods

Methods	w_1	w_2	w_3	w_4
Entropy	0.10444	0.13603	0.64551	0.11401
Critic	0.36515	0.18964	0.28223	0.16296
St. dev.	0.2186	0.28373	0.26211	0.23555

The ranking of alternatives by different scenarios and weight determination methods is shown in Table 14. It can be easily observed from Table 14 that although the weights differ greatly, a very small change in ranking results is seen. Thus, the proposed approach is stable in terms of ranking. To further verify the results the SSCs between the ranking obtained is calculated. From Table 15 it is observed that the SSCs between the ranking is greater than 0.8 under different weights. Thus, the proposed approach is stable under different weights.

Table 14: Ranking of alternatives by different scenarios

Alternative	**Original**	**Critic**	**St. Dev.**	**S1**	**S2**	**S3**	**S4**	**S5**
Hindustan unilever	3	3	3	3	3	3	3	3
Asian Paints	4	4	4	4	4	4	4	4
TCS	2	1	2	2	2	2	2	1
Reliance industries	1	2	1	1	1	1	1	2

Table 15: The SSCs between the ranking results

	Original	**Critic**	**St. Dev.**	**S1**	**S2**	**S3**	**S4**	**S5**
Original	1	0.8	1	1	1	1	1	0.8
Critic	-	1	0.8	0.8	0.8	0.8	0.8	1
St. Dev.	-	-	1	1	1	1	1	0.8
S1	-	-	-	1	1	1	1	0.8
S2	-	-	-	-	1	1	1	0.8
S3	-	-	-	-	-	1	1	0.8
S4	-	-	-	-	-	-	1	0.8
S5	-	-	-	-	-	-	-	1

7 Conclusions

This paper discusses the limitations of the RAFSI method and endeavors to address these deficiencies by introducing a modified RAFSI method (MRAFSI). To assess the efficacy of the proposed method, a real case study is conducted to rank five indices of the Bombay Stock Exchange (BSE) for the fiscal year 2020-21. Comparative analysis with established MCDM methods is performed to validate the modified approach, confirming the consistency in results and affirming the validity of the modified method.

In recognition of uncertainties prevalent in real-world scenarios, the MRAFSI method undergoes fuzzification using the triangular fuzzy numbers. The fuzzy modified RAFSI (FMRAFSI) is applied to a supplier selection problem. Comparative validation with traditional fuzzy methods is conducted, revealing congruent outcomes and thus affirming the validity of the FMRAFSI method. Additionally, a sensitivity analysis is carried out to showcase the resilience and reliability of the proposed approach.

For the future work, the proposed framework can be integrated to leverage hybrid models [29-30], thereby achieving more effective outcomes. It would be fascinating to use the proposed method to address a variety of further real-world decision-making issues.

References

[1] Triantaphyllou, E. (2001). Two new cases of rank reversals when the AHP and some of its additive variants are used that do not occur with multiplicative AHP. *Journal of Multi-criteria Decision Analysis*, 10(1), 11–25. https://doi.org/10.1002/mcda.284

[2] Barzilai, J., & Golany, B. (2017). AHP Rank Reversal, Normalization and Aggregation Rules. *Information Systems and Operational Research*, 32(2), 57-64. https://doi.org/10.1080/03155986.1994.11732238

[3] Belton, V., & Gear, T. (1983). On a Short-Coming of Saaty's Method of Analytic Hierarchies. *Omega*, 11, 228–230. https://doi.org/10.1016/0305-0483(83)90047-6

[4] Triantaphyllou, E., & Mann, S.H. (1989). An examination of the electiveness of multi-dimensional decision-making methods: A decision-making paradox. *Decision Support Systems*, 5(3), 303–312. https://doi.org/10.1016/0167-9236(89)90037-7

[5] Saaty, T. L., & Vargas, L. G. (1984). Inconsistency and rank preservation. *Journal of mathematical psychology*, *28*(2), 205-214. https://doi.org/10.1016/0022-2496(84)90027-0

[6] Fedrizzi, M., Giove, S., & Predella, N. (2018). Rank reversal in the AHP with consistent judgements: A numerical study in single and group decision making. Collan M., Kacprzyk J. (eds). In *Soft Computing Applications for Group Decision-making and Consensus Modeling* (pp 213-225). Studies in Fuzziness and Soft Computing, 357. Springer, Cham. https://doi.org/10.1007/978-3-319-60207-3_14

[7] Wang, Y., & Luo, Y. (2009). On rank reversal in decision analysis. *Mathematical and Computer Modelling*, 49, 1221–9. https://doi.org/10.1016/j.mcm.2008.06.019

[8] Wang, Y. M., & Elhag, T.M.S. (2006). An approach to avoiding rank reversal in AHP. *Decision Support System*, 42, 1474–80. https://doi.org/10.1016/j.dss.2005.12.002

[9] Mufazzal, S., & Muzakkir, S. M. (2018). A new multi-criterion decision making (MCDM) method based on proximity indexed value for minimizing rank reversals. *Computers & Industrial Engineering*, 119, 427–38. https://doi.org/10.1016/j.cie.2018.03.045

[10] Sałabun, W., Ziemba, P., & Wątróbski, J. (2016). The rank reversals paradox in management decisions: The comparison of the ahp and comet methods. In *Intelligent Decision Technologies 2016: Proceedings of the 8th KES International Conference on Intelligent Decision Technologies (KES-IDT 2016)–Part I* (pp. 181-191). Springer International Publishing. https://doi.org/10.1007/978-3-319-39630-9_15

[11] de Farias Aires, R. F., & Ferreira, L. (2019). A new approach to avoid rank reversal cases in the TOPSIS method. *Computers & Industrial Engineering*, *132*, 84-97. https://doi.org/10.1016/j.cie.2019.04.023

[12] Yang, W., & Wu, Y. (2020). A new improvement method to avoid rank reversal in VIKOR. *IEEE Access*, 8, 21261-21271. https://doi.org/10.1109/access.2020.2969681

[13] Majumdar, A., Tiwari, M. K., Agarwal, A., & Prajapat, K. (2021). A new case of rank reversal in analytic hierarchy process due to aggregation of cost and benefit criteria. *Operations Research Perspectives*, 8, 100185. https://doi.org/10.1016/j.orp.2021.100185

[14] Liu, X., & Ma, Y. (2021). A method to analyze the rank reversal problem in the ELECTRE II

method. *Omega*, 102, 102317. https://doi.org/10.1016/j.omega.2020.102317

[15] Tiwari, R. K., & Kumar, R. (2021). G-TOPSIS: a cloud service selection framework using Gaussian TOPSIS for rank reversal problem. The Journal of Supercomputing, 77, 523-562. https://doi.org/10.1007/s11227-020-03284-0

[16] Yang, B., Zhao, J., & Zhao, H. (2022). A robust method for avoiding rank reversal in the TOPSIS. *Computers & Industrial Engineering*, 174, 108776. https://doi.org/10.1016/j.cie.2022.108776

[17] Stevi´c, Z., Pamucar, D., Puška, A., & Chatterjee, P. (2020). Sustainable supplier selection in healthcare industries using a new MCDM method: Measurement of alternatives and ranking according to Compromise solution (MARCOS). *Computers & Industrial Engineering*, 140, 106231. https://doi.org/10.1016/j.cie.2019.106231

[18] Žižović, M., Pamučar, D., Albijanić, M., Chatterjee, P., & Pribićević, I. (2020). Eliminating rank reversal problem using a new multi-attribute model – the RAFSI method, *Mathematics*, 8(6), 1015. https://doi.org/10.3390/math8061015

[19] Pamučar, D., Žižović, M., Marinković, D., Doljanica, D., Jovanović, S. V., & Brzaković, P. (2020). Development of a multi-criteria model for sustainable reorganization of a healthcare system in an emergency situation caused by the COVID-19 pandemic, *Sustainability*, 12(18), 7504. https://doi.org/10.3390/su12187504

[20] Božanić, D., Milić, A., Tešić, D., Salabun, W., & Pamučar, D. (2021). D numbers–FUCOM–fuzzy RAFSI model for selecting the group of construction machines for enabling mobility. *Facta Universitatis, Series: Mechanical Engineering*, *19*(3), 447-471. https://doi.org/10.22190/fume210318047b

[21] Saluja, R. S., Mathew, M., & Singh, V. (2023). Improved proximity indexed value MCDM method for solving the rank reversal problem: A simulation-based approach. *Arabian Journal for Science and Engineering*, 48(9), 11679-11694. https://doi.org/10.1007/s13369-022-07553-3

[22] Tu, J., & Wu, Z. (2023). Analytic hierarchy process rank reversals: causes and solutions. *Annals of Operations Research*, 1-25. https://doi.org/10.1007/s10479-023-05278-6

[23] Dehshiri, S. S. H., & Firoozabadi, B. (2023). A new multi-criteria decision making approach based on wins in league to avoid rank reversal: A case study on prioritizing environmental deterioration strategies in arid urban areas. *Journal of Cleaner Production*, 383, 135438. https://doi.org/10.1016/j.jclepro.2022.135438

[24] Kizielewicz, B., Shekhovtsov, A., & Sałabun, W. (2021, June). A new approach to eliminate rank reversal in the mcda problems. *In International Conference on Computational Science* (pp. 338-351). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-77961-0_29

[25] Pamučar, D., Božanić, D., & Ranđelović, A. (2017). Multi-criteria decision making: An example of sensitivity analysis. *Serbian Journal of Management*,12(1), 1-27. https://doi.org/10.5937/sjm12-9464

[26] Bobar, Z., Božanić, D., Đurić-Atanasievski, K., & Pamučar, D. (2020). Ranking and Assessment of the Efficiency of Social Media using the Fuzzy AHP-Z Number Model - Fuzzy MABAC. *Acta Polytechnica Hungarica*, 17(3), 43-70. https://doi.org/10.12700/aph.17.3.2020.3.3

[27] Pamučar, D., Behzad, M., Božanić, D., & Behzad, M. (2021). Decision making to support sustainable energy policies corresponding to agriculture sector: Case study in Iran's Caspian Sea coastline. *Journal of Cleaner Production*, 292, 125302. https://doi.org/10.1016/j.jclepro.2020.125302

[28] Pamučar, D., Deveci, M., Canıtez, F., & Božanić, D. (2020). A fuzzy full consistency method-dombi bonferroni model for prioritizing transportation demand management measures. *Applied Soft Computing Journal*, 87, 105952. https://doi.org/10.1016/j.asoc.2019.105952

[29] Azeroual, O., Ershadi, M. J., Azizi, A., Banihashemi, M., & Abadi, R. E. (2021). Data quality strategy selection in CRIS: using a hybrid method of SWOT and BWM. *Informatica*, *45*(1). https://doi.org/10.31449/inf.v45i1.2995

[30] Ahmad, S., Khan, Z. A., Ali, M., & Asjad, M. (2023). A Novel Framework Based on Integration of Simulation Modelling and Mcdm Methods for Solving Fms Scheduling Problems. *Informatica*, *47*(4). https://doi.org/10.31449/inf.v47i4.3480

https://doi.org/10.31449/inf.v48i1.4475 Informatica **48** (2024) 31–44 **31**

A Deep Learning Model for Context Understanding in Recommendation Systems

Ngo Le Huy Hien[1], Luu Van Huy[2], Hoang Huu[2] Manh, Nguyen Van Hieu[2,*]
[1]Leeds Beckett University, Leeds, United Kingdom
[2]The University of Danang - University of Science and Technology, Danang, Vietnam
E-mail: n.hien2994@student.leedsbeckett.ac.uk, lvhuy@dut.udn.vn, hoanghuumanh54@gmail.com, nvhieuqt@dut.udn.vn
*Corresponding author

Keywords: recommendation system, context understanding, convolutional neural network, matrix factorization, deep learning, text processing

Received: November 2, 2022

Due to the robust growth in the amount of data and Internet users, there has been a significant rise in information overload, hindering timely access to user demand. While information retrieval systems, such as Google, Bing, and Altavista have partially addressed this challenge, prioritization and personalization of information have yet to be fully implemented. Therefore, recommendation systems are developed to resolve the issue by filtering and segmenting important information from an enormous volume of data based on different criteria such as preferences, interests, and user behaviors. By collecting data on users' interests and purchased products, the system can predict whether a particular user would enjoy an item, thus delivering an appropriate suggestion strategy. However, the increased number of Internet users and items has resulted in sparseness in increasingly vast datasets, reducing the performance of recommendation algorithms. Therefore, this study developed a model integrating Convolutional Neural Network (CNN) and Matrix Factorization (MF) to add extra product and user information, extract contexts, and add bias to the observed ratings in the training process, attempting to enhance the recommendation accuracy and context understanding. This approach can take advantage of CNN to efficiently capture an image's or document's local features, with the combination of MF to create relationships between 2 main entities, users and items. The proposed model obtained the highest RMSE of 0.93 when predicting favorable movies for 4,000 users, with an ability to learn complex contextual features and suggest more relevant content. The results are promising and can act as a reference for developing context understanding in recommendation systems, and future work may focus on optimizing the performance and developing more text-processing techniques.

Povzetek: Razvit je nov model globokega učenja, ki združuje konvolucijske nevronske mreže (CNN) in matrično faktorizacijo (MF) za izboljšanje natančnosti in razumevanja konteksta v priporočilnih sistemih.

1 Introduction

Recommendation systems (also known as recommender systems [1]) are algorithms designed to deliver suggestions for the most pertinent items to a certain user by filtering out information from a pool of data using various factors [2]. Normally, the recommendations pertain to different decision-making processes, including what movies to watch, books to read, products to buy, music to listen to, online news to read, or other products based on the desired industry [3]. Recommendation systems are substantially beneficial when a person has to pick an item from an overwhelming number of options provided by a service [4]. Netflix [5, 6] and Amazon [7], for example, employ recommendation systems to assist their consumers in choosing a suitable product or movie. The recommendation system handles a huge amount of data by filtering the most significant information from data given by a user and other criteria that correlate to their interests and preferences [3]. It determines the match between the user and the item, then infers the similarities among them for suggestions [4].

Recommendation systems have been proven to provide decent benefits to both users and supplied services. They were characterized from the standpoint of E-commerce as a tool that assists users in searching through a source of data associated with users' preferences [8]. Especially, under a complex and large accumulation of information, recommendation systems might showcase their advantage to enhance the quality of decision-making strategies [9]. This utility may result in decreasing transaction costs associated with locating and selecting products in the E-commerce sector [10]. Even in several companies, an efficient recommendation system can generate colossal revenue, and serve as a means to differ considerably from their rivals [11].

It is prevailing to apply recommendation systems when having insufficient personal knowledge or expertise with

the alternatives since the systems may support and enrich the social process of making decisions based on the [9]. For instance, recommender systems are utilized in scientific libraries to assist users by enabling them to go beyond catalog searches [3]. Therefore, these types of systems can address the information overloading issue, which is commonly encountered in recent years [12], by operating accurate and efficient recommendation algorithms to deliver individualized, distinctive service and content suggestions [13].

There are several recent techniques have been developed for constructing recommendation systems, including collaborative filtering, content-based filtering, and hybrid filtering [14]. The most developed and widely used technique is collaborative filtering, which finds users who own similar preferences and utilizes their views to suggest to another user [15]. Contrarily, the content-based approach links user attributes to content resources. It hence often disregards inputs from other users and delivers recommendations solely based on the information provided by the user [16]. Notwithstanding, hybrid filtering can improve the effectiveness and accuracy of recommendation systems, by combining two or more filtering approaches in various methods. It balances out the corresponding deficiencies of different filtering techniques while using their respective strengths. The methods can be weighted, switching, cascade, mixed, feature-combination, feature-augmented, or meta-level hybrid depending on the operations of the combined techniques [17].

However, the aforementioned filtering techniques retain a few drawbacks, notwithstanding their success. Overspecialization, limited content analysis, and data scarcity are a few issues with content-based filtering algorithms. In addition, cold-start, scalability, and sparsity issues remain to exist in collaborative techniques, reducing the effectiveness of recommendations [18]. It can be seen that the common problem with such filtering techniques is data sparsity. It is because of the explosive growth in the number of users and items in the fast-growing service market, which increased the sparseness of product review data from users [19]. This sparseness diminishes the prediction accuracy of traditional filtering techniques [20].

In order to address the above data sparseness limitation, in this paper, different factors have been added to the recommendation system such as user information, user interactions, and product description documents instead of only using review data, attempting to enhance the accuracy of the system. Moreover, traditional information retrieval methods mostly use the bag-of-words model, which ignores the context information of the text document [21]. To address this, the study proposed a model to apply a Convolutional Neural Network (CNN) in the recommendation system to better understand the text document. Owing to the fact that CNN can efficiently capture local features of documents or images through local receptive fields, shared weights, and pooling [22]. However, since CNN is primarily used in classification problems, this study proposed an approach to integrate it into Matrix Factorization (MF) to define relationships between users and items. The combination makes it possible to take full advantage of both CNN and MF [23]. Inspired by the work of Donghyun and colleagues [24], this study aims to enhance the model by adding bias for the training more objectively; and supplementing extra information from description documents of both users and items. The research outcomes are promising and can be used as a reference for further developing context understanding in recommendation systems.

2 Literature review

2.1 The development of recommendation systems

Recommendation systems have gained considerable interest since their initial introduction and have been widely utilized in various sectors, including e-commerce [8], e-library [31], e-tourism [32], education [33], news [34], information retrieval, and digital content services [35]. Table 1 indicates the eminent applications of recommendation systems in different domains.

Item Type	Recommendation Systems
E-commerce Products	Amazon [7], eBay [36], Shopify, Flipkart [37]
Videos	Netflix [5], YouTube [38], Dailymotion, Hulu [39], MovieLens, Nanocrowd, Jinni [40]
Online News	Google News, Yahoo! News, BBC, New York Times [41], Findory [42], Digg, Zite [43]
Music	Spotify, Apple Music, Amazon Music, Soundcloud, Pandora, Mufin [44]
Social Networking Contents	Facebook, TikTok, Twitter, LinkedIn, Instagram [45]

Table 1: Current eminent recommendation systems in different domains

Leading e-commerce company Amazon applies a collaborative filtering technique to address scalability challenges by offline generating a table of related items using an item-to-item matrix [7]. To enhance suggestion quality, it employs topic diversity algorithms. Following that, the algorithm suggests items that are comparable online based on the customers' past purchases [46]. Thanks to this, items that are not among the shop's 100,000 best-selling items have helped Amazon gain 20% to 40% of sales [47].

Netflix Recommendation Engine uses algorithms that filter its contents using each user's unique profile. The system uses 1,300 clusters based on user choices to filter over 3,000 titles at once [48]. Cinematch, a proprietary recommendation system used by Netflix, has a root mean squared error (RMSE) of 0.9525. In 2009, Netflix held a compe-

tition called 'Netflix Prize', attempting to produce a recommender system that outperformed its algorithm, with a million-dollar prize for the winner [6]. For that reason, 60% of Netflix's DVDs are rented thanks to recommendation algorithms, and 47% of North Americans prefer Netflix with a retention rate of 93%. [49]

TikTok, one of the most popular and rapidly expanding social media networks in the world, has its secret strength as a unique recommendation system for discovering and distributing content [50]. TikTok blends videos from newbies and celebrities in the 'For You' feed, rewards high-quality creative content based on page views, and encourages emerging users to share videos with other viewers. Therefore, every user has the opportunity to become famous on the platform, regardless of their fanbase or level of popularity. High-quality creative work may be easily shared thanks to TikTok's recommendation system, which regularly suggests videos to individuals with similar interests [51].

It can be seen that recommendation systems have been applied in numerous domains and have helped businesses not only generate colossal revenue but also serve as a means to differ considerably from their competitors.

2.2 Related works

For a system to deliver its customers reliable and helpful recommendations, the usage of accurate and efficient recommendation algorithms is essential. Therefore, it is critical to clarify the advantages and limitations of various recommendation approaches. There are several recent techniques for constructing recommendation systems, which are content-based filtering, collaborative filtering, and hybrid filtering, as depicted in Figure 2.1 [14].

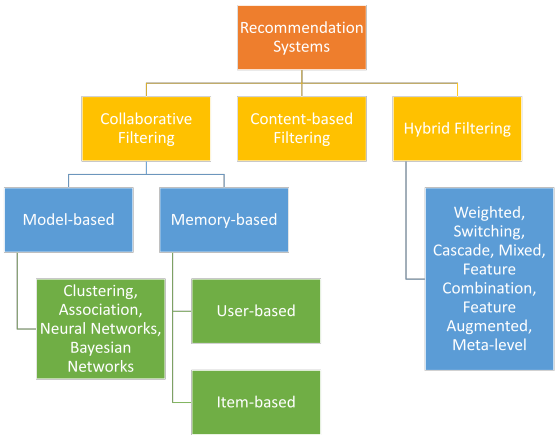


Figure 2.1: Different recommendation filtering techniques.

First of all, collaborative filtering is a technique to find users who own similar preferences and utilize their views to suggest to another user. It has become the most developed and widely used filtering technique in recommendation systems [15]. Collaborative filtering is prominent when the content cannot be accurately and simply represented by metadata, like music and movies [25]. This technique aims to build a database of user preferences for things called a user-item matrix. By comparing the commonalities between users' profiles, it connects people with shared interests and preferences in a so-called neighborhood to provide suggestions. The user then receives suggestions for unseen items that received favorable reviews from others in the neighborhood [26]. The suggestions can be in the form of recommendations or predictions. A recommendation is a list of the top items that the user would enjoy the best, whereas a prediction is an estimated favorable score of an item for the target user [27].

In contrast, content-based filtering links user characteristics to the attributes of items. It hence often disregards inputs from other users and delivers recommendations solely based on the information provided by the user [16]. This filtering technique is significant when the suggested documents can be metadata-represented, which could be books, news, and web pages. Content-based filtering extracts characteristics from the content of items previously rated by different users and then merges them into a training set. From there, the system recommends items that are greatly related to a user's favorability to them. The technique can deliver recommendations even when a user never offered ratings before [28]. As a result, users may receive suggestions without disclosing their profiles, ensuring their privacy. Furthermore, content-based filtering could handle circumstances in which different users might not have identical items, but only similar items that shared common characteristics [29].

Nevertheless, by integrating two or more filtering algorithms diversely, hybrid filtering can increase the efficacy and accuracy of recommendation systems. It compensates for the inadequacies of various filtering systems while maximizing their unique strengths [17]. Depending on the operations of the combined approaches, the methods can be weighted, switching, cascade, mixed, feature-combination, feature-augmented, or meta-level hybrid. Collaborative filtering and content-based filtering approaches can be used differently before being combined. Thereafter a unified model was formed that encompasses both content-based and collaborative filtering capabilities. Consequently, the data sparsity and cold-start issues could be solved by merging item ratings, characteristics, and demographic information [30].

Despite the success of the aforementioned filtering techniques, they come with certain drawbacks. Issues like over-specialization, limited content analysis, and data scarcity pose challenges for content-based filtering algorithms. Collaborative techniques also grapple with problems such as cold-start, scalability, and sparsity, ultimately hampering the effectiveness of recommendations [18]. A common underlying problem in these filtering techniques is data sparsity, which stems from the rapid expansion of users and items in the dynamic service market. This proliferation has increased the sparseness of product review data from users,

leading to a decline in the prediction accuracy of traditional filtering methods [19, 20].

3 Methodology

To overcome the above limitation of data sparseness, this study aims to develop a model integrating Convolutional Neural Network (CNN) and Matrix Factorization (MF) to add extra product and user information and extract contexts before training, attempting to enhance the recommendation accuracy. In this section, the architecture of CNN and MF is briefly presented.

3.1 Convolutional neural network

Convolutional Neural Network (CNN/ ConvNet - proposed by Fukushima Kunihiko) is a variant of a feedforward neural network. Convolutional Neural Networks represent significant progress and influence in the development of Deep Learning [52]. Many CNN variations, including VGGNet, MobileNet, Inceptions, ResNet, RegNet, DenseNet, and EfficientNet have been developed robustly. These variants emphasize different facets of accuracy, efficiency, and scalability. The field of computer vision is mostly dominated by ConvNets models [53].

The organization of the visual cortex and the human brain's neural network both had an influence on CNN's architecture [54]. Individual neurons can only respond to stimuli in the restricted visual field region known as the Receptive Field. A succession of similar fields that overlap encompasses the entire visual field [55]. There are four main types of layers for a convolutional neural network: the convolutional layer (to extract local features), the pooling layer (representing data of the previous layer in a more concise form, i.e., select only the typical features with the highest scores through activation functions), the ReLU correction layer and the fully-connected layer [56], as indicated in Figure 3.1.

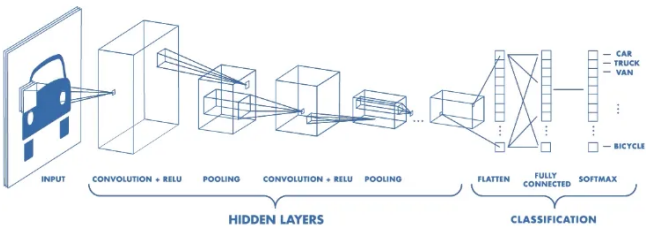


Figure 3.1: The Architecture of CNN. [57]

As shown in Figure 3.1, a CNN normally consists of two main components:

1. Hidden layers or feature extraction layers: in this component, the network will perform a series of convolution and pooling computations to detect features. For example, if an image of a zebra is inputted, in this component, the network will recognize its stripes, two ears, and four legs.
2. Classification: in this component, a class with full associations will act as a classifier of previously extracted features.

The CNN model in natural language processing often considers the local context aspect of the corpus [58]. These contexts are extracted through filters or the kernel and aggregated at the pooling layer [59]. However, since the CNN model is often used for classification problems, it is challenging to apply CNN directly to the recommendation system.

3.2 Matrix factorization

Matrix Factorization (MF) is a commonly used collaborative filtering method in recommendation systems proposed by Simon Funk [60]. Matrix Factorization decomposes the performance evaluation matrix into a product of two matrices U and V. While U represents the correlation between users, V represents the relationship between items, described in Figure 3.2.

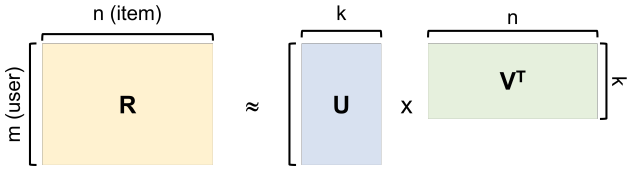


Figure 3.2: The concept of matrix factorization.

As shown in Figure 3.2, the Matrix Factorization technique involves decomposing a large matrix R into two smaller matrices U and V, such that the reconstruction of R from these smaller matrices is as accurate as possible, i.e., $R \approx U \times V^T$.

In which:

- U is a matrix of size $m \times k$, where each row represents k latent factors describing user m.
- V is a matrix of size $n \times k$, with each row being a vector comprising k latent factors describing item i (typically $k << m$ and $k << n$).
- V^T denotes the transpose matrix of V.

The key challenge in the MF technique lies in determining the values of the two parameters (matrices) U and V. These parameters are identified by optimizing an objective function. In the context of rating prediction, the objective function, denoted as L, is expounded upon in the subsequent section.

The concept of latent features that reflect the relationship between objects and users is fundamental in Matrix Factorization for Recommendation Systems. For example, in a

movie recommendation system, the latent features can be criminal, political, action, comedy, etc.; may also be a combination of these features or anything that may not need to be named [61]. Each item can bring some latent features to some extent corresponding to the coefficients in its vector v. The higher the coefficient, the higher the possibility of having that feature. Similarly, each user will also tend to prefer certain latent features described by the coefficients in its vector u. The higher the coefficient, the more likely users prefer the movies with that latent feature. The value of the expression uv will be high if the corresponding components of v and u are both high. This means that the item has latent features that the user likes, thus the system recommends this item to that user.

Assume that there are m users and n items, with a user-item rating matrix R, in which R $\in R^{m\times n}$. In Matrix Factorization, latent models of user i and item j can be represented as k-dimensional models, $u_i \in R_k$ and $v_j \in R_k$. The observed rating r_{ij} of user i on item j is calculated by the inner product of respective latent models of user i and item j. A common approach to training latent models is minimizing a loss function L, which comprises sum-of-squared-error terms among the observed ratings and the predicted ratings. Therefore, the loss function in this situation can be expressed as:

$$L = \sum_{i}^{m}\sum_{j}^{n} I_{ij}(r_{ij} - u_i^T v_j)^2 + \\ + \lambda_u \sum_{i}^{m} ||u_i||^2 + \lambda_v \sum_{j}^{n} ||v_j||^2 \quad (1)$$

in which:

- I_{ij} is an indicator function that becomes 1 if user i rated item j and equals 0 if not.
- λ denotes the regularization term. When λ is excessively large, the model tends to underfit the data; conversely, if λ is overly small, the model may become overly complex, leading to overfitting. The fine-tuning of the λ value is a crucial aspect in optimizing the performance of the MF model.
- λ_u is the regularization parameter associated with user vectors u_i. Regularization serves as a technique to prevent overfitting in machine learning models. It is applied in the loss function by penalizing the squared Euclidean norm (L2 norm) of user vectors. This regularization constrains user vectors from becoming excessively large during the training process, mitigating the risk of overfitting to the training data and potentially enhancing the model's generalization ability to unseen data.
- Similarly, λ_v represents the regularization parameter for item vectors u_i. This regularization parameter is essential for preventing overfitting in the context of item vectors, analogous to its role in the regularization of user vectors (λ_u).

4 Proposed model

4.1 General architecture

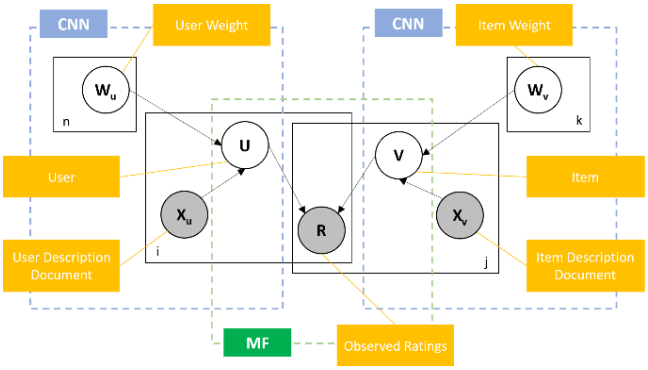


Figure 4.1: General Architecture.

As depicted in Figure 4.1, MF (Matrix Factorization, in the green box) is the decomposition of the observed rating matrix R of user-item into two matrices with lower weights. Matrix U represents the relationship between users, while matrix V represents the correlation between the items. The model aims to add product features to the recommendation system. CNN in natural language processing often considers the local context aspect of the text. Therefore, CNN is used to extract features with local contexts of the user and item description sets and then add the information to matrix U (matrix containing vectors describing characteristics of the user, such as age, gender, and occupation) and V (matrix containing vectors describing features of the item) respectively. This technique can complement and clarify the properties of the vectors in matrix U and V.

In Figure 4.1, X_u and X_v act as the set of documents describing the user and item respectively, and W_u and W_v are the weights of the CNN model for the user and item correspondingly. The outputs of the CNN are latent feature vectors of those input documents. The difference between those latent feature vectors with matrix U and V is the integration between CNN and MF in fully analyzing descriptive documents and evaluation data.

This research employs a Convolutional Neural Network (CNN) to extract local features from embedding vectors, consisting of the following layers:

- Input layer: receives embedding vectors describing product narratives with a length of 100 tokens.
- The token and position embedding layer comprises two main components:
  - Token embedding: transforms each word in the product narrative into a dense vector representation. This representation captures the semantic meaning of the word as well as its relationships with other words in the vocabulary.

  - Position embedding: encodes the position of each word in the product narrative into a vector representation. This representation helps the model understand the context of each word and its relationships with other words in the product narrative.

- The output of the embedding layer, comprising token and position information, is a sequence of embedding vectors, where each embedding vector represents a token (word) in the product narrative and incorporates its position in the product narrative.

- Subsequently, the embedding vectors are fed into a CNN layer, consisting of fundamental layers such as Convolutional, pooling, and incorporating dropout techniques to extract more complex features from the text. The CNN layer learns to identify patterns and relationships among the embedding vectors, which are then utilized to predict user rankings for different products.

Details of the CNN model architecture are illustrated in Figure 4.2.

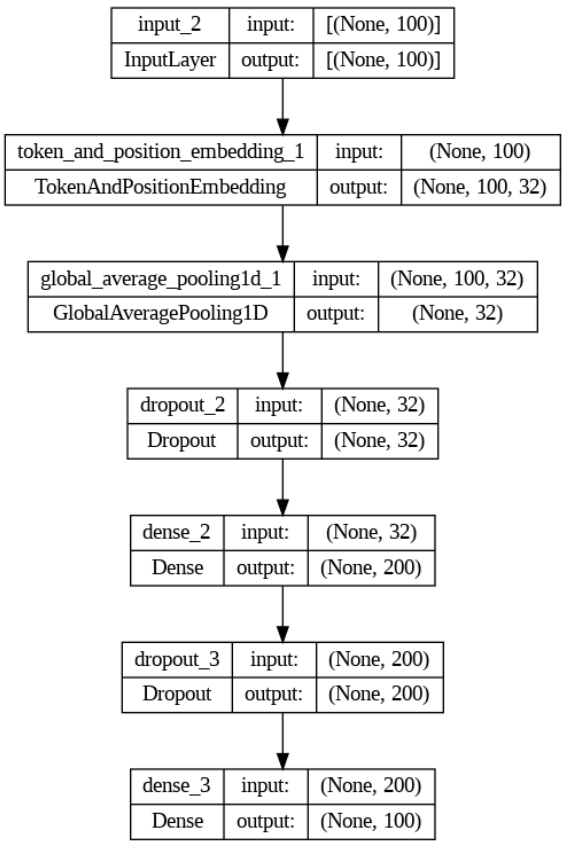


Figure 4.2: The Used CNN Architecture.

The rationale behind the utilization of this CNN structure is predicated upon the model's input being comprised of embedded vectors used to depict products, typically of relatively modest dimensionality (dim = 100). Consequently, a CNN architecture with fundamental layers, as expounded above, is employed in this study to extract local features from the embedded vector.

4.2 Adding bias

As mentioned in Section 3.2, the observed rating r_ij of user i on item j is calculated by the inner-product of respective latent models of user i and item j, which can be indicated as:

$$r_{ij} \approx \hat{r}_{ij} = u_i^T v_j \tag{2}$$

However, to avoid overfitting issues, this study adds bias to the observed rating:

$$\hat{r}_{ij} = u_i^T v_j + d_i + b_j \tag{3}$$

in which:

- d_i is a coefficient representing the pleasantness of user i. The higher the coefficient, the better the user i tends to rate the products.

- b_j is a coefficient illustrating product quality, the higher the coefficient. The more users tend to rate that product better.

4.3 Loss function

From there, the loss function now can be depicted as:

$$\begin{aligned} L(U,V,W) = &\sum_{i}^{m}\sum_{j}^{n}\frac{I_{ij}}{2}(r_{ij}-\hat{r}_{i_j})^2 \\ &+ \frac{\lambda_U}{2}\sum_{j}^{m}||v_j - cnn(Wv, Xv_j)||_2 \\ &+ \frac{\lambda_{Wu}}{2}\sum_{k}^{|w_{u_k}|}||w_{u_k}||_2 \\ &+ \frac{\lambda_{Wv}}{2}\sum_{n}^{|w_{v_n}|}||w_{v_n}||_2 \end{aligned} \tag{4}$$

The loss function is minimal when the derivative of the above equation is 0. The loss function uses coordinate descent to find the function that updates u and v. This optimizes having to iterate over and over one variable while correcting the others.

Assuming W_u, W_j, and V (or U) are constants, the above equation becomes a quadratic function with respect to U (or V). Therefore:

$$\begin{aligned} u_i &\leftarrow (VIiV^T + \lambda_u I_k)^{-1}(VR_i + \lambda_u nn(Wu, Xu_i)) \\ d_i &\leftarrow (r_{ij} - u_i^T v_j - b_j) \\ v_j &\leftarrow (UI_jU^T + \lambda VI_k)^{-1}(UR_j + \lambda_v cnn(Wv, Xv_j)) \\ b_j &\leftarrow (r_{ij} - u_i^T v_j - d_j) \end{aligned} \tag{5}$$

W_u and W_j will be updated through the backpropagation of the CNN.

5 Experiment and results

5.1 Dataset

This research utilizes Movielens 1M [62], a user's movie review dataset, which contains 6000 users and 4000 movies. It was released in 2003 with a rating rate of 4.6%. This dataset includes:

- Movie information: id, movie name, genre, release year;
- User information: gender, age, occupation;
- List of user reviews corresponding to movies (1 million samples).

The training was conducted on Google Colab with the configuration specified in Table 2.

Type	Specifications
CPU	Intel(R) Xeon(R) CPU @ 2.20GHz
Number of CPUs	2
RAM	12.0 GB
Memory	108.0 GB [44]
GPU	Nvidia Tesla K80

Table 2: Device Specification.

5.2 Dataset pre-processing

The input of the model is the item description document set. Particularly in this experiment, it contains 4000 movie description texts corresponding to 4000 movies in the dataset. A sample data used in the dataset is presented in Figure 5.1.

The user quantity within the dataset was partitioned for experimental purposes, comprising subsets of 1000 users, 2000 users, and so forth. This approach facilitated the evaluation of the model across varying dataset scales, allowing an examination of potential impacts. Statistics of the number of users, items, and ratings are presented in Table 3 for reference and analysis.

From the description text of the movies, latent features were extracted to add to the training model. The input text set of movie descriptions has been through different preprocessing steps, as shown in Figure 5.2, starting with cleaning to remove the noise in the text like HTML tags. The next step is word splitting, meaning splitting the sentences into single words. Those words were then normalized to the same font and type. And finally, stopwords will be eliminated, which are words that appear frequently but contain trivial meanings, such as 'is', 'that', or 'this' in English. A sample of a movie description after the pre-processing process is presented in Figure 5.3.

movie id: 39

movie name: Toy Story

description: A little boy named Andy loves to be in his room, playing with his toys, especially his doll named "Woody". But, what do the toys do when Andy is not with them, they come to life. Woody believes that his life (as a toy) is good. However, he must worry about Andy's family moving, and what Woody does not know is about Andy's birthday party. Woody does not realize that Andy's mother gave him an action figure known as Buzz Lightyear, who does not believe that he is a toy, and quickly becomes Andy's new favorite toy. Woody, who is now consumed with jealousy, tries to get rid of Buzz. Then, both Woody and Buzz are now lost. They must find a way to get back to Andy before he moves without them, but they will have to pass through a ruthless toy killer, Sid Phillips.

user rating: [5. 0. 0. ... 0. 5. 4.]

vector movie: [4289, 940, 4912, …, 4173, 396, 7352]

Figure 5.1: Sample data used in the dataset.

Number of Users	Number of Items	Number of Ratings
1000	3280	154212
2001	3452	337262
3001	3477	484775
4001	3505	660411
5001	3532	826438

Table 3: Statistics of the number of users, items, and ratings.

5.3 Training

The dataset was divided into 3 subsets, which are training, validation, and testing sets. Corresponding to each user, the number of user reviews will be divided by the ratio of 80% for the training set, 10% for the test set, and 10% for the validation set.

λU	λV	Dimension	Train. Loss	Val. Loss	Test. Loss
10	40	500	0.76	0.88	0.88
10	60	500	0.77	0.88	0.88
10	50	50	0.78	0.89	0.88
10	10	50	0.7	0.90	0.90
100	10	100	0.87	0.90	0.90
50	100	100	0.88	0.91	0.91

Table 4: Loss results in different hyperparameters.

From Table 4, it can be seen that the ratio between λU and λV significantly affects the results. If λU is much larger than λV, meaning a higher priority is given to learning the parameters of U, a good result could not be attained. While the goal of the problem is to use data from the item, it

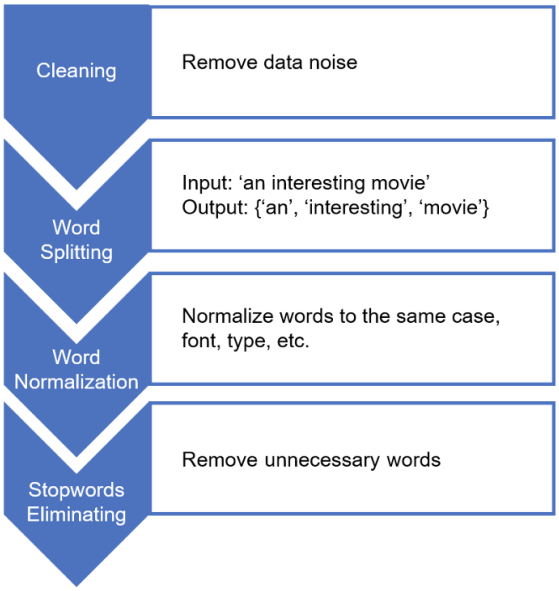


Figure 5.2: Text Pre-processing Process.

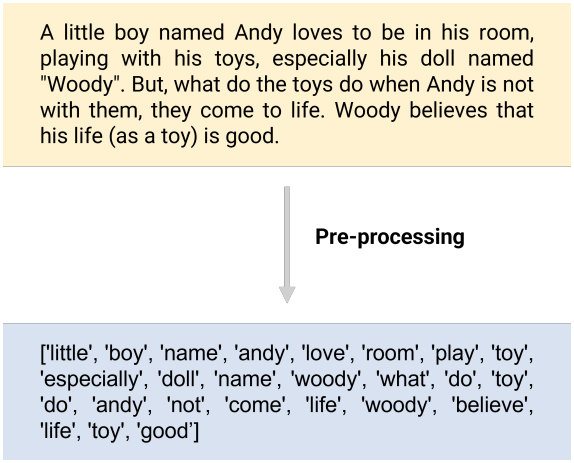


Figure 5.3: Sample Movie Description Text After Pre-processing Process.

is better to give preference to λV, making it slightly higher than λU, to obtain a better result.

5.4 Evaluation

To evaluate the model's general performance, this study uses Root-mean-square error (RMSE) and mean-square error (MSE), which represent the dispersion of the predicted data relative to the actual data.

$$RMSE = \sqrt{\frac{\sum_i^m (\hat{r}_i - r_i)^2)}{m}} \tag{6}$$

$$MSE = \frac{1}{m}\sum_i^m (\widehat{r_i} - r_i)^2 \tag{7}$$

The RMSE function evaluates the results after each iteration for all 3 training, validation, and testing sets. The model training process was repeated for about 100-200 iterations until the loss function gave the smallest value on the validating and testing sets. RMSE results of the model on the training, validating, and testing sets are illustrated in Figure 5.4.

As can be seen from Figure 5.4, in the 8^{th} iteration, the results began to deteriorate, and the validation RMSE increased while the training RMSE continued to be overfitting. Therefore, the result was obtained in the 8th iteration. The evaluation of results for the entire data is shown in Table 5.

Table 5 evaluates the proposed model using two metrics: Root Mean Square Error (RMSE) and Mean Squared Error (MSE). These metrics gauge the disparity between predicted rankings and actual rankings. Based on the tabulated data, it is evident that the proposed model demonstrates strong performance on the test set, yielding an RMSE of

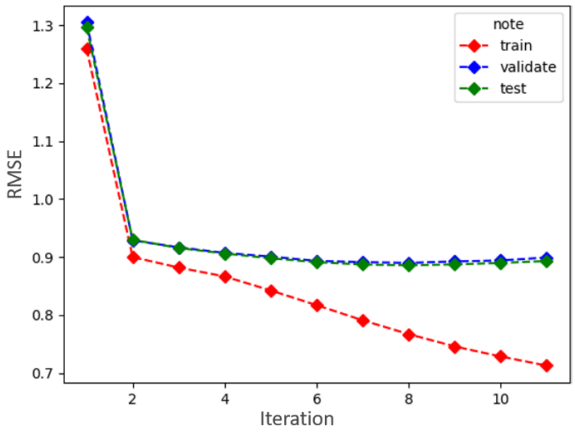


Figure 5.4: Plotting RMSE results.

0.89 and MSE of 0.78. This signifies the model's ability to accurately predict user rankings for diverse products.

The RMSE and MSE values across all three sets—training, validation, and testing—indicate that the model exhibits robust predictive capabilities on the test dataset. Both RMSE and MSE values remain stable, with minimal deviation observed between the validation and test datasets. This suggests that the model does not encounter issues related to overfitting or underfitting.

To determine how the results correlate with the user amount, a comparison of RMSE with different numbers of users is presented in Table 6.

Evaluation metric	Training	Validation	Testing
RMSE	0.76695	0.88974	0.88563
MSE	0.58821	0.79163	0.78435

Table 5: Result Evaluation in different metrics.

No. of users	Train. RMSE	Val. RMSE	Test. RMSE	Exec. time (s)	Train. time (s)
1000	0.87865	0.91478	0.90093	0.0062	110
2000	0.87205	0.91791	0.93004	0.0052	75
3000	0.87168	0.91896	0.92671	0.0053	91
4000	0.86955	0.91383	0.92973	0.005	159
5000	0.87865	0.91478	0.90093	0.0062	110

Table 6: Comparison of the RMSE with different numbers of users.

Table 6 demonstrates when increasing the number of users in the dataset, from 1000 to 5000, the accuracy increases, but with a longer convergence time. Therefore, in order to produce appropriate recommendations, recommendation system applications need to employ a large dataset.

5.5 Utilizing the training results

The results obtained after training the model are 2 matrices U and V. An evaluation matrix Y[i,j] can be generated as:

$$Y[i,j] = U[i] * V[j]^T \tag{8}$$

in which:
- i: i-th user
- j: j-th item

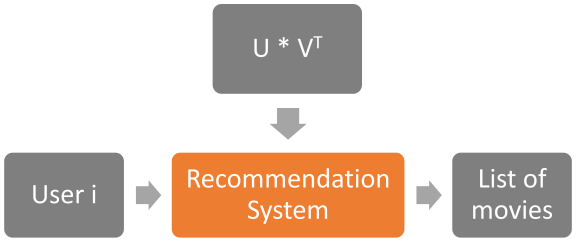


Figure 5.5: Using the training results for creating recommendations.

As depicted in Figure 5.5, the evaluation matrix can be applied in the recommendation system for further usage, which outputs a list of recommended movies for the i^{th} user.

6 Conclusions and future work

In this research, a deep learning model for recommendation systems is proposed by integrating Convolutional Neural Network and Matrix Factorization to add extra information and extract contexts before training, attempting to enhance recommendation accuracy and context understanding. Despite substantial previous efforts [21, 63, 64], this study adds additional information on both user and item description documents and applied Convolutional Neural Networks to efficiently capture their local features. Furthermore, this research adds bias to the observed ratings to avoid overfitting issues and uses Matrix Factorization to create relationships between users and items. The proposed model can be further used as a benchmark for developing context comprehension in recommendation systems, hence delivering more relevant recommendations for users.

It is observed that the model obtained a very good RMSE of 0.89 in the testing set, which means the model can relatively predict favorable movies of users accurately. Testing on different amounts of users reveals that the more users, the higher the accuracy, but the longer the convergence time. It is noted that this study subdivides the dataset to assess each subset independently, as opposed to providing a comprehensive evaluation of the entire dataset. Consequently, the rationale for refraining from comparing with other models stems from the divergence in data partitioning strategies. Hence, the evaluation process becomes inherently untenable due to the dissimilarity in data distribution methodologies across models.

Future research may aim to overcome the scant user information (e.g., hobbies, location, marital status) by looking for a large dataset with more user information, including more features in the user description documents, leading to a higher impact on the prediction. Moreover, the proposed model could be developed further by swapping out Matrix Factorization with more efficient techniques, such as singular value decomposition (SVD).

Acknowledgement

This research was funded and implemented for the Rising-Star project of the University of Science and Technology, The University of Danang, Vietnam.

References

[1] L. Lü, M. Medo, C. H. Yeung, Y.C. Zhang, Z.K. Zhang and T. Zhou (2012) Recommender systems, *Physics reports* , vol. 519, no. 1, pp. 1-49. https://doi.org/10.1016/j.physrep.2012.02.006

[2] H. Ko, S. Lee, Y. Park, A. Choi (2022) A survey of recommendation systems: recommendation models, techniques, and application fields, *Electronics*, vol. 11, no. 1, p. 141. https://doi.org/10.3390/electronics11010141

[3] F. O. Isinkaye, Y. O. Folajimi and B. A. Ojokoh (2015) Recommendation systems: Principles, methods and evaluation, *Egyptian informatics journal*, vol. 16, no. 3, pp. 261-273. https://doi.org/10.1016/j.eij.2015.06.005

[4] G. Shani and A. Gunawardana (2011) Recommender systems handbook, *Recommender systems hand-*

book, pp. 257-297. https://doi.org/10.1007/978-0-387-85820-3_8

[5] Gomez-Uribe, C. A. and N. Hunt (2015) The netflix recommender system: Algorithms, business value, and innovation, *ACM Transactions on Management Information Systems (TMIS)*, vol. 6, no. 4, pp. 1-19. https://doi.org/10.1145/2843948

[6] J. Bennett, S. Lanning (2007) The netflix prize, *in KDD cup and workshop*, p.35.

[7] B. Smith, G. Linden (2017) Two decades of recommender systems at Amazon. com, *IEEE internet computing*, , vol. 21, no. 3, pp. 12-18. https://doi.org/10.1109/MIC.2017.72

[8] S. S. Li, E. Karahanna (2015) Online recommendation systems in a B2C E-commerce context: a review and future directions, *Journal of the association for information systems*, vol. 16, no. 2, p. 2. https://doi.org/10.17705/1jais.00389

[9] K. Al Fararni, B. Aghoutane, J. Riffi, A. Sabri and A. Yahyaouy (2020) Comparative study on approaches of recommendation systems, *Embedded Systems and Artificial Intelligence*, pp. 753-764. https://doi.org/10.1007/978-981-15-0947-6_72

[10] C.-S. Juan-Pedro, I. Ramos-de-Luna, E. Carvajal-Trujillo and Á. F. Villarejo-Ramos (2020) Online recommendation systems: Factors influencing use in e-commerce, *Sustainability Publisher*, pp. 01-15. https://doi.org/10.3390/su12218888

[11] L. Ebrahimi, V. R. Mirabi, M. H. Ranjbar and E. H. Pour (2019) A customer loyalty model for e-commerce recommendation systems, *Journal of Information Knowledge Management* , vol. 18, no. 3 pp. 12-18. https://doi.org/10.1142/S0219649219500369

[12] Z. Wang, X. Yu, N. Feng and Z. Wang (2014) An improved collaborative movie recommendation system using computational intelligence, *Journal of Visual Languages Computing*, vol. 25, no. 6, pp. 667-675. https://doi.org/10.1016/j.jvlc.2014.09.011

[13] M. Robillard, R. Walker and T. Zimmermann (2009) Recommendation systems for software engineering, *IEEE software*, vol. 27, no. 4, pp. 80-86. https://doi.org/10.1109/MS.2009.161

[14] L. Shah, H. Gaudani and P. Balani (2016) Survey on recommendation system, *International Journal of Computer Applications*, vol. 137, no. 7, pp. 43-49. https://doi.org/10.5120/ijca2016908821

[15] S. K. Raghuwanshi and R. K. Pateriya(2019) Collaborative filtering techniques in recommendation systems, *Data, Engineering and Applications*, pp. 11-21. https://doi.org/10.1007/978-981-13-6347-4_2

[16] S. Eliyas and P. Ranjana (2022) Recommendation Systems: Content-Based Filtering vs Collaborative Filtering, *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, pp. 1360-1365. https://doi.org/10.1109/ICACITE53722.2022.9823730

[17] P. B. Thorat, R. M. Goudar and S. Barve (2015) Survey on collaborative filtering, content-based filtering and hybrid recommendation system, *International Journal of Computer Applications*, vol. 110, no. 4, pp. 31-36. https://doi.org/10.5120/19308-0760

[18] G. Suganeshwari and S. P. S. Ibrahim (2016) A survey on collaborative filtering based recommendation system *3rd international symposium on big data and cloud computing challenges* , pp. 503-518. https://doi.org/10.1007/978-3-319-30348-2_42

[19] N. L. H. Hien, T. Q. Tien and N. V. Hieu (2020) Web crawler: Design and implementation for extracting article-like contents,*Cybernetics and Physics*, vol. 9, no. 3, pp. 144-151. https://doi.org/10.35470/2226-4116-2020-9-3-144-151

[20] P. Kumar and R. S. Thakur (2018) Recommendation system techniques and related issues: a survey *International Journal of Information Technology*, vol. 10, no. 4, pp. 495-501. https://doi.org/10.1007/s41870-018-0138-8

[21] S. Bhattacharya and L. Ankit (2019) Movie recommendation system using bag of words and scikit-learn, *Int J Eng Appl Sci Technol*, vol. 4, pp. 526-528. http://doi.org/10.33564/IJEAST.2019.v04i05.076

[22] M. Sheikh Fathollahi and F. Razzazi(2021) Music similarity measurement and recommendation system using convolutional neural networks,*International Journal of Multimedia Information Retrieval*, vol. 10, no. 1, pp. 43-53. https://doi.org/10.1007/s13735-021-00206-5

[23] A. F. Agarap(2017) An architecture combining convolutional neural network (CNN) and support vector machine (SVM) for image classification,*arXiv preprint*, p. arXiv:1712.03541. https://doi.org/10.48550/arXiv.1712.03541

[24] D. Kim, C. Park, J. Oh, S. Lee and H. Yu (2016) Convolutional matrix factorization for document context-aware recommendation,*Proceedings of the 10th ACM*

Conference on Recommender Systems " pp. 233-240. https://doi.org/10.1145/2959100.2959165

[25] B. S. Neysiani, N. Soltani, R. Mofidi and M. H. Nadimi-Shahraki (2019) Improve performance of association rule-based collaborative filtering recommendation systems using genetic algorithm,*International Journal of Information Technology and Computer Science*, vol. 11, no. 2, pp. 48-55. 10.5815/ijitcs.2019.02.06

[26] Z. Cui, X. Xu, X. U. E. Fei, X. Cai, Y. Cao, W. Zhang and J. Chen(2020), Personalized recommendation system based on collaborative filtering for IoT scenarios,*IEEE Transactions on Services Computing*, vol. 13, no. 4, pp. 685-695. https://doi.org/10.1109/TSC.2020.2964552

[27] W. Zhang and J. Wang(2018), Content-bootstrapped collaborative filtering for medical article recommendations,*2018 IEEE International Conference on Bioinformatics*, pp. 1184-1188. https://doi.org/10.1109/BIBM.2018.8621180

[28] U. Javed, K. Shaukat, I. A. Hameed, F. Iqbal, T. M. Alam and S. Luo (2021), A review of content-based and context-based recommendation systems,*International Journal of Emerging Technologies in Learning (iJET)*, vol. 16, no. 3, pp. 274-306. http://doi.org/10.3991/ijet.v16i03.18851

[29] J. Son and S. B. Kim (2017) Content-based filtering for recommendation systems using multiattribute networks,*Expert Systems with Applications*, vol. 89, pp. 404-41. https://doi.org/10.1016/j.eswa.2017.08.008

[30] A. A. Kardan and M. Ebrahimi(2013) A novel approach to hybrid recommendation systems based on association rules mining for content recommendation in asynchronous discussion groups, *Information Sciences*, vol. 219, pp. 93-110. https://doi.org/10.1016/j.ins.2012.07.011

[31] C. Tsai and M. Chen(2008) Using adaptive resonance theory and data□mining techniques for materials recommendation based on the e□library environment,*The Electronic Library*, vol. 26, no. 3, pp. 287-302. https://doi.org/10.1108/02640470810879455

[32] R. Khan, H. Ur, C. K. Lim, M. F. Ahmed, K. L. Tan and M. B. Mokhtar(2021) Systematic review of contextual suggestion and recommendation systems for sustainable e-tourism,*Sustainability*, vol. 13, no. 15, p. 8141. https://doi.org/10.3390/su13158141

[33] A. C. Rivera, M. Tapia-Leon and S. Lujan-Mora (2018) Recommendation systems in education: a systematic mapping study,*International Conference on Information Technology Systems*, pp. 937-947. https://doi.org/10.1007/978-3-319-73450-7_89

[34] C. Feng, M. Khan, A. U. Rahman and A. Ahmad (2020) News recommendation systems-accomplishments, challenges future directions,it IEEE Access , vol. 8, pp. 16702-16725. https://doi.org/10.1109/ACCESS.2020.2967792

[35] M. C. Kim and C. Chen (2015) A scientometric review of emerging trends and new developments in recommendation systems,*Scientometrics*, vol. 1, pp. 239-263. https://doi.org/10.1007/s11192-015-1595-5

[36] G. Wei, Q. Wu and M. Zhou (2021) A hybrid probabilistic multiobjective evolutionary algorithm for commercial recommendation systems, *IEEE Transactions on Computational Social Systems*, vol. 8, no. 3, pp. 589-598. https://doi.org/10.1109/TCSS.2021.3055823

[37] V. T. R. M. Vivek, C. Saravanan and K. V. Kumar(2022) A Novel Technique for User Decision Prediction and Assistance Using Machine Learning and NLP: A Model to Transform the E-commerce System, *Big data management in Sensing: Applications in AI and IoT*, p. 61.

[38] J. Davidson, B. Liebald, J. Liu, P. Nandy, T. V. Vleet, U. Gargi and S. Gupta (2010) The YouTube video recommendation system,*Proceedings of the fourth ACM conference on Recommender systems*, pp. 293-296. https://doi.org/10.1145/1864708.1864770

[39] W. Bellante, R. Vilardi and D. Rossi(2013) On Netflix catalog dynamics and caching performance, /it 2013 IEEE 18th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), pp. 89-93. https://doi.org/10.1109/CAMAD.2013.6708095

[40] X. H. Pham, T. N. Luong and J. J. Jung (2013) An black-box testing approach on user modeling in practical movie recommendation systems, *International Conference on Computational Collective Intelligence*, pp. 72-79. https://doi.org/10.1007/978-3-642-40495-5_8

[41] S. Raza and C. Ding (2021) News recommender system: a review of recent progress, challenges, and opportunities,*Artificial Intelligence Review*, pp. 1-52. https://doi.org/10.1007/s10462-021-10043-x

[42] F. Carmagnola, F. Vernero and P. Grillo(2009) Sonars: A social networks-based algorithm for social recommender systems,*International Conference on User Modeling, Adaptation, and Personalization*, pp. 223-234. https://doi.org/10.1007/978-3-642-02247-0_22

[43] W. Zou (2018) Design and application of incremental music recommendation system based on Slope one algorithm,*Wireless Personal Communications*, vol. 102, no. 4, pp. 2785-2795. https://doi.org/10.1007/s11277-018-5303-7

[44] W. Strank (2021) Analyzing Networks of Musical Context in the Digital Age1, *119: The Oxford Handbook of Music and Advertising*. pp.6080-6088 https://doi.org/10.1093/oxfordhb/9780190691240.013.48

[45] J. Sanz-Cruzado Puig(2021) Contact recommendation in social networks: algorithmic models,*diversity and network evolution*,pp. 519-569

[46] L. Lü, M. Medo, C. H. Yeung, Y.-C. Zhang and a. T. Z. Zi-Ke Zhang (2012) Recommender systems, *Physics reports*, vol. 519, no. 1, pp. 1-49. https://doi.org/10.1016/j.physrep.2012.02.006

[47] E. Brynjolfsson, Y. Hu and M. D. Smith (2003) Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers, *Management science*, vol. 49, no. 11, pp. 1580-1596. https://doi.org/10.1287/mnsc.49.11.1580.20580

[48] X. Amatriain and J. Basilico (2015) Recommender systems in industry: A netflix case study, *Recommender systems handbook*, pp. 385-419. https://doi.org/10.1007/978-1-4899-7637-6

[49] H. Verma (2022) Netflix Recommendation Engine - How Netflix uses Big data and Analytics to Recommend you your Favourite Shows, *startuptalky* [Online]. Available: https://startuptalky.com/netflix-recommendation-engine/ [Accessed 29 10 2022].

[50] M. Zhang and Y. Liu (2021) A commentary of TikTok recommendation algorithms in MIT Technology Review 2021, *Fundamental Research*, vol. 1, no. 6, pp. 846-847. https://doi.org/10.1016/j.fmre.2021.11.015

[51] Z. Chen and C. Shi (2022) Analysis of Algorithm Recommendation Mechanism of TikTok, *International Journal of Education and Humanities*, vol. 4, no. 1, pp. 12-14.

[52] S. Albawi, T. A. Mohammed and S. Al-Zawi (2017) Understanding of a convolutional neural network, *international conference on engineering and technology*, pp. 1-6. https://doi.org/10.1109/ICEngTechnol.2017.8308186

[53] J. Chai, H. Zeng, A. Li and E. W. Ngai (2021) Deep learning in computer vision: A critical review of emerging techniques and application scenarios, *Machine Learning with Applications*, vol. 6, p. 100134. https://doi.org/10.1016/j.mlwa.2021.100134

[54] N. L. H. Hien, L. V. Huy and N. V. Hieu (2021) Artwork style transfer model using deep learning approach, *Cybernetics and Physics*, vol. 10, no. 3, pp. 127-137. http://doi.org/10.35470/2226-4116-2021-10-3-127-137

[55] L. Liu, F.-X. Wu, Y.-P. Wang and J. Wang (2020) Multi-receptive-field CNN for semantic segmentation of medical images, *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 11, pp. 3215-322. https://doi.org/10.1109/JBHI.2020.3016306

[56] N. L. H. Hien and A.-L. Kor (year) Analysis and Prediction Model of Fuel Consumption and Carbon Dioxide Emissions of Light-Duty Vehicles, *Applied Sciences*, vol. 12, no. 2, p. 803. https://doi.org/10.3390/app12020803

[57] Prabhu (2018) Understanding of Convolutional Neural Network (CNN) — Deep Learning, *Medium* [Online]. Available: https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-d [Accessed 29 10 2022].

[58] N. L. H. Hien, H. M. Hoang, N. V. Hieu and N. V. Tien (2021) Keyphrase Extraction Model: A New Design and Application on Tourism Information, *Informatica*, vol. 45, no. 4. https://doi.org/10.31449/inf.v45i4.3493

[59] P. J. L. and G. W. Li (2018) Application of convolutional neural network in natural language processing, *15th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*, pp. 120-122. https://doi.org/10.1109/ICCWAMTIP.2018.8632576

[60] Y. R. B. and C. V. Koren (2009) Matrix factorization techniques for recommender systems, *Computer*, vol. 42, no. 8, pp. 30-37. https://doi.org/10.1109/MC.2009.263

[61] H.-J. Xue, X. Dai, J. Zhang, S. Huang and J. Chen (2017) Deep matrix factorization models for recommender systems, *IJCAI*, vol. 17, pp. 3203-3209. https://www.ijcai.org/proceedings/2017/447

[62] grouplens (2003) MovieLens 1M Dataset [Online]. Available: https://grouplens.org/datasets/movielens/1m/ [Accessed 29 10 2022].

[63] H. Liu, C. Zheng, D. Li, X. Shen, K. Lin, J. Wang, Z. Zhang, Z. Zhang and N. N. Xiong (2021) EDMF: Efficient deep matrix factorization with review feature learning for industrial recommender system, *IEEE*

Transactions on Industrial Informatics, vol. 18, no. 7, pp. 4361-4371. https://doi.org/10.1109/TII.2021.3128240

[64] H. Wu, Z. Zhang, K. Yue, B. Zhang, J. He and L. Sun(2018) Dual-regularized matrix factorization with deep neural networks for recommender systems, *Knowledge-Based Systems*, vol. 145, pp. 46-58. https://doi.org/10.1016/j.knosys.2018.01.003

Identification of Students' Confusion in Classes from EEG Signals Using Convolution Neural Network

Rekha Sahu[1], Satya Ranjan Dash[2,*], and Amarendra Baral[3]
[1]Silicon Institute of Technology, Bhubaneswar, Odisha, India
[2]Kalinga Institute of Industrial Technology, Bhubaneswar, Odisha, India
[3]Trident Academy of Technology, Bhubaneswar, Odisha, India
E-mail:sahu_r@rediffmail.com, sdashfca@kiit.ac.in, hodmath@tat.ac.in
*Corresponding author

Keywords: students' confusion, predefined confusion, user-defined confusion, mismatch of confusion label, one-dimensional convolution neural network, electroencephalography signals

Received: January 7, 2023

For a student, classes are vital factors for gaining knowledge. The lectures may be online or offline, but getting knowledge without confusion is a major issue. The confusion labels can be measured from the electroencephalography signals and the confusion can be solved after knowing that students are suffering from confusion. Different machine learning approaches were implemented on electroencephalography signals to identify the suffering of students from confusion. The performance of traditional machine learning approaches in predicting confusion status is found as poor. In this paper, the one-dimensional convolution neural network is implemented on the electroencephalography signals to detect confusion of the students at the time of watching video classes. Students' attention, mediation, electroencephalography signals, delta, theta, alpha1, alpha2, beta1, beta2, gamma1 and gamma2 are taken into consideration to train a one-dimensional convolution neural network classifier. The one-dimensional convolution neural network approach has achieved better accuracy in detecting the confusion of the students. Besides finding confusion labels of students, the experiment is performed when understandable classes are creating confusion and the difficult classes are understandable for the students. This second experiment is also performed on electroencephalography signals of students and after identification of confusion status, the improvement of students' deficiencies can be possible. For future work, more data and different aspects of the students can be taken into consideration for identifying confusion and different obstacles respectively which helps to improve in achieving perfect knowledge from the classes.

Povzetek: Raziskava obravnava identifikacijo zmedenosti študentov med predavanji z uporabo EEG signalov in enodimenzionalne konvolucijske nevronske mreže, kar omogoča boljše razumevanje in obravnavo učnih ovir za izboljšanje pedagoškega procesa.

1 Introduction

Education influences society significantly and education is an essential aspect of a better society and comfortable life. To spread education all over society, the proper way of teaching, as well as students' perception levels, should be analyzed and emphasized for adopting the improvement approaches in teaching procedures. The teaching procedures and student perceptions are vital factors in creating an educated society. Investigations show that students are facing problems when learning from lectures. They suffer from confusion and are unable to understand the lectures. It is found that students can better learn only if the teaching procedure, as well as student perception, is better [8]. Further, the teaching process influences the educational system drastically and delivering a better lecture, appreciated by students, influences the educational system positively [30, 34]. By the way, students' attitudes in perceiving the contents of the lecture are also influencing the learning strategy [29]. Different observation shows that student confusion level is an important factor to certify whether a class is appreciable for better education or not. Again, the lectures are delivered either online or offline. Whatever may be the procedure of delivery of the lectures, mainly the students should understand and be clear on the concepts behind the lessons. Otherwise, the lectures are unnecessary, wastage of time, and meaningless. Since, nowadays, education is provided through online classes, experiments have been performed on the impact of online classes [6].During the pandemic, online classes were taken to overcome from discontinuous classes of the students. But, the students were suffering from confusion, down, sad, upset, excitement etc. in the classes [24]. Because of the online classes during covid-19, the perception of the students was less [7, 34]. Besides the deficiencies in students' perception of online classes, the instructors also showed their deficiencies in teaching, behaviours, emotions, attention, cog-

nitive workload and trust [18]. Moreover, the relationship between the students' and instructors' behaviours, emotions, attentional, cognitive workloads, trust and collaboration was required to enhance the clarity, and understanding of the lectures in the classes [18]. By the way, online lectures are more useful since the lectures can be attended at any time and anywhere according to the flexibility of students. Even after the pandemic, online classes are appreciated for higher education along with the cognizance of staff and students which is essential [10]. It is needed to observe the impact on the understanding and confusion level of students automatically during classes for taking appropriate actions.

During the online classes, whether the student is in confusion or not, was a vital matter. In an experiment, the online class was shown as poor in participation, emotional, skill and performance engagement in contrast to face-to-face classes [37]. But, Ram´ırez-Moreno et.al. has found from electroencephalography (EEG) signals that online teaching is better than classroom teaching [26]. The EEG signals from the frontal lobes visualize the confusion level of a human. Hence, the EEG signals from the frontal lobes of students could state whether the student is in confusion or not during online teaching. Further, the experiment stated that the Fp1 channel is placed on the frontal lobe and it can be used to measure the concentration and confusion level of a subject [22]. Again, by manipulating raw EEG signals of the Fp1 channel, delta, theta, alpha, beta and gamma frequencies have been extracted for deep analysis [1].For classifying the EEG signals' pattern, traditional machine learning and deep learning have been applied to EEG signals datasets to find the pattern of EEG signals in recognizing the student state features [16].

The confusion labels of students can be measured from EEG signals and the deep learning approach implementation on EEG signals can find out the specific pattern for a specific target class [22, 16]. These influence to implementation of a deep learning approach on the EEG signals of students for detecting students' confusion labels. In our experiment, the EEG signals of students had been collected at the time of watching videos in online classes[35]. Intentionally, the videos were created as confused videos and non-confused videos. Those are called predefined confused and non-confused labels. After watching the videos for learning, the students labelled whether the videos are creating confusion or non-confusion in understanding the lessons. Those are called user-defined confusion and non-confusion labels. Both pre-defined labels and user-defined labels are mismatched in some cases. Hence, we have created three questions. Firstly, for which pattern of EEG signals, the students are suffering from confusion. Secondly, for which pattern of EEG signals, the students are not in confusion. Since in some cases, pre-defined labels and user-defined labels are mismatching, so thirdly, we have analysed the pattern of EEG signals for which signals are mismatched. The collected EEG signals are raw Fp1 EEG signals. From the raw Fp1 EEG signals, different features like, Attention, Meditation, Raw EEG signals, Delta frequency, Theta frequency, Alpha1, Alpha 2, Beta1, Beta 2, Gamma1, and Gamma2 are extracted for confused and non-confused students. Since deep learning approaches are implemented for finding the pattern for classification tasks [16], so we have applied a one-dimensional convolution neural network (1DCNN) on our extracted dataset to classify the EEG signals for confusion, non-confusion and mismatching labels of user-defined and pre-defined labels. The overall work performed in this paper is represented in Fig. 1.

The rest of the paper is as follows. In section 2, related work is stated. Our experiment details are represented in section 3. The description of the dataset is presented in section 3.1, the technology applied is elaborated in section 3.2 and the result analysis is presented in section 3.3. Finally, in section 4, a conclusion and possible future work are stated.

2 Related work

For developing teaching-learning procedures, different experiments and surveys are performed. Some surveys have concluded that students are suffering from academic stress drastically. Even achieving knowledge from the lectures of reputed universities is becoming hard for them [3]. Sometimes for improving learning, students were specially trained with some teaching-learning techniques and got good scores in comparison to direct attending the lecture [19]. Moreover, student confusion is a major factor in college lectures and the detection of confusion depends on attention and meditation [23]. It is hard to measure the attention of the student through self-report or from the behaviour of the students. The state of the students' minds can be analyzed and found from the EEG signals [20, 9].

Since the report of students or observers is not sufficient to measure mind state and the mind state of a student can be measured from EEG recording [20, 9], so we have experimented with EEG recording to find out the confused students. Our survey helps to find how different factors like Attention, Meditation, Raw EEG signals, Delta frequency, Theta frequency, Alpha1, Alpha 2, Beta1, Beta 2, Gamma1, and Gamma2 are extracted from EEG signals. J. K. Grammer, et.al. stated that from EEG signals, the measurement of student attention can be quantified [13]. Moreover, from the channel Fp1 EEG signals, the attentive & inattentive students are classified and Ning-Han Liu, et.al. implemented Support Vector Machine (SVM) approach to classify the EEG signals pattern to visualize the attention of the student [17]. It is found out Meditation describes the state of calmness and focused attention of mental activity and this can be identified from EEG signals[32] and it is observed that Mindfulness meditation can be quantified from the frequency of EEG signals [2]. The above-mentioned Fp1 channel is placed on the frontal lobe and it can be used to measure the concentration of a subject. Again, memory retrieval, decision-making, planning, response evalua-

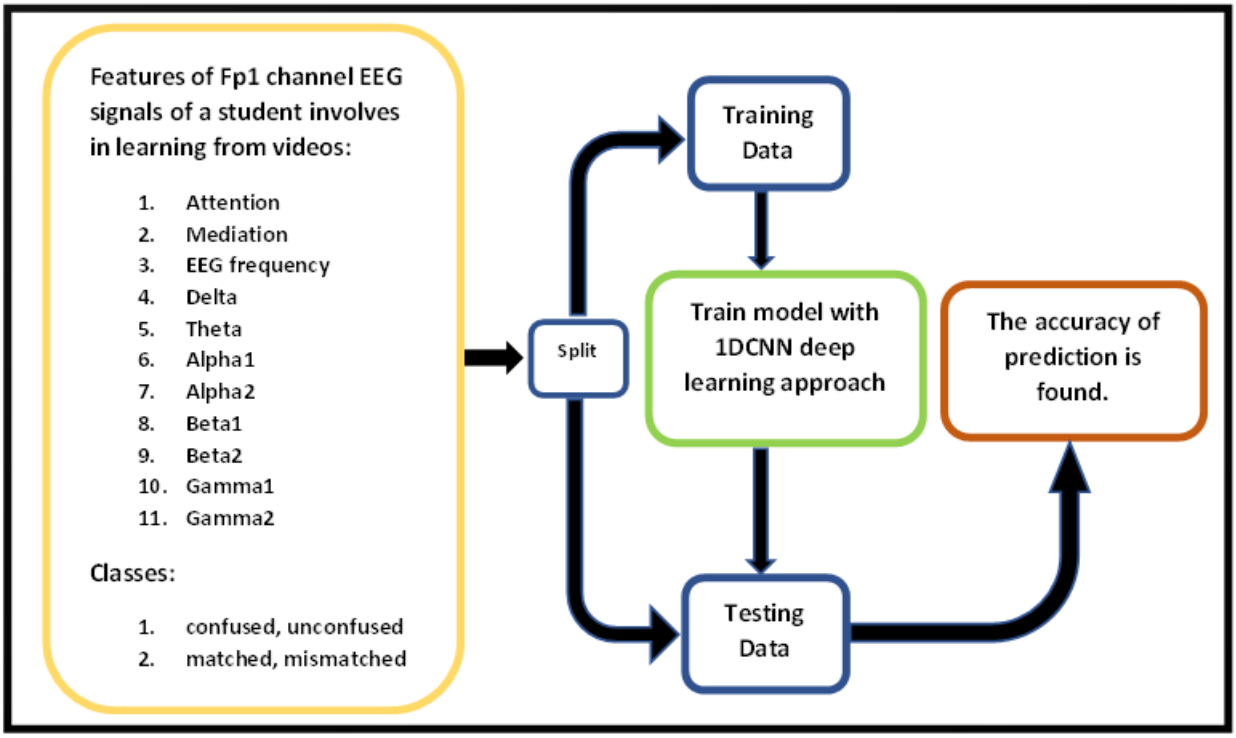


Figure 1: Overall workflow diagram: It is representing different attribute values generated from EEG signals taken as input. The class values as confused vs unconfused and matched vs mismatched are included. Input data are split as training dataset and testing dataset. 1DCNN classifier is trained using a training dataset and implemented on a testing dataset to find the accuracy of prediction.

tion, and reflection of a subject are studied through channel frequency[22]. At the same time, EEG signals can display five types of EEG waves i.e., gamma, beta, alpha, theta, and delta [1]. Generally in the case of the gamma wave, higher processing tasks and cognitive functioning are performed. The gamma waves are responsible for cognitive functioning, learning, memory, information processing, attention, focus, consciousness, mental processing, and perception. In other sites, Beta waves are related to conscious thought, logical thinking, stimulating effect, conscious focus, memory, and problem-solving. Again, Alpha waves lead to the feeling of deep relaxation and calm down whereas, Theta waves involve improving intuition, creativity and a more natural feel. Lastly, Delta waves involve feeling rejuvenated, promoting the immune system, natural healing, and restorative/deep sleep[25]. To find out the delta, theta, alpha, beta and gamma frequency, we manipulate raw EEG signals [1] and hence by manipulating Fp1 EEG recording, we can find the delta, theta, alpha, beta and gamma band frequency for Fp1 EEG channel. To study the pattern of EEG signals to recognize the student state features, both traditional machine learning and deep learning can be applied to EEG signal datasets [16].

From the literature survey, we have found that machine learning and deep learning approaches are applied to EEG signals to find different patterns [14, 27, 28]. The Machine learning approaches like logistic regression, random forest, decision tree, K- nearest neighbour (KNN) and SVM are applied to Brain-Computer Interface (BCI) data set and found out logistic regression has given better performance in the detection of students' confusion in Massive Open Online Course (MOOC)[5]. Again, the attention of students is studied from the EEG signals when the students were involved in MOOC and traditional classrooms and the SVM approach was implemented on the EEG data [32]. The experiment result concluded that the MOOC learning process maintains higher attention. Besides, different traditional machine learning approaches like the random forest, SVM and KNN are applied to the EEG signals dataset to classify students' attention levels when involve in online classes [4]. Not only traditional machine learning, but deep learning approaches have also given better performance in identifying a specific EEG signal pattern. The experiment on EEG signals of nineteen students is performed to identify their emotions like happiness, sadness, anger, fear, disgust, and surprise. In this experiment, the deep learning approaches i.e., Long Short Term Machine (LSTM) and Convolution Neural Network (CNN) are applied to the EEG signals to identify the emotions and found 99.8% classification accuracy with implementing CNN [14]. Again, the Students' attentiveness towards the lectures is measured from EEG signals patterns, and it was fruitful by analysing EEG signals data using three-dimensional CNN [15]. With the above survey, we also found out that Bidirectional LSTM Recurrent

Neural Networks were implemented on the EEG signals dataset to identify the confused and non-confused students when involve in online courses. It was observed that the classification accuracy was 73.3% and the gamma 1 wave can be used to identify the confusion [23]. A deep learning approach can also be implemented on EEG signals to find out the attention level of a student[33]. Thus, the survey concludes that the traditional machine learning, deep learning and spiking neural network analysed and classified the EEG signals for extracting specific patterns [27, 28]. It is observed that the one-dimensional convolution neural network (1DCNN) is implemented on the EEG signals and given higher accuracy in detecting the different pattern EEG signals [27]. Again, the CNN approach is implemented on the raw EEG signals of one channel to detect sleep disorders[31]. Besides, fear, fun and sad emotions are identified from the EEG signals using the CNN approach [12]. After going through the above literature, we have proposed a 1DCNN model applied to EEG signals data set to detect confusion of students when involve in video classes. In the next section, we have stated our experiment and compared our novel approach with other works and also the different aspect, we have experimented, with is elaborated.

3 Experiment

For fair teaching procedure, emphasis should be given to observing how fairly lecturers are delivered and how much students can able to perceive from lectures. Hence, the student's understanding and confusion status is essential to observe. Our experiment is performed to find out whether a student is in a confused or non-confused state when watching online lectures. Therefore, EEG signals are collected from the students when they were watching the lectures. Those signals are used to train the models for classification tasks. Here, confusion and non-confusion of a student are interpreted according to predefined or user-defined labels. Predefined implies the videos of the lecture are recorded intentionally as either confused or not confused lectures. User-defined implies students practically labelled that the lecture is either confusing or not confusing. With this dataset, a deep learning model is trained. The model predicts whether the student is in confusion according to the predefined or confusion according to the user-defined. Also, a model is trained to find out the pattern of signals for which predefined opinions and user-defined opinions are the same and for which they have mismatched. The explanation of experiments is as follows. We have described the dataset in section 3.1, the description of the method applied to the dataset is represented in section 3.2 and finally in section 3.3 result of the experiment is discussed.

3.1 Dataset and its analysis

For finding whether the students suffering from confusion, the EEG signals pattern is required to study when they are involved in watching MOOC video clips. We have collected EEG brain wave dataset from the Kaggle database [35]. To collect the dataset of students' EEG signals, twenty videos were prepared and each video was of two minutes. Again, a two-minute clip in the middle of a topic is chopped to make the videos more confusing. Out of twenty videos, ten videos are prepared to confuse a normal student and ten videos are prepared to not confuse a normal student. These videos are shown to ten students to test their confusion labels. However, one student is not considered for missing data due to a technical defect. Among twenty videos, randomly five videos of each category are picked and those are presented to a student in random sequence. This was the procedure that was followed for each student. Then, the students were instructed to learn as much as possible from the video clip. When the students were watching the video clip, the body language of the students was observed and the confused state of the students was noted. In general, after each video, the student rated the confusion label as well as an observer of the student rated the corresponding confusion label. The confusion label was defined on a scale of 1-7, where 1 stands for least confusing and 7 stands for most confusing.

EEG signals from each student were collected from the frontal lobe (Fp1) that lies between the left eyebrow and hairline. Using a wireless single-channel Mindset, EEG signals of Fp1 were collected and those are depicted in figure. 2. Besides, using NeuroSky's API, the following signals' information is collected.

1. The raw EEG signal, sampled at 512 Hz
2. An indicator of signal quality, reported at 1 Hz
3. MindSet's proprietary "attention" and "meditation" signals are said to measure the user's level of mental focus and calmness, reported at 1 Hz
4. A power spectrum, reported at 8 Hz, clustered into the standard named frequency bands: delta (1-3Hz), theta (4-7 Hz), alpha (8-11 Hz), beta (12-29 Hz), and gamma (30-100 Hz)

Finally, from the Fp1 channels recording, the attributes Attention, Meditation, Raw EEG signals, Delta frequency, Theta frequency, Alpha1, Alpha 2, Beta1, Beta 2, Gamma1, and Gamma2 are taken into consideration. To characterize the overall values of the attributes, the mean statistic is calculated. We have 100 data points for 9 subjects and each watch 10 videos. The class value for the corresponding instance is the label based on a predefined confusion label as the experiment designed and the user-defined confusion label as the user's subjective rating. Hence, for one instance we have two labels one is a predefined confusion label and another is user defined confusion label. Besides, a mismatch label is generated to differentiate the predefined confusion label and the user-defined confusion label. In the dataset, the number of instances is 12811 and the number of attributes is 16. The attributes are the serial number of subjects, the serial number of videos, Attention, Meditation, Raw EEG signals, Delta frequency, Theta frequency, Alpha1, Alpha 2, Beta1, Beta 2, Gamma1, Gamma2, the pre-

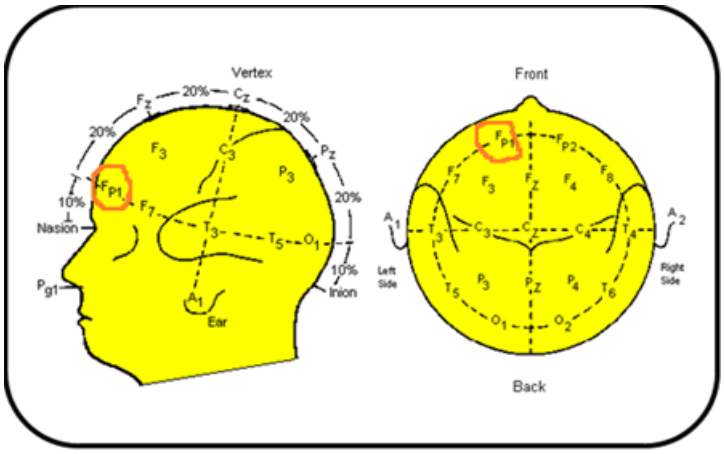


Figure 2: Fp1 channel location is shown on the head which is between the left eyebrow and hairline.

defined, user-defined and the mismatched labels. Attention, Meditation, Raw EEG signals, Delta frequency, Theta frequency, Alpha1, Alpha 2, Beta1, Beta 2, Gamma1, and Gamma2 are the frequency values and the pre-defined and user-defined attributes contain either 0 or 1, where 0 stands for the student is not confused and 1 stand for the student is confused. Again, the mismatch attribute contains 1 or -1 or 0, where 1 implies confused according to predefined but not confused according to the user-defined, -1 in mismatch implies not confused according to predefined but confused according to the user-defined and 0 implies both have the same label. All information about the dataset is summarized in the table. 1.

The graphical analysis of 11 attributes of three types of class i.e., predefined confused, user-defined confused and mismatched labels, are depicted in figs. 3, 4, 5, 6, 7, 8, 9, 10.

3.2 One-dimensional convolution neural network approach (1DCNN)

We have proposed a variant of the CNN approach called 1DCNN to identify the confused student against the unconfused. 1DCNN is a sequence of layers: convolution layer, pooling layer, flatten layer and dense layer followed by activation function [27]. The purpose of the convolution layer is to filter the data. For the convolution operation, we have the kernel, the dot product is performed between the input data and kernel. The stride and padding are performed and finally get a new filter dataset. Then, the dataset is reduced by doing the max pooling operation in the pooling layer. After pooling, we flatten the pooling data into a column. Then those column data are the input for the artificial neural network that is the dense layer of the proposed approach. On the output of the dense layer, we use the activation functions like the Re LU function and soft-max function, which are defined in equations 1 and 2 respectively.

$$f(x) = \begin{cases} 0 & \text{when } x < 0 \\ 1 & \text{when } x \geq 0 \end{cases} \tag{1}$$

$$S(x) = \frac{e^x}{\Sigma_{x=1}^{n} e^x} \tag{2}$$

In the convolution layer, one-row data (1×n) is filtered using the convolution operation with a one-dimensional filter (1×m). The maximum value of one pad is taken for max pooling. Besides, the ReLU function gives the output value when the value is positive otherwise it gives zero and the softmax function predicts the probability of input data belonging to a class. The diagrammatical representation of the CNN model is represented in Fig. 11.

3.3 Experiment result and discussion

1DCNN approach applies to predefined confusion EEG signal datasets and user-defined confusion EEG signal datasets to identify the confused students according to predefined confusion and user-defined confusion of students respectively. Videos are intentionally recorded as confused videos and unconfused videos. Some confused videos are rated as unconfused by the students and some unconfused videos are rated as confused by the students. 1DCNN is also applied to find the signal pattern for the mismatch of the user-defined and predefined class labels.

For the predefined confused EEG signals dataset, the structure of 1DCNN is as follows. The kernel size is 1×3, the number of filters is 10, and the input shape is 1×12. The Max pooling size is 4. After flattening, two dense layers are structured with 500 neurons with a ReLU activation function followed by 2 neurons with a SoftMax activation function. For optimization, Adam's version of the gradient descent learning approach is implemented. 80% data is used for training and 20% is used for testing. With one epoch, we have got 100% classification accuracy in finding confused students' EEG patterns in contrast to unconfused ones.

For the user-defined confused EEG signals dataset, the

Number of subjects	9
Number of Videos	20 (10 for confused and 10 for not confused)
EEG recording duration per subject and video	2 min (total 6 hours recording)
Channel recorded	One channel Fp1
Number of attributes	17
Number of instances	12811
Class label	Confused and not confused mismatched of pre-defined and user-defined opinion

Table 1: Dataset descriptions of the students' online classes and their confusion labels.

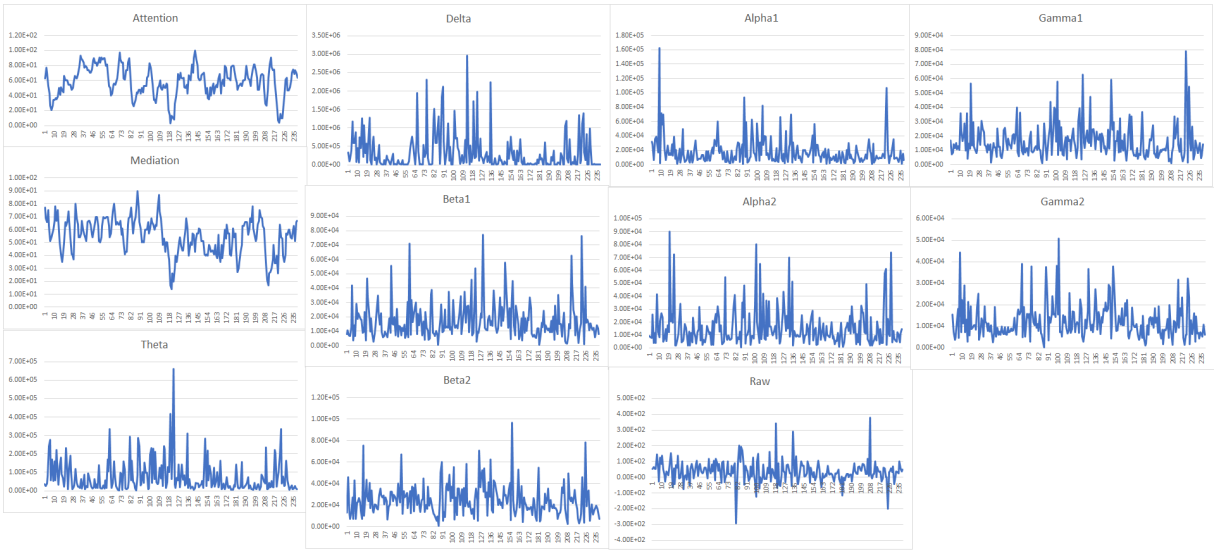


Figure 3: Attributes value representation for user defined non-confused labels.

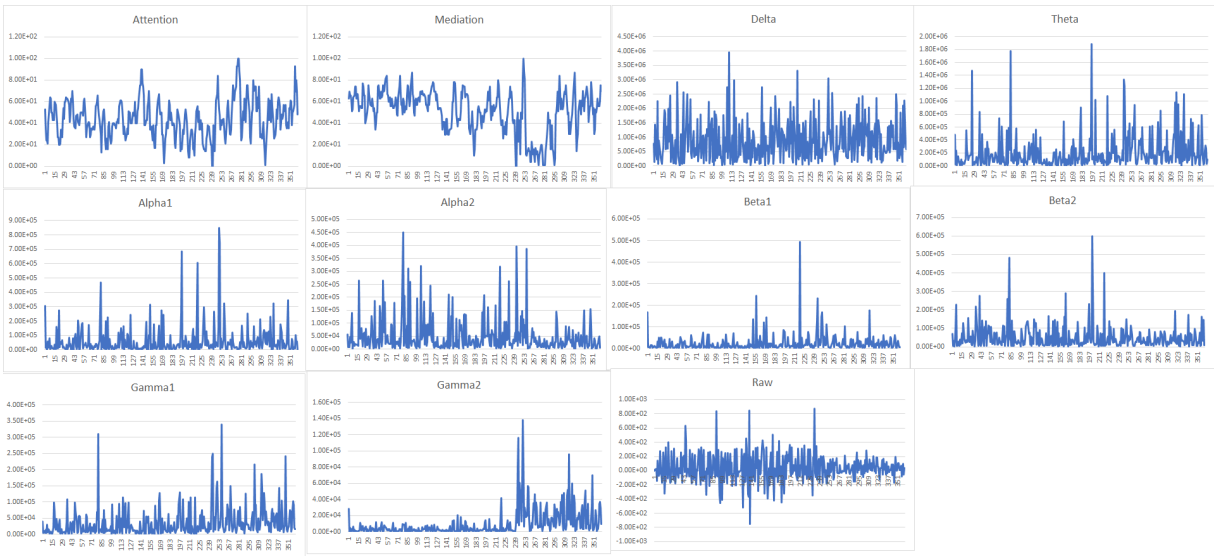


Figure 4: Attributes value representation for user-defined confused labels.

structure of 1DCNN is as follows. The kernel size is 1×3, the number of filters is 10, and the input shape is 1×11. The max pooling is 4. After flattening, two dense layers are structured with 1000 neurons with a ReLU activation function followed by 2 neurons with a SoftMax activation function. For optimization, Adam's version of the gradient descent learning approach is implemented. 80% data is used for training and 20% is used for testing. With 1500 epochs, we have got 99% classification accuracy in finding confused students' EEG patterns in contrast to unconfused

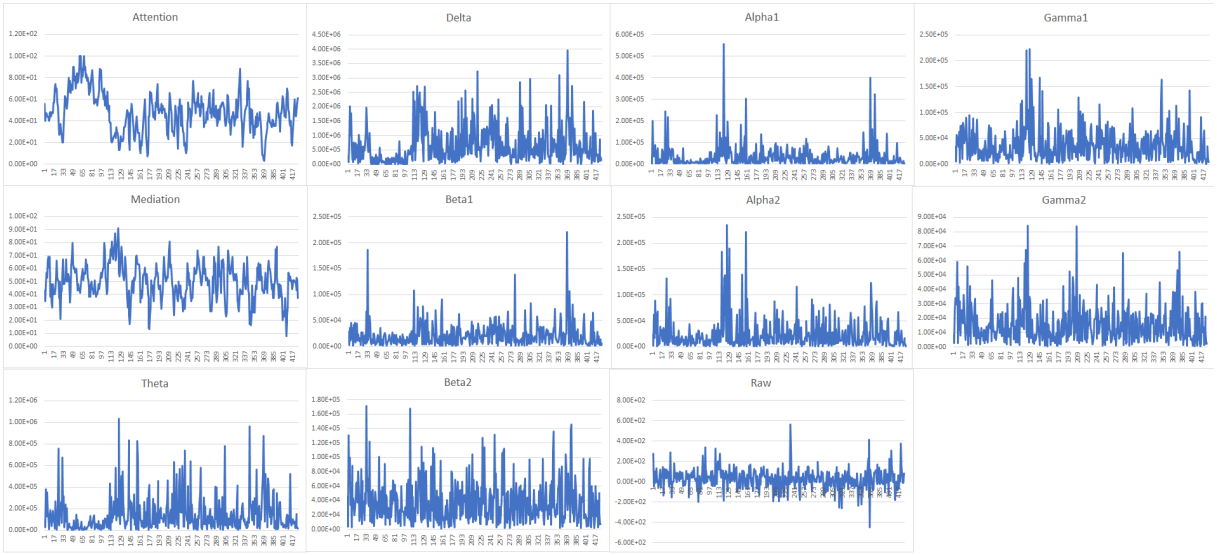


Figure 5: Attributes value representation for predefined non-confused labels.

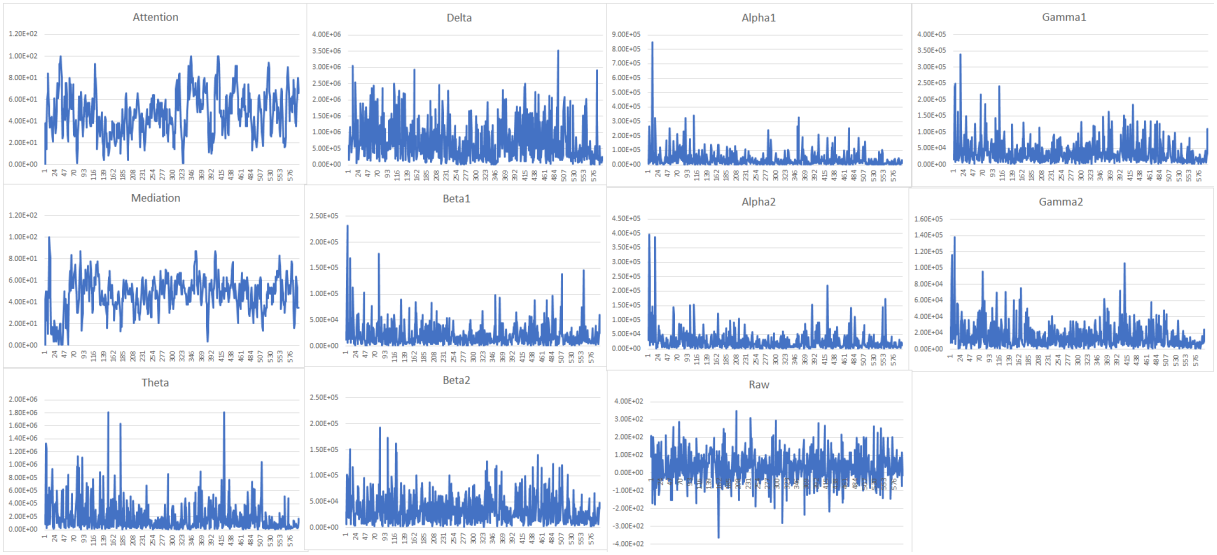


Figure 6: Attributes value representation for predefined confused labels.

ones.

For mismatched user-defined and pre-defined confused rate EEG signals dataset, the structure of 1DCNN is as follows. The kernel size is 1×3, the number of filters is 10, and the input shape is 1×11. The max pooling is 4. After flattening, two dense layers are structured with 500 neurons with a ReLU activation function followed by 3 neurons with a SoftMax activation function. For optimization, Adam's version of the gradient descent learning approach is implemented. 80% data is used for training and 20% is used for testing. With 10000 epochs, we have got 99% classification accuracy in finding mismatches.

Some works are performed on the EEG signals confused dataset [23, 21, 11]. The probability-based features approach utilizes the probabilistic output from the random forest and gradient-boosting machine to train machine learning models to detect the confused student [11]. Again, Gaussian Naïve Bayes classifiers are trained with the dataset to find out the confused students. The accuracy of the classification pattern of EEG signals for the confused student was less than 70% [36]. The bidirectional LSTM Recurrent Neural Networks approach is applied to the confused EEG signal data to detect the confused student and the classification accuracy is found to as 73.3% [23]. The experiments with different traditional machine learning approaches and deep learning approaches on the dataset have given less accuracy in comparison to our experiment except for the probability feature-based approach and the performances are summarized in table. 2.

Thus, from the summary in table. 2, it is concluded our proposed approach has efficiency to identify the confused students. Besides, the experiments with different traditional

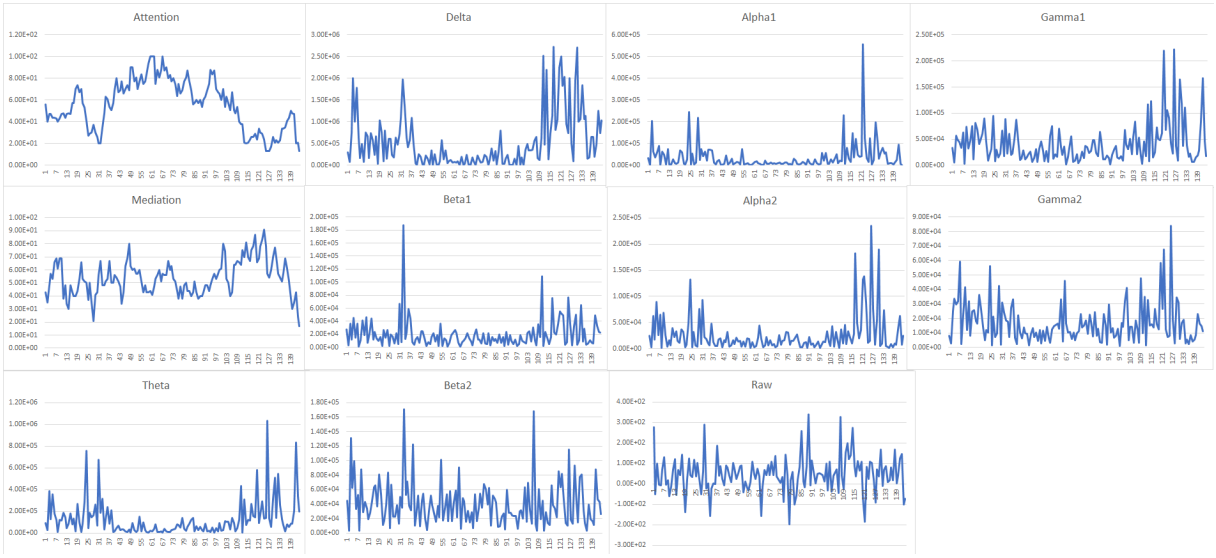


Figure 7: Attributes value representation for user-defined and pre-defined labels are matched (not confused).

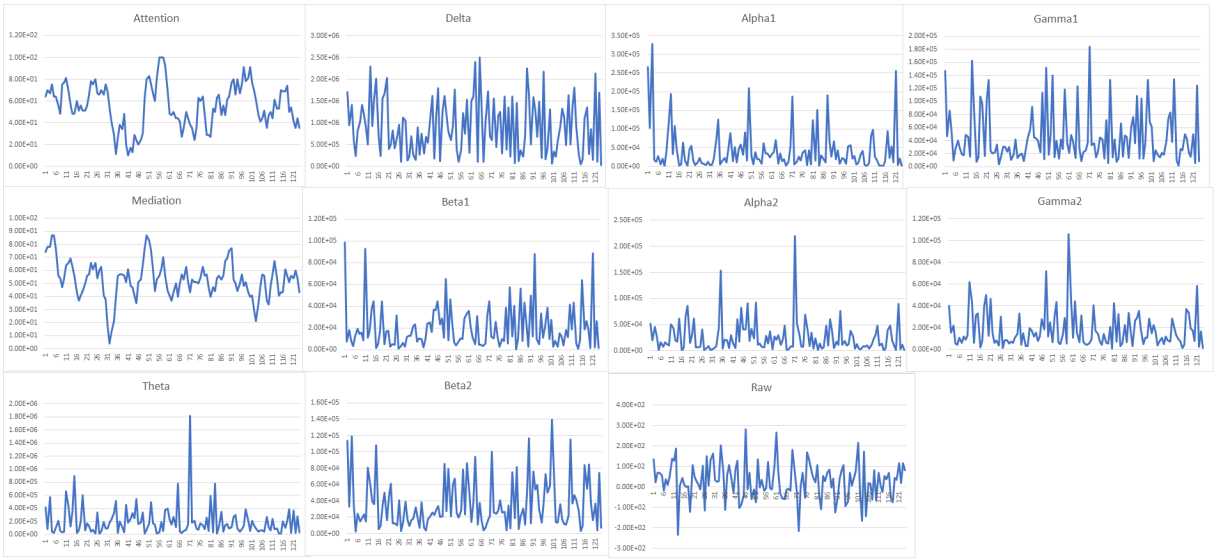


Figure 8: Attributes value representation for user-defined and pre-defined labels are matched (confused).

Approach Implemented	Purpose of the Approach	Accuracy
1DCNN	Detect confused student (according to predefined)	100%
1DCNN	Detect confused student (according to user-defined)	99%
1DCNN	Detection of mismatch of user-defined and pre-defined confused label	99%
The probability-based features approach utilizes the probabilistic output from the random forest and gradient-boosting method	Detect confused student	99%
Gaussian Naïve Bayes method	Detect confused student	70%
The bidirectional LSTM Recurrent Neural Networks approach Neural Networks approach	Detect confused student	73.3%

Table 2: Summary of the performances of different approaches.

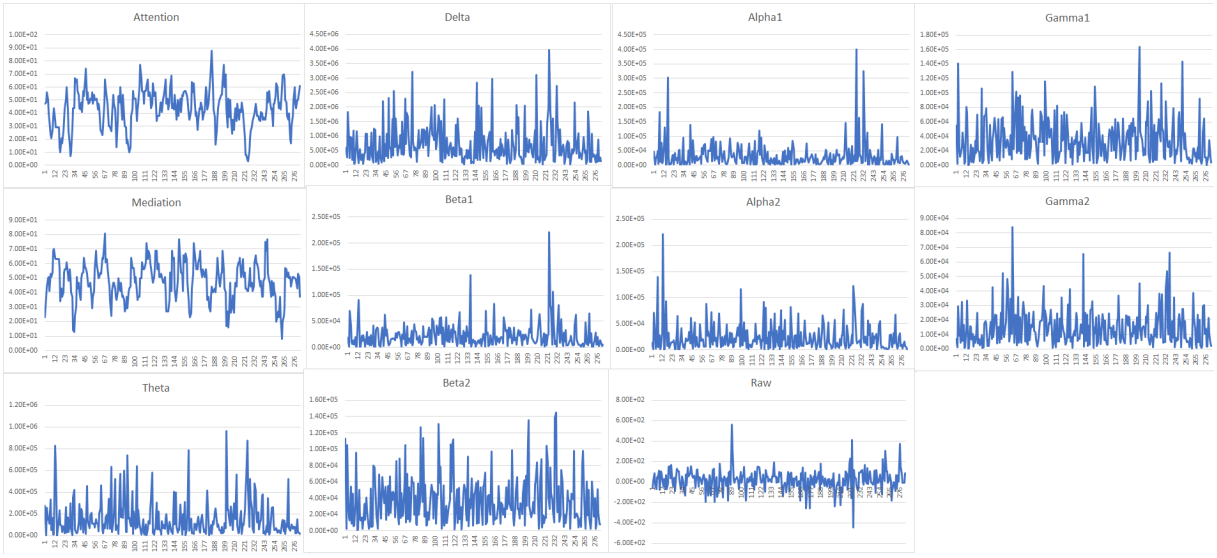


Figure 9: Attributes value representation for user-defined and pre-defined labels are matched (when predefined is not confused and user-defined is confused).

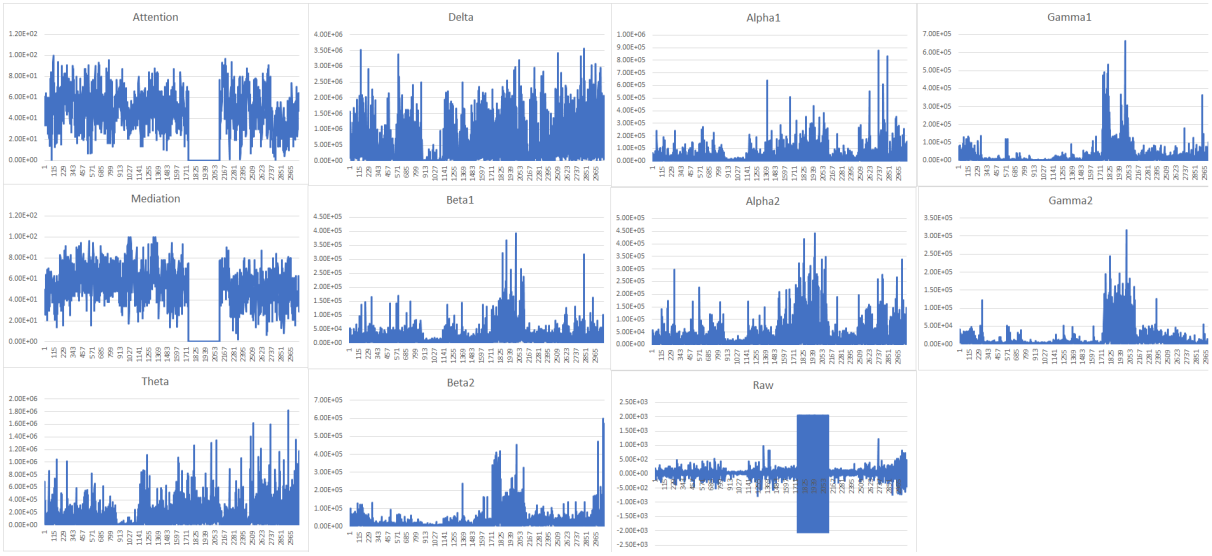


Figure 10: Attributes value representation for user-defined and pre-defined labels are matched (when predefined is confused and user-defined is not confused).

machine learning approaches and deep learning approaches on the dataset have given less accuracy in comparison to our experiment, except for the probability feature-based approach. The probability feature-based approach and other machine learning approaches have emphasized the finding of confused students from the signals whereas our experiment has performed on more than finding confused students i.e., when user-defined confusion is found, when predefined confusion is found and when predefined & user-defined labels are mismatched. For all three cases, EEG signals' patterns are trained using the 1DCNN model and have given 100%, 99% and 99% classification accuracies respectively. Besides, no discussion is shown in any paper still now on mismatched labels of user-defined and predefined labels. In finding a mismatch, it is possible to analyze more on the reason for the mismatch. The reason for the mismatch may be due to misinterpretation or more talented students. If the predefined confusion level is 0 but the user-defined confusion level is 1, then it will be assumed the student is more talented or had knowledge of the lecture before. If the predefined confusion level is 1 but the user-defined confusion level is 0, then those students should be analyzed to study the reason for confusion and their EEG signals pattern predict the student is in confusion although the lecture is very simple to understand. This issue can be analyzed more to treat the student's deficiency.

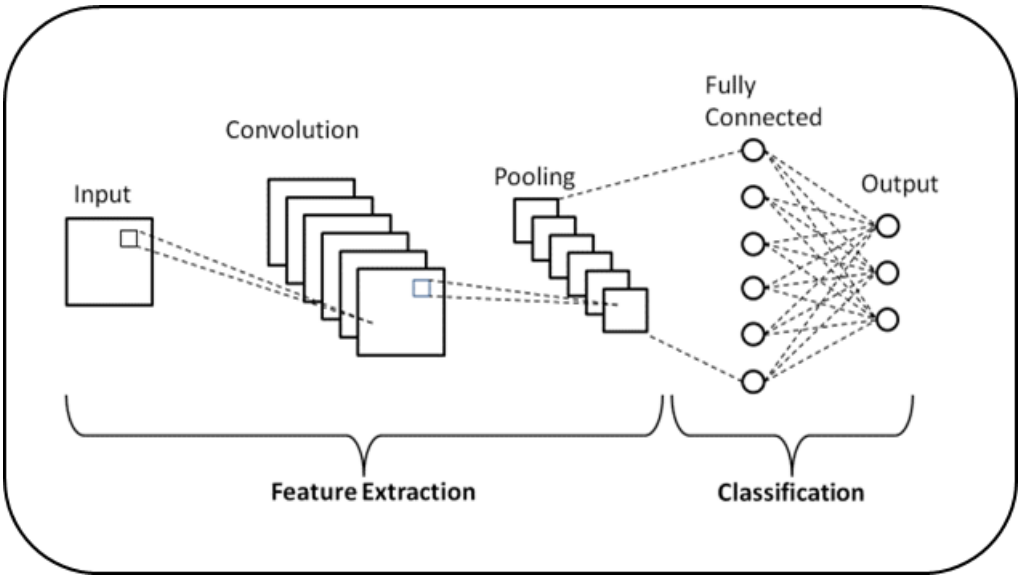


Figure 11: Convolution Neural Network Architecture with the input layer, convolution layer, pooling layer, dense layer followed by output layer.

4 Conclusion and future work

Students learn from the lectures in the classes and so the lectures should be understandable without confusion. Due to covid 19 pandemic, classes were online mode and nowadays also video lectures are influencing students. In this work, the confusion labels were studied when the student was watching video lectures. Twenty videos were collected out of which, ten were confused videos and ten were non-confused videos. Nine students' EEG recordings were collected and the attributes' values were extracted to find the patterns for confused students according to predefined, non-confused students according to predefined, confused students according to the user-defined, and non-confused students according to the user-defined. Besides, the mismatched patterns of user-defined and predefined are extracted. For extracting the patterns, 1DCNN is implemented and found to have better classification accuracy. For pre-defined labels, it has given 100% classification accuracy. For user-defined labels, it has given 99% classification accuracy. Finally, the mismatched confusion label of user-defined and predefined has shown classification accuracy as 99%. In all three cases, 80% data is used for training with 1DCNN and 20% data is used for testing. Thus, the proposed deep learning approach has given better accuracy in finding confused students when pre-defined confused labels are mismatched with the student-defined confusing label. The experiments were performed to identify the pattern of EEG signals for confused students but no discussion was emphasized for the pattern that causes mismatched and our paper has discussed mismatch in confusion labels. By applying the approach to more datasets, we can extract more information for analyzing students' confusion. As a result, the deliberation of lectures can be improved and the students can be treated accordingly.

More research can be performed relating to confusion and other problems of the students when involved in offline or online classes or watching videos. We have taken less amount of EEG datasets, and more experiments with more datasets can give better conclusions regarding the confusion of students and correspondingly we may treat the students for better achievement in education. The major study of mismatches of user-defined confusion and pre-defined confusion labels tends to analyze the different characteristics of the students to check whether the student is more talented (user defined is 0 but predefined is 1) or not talented (user defined is 1 but predefined is 0) or any other issues (previously know about the contains of lectures). Hence, mismatch leads to more analysis on the features of students and this can be kept as feature work. Moreover, if the user-defined label is the same as the predefined label, then there will not require more analysis, otherwise, more analysis will require on the attribute values or some other criteria are taken into consideration to find the reason for the mismatch like a student is more talented. Besides confusion, researchers focus on other attributes for finding deficiencies like attention, interest etc. for better improvement of the students in classes (online/offline) or watching videos.

References

[1] 5 types of brain waves frequencies: gamma, beta, alpha, theta, delta. Access: November 2022. `https://mentalhealthdaily.com/2014/04/15`.

[2] S. Aggarwal, M. Lamba, K. Verma, S. Khuttan, and H. Gautam. A preliminary investigation for assessing attention levels for massive online open courses learning environment using eeg signals: An experimental

study. *Human Behavior and Emerging Technologies*, 3(5):933–941, 2021.

[3] J. Agolla and H. Ongori. Assesment of academic stress among undergraduate students: The case of university of botswana. pages 63–70, 2009.

[4] A. Al-Nafjan and M. Aldayel. Predict students' attention in online learning using eeg data. *Sustainability*, 14(11):6553, 2022. http://dx.doi.org/10.3390/su14116553.

[5] V. A. S. M. Anala and G. Bhumireddy. Comparison of machine learning algorithms on detecting the confusion of students while watching moocs, 2022.

[6] R. P. Baral. The digital divide in online learning: A case study of university students in nepal. *Prithvi Academic Journal*, pages 88–99, 2022. http://dx.doi.org/10.3126/paj.v5i1.45043.

[7] H. Beyari. Predicting the saudi student perception of benefits of online classes during the covid-19 pandemic using artificial neural network modelling. *IJCSNS*, 22(2):145, 2022. http://dx.doi.org/10.31235/osf.io/3vwcu.

[8] B. Cerbin. Improving student learning from lecture. November 2019. https://takinglearningseriously.com/2019/11/16/improving-student-learning-from-lecture/.

[9] K.-m. Chang, J. Nelson, U. Pant, and J. Mostow. Toward exploiting eeg input in a reading tutor. *International Journal of Artificial Intelligence in Education*, 22(1-2):19–38, 2013. http://dx.doi.org/10.1007/978-3-642-21869-9_31.

[10] N. Connon and E. Pirie. Home-based learning (hbl) in higher education post covid: an analysis from staff and student perspectives. *Journal of innovation in polytechnic education*, 4(1), 2022.

[11] T. Daghriri, F. Rustam, W. Aljedaani, A. H. Bashiri, and I. Ashraf. Electroencephalogram signals for detecting confused students in online education platforms with probability-based features. *Electronics*, 11 (18):2855, 2022. http://dx.doi.org/10.3390/electronics11182855.

[12] H. Donmez and N. Ozkurt. Emotion classification from eeg signals in convolutional neural networks. In *2019 Innovations in Intelligent Systems and Applications Conference (ASYU)*, pages 1–6. IEEE, 2019. http://dx.doi.org/10.1109/asyu48272.2019.8946364.

[13] J. K. Grammer and A. Lenartowicz. What do we know about student attention in the classroom? July 2021. https://www.sciencedirect.com/topics/neuroscience/gamma-wave.

[14] A. Hassouneh, A. Mutawa, and M. Murugappan. Development of a real-time emotion recognition system using facial expressions and eeg based on machine learning and deep neural network methods. *Informatics in Medicine Unlocked*, 20:100372, 2020. http://dx.doi.org/10.1016/j.imu.2020.100372.

[15] J. Hemphill, A. Myers, and M. K. Warman. Uc-19 comparison of active and passive attention based tasks using eeg with convolutional neural network. 2021.

[16] H. Jingchao and H. Zhang. Recognition of classroom student state features based on deep learning algorithms and machine learning. *Journal of Intelligent & Fuzzy Systems*, 40(2):2361–2372, 2021. http://dx.doi.org/10.3233/jifs-189232.

[17] N.-H. Liu, C.-Y. Chiang, and H.-C. Chu. Recognizing the degree of human attention using eeg signals from mobile sensors. *sensors*, 13(8):10273–10286, 2013. http://dx.doi.org/10.3390/s130810273.

[18] S. Ma, T. Zhou, F. Nie, and X. Ma. Glancee: An adaptable system for instructors to grasp student learning status in synchronous online classes. In *CHI Conference on Human Factors in Computing Systems*, pages 1–25, 2022. http://dx.doi.org/10.1145/3491102.3517482.

[19] L. A. Moreno López, M. Somacarrera Pérez, M. Díaz Rodríguez, J. Campo Trapero, and J. Cano Sánchez. Problem-based learning versus lectures: Comparison of academic results and time devoted by teachers in a course on dentistry in special patients. 2009. http://dx.doi.org/10.4317/medoral.14.e583.

[20] J. Mostow, K.-m. Chang, and J. Nelson. Toward exploiting eeg input in a reading tutor. In *Artificial Intelligence in Education: 15th International Conference, AIED 2011, Auckland, New Zealand, June 28–July 2011 15*, pages 230–237. Springer, 2011. http://dx.doi.org/10.1007/978-3-642-21869-9.

[21] J. Murphy. An overview of convolutional neural network architectures for deep learning. *Microway Inc*, pages 1–22, 2016.

[22] T. A. Nguyen and Y. Zeng. Analysis of design activities using eeg signals. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, volume 44137, pages 277–286, 2010. http://dx.doi.org/10.1115/detc2010-28477, DOI=10.1115/detc2010-28477.

[23] Z. Ni, A. C. Yuksel, X. Ni, M. I. Mandel, and L. Xie. Confused or not confused? disentangling brain activity from eeg data using bidirectional lstm recurrent neural networks. In *Proceedings of the 8th acm*

international conference on bioinformatics, computational biology, and health informatics, pages 241–246, 2017.

[24] D. L. A. Prastini, S. Supiani, and R. Ratna. The emotional experiences of english student teachers'practicum in learning to teach during covid-19 pandemic. *Proceeding: Islamic University of Kalimantan*, 2022.

[25] B. W. G. Priyanka A. Abhang and S. C. Mehrotra. Technical aspects of brain rhythms and speech parameters. 2016. https://www.sciencedirect.com/topics/neuroscience/gamma-wave.

[26] M. A. Ramírez-Moreno, M. Díaz-Padilla, K. D. Valenzuela-Gómez, A. Vargas-Martínez, J. C. Tudón-Martínez, R. Morales-Menendez, R. A. Ramírez-Mendoza, B. L. Pérez-Henríquez, and J. d. J. Lozoya-Santos. Eeg-based tool for prediction of university students' cognitive performance in the classroom. *Brain sciences*, 11(6):698, 2021. http://dx.doi.org/10.3390/brainsci11060698.

[27] R. Sahu, S. R. Dash, L. A. Cacha, R. R. Poznanski, and S. Parida. Epileptic seizure detection: a comparative study between deep and traditional machine learning techniques. *Journal of integrative neuroscience*, 19(1):1–9, 2020. http://dx.doi.org/10.31083/j.jin.2020.01.24.

[28] R. Sahu, S. R. Dash, L. A. Cacha, R. Poznanski, and S. Parida. Classifier implementation for spontaneous eeg activity during schizophrenic psychosis. *Computación y Sistemas*, 25(3):493–514, 2021. http://dx.doi.org/10.13053/cys-25-3-3874.

[29] R. Sahu, S. R. Dash, and S. Das. Career selection of students using hybridized distance measure based on picture fuzzy set and rough set theory. *Decision Making: Applications in Management and Engineering*, 4(1):104–126, 2021. http://dx.doi.org/10.31181/dmame2104104s.

[30] A. H. Sequeira. Introduction to concepts of teaching and learning. *Available at SSRN 2150166*, 2012. urlhttp://dx.doi.org/10.2139/ssrn.2150166.

[31] A. Sors, S. Bonnet, S. Mirek, L. Vercueil, and J.-F. Payen. A convolutional neural network for sleep stage scoring from raw single-channel eeg. *Biomedical Signal Processing and Control*, 42:107–114, 2018. http://dx.doi.org/10.1016/j.bspc.2017.12.001.

[32] J. Tee and W. Leong. Eeg extraction for meditation. *Journal of Engineering Science and Technology*, 13(7):2125–2135, 2018. http://dx.doi.org/10.1049/pbhe016e_ch1.

[33] C. K. Toa, K. S. Sim, and S. C. Tan. Electroencephalogram-based attention level classification using convolution attention memory neural network. *IEEE Access*, 9:58870–58881, 2021. http://dx.doi.org/10.1109/access.2021.3072731.

[34] S. Vikas and A. Mathur. An empirical study of student perception towards pedagogy, teaching style and effectiveness of online classes. *Education and Information Technologies*, 27(1):589–610, 2022. http://dx.doi.org/10.1007/s10639-021-10793-9.

[35] H. Wang. Confused student eeg brainwave data. Access: November 2022. https://www.kaggle.com/datasets/wanghaohan/confused-eeg.

[36] H. Wang, Y. Li, X. Hu, Y. Yang, Z. Meng, and K.-m. Chang. Using eeg to improve massive open online courses feedback interaction. In *AIED Workshops*, 2013.

[37] A. Whiting. Investigating the impact on student engagement from converting face-to-face classes to online in response to covid-19. *Atlantic Marketing Journal*, 11(1):9, 2022.

A Hybrid Feature Selection Based on Fisher Score and SVM-RFE for Microarray Data

Hind Hamla[1], Khadoudja Ghanem[2]
[1]Laboratory of Modelling and Implementation of Complex System, Department of Computer Science, University of Abdelhamid Mehri Constantine 2, Constantine, Algeria
[2]Laboratory of Modelling and Implementation of Complex System, Department of Computer Science, University of Abdelhamid Mehri Constantine 2, Constantine, Algeria
E-mail: hind.hamla@univ-constantine2.dz , khadoudja.ghanem@univ-constantine2.dz

Keywords: SVM-RFE, Fisher score, gene selection, microarray data

Received: March 22, 2023

Microarray data analysis has played a significant role in disease diagnosis and tumor type identification over the last two decades. However, due to the curse of dimensionality issues, microarray data classification remains a challenging task. This issue arises from a situation where the number of features is large, but the number of samples is small. As a result, dimension reduction techniques, specifically feature selection methods, are critical for removing non-informative features and improving cancer classification. This paper presents a Filter-embedded hybrid feature selection method to address the gene selection challenge in microarray data analysis. First, it selects the features with the highest Fisher score to create a candidate subset for the next embedded stage. Second, the proposed method employs support vector machine-recursive feature elimination (SVM-RFE) on the candidate subset to identify the optimal set of features to enhance cancer classification. Extensive experiments were conducted with ten high-dimensional microarray datasets to assess the efficacy of the proposed approach. The results show that the proposed method improves classifier performance significantly regarding classification accuracy, number of selected features, and computational efficiency.

Povzetek: Predstavljena je hibridna metoda izbire značilk z uporabo Fisherjeve ocene in SVM-RFE za izboljšanje natančnosti klasifikacije raka z analizo mikromrežnih podatkov.

1 Introduction

Over the last two decades, advances in microarray technology have enabled researchers to analyze thousands of genes simultaneously, which has been used in various applications such as disease classification [3]. Microarray data classification is an effective tool for early disease diagnosis and determining disease subtypes [9]. However, due to the curse of dimensionality, where the number of features is remarkably large (often thousands of features) while the number of samples is limited (often tens of samples), this task poses a significant challenge for machine learning algorithms [5]. In addition, a significant proportion of genes are irrelevant or redundant, affecting classifier performance [4]. Thus, gene selection methods have emerged as effective approaches for reducing dimensionality in microarray data. Gene selection methods seek to identify and eliminate redundant and irrelevant features to obtain a subset of the most informative features [32]. These methods have improved classification accuracy while reducing computational costs associated with classifiers [34].

Gene selection methods are broadly classified as filter, wrapper, and embedded methods. Filter methods select features independently from thelearning classifier, based on statistical properties [3]. These methods are fast, but they produce a low classification accuracy [15]. The wrapper methods use the learning algorithm to evaluate a subset of selected features [3].Although they produce higher classification accuracy, they are computationally expensive. Therefore, when dealing with high-dimensional data, these methods are avoided [6]. Embedded methods select features during the learning process [31]. They are appropriate for analyzing microarray data due to their reduced computational demands compared to wrapper methods and enhanced efficiency compared to filter methods. [5]. Hybrid methods, which sequentially combine two or more feature selection methods from the same or different conceptual origins, have recently emerged [6] to leverage the strengths of diverse methodologies.

Many feature selection (FS) surveys for microarray data processing have been conducted. [2] compares feature selection methods including information gain, twoing rule, sum minority, max minority, Gini index, sum of variances, t-statistics, and one-dimension support vector machines. This study use two publicly available glioma gene expression datasets for evaluation. It was discovered that feature selection is important in the classification of gene expression data. In [7], the authors examined the importance and

challenges of feature selection methods when dealing with high-dimensional data such as microarray and instruction detection. The paper emphasized the importance of efficient techniques for managing the computational complexity of high-dimensional data. Furthermore, open issues in feature selection are addressed, particularly in the context of big data and high-dimensional datasets.

The authors of [8] compared five filter methods: the F test, the T-test, the signal-to-noise ratio (S/R), ReliefF, and the Pearson product-moment correlation coefficient (CC). The study used five microarray datasets: leukemia, lung cancer, lymphoma, central nervous system cancer, and ovarian cancer. The results showed that combining the signal-to-noise ratio (S/R) with KNN classifiers produced the best classification accuracy. In [13], the researchers investigated the effect of popular filter methods (ReliefF, Mutual information, Chi-square, F-score, Fisher score, Laplacian, MRMR, and CMIM) on six well-known classifiers (random forest, logistic regression, K-Nearest Neighbour, decision tree, and Support Vector Machine). The experiment was carried out on ten high-dimensional microarray datasets, and the results revealed a distinct trend. Univariate filter feature selection techniques such as Mutual Information, F-score, and Fisher score outperformed multivariate techniques such as MRMR and CMIM. Only a few studies on embedded methods have been conducted. [12] assessed the efficacy of five embedded feature selection techniques: decision trees, random forests, lassos, ridges, and SVM-RFE. The experiment employed ten high-dimensional microarray datasets. The results highlight the SVM-RFE's superior accuracy performance.

This paper combines the embedded method's performance with the filter method's computational efficiency. the proposed method is divided into two stages: The Fisher score filter method is used in the first stage to select the most relevant features due to its effective performance with high-dimensional data [10]. Second, the selected subset is input for the embedded Support Vector Machine Recursive Feature Elimination SVM-RFE method. This combination improves classification accuracy while significantly reducing the number of selected features. Experiments were conducted on ten high-dimensional microarray datasets, including Colon, Central Nervous System CNS, Leukemia, Breast cancer, Lung cancer, Leukemia3-Classes, Leukemia4-Classes, Ovarian, Lymphoma, and MLL. The experimental setup consists of three major components:

- A comparative analysis of the proposed method with other filter methods combined with the same embedded method, SVM-RFE, specifically ReliefF_SVM-RFE and Mutual Information (MI)_SVM-RFE. In addition, we present SVM-RFE results without using a filter method. We avoid comparing the proposed method to the Minimum-Redundancy Maximum-Relevancy (MRMR) and Chi-square filter methods because they have already been studied [19] and [4].
- Investigation the impact of employing six well-established classifiers: Support Vector Machine (SVM), Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), Naïve Bayes (NB), and K-Nearest Neighbour (KNN) on the feature subset selected by our proposed method.
- Finally, to highlight the effectiveness of the proposed method, we compared it with filter-wrapper methods ([30], [34], [23], [21], [24]) and with filter-embedded ones [19] and [4].

The paper is structured as follows: Section 2 examines related works on hybrid feature selection methods. Section 3 briefly describes the Fisher score algorithm and the SVM-RFE algorithm. Section 4 describes the proposed method in depth. Section 5 presents a comprehensive analysis of the experimental findings. Finally, Section 6 provides the conclusion and outlines potential future directions.

2 Related work

Numerous hybrid feature selection methods have been proposed to address the dimensionality reduction challenge and eliminate irrelevant and redundant features from microarray data. While most existing studies in the literature combine filter methods and wrapper methods [1], only a few works investigate the combination of embedded methods and filter methods. In this section, we will review some recent hybrid feature selection methods that have been published in the literature.

2.1 Hybrid wrapper-filter methods

Given their adaptability and efficiency in dealing with large-scale issues, meta-heuristics methods have attracted attention for solving gene selection problems [26]. However, these methods frequently necessitate a significant amount of computational time. Therefore, meta-heuristics have been combined with filter methods to narrow the search space and speed up the feature selection process [21]. Naik et al. [20] proposed a hybrid feature selection method combining the filter and wrapper methods. The Fisher score filter method was used to select a subset of features. The Binary Dragonfly Algorithm was used in the wrapper method to search for an informative subset of features, and the Radial Basis Function Neural Network was used as the learning model that evaluates the selected subset. Shukla [24] designed HMPAGA, a hybrid feature selection method that used an ensemble gene selection method to filter out noisy and redundant genes. It also used a multi-population adaptive genetic algorithm to identify high-risk difference genes. SVM and NB classifiers were used as objective functions.

Shulka et al. [25] proposed a two-stage feature selection method for microarray data recognition. In the first stage, noisy and redundant features were removed using a

multi-layer approach and f-score filter methods. An adaptive genetic algorithm selected the most important features in the second stage. Zhang et al. [30] proposed IG-MBKH, a hybrid feature selection method that combines Information Gain and Modified Binary Krill Herd. The method was validated using nine high-dimensional microarray datasets, improving classification accuracy with fewer features. Zheng et al. [34] presented the K Value Maximum Reliability Minimum Redundancy Improved Grey Wolf Optimizer (KMR2IGWO), a hybrid feature selection method. MRMR was used in the filter stage to select K features, with K determined by the dataset's message. These features were then used as input for the IGHO algorithm, with the SVM classifier used to assess classification accuracy. KMR2IGWO's performance was validated using 14 microarray datasets, highlighting its superiority.

MIMAGA, a combination of mutual information maximization and adaptive genetic algorithm (AGA), was introduced by Lu et al. [17]. MIM was used to choose a subset of 300 features. Then, AGA was applied with the accuracy of ELM classifier serving as the fitness function. Sadeghian et al. [23]introduced a three-stage hybrid feature selection method named Ensemble Information Theory-based binary Butterfly Optimization Algorithm (EIT-bBOA). The method employed Minimal Redundancy-Maximal New Classification Information (MR-MNCI) in the initial phase to eliminate 80% irrelevant features. Subsequently, the Information Gain-binary butterfly optimization algorithm (IG-bBOA) optimized the first phase. In the final phase, an ensemble of ReliefF and the Fisher Score method was applied to the final feature subset. The method was evaluated using six well-known datasets. Ouadfel et al. [21] developed a two-stage feature selection method that used the ReliefF filter method to estimate feature relevance in the first stage. The top-ranked M features where then preselected. The second stage combined the binary Equilibrium Optimizer with a local search strategy based on Pearson coefficient correlation. The proposed method was evaluated on 16 UCI datasets and ten high-dimensional biological datasets.

2.2 Hybrid embedded-filter methods

In terms of computational time, embedded feature selection methods outperform wrapper methods. Though only a few embedded methods have been presented in the literature, [12] conducted a comparative study of the most common ones. SVM-RFE emerged as the most accurate method, with comparable execution time and selected features,. Furthermore, SVM-RFE has consistently demonstrated its efficacy [16]. Thus, many studies have proposed hybridization between filter and embedded methods that concentrate on combining SVM-RFE with filter methods. SVM-RFE has been shown to be effective in identifying informative genes in microarray data [33]. Mundra et al. [19] proposed a hybrid feature selection method combining MRMR and SVM-RFE. The approach's performance was assessed on four well-known microarray datasets. Almutiri and Saeed [4] introduced the ChiSVMRFE feature selection method based on the Chi Square Statistic and SVM-RFE. On ten microarray datasets, the proposed method was evaluated. Mishra et al. [18] combined SVM-RFE with the Bayesian T-test for gene selection, which resulted in improved classification accuracy, fewer selected genes, and a lower classification error rate.

Huang et al. [14] enhanced the SVM-RFE's performance for gene selection by incorporating feature clustering, thereby reducing computational complexity and gene redundancy. Li et al. [16] proposed VSSRFE, an improved version of SVM-RFE that aimed to reduce time using a more efficient SVM classifier implementation. The results demonstrated the proposed method's efficiency in terms of time reduction. Combining wrapper or embedded methods with filter methods consistently improves classifier performance in terms of classification accuracy and computational efficiency, according to the aforementioned works. SVM-RFE, in particular, has demonstrated its ability to improve classification accuracy while optimizing feature dataset. This paper combines SVM-RFE, a leading embedded method, with the best filter method to further improve the results.

3 Background

This section describes the Fisher score and SVM-RFE methods.

3.1 Fisher score

The Fisher score algorithm is a well-known filter feature selection method that is applied to a subset of discriminative features. In summary, the algorithm works as follows: It begins by calculating the average and variance of each feature for each class. Then, it calculates scatter matrices between and within classes to assess the effectiveness of the features in differentiating various classes. The Fisher Scores are then calculated using these matrices, allowing for comparing different features. Features with higher Fisher Scores are considered more important for distinguishing between classes. We can rank the features and select the best based on their scores. The goal is to minimize the distances between samples in the same class while increasing the distances between samples in different classes [29]. Fisher scores f_i are calculated as follows:

$$SC_F\left(f_i\right) = \frac{\sum_{j=1}^{c} n_j \left(\mu_{i,j} - \mu_i\right)^2}{\sum_{j=1}^{c} n_j \sigma_{i,j}^2} \tag{1}$$

where, u_i is the mean of f_i feature, $n_j is$ the number of samples in the class j^{th}, $u_i j$ is the mean of f_i in the j^{th} class, and σ_{ij} is the variance of f_i in the j^{th} class. Usually, a higher Fisher score means the feature is vital for classification.

3.2 Support Vector Machine Recursive Feature Elimination (SVM-RFE)

SVM-RFE is an embedded feature selection method introduced by Guyon et al. [11]. This method employs a weight vector as a criterion for splitting, calculated as follows:

$$\mathrm{W} = \sum_{i=1}^{n} (y_i, x_i, \alpha_i) \tag{2}$$

where, i represents the number of features ranging from 1 to n, y_i is the labeled class of the sample x_i. α_i is the maximum class separation margin estimated from the training set. SVM-RFE works in a recursive manner, similar to iterative refinement. The entire feature set is initially used to train an SVM classifier. The algorithm then iteratively eliminates features with the lowest discriminative power, reducing the risk of the curse of dimensionality and overfitting. The features are then ranked according to their contribution to the classification task. The i^{th} ranking criterion is calculated as follows:

$$\mathrm{R} = W^2 \tag{3}$$

The higher the value of the ranking criterion, the more important the feature. Algorithm 1 depicts the detailed SVM-RFE algorithm.

Algorithm 1 Pseudocode of SVM-RFE

Input: F initial feature set
Output: R rank list

$R = \emptyset$
while $F \neq \varnothing$ **do**
    Train SVM with F
    Compute the weight vector using Equation 2
    Compute the ranking criterion using Equation 3
    Find feature with the lowest ranking criterion
    Update the Ranked list of features
    $\mathrm{R} = \mathrm{R} + \mathrm{F_i}$
    Update set of features
    $\mathrm{F} = \mathrm{F} - \mathrm{F_i}$
end while

4 Proposed method

Because of its low computational requirement, the Fisher score is a simple and efficient feature selection method that is particularly suitable for high-dimensional microarray data classification [28]. However, the Fisher score does not achieve satisfactory classification accuracy. SVM-RFE, on the other hand, has been successfully applied to gene selection problems. It has consistently outperformed several other embedded methods regarding classification accuracy while using a smaller feature set [12]. Nonetheless, one major disadvantage of SVM-RFE is the lengthy feature selection process, especially when dealing with high-dimensional data such as microarray [16]. This work proposes a hybrid feature selection method that combines the computational efficiency of the Fisher score filter method and the high performance of the SVM-RFE embedded method to capitalize on the strengths of both. Fig. 1 shows the flowchart of the hybrid filter-embedded method.

The following are the specifics of the proposed method:

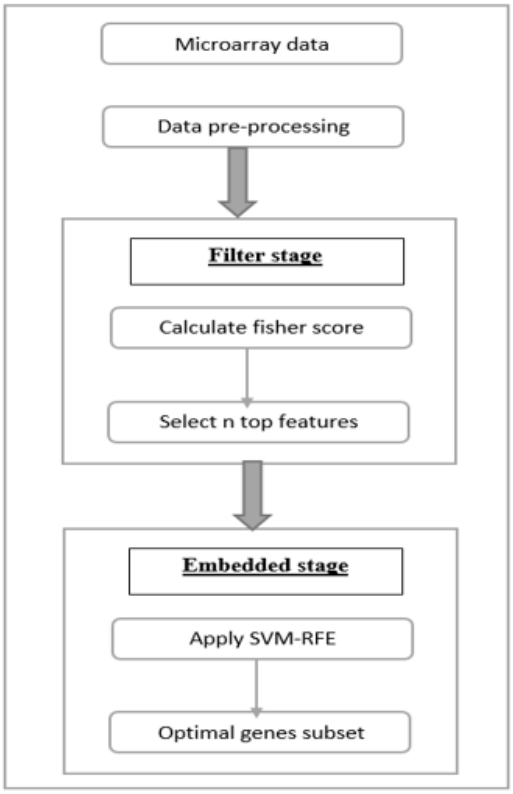


Figure 1: Flowchart of the proposed method.

1. Data pre-processing
   This first step involves replacing missing values with the mean value derived from all known gene values.

2. Filter stage

   Calculate Fisher score
   The Fisher score is used at this stage to eliminate redundant and irrelevant features. Eq. (1) calculates the Fisher score value for each feature, and the features are then sorted based on these values. The higher the Fisher score value, the more informative the feature is for classification.

   Select n top features
   The top n features the Fisher score method indicates are selected as candidate input for the embedded stage.

3. Embedded stage
   SVM-RFE is applied to the previously selected candidate inputs. SVM-RFE uses all the selected features

to train the SVM classifier. Each iteration removes the features with the lowest ranking criterion from the features set. This process is repeated until all features have been removed. The features are sorted in reverse order of removal, with the most recently removed features considered the most important.

4. Select optimal subset
   Finally, SVM-RFE selects a subset of m most important features. The value of n and m is determined through experimentation, with m always being less than n (m<n). This selected subset constitutes the set of informative genes for classification.

The proposed hybrid feature selection approach effectively addresses the challenges of high-dimensional microarray data by combining the Fisher score and SVM-RFE methods. The classification accuracy and interpretability can be improved by selecting a small but informative subset of genes. This has the potential to greatly aid in disease diagnosis and tumor classification. Furthermore, the proposed method balances computational efficiency with classification performance, thereby contributing to bioinformatics and microarray data analysis.

5 Experimental results

In this section, we describe the experimental setup employed to evaluate our hybrid method'efficacy for genes selection from high-dimensional microarray datasets. The goal is to evaluate the efficacy of SVM-RFE when combined with MI, ReliefF, and Fisher scores to determine the best filter method for a microarray dataset using SVM-RFE. Furthermore, the selected gene subset will be tested using a variety of classifiers, including SVM, LR, DT, RF, NB, and KNN. The proposed method is then compared to other existing hybrid feature selection methods. We used a personal computer with an Intel Core i7 processor, 2.9 GHz, and 8 GB of RAM to conduct the experiments. The results presented in this paper are an average of five runs.

5.1 Datasets description

The proposed method is evaluated on ten high-dimensional microarray datasets [35]. The datasets include 2-classes, 3-classes, 4-classes, 5-classes. The number of samples in these datasets is ranged from 60 to 253, while the number of features in these datasets is ranged from 2,000 to 24,481. Table 1 presents detailed information about these datasets. For the evaluation step, we employ 10-fold cross-validation. In this procedure, the datasets are randomly divided into training and testing data subsets, with an 80% and 20% proportion, respectively. The final results are obtained through averaging fold outcomes, a practice employed to address potential issues related to class imbalance.

5.2 Performance measure

Cross-validation [27] is a well-known method for determining the misclassification rate. The data is randomly divided into k subsets of approximately equal size in k-fold cross-validation. The classifier is trained on k-1 folds and then tested on the last fold. This procedure is repeated until every k-fold is used as the test sub-set. The average of the recorded scores is used as the performance metric. In this work, we use several performance metrics, including accuracy, recall, precision, and F-measure, in addition to execution time, to assess the effectiveness of the proposed method.

Accuracy: the ratio of samples that are correctly predicted:

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + FP + TN} \tag{4}$$

Recall: the ratio of the positive samples that are predicted as positive:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{5}$$

Precision: the ratio of the positive prediction that is correct:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{6}$$

F-measure: is a harmonic mean of the precision and recall:

$$\text{F-measure} = \frac{2 * \text{ Precision } * \text{ Recall}}{\text{Presision } + \text{ Recall}} \tag{7}$$

The indicators for evaluation are:
Acc: Accuracy
Rec: Recall
Pre: Precision
Fmes: F-measure
Nb-FS: number of selected features

5.3 Combination of filter methods with SVM-RFE

Table 2 displays the results of various feature selection methods combined with SVM-RFE. Using the gene subsets selected by these methods, we assess each classifier's accuracy, recall, precision, F-measure, and execution time. According to Table 2, the classification accuracy of the SVM classifier on the original dataset is not very interesting, especially for the breast and CNS datasets, where the classification accuracy did not reach 70%. However, feature selection methods enhance classification performance regarding accuracy, recall, precision, and F-measure. The proposed method consistently performs comparable or better than other feature selection methods. Notably, the execution time was reduced for all datasets after using feature selection methods. Moreover, the proposed method demonstrates remarkable efficacy by achieving 100% accuracy for nine out of ten datasets, using less than 1%

of the original genes. This finding demonstrates our proposed method's ability to identify informative genes for microarray data analysis. In some cases, SVM-RFE outperforms other feature selection methods, implying that the filter methods have eliminated some important features. Fig. 3 presents the number of selected features. The proposed method clearly achieves higher classification accuracy with fewer than 20 features.

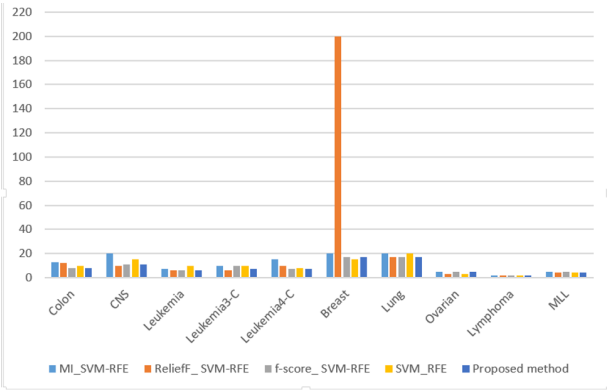


Figure 2: The number of selected features.

5.4 Evaluation of the application of different classifiers on the subset selected by the proposed method

Using the subset of features selected by our proposed method, we compare six popular classifiers: SVM, LR, DT, RF, NB, and KNN. The results in Table 3 indicate that:

- The six classifiers SVM, LR, DT, RF, NB, and KNN achieve comparable classification performance when using the subset of features selected by the proposed method. Based on this result, various classifiers can perform well when using the selected gene subset, indicating that the subset contains relevant and discriminative information.
- SVM consistently outperforms other classifiers regarding accuracy, recall, precision, and F-measure across all datasets. Due to its ability to find optimal hyperplanes for separating data points, SVM is effective for various datasets.
- DT has generally lower accuracy than other classifiers. This may be because it tends to overfit training data, especially when dealing with high-dimensional data sets.
- The fastest classifier is KNN, while RF is the slowest. However, RF still delivers competitive results despite its longer execution time, indicating its ability to handle high-dimensional data effectively.
- On the selected gene subset, the results obtained by the proposed method match those of the SVM classifier. Thus, the proposed method is valid and reliable due to this consistency.

5.5 Comparison of the proposed method with other hybrid methods

The performance of the proposed method was compared with several hybrid feature selection methods available in the literature, including filter-wrapper methods (IG-MBKH [30], KMR2IGWO [34], EIT-bBOA [23], RBEO-LS [21], and HMPAGA [24]). And filter-embedded methods (ChiSVM-RFE [4] and SVM-RFE with MRMR [19]). The comparison is based on classification accuracy and the number of selected features, as shown in Table 4. The symbol "-" means that information is unavailable.

The results Table 4 indicate that the proposed method achieves a comparable classification accuracy while selecting a reduced subset of features. It attains the highest classification accuracy for all datasets except the Colon dataset, with a small number of genes. Moreover, though a direct execution time comparison was not performed, embedded methods consume less time than wrapper methods, as demonstrated in [22]. This finding suggests that the proposed method is more efficient considering the execution time.

6 Conclusion and future work

Microarray data is well known for being high-dimensional and highly redundant. Thus, feature selection methods are critical in removing irrelevant and redundant features. This paper proposes a hybrid feature selection method that combines the Fisher score and SVM-RFE. The proposed method is divided into two stages. The Fisher score filter method selects a candidate subset of features in the first stage. The subset is then used as input for the SVM-RFE to further reduce the number of features to less than 20. The proposed method outperforms other methods such MI_-SVM-RFE, ReliefF_SVM-RFE, and SVM-RFE in terms of accuracy, recall, precision, F-measure, number of selected features, and runtime in experimental evaluations on ten high-dimensional datasets, some of which had over 20,000 features. In addition, we compared the proposed method to several methods proposed in the literature. According to the results, the proposed method consistently achieved higher classification accuracy and selected a smaller number of features for most datasets. These findings demonstrate the efficacy of the proposed method in addressing the challenges of high-dimensional microarray data analysis.

Table 1: Datasets description.

Datasets	Number of instances	Number of features	Number of classes
Colon tumor	62	2000	2
CNS	60	7129	2
Leukemia	72	7129	2
Breast cancer	97	24481	2
Lung_cancer	203	12600	5
Ovarian cancer	253	15154	2
Leukemia 3 classes	72	7129	3
Leukemia 4 classes	72	7129	4
Lymphoma	62	4026	3
MLL	72	12582	3

Table 2: The experimental result of filter-embedded methods.

Dataset	Methods	Acc	Rec	Pre	Fmes	Time	Nb-FS
Colon	No_FS	79.16%	83.33%	77.66%	79.30%	0.0075	-
	SVM-RFE	96.0%	96.66%	97.5%	96.57%	0.001	10
	ReliefF_SVM-RFE	95.5%	91.66%	100.0%	94.66%	0.002	12
	MI_SVM-RFE	98.00%	96.66%	100.0%	98.00%	0.002	10
	Proposed	98.00%	100.0%	97.5%	98.57%	0.001	8
CNS	No_FS	68.83%	60.0%	63.33%	56.66%	0.0322	-
	SVM-RFE	98.00%	95.0%	100.0%	96.66%	0.001	15
	ReliefF_SVM-RFE	98.00%	95.0%	100.0%	96.66%	0.003	9
	MI_SVM-RFE	98.00%	95.0%	100.0%	96.66%	0.003	15
	Proposed	100.0%	100.0%	100.0%	100.0%	0.003	11
Leukemia	No_FS	96.33%	95.0%	96.66%	94.66%	0.061	-
	SVM-RFE	94.57%	86.66%	100.0%	91.33%	0.001	8
	ReliefF_SVM-RFE	100.0%	100.0%	100.0%	100.0%	0.001	6
	MI_SVM-RFE	100.0%	100.0%	100.0%	100.0%	0.002	6
	Proposed	100.0%	100.0%	100.0%	100.0%	0.001	6
Leukemia 3-C	No_FS	94.57%	97.0%	94.0%	96.0%	0.056	-
	SVM-RFE	100.0%	98.0%	99.0%	99.0%	0.001	10
	ReliefF_SVM-RFE	100.0%	100.0%	100.0%	100.0%	0.004	7
	MI_SVM-RFE	98.57%	99.0%	95.0%	97.0%	0.005	7
	Proposed	100.0%	98.0%	99.0%	99.0%	0.0	7
Leukemia 4-C	No_FS	91.29%	98.0%	96.0%	96.0%	0.0663	-
	SVM-RFE	100.0%	99.0%	99.0%	99.0%	0.0009	8
	ReliefF_SVM-RFE	98.57%	99.0%	99.0%	99.0%	0.001	8
	MI_SVM-RFE	100.0%	99.0%	99.0%	99.0%	0.0	7
	Proposed	100.0%	99.0%	99.0%	99.0%	0.0	7
Breast	No_FS	65.47%	50.0%	57.49%	52.66%	0.5012	-
	SVM-RFE	98.57%	100.0%	97.5%	98.57%	0.001	15
	ReliefF_SVM-RFE	94.64%	93.33%	95.0%	93.14%	0.0	200
	MI_SVM-RFE	98.57%	96.66%	100.0%	98.0%	0.0	20¨
	Proposed	100.0%	100.0%	100.0%	100.0%	0.002	17
Lung cancer	No_FS	93.90%	91.0%	96.0%	93.0%	0.2905	-
	SVM-RFE	100.0%	99.0%	99.0%	99.0%	0.003	20
	ReliefF_SVM-RFE	98.75%	96.0%	97.0%	97.0%	0.002	17
	MI_SVM-RFE	98.75%	98.0%	96.0%	97.0%	0.003	20
	Proposed	100.0%	97.0%	97.0%	97.0%	0.003	17
Ovarian	No_FS	100.0%	100.0%	100.0%	100.0%	0.4832	-
	SVM-RFE	100.0%	100.0%	100.0%	100.0%	0.0	3
	ReliefF_SVM-RFE	100.0%	100.0%	100.0%	100.0%	0.0	3
	MI_SVM-RFE	100.0%	100.0%	100.0%	100.0%	0.0	5
	Proposed	100.0%	100.0%	100.0%	100.0%	0.001	5

Lymphoma	No_FS	100.0%	98.0%	90.0%	94.0%	0.0507	-
	SVM-RFE	100.0%	100.0%	100.0%	100.0%	0.0005	2
	ReliefF_SVM-RFE	100.0%	100.0%	100.0%	100.0%	0.0	2
	MI_SVM-RFE	100.0%	100.0%	100.0%	100.0%	0.0	2
	Proposed	100.0%	100.0%	100.0%	100.0%	0.001	2
MLL	No_FS	98.57%	100.0%	100.0%	100.0%	0.1222	-
	SVM-RFE	100.0%	100.0%	100.0%	100.0%	0.0005	4
	ReliefF_SVM-RFE	100.0%	98.0%	99.0%	98.0%	0.0	4
	MI_SVM-RFE	100.0%	98.0%	98.0%	98.0%	0.0	5
	Proposed	100.0%	100.0%	100.0%	100.0%	0.001	4

Table 3: The experimental results of applying different classifiers on the selected subset.

		SVM	LR	DT	RF	NB	KNN
Colon	Acc	98.00%	96.0%	79.23%	84.6%	79.0%	85.33%
	Rec	100.0%	96.66%	80.66%	86.99%	73.33%	91.66%
	Pre	97.5%	97.5%	71.93%	81.66%	80.0%	85.83%
	Fmes	98.5	96.57	76.21%	81.76%	75.14%	87.04%
	Time	0.001	0.001	0.001	0.061	0.0006	0.0001
CNS	Acc	100.0%	96.33%	77.83%	86.33%	85.50%	90.33%
	Rec	100.0%	90.0%	59.0%	69.0%	85.0%	85.0%
	Pre	100.0%	100.0%	96.66%	79.0%	81.66%	93.33%
	Fmes	100.0%	93.33%	60.73%	72.99%	81.33%	86.0%
	Time	0.0016	0.001	0.001	0.0970	0.001	0.0
Leukemia	Acc	100.0%	96.33%	92.80%	98.29%	100.0%	100.0%
	Rec	100.0%	95.0%	87.32%	96.0%	100.0%	100.0%
	Pre	100.0%	96.66%	95.32%	100.0%	100.0%	100.0%
	Fmes	100.0%	94.66%	90.37%	96.66%	100.0%	100.0%
	Time	0.0019	0.0013	0.0016	0.1091	0.0006	0.0005
Leukemia3-C	Acc	100.0%	98.00%	95.29%	96.73%	96.33%	96.90%
	Rec	98.0%	98.0%	94.4%	93.20%	93.0%	
	Pre	99.0%	99.0%	93.60%	97.20	97.0%	86.0%
	Fmes	99.0%	99.0%	93.40%	95.20	95.0%	96.0%
	Time	0.0004	0.001	0.001	0.0660	0.0011	90.0%
							0.0004
Leukemia4-C	Acc	100.0%	97.14%	87.41%	89.68%	88.35%	93.21%
	Rec	99.0%		83.2%	79.4%	66.0%	87.0%
	Pre	99.0%	90.0%	90.2%	90.0%	64.0%	97.0%
	Fmes	99.0%	98.0%	85.8%	82.80%	64.0%	91.0%
	Time	0.0013	93.0%	0.0009	0.085	0.0006	0.0005
			0.0015				
Breast	Acc	100.0%	95.0%	72.61%	84.09%	69.36%	85.95%
	Rec	100.0%	90.0%	60.16%	71.15%	35.0%	75.83%
	Pre	100.0%	97.5%	53.76%	87.83%	61.66%	92.66%
	Fmes	100.0%	91.57%	54.26	72.98%	43.33%	80.40%
	Time	0.002	0.001	0.001	0.070	0.001	0.0002
Lung	Acc	100.0%	92.54%	90.98%	91.72%	92.12%	92.56%
	Rec	97.0%	80.0%	78.20%	89.6%	75.0%	87.0%
	Pre	97.0%	91.0%	80.8%	93.0%	69.0%	95.0%
	Fmes	97.0%	84.0%	77.4%	91.0%	72.0%	90.0%
	Time	0.0021	0.0063	0.0030	0.1100	0.0009	0.0007

Ovarian	Acc	100.0%	100.0%	98.10%	99.04%	100.0%	100.0%
	Rec	100.0%	100.0%	97.21%	99.25%	100.0%	100.0%
	Pre	100.0%	100.0%	98.43%	99.55%	100.0%	100.0%
	Fmes	100.0%	100.0%	98.17%	98.92%	100.0%	100.0%
	Time	0.0014	0.0029	0.0016	0.0640	0.0014	0.0019
Lymphoma	Acc	100.0%	96.66%	96.33%	100.0%	98.33%	100.0%
	Rec	100.0%	87.0%	100.0%	98.8%	93.0%	100.0%
	Pre	100.0%	94.0%	100.0%	99.4%	99.0%	100.0%
	Fmes	100.0%	88.0%	100.0%	99.6%	96.0%	100.0%
	Time	0.0012	0.0010	0.0013	0.0873	0.0009	0.0008
MLL	Acc	100.0%	90.97%	93.95%	97.57%	96.90%	93.47%
	Rec	100.0%	87.0%	97.0%	97.2%	97.0%	91.0%
	Pre	100.0%	86.0%	97.0%	98.0%	96.0%	91.0%
	Fmes	100.0%	87.0%	97.0%	97.8%	96.0%	91.0%
	Time	0.0009	0.0008	0.0008	0.0730	0.0011	0.0001

Table 4: Comparison of the proposed method with other hybrid methods.

	Colon		CNS		Leukemia		Leukemia3-C		Leukemia4-C	
	ACC	NB_FS	ACC	NB_FS	ACC	NB_FS	ACC	NB_FS	ACC	NB_FS
[30]	96.47%	17.1	90.34%	14.7	100.0%	4.2	99.44	15.8	99.44	15.8
[34]	98.80%	8	-	-	-	-	-	-	-	-
[23]	92.0%	30	84.0%	30	-	-	-	-	-	-
[21]	100.%	6.13	-	-	100.0%	7.7	-	-	-	-
[24]	98.87%	16	-	-	98.84%	12.9	-	-	-	-
[4]	96.67%	10	-	-	-	-	-	-	-	-
[19]	91.68%	78	-	-	98.35%	37	-	-	-	-
Proposed	98.00%	8	100.0%	11	100.0%	6	100.0%	7	100.0%	7
	Breast		Lung cance		Ovarian		Lymphoma		MLL	
	ACC	NB_FS	ACC	NB_FS	ACC	NB_FS	ACC	NB_FS	ACC	NB_FS
[30]	-	-	96.12%	23.8	100.0%	3.4	-	-	99.72%	11.1
[34]	-	-	99.3%	12	-	-	99.9%	10.6	/	-
[23]	-	-	-	-	-	-	94.0%	30	-	-
[21]	-	-	99.35%	9.1	-	-	-	-	-	-
[24]	94.15%	16.8	99.52%	12.9	-	-	-	-	-	-
[4]	-	-	-	-	100.0%	10	-	-	-	-
[19]	-	-	-	-	-	-	-	-	-	-
Proposed	100.0%	17	100.0%	17	100.0%	5	100.0%	2	100.0%	4

However, the number of selected features in the filter stage is determined empirically. In the future, we aim to create a mathematical function to determine the threshold based on the input dataset. Another objective is to improve the overall performance by incorporating more filter methods into the proposed method to eliminate irrelevant and redundant features.

References

[1] Muhammed Abd-Elnaby, Marco Alfonse, and Mohamed Roushdy. Classification of breast cancer using microarray gene expression data: A survey. *Journal of Biomedical Informatics*, 117:103764, 2021. URL http://dx.doi.org/10.1016/j.jbi.2021.103764.

[2] Heba Abusamra. A comparative study of feature selection and classification methods for gene expression data of glioma. *Procedia Computer Science*, 23:5–14, 2013. URL http://dx.doi.org/10.1016/j.procs.2013.10.003.

[3] Russul Alanni, Jingyu Hou, Hasseeb Azzawi, and Yong Xiang. A novel gene selection algorithm for cancer classification using microarray datasets. *BMC medical genomics*, 12(1):1–12, 2019. URL http://dx.doi.org/10.1186/s12920-018-0447-6.

[4] Talal Almutiri and Faisal Saeed. Chi square and support vector machine with recursive feature elimination for gene expression data classification. In *2019 First International Conference of Intelligent Computing and Engineering (ICOICE)*, pages 1–6. IEEE, 2019. URL http://dx.doi.org/10.1109/icoice48418.2019.9035165.

[5] Verónica Bolón-Canedo and Amparo Alonso-Betanzos. *Microarray Bioinformatics*. Springer, 2019. URL http://dx.doi.org/10.1007/978-1-4939-9442-7.

[6] Verónica Bolón-Canedo, Noelia Sánchez-Marono, Amparo Alonso-Betanzos, José Manuel Benítez, and Francisco Herrera. A review of microarray datasets and applied feature selection methods. *Information Sciences*, 282:111–135, 2014. URL http://dx.doi.org/10.1016/j.ins.2014.05.042.

[7] Verónica Bolón-Canedo, Noelia Sánchez-Maroño, and Amparo Alonso-Betanzos. Feature selection for high-dimensional data. *Progress in Artificial Intelligence*, 5:65–75, 2016. URL http://dx.doi.org/10.1007/s13748-015-0080-y.

[8] Sara Haddou Bouazza, Khalid Auhmani, Abdelouhab Zeroual, and Nezha Hamdi. Selecting significant marker genes from microarray data by filter approach for cancer diagnosis. *Procedia Computer Science*, 127:300–309, 2018. URL http://dx.doi.org/10.1016/j.procs.2018.01.126.

[9] Zhipeng Cai, Randy Goebel, Mohammad R Salavatipour, and Guohui Lin. Selecting dissimilar genes for multi-class classification, an application in cancer subtyping. *BMC bioinformatics*, 8(1):1–15, 2007. URL http://dx.doi.org/10.1186/1471-2105-8-206.

[10] Hakan Gunduz. An efficient dimensionality reduction method using filter-based feature selection and variational autoencoders on parkinson's disease classification. *Biomedical Signal Processing and Control*, 66:102452, 2021. URL http://dx.doi.org/10.1016/j.bspc.2021.102452.

[11] Isabelle Guyon, Jason Weston, Stephen Barnhill, and Vladimir Vapnik. Gene selection for cancer classification using support vector machines. *Machine learning*, 46(1):389–422, 2002. URL https://link.springer.com/article/10.1023/A:1012487302797.

[12] Hind Hamla and Khadoudja Ghanem. Comparative study of embedded feature selection methods on microarray data. In *IFIP International Conference on Artificial Intelligence Applications and Innovations*, pages 69–77. Springer, 2021. URL http://dx.doi.org/10.1007/978-3-030-79150-6_6.

[13] Hind Hamla and Khadoudja Ghanem. A comparative study of filter feature selection methods on microarray data. In *International Conference on Computing and Information Technology*, pages 186–201. Springer, 2022. URL http://dx.doi.org/10.1007/978-3-031-25344-7_18.

[14] Xiaojuan Huang, Li Zhang, Bangjun Wang, Fanzhang Li, and Zhao Zhang. Feature clustering based support vector machine recursive feature elimination for gene selection. *Applied Intelligence*, 48(3):594–607, 2018. URL http://dx.doi.org/10.1007/s10489-017-0992-2.

[15] Hengxun Li, Wei Guo, Guoying Wu, and Yanxia Li. A rf-pso based hybrid feature selection model in intrusion detection system. In *2018 IEEE Third International Conference on Data Science in Cyberspace (DSC)*, pages 795–802. IEEE, 2018. URL http://dx.doi.org/10.1109/dsc.2018.00128.

[16] Zifa Li, Weibo Xie, and Tao Liu. Efficient feature selection and classification for microarray data. *PloS one*, 13(8):e0202167, 2018. URL http://dx.doi.org/10.1371/journal.pone.0202167.

[17] Huijuan Lu, Junying Chen, Ke Yan, Qun Jin, Yu Xue, and Zhigang Gao. A hybrid feature selection algorithm for gene expression data classification. *Neurocomputing*, 256:56–62, 2017. URL http://dx.doi.org/10.1016/j.neucom.2016.07.080.

[18] Shruti Mishra and Debahuti Mishra. Svm-bt-rfe: An improved gene selection framework using bayesian t-test embedded in support vector machine (recursive feature elimination) algorithm. *Karbala International Journal of Modern Science*, 1(2):86–96, 2015. URL http://dx.doi.org/10.1016/j.kijoms.2015.10.002.

[19] Piyushkumar A Mundra and Jagath C Rajapakse. Svm-rfe with mrmr filter for gene selection. *IEEE transactions on nanobioscience*, 9(1):31–37, 2009. URL http://dx.doi.org/10.1109/tnb.2009.2035284.

[20] Akshata Naik, Venkatanareshbabu Kuppili, and Damodar Reddy Edla. Binary dragonfly algorithm and fisher score based hybrid feature selection adopting a novel fitness function applied to microarray data. In *2019 International Conference on Applied Machine Learning (ICAML)*, pages 40–43. IEEE, 2019. URL http://dx.doi.org/10.1109/icaml48257.2019.00015.

[21] Salima Ouadfel and Mohamed Abd Elaziz. Efficient high-dimension feature selection based on enhanced equilibrium optimizer. *Expert Systems with Applications*, 187:115882, 2022. URL http://dx.doi.org/10.1016/j.eswa.2021.115882.

[22] Beatriz Remeseiro and Veronica Bolon-Canedo. A review of feature selection methods in medical applications. *Computers in biology and medicine*, 112:103375, 2019. URL http://dx.doi.org/10.1016/j.compbiomed.2019.103375.

[23] Zohre Sadeghian, Ebrahim Akbari, and Hossein Nematzadeh. A hybrid feature selection method based on information theory and binary butterfly optimization algorithm. *Engineering Applications of Artificial Intelligence*, 97:104079, 2021. URL http://dx.doi.org/10.1016/j.engappai.2020.104079.

[24] Alok Kumar Shukla. Multi-population adaptive genetic algorithm for selection of microarray biomarkers. *Neural Computing and Applications*, 32(15):11897–11918, 2020. URL http://dx.doi.org/10.1007/s00521-019-04671-2.

[25] Alok Kumar Shukla, Pradeep Singh, and Manu Vardhan. A hybrid gene selection method for microarray recognition. *Biocybernetics and Biomedical Engineering*, 38(4):975–991, 2018. URL http://dx.doi.org/10.1016/j.bbe.2018.08.004.

[26] Alok Kumar Shukla, Diwakar Tripathi, B Ramachandra Reddy, and D Chandramohan. A study on metaheuristics approaches for gene selection in microarray data: algorithms, applications and open challenges. *Evolutionary Intelligence*, 13(3):309–329, 2020. URL http://dx.doi.org/10.1007/s12065-019-00306-6.

[27] Mervyn Stone. Cross-validatory choice and assessment of statistical predictions. *Journal of the royal statistical society: Series B (Methodological)*, 36(2):111–133, 1974. URL http://dx.doi.org/10.1111/j.2517-6161.1974.tb00994.x.

[28] Lin Sun, Xiao-Yu Zhang, Yu-Hua Qian, Jiu-Cheng Xu, Shi-Guang Zhang, and Yun Tian. Joint neighborhood entropy-based gene selection method with fisher score for tumor classification. *Applied Intelligence*, 49(4):1245–1259, 2019. URL http://dx.doi.org/10.1007/s10489-018-1320-1.

[29] J Yang, YL Liu, CS Feng, and GQ Zhu. Applying the fisher score to identify alzheimer's disease-related genes. *Genet Mol Res*, 15(2), 2016. URL http://dx.doi.org/10.4238/gmr.15028798.

[30] Ge Zhang, Jincui Hou, Jianlin Wang, Chaokun Yan, and Junwei Luo. Feature selection for microarray data classification using hybrid information gain and a modified binary krill herd algorithm. *Interdisciplinary Sciences: Computational Life Sciences*, 12:288–301, 2020. URL http://dx.doi.org/10.1007/s12539-020-00372-w.

[31] Huaqing Zhang, Jian Wang, Zhanquan Sun, Jacek M Zurada, and Nikhil R Pal. Feature selection for neural networks using group lasso regularization. *IEEE Transactions on Knowledge and Data Engineering*, 32(4):659–673, 2019. URL http://dx.doi.org/10.1109/tkde.2019.2893266.

[32] Xue Zhang, Zhiguo Shi, Xuan Liu, and Xueni Li. A hybrid feature selection algorithm for classification unbalanced data processsing. In *2018 IEEE International Conference on Smart Internet of Things (SmartIoT)*, pages 269–275. IEEE_, 2018. URL http://dx.doi.org/10.1109/smartiot.2018.00055.

[33] Ying Zhang, Qingchun Deng, Wenbin Liang, and Xianchun Zou. An efficient feature selection strategy based on multiple support vector machine technology with gene expression data. *BioMed research international*, 2018, 2018. URL http://dx.doi.org/10.1155/2018/7538204.

[34] Yuefeng Zheng, Ying Li, Gang Wang, Yupeng Chen, Qian Xu, Jiahao Fan, and Xueting Cui. Retracted: A hybrid feature selection algorithm for microarray data. *Concurrency and Computation: Practice and Experience*, 31(12):e4716, 2019. URL http://dx.doi.org/10.1002/cpe.4716.

[35] Zexuan Zhu, Yew-Soon Ong, and Manoranjan Dash. Markov blanket-embedded genetic algorithm for gene selection. *Pattern Recognition*, 40(11):3236–3248, 2007. URL http://dx.doi.org/10.1016/j.patcog.2007.02.007.

https://doi.org/10.31449/inf.v48i1.4839 Informatica **48** (2024) 69–78 **69**

Prediction of Author's Profile Basing on Fine-Tuning BERT Model

Bassem Bsir[1,2], Nabil Khoufi[3], Mounir Zrigui[1,2]
[1] ISITCom, University of Sousse, 4011 Hammam Sousse,
[2] Laboratory in Algebra, Numbers Theory and Intelligent Systems, University of Monastir, Monastir, Tunisia
[3] ANLP Research Group, FSEGS, Sfax, Tunisia
E-mail : Bsir.bassem@yahoo.fr, nabil.khoufi@outlook.com, mounir.zrigui@fsm.rnu.tn

Keywords: BERT, Author profiling (AP), fine tuning, PAN 2018 Corpus dataset, NLP, Transformer-model, deep learning, Self-attention Transformers.

Received: Februar 5, 2023

The task of author profiling consists in specifying the infer-demographic features of the social networks' users by studying their published content or the interactions between them. In the literature, many research works were conducted to enhance the accuracy of the techniques used in this process. In fact, the existing methods can be divided into two types: simple linear models and complex deep neural network models. Among them, the transformer-based model exhibited the highest efficiency in NLP analysis in several languages (English, German, French, Turk, Arabic, etc.). Despite their good performance, these approaches do not cover author profiling analysis and, thus, should be further enhanced. So, we propose in this paper a new deep learning strategy by training a customized transformer-model to learn the optimal features of our dataset. In this direction, we fine-tune the model by using the transfer learning approach to improve the results with random initialization. We have achieved about 79% of accuracy by modifying model to apply the retraining process using PAN 2018 authorship dataset.

Povzetek: Članek predstavlja novo metodo za napovedovanje avtorjevega profila, ki temelji na modelu Fine-Tuning BERT.

1 Introduction

As defined by [6], author profiling (AP) is a Natural Language Processing (NLP) research domain that aims at deducing social-demographic data on the author or user of a specific application or software service. It consists first in extracting automatically, from the text, information showing the authors' gender, age and other demographic features. These data are used in several fields such as in forensics, security and marketing.

In the last decades, the main methods utilized in Natural Language Processing (NLP) are deep neural networks relying on Transformers. As instance of these techniques, Self-attention Transformers and, particularly, the self-supervised-trained variants, also called BERT (Bidirectional Encoder Representations from Transformers) models [26], showed high performance in several tasks such as text classification [18], Sentiment Analysis [14], question answering [38], natural language inference [45][40], etc. In fact, these novel methods have revolutionized NLP tasks by dropping the recurrent part and only keeping attention mechanisms.

Indeed, transformers-based pre-trained language models, such as OpenAI GPT [3], BERT [26], RoBERTa [41], have proven their good performance in learning language representation by employing huge quantity of unlabeled data [4]. Nevertheless, their training is often performed on large monolingual English corpora or on multi-lingual corpora involving more than one hundred languages. Recent study has demonstrated that the performance of the fine-tuning from multi-lingual models is almost similar to that of monolingual models for low resource languages [1].

Despite the wide use of the afore-mentioned methods, their accuracy for Arabic Author profiling should be further enhanced, particularly in tokenization level in the task of data processing. For this reason and as no previous has focused on the identification of the gender of the author form Arabic texts published on social networks using these models, we examine the efficacy of several multilingual models for AP tasks. Then we choose to fine-tuned model. We are focusing on the Ara-BERTv2-base model in order to change its parameters and search for the most suitable ones for the gender identification task.

The present manuscript is structured as follows. Section II presents the works deal-ing with author profiling. In Section III, we depict the introduced approach, the employed datasets, and the training details. Section IV shows a comparative study of the obtained findings with those obtained in the stat-of- the art and discusses the experimental results. We end the paper with a short conclusion.

2 State of the art

Several approaches and methods have been recently developed and applied in AP.

We can classify these approaches into two categories. The first category includes traditional machine learning

methods [2]. The second category includes deep learning techniques [11] [37] [14].

Traditional machine learning methods have been explored by researchers for the task of gender prediction in author profiling. Indeed, Poulston et al. in 2017 used the genism Python library for LDA topic extraction with SVM classifiers. Their results proved that the topic models are useful in developing author's profiling systems. Argamon et al. in 2012 analyzed an analogous sample taken from the BNC consist-ing of fiction and non-fiction documents. Their corpus includes 604 texts equally divided by genre and controlled for authorial origin for a total size of 25 million words. Their anal ysis consists in a frequency count of basic and most frequent function words, part-of speech tags and part-of-speech two-grams and three-grams. The counts were processed by a machine-learning algorithm used to classify the texts according to the author's gender. They obtained an accuracy of 80%.

In 2017, Martinc et al based on the corpus collected from Twitter text written by four different languages (Arabic, English, Portuguese and Spanish), they obtained 70.02 by using logistic regression by combining character, word POS n-grams, emo-ji's, senti ments, character flood in gland lists of words per variety in PAN 2017 competition.

González-Gallardo et al. predicted the gender, age and personality traits of Twitter users. They accounted stylistic features represented by character N grams and POS N-grams to classify tweets. They applied Support Vector Machine (SVM) with a linear kernel called LinearSVC and obtained 83.46% for gender detection [16].

While these methods have shown some success, they are often limited by the quality of the features used and the complexity of the task, which can lead to lower accuracy compared to deep learning methods.

In the last three years have been many recent modeling improvements on NLP tasks. These models have largely focused on building separate models for each language or for a small group of related languages. However, Transfer Learning from large-scale pre-trained models in Natural Language Processing (NLP) becomes more prevalent they often have several hundred million parameters and current research on pre-trained models indicates that training even larger models still leads to better perfor-mances on NLP tasks [5][41].

Indeed, Devlin and Chang proved that the main challenge in NLP consists of the small quantity of the training data [26]. To deal with this issue, they have suggested transformer-based models trained on huge unlabeled datasets (e.g., Wikipedia's dataset). The authors were able to apply the pretrained models on smaller datasets without the need for developing training models from scratch. De-spite the fact that the proposed technique provided high accuracy in executing various NLP tasks [26][29], "fine-tuning" should be performed on the pretrained models before being applied on smaller da-tasets. As example of the pretrained models, we can mention the bidirectional en-coder representations from transformers (BERT) characterized by its bidirectionality.

In 2017, Vaswani et al introduce a new language representation model called BERT, which stands for Bidirectional Encoder Representations from Transformers. It's designed to pre-train deep bidirectional representations from unlabeled text by joint-ly conditioning on both right and left context in all layers. It obtains new state-of-the-art results on eleven natural language processing tasks, including pushing the GLUE score to 80.5, MultiNLI accuracy to 86.7%, SQuAD v1.1 question answering Test F1 to 93.2 and SQuAD v2.0 Test F1 to 83.1[28]. Unlike Radford et al. [3], which uses unidirectional language models for pre-training, BERT uses masked language models to enable pretrained deep bidirectional repre-sentations. It's also reducing the need for many heavily engineered task specific architectures. BERT is the first finetuning based representation model that achieves state-of-the-art performance on a large suite of sentence-level and token-level tasks, outperforming much task-specific architecture [3].

Ai, M. proposed the tasks of Russian news event detection. They present datasets for the Russian news event clustering, headline selection, and headline generation tasks along with baselines. Authors demonstrated the successful models were classifica-tion-based BERT models. However, it turns out clustering embeddings can be almost as effective when trained with correct pooling and loss function [2].

Rangel et al. in 2021 explored the use of the BERT language model for author pro-filing in multiple languages. The authors found that the BERT model achieved high accuracy rates for gender and age prediction in several languages. they achieved an accuracy of 96.4% for gender prediction and 77.5% for age prediction on the English dataset, and an accuracy of 92.3% for gender prediction and 62.5% for age predic-tion on the Spanish dataset. The study also evaluated author profiling in French, Portuguese, and Italian, achieving similarly high accuracy rates. [10].

In the same year, 2021, other study used a combination of n-gram-based features and a random forest classifier to predict the gender and age of authors was presented by Khader and Al-Ani. The results showed that the approach achieved high accuracy rates, particularly for gender prediction. Indeed, for gender prediction, the approach achieved an accuracy rate of 97.9%, while for age prediction, the accuracy rate was 91.3%.

In 2019, Victor SANH et al. Show that it is possible to reach similar performances on many downstream-tasks using much smaller language models pre-trained with knowledge distillation, resulting in models that are lighter and faster at inference time. It is possible to reduce the size of a BERT model by 40%, while retaining 97% of its language understanding capabilities and being 60% faster [18].

For instance, while processing the word bank (which have two meanings (financial institution or the shore of a river), the BERT model analyzes all words in the sen-tence at both valences and produces a score showing the best representation of the meaning of the words in a specific context.

The main objective of this research work is to study the impact of the common pre-processing methods utilized

to determine the author's age and gender in case of using the pretrained model called BERT.

The following section presents first the existing research works based on the preprocessing techniques applied in author profiling. Then, the implementation of the five considered cases of the preprocessing methods and the different steps of each con-ducted experiment is detailed. Subsequently, the findings obtained in the experiments are described and the impacts of each preprocessing method on the accuracy of the model in predicting the authors' gender are discussed. Finally, in the conclusion, we show briefly the important results of the current work study and highlight the directions of our future work.

[13] employed the NN model with GRU to determine the writer's by examining Facebook's and Twitter posts. The used NNP model input was prepro-cessed and divided into two layers: embedding layer and stylometric features extrac-tion phases. In fact, the embedding layer output was linked to a bidirectional GRU layer and then, to an activation layer. However, the stylometric features were nor-malized and attached directly to the same activation layer utilized after the GRU layer. The authors compared the obtained findings, which are inferior to the best result in PAN'AP (2017), to the best findings provided by Basile et al. in PAN' AP (2017).

In 2017, Estruch et al. enhanced an early fusion model, which was based on perform-ing fusion after the decision level single source classification. The authors achieved 91% GI accuracy on an English dataset in Singapore retrieved from Foursquare, Instagram and Twitter.

The approach developed by Sebastian Sierra and al. in 2018 was applied to assess the authors' gender employing multi-modal information (texts and images). The multi-modal representation was learned using GMUs. Indeed, accuracy rates equal to 0.74 and 0.81 were obtained in the multi-modal scenario for the test partition for English, Spanish and Arabic, respectively.

Moreover, the gold standard data was translated by Veenhoven et al in 2018 into the language of interest. Bi-LSTM and CNN architectures were also utilized to solve the GI problem by considering PAN-AP (2018) dataset. By considering the RNN, the highest obtained GI accuracy was equal to 79.3%, 80.4% and 74.9% for English, Spanish and Arabic languages, respectively.

The deep learning approach introduced by Yasuhide Miuraand et al. (2017) provided the best result when applied on the Portuguese language. The authors used the Re-current Neural Networks (RNN) for words and Convolutional Neural Networks (CNN) for characters. Therefore, they obtained two representations of various levels for a single message. The representations were, then, classified according to the writers' gender by employing attention mechanism, max-pooling layer and fully connected layer. More precisely, the word embeddings layer was first trained by the skip-gram. On the other hand, in the character embedding layer, weights were arbi-trarily initialized using the uniform distribution.

In 2013 [20], 2014 [21] and 2015 [22] PAN competition, the age and gender profiling was performed

by analyz-ing the English and Spanish datasets and applying the traditional supervised machine learning approaches, namely Logistic Regression, Random Forest, SVMs, etc. The objective of PAN competition organized in in 2016 [23] consists in validating the robustness of techniques from the cross-genre perspective. The obtained results showed that SVMs were the dominant paradigm. Then, F. Rangel et al in 2017 added, in 2017, two more languages (Arabic and Portuguese) to the dataset. Although SVMs were selected by several participants, deep neural networks (i.e., Windowed Recurrent Convolutional Neural Network as an extension of the Recur-rent Convolutional Neural Network) attained the state-of-the-art performance in terms of gender identification.

Among the PAN 2017 tasks, we cite the gender identification from Twitter texts. Concerning the Arabic language, the best model relied on representing the text as a vector including the combinations of character, word and POS n-grams with emojis, character flooding, and sentiment words. Besides, logistic regression was employed to train the classifier [34]. Approaches for predicting an AP can be broadly categorized into three types of methods as shown in table1.

Since the task of determining an author's profile can be seen as a classification task, we can benefit from pre-trained models. Indeed, pre-trained language models like BERT, GPT, ELMo, etc., capture extensive linguistic knowledge from large amounts of textual data. These models can be fine-tuned for specific authorship profiling tasks. on one side, by employing pre-trained models, transfer learning enables the transfer of general language and contextual knowledge to more specific author attribution tasks. This can enhance the models' ability to grasp subtle characteristics of an author's writing style. On the other side, Transfer learning is particularly useful when specific datasets for author attribution are limited. By fine-tuning pre-trained models on smaller datasets, better performance can be achieved with fewer specific data.

3 The proposed approach

The present work presents a fine tune model Approach. More specifically, we build an Arabic pretrained model based on the Ara-BERTv2-large model which is an improved version of BERT model [26] To design the proposed model, multi-lingual transformer models trained on large corpus, were fine tuned. The general architecture of this model is shown in Figure 1.

Table 1: Three types of methods for authorship identification task

Approaches	Example of features	Example of authors and Results
stylometry methods	The total number of characters The number of capitalized letters Character N-Grams The ratio of capital letters to total number of characters The ratio of white-space characters to total number of characters	Corney & al 2002 described an investigation of authorship gender attribution mining from e-mail text documents. They obtained 70.2 % precision rate for gender detection. Koppel and Peneebaker, analyzed a corpus of 71,000 blogs incorporating almost 300 million words. They obtained 43.8% and 86% for age and gender accuracy prediction, respectively (Schler & al 2006). In 2016 Bilan & al 2016, built a Cross-genre Author Profiling System (CAPS). Their system attained 74.36% accuracy for gender identification.
Content-based methods	Frequency of Function words The Number of contraction words Frequency of punctuations Stopwords The proportion ratio of singular to plural nouns and proper nouns and pronouns	Busger et al., 2016 obteined 0.5575 accuracy for gender identification in English data PAN 2016 competition. Dichiu & al 2016, applied SVM classifier and neural network on TF-IDF and verbosity features. Their results are almost similar to those provided in Bayot & al 2016. They got 61.5% gender accuracy and 41.03% age accuracy.
Deep learning models	- subword character embedding (word - n-gram embeddings - GloVe - FastText - ELMo (Embeddings from Language Models)	Nils Schaetti et al., 2017 used TF-IDF and a Deep-Learning model based on Convolutional Neural Networks. They obtained 0.66%, 0.73%, 0.81% and 0.57% of accuracy in the test partition for English, Spanish, Portuguese and Arabic respectively in PAN 2017 competition. Salvador et al., 2017 generated embeddings of the authors' text based on subword character n-grams. They got 0.7919 for gender identification in PAN 2017 competition[1]. Victor SANH et al, 2019 showed that it is possible to reach similar performances on many downstream-tasks using much smaller language models pre-trained with knowledge distillation, resulting in models that are lighter and faster at inference time. It is possible to reduce the size of a BERT model by 40%, while retaining 97% of its language understanding capabilities and being 60% faster.

The PERT tokenizer, trained by the WordPiece tokenization, was used to split the input text into a list of tokens. Such division reveals that means that a word can be broken down into several sub-words. The BERT vector assigned to a word is a function of the entire sentence. Therefore, a word can have different vectors according to the contexts in which they are used. There are different built-in tokenizers. The basic one is character tokenizer. However, the pretrained Arabic BERT utilizes a word-by-word tokenizer.

3.1 Corpus

The dataset, collected from Twitter, is part of the author's profiling task of PAN@CLEF 2018. For each tweet collection, Arabic texts are composed of tweets written by 2400 authors: 100 tweets per authors. Four varieties of the Arabic language were used in this corpus: Egypt, Gulf, Levantine and Maghrebi.

3.2 Pre-training model

Ara-BERTv2-large was trained to learn the distributed representation from the unlabeled texts by jointly conditioning on the left and right contexts of a certain token. The models were trained for 10 epochs with learning rate 1e-5 employing cross entropy loss criterion and Adam optimization algorithm. 32 samples were utilized in each mini batch, except when this did not fit in memory. During the training phase, a sequence of fixed length was used and padding or truncate was applied when necessary. The sequence length consists of 30, 100. Besides, all model parameters were fine-tuned during training, i.e., no layer was kept frozen. The model with the best validation set performance was evaluated on the test dataset.

[1] https://pan.webis.de/clef17/pan17-web/

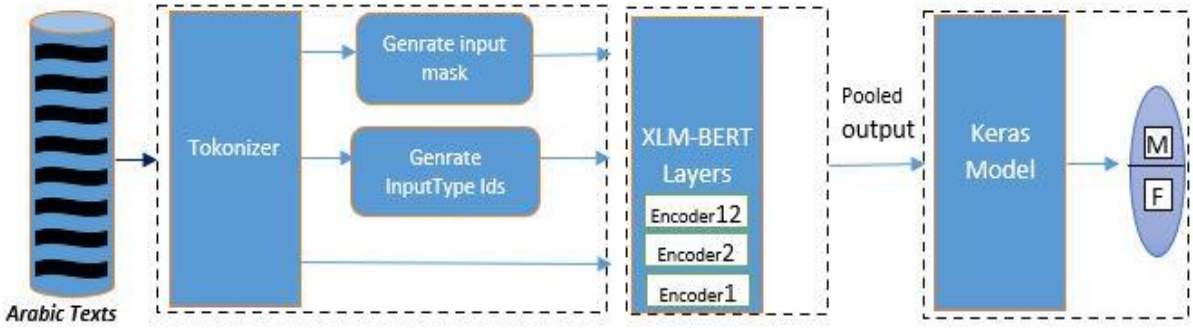


Figure 1 : The architecture of XLM-RoBERTa Approach.

3.3 Regularization of hyperparameters

This section details the fine-tuning of hyperparameters. The selected values and hyperparameters were determined through various tests, considering only those values that demonstrated optimal performance for the introduced model.

During the pretraining phase, we employed an Adam optimizer with a learning rate of 1e-8, a batch size of 64, a maximum sequence length of 512, and a masking probability of 15%. Additionally, a dropout rate of 0.1 was utilized to prevent model overfitting. All models were trained using a batch size of 16 and 5 epochs.

We leveraged Python within the Google Colab environment, a cloud service by Alphabet Inc., for implementing deep learning algorithms. Colab provides access to accelerated cloud tensor processing units (TPUs) developed by Google, each boast-ing up to 180 teraflops of computation power and high-bandwidth memory on a sin-gle board. In this study, Colab Pro was used, providing virtual machines (VMs) with doubled memory compared to standard Colab VMs.

To ensure uniform input sizes for Ara-BERTv2-large, we set a maximum sentence length of 128. Inputs were adjusted by padding and truncating until all sequences reached this length, employing the "pad_sequences" Python function with the "post" value for both padding and truncation, ensuring these operations occurred at the end of the sequences.

4 Experimentation

4.1 Preprocessing

We apply our preprocessing function before training/testing on any dataset. We used the library farasapy for segmentation, stemming, Part Of Speech tagging (POS tagging) and diacritization. Also use the unpreprocess function to reverse the preprocessing changes, by fixing the spacing around non alphabetical characters, and also de-segmenting if the model selected need pre-segmentation.

```
output_text = " و+ لن نبالغ إذا قل +نا : إن ' هاتف ' أو ' كمبيوتر ال+ مكتب '
في زمن +نا هذا ضروري"

arabert_prep.unpreprocess(output_text)
```

Figure 2: Example of the unpreprocess function.

4.2 Results

We approached this phase in 2 steps. The first serves as a pre-selection step. In this step, we choose the best-performing model for gender detection, in order to compare it in a second step with the XLM-RoBERTa model. Indeed, AraBERT comes in 6 variants. Each variant of AraBERT is pre-trained on a large cor-pus of Arabic text and can be fine-tuned on specific downstream tasks. In order to explore the most adopted variant for authorship detection task, we experiment with the six variants presented in the table below. We rely on the PAN 2018 corpus da-taset to test these models.

For hyperparameters setting, we used the same parameters in the different experiments for all the models: a peak learning rate to 1 × 10−5, maximum sequence length 128 tokens, batch size 64, 10 training epochs. Two objective functions are used during the language model pretraining step. The bidirectional nature ensures that the model can effectively make use of both past and future to-kens for this. The second objective is the next sentence prediction (NSP) task.

The results show that Ara-BERTv2-large achieved the highest accuracy, 79.7%, while Ara-BERTv2-base and Ara-BERTv1-base followed closely with an accuracy rate of 78.1%. Our experiments demonstrate that Ara-BERTv2-base, the largest model with 345 million parameters, is the most effective for the gender detection task in the Arabic language.

Upon exploring the obtained results, we notice that the difference in accuracy between the different tested models does not exceed 1%. These models, despite being trained using different pre-training objectives, such as masked language for Ara-BERTv2-large or masked language combined with next sentence prediction objectives for Ara-BERTv0.1-base, are highly effective for NLP tasks and outperform ML algorithms or N-Gram models.

Table 2: Performance of different models for gender identification task

Model	HuggingFace Model	Size (MB/Params)	Pre-Segmentation	DataSet (Sentences/Size/nWords)	Accuracy
AraBERT v0.2-Twitter-base	bert-base-arabertv02-twitter	543MB / 136M	No	Same as v02 + 60M Multi-Dialect Tweets	0,763
AraBERT v0.2-Twitter-large	bert-large-arabertv02-twitter	1.38G / 371M	No	Same as v02 + 60M Multi-Dialect Tweets	0,771
AraBERT v0.2-base	bert-base-arabertv02	543MB / 136M	No	200M / 77GB / 8.6B	0,767
AraBERT v0.2-large	bert-large-arabertv02	1.38G / 371M	No	200M / 77GB / 8.6B	0,765
AraBERT v2-base	bert-base-arabertv2	543MB / 136M	Yes	200M / 77GB / 8.6B	0,781
AraBERT v2-large	bert-large-arabertv2	1.38G / 371M	Yes	200M / 77GB / 8.6B	0,797
AraBERT v0.1-base	bert-base-arabertv01	543MB / 136M	No	77M / 23GB / 2.7B	0,771
AraBERT v1-base	bert-base-arabert	543MB / 136M	Yes	77M / 23GB / 2.7B	0,781

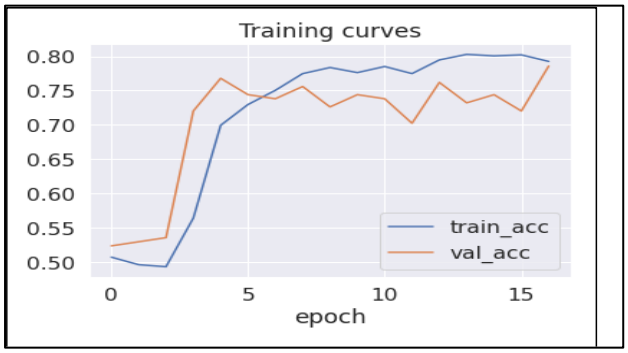


Figure 3: Accuracy of Ara-BERTv2-large model.

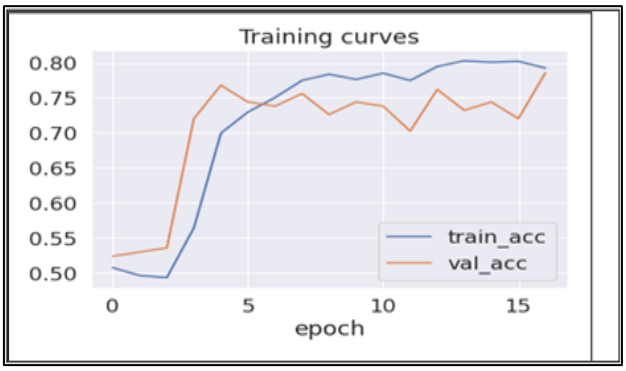


Figure 4: Fine tuning of Ara-BERTv2-large model.

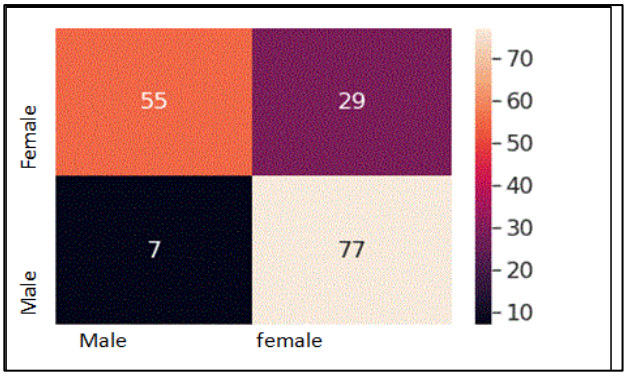


Figure 4 : Confusion matrix of the trained Ara-BERTv2-large

In terms of the adopted model's generalization, we will then compare it with XLM-RoBERTa, which is a multilingual model. This comparison can provide in-sights into the effectiveness of the Ara-BERTv2-large model for Arabic language tasks and whether a more general multilingual model is more suitable for the task at hand.

This comparison can also help researchers and practitioners to choose the most appropriate model for their specific use case and language. Indeed, XLM-RoBERTa is the multilingual variant of RoBERTa trained with a multilingual MLM on one hundred languages, with more than two terabytes of filtered Common Crawl da-ta.XLM-RoBERTa showed its superiority over BERT by its trainability on larger datasets, using larger vocabulary

as well as on longer sequences with larger batches in some cases. NSP task was removed and only MLM loss was used for pretraining. XLM-RoBERTa exhibited impressive performance in several multilingual NLP tasks and can perform comparably to monolingual language models (A. Conneau and al, 2019).

To fine-tune XLM-RoBERTa model for gender detection, we run fine-tuning experiments on single GPUs using Transformers software. Then, we fixed the peak learning rate to 1 × 10−5, maximum sequence length 128 tokens, batch size 64, 15 training epochs. We were also experimenting with other hyperparameters setting, but this one gave consistently the best results across all models and datasets. After each training epoch, we evaluated the model on the development dataset (if not pre-sent in the dataset, we randomly held out 10% of the training samples as a develop-ment data) and at the end, we used the best model for evaluation on the test dataset.

Model	Accuracy	Recall	F1-Score
XLM-RoBERTa	0,768	0,75	0,84
Ara-BERTv2-large	0,797	0,76	0,82

Table 3 : Performance of Ara-BERTv2-large and XLM-RoBERTa

4.3 Discussion

Our fine-tuned Ara-BERTv2-large model has achieved higher accuracy (79.7%) on the test data set compared to XLM-RoBERTa (76.8%), as shown in Figure 3 and Table 2. This confirms the suggestion that a model specifically designed for a particular language can perform better than a more general multilingual model in tasks related to that language. Indeed, even when compared to the other models trained specifically on the Arabic language, as shown in the table1, we can see that XLM-RoBERTa remains less competitive and only outperforms 3 out of 8 models.

The reason for this is found in the scale of the Ara-BERTv2-large model, encompassing an increased number of parameters, layers, and a larger hidden size. Conse-quently, this heightened capacity enables it to capture finer and more intricate patterns, particularly in the context of social media platforms and Arabic datasets, which often include dialectal variations.

As a baseline, we used three results obtained in PAN@CLEF as a baseline method to assess our technique and show its efficiency. Those obtained by applying deep learn-ing method in PAN@CLEF2017: the result of [21] based on using BRNN Gated Recurrent Unit and [24] relying on CNN architecture as well as the best results of gender identification obtained in PAN@CLEF2017 [27].

Kodiyanand et al. in 2017, utilizing GRU, achieving a 71.50% result, and another by Miura et al. in 2017, employing CNN and obtaining a 76.44%.

The results from the conducted experiments demonstrated that the developed Ara-BERTv2-large achieved state-of-the-art performance on Arabic datasets. This vali-dates our initial hypothesis and underscores the effectiveness of pre-trained models for the Arabic language, particularly when dealing with dialectal variations. It emphasizes how these models, with their sophisticated architecture and extensive training on Arabic text, can comprehend the intricate nuances and complexities within the language. This success reinforces the value and potential of deep learning techniques specifically designed for Arabic NLP tasks, showcasing their ability to achieve state-of-the-art results in handling various linguistic challenges present in Arabic datasets.

5 Conclusions

The novelty of this work consists in conducting the Arabic author profiling experi-ments by focusing on the gender of the social networks' users. In this task, the trans-fer learning methods were utilized, for the first time, on the Arabic language.

Three deep learning models were applied with words embeddings for the prediction of Twitter Arabic authors' gender.

The experimental results revealed that the suggested model, called XLM-RoBERTa and used as a contextual embedding technique outperforms the models Ara-BERTv2-base. To sum up, deep learning techniques are not very efficient in detecting the authors' profile and, more precisely, his/her age and gender. As future work, we will explore and enhance the performance of deep learning approaches in author's profiling by augmenting the size of the training set, using different tuning parameters, and employing various types of word embeddings).

References

[1] A. Conneau, K. Khandelwal, N. Goyal, V. Chaudhary, G. Wenzek, F. Guzmán, E. Grave, M. Ott, L. Zettlemoyer, and V. Stoyanov, "Unsupervised cross-lingual representation learning at scale," arXiv preprint arXiv:1911.02116, 2020. https://doi.org/10.18653/v1/2020.acl-main.747

[2] Ai, M.: BERT for Russian news clustering. In: Proceedings of the International Conference "Dialogue 2021", p. 6. Moscow, Russia.2021. https://doi.org/10.28995/2075-7182-2021-20-385-390

[3] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, "Improving language understanding by generative pre-training," URL https://s3- us-west-2. amazonaws.com/openai-assets/research covers/languageunsupervised/language_ understanding_paper. pdf, 2018.

[4] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. u. Kaiser, and I. Polosukhin, "Attention is all you need," in Advances in Neural Information Processing Systems 30. Curran Associates, Inc., 2017, pp. 5998–6008. [Online]. Available: http://papers.nips.cc/paper/7181-attention-is-all-you-need.pdf

[5] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. Language models are unsupervised multitask learners. 2019.

[6] Alvarez-Carmona, M. A., Lopez-Monroy, A. P., Montes- ´ y Gomez, M., Villase ´ nor-Pineda, L., and Meza, I. ˜ (2016). Evaluating topic-based representations for author profiling in social media. In Montes y Gomez, M.

[7] Antoun, Wissam, Fady Baly, and Hazem Hajj. "Arabert: Transformer-based model for arabic language understanding." arXiv preprint arXiv:2003.00104 (2020).

[8] Bernard, G.: Resources to compute TF-IDF weightings on press articles and tweets (2022). https://doi.org/10.5281/zenodo.6610406.2022

[9] Bassem, B., & Zrigui, M. (2020). Gender identification: a comparative study of deep learning architectures. In Intelligent Systems Design and Applications: 18th International Conference on Intelligent Systems Design and Applications (ISDA 2018) held in Vellore, India, December 6-8, 2018, Volume 2 (pp. 792-800). Springer International Publishing. https://doi.org/10.1007/978-3-030-16660-1_77

[10] Butt, S., Ashraf, N., Sidorov, G., & Gelbukh, A. F. (2021, September). Sexism Identification using BERT and Data Augmentation-EXIST2021. In IberLEF@ SEPLN (pp. 381-389).2021.

[11] Bsir, B., & Zrigui, M. (2018). Enhancing deep learning gender identification with gated recurrent units architecture in social text. Computación y Sistemas, 22(3), 757-766. https://doi.org/10.13053/cys-22-3-3036

[12] Bsir, B., & Zrigui, M. (2018). Bidirectional LSTM for author gender identification. In Computational Collective Intelligence: 10th International Conference, ICCCI 2018, Bristol, UK, September 5-7, 2018, Proceedings, Part I 10 (pp. 393-402). Springer International Publishing.

[13] Bsir, B., & Zrigui, M. (2019). Document model with attention bidirectional recurrent network for gender identification. In Advances in Computational Intelligence: 15th International Work-Conference on Artificial Neural Networks, IWANN 2019, Gran Canaria, Spain, June 12-14, 2019, Proceedings, Part I 15 (pp. 621-631). Springer International Publishing. https://doi.org/10.1007/978-3-030-20521-8_51

[14] Catelli, R., Pelosi, S., & Esposito, M. (2022). Lexicon-based vs. Bert-based sentiment analysis: A comparative study in Italian. Electronics, 11(3), 374.2022.https://doi.org/10.3390/electronics11030374

[15] C. Sun, X. Qiu, Y. Xu, and X. Huang, "How to fine-tune BERT for text classification?" in China National Conference on Chinese Computational Linguistics. Springer, 2019, pp. 194–206.

[16] González-Gallardo, C. E., Montes, A., Sierra, G., Núnez-Juárez, J. A., Salinas-López, A. J., & Ek, J. (2015, September). Tweets Classification using Corpus Dependent Tags, Character and POS N-grams. In CLEF (working notes).2015

[17] Chi Sun, Xipeng Qiu, Yige Xu, and Xuanjing Huang. How to fine-tune BERT for text classification? In China National Conference on Chinese Computational Linguistics, pages 194–206. Springer, 2019.

[18] DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter / Victor Sanh, Lysandre Debut, Julien Chaumond, Thomas Wolf // arXiv preprint arXiv:1910.01108. — 2019.

[19] Estruch, C. P., Paredes, R. Rosso, P., Learning Multimodal Gender Profile using Neural Networks. Rec Adv Nat Language Process 2017; Varna: Bulgaria, pp: 577-582. Available from: http://users.dsic.upv.es/~prosso/resources/PerezEtAl_RANLP17.pdf. https://doi.org/10.26615/978-954-452-049-6_075

[20] F. Rangel, P. Rosso, M. Koppel, E. Stamatatos, and G. Inches. "Overview of the author profiling task at PAN 2013". CLEF 2013 Evaluation Labs and Workshop – Working Notes Papers, 2013.

[21] F. Rangel, P. Rosso, I. Chugur, M. Potthast, M. Trenkmann, B. Stein, B. Verhoeven, and W. Daelemans. "Overview of the 2nd author profiling task at PAN 2014". CLEF 2014 Evaluation Labs and Workshop – Working Notes Papers, 2014.

[22] F. Rangel, F. Celli, P. Rosso, M. Potthast, B. Stein, and W. Daelemans. "Overview of the 3rd author profiling task at PAN 2015". CLEF 2015 Evaluation Labs and Workshop – Working Notes Papers, 2015.

[23] F. Rangel, M. Francisco, P. Rosso, B. Verhoeven, W. Daelemans, M. Potthast, Martin, and B. Stein. "Overview of the 4th author profiling task at PAN 2016: Cross-Genre Evaluations". Working Notes Papers of the CLEF 2016 Evaluation Labs, 2016.

[24] F. Rangel, M. Francisco, P. Rosso, M. Potthast, and B. Stein. "Overview of the 5th author profiling task at PAN 2017: gender and language variety identification in Twitter". Working Notes Papers of the CLEF 2017 Evaluation Labs, 2017.

[25] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics. 2019.

[26] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," arXiv preprint arXiv:1810.04805, 2018.

[27] J. Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pages 4171–4186, Minneapolis, Minnesota, 2019. Association for Computational Linguistics.

[28] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," NAACL HLT 2019 - 2019 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. - Proc. Conf., vol. 1, no. Mlm, pp. 4171–4186, 2019.

[29] J. Howard and S. Ruder, "Universal language model fine-tuning for text classification," ACL 2018 - 56th Annu. Meet. Assoc. Comput. Linguist. Proc. Conf. (Long Pap., vol. 1, pp. 328–339, 2018, doi: 10.18653/v1/p18-1031. https://doi.org/10.18653/v1/p18-1031

[30] Haffar N., Ayadi R., Hkiri E., Zrigui M. (2021) Temporal Ordering of Events via Deep Neural Networks. In: Lladós J., Lopresti D., Uchida S. (eds) Document Analysis and Recognition – ICDAR 2021. ICDAR 2021. Lecture Notes in Computer Science, vol 12822. Springer, Cham. https://doi.org/10.1007/978-3-030-86331-9_49.

[31] Abdellaoui, H., & Zrigui, M. (2018). Using tweets and emojis to build tead: an Arabic dataset for sentiment analysis. Computación y Sistemas, 22(3), 777-786. https://doi.org/10.13053/cys-22-3-3031

[32] Henry Tsai, Jason Riesa, Melvin Johnson, Naveen Arivazhagan, Xin Li, and Amelia Archer. 2019. Small and Practical BERT Models for Sequence Labeling. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), pages 3632–3636, Hong Kong, China. Association for Computational Linguistics.2019.

[33] Kim, J., Aum, J., Lee, S., Jang, Y., Park, E., Choi, D.: FibVID: comprehensive fake news diffusion dataset during the COVID-19 period. Telemat. Inform. 64, 101688 (2021). https://doi.org/10.1016/j.tele.2021.101688.2022.

[34] Matej Martinc, Iza Skrjanec, Katja Zupan, and Senja Pollak. Pan 2017: Author profiling-gender ˇ and language variety prediction. Cappellato et al.[13], 2017.

[35] Martinc, M., Škrjanec, I., Zupan, K., & Pollak, S. Pan 2017: Author profilinggender and language variety prediction. CLEF (Working Notes) 2017. CEUR Workshop Pro ceedings 1866, CEUR-WS.org (2017).

[36] R. Sennrich, B. Haddow, and A. Birch, "Neural machine translation of rare words with subword units," arXiv preprint arXiv:1508.07909, 2015.

[37] Suman, C., Kumar, P., Saha, S., & Bhattacharyya, P. (2019, December). Gender Age and Dialect Recognition using Tweets in a Deep Learning Framework-Notebook for FIRE 2019. In FIRE (Working Notes) (pp. 160-166).

[38] S. Garg, T. Vu, and A. Moschitti, "TANDA: Transfer and adapt pre-trained transformer models for answer sentence selection," arXiv preprint arXiv:1911.04118, 2019.

[39] Schlicht, I. B., & de Paula, A. F. M. (2021). Unified and multilingual author profiling for detecting haters. arXiv preprint arXiv:2109.09233.

[40] Velankar, A., Patil, H., & Joshi, R. (2022, November). Mono vs multilingual bert for hate speech detection and text classification: A case study in marathi. In Artificial Neural Networks in Pattern Recognition: 10th IAPR TC3 Workshop, ANNPR 2022, Dubai, United Arab Emirates, November 24–26, 2022, Proceedings (pp. 121-128). Cham: Springer International Publishing.2002. https://doi.org/10.1007/978-3-031-20650-4_10

[41] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov, "RoBERTa: A robustly optimized BERT pretraining approach," CoRR, vol. abs/1907.11692, 2019.

[42] Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R., Le, Q.V.: XLNet: generalized autoregressive pretraining for language understanding. arXiv:1906.08237 [cs] (2020)

[43] Yasuhide Miura, Tomoki Taniguchi, Motoki Taniguchi, and Tomoko Ohkuma. Author profiling with word+ character neural attention network. Cappellato et al.[13].2017.

[44] Z. Zhang, Y. Wu, H. Zhao, Z. Li, S. Zhang, X. Zhou, and X. Zhou, "Semantics-aware BERT for language understanding," arXiv preprint arXiv:1909.02209, 2019. https://doi.org/10.1609/aaai.v34i05.6510

[45] Zrigui, M., Ayadi, R., Mars, M., & Maraoui, M. (2012). Arabic text classification framework based on latent dirichlet allocation. Journal of computing and information technology, 20(2), 125-140. https://doi.org/10.2498/cit.1001770

Liver Disease Classification - An XAI Approach to Biomedical AI

Ebenezer Agbozo, Daniel Musafiri Balungu
Ural Federal University, Russian Federation
E-mail: eagbozo@urfu.ru, danielbal03.db@gmail.com

Keywords: deep learning, biomedical science, explainable ai, data-driven decision making, shapley values, prescriptive analytics

Received: Januar 12, 2023

Explosive amounts of biological and physiological data, including medical images, electroencephalograms, genomic information, and protein sequences, have been made available to us thanks to advances in biological and medical technologies. Understanding human health and disease is made easier by using this data for learning. Deep learning-based algorithms, which were developed from artificial neural networks, have significant potential for identifying patterns and extracting features from large amounts of complex data. However, these recent advancements involve blackbox models: algorithms that do not provide human-understandable explanations in support of their decisions. This limitation hampers the fairness, accountability and transparency of these models; the field of XAI tries to solve this problem providing human-understandable explanations for black-box models. This paper focuses on the requirement for XAI to be able to explain in detail the decisions made by an AI in a biomedical setting to the expert in the domain, e.g., the physician in the case of AI-based clinical decisions related to diagnosis, treatment, or prognosis of a disease. In this paper, we made use of the Indian Patient Liver Dataset (IPLD) collected from Andhra Pradesh region. The deep learning model with a 0.81 accuracy score (0.82 for the hyperparameter- tuned model) is built on Keras-Tensorflow and due to the imbalance in the target values, we integrated GANs as a means of oversampling the dataset. This study integrated the XAI concept of Shapley Values to shed light on the predictive results obtained by the liver disease detection model.

Povzetek: Študija obravnava klasifikacijo jetrnih bolezni z uporabo razložljive umetne inteligence (XAI), ki omogoča razumevanje odločitev modelov globokega učenja z integracijo Shapley vrednosti za razlago prediktivnih rezultatov.

1 Introduction

For most of its history, medicine was practiced on artistic principles rather than according to modern definitions of science. In the past two centuries, the practice of medicine has been more closely aligned with scientific method principles, particularly in regards to comprehending the molecular causes of disease. Advances in anatomy, physiology, genetics, immunology, and other scientific sub-disciplines have helped to define and broaden the scope of contemporary medicine from the beginning of a research tradition in the modern era.

Medical science benefits from biomedical science because it enables doctors to comprehend the crucial steps involved in infectious diseases brought on by bacteria, viruses, protozoa, and other microorganisms, the impact of body physiology and biochemistry on maintaining health, and the immune system's tolerance or rejection of transplanted tissues. It provides a framework for developing novel methods of maintaining health as well as for testing someone's blood, urine, or tissue for the presence of disease.

The goal of biomedical science is to identify diseases using various techniques. Early diagnosis can save a person's life in many conditions, including cancer. Over the last decade, technologies have been driving the healthcare industry through various innovations in how we find, prevent, and cure diseases. This shouldn't have happened without the massive growth of AI-driven technologies and digitization of healthcare workflows, as a response to more savage global conditions, as well as the rising demand on accessible and quality medical service. Those medical innovations have pushed the envelope of possibility and increased the well-being of millions. This year is no different. Doctors and researchers on the forefront of medicine and technology are enhancing patient care in a number of ways with technology spearheading the initiatives. Here are some medical innovations: bringing diseases to an end with CRISPR Technology, UAV technology for medical supply distribution, IoT for healthcare, and remote patient monitoring.

Recent ML developments promise to significantly enhance the accuracy of diagnosis and the screening for retinal disorders. Systems created using these techniques have shown expert-level accuracy in the detection of a variety of eye disorders, including glaucoma, age-related macular degeneration (AMD), diabetic retinopathy, and other anomalies related to retinal diseases[1]–[3]. But it's

not entirely clear how these models affect clinical settings. Many difficulties have been encountered in the past when ML algorithms have been used in computer-assisted diagnosis settings, including over reliance (repeating model errors) and under reliance (ignoring accurate algorithm predictions) [4], [5]. If the computer assisted diagnosis system can explain its black box AI predictions, some of these problems might be avoided [6]). Explainable AI (XAI) aims at decoding the decision of AI (Deep learning/Machine learning) black box to the extent of human-interpretable level. As such we pose the following research questions: (RQ1). How has explainable AI been applied in the sphere of biomedical science? (RQ2) How can deep learning algorithms to classify Liver Disease from a set of patients' records generate further interpretable justification for its prediction results? (RQ3) Can the justification of explainable AI results for predicting the presence or absence of liver disease be visually presented? Our paper aims to contribute to the ongoing research on explainability in line with the desire for understanding of AI predictions in industry.

The subsequent sections of the article are structured as follows, related works section where we delve into AI's diffusion in biomedical science; followed by the next section which explores explainable AI (XAI); followed by the data and method section where we train neural network models and apply XAI algorithms on the results (revealed in the results section); finally, we conclude on the study and summarize our findings.

2 Related works

Many important problems in biomedical decision making can be expressed as binary classification problems. For example, one may wish to identify infants infected with hepatitis C virus from a sample of infants born to infected mothers [7], screen for prostate cancer using prostate-specific antigen [8], or predict which breast cancer patients will respond to treatment based on genetic characteristics [9].

In order to address the methods, techniques and algorithms used for making decisions in biomedicine, let us take into account the following aspects of medical data processing: missing data imputation, diagnostics (classification and prediction), clustering and personalizing the treatment. A previous study predicted missing data, analyzed the nature of data gaps, and filled these gaps using decision tree-based computation techniques and regression approach [10]. Similar outcomes for associative rules mining in medical data were found by another study [11]. In addition, a study adopted Bayesian networks, ANN, and k-means algorithms to predict cardiac disease [12]. However, Bayesian networks are too sluggish for both online diagnostics and processing the vast amounts of data. Y. Tang created a method for paralleling Bayesian networks in response to this [13]. For multi-parameter, massive, and dynamic medical data flows, Bayesian networks should still be used in conjunction with other machine learning techniques, even in the presence of parallelism. Fuzzy logic-based artificial neural network technology is actively employed to analyze a variety of medical data. Thus, a system of quick medical diagnosis based on auto-associative neuro-fuzzy memory was developed in the works [14]–[17]. To increase the accuracy of the classification problem's results, however, is still of uttermost priority. The use of existing techniques and computational intelligence tools to address such issues is further constrained by the issue of imbalanced input data as well as the tiny samples of data manually collected by medical professionals [18].

The cluster analysis is frequently used to identify outliers. In the medical field, outliers refer to variations from the ideal patient circumstances based on the regional protocol and unique traits. Partitioning techniques are among the simplest clustering algorithms. The K-means algorithm creates k clusters that are spread far apart from one another. The assumption (hypothesis) regarding the number of clusters and the variety of the instances in various clusters is the fundamental sort of problem that the k-means method solves. The results of prior research and theoretical considerations may be used to guide the selection of the k number [19].

The decisions made in the healthcare industry generally involve a variety of criteria, many options, flawed data, and varying stakeholder preferences. However, the systemic assessment and the processing of pertinent information, a process that involves the flow of data between numerous components, frequently present problems for the decision-makers. Because of this, decision-makers' reliance on informal judgments or processes can result in poor choices in these situations [20]. The widespread availability of data has sparked a growing interest in methods for extracting useful facts and information from data and decision-making that is data-driven. As a result, the data science field seeks to learn from data and frequently impact decisions to make them increasingly dependable. The Decision Support System (DSS) is a flexible framework used in the artificial intelligence (AI) industry for managing the formalization of human problem-solving and contemplation techniques.

DSS can support the problem-solving process based on two principles, including knowledge and the capacity for reasoning. Overall, the consideration of AI is based on a variety of justifications, including an input and operational point of view, an output and behavioral viewpoint, an evaluation of its relevance, i.e., its ideal performance, and a comparison of its consistency and quality with human performance [21]. In order to represent the framework under consideration, distinct AI methodologies lead to different approaches, for instance, for the management of complex problems, such as the significantly complicated decision-making in the healthcare industry. Another important aspect that was emphasized is the idea of distributing processing power and intelligence among network systems. According to Urdea et al., combining patient statistical data with test results data generated at the point of care can result in a complete dataset that can be effectively used to concentrate fine-grained observation data about a variety of diseases using data analysis at both the individual and

population levels [22]. According to research, demographic databases combined with test results might be used to obtain a single dimension, which is equivalent to the population's overall health [23]. The large healthcare data may also be retrieved and applied in prediction-based tasks, which is of extreme significance to decision-making in healthcare. This is done by integrating the aforementioned datasets with mobility patterns, location data, and trends in disease pervasiveness.

Table 1. ML and DL applications

Detection	**Prediction**	**Generation**
Image interpretation	Classification	Design
Text & Speech	Analysis	Visual Art
Abuse and Fraud	Recommendations	Text
Human behavior & Identity	Collective behavior	Music

In recent times, deep learning (DL) has been one of the fast-growing ML fields. It attempts to model abstraction from large-scale data by employing multi-layered deep neural networks (DNNs), thus making sense of data such as images, sounds, and texts. Deep Learning helps provide intelligent answers to complex issues. The structure and operation of the human brain serve as its foundation. Artificial neural networks are used by deep learning to analyze data and make predictions. It has applications in practically every business industry.

Deep Learning is used in a large number of applications that are used on a daily basis, such as the Google translator; in virtual assistants such as Yandex Alice, Apple's Siri, Microsoft's Cortana and Google Assistant, which use Deep Learning algorithms for voice recognition; classification of emails and even for security systems that make use of facial recognition. Another of the areas where Deep Learning is applied, is in something as complex as autonomous cars, which every day are closer to becoming a reality.

In the case of factories, for example, it can be used to recognize new parts that have not been previously introduced into the system, since the Deep Learning algorithm has studied other previous photos in which it has been indicated what it is a piece and when a new part has been introduced into the system, it has been recognized as such without having to indicate it.

Another very important application in factories is the intelligent recognition of defects. Once the system has been trained with different defects (shape, size, geometry, etc.), it is possible that the system could recognize new defects because it has learned what it is. It is a very interesting application because of the variability of defects it is common not to be able to categorize all at first.

A flood of biological and medical data, including information about medical imaging, biological sequences, and protein structures, has been amassed in recent decades as a result of advancements in high-throughput technology. This section reviews some effective deep learning applications in the biomedical domains.

- ***Medical image classification and segmentation***

Machine learning has long been a potent tool in the diagnosis or assessment of diseases using medical images. Traditionally, classification (identification of diseases or abnormalities) and segmentation of regions of interest (tissues and organs) in various medical applications rely on manually created discriminative characteristics. Participation of skilled physicians is required in this. The widespread use of machine learning in the medical image domain has been hampered by the complexity and ambiguity of medical images, limited expertise in medical image interpretation, and the demand of vast amounts of annotated data. A number of computer vision tasks, including object detection, localization, and segmentation in natural images, have been successfully completed using deep learning techniques.

For the qualitative and quantitative assessment of medical imaging, the segmentation of tissues and organs is essential. To accomplish precise brain tumor segmentation, Pereira et al. used data augmentation, tiny convolutional kernels, and a pre-processing stage [24]. In 2013 and 2015, their CNN-based segmentation technique took first place and second place in the Brain Tumor Segmentation (BRATS) Challenge. Magnetic resonance images (MRI) and a two-phase training process was used by a study to. demonstrated brain tumor segmentation approach (fully automatic) which took the 2nd place in BRATS 2013 [25]. By using the INbreast and Digital Database for Screening Mammography (DDSM) datasets, their methodology outperformed SOTA techniques at the time in terms of model accuracy and effectiveness [26], [27]. Additionally, deep learning architecture in medical research have been shown to segment the heart's left ventricle from MR data [28], the pancreas from computed tomography [29], the prostate from MRI [30], the tibial cartilage from magnetic resonance imaging [31], and the hippocampus from MR brain images [32], [33]. Through semantic segmentation (the process of classifying or labeling each pixel of an image in order to distinguish various tissues or organs [34], [35]) based on a deep neural network architecture where organs, skeletal muscles, as well as fat in CT scans are vividly distinguished [36]. Also, accurate segmentation findings were achieved by semantic segmentation of MRIs [37], [38].

- ***Genomic sequencing and gene expression analysis***

Genomic sequencing, which establishes the precise arrangement of nucleotides within a DNA molecule, is increasingly essential for many applications, including fundamental biological study, medical diagnostics, biotechnology, forensic biology, virology, and biological systematics. Deep learning application in genomic sequencing is divided into two fields: learning the functional activity of DNA sequencing and DNA methylation.

Three processes make up the biological process of gene expression: transcription, RNA processing, and translation. An RNA molecule called precursor messenger RNA (pre-mRNA), which is a copy of the DNA in the transcribed gene, is produced as a result of transcription. The pre-mRNA is then altered by RNA processing to create a new RNA molecule termed messenger RNA

(mRNA). Reading the three-letter (codes) in the mRNA sequence during translation results in the creation of a protein molecule (an amino acid chain) [39]. The alternative splicing field and the prediction of gene expression are the two directions in which deep learning techniques are utilized in the field of gene expression.

3 Explainable AI (XAI)

The goal of XAI is to improve the human understanding of the output of AI systems. The term was initially used in previous studies to indicate how well their system could account for the actions of AI-controlled characters in simulation games [40]. Since researchers began looking at explanation for expert systems in the middle of the 1970s, the explainability problem has been a challenge. The unstoppable spread of AI/ML across all spheres and its critical influence in decision-making processes, while not being able to deliver comprehensive details regarding the chain of reasoning leading to some decisions, predictions, recommendations or actions made by it, are directly responsible for the resurgence of this research topic. Therefore, new AI strategies that can make decisions comprehensible and explicable are required due to societal, ethical, and legal demands.

Demystification of the black-box models is at the heart of XAI, which also implies responsible AI because it can aid in the creation of transparent models. This should take place without affecting the accuracy of the AI models; as a result, accuracy and interpretability must frequently be traded off in AI in general and in ML in particular. Accuracy is intimately related to the quality and amount of the training data, which naturally draws a connection to the data science discipline.

Explainability plays a fundamental role in the justification of AI-based predictions or classifications. It aids in prediction verification, model modification, and for unveiling insights into the problem at hand, thereby leading to more dependable AI systems. The need for explaining AI systems is purported to stem from four (4) reasons. In spite of the fact that the four (4) reasons may appear to overlap, it is believed to capture the core motivations of model explainability. These include: Explaining to Justify (the reason for the specific outcome(s)); Explaining to Control (gain insight into vulnerabilities or defects - debugging); Explaining to Improve (a comprehensible model makes improvement possible by focusing on desired constructs); and Explaining to Discover (revealing the unforeseen) [41].

As purported by research the goals of XAI have been summarized into the concepts evident in figure 1. Literature clearly distinguishes between models that can be understood using external XAI approaches and those that are interpretable by design. This distinction between transparent models and post-hoc explainability is more widely understood than the distinction between interpretable models and model interpretability methodologies. This same dichotomy can be seen in the paper discussed in a previous study, where the authors contrast the approaches used to address the transparent box design problem with those used to address the black-box problem's explanation [43].

Figure 1: XAI goals [42]

By using a variety of techniques to improve their interpretability, such as text explanations, visual explanations, local explanations, explanations by example, explanations by simplification, and explanations based on feature relevance, post-hoc explainability aims to target models that are not easily interpretable by design techniques.

Here are some XAI methods that have been applied in some real-world tasks, such as autonomous driving and healthcare. These methods develop explainable algorithms to interpret results and improve their decisions or actions according to the task. Recent self-driving systems have adopted interpretation techniques to improve the actions of the autonomous driving system and reduce the risk of a crash. This is also important to increase the trust between humans and AI machines.

- **Explainable decisions for autonomous cars**

In [44], the authors suggested a novel, comprehensible self-driving system that was motivated by human drivers' responses and choices. The suggested solution uses a CNN to extract features from the input image, and a global module to create the scene context and offer information on where the items are in relation to each other. To create the actions and explanations, a local branch is used to pick the scene's most crucial elements and link them to the scene's context. Finally, explanations in visual form are created for the input image. Similar to [45], the authors suggested an architecture for autonomous driving that is aided and trained by humans.

In order to separate the objects from the incoming video stream, the system uses a visual encoder. A vehicle controller is trained to speak commands, such as stopping the automobile when the traffic light turns red, verbally. The controller also creates attention maps to emphasize the key areas and justify their choices. An observation

generator is used to aggregate video frames and provide general observations, which must be taken into account while driving, further enhancing the system's robustness. The vehicle controller also receives these observations to help it make better decisions.

- **Explainable medical systems**

AI-based systems have also been used in medical settings in the fields such as drug development and medical imaging, thus produced notable breakthroughs. To help medical professionals by offering helpful explanations so that any expert may grasp a system's predictions, researchers have recently concentrated on explainable medical systems. The authors of concentrated on coronavirus detection from x-ray images [46]. To extract information from the images and determine whether a patient has pneumonia or coronavirus, researchers suggested using a deep convolutional neural network. The infected areas from the x-ray are then highlighted and visual explanations are provided through Grad-CAM [44].

4 Data and method

In this paper, we made use of the Indian Patient Liver Dataset (IPLD) collected from Andhra Pradesh region, a widely known dataset within the ML research community, which comprises observations with 416 liver patient records and 167 non liver patient records [47]. As highlighted in figure 2, the dataset was pre-processed, dropping four (4) unavailable observations, as well as normalization (Min-Max Scalar). The deep learning model is built on Keras-Tensorflow and due to the imbalance in the target values, we integrated Tabular GANs (Generative Adversarial Networks) as a means of oversampling the dataset due to the small sample size.

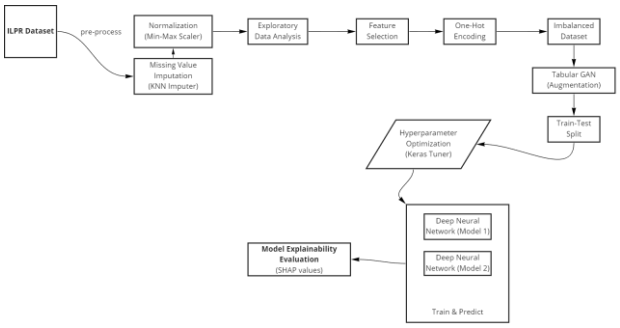


Figure 2: Research workflow

For model interpretability purposes, our study incorporated the SHAP package [48] which was developed as an offspring of research from the University of Washington, and Microsoft Research. Model interpretability is extremely important in AI and it produces end-user trust, delivers insight as to how a model may be improved, as well as supports understanding of the process being modelled [48]. Our study integrated the XAI concept of Shapley Values to shed light on the predictive results obtained by the liver disease detection model. The concept of Shapley Values hails from cooperative game theory and was introduced in 1953 [49]. It is defined as the sum (weighted) of the agents' marginal contributions to coalitions [50]. The three theoretical properties of Shapley values are local accuracy, missingness, and consistency [48]. Marginal contribution is a central component to understanding Shapley values and is defined as the amount by which the evaluation of a submodel increases when a given feature is introduced to the submodel [49]. To formally represent Shapley values as marginal contribution, the formula below is indicated:

$$\phi_i(N,v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N|-|S|-1)!}{N!} [v(S \cup \{i\}) - v(S)]$$

Where ϕi denotes the average marginal contribution for a player i; N denotes the number of players; v is the game; and S denotes the sets of different coalitions [51]. A Shapley value is representative of a unique quantity that is capable of constructing an explanatory model that locally linearly approximates the original model, given a specific input [52]. From an ML perspective, some studies have adopted Shapley values as a feature selection tool due to its appealing nature with regards to highlighting which features contribute to an obtained output, but in their study, Fryer et al. noted that, in general, the axioms (Efficiency, Null Player, Symmetry, Additivity, and Balanced Contributions) do not provide any guarantee that Shapley values are suitable for feature selection, and may most likely in some cases imply the opposite [49]. They also highlighted that the favorability of Shapley value axioms depends non-trivially on how the Shapley value is appropriated within a particular XAI application.

Shapley values, when applied within a human-centric ML perspective, are capable of shifting the perspective and obtaining insights into client behaviour as well as desires, thereby creating relevant persona profiles which leads towards the trajectory of prescriptive analytics [53]. Shapley values have been applied by previous studies to interpret log anomaly detection systems; to understand client creditworthiness prediction; understand the propensity of clients to buy an insurance policy as well as the risk of churn with respect to an existing customer [52]–[54]. The next section discusses the results of our study.

5 Results

As a means of explaining model predictions, our study utilised SHapley Additive exPlanations (SHAP) and visualises interpretations as SHAP summary plots and SHAP dependence plots. SHAP approximation techniques that exist include Kernel, Deep, and Tree SHAP which are used for kernel-based, deep neural network based, and tree-based models respectively. In order to establish the relationship between features and target variable, initial results from the exploratory data analysis are highlighted in figure 3, via a correlation matrix plot. The following strong positive correlations were established: (1) "Direct_Bilirubin" and "Total_Bilirubin"; (2) "Aspartate_Aminotransferase" and "Alamine_Aminotransferase"; (3) "Albumin" and

"Total_Proteins"; (4) "Albumin and Globulin Ratio" and "Albumin". A negative correlation between the target variable and three features (1) "Total Proteins", (2) "Albumin", and (3) "Albumin and Globulin Ratio", while having a weak positive correlation with the other 8 features.

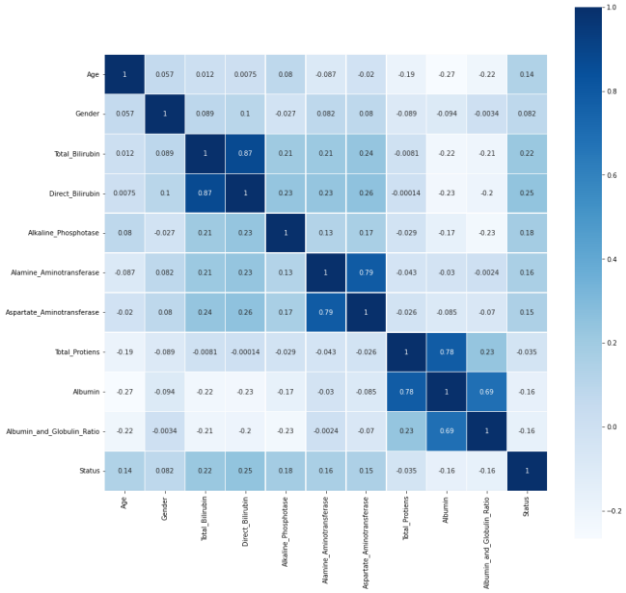


Figure 3: Correlation plot of features and output

Figures 4 and 5 highlight the deep neural network architectures built on Keras-Tensorflow. In order to obtain the best possible training hyperparameters (learning rate, dropout rate, bias vector, neurons, and activation functions), we utilised the RandomSearch feature of the Keras Tuner package (5 and 10 max trials respectively for Models 1 and 2). Figure 3 illustrates the DNN architecture of model 1 and figure 4 highlights that of the hyperparameter tuned model (model 2). The parameter spaces for the hyperparameter tuning process were as follows:

a. Number of Layers – 4

b. Number of Units (Neurons) – value domain = [16 – 512]; step = 16

d. Activation Function – value domain = [ReLU, tanh]; choice step

e. Learning Rate – value domain = [1e-2, 1e-3, 1e-4] – choice step

The binary cross entropy loss as well as the mean absolute error and accuracy metrics were utilized within the Adam optimizer.

Table 3. Sampled data for XAI analysis

Age Category	Gender	Age	Status
Young	F	26	1
	M	18	
	F	29	0
	M	25	
Old	F	58	1
	M	51	
	F	48	0
	M	64	

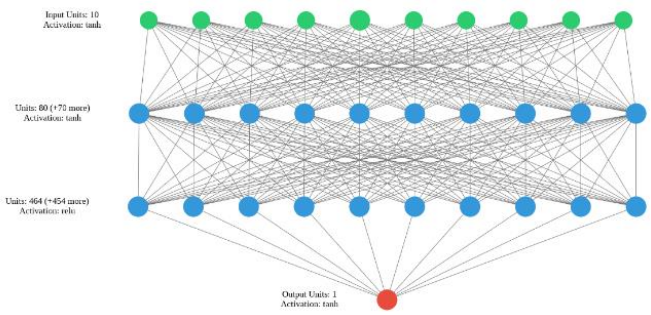


Figure 4: Model 1 - deep neural network architecture

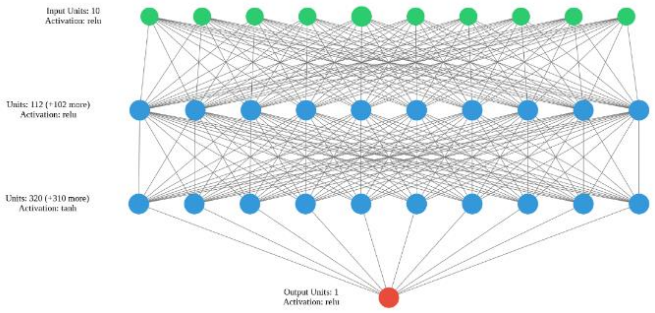


Figure 5: Deep neural network architecture (hyperparameter tuned)

Table 2. Classification model evaluation results

Model	Accuracy	Precision	Recall	F-Measure
1	0.81	**0.74**	**0.89**	**0.81**
2	**0.82**	0.72	0.81	0.76

Table 2 highlights the model evaluation results of the deep learning models for classifying liver disease patients. Based on Accuracy Model 2 had a slightly higher accuracy but lower precision, recall, and f-measure.

The utilized metrics are calculated as follows (where TP denotes True Positive; TN = True Negative; FP = False Positive; FN = False Negative:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F - Measure = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

Table 3 highlights the sampled (using purposive sampling), we selected four (4) of the youngest males (2) and female (2) patients, as well as four (4) of the oldest males (2) and females (2) – with one (1) of each sex being an individual with liver disease and the others with no liver disease. The aim of this was to describe the application of SHAP to deep learning models and inferring from the results based on observations within the dataset. In summary, why the model predicted what it predicted.

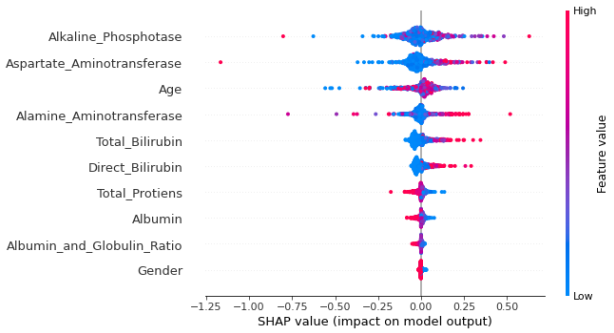


Figure 6: SHAP plots - deep neural network model (best model) - beeswarm plot

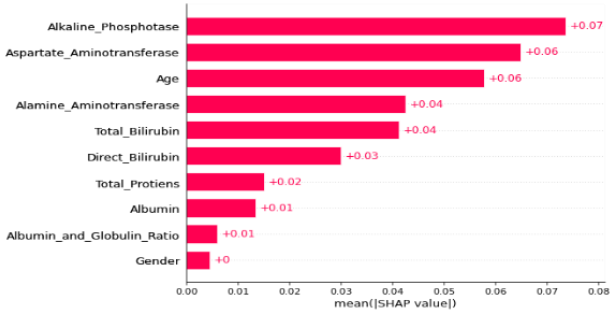
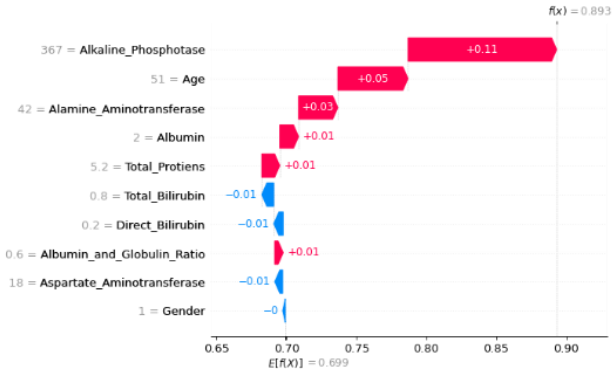


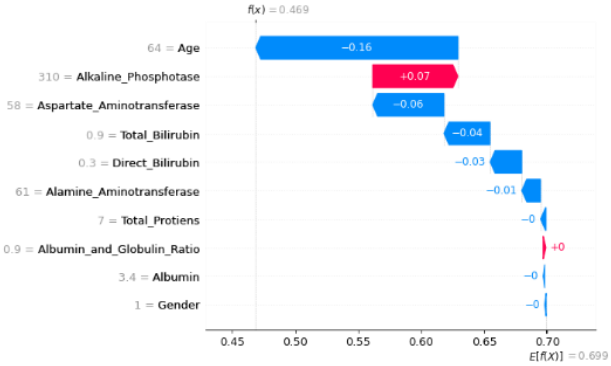
Figure 7: SHAP plots - deep neural network model (best model) - bar plot

In our study we utilized the DeepSHAP functionality due to the fact that it is tailor-made for deep learning models just like ours. Figures 6 and 7 highlight a Beeswarm plot and Bar plot respectively indicating the influence of predictors on the best deep learning model.

The results from our selected sample (from table 4) are presented in figures 8, and 9 for old as well as 10, and 11 for young individuals respectively. The results reveal the impact of certain features on the overall prediction for each selected sample observation; red indicative of the positive contribution and blue indicative of none or negative contribution to the overall outcome. Such results can aid medical staff in understanding how each individual patient's body may react to certain dosage of treatment, thus creating space for personalized treatment.

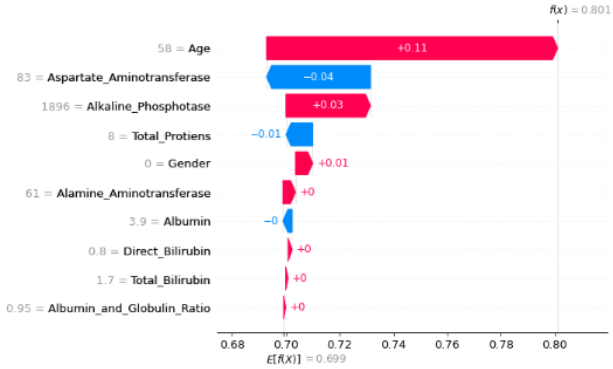


(a) Old male (status = liver disease)

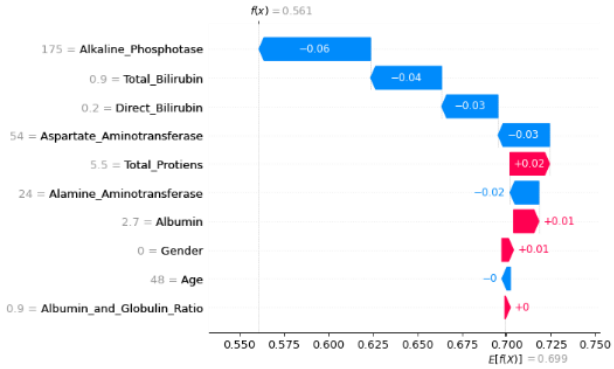


(b) Old Male (Status = No Liver Disease)

Figure 8. Comparative analysis of SHAP plots for two old males (a – with disease and b – no disease)

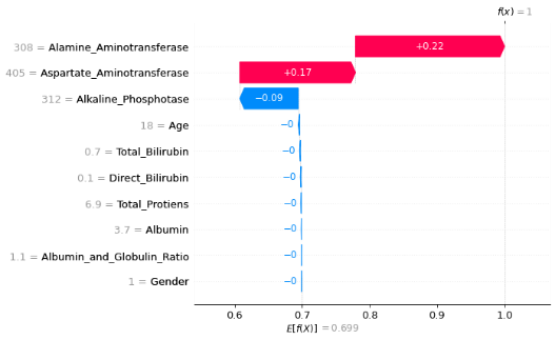


(a) Old female (status = liver disease)



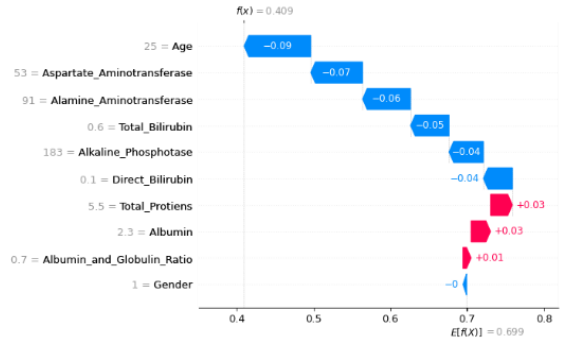
(b) Old Female (Status = No Liver Disease)

Figure 9: Comparative analysis of SHAP plots for two old females (a – with disease and b – no disease)



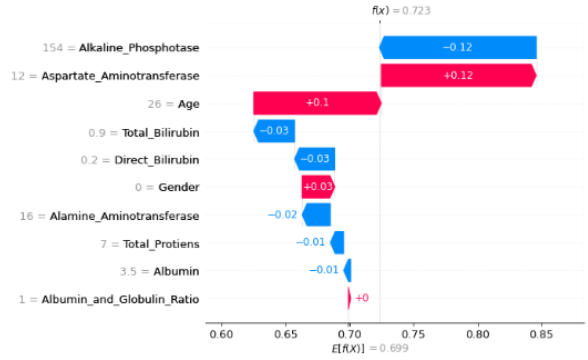
(a) Young male (status = liver disease)

(b)

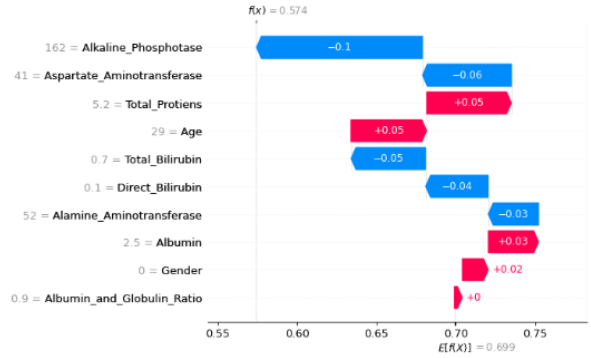


(b) Young male (status = no liver disease)

Figure 10: Comparative analysis of SHAP plots for two young males (a – with disease and b – no disease)



(a) Young Female (Status = Liver Disease)



(b) Young female (status = no liver disease)

Figure 11: Comparative analysis of SHAP plots for two young females

It can be observed that SHAP plots (from figures 8 and 9, as well as 10 and 11) vary by individual and this gives a more nuanced perspective to model prediction outcomes due to the ability to interpret each predicted outcome and provide personalized solutions to each patient (be it dietary, lifestyle or medical).

In more recent times, with the gradual growth in XAI, there has been some pushback (especially high-stakes decision making) [55]. These conversations will continue as AI research further develops. From our perspective, we conclude that, XAI can be used as a decision support tool provided the model is tested and meets robust real-world and ethical requirements of whichever industry it is needed for. Our research does not claim to propose XAI as the optimal decision support system within healthcare where models play high-stake roles because, in simple terms, XAI is not the remedy for a low performance model within the real world. As such, we recommend end-to-end machine learning which follows current MLOps industry guidelines such as: (a) Efficient Pipelines, Model Re-Training, and Monitoring (Symeonidis et al., 2022); (b) MLOps Maturity Model proposed by John et al. which encompasses Automated Data Collection, Automated Model Deployment, Semi-automated Model Monitoring, Fully-automated Model Monitoring, and as well incorporates governance and security protocols; (c) Responsible AI - Openness to Learning and Changing the Culture, Model Development Preparation, Selection of the Right Tools, Automating the Pipelines, and Monitoring [56], [57]. In summary, the power eXplainable Artificial Intelligence can be experienced, when intrinsically end-to-end AI implementation is done following appropriate MLOps guidelines.

6 Conclusion

As AI continues to gain ubiquity, Explainable AI's relevance is now more than ever essential in all spheres. Primarily in safety-critical domains such as healthcare, the need to interpret AI model predictions will go a long way to support medical treatment as well as personalized medicine.

This study sought to present the applicability of explainability within deep learning models, which have been known as black-box models within the AI sphere. We conducted a research summary on the applications of explainable AI (XAI) in biomedical research and utilized the Indian Liver Patient Dataset as a case study. Furthermore, making use of data-preprocessing, feature selection, data augmentation (with Generative Adversarial Network techniques for Tabular Data), and hyperparameter optimization, we developed deep learning classification models to classify liver disease. In addition, we integrated SHAP (Shapley Values) in interpreting the models, thus establishing model explainability. Finally, we discussed XAI and its implications and made recommendations.

With respect to theoretical implications, our work contributes to the extant literature and conversations on the explainable and interpretable AI paradigm primarily within the healthcare research sphere, i.e. adopting SHAP values. In like manner, our study serves as a contribution to research on data augmentation in the face of inadequate observations for deep learning models. It must be noted that, our research provides practical implications for researchers and health workers to adopt explainable models in supporting decision making process of medical diagnostics and prescription. Practically, our work is relevant to healthcare in deprived areas where trusted AI models (with explainable features) can be deployed on the edge to aid in affordable and mobile healthcare provision.

We recommend future research to reproduce our study within other medical contexts, as well as explore alternative explainable approaches to biomedical healthcare deep learning models. In addition, we recommend future research to delve into developing XAI frameworks or guideliness for healthcare implementation.

Conflict of interests: The authors declare no conflict of interest.

Author contributions: E.A.: Research formulation, Article Writing, and Analysis - Model Explainability Experimentation. D.M.B.: Analysis – ML Model Development, and Article Writing.

References

[1] Grassmann, F., Mengelkamp, J., Brandl, C., Harsch, S., Zimmermann, M. E., Linkohr, B., Peters, A., Heid, I. M., Palm, C., & Weber, B. H. (2018). A deep learning algorithm for prediction of age-related eye disease study severity scale for age-related macular degeneration from color fundus photography. Ophthalmology, 125(9), 1410–1420. https://doi.org/10.1016/j.ophtha.2018.02.037

[2] Nayak, J., Acharya U, R., Bhat, P. S., Shetty, N., Lim, T.-C., & others. (2009). Automated diagnosis of glaucoma using digital fundus images. Journal of Medical Systems, 33(5), 337–346. https://doi.org/10.1007/s10916-008-9195-z

[3] Raman, R., Dasgupta, D., Ramasamy, K., George, R., Mohan, V., & Ting, D. (2021). Using artificial intelligence for diabetic retinopathy screening: Policy implications. Indian Journal of Ophthalmology, 69(11), 2993–2998. https://doi.org/10.4103/ijo.IJO_1420_21

[4] Kohli, A., & Jha, S. (2018). Why CAD failed in mammography. Journal of the American College of Radiology, 15(3), 535–537. https://doi.org/10.1016/j.jacr.2017.12.029

[5] Sayres, R., Taly, A., Rahimy, E., Blumer, K., Coz, D., Hammel, N., Krause, J., Narayanaswamy, A., Rastegar, Z., Wu, D., & others. (2019). Using a deep learning algorithm and integrated gradients explanation to assist grading for diabetic retinopathy. Ophthalmology, 126(4), 552–564. https://doi.org/10.1016/j.ophtha.2018.11.016

[6] Cabitza, F., Rasoini, R., & Gensini, G. F. (2017). Unintended consequences of machine learning in medicine. Jama, 318(6), 517–518. https://doi.org/10.1001/jama.2017.7797

[7] Shebl, F. M., El-Kamary, S. S., Saleh, D. A., Abdel-Hamid, M., Mikhail, N., Allam, A., El-Arabi, H., Elhenawy, I., El-Kafrawy, S., El-Daly, M., & others. (2009). Prospective cohort study of mother-to-infant infection and clearance of hepatitis C in rural Egyptian villages. Journal of Medical Virology, 81(6), 1024–1031. https://doi.org/10.1002/jmv.21480

[8] Etzioni, R., Pepe, M., Longton, G., Hu, C., & Goodman, G. (1999). Incorporating the time dimension in receiver operating characteristic curves: A case study of prostate cancer. Medical Decision Making, 19(3), 242–251. https://doi.org/10.1177/0272989x9901900303

[9] Fan, C., Prat, A., Parker, J. S., Liu, Y., Carey, L. A., Troester, M. A., & Perou, C. M. (2011). Building prognostic models for breast cancer patients using clinical variables and hundreds of gene expression signatures. BMC Medical Genomics, 4(1), 1–15. https://doi.org/10.1186/1755-8794-4-3

[10] Telenyk, S., Czajkowski, K., Bidiuk, P., & Zharikov, E. (2019). Method of assessing the state of monuments based on fuzzy logic. 2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), 1, 500–506. https://doi.org/10.1109/idaacs.2019.8924315

[11] Dangare, C. S., & Apte, S. S. (2012). Improved study of heart disease prediction system using data mining classification techniques. International Journal of Computer Applications, 47(10), 44–48. https://doi.org/10.5120/7228-0076

[12] Vijiyarani, S., & Sudha, S. (2013). Disease prediction in data mining technique–a survey. International Journal of Computer Applications & Information Technology, 2(1), 17–21.

[13] Tang, Y., Wang, Y., Cooper, K. M., & Li, L. (2014). Towards big data Bayesian network learning-an ensemble learning based approach. 2014 IEEE International Congress on Big Data, 355–357. https://doi.org/10.1109/bigdata.congress.2014.58

[14] Bodyanskiy, Y., Perova, I., Vynokurova, O., & Izonin, I. (2018). Adaptive wavelet diagnostic neuro-fuzzy network for biomedical tasks. 2018 14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET), 711–715. https://doi.org/10.1109/tcset.2018.8336299

[15] Perova, I., Brazhnykova, Y., Bodyanskiy, Y., & Mulesa, P. (2018). Neural network for online principal component analysis in medical data mining tasks. 2018 IEEE First International Conference on System Analysis & Intelligent Computing (SAIC), 1–5. https://doi.org/10.1109/tcset.2018.8336299

[16] Perova, I., Litovchenko, O., Bodvanskiy, Y., Brazhnykova, Y., Zavgorodnii, I., & Mulesa, P. (2018). Medical data-stream mining in the area of electromagnetic radiation and low temperature influence on biological objects. 2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP), 3–6. https://doi.org/10.1109/dsmp.2018.8478577

[17] Perova, I., & Mulesa, P. (2015). Fuzzy spacial extrapolation method using Manhattan metrics for tasks of Medical Data mining. 2015 Xth International Scientific and Technical Conference" Computer Sciences and Information Technologies"(CSIT), 104–106. https://doi.org/10.1109/stc-csit.2015.7325443

[18] Izonin, I., Trostianchyn, A., Duriagina, Z., Tkachenko, R., Tepla, T., & Lotoshynska, N. (2018). The combined use of the wiener polynomial and SVM for material classification task in medical implants production. International Journal of Intelligent Systems and Applications, 10(9), 40–47. https://doi.org/10.5815/ijisa.2018.09.05

[19] Melnykova, N., Shakhovska, N., & Sviridova, T. (2017). The personalized approach in a medical decentralized diagnostic and treatment. 2017 14th

International Conference The Experience of Designing and Application of CAD Systems in Microelectronics (CADSM), 295–297. https://doi.org/10.1109/cadsm.2017.7916139

[20] Baltussen, R., & Niessen, L. (2006). Priority setting of health interventions: The need for multi-criteria decision analysis. Cost Effectiveness and Resource Allocation, 4(1), 1–9. https://doi.org/10.2139/ssrn.943814

[21] Russell, S. J. (2010). Artificial intelligence a modern approach. Pearson Education, Inc. https://doi.org/10.1016/0004-3702(96)00007-0

[22] Urdea, M., Penny, L. A., Olmsted, S. S., Giovanni, M. Y., Kaspar, P., Shepherd, A., Wilson, P., Dahl, C. A., Buchsbaum, S., Moeller, G., & others. (2006). Requirements for high impact diagnostics in the developing world. Nature, 444(1), 73–79. https://doi.org/10.1038/nature05448

[23] Drain, P. K., Hyle, E. P., Noubary, F., Freedberg, K. A., Wilson, D., Bishai, W. R., Rodriguez, W., & Bassett, I. V. (2014). Diagnostic point-of-care tests in resource-limited settings. The Lancet Infectious Diseases, 14(3), 239–249. https://doi.org/10.1016/s1473-3099(13)70250-0

[24] Pereira, S., Pinto, A., Alves, V., & Silva, C. A. (2016). Brain tumor segmentation using convolutional neural networks in MRI images. IEEE Transactions on Medical Imaging, 35(5), 1240–1251.

[25] Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P.-M., & Larochelle, H. (2017). Brain tumor segmentation with deep neural networks. Medical Image Analysis, 35, 18–31.

[26] Moreira, I. C., Amaral, I., Domingues, I., Cardoso, A., Cardoso, M. J., & Cardoso, J. S. (2012). Inbreast: Toward a full-field digital mammographic database. Academic Radiology, 19(2), 236–248.

[27] PUB, M. H., Bowyer, K., Kopans, D., Moore, R., & Kegelmeyer, P. (n.d.). The digital database for screening mammography. Proceedings of the Fifth International Workshop on Digital Mammography, 212–218.

[28] Ngo, T. A., Lu, Z., & Carneiro, G. (2017). Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance. Medical Image Analysis, 35, 159–171. https://doi.org/10.1016/j.media.2016.05.009

[29] Roth, H. R., Farag, A., Lu, L., Turkbey, E. B., & Summers, R. M. (2015). Deep convolutional networks for pancreas segmentation in CT imaging. Medical Imaging 2015: Image Processing, 9413, 378–385. https://doi.org/10.1117/12.2081420

[30] Prasoon, A., Petersen, K., Igel, C., Lauze, F., Dam, E., & Nielsen, M. (2013). Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network. International Conference on Medical Image Computing and Computer-Assisted Intervention, 246–253. https://doi.org/10.1007/978-3-642-40763-5_31

[31] Liao, S., Gao, Y., Oto, A., & Shen, D. (2013). Representation learning: A unified deep learning framework for automatic prostate MR segmentation. International Conference on Medical Image Computing and Computer-Assisted Intervention, 254–261. https://doi.org/10.1007/978-3-642-40763-5_32

[32] Guo, Y., Wu, G., Commander, L. A., Szary, S., Jewells, V., Lin, W., & Shen, D. (2014). Segmenting hippocampus from infant brains by sparse patch matching with deep-learned features. International Conference on Medical Image Computing and Computer-Assisted Intervention, 308–315. https://doi.org/10.1007/978-3-319-10470-6_39

[33] Kim, M., Wu, G., & Shen, D. (2013). Unsupervised deep learning for hippocampus segmentation in 7.0 Tesla MR images. International Workshop on Machine Learning in Medical Imaging, 1–8. https://doi.org/10.1007/978-3-319-02267-3_1

[34] Schlegl, T., Waldstein, S. M., Vogl, W.-D., Schmidt-Erfurth, U., & Langs, G. (2015). Predicting semantic descriptions from medical images with convolutional neural networks. International Conference on Information Processing in Medical Imaging, 437–448. https://doi.org/10.1007/978-3-319-19992-4_34

[35] Xu, Y., Li, Y., Wang, Y., Liu, M., Fan, Y., Lai, M., Eric, I., & Chang, C. (2017). Gland instance segmentation using deep multichannel neural networks. IEEE Transactions on Biomedical Engineering, 64(12), 2901–2912. https://doi.org/10.1109/tbme.2017.2686418

[36] Lerouge, J., Hérault, R., Chatelain, C., Jardin, F., & Modzelewski, R. (2015). IODA: An input/output deep architecture for image labeling. Pattern Recognition, 48(9), 2847–2858. https://doi.org/10.1016/j.patcog.2015.03.017

[37] Moeskops, P., Viergever, M. A., Mendrik, A. M., De Vries, L. S., Benders, M. J., & Išgum, I. (2016). Automatic segmentation of MR brain images with a convolutional neural network. IEEE Transactions on Medical Imaging, 35(5), 1252–1261. https://doi.org/10.1109/tmi.2016.2548501

[38] Shin, H.-C., Orton, M. R., Collins, D. J., Doran, S. J., & Leach, M. O. (2012). Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4D patient data. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8), 1930–1943. https://doi.org/10.1109/tpami.2012.277

[39] Leung, M. K., Delong, A., Alipanahi, B., & Frey, B. J. (2015). Machine learning in genomic medicine: A review of computational problems and data sets. Proceedings of the IEEE, 104(1), 176–197. https://doi.org/10.1109/jproc.2015.2494198

[40] Van Lent, M., Fisher, W., & Mancuso, M. (2004). An explainable artificial intelligence system for small-unit tactical behavior. Proceedings of the National Conference on Artificial Intelligence, 900–907.

[41] Adadi, A., & Berrada, M. (2020). Explainable AI for Healthcare: From Black Box to Interpretable Models.

Scopus. https://doi.org/10.1007/978-981-15-0947-6_31

[42] Nazar, M., Alam, M. M., Yafi, E., & Su'Ud, M. M. (2021). A Systematic Review of Human-Computer Interaction and Explainable Artificial Intelligence in Healthcare with Artificial Intelligence Techniques. IEEE Access. https://doi.org/10.1109/ACCESS.2021.3127881

[43] Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. ACM Computing Surveys (CSUR), 51(5), 1–42. https://doi.org/10.1145/3236009

[44] Xu, Y., Yang, X., Gong, L., Lin, H.-C., Wu, T.-Y., Li, Y., & Vasconcelos, N. (2020). Explainable object-induced action decision for autonomous vehicles. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 9523–9532. https://doi.org/10.1109/cvpr42600.2020.00954

[45] Kim, J., Moon, S., Rohrbach, A., Darrell, T., & Canny, J. (2020). Advisable learning for self-driving vehicles by internalizing observation-to-action rules. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 9661–9670. https://doi.org/10.1109/cvpr42600.2020.00968

[46] Brunese, L., Mercaldo, F., Reginelli, A., & Santone, A. (2020). Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays. Computer Methods and Programs in Biomedicine, 196, 105608. https://doi.org/10.1016/j.cmpb.2020.105608

[47] Ramana, B., Babu, M., & Venkateswarlu, N. (2012). ILPD (Indian Liver Patient Dataset) Data Set.

[48] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30.

[49] Fryer, D., Strümke, I., & Nguyen, H. (2021). Shapley values for feature selection: The good, the bad, and the axioms. IEEE Access, 9, 144352–144360. https://doi.org/10.1109/access.2021.3119110

[50] Maniquet, F. (2003). A characterization of the Shapley value in queueing problems. Journal of Economic Theory, 109(1), 90–103. https://doi.org/10.1016/s0022-0531(02)00036-4

[51] Souza, J., & Leung, C. K. (2021). Explainable Artificial Intelligence for Predictive Analytics on Customer Turnover: A User-Friendly Interface for Non-expert Users. In Explainable AI Within the Digital Transformation and Cyber Physical Systems (pp. 47–67). Springer. https://doi.org/10.1007/978-3-030-76409-8_4

[52] Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J. (2021). Explainable machine learning in credit risk management. Computational Economics, 57(1), 203–216. https://doi.org/10.1007/s10614-020-10042-0

[53] Gramegna, A., & Giudici, P. (2020). Why to buy insurance? An explainable artificial intelligence approach. Risks, 8(4), 137. https://doi.org/10.3390/risks8040137

[54] Tallón-Ballesteros, A., & Chen, C. (2020). Explainable AI: Using Shapley value to explain complex anomaly detection ML-based systems. Machine Learning and Artificial Intelligence, 332, 152. https://doi.org/10.3233/faia200777

[55] Rudin, C. (2019). Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. Nature Machine Intelligence, 1(5), 206–215. https://doi.org/10.1038/s42256-019-0048-x

[56] John, M. M., Olsson, H. H., & Bosch, J. (2021). Towards mlops: A framework and maturity model. 2021 47th Euromicro Conference on Software Engineering and Advanced Applications (SEAA), 1–8. https://doi.org/10.1109/seaa53835.2021.00050

[57] Matsui, B. M., & Goya, D. H. (2022). MLOps: A Guide to its Adoption in the Context of Responsible AI. 2022 IEEE/ACM 1st International Workshop on Software Engineering for Responsible Artificial Intelligence (SE4RAI), 45–49. https://doi.org/10.1145/3526073.3527591

Simulation for Dynamic Patients Scheduling Based on Many Objective Optimization and Coordinator

Ali Nader Mahmed, M. N. M. Kahar*
First Faculty of Computing, College of Computing and Applied Sciences Universiti Malaysia Pahang, Malaysia
E-mail: mnizam@ump.edu.my
*Corresponding author

Keywords: dynamic, hospital admission and scheduling, patients' admission scheduling, multi-objective optimization, many objective optimizations, non-dominated optimization

Received: October 4, 2023

The Patient Admission Scheduling Problem (PASP) involves scheduling patient admissions, hospital time locations, to achieve certain quality of service and cost objectives, making it a multi-objective combinatorial optimization problem and NP-hard in nature. In addition, PASP is used in dynamic scenarios where patients are expected to arrive at the hospital sequentially, which requires dynamic optimization handling. Taking both aspects, optimization and dynamic utilization, we propose a simulation for dynamic patient scheduling based on multi-objective optimization, window, and coordinator. The role of multi-objective optimization deals with many soft constraints and providing a set of non-dominated solution coordinators. The role of the counter is to collect newly arrived patients and previously unconfirmed patients with the aim of passing them on to the coordinator. Finally, the role of the coordinator is to select a subset of patients from the window and pass them to the optimization algorithm. On the other hand, the coordinator is also responsible for those selected from the non-dominant solutions to activate it in the hospital and decide on unconfirmed employees to place them in the window for the next round. Simulator evaluation and comparison between several optimization algorithms show the superiority of NSGA-III in terms of set criticality and soft constraint values. Therefore, it treats PASP as a multi-objective dynamic optimization of a useful solution. NSGA-II is guaranteed 0.96 percent dominance over NSGA-II and 100 percent dominance of all other algorithms.

Povzetek: Gre za dinamično razporejanje pacientov z uporabo večciljne optimizacije, ki obravnava kompleksni problem razporejanja sprejema pacientov v bolnišnico, izboljšuje kakovost storitev in učinkovitost z uporabo NSGA-II algoritma za optimizacijo.

1 Introduction

In the 21st century, life expectancy doubled globally, and new health delivery models and technologies are predicted to considerably extend healthy life expectancy [1]. The demand for healthcare services has risen in recent decades because of an ageing population and advancements in preventative care [2], yet the healthcare sector is still under pressure. to reduce costs and raise standards of treatment. The healthcare industry has mainly shifted its focus to a value-based strategy to offset a potential increase in clinical medicine expenditures that are not accompanied by appreciable improvements in health outcomes [3]. In this situation, achieving the greatest results at the lowest cost is the main objective, making effective resource management and patient happiness crucial but competing goals that health care administrators must meet. Practical concerns including admissions control, process design, aggregate planning, capacity distribution, and appointment scheduling must be taken care of in order to address this obstacle. The Patient Admission Scheduling Problem is one of these issues (PASP). Patients' admission scheduling problem (PASP) is how to plan patient's admission and their location and time in the hospital in order to meet certain quality of service and cost objectives [4]. It is considered as complex combinatorial optimization process with many constraints [5]. This is because it involves allocating resources for patients according to the condition of the hospital and the condition of the patient in order to meet the satisfactory level of the patient within the time limit for scheduling. Choosing an appropriate room to allocate to patients while taking into account medical needs, patient demands, and hospital resource availability is the focus of the patient bed assignment problem (PBAP), a PASP sub-task [6]. It is considered as a paramount problem for hospitals and medical centers. PBAP is an NP-hard problem [7]. For solving PBAP, it is needed to create an autonomous system that receives patients requests online or through phone and automatically assign them to beds without the need of human intervention. A conceptual representation of this process is depicted in Figure 1 and the result is mapping patients to the best beds inside the rooms for meeting both the health and satisfaction requirements.

Patient Scheduling is regarded as constrained combinatorial optimization problem with NP-hard nature. Adding the dynamic in terms of patient's arrival and change of preference to the problem makes more complex.

In addition, the problem has a limitation in terms of capacity of the room which leads to a condition of over-crowding that needs to be minimized. Another added complexity to the problem is the need to identify various information of the patients' conditions, their special need and the criticality of their cases before performing the mapping. The process should be automated in order to facilitate the management of the hospitals and to increase the quality of service within the allocated cost.

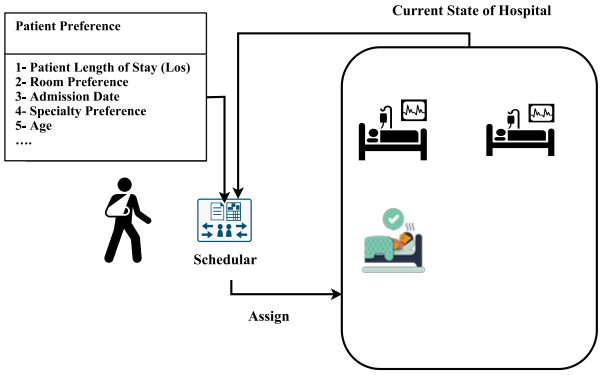


Figure 1: Conceptual representation of the process of automatic PBAP in hospitals.

Meta-heuristic searching optimization algorithms are set of optimization algorithms with capability of solving complex optimization problem based on generating candidate solutions randomly and enabling an evolving of them based on heuristics [8]. The literature contains wide range of meta-heuristic optimization algorithms, some of them are inspired from biological phenomena such as genetic algorithm [9], others are inspired from physical phenomena such as simulation annealing [10]. In addition, there is numerous metaphors used for deriving meta-heuristic algorithm such as ant colony [11], artificial bee colony [12], particle swarm optimization [13]…etc. Despite the type of the metaphor, we can classify the meta-heuristic optimization algorithms into two main categories: single objective and multi-objective [14]. In the single objective, the algorithm aims at optimizing a formulated a single objective function from the problem definition while in the multi-objective, the algorithm aims at optimizing simultaneously multi-objective functions using the concept of Pareto domination. The latter type can be utilised to solve PBAP by treating soft-constraint violations as multi-objective functions that must be minimised during optimization [15]. A strategy for enabling the algorithm to take into account the dynamical nature of the problem must be developed before a multi-objective meta-heuristic optimization algorithm may be used directly. In this article, we propose a simulation that extends the optimization with additional steps in order to enable dynamic scheduling for PBAP using multi-objective optimization. In addition, we provide an algorithm for selecting one solution of the pareto front to use it for providing the allocation decision under two sets: confirmed allocated patients and non-confirmed allocated patients.

The rest of the article is divided into the following sections. We present the contribution in section 2. Next, the literature survey is presented in section 3. Afterwards, a background is provided in section 4. In addition, we present the methodology in section 5. Next, experimental works and evaluation are provided. Lastly, the conclusion and future work are provided in section 7.

2 Contribution

The development of dynamic patient scheduling that supports many patient objectives is the ultimate purpose of this study. The contributions listed below are presented in this article.

1. To the best of our knowledge, this article provides the first in terms of simulating arrival of patients to hospital and an algorithm for scheduling using multi-objective optimization, solutions selection, and scheduler.
2. The scheduling in this article avoids implicit constraint that causes greedy behavior by using the concept of non-confirmed patients. More specifically, it provides list of non-confirmed patients which automatically feeds another list of confirmed patients when their scheduling day is within less than D days on one side and provides the remaining patients inside list of non-confirmed patients to a new call of optimization on the other side.
3. This article enables dynamic multi-objective optimization through solution selection. More specifically, considering that multi-objective optimization algorithm provides Pareto front which represents set of non-dominated solutions, one solution is to be selected for enabling or scheduling. In order to do so, the algorithm performs solution selection using weighted summation of the objectives with respect to their corresponding preference.
4. In order to distinguish between patients that are allowed for rescheduling from new arrived patients, we use variable length optimization (VLO). In VLO, different lengths of solutions are used where each solution allow for rescheduling of different sub-sets of the non-confirmed patients.

3 Literature survey

There are two subsections in the literature. The first is the patient admission scheduling literature, which is discussed in sub-section 2.1. The second is discussed in sub-section 2.2 and is about the application of multi-objective optimization methods to scheduling problems.

3.1 Patient admission scheduling

In the work of [16] which has aimed at solving the problem of PAS based on offline perspective. His proposed combinatorial formulation of the optimization problem of PAS using integer linear programming and

proposed Tabu search algorithm for solving it. They aimed at finding the optimal bed assignments for elective patients based on pre-knowledge of the hospital departments, rooms capacity, beds availability, equipment, technical issues, and qualitative elements like the patient's choice for gender, age, and room compatibility. Their work has drawn criticism from a number of angles, including the impracticality of an offline solution given the dynamic nature of the issue, and considering optimizing a weighted average of the soft-constraints which can cause sub-optimality due to the non-convexity of the model or limit the choices to the decision maker due to providing only single optimal solution. In the work of [17], Fix-and-Relax (F&R) and Fix-and-Optimize (F&O) are techniques based on Mixed Integer Programming (MIP) that break down PAS problem instances into smaller chunks before optimising the smaller chunks. More specifically, iteratively improved Quick solutions produced by the F&R heuristic are fed into the F&O heuristic. Patient length of stay (LoS), room preference, admission date, specialty preference, age, as well as time decomposition taking different optimization window sizes, are the factors that have employed decompositions. Ceschia and Schaerf (2016) suggested a different formulation for the demester problem known as a dynamic patient-to-room assignment problem that helped reduce the number of decision variables, compute different lower bound values by omitting some constraints, and adapt simulated annealing to find the best solution. The work of [16] has also been improved by [18] by including local search moves into two tiers of heuristics or hyper-heuristics. The great deluge algorithm was used in this work as a component of the hyper-heuristic, but it was criticised in the work of [19] due to the linear decay rate of its deluge algorithm, which was improved to non-linear adaptive decay rate using the same soft and hard constraints of demester [16] . The scheduling goals in the work [20] were divided into short- and long-term goals, and periodic re-optimization was employed. Using column generation and Dantzig-Wolfe decomposition, the lower bounds are computed. A scheduling algorithm is used in the research of [21] to schedule tourist travel to destination medical centres. The goals are to keep patients' preferred commencement days and flow times as close to real time as possible. They scheduled everything using a flow-shop system. Simulated annealing and tabu search were used with simulation for optimization. The simulation is based on discrete event simulation, which assesses the solution considering the admission day, admission time, and patient sequence as decision factors on each day. The current deterministic model created by [16] was modified in the work of [22] to become stochastic. To represent the arrivals and departures, they employed discrete phase type distribution and a Poisson distribution, respectively. Hence, their model has evolved from the previous deterministic one into a stochastic one. The work of [9] involved the modelling of appointment times that depend on both the needs of the patients and the speed factor of the doctors' performance. Their model is solved utilising a genetic algorithm for large-scale problems and a single solver for small-scale problems. Overall, the literature has addressed the PAS issue from a variety of angles and levels of practicality, including the addition of soft limits, the unpredictability of LoS, and the acceptance of urgent patients. However, the non-domination component of the issue has not been addressed by any of the prior solutions. When dealing with the soft restrictions as separate objectives, the PAS problem is a multi-objective optimization problem. In this manner, By using the penalty concept, we give the decision-maker more options and reduce the disadvantage of the linear combination of soft constraints under the weighted average formula, taking into account that the latter has no application in the problem and the linear combination of constraints does not correspond to the real-world model. Table 1 presents an overview of the existing approaches of PAS in the literature with the key features and criticism and improvements. Table 2 provides a summary of the various methods for patient bed scheduling.

Table 1: Summary of patient admission scheduling (PAS) approaches in literature

Reference	**Approach/Technique**	**Key Features**	**Criticisms/Improvements**
[9]	Genetic Algorithm	Appointment times based on patient needs and doctor performance	Provides single solution
[16]	Integer Linear Programming & Tabu Search	Offline solution, optimal bed assignments considering various constraints	Criticized for impracticality in dynamic settings; limited by single optimal solution
[17]	Fix-and-Relax (F&R), Fix-and-Optimize (F&O)	Decomposition of PAS problem, Mixed Integer Programming	Subject to local minima because of decomposition
[18]	Hyper-heuristics & Great Deluge Algorithm	Improved [16] by local search moves	Criticized for linear decay rate in deluge algorithm
[19]	Non-linear Adaptive Decay Rate	Improved [18] using non-linear adaptive decay rate	Not handling dynamic environment
[20]	Column Generation & Dantzig-Wolfe Decomposition	Short- and long-term scheduling goals, periodic re-optimization	Not handling dynamic environment
[21]	Flow-shop System, Simulated Annealing & Tabu Search	Scheduling for tourist travel to medical centres	Local search capability only
[22]	Stochastic Model	Discrete phase type distribution, Poisson distribution	Evolution from deterministic to stochastic model
[23]	Dynamic patient-to-room assignment	Reduced decision variables, simulated annealing	It does not have global search capability

Table 1: Overview of the various approaches for patient, bed scheduling.

Author	Hard constraints	Soft-constraint	Objective function	Algorithm	Limitation
[16]	8	5	Weighted average	Tabu Search	Sub-optimality due to weighted average and non-convexity
[24]	2	4	Weighted average	Simulated annealing	Weighted average causes sub-optimal result
[25]	3	6	Weighted sum	Hyper-heuristic	Weighted average causes sub-optimal result
[26]	5	3	Weighted sum	deluge algorithm	Weighted average causes sub-optimal result
[27]	-	8	Weighted sum	Mixed Integer Programming (MIP)	more computational time
[28]	12	2	Weighted sum	tabu search (TS) and simulated annealing (SA) with simulation	Not including resource utilization, age and gender
[9]	15	4	Weighted sum	Genetic algorithm	Concern about convergence, sub-optimality due to weighted sum

3.2 Multi-objective optimization for scheduling

Various scheduling issues and applications have been solved using the multi-objective particle swarm optimization technique. Modified multiple-objective particle swarm optimization (MMOPSO), which was proposed by Ghasemi, Khalili-Damghani, et al. in 2019, was used to solve a mixed-integer mathematical programming model for the earthquake reaction phase. Two local search operations are included in the improved multi objective particle swarm optimization. The model considers two target functions: lowering the total cost of facility location and allocation, as well as decreasing the amount of supply deficit. This method beat out the two well-known non-dominated sorting genetic algorithms, NSGA-II and epsilon constraint method, in tests. In the study of Adhikari and Srirama (2019), a modified variation of multi-objective particle swarm optimization was used to optimise the problem of container-based scheduling for the Internet of Things in a cloud context. Energy usage and computing time are the two optimization goals that the writers have considered. To assess the quality of the solution, the weighted sum approach-based fitness function is used to cope with the multi-objective elements.

The acceleration component of multi-objective particle swarm optimization changed the convergence speed. Considering that the typical PSO looks for the best possible solution by combining the individual and current global bests of the particles the acceleration PSO (APSO) approach, which is a modification of the PSO algorithm based on its velocity and displacement, was developed in (Yang, Deb et al. 2011) due to the limits of convergence speed and accuracy. The APSO approach lowers unpredictability as iterations continue by using the individuals that perform best globally. In the study by Fang and Popole (2019), which generated neighborhoods for each particle and used the self-organizing mapping (SOM) approach to select the neighborhood best solution, the particle swarm optimization was modified once again to enhance its searching performance. Analytical research of the convergence of self-adaptive PSO (APSO) with the purpose of presenting a parameter selection method that ensures the convergence was carried out in the work of [29]. Using the suggested SAPSO, they created the SAMOPSO MOO framework, which is based on SAPSO. They also create an external repository that stores the non-dominated solutions in order to obtain a well-distributed Pareto front. The proposed MOO system then uses a cyclic sorting mechanism to update the external repository while integrating elitist-preserving principles. Particle swarm optimization has been modified in the work of [30] to tackle large dimensional discrete variables. To enhance the performance, the method included stretching and changing neighborhood search techniques. Jumping PSO, variable neighborhood search, and the stretching approach are all included in their whole integrated model. Non-dominated sorting genetic algorithm was slightly adjusted and used to solve the scheduling of surgeries in operating rooms in the work of [31]. This work shows that the modification of the searching algorithm is not limited to particle swarm optimization method. The resolved model is a resource allocation methodology that primarily concentrates on allocating operating rooms (ORs) for each surgical specialty (SS). The initialization of the population and the selection using the tournament comprised the first part of the change to NSGA-II. An idea for a multi-patent crossover genetic algorithm appeared in the publication [32]. When it functions for n parents, their definition of the multi-parent operator is to define the cross operator with n string division points. Overall, scheduling problems with a multi-objective nature may be solved well using meta-heuristic search optimization techniques. However, the bulk of methods for resolving issues with a limited number of objectives employed algorithms. Given that changing the PAS problem to a mulz3ti-objective problem entails a large number of objectives derived from soft constraints, in order to ensure convergence behavior, the addition of a large number of objectives necessitates particular adjustment to the searching criteria. Aside from that We can observe that the scheduling programme made use of a meta-heuristic multi-objective optimization approach that included particle and genetic based searches. Additionally, the bulk of them require special operator designs depending on the application's nature and cannot be used directly. Table 2 lists all of the papers that addressed the PAS/NRP dilemma. It is observed from Table 1 that the literature contains many multi-objective metaheuristic algorithms, however, all of them have dealt with the multi-objective as single objective based on

weighted average of the objectives which subject to local minima. To handle this, it is needed to propose non-dominated sorting based multi-objective optimization. On the other side, we observe from Table 2 that the number of soft-constraints ranges between 5 to 10 which makes the problem as candidate many objective optimizations instead of traditional multi-objective optimization when we consider the soft-constraints as objectives of the problem.

Table 2: Pseudocode of the process of selecting non-dominated solutions based on the process of NSGA-III.

Input:
- H structured reference points Zs or supplied aspiration
- points Za,
- parent population Pt

Output:
- P(t+1)

Start
1: St=∅,i=1
2: Qt = Recombination+Mutation(Pt)
3: Rt = Pt ∪ Qt
4: (F1,F2,...)=Non-dominated-sort(Rt)
5: repeat
6: (St = St ∪ Fi and i = i + 1
7: until|St|≥N)
8: Last front to be included: Fl = Fi
9: if|St|=N then
10: P(t+1) = St, break;
11: else
12: P = all previous fronts
13: Points to be chosen from Fl: K = N − |Pt+1|
14: Normalize objectives and create reference set Zr: Normalize(fn,St,Zr,Zs,Za)
15: Associate each member s of St with a reference point: [π(s),d(s)] =Associate(St,Zr) % π(s): closest reference point, d: distance between s and π(s)
16: Compute niche count of reference point
17: Choose K members one at a time from Fl to construct P(t+1): Niching(K, ρj, π, d, Zr, Fl, P(t+1))
18: End If
End

Table 3: Review of articles worked on solving PAS problem

Author	**Application**	**Hard constraints**	**Soft constraints**	**Optimization method**	**Type**
Demester [16]	PBAS	8	5	Hybrid Tabu search with heuristics	Static
Sara [33]	PBAS	2	10	Tabu local search	Dynamic
Saif [19]	PBAS	5	6	Adaptive deluge algorithm	Static

Table 2: Overview of multi-objective optimization in scheduling problems

Reference	Method/Technique	Key Features	Application	Limitations/Improvements
Ghasemi, Khalili-Damghani, et al. (2019)	MMOPSO	Mixed-integer model, focus on cost and supply deficit	Earthquake response	Superior to NSGA-II and epsilon constraint method
Adhikari and Srirama (2019)	Modified PSO	Optimizes energy use and computing time	IoT scheduling in cloud	Weighted sum approach for multi-objective handling
Yang, Deb et al. (2011)	APSO	Improved convergence through individual and global bests	General optimization	Reduces unpredictability, addresses speed and accuracy limits
Fang and Popole (2019)	Modified PSO with SOM	Neighborhood generation, neighborhood best solution selection	PSO performance enhancement	Provides only single solution
[23]	SAPSO & SAMOPSO	Self-adaptive PSO, external repository for Pareto front	Multi-objective optimization framework	Cyclic sorting, elitist-preserving principles
[24]	Modified PSO	Addresses large dimensional discrete variables	General optimization	Uses stretching, neighborhood search techniques
[25]	Modified NSGA-II	Resource allocation in operating room scheduling	Surgery scheduling	Focuses on allocating ORs to surgical specialties
[26]	Multi-parent crossover genetic algorithm	Multi-parent operator for n parents	Genetic algorithm variation	Does not have non-domination sorting perspective

4 Research gap

It is observed that in the domain of Patient Admission Scheduling (PAS) and similar scheduling challenges, most studies predominantly utilize techniques that manage multiple objectives through a weighted average approach. While this method is widely accepted, it is often prone to leading to local minima, thereby potentially yielding suboptimal solutions.

Furthermore, the literature demonstrates a significant absence of non-dominated sorting approaches in multi-objective optimization for scheduling problems. Non-dominated sorting plays a crucial role in identifying truly optimal solutions across a range of objectives, without unfairly favoring any single one. This aspect of optimization is particularly important in scenarios where a balanced consideration of multiple factors is essential.

Additionally, the current methodologies in the field largely concentrate on traditional multi-objective optimization. However, in scenarios such as PAS, where the number of soft constraints is considerable, ranging between 5 to 10, the issue becomes more aligned with many-objective optimization. This transition from multi-objective to many-objective optimization is not sufficiently addressed in the existing research, indicating a gap in the approach to handling complex scheduling problems with a multitude of objectives.

5 Methodology

This section presents the developed methodology for our dynamic patient's admission scheduling. It starts with presenting the pre-processing in sub-section 6.1. Next, the window- based NSGA-III in sub-section 6.2. Next, we present the selection of confirmed and non-confirmed patients in sub-section 6.3. Afterwards, the variable length optimization of window- based NSGA-III is given in sub-section 6.4. Lastly, the evaluation metrics are provided in sub-section 6.5.

1.1 Problem formulation

Assuming that we have a hospital combined of set of departments D under various specialisms S and each department contains set of rooms under the department R. In addition, we assume that we have an arrival rate of patients to the hospital where each patient requires serving it within certain number of nights inside a preferred department and by type of specialism. In addition, each room has certain capacity for accommodating pre-defined number patients at once. Our problem is about allocating the patients inside the rooms within period of time (number of nights) using solution vector x with minimizing the violation of soft-constraints $(f_1, f_2, \dots f_n)$ and preventing the violation of hard-constraint $(h_1, h_2, \dots h_m, g_1, g_2, \dots g_k)$.

The solution is combined of set of components that defines the allocation of each patient at each night for the selected room. In other words, the solutions length equals to the number of patients, and each component inside the solutions is a tuple of tree values, namely, the index of the bed that is assigned to the patient, the starting night, and the ending night. This problem is formulated as multi-objective optimization problem as:

$$x = argmin\ (f_1, f_2, \dots f_n) \quad (1)$$

$$s.t.\ \ g_1 = 0, g_2 = 0, \dots, g_k = 0$$
$$h_1 \geq 0\ , h_2 \geq 0, \dots h_m \geq 0$$

Hence, the problem is formulated mathematically as multi-objective optimization problem with many objective functions, many hard and soft-constraint. According to [17], this is regarded as NP-hard problem.

Assuming that the outcome of the optimization after running at time t it is PF_t. We use the penalties of the soft-constraint to provide ranking of the solutions based on the overall cost in a descending manner. This is done using this Equation (2)

$$y_j = \sum_{i=1}^{NSC} w_i f_i(x_j) \quad (2)$$

Where:

x_j is a solution selected from the Pareto Front

w_i is the penalty that is associated with the soft-constraint i

y_j is the overall cost of the solution x_j

Next, we select the solution that has the lowest cost as the activated solution. From the activated solution, the algorithm selects the patients that are scheduled within three days as confirmed patients and the patients that are scheduled later than three days as non-confirmed patients.

The optimization problem is repeated in different days with different number of patients. The changing of the number of patients implies changing the length of the solution space. The algorithm will work on allocating selected patients of the non-confirmed list of patients.

1.2 Simulator

The simulator is presented in Figure 2. The newly arrived patients are fed into the scheduler which is responsible on receiving a solution from the solution selection block, and providing it to the list of non-confirmed patients. The list of non-confirmed patients provides its non-confirmed patients to a new call of the optimization algorithm and provides the patients that have their scheduled day within less than D days to the confirmed patients list through sub-block named confirm. The optimization algorithm operates on different lengths of solutions because of the change number of patients, consequently, the algorithm is named as variable length non-dominated sorting genetic algorithm.

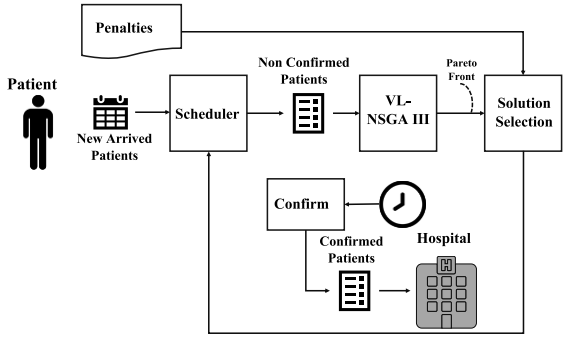


Figure 2: Simulation of dynamic patients scheduling using multi-objective optimization.

The following assumptions are inherent in the simulation model for the dynamic scheduling of patients in a hospital environment:

1. Hospital Structure: The hospital is composed of a set of departments D each specializing in various fields S, and containing a set of rooms R.
2. Room Capacity: Each room within a department can accommodate a pre-defined number of patients simultaneously.
3. Patient Arrival Rate: There is a specific rate at which patients arrive at the hospital, and each patient requires a certain number of nights within a preferred department and specialization.
4. Service Duration: Each patient is to be served within a specified number of nights.
5. Dynamic Solution Space: The optimization problem is dynamic, with the solution space changing in length due to the varying number of patients on different days, affecting the allocation of patients from the non-confirmed list.
6. Time-Dependent Optimization Outcome: The outcome of the optimization process at time t is denoted as PF_t indicating a time-dependent Pareto Front.

1.3 General algorithm

The algorithm of the scheduling combines the optimization with additional steps in order to enable dynamic scheduling. Firstly, there is a pre-processing step with the goal of preparing prior calculation of the various soft-constraints values. This enables shorter execution time of the optimization throughout the time interval of scheduling. Secondly, the new arrived patients are entered to queue according to their arriving time and the queue has a certain length so when the queue if full again the optimization is conducted and the new patients are located and the non-confirmed patients are allowed to be re-located. Thirdly, an algorithm for selecting one solution from the pareto front is enabled after running the optimization. This algorithm uses a weighted average formula of the soft-constraint according to a penalty entered from the user. Fourthly, the solution is activated and patients from the queue are decomposed into two sets: the first one is the confirmed patients and the second one is the non-confirmed patients. The difference between the two sets is that the confirmed patients are the patients that are scheduled with three days from the current date while the non-confirmed patients are the patients that are scheduled later than three days as long as their scheduling does not exceed the permitted period. A pseudocode of the general algorithm is given in Table 4.

Table 4: Pseudocode of the general scheduling algorithm using queue, multi-objective optimization and solution selection algorithm.

Input:
- w: Weights of the soft-constraints penalties
- Q: Queue used for storing new patients before re-running the MOO optimization
- timeInterval: Time interval for scheduling
- It: Number of iterations for the MOO optimization
- popSize: Size of the population in the optimization
- Rooms: Room matrix with information about supported departments, specialisms, and capacities

Output:
- schDecision: Scheduling decision, assigning each patient to a room

Start:
1: Pre-calculate soft-constraints using preProcessing (Rooms, w)
2: For each time interval in timeInterval
3: While Q is not full
4: Add new patient to Q
5: End while
6: Run MOO optimization using Optimization (popSize, It)
7: Select solution using selectSolution (paretoFront, w, soft-constraints)
8: Divide patients into confirmed and non-confirmed using assignFrom (Q, solution)
9: Remove confirmed patients from Q and add them to schDecision
10: Add non-confirmed patients to Q
11: End for
12: Return schDecision
End

1.4 Pre-processing

The goal of the pre-processing is to execute pre-calculation of the possible values of soft-constraints penalties in advanced according to all possible values of violations. As an example, For the gender constraint violation, assuming that we have n patients inside a room, it is possible to have mixed gender violation. This violation takes certain value if the majority are female and different value if the majority are males. Another example is the violation of the room capacity constraint, which takes different value according to the number of patients that exceed the room capacity. Assuming that the set of patients is denoted as $P = \{p_i\}, i = 1, \dots n$ and the set of rooms is denoted as $R = \{r_j\}, j = 1, \dots m$ where $n \gg m$. However, the patients arrive based on an arrival rate λ. Instead of calculating the soft-constraint based on the patient using function $f(p_i, r_j)$, we map the patient to a class or category according to his gender, needs or preference $C_p(p_i)$, and the room to a class or category according to its occupied patients, department and supported specialism $C_r(r_i)$ and we apply pre-calculated function for providing the soft-constraint or violation $f(C_p(p_i), C_r(r_i))$. Considering that the number of values of $C_p(p_i)$ and $C_r(r_i)$is limited then the generating the of

the corresponding soft-constraint is more efficient by using $f(C_p(p_i), C_r(r_i))$ instead of $f(p_i, r_j)$.

1.5 Initialization algorithm

The initialization algorithm is in charge of creating the primary arrangement interior the window, which signifies the number of days which will handle a specific number of unused understanding candidates. S_pre, which stands for the arrangement decided based on the past window, and Information, which stands for the information that comprises numerous sorts of data, essentially a list of rooms, an overhauled list of patients, and the fittingness of the patients for the rooms, are the inputs for this strategy. The arrangement after optimization based on the current window and upgraded persistent list is demonstrated by the yield, S current. The strategy cycles through the List-new-patients and begins a variable called Room with the esteem of -1, showing that a appropriate room has not however been found for this quiet. A deferred persistent or a patient who wasn't deferred is the persistent in address. Within the previous situation, it decides whether or not the room from the earlier arrangement is suitable by checking it. The quiet is put in this room since it is appropriate and open. Something else, in case there are any open rooms, a irregular room is chosen for this persistent. The understanding is designated to his room from a earlier arrangement or at arbitrary within the occasion that no open rooms are accessible, and it receives a delay, giving the hail delay a esteem of 1.

Table 5: The generation of the initial solution.

Input
- S_{pre} // *previous solution*
- $Data$ //*includes rooms and patients and Room-Patient-Suitability*
- W // *the current window of performing the new optimization*

Output
- $S_{current}$ // *initial solution for current window*

Start Algorithm
1: for patient in List of patients from solution
2: Room ← -1 //*initialization*
3: if the patient is delayed (not new)
4: if initial room still has space **AND** this room is suited for this patient
5: $Room \leftarrow previous\ day\ solution\ room$
6: end if
7: end if
8: while not (Room is suited and has space) and there is more Rooms
9: $Room \leftarrow random\ (Rooms)$
10: end while
11: if Room not equal to -1
12: Assign patient to Room.
13: Set his delay value to zero.
14: else // *the case the room is still* -1
15: assign patient to Room // *if it's delayed, we can use a not suited room.*
16: set his delay value to one.
17: end if
18: end for
End Algorithm

1.6 Crossover

Crossover's function is to create a new generation from an existing one, which promotes exploitation, while mutation's function is to tweak an existing solution in some way, which promotes exploration. In genetics, both crossover and mutation exist. The algorithm for the crossover is shown in Table 6. The input consists of the entire population and IN, which denotes the proportion of the population where crossover is carried out. The elites, who stand for the generation's best answers, are typically subject to the crossover.

The population after crossover is the output. The algorithm chooses two random crossover solutions for each crossover iteration and creates a random fraction of patients to shift their rooms and assign them to DeltaRooms from each crossover solution. Additionally, it creates a random sample of patients and sends them to DeltaDelay in order to adjust their delay. Then it makes the necessary changes to the initially chosen two parents and includes the off-springs in the new generation.

Table 6: The crossover operation for the genetic design.

Input:
- current generation,

Output:
- new generation

Start Algorithm
1: Choose a random portion of the generation to apply crossover to.
2: for counter IN portion size
3: Choose two parents x,y from the current generation
4: DeltaRooms ← random portion of patients to change their rooms from solution x to solution y.
5: DeltaDelay ← random portion of patients to change their delay from solution x to solution y.
6: Child 1=change (x, y, DeltaRooms, DeltaDelay)
7: Child 2=change (y, x, DeltaRooms, DeltaDelay)
8: Add child 1 and child 2 to new the generation
end for.
End Algorithm

1.7 Mutation

For the mutation, the pseudocode is presented in Table 7. The input of the algorithm is the individual or solution that will be selected for mutation, the mutation rate which indicates to how many patients in the Individual receptivity to change and acceptance rate ap determine whether or not we adopt the dominating solution following mutation. This step is taken to make it possible to avoid local minima.

After mutation, the output is altered individually. As can be seen from the pseudocode, the algorithm chooses at random either the type 1 or type 2 neighborhood type before performing the mutation on the chosen person. The algorithm then verifies domination and accepts the solution if it is the dominant one. It accepts non-dominance with a probability known as the acceptance rate. The objective is to make the algorithm more explorable.

Table 7: The mutation operation for the genetic design.

Input:
- Solution
- Mutation rate: how many patients in the individual to change.
- ap: acceptance rate

Output:
- new Solution with mutated individuals

Start Algorithm
1: select random neighborhood
2: $new - Solution \leftarrow neighborhood$ (Solution, Mutation rate)
3: If new- Solution Dominates the current Solution

4: $current\ Solution \leftarrow new - Solution$
5: Else
6: Generate a probability to allow bad Solutions
7: if $generated\ probability > ap$
8: current Solution ← new- Solution
9: End for
End Algorithm

Neighborhood 1 or Neighborhood 2—shown in Tables 8 and 9 respectively—are the bases for the neighborhood operation. While neighborhood 2 focuses on changing the delay of random patients randomly, neighborhood 1 focuses on changing the location or room of random patients at random. In order to provide the searching method more latitude, both of them must be employed in the mutation.

In Table 8, the mutation rate and the current solution.

Table 8: Pseudocode of neighborhood 1 operator used in the mutation.

Input:
- Mutation rate
- Current Solution
Output:
- new Solution after the change
Start Algorithm
1: While Mutation rate
2: patient ← random (current Solution patients)
3: new-room ← random (current Solution rooms)
4: if the new-room is suited for this patient
5: set the patients room to the new-room.
6: end if
7: end while
End Algorithm

Table 9: Pseudocode of neighborhood 2 operator used in the mutation.

Input:
- Mutation rate
- Current Solution
- Window
Output:
- new Solution after the change
Start Algorithm
1: while Mutation rate
2: patient ← random (current individual patients)
3: new-delay ← random (1 ← 0)
4: if the new-delay + day is in the patients staying range
5: set the patients delay to the new-delay.
6: end if
7: end while
End Algorithm

1.8 Solution sorting

For sorting solutions, we use domination operators. The only domination operator is non-dominated sorting which has the role of sorting the solutions into ranks, the first rank includes the non-dominated solutions over the entire population. The second rank includes the solutions that are dominated by the first rank and dominating other ranks and so on. The algorithm is divided into a main. The algorithm of solutions ranking is tasked with orchestrating the entire sorting process, where fronts are initialized, and each solution in the population is systematically evaluated and ranked. The algorithm commences by initializing separate fronts, each intended to group solutions of equivalent non-domination levels. The core of the algorithm involves a thorough evaluation of each solution in the population to determine its dominance relationships. Solutions are compared pairwise, leading to the identification of those dominated by and dominating each solution. The first front is populated with solutions that are not dominated by any other, representing the optimal trade-offs. Subsequent fronts are iteratively constructed, where each front consists of solutions only dominated by those in the preceding front. This iterative process continues until all solutions are assigned to a rank, effectively segregating the population into distinct layers of non-dominated sets. The outcome is a hierarchically structured set of solutions, providing a clear perspective on their relative quality and guiding the selection process in the evolutionary algorithm.

Table 10: Pseudocode of solutions ranking

Inputs:
- **Population P: A set of N solutions.**

Outputs:
- **Ranked Fronts**: Sets of solutions sorted into different ranks based on non-domination.

Start Algorithm
1. **Initialize Fronts:**
   Create empty lists for each front (Front 1, Front 2, ...).
2. **Evaluate and Rank Each Solution:**
   for each solution **p** in **Population P**:
   Initialize **dominatedByP** (list of solutions dominated by **p**) as an empty list.
   Initialize **dominatesP** (count of solutions that dominate **p**) as zero.
   for each solution **q** in **Population P**:
   if **p** dominates **q**, add **q** to **dominatedByP**.
   if **q** dominates **p**, increment **dominatesP**.
   if **dominatesP** is zero (i.e., **p** is not dominated by any other solution):
   Assign **p** to Front 1.
3. **Construct Subsequent Fronts:**
   Initialize **Current Front** as Front 1.
   while **Current Front** is not empty:
   Initialize **Next Front** as an empty list.
   for each solution **p** in **Current Front**:
   for each solution **q** in **dominatedByP** of **p**:
   Decrement **dominatesP** counter for **q**.
   if **dominatesP** for **q** becomes zero:
   Assign **q** to **Next Front**.
   Replace **Current Front** with **Next Front**.
4. **Return the Ranked Fronts:**
   The fronts are ranked such that Front 1 contains solutions not dominated by any other, and each subsequent front contains solutions only dominated by those in the previous front.

End Algorithm

procedure and two sub-procedures, each fulfilling distinct roles

1.9 Selection of solution

The result of the optimization when it is applied is a Pareto front which represents set of non-dominated solutions. Thus, we need an algorithm that selects solution out of the Pareto front for enabling it in the scheduling. Assuming that the weights of the soft-constraints or the objectives are represented by a vector $w = [w_1\ w_2 \dots w_m]$ where $w_1 + w_2 \dots + w_m = 1$. The solutions will be ranked based on linear production between the weights and the values of the objective function. In other words, each solution x_i from the pareto front will be mapped to one cost value based on the Equation (3)

$$x = argmin\ (f_1, f_2, \dots f_n) \qquad (3)$$

$$f(x_i) = w{y_i}^T$$

where

$$w = [w_1\ w_2\ \dots w_m]$$

$$y = \begin{bmatrix} y_{i,1} \\ y_{i,2} \\ \\ \\ y_{i,m} \end{bmatrix}$$

After that, the solutions are sorted in an ascending manner according to the cost values or $f(x_i)$ and the first solution or the solution that has the least cost value is selected and enabled. The result of enabling the solutions is two set of patients: the first one is confirmed set S_{conf} and it includes patients that are scheduled within three days and the second one is the non-confirmed set $S_{non-conf}$ and it includes patients that are scheduled later than three days. For S_{conf}, we remove them from the queue so they will not be used again for re-scheduling while for $S_{non-conf}$ we keep them in the queue so they are allowed for rescheduling in the next execution of the algorithm.

1.10 Variable length optimization of Window Based NSGA-III

In order to distinguish between patients that are allowed for rescheduling from new arrived patients, we use variable length optimization (VLO). In VLO, different lengths of solutions are used where each solution allow for rescheduling of different sub-sets of the non-confirmed patients. The goal of this is to conduct optimization with giving more importance to rescheduling of later scheduled patients and less importance of earlier scheduling patients.
The optimization in this case, will generate different number of solutions according to the number of patients where the solutions that contains earlier scheduled patients are less than the solutions of later scheduled patients. We call this algorithm variable length NSGA-III or VL-NSGA-III.

1.11 Evaluation metrics

The evaluation metrics that were employed to assess our created strategy are provided in this subsection. It has broken down.

- **Set coverage:**

This metric compares the Pareto sets P_{s1} and P_{s2} as follows

$$c(P_{s1}, P_{s2}) = \frac{|\{y \in P_{s2} \mid \exists x \in P_{s1} : x > y\}|}{|P_{s2}|} \qquad (4)$$

C is equal to the number of solutions in Ps2 divided by the proportion of non-dominated solutions in Ps2 that are dominated by non-dominated solutions in P s1. Therefore, it is crucial to reduce the value of C (X, P s) for all pareto sets X while assessing a set Ps.

- **Hyper-Volume**

The HV-metric has been used widely in evolutionary multi- objective optimization to evaluate the performance of search algorithms. It computes the volume of the dominated portion of the objective space relative to a worst solution (reference point); this region is the union of the hypercube whose diagonal is the distance between the reference point and a solution x from the Pareto set PS. Higher values of this measure indicates to more desirable solutions. HV is given by the Equation (5).

$$HV = \text{volume}\left(\bigcup_{x \in P_s} \text{HyperCube}\ (x)\right) \qquad \textbf{(5)}$$

6 Experimental works And evaluation

The assessment is a simulator-based assessment. For this stage, we utilized the simulator's data, which covered a total of 36 days. The data has similar layout to the data provided in the work of [34]. We contrasted NSGA-3, which incorporates numerous objective optimizations based on our created operators, with the following benchmarks: particle swarm optimization (PSO), multi-objective particle swarm optimization (MOPSO), and objective decomposition particle swarm optimization (ODPSO). The set coverage, hyper-volume, and convergence curves were produced.

1.12 Set-Coverage

The results of the set-coverage reveal the superiority of NSGA-III over the benchmarks. More specifically, NSGA-III has accomplished full domination over PSO which is single optimization algorithm, full domination over both MOPSO and ODPSO which are multi-objective algorithms, and 0.66 domination over NSGA-II. On the other side, non-of the algorithms of ODPSO, MOPSO, and PSO were capable of dominating NSGA-III. However, NSGA-II has provided 0.96 percentage of domination over NSGA-II.

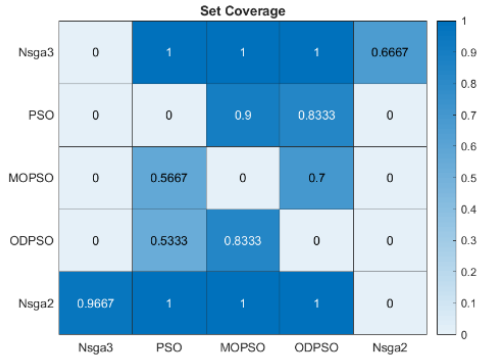


Figure 3: Set coverage of our developed WB approach and it is comparison with the benchmarks.

Hyper-volume

The results of the hyper-volume are presented in Figure 4. We find that the hyper-volume generated from NSAG-III and NSGA-II were the highest compared with the other approaches

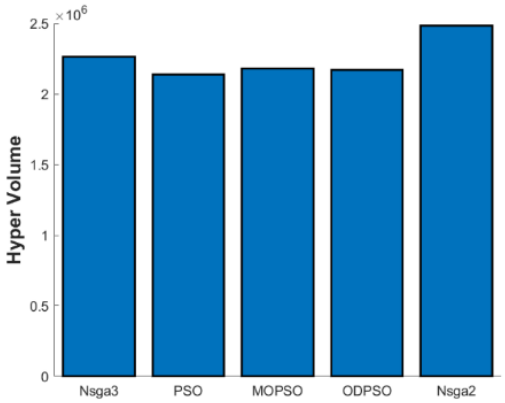
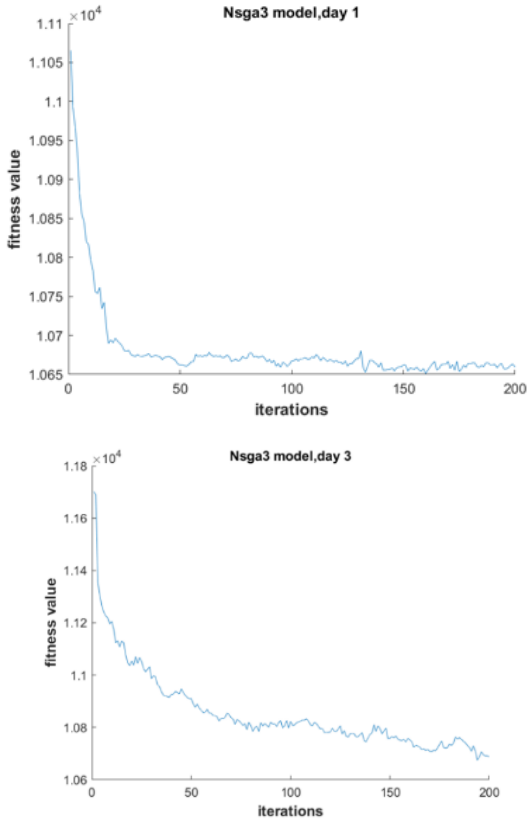


Figure 4: Hyper-volume of our developed algorithm and its comparison with the benchmarks.

1.13 Convergence curve

Considering that the optimization is reapplied in every day, the convergence curve is plotted to show the effectiveness of the optimization. The convergence curve is plotted based on fitness value equals to the average of the objectives. For plotting the convergence curve, we use calculate a fitness value as weighted average of the soft constraints based on the penalties of them. In Figure 5, we present the convergence of days 1, 2, 3 and the last day 36.



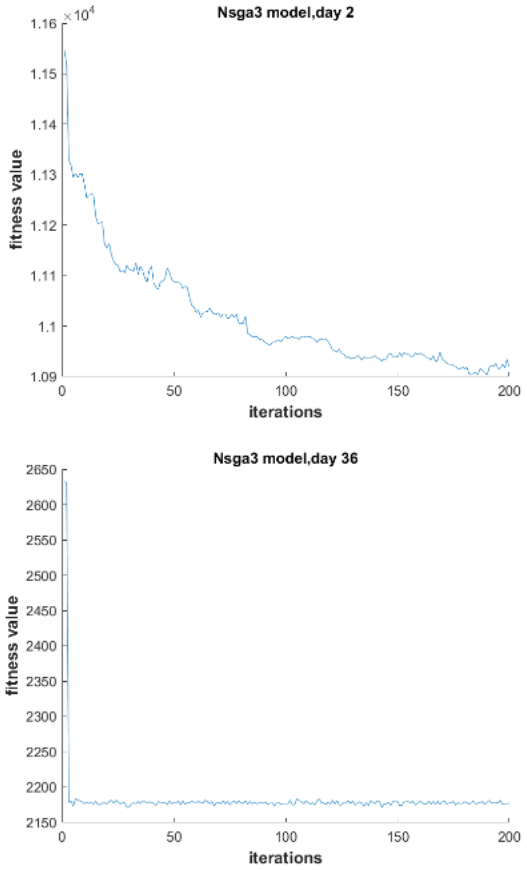
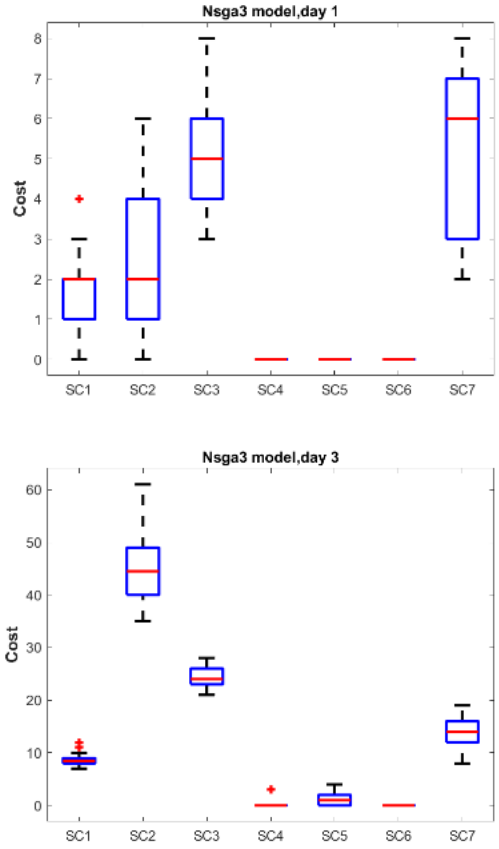


Figure 5: The convergence curve of NSGA-III of some of the optimization days.



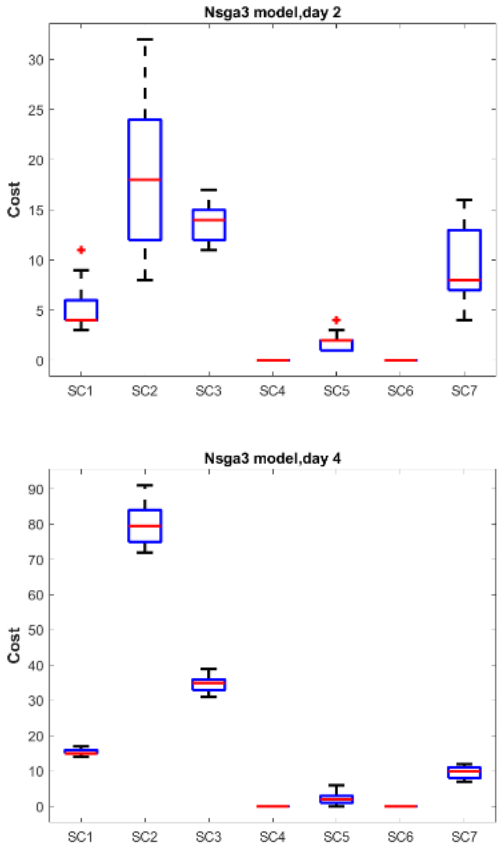


Figure 6:The boxplot of soft-constraints of NSGA-III of some of the optimization days.

1.14 Soft-constraints-values

In addition to the set-coverage, hyper-volume and convergence curve, we present the soft-constraints of each day Pareto front as boxplot diagram in Figure 6. The soft-constraints are encoded according to the symbols provided in Table 10.

Table 11: Coding for the soft-constraints used in the optimization.

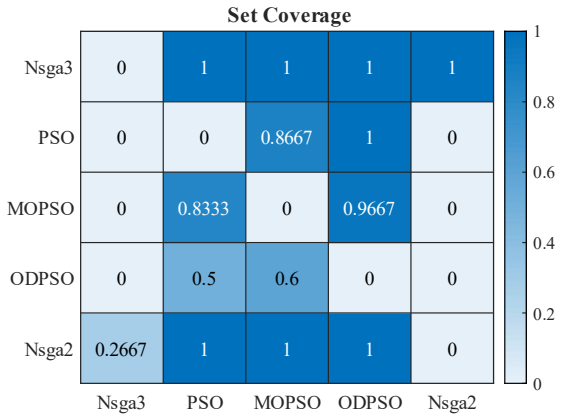
Code	Meaning
SC1	Missing Room Equipment
SC2	Unsatisfied Room Preference
SC3	Partial Specialty Level
SC4	Unsatisfied Gender Policy
SC5	Over -Crowd Risk
SC6	Delay
SC7	Transfer

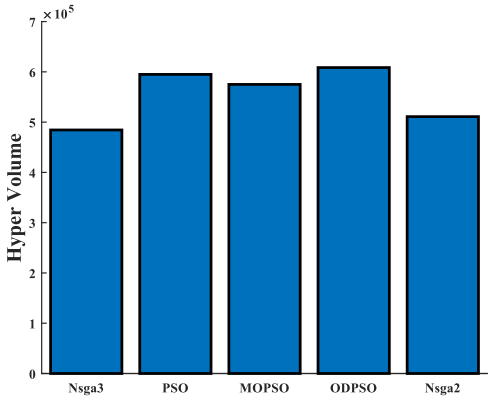
The visualization shows a similar performance between the various days in the relative relation between the soft-constraints with changing in the values obtained from one day to another.

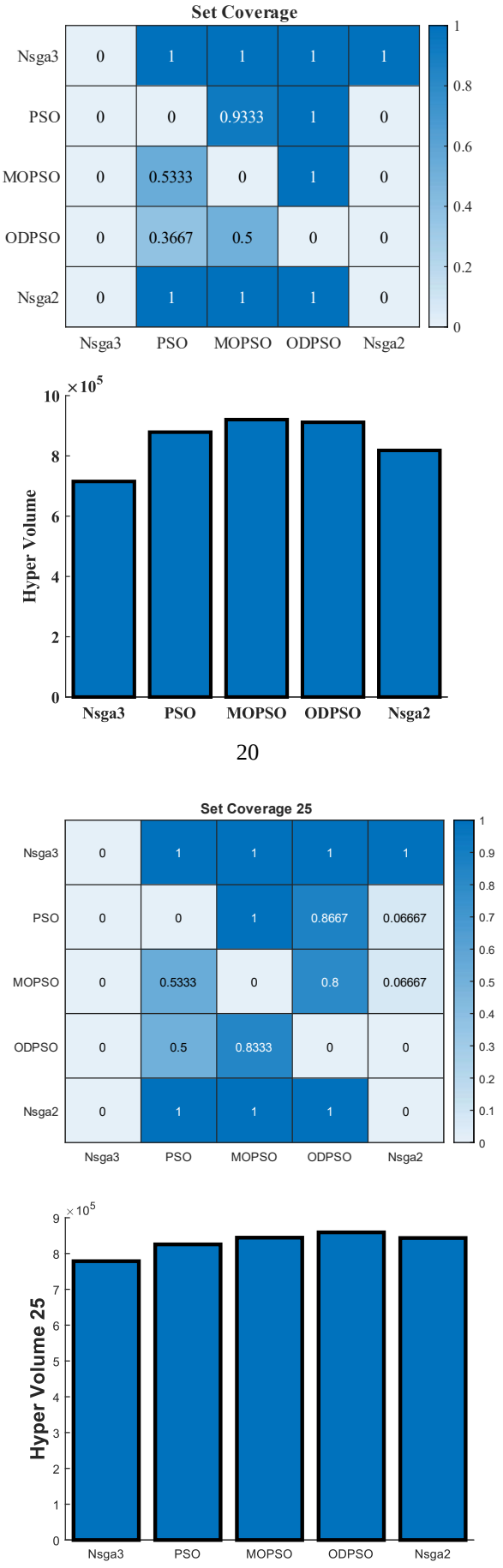
This is interpreted by the effect of the dynamic in the performance that changes from one day to another. However, associating this graph with the convergence graph given earlier shows that the algorithm was capable of handling the dynamics and brining the cost to a lower value.

1.15 Robustness evaluation scenarios

For evaluating our algorithm more comprehensively, we conducted a robustness evaluation by increasing the arrival rate of patients in the range of 15, 20, 25, and 30 patients per day. For each scenario, we generated the values of set coverage and hyper-volume. Observing the results of the set coverage as depicted in figure – confirms our finding of the superiority of of NSGA-III over other benchmarks. This is concluded from the domination of NSGA-III compared with the other optimization algorithms. It is found that a full domination was obtained when the arrival rate was 15. This is associated with high values of hyper-volume and competitive to other methods. Hence, it is found that increasing the arrival rates of patients has not only maintained the superiority but also the diversity of decision making.

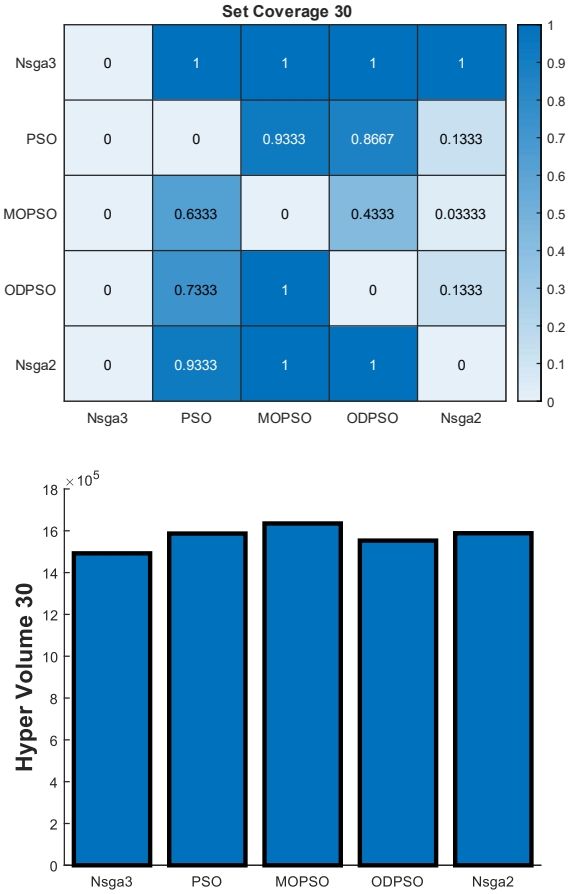






20

25



30

Figure 7:Set coverage and hyper-volume for different values of arrival rates ranging from 15 until 30

7 Conclusion and future work

Dynamic patient scheduling for hospital admission is challenging combinatorial problem with dynamical nature and many soft-constraints. An effective approach for solving it is using many-objective optimization MOO algorithms. However, direct application of them is not feasible due to the static nature of MOO algorithms. Hence, handling this application requires incorporation of other assisting blocks.

In this article, we have developed a novel simulator for dynamic scheduling of patients with window and coordinator. The role of the window is to accumulate both newly arrived patients and non-o patients.

The coordinator's duties include choosing a subset of patients from the window, placing them in the optimization block on one side, and choosing a non-dominated solution, activating it in the hospital on the other. A rigorous 36-day evaluation using PSO, ODPSO, MOPSO, NSGA-II, and NSGA-III has shown that NSGA-III is superior based on set-coverage and soft-constraints.

The practical implications of the findings from this proposed solution have been deemed to hold significant promise for enhancing the efficiency of hospitals and healthcare systems. Improved resource utilization, reduced patient wait times, and elevated overall care

quality could be achieved through the implementation of a dynamic scheduling system based on multi-objective optimization. Despite these benefits, challenges such as the integration with existing healthcare systems, staff training, and the need for robust data privacy and security measures have been identified as potential obstacles. Furthermore, the scalability and customization required for the system to be successfully adopted across various healthcare settings present additional complexities. A gradual, phased approach to implementation, involving pilot testing and stakeholder engagement, can be suggested to mitigate these challenges and facilitate smoother adoption.

Future research is to explore the adaptability of the methodology used in the healthcare scheduling system to other complex scheduling problems across different domains. The manufacturing sector, transportation and logistics, energy management, education, event management, and urban planning have been identified as areas where similar optimization techniques could be applied. Each domain presents its unique set of challenges and constraints, necessitating the customization of the optimization framework. The extension of this research into varied domains is expected to account for specific requirements and challenges while considering the effects on human behavior, regulatory standards, and economic considerations.

References

[1] I. Papanicolas, L. R. Woskie, and A. K. Jha, "Health care spending in the United States and other high-income countries," *Jama*, vol. 319, no. 10, pp. 1024-1039, 2018.

[2] N. Fares, R. S. Sherratt, and I. H. Elhajj, "Directing and orienting ICT healthcare solutions to address the needs of the aging population," in *Healthcare*, 2021, vol. 9, no. 2, p. 147: MDPI.

[3] J. Meehan, L. Menzies, and R. Michaelides, "The long shadow of public policy; Barriers to a value-based approach in healthcare procurement," *Journal of Purchasing Supply Management*, vol. 23, no. 4, pp. 229-241, 2017.

[4] R. Guido, V. Solina, and D. Conforti, "Offline patient admission scheduling problems," in *International Conference on Optimization and Decision Science*, 2017, pp. 129-137: Springer.

[5] A. N. Mahmed and M. Kahar, "Window-Based Multi-Objective Optimization for Dynamic Patient Scheduling with Problem-Specific Operators," *Computers*, vol. 11, no. 5, p. 63, 2022.

[6] C. Taramasco, B. Crawford, R. Soto, E. M. Cortés-Toro, and R. Olivares, "A new metaheuristic based on vapor-liquid equilibrium for solving a new patient bed assignment problem," *Expert Systems with Applications*, vol. 158, p. 113506, 2020.

[7] R. Guido, M. C. Groccia, and D. Conforti, "An efficient matheuristic for offline patient-to-bed assignment problems," *European Journal of Operational Research*, vol. 268, no. 2, pp. 486-503, 2018.

[8] K. Hussain, M. N. M. Salleh, S. Cheng, and Y. Shi, "Metaheuristic research: a comprehensive survey," *Artificial Intelligence Review*, vol. 52, no. 4, pp. 2191-2233, 2019.

[9] R. Alizadeh, J. Rezaeian, M. Abedi, and R. Chiong, "A modified genetic algorithm for non-emergency outpatient appointment scheduling with highly demanded medical services considering patient priorities," *Computers Industrial Engineering*, vol. 139, p. 106106, 2020.

[10] K. Dorgham, I. Nouaouri, H. Ben-Romdhane, and S. Krichen, "A hybrid simulated annealing approach for the patient bed assignment problem," *Procedia Computer Science*, vol. 159, pp. 408-417, 2019.

[11] A. Hammouri, "A modified biogeography-based optimization algorithm with guided bed selection mechanism for patient admission scheduling problems," *Journal of King Saud University-Computer Information Sciences*, 2020.

[12] J. Luo, Q. Liu, Y. Yang, X. Li, M.-r. Chen, and W. Cao, "An artificial bee colony algorithm for multi-objective optimisation," *Applied Soft Computing*, vol. 50, pp. 235-251, 2017.

[13] D. Wang, D. Tan, and L. Liu, "Particle swarm optimization algorithm: an overview," *Soft Computing*, vol. 22, no. 2, pp. 387-408, 2018.

[14] R. Tanabe and H. Ishibuchi, "An easy-to-use real-world multi-objective optimization problem suite," *Applied Soft Computing*, vol. 89, p. 106078, 2020.

[15] H. R. Maier, S. Razavi, Z. Kapelan, L. S. Matott, J. Kasprzyk, and B. A. Tolson, "Introductory overview: Optimization using evolutionary algorithms and other metaheuristics," *Environmental modelling*
software, vol. 114, pp. 195-213, 2019.

[16] P. Demeester, W. Souffriau, P. De Causmaecker, and G. V. Berghe, "A hybrid tabu search algorithm for automatically assigning patients to beds," *Artificial Intelligence in Medicine*, vol. 48, no. 1, pp. 61-70, 2010.

[17] A. M. Turhan and B. Bilgen, "Mixed integer programming based heuristics for the Patient Admission Scheduling problem," *Computers Operations Research*, vol. 80, pp. 38-49, 2017.

[18] B. Bilgin, P. Demeester, M. Misir, W. Vancroonenburg, and G. V. Berghe, "One hyper-heuristic approach to two timetabling problems in health care," *Journal of Heuristics*, vol. 18, no. 3, pp. 401-434, 2012.

[19] S. Kifah and S. Abdullah, "An adaptive non-linear great deluge algorithm for the patient-admission problem," *Information Sciences*, vol. 295, pp. 573-585, 2015.

[20] Y.-H. Zhu, T. A. Toffolo, W. Vancroonenburg, and G. V. Berghe, "Compatibility of short and long term objectives for dynamic patient admission scheduling," *Computers Operations Research*, vol. 104, pp. 98-112, 2019.

[21] M. Rezaeiahari and M. T. Khasawneh, "Simulation optimization approach for patient scheduling at destination medical centers," *Expert Systems with Applications*, vol. 140, p. 112881, 2020.
[22] A. K. Abera, M. M. O'Reilly, M. Fackrell, B. R. Holland, and M. Heydar, "On the decision support model for the patient admission scheduling problem with random arrivals and departures: A solution approach," *Stochastic Models*, vol. 36, no. 2, pp. 312-336, 2020.
[23] S. Ceschia and A. Schaerf, "Dynamic patient admission scheduling with operating room constraints, flexible horizons, and patient delays," *Journal of Scheduling*, vol. 19, pp. 377-389, 2016.
[24] S. Ceschia and A. Schaerf, "Dynamic patient admission scheduling with operating room constraints, flexible horizons, and patient delays," *Journal of Scheduling*, vol. 19, no. 4, pp. 377-389, 2016.
[25] B. Bilgin, P. Demeester, M. Misir, W. Vancroonenburg, and G. V. Berghe, "One hyper-heuristic approach to two timetabling problems in health care," *Journal of Heuristics*, vol. 18, no. 3, pp. 401-434, 2012.
[26] S. Kifah and S. Abdullah, "An adaptive non-linear great deluge algorithm for the patient-admission problem," *Information Sciences*, vol. 295, pp. 573-585, 2015.
[27] Y.-H. Zhu, T. A. Toffolo, W. Vancroonenburg, and G. V. Berghe, "Compatibility of short and long term objectives for dynamic patient admission scheduling," *Computers Operations Research for Health Care*, vol. 104, pp. 98-112, 2019.
[28] M. Rezaeiahari and M. T. Khasawneh, "Simulation optimization approach for patient scheduling at destination medical centers," *Expert Systems with Applications*, vol. 140, p. 112881, 2020.
[29] B. Tang, Z. Zhu, H.-S. Shin, A. Tsourdos, and J. Luo, "A framework for multi-objective optimisation based on a new self-adaptive particle swarm optimisation algorithm," *Information Sciences*, vol. 420, pp. 364-385, 2017.
[30] C. Seren, "A hybrid jumping particle swarm optimization method for high dimensional unconstrained discrete problems," in *2011 IEEE Congress of Evolutionary Computation (CEC)*, 2011, pp. 1649-1656: IEEE.
[31] Q. Lu, X. Zhu, D. Wei, K. Bai, J. Gao, and R. Zhang, "Multi-phase and integrated multi-objective cyclic operating room scheduling based on an improved NSGA-II approach," *Symmetry*, vol. 11, no. 5, p. 599, 2019.
[32] A. Arram and M. Ayob, "A novel multi-parent order crossover in genetic algorithm for combinatorial optimization problems," *Computers Industrial Engineering*, vol. 133, pp. 267-274, 2019.
[33] S. Ceschia and A. Schaerf, "Local search and lower bounds for the patient admission scheduling problem," *Computers Operations Research for Health Care*, vol. 38, no. 10, pp. 1452-1463, 2011.
[34] S. Ceschia and A. Schaerf, "Modeling and solving the dynamic patient admission scheduling problem under uncertainty," *Artificial intelligence in medicine*, vol. 56, no. 3, pp. 199-205, 2012.

Multimedia VR Image Improvement and Simulation Analysis Based on Visual VR Restructuring Algorithm

Xiangyang Xu
Henan Police College, Zhengzhou, Henan, 450046, China
E-mail: xxy@hnp.edu.cn

Keywords: visual VR reconstruction algorithm, image optimization, virtual reality technology, median filter

Received: October 25, 2023

Due to the advancement of science and technology, the application of virtual reality (VR) technology is more and more extensive, and people can truly immerse themselves in the virtual space through virtual reality. Relying on the visual VR reconstruction algorithm, this paper deals with the problems of "burring" and insufficient compression of relatively simple video imaging devices. Using virtual reality as a foundation, the multimedia effect of the video image is processed, and according to six operation modules, a system combining virtual reality technology is designed. From the aspect of determining the relationship between video image data and color, it is classified into three types: binary image, pseudo-color image, and grayscale image, and the grid of each point is defined and quantified. The extreme value filtering algorithm is used to perform a sorting calculation on the image pixels in the filtering window so as to improve the image effect with the threshold value suitable for filtering processing. Simulation results show that the VR visual restoration algorithm has a higher compression ratio and higher optical efficiency and can effectively support multimedia VR image improvement and simulation analysis.

Povzetek: Študija se ukvarja z analizo multimedijskih VR slik z uporabo algoritma za vizualno rekonstrukcijo VR, ki naslavlja težave z zamegljenostjo in nezadostno kompresijo slik, povečuje kompresijsko razmerje in optično učinkovitost za podporo izboljšanju VR slik.

1 Introduction

Under the influence of rapid social and economic development, people's pursuit of virtual space is getting higher and higher [1-4]. On the one hand, the immersion type makes users be in a virtual environment; that is, the physical environment is no longer felt but integrated into a new virtual environment; on the other hand, the interactive type enables users to control the surrounding environment in real-time, making people interactive with the surrounding virtual environment and be creative. Therefore, in the environment of virtual reality, it is first necessary to solve the construction of the virtual environment, that is, how to construct it. At the same time, the perfection of the virtual 3D space is the first sign of the user's experience [5-12].

The related technology of virtual reality is based on the human-centered perspective, viewing the video from 360° to the surroundings, which is no longer limited by time and space, so that people can fully experience the real virtual space. Different from traditional video images, the resolution of virtual reality video images is higher, and the bandwidth resources and storage resources occupied are higher. Therefore, its transmission, storage, anti-interference, and other capabilities are poor, and complex digital processing cannot be conducted [13-15]. How to effectively process multimedia VR images, this paper relies on the visual VR reconstruction algorithm; facing the problem of processing the related effects of virtual and real video graphics images, we can use filtering to calculate the threshold, which can effectively remove the noise in the video image, to explore the improvement and simulation analysis of multimedia VR images.

2 Related works

Table 1: Literature survey

Reference	Key Findings	Methodologies	Outcomes
[16]	The paper examined the main findings of designing interactive VR classrooms, with a specific focus on categorizing educational activities.	They investigated the incorporation of deep learning algorithms and utilized a quantitative regression analysis methodology.	They demonstrated the impact of assessing teaching quality on improving the learning experience.

[17]	The study examined significant discoveries concerning occlusion in hand posture estimation, emphasizing the use of the Skeleton-Difference Loss Function and the Object-Manipulating Loss Function.	The research utilized approaches that specifically targeted the training of deep learning models.	The experimental results indicated the flexibility and exceptional efficiency of the suggested system across many circumstances.
[18]	The article examined the advancement of visual effects in landscaping graphics, emphasizing the application of deep belief networks as classifiers.	The article employed a three-fold cross-validation methodology, a deep belief network learning process, a wavelet deep belief network model, and a weighted k-nearest neighbor algorithm.	The results indicated an improvement in recognition accuracy and classification effectiveness, suggesting possible applications in garden image recognition technologies.
[19]	The paper presented significant findings on the development of virtual interactive models, with a focus on improving the user experience.	The article utilized approaches focused on the integration of technology to research cultural heritage.	The research revealed advancements in user experience, cultural investigation, and conservation.
[20]	The paper included deep learning integration, multidimensional assistance, and the broad commercial use of a particular system or technology.	The paper examined the approaches associated with the use of supporting technologies, with a specific emphasis on the impact of deep learning.	The paper examined the results of widespread commercial applications, the growing popularity of VR products, and the incorporation of deep learning theory.
[21]	The paper examined significant discoveries on the broader uses of VR in architecture design, with a particular focus on advanced rendering techniques.	The research examined techniques that relate to Immersive Rendering and Deep Learning Training, explicitly emphasizing the Camera Velocity Rendering Method.	The study investigated the practicality and efficacy of enhanced animation routes in mitigating VR sickness, demonstrating improved results in the domain of VR.
[22]	The integration of VR technology for complete quality improvement was presented.	The methods for building digitization and VR technology applications in high-end construction projects were examined in this study.	The consequence of the article focused on "Enhanced Designer-User Interaction" and offered recommendations for the advancement of the sector.
[23]	The study examined significant discoveries concerning the emotional influence of video games, highlighted the advancements in design approaches, and discussed the challenges encountered in VR.	The paper described the techniques used to develop emotionally intelligent virtual avatars, with a specific emphasis on the implementation of emotional avatars from Bernardo Agents.	The article demonstrated enhanced narrative perception, a favorable influence on presence, and versatility in its application to different virtual worlds.
[24]	The research presented significant discoveries about image transformation technology, specifically focusing on the construction and equalization of Grey Level Histograms.	The study examined approaches connected with modeling and 3D Technology in VR, with a specific emphasis on Image Transformation Technology.	The paper's result involved addressing interface difficulties and improving visual effects that were in line with human features.
[25	The study examined essential findings regarding the incorporation and	The paper explored techniques for reducing	The research exhibited improved detection performance achieved by

	interaction of multimedia data, with a specific focus on resolving interference problems in multimedia networks.	interference in multimedia networks by utilizing compressed coding and decoding technology.	effectively implementing compression technology, supported by positive testing outcomes.

The need for standardized tools and workflows for creating high-quality VR content, along with difficulties in capturing and producing content that maximizes VR capabilities, contribute to challenges such as low resolution leading to pixelation. To address these issues, we propose the implementation of the Visual VR Restructuring Algorithm.

Dataset

Everyday Objects in Context database [26]: The COCO Dataset, an extensive image collection, consists of over 20,000 carefully annotated images across 81 distinct image types. This dataset was utilized for both training as well as assessment purposes. This step can be regarded as a top-down data training stage, in which user-labeled information is used for supervised salient recognition of objects. Furthermore, our issue is due to the need for more knowledge regarding the specific item or object class. Therefore, we must establish a general conspicuous object by relying on global properties. Figure 1 depicts the Dataset sample images.

Figure 1: Dataset sample images

3 Research methods

3.1 Virtual reality technology

VR technology is composed of modules such as feedback, detection, sensors, controllers, modeling, etc. The specific composition is depicted in Figure 2.

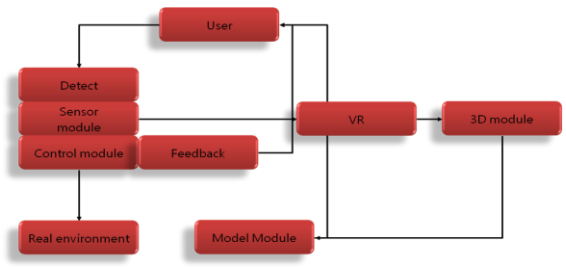


Figure 2: System composition of virtual reality technology

Among these six modules, which are entirely different but related, the sensor module is linked to the user by the detection and feedback modules. It communicates with the 3D module via the control module.

3.2 Visual VR reconstruction algorithm

3D Stereo Matching Correlated with Phase: If two images with a size of m in $E(m)$ and $H(m)$sequences have the same time, the calculation of the discrete Fourier function is shown in formula (1) and formula (2):

$$E(l) = \sum_{m=-N}^{N} e(m) X_M^{lm} = B_E(l) f^{i\theta E(l)}, \quad (1)$$

$$H(l) = \sum_{m=-N}^{N} h(m) X_M^{lm} = B_H(l) f^{i\theta H(l)}. \quad (2)$$

where $m = -N, \ldots, N,$, $M = 2N + 1, X_M = f^{-\frac{l2m}{M}}, B_E(l)\&(l)$ are the "image" amplitude data of the description; the two values $f^{i\theta E(l)}$ & $f^{i\theta E(l)}$ denote the phase areas of the u, then it can be obtained that the calculation of the image phase difference after normalization processing is as shown in the formula (3):

$$\hat{Q}(l) = \frac{E(l)\overline{H(l)}}{|E(l)H(l)|} = f^{i\theta(l)} \quad (3)$$

Among them is a conjugate complicated number descri$H(l)\ \theta(l) = \theta_e(l) - \theta_H\theta(l)$ption. The inverse transformation of the discrete Fourier function can be obtained by Q, as shown in formula (4):

$$\hat{q}(m) = \frac{1}{M}\sum_{m=-N}^{N} \hat{Q}(l) X_M^{-lm} \quad (4)$$

Phase correlation visual VR reconstruction under the averaging method: When images of low quality are used, the accuracy of the corresponding 3D stereo adaptation method is biased [7-9]. At this time, the averaging method fun is used to improve the accuracy of binocular vision

adaptation, utilizing the particular procedure illustrated in Figure 3.

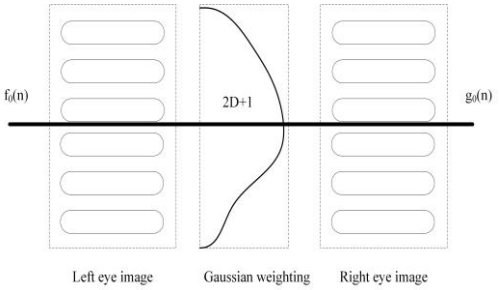


Figure 3: Visual VR 3D matching under phase correlation

As described in Figure 3 above, the 3D VR matching of visual images based on phase correlation, we can conclude that the human eyes are symmetrical. The lateral configuration can be applied to the optical system. The obtained sequence of sample images f (m) of the numerical value is shifted by L distance units along the y-axis to get $f_e(m)$. Subsequently, a line $h_e(m)$ corresponding to this value is calculated, and a specific phase-related numerical processing is performed on $f_e(m)$ and $h_e(m) r_{e(m)}$ obtained so that the final mean sequence value is shown in formula (5):

$$\hat{q}_{ave}(m) = \frac{1}{2C+1} \sum_{k=-C}^{K=C} \hat{q}_K(m) \quad (5)$$

3D Matching Structure of CTF: e_s represents the binocular image, displayed on the idea of the first level, in formula (6), it corresponds to the average value of the pixels illustrated in the four relatively close areas of the previous level e_{s-1}, as shown in formula (6):

$$e_{s-1}(w.z) = \frac{1}{4} \sum_{j=0}^{1} \sum_{i=0}^{1} e_{s-1}(2w + j, 2z + i) \quad (6)$$

After the denoising method, J_w^i the grayscale arrangement of the image in the visual VR remodeling is made. To obtain the dispersion of the characteristics of the surrounding area of the target individual, w=p,e, and the grayscale transformation formula is shown in Equation (7), Equation (8):

$$T_d = [T_0, T_1, \dots, T_{R-1}]_{binary} = \left[\sum_j^{R-1} T_j \times 3^j\right]_{dce} \quad (7)$$

$$T_j = \sum_i^{X \times X} J_w^i \quad (8)$$

It can be seen that W is the transformation step size, Y is the gray value of the area near the target range, and the denoising function is constructed as shown in formula (9, 10):

$$w(l+1) = T_j(l)_w + x_j(l),\ j = 1,2,\dots,n, \quad (9)$$

$$y(l) = G_j(l)_w(l) + u_j(l),\ j = 1,2,\dots,n \quad (10)$$

The above formula (9) $u_j(l)x_j(l)$ represents the pixel noise value of different target individuals in the range area, and the mean value of $G_j(l)_w$ and $T_j(l)_w$ is 0, and the variance represents the state that this value $T_{j\,(l)}$ is uniformly distributed on the image and constitutes a fuzzy set $v = \{v_{jk}\}$. After that, the noise-removed output is obtained, and the texture characteristics are analyzed, as shown in formula (9, 10):

$$J_{GSM} = J\left(D^M; C^M | t^M\right) = \sum_{j=1}^{M} J\left(D_j; C_j | t_j\right) = \sum_{j=1}^{M} \left(g\left(C_j | t_j\right) - g\left(C_j | D_j t_j\right)\right) = \sum_{j=1}^{M} \left(g\left(h_j D_j + U_j | t_j\right) - g\left(U_j\right)\right) \quad (11)$$

Considering the difference between the reconstructed image ranges, the pixel grayscale range is formed in the direction of the gradient to obtain the alternative formula for image denoising by describing the spatial texture characteristics $v^{(m)}(w, z, c)$, as shown in formula (12) and formula (13):

$$v^{(m+1)}(w, z) = v^{(m)}(w, z) + \delta v_1^{(m)}(w, z), \quad (12)$$

$$v_1^{(m)}(w, z) = N\Delta_w v^{(m)}(w, z) + M\Delta_s v^{(m)}(w, z, c) \quad (13)$$

Then, formula (13) is applied to the “original image for high-pass filtering." After processing, the texture of the “image” is enhanced again, as shown in formula (14):

$$d(w, z) = \sum_{j=-m}^{m} \sum_{i=-n}^{n} g(j, i)\, h(w - j, z - i) \quad (14)$$

Among them, d is the high-frequency range selected by the feature, h denotes the image reconstructed, and g represents a 3×3 rectangular matrix based on a high-pass filter, as shown in Equation (15):

$$h = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \quad (15)$$

After that, an image e_1 with enhanced texture quality can be obtained according to equation (16):

$$e_1(w,z) = e(w,z) + d(w,z) \quad (16)$$

Where(w,z) denotes the coordinates for the direction of the image, e_1 denotes the pattern image after details are enhanced, e denotes the original image, and d denotes the high-frequency range selected.

3.3 Modification of multimedia vision vr reconstruction algorithm

(1) Algorithm peak search

In order to reduce the maximum peak range, recalculate the link representing the strengthening of low-pass filtering V_1, V_2 and calculate a peak value, as shown in formula (17):

$$g(l_1,l_2) = \begin{cases} 1, & |l_1| \le V_1, |l_2| \le V_2 \\ 0, & otherwise \end{cases} \quad (17)$$

(2) Error Identification and peak relocation

The peak α in the random construction level is collected and sorted into a set. If the edge of the limited range of level l is the threshold α_{th} , then $r_k(n_k)$ it can be calculated. To confirm that this point is a cluster point, the phase correlation peak α must be ≥ the threshold α_{th} , as shown in equations (18) and (19):

$$r_{kd}(n_k) = r_k(n_k), \quad (18)$$

$$c_{kd}(n_k) = n_k - r_k(n_k) \quad (19)$$

In order to confirm that this point is an outlier, $\alpha < \alpha_{th}$ it is assumed that the arrangement is arranged according to the size of the middle value taken. All issues between the 5×5 area in the near range excluding this point are selected, and the c value of this point is defined; see Formulas (20) and (21) are shown as:

$$c_l'(n_k) = \left(c_1^{med}, c_2^{med}\right), \quad (20)$$

$$q_k'(n_k) = n_k - c_k'(n_k) \quad (21)$$

Secondly, a new peak value is obtained by using the phase correlation calculation, assuming $\alpha > \alpha_{th}$, as shown in formula (22) and formula (23):

$$r_{kd}(n_k) = r_k'(n_k) \quad (22)$$

$$c_{kd}(n_k) = n_k - c_k'(n_k) \quad (23)$$

On the contrary, $c_{kd}(n_k) = c_k'(n_k)$it is made

(3) Algorithm correction

Let the relative reconstruction point be(w_r, z_r), the sample direction has the coordinates (x_l, y_l) , the calculated peak value is α, and the visual aberration of the two eyes is$c = y_1 - y_r$.

If $C_{min} < C < C_{max}$ and $\alpha \ge \alpha_{th}$ are entered into the stereo model, where $C_{max}$and $C_{min}$are the maximum and minimum parallax values in sequence?

3.4 Video image digitization

In this design, video image digitization mainly uses the quantization method. And make corresponding assumptions, use equidistant sampling to obtain a nearly coherent image$e(w,z)$, and set it as a rectangular array of N*M so the following formula (24) can be obtained:

$$e(w,z) = \begin{bmatrix} e(0,0) & e(1,0) & \cdots & e(0,N-1) \\ e(1,0) & e(1,1) & \cdots & e(1,N-1) \\ \cdots & \cdots & \cdots & \cdots \\ e(M-1,0) & e(M-1,1) & \cdots & e(M-1,N-1) \end{bmatrix} \quad (24)$$

Every element is an independent discrete variable; the right side of formula (4) shows a video image of a number, and each element in the data set is described as a corresponding pixel [10].

In the actual calculation process, in order to make Z and r an array of real numbers and integers, during the collection of samples, they are converted into network format as the flattened data of the image. The grid of each node is calculated and finally determined according to the Cartesian coordinate system.

During the actual calculation process, Y and R are transformed into arrays of real numbers and integers. This conversion occurs while collecting samples, where they are represented in network format as the flattened data of the image. Subsequently, the grid of each node is computed and ultimately determined based on the Cartesian coordinate system

During the whole process of digitally converting the idea, first, we determine the N and M dimensions of the image and the distinct grayscale values. H in pixels. When calculated, these values are usually rounded to an integer power of 2. Therefore, the expression tested in the snapshot can be expressed by Equation (25):

$$\begin{cases} N = 2^n, \\ M = 2^m, \\ H = 2^l \end{cases} \quad (25)$$

Suppose the value range of the discrete grayscale is set to be between 0 and 10, and the distribution is in a uniform state. In that case, the bits required to store the digital video image can be retrieved by the following formula, as shown in formula (26):

$$b = M \times N \times K \quad (26)$$

If M=N, then it can be shown as formula (27):

$$b = N^2 k \quad (27)$$

3.5 Video image processing effect optimization

Analysis of one-way multi-stage median filter algorithm: The median filter should replace the median value of the area between two adjacent points with a point in the numerical interval and finally calculate the median value using the following method [11-12]:

Let $w_1, w_2.., w_n$ be a set of m values, which are sorted according to their size, and get the following formula (28):

$$w_{j1} \leq w_{j2} \leq w_{j3} \leq \cdots \leq w_m \quad (28)$$

Among them, when the odd number is m, the following formula (29) is obtained:

$$z = Med\ (w_1, w_2, w_3, \ldots, w_n) = w_j \frac{(m+1)}{2} \quad (29)$$

When the even number is m, see the following formula (30):

$$z = Me\ (w_1, w_2, w_3, \ldots, w_m) = \frac{1}{2}\left[w_j \frac{(m+1)}{2} + w_j \frac{(m)}{3} + 1\right] \quad (30)$$

From the calculation results of the above formulas, it can be found that the result of the effect of the image is similar to the simple root mean square in the results obtained from the 3*3 window, and y represents the median value of the sequence.

During the process of this calculation, let w (n, m) be a complex video image in the original noise and represented as a grayscale point value of (n, m) in the coordinates. First, select the rectangle where the window L=2m+1 is located and divide the window into four independent windows, in which integer m is positive. The process is described as follows, see calculation equations (31)-(34):

$$X_1(n,m) = \{w(n, m+j), -M \leq j \leq M\} \quad (31)$$

$$X_2(n,m) = \{w(n+j, m), -M \leq j \leq M\} \quad (32)$$

$$X_3(n,m) = \{w(n+j, m-j), -M \leq j \leq M\} \quad (33)$$

$$X_4(n,m) = \{w(n+j, m-j), -M \leq j \leq M\} \quad (34)$$

According to the calculation, the schematic diagram of the MLM filter in Figure 4 can be obtained:

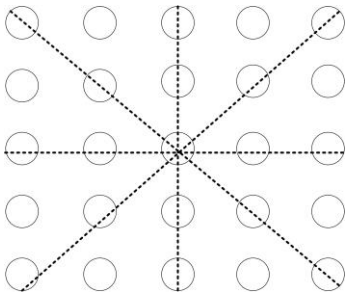


Figure 4: MLM filter

A set of twisted one-dimensional image windows along the horizontal or vertical direction is represented by X_1, X_2, X_3, X_4 respectively, and then $Y_1(n,m)$,$Y_2(n,m)Y_3(n,m)$ and $Y_4(n,m)$denote the median of X_1, X_2, X_3, X_4 (4 windows), as shown in equations (35)-(38):

$$Y_1(n,m) = med\ [w(j,i) \in X_1(n,m)], \quad (35)$$

$$Y_2(n,m) = med\ [w(j,i) \in X_2(n,m)], \quad (36)$$

$$Y_3(n,m) = med\ [w(j,i) \in X_3(n,m)], \quad (37)$$

$$Y_4(n,m) = med\ [w(j,I) \in X_4(n,m)] \quad (38)$$

Where $V_{min}(n,m)V_{max}(n,m)$and denote the max and min of$Y_1(n,m)Y_2(n,m)Y_3(n,m)$, and $Y_4(n,m)$ (mean). Thus, equations (39) and (40) can be obtained:

$$V_{max}(n,m) = \max[Y_1(n,m), Y_2(n,m), Y_3(n,m), Y_4(n,m)], \quad (39)$$

$$V_{min}(n,m) = \min[Y_1(n,m), Y_2(n,m), Y_3(n,m), Y_4(n,m)], \quad (40)$$

Based on the above formula, the derivation method of the multi-stage median filter of the single term is shown in formula (41):

$$z(n,m) = med\,[V_{min}(n,m), V_{max}(n,m), w(n,m)] \quad (41)$$

Image optimization by extreme median filtering based on thresholds: In this paper, points are introduced to establish median extremum filtering and enhance the arrangement of image pixels in the window used for filtering. The image range is pre-determined and segmented into the fine details of the image edges, noise influence, and flat range optimization [13]:

First, arrange the pixel points $W[x_{j,i}]$ in the window to find the $\min\,(W[x_{j,i}])$ point and the $\max(W[x_{j,i}])$ point; this point $\max(W[x_{j,i}])$represents the maximum value point and $\min\,(W[x_{j,i}])$refers to the minimum value point. After that, compare this point $x_{j,i}$ with $\max(W[x_{j,i}])\ \min\,(W[x_{j,i}])$ and. The results suggest that if the two points are entirely different, no filtering will be performed on the original value; on the contrary, if the two points get the same value state, the program can be started using a pre-judged calculation method.

If f (w, z) is the grayscale of point (w, z) in the image, and h (w, z) is that of the pixel in (w, z) (adjacent range), operator Z is selected and applied to e (w, z) and h (w, z). Z=Z (e, h) can be obtained. Then, continue with the next step according to the different Z. Here the way of Z is shown in formula (42):

$$Z = \sum_{j=0}^{s} z\left(e(w,z) - h_j(w,z)\right) \quad (42)$$

The following formula (43) can be obtained:

$$Z(w) = \begin{cases} 1, & |w| \le S, \\ 0, & |w| > S \end{cases} \quad (43)$$

After calculation, the value of i in the formula is shown in Figure 5.

In Figure 5, we can see that the point f (w+1, z) denotes point 0 adjacent to end (w, z), and these values circle the point (w, z) in turn until the point f (w+1, z-1) becomes the seventh point.

5	4	3
6	(w,z)	5
7	0	1

Figure 5: i value distribution.

According to the above description, it can be obtained that the threshold value T represents a constant threshold value. If the visual effect of the image is in a good state, it will not be affected by significant noise pollution, and the distribution change has not become large, then in the calculation, the minimum value of T value should be taken in the process. Otherwise, it may cause errors because the threshold selection result is not accurate [14-15].
Suppose the final selected result is too high. In that case, the noise will be misunderstood as a helpful signal point during image processing, and most of the noise is retained during processing, reducing filtering efficiency and visual effects. Instead, this can happen if the threshold (f) is too low. A helpful signal point would be seen as contamination noise, which would make the image even more blurry, and signal noise would make the visual effect much lower.
According to the above calculations, it can be judged that the following conclusions can be further obtained:

1) If the gray value in pixels is infinitely equal to or close to it, i.e., y is equal to 0, then this point can be regarded as an isolated point of the median noise filter.

2) When 1≤y≤4, that is, the grayscale values of 1 to 4 pixels are equal to or very close, this point is regarded as the peripheral detail node of the part and is not processed.

3) When the Y value is no less than 4, i.e., more significant than the grayscale of four pixels, and is equal to or very close to this value, it can be considered that the point is in a flat area, and the issue has not been processed. The whole operation process is shown in formula (44):

$$y_{ij} = \begin{cases} med\left(W\left[x_{ij}\right]\right), Y = 0 \\ x_{ij}, Y > 0 \end{cases} \quad (44)$$

4 Experiment result texture feature analysis noise ratio

If the gray level co-occurrence matrix is defined by a value $\rho(d,\varphi)$ as the probability that the gray value is set to j from a point whose gray value is set to i, the possibility that the gray value at a point leaving a relatively unchanged position is set to j. In this position, d is selected to be equal to 1, and the importance of φ is set to 0°, 45°, 90°, and 135°, which are representative angles. In this way, a contrast ratio and a calculation of the entropy value can be established, as shown in Equation (45):

$$\begin{cases} I(d,\varphi)=\sum_{k} k^2\left[\sum_{i}\sum_{j}\rho(i,j|d,\varphi)\right] \\ H(d,\varphi)=-\sum_{i}\sum_{j}\rho(i,j|d,\varphi)\log\sum_{i}\sum_{j}\rho(i,j|d,\varphi) \end{cases} \tag{45}$$

Based on the results of the calculation of k=I-J, in the process of optimizing the entire video image, the value of contrast is a significant factor used to measure the texture grooves of the image. The deeper the groove texture, the correspondingly greater the contrast between the images. Entropy can measure image information. If there are many fine textures, the value of entropy becomes larger.

For the evaluation after video image optimization processing, the evaluation index represents the contrast to the Noise Ratio (CNR) that may be contained in the image, as shown in formula (46):

$$CNR=\frac{2(\mu_t-\mu_b)^2}{\sigma_t^2+\sigma_b^2} \tag{46}$$

In formula (45), the noise target area value in the image is represented by μ_t and μ_b represents a mean value described as the background area. The noise target value and the standard deviation of the background are represented by these two data σ_t σ_b, respectively. From the value of the calculation result, if the CNR value obtained above is higher, it means that the image has yet to reach the optimal effect.

4.1 Experimental result

The recommended task is executed on CUDA 9.0, Python 3.6, and Tensor Flow 1.9.0, Python software, and is required to be installed alongside Python to carry out the procedure.

In this study, the field-programmable gate array (FPGA), digital signal processing (DSP), and simulation system-based video image optimization [27] are compared with image processing techniques. The simulation results are shown in Figure 6. The proposed algorithm in this study has the best filtering effect, as shown in Figure 6 because it can determine the most appropriate threshold value. Other algorithms are capable of optimizing, but they are unable to produce outcomes that meet expectations. It could be because calculating the image enhancement standard value is challenging.

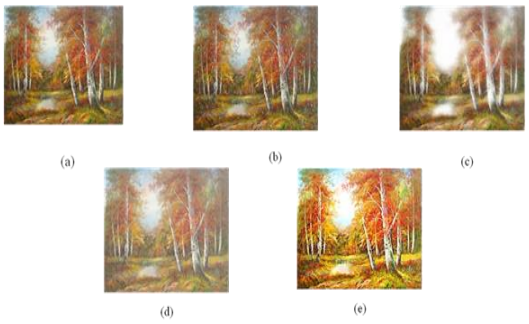


Figure 6 : An analysis of the impact of several techniques on image optimization: (a) source image; (b) simulator; (c) FPGA; (d) DSP; (d) the proposed algorithm

The results of comparing the proposed algorithm's performance are shown in Figure 7. Compared to other techniques, the images processed using the algorithm described here have lower contrast, entropy, and noise. These outcomes demonstrate the effectiveness of this strategy in adjusting contrast.

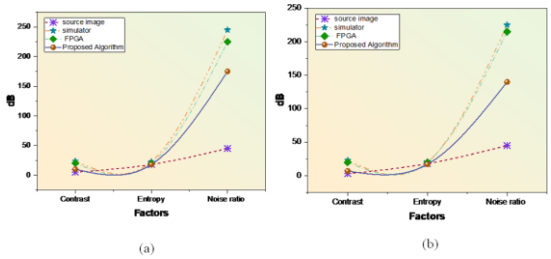


Figure 7: Results of comparing algorithm performance: (a) initial test outcome; (b) second-round test result.

The presented VR-based simulation system for video image processing optimization is compared to DSP-based and FPGA-based techniques in high-brightness images. The simulation results are shown in Figure 8.

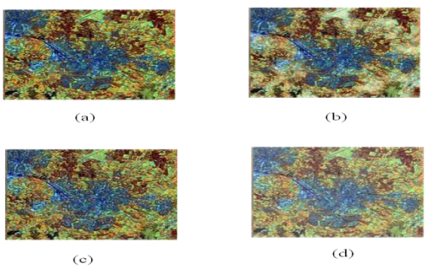


Figure 8: Evaluation of enhanced processing for brighter images: (a) source image; (b) simulator; (c) FPGA; (d) the proposed algorithm

Figure 9 depicts the Comparative algorithm performance outcomes. The proposed algorithm presented in this paper has lower contrast, entropy, and noise values compared to previous approaches. This suggests that it is capable of effectively adjusting contrast parameters, preserving image details, and reducing glare intensity. The proposed algorithm developed in this paper outperforms existing methods in terms of image quality.

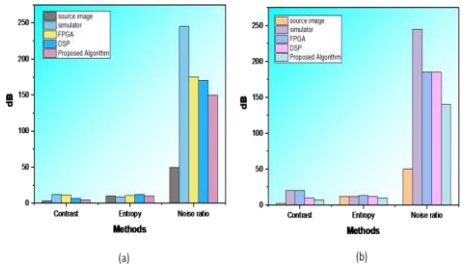


Figure 9: Comparative algorithm performance outcomes: (a) initial test outcome; (b) second-round test result.

The loss metric is used to quantify the prediction error of a model with the goal of reducing the difference between predicted and actual values. Figure 10 depicts the outcome of loss.

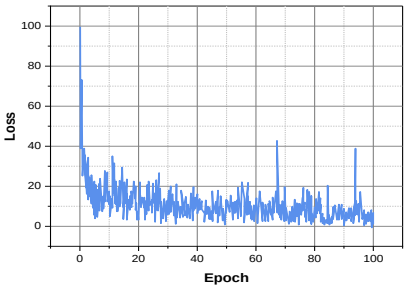


Figure 10: Outcome of loss.

Discussion

Field-Programmable Gate Array (FPGA) algorithms and Digital Signal Processing (DSP) algorithms face limitations in terms of resources, including logic cells, memory, and interconnects, making scalability and complexity challenging. When implementing a proposed visual restoration algorithm, achieving a balance between computational efficiency and preserving image quality poses challenges.

5 Conclusion

The study introduced a VR reconstruction algorithm for increasing the resolution of VR content to provide a clearer and more detailed visual experience. We gathered a COCO dataset for training models to detect and segment objects in VR environments. This can be useful for applications like virtual object manipulation or scene understanding—the proposed method results in an image with the best visual effect. Throughout the future rounds, the discriminant and generator will mutually enhance their learning process, resulting in improved quality of samples and resolution, as well as enhanced VR image improvement, often involving high-resolution content and complex data. Streaming such content in real-time may face challenges related to data transmission speeds and bandwidth limitations, affecting the overall user experience. In future research, improvements may focus on enabling collaborative VR experiences, allowing multiple users to interact seamlessly within a shared virtual environment, enhancing the social aspect of multimedia VR.

Data availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of interest

The authors declare no conflicts of interest

Funding statement

This study did not receive any funding in any form.

References

[1] Bannas P, Li Y, Motosugi U, et al. Prior Image Constrained Compressed Sensing Metal Artifact Reduction (PICCS-MAR): 2D and 3D Image Quality Improvement with Hip Prostheses at CT Colonography[J]. European Radiology, 2016, 26(7):2039-2046.

[2] Garg M, Naik T R, Pathak C S, et al. Significant improvement in the electrical characteristics of Schottky barrier diodes on molecularly modified Gallium Nitride surfaces[J]. Applied Physics Letters, 2018, 112(16):163-173.

[3] Vincenzo P, Brodie J P, Terry B, et al. A SLUGGS and Gemini/GMOS combined study of the elliptical galaxy M60: wide-field photometry and kinematics of the globular cluster system[J]. Monthly Notices of the Royal Astronomical Society, 2015, 450(2):12-20.

[4] Vrchota P , Prachar A , Smid M . Improvement of Computational Results of NASA Airliner Model

by Wing Modal Analysis[J]. Journal of Aircraft, 2017, 54(4):1-9.

[5] Hak, Gu, Kim, et al. VRSA Net: VR Sickness Assessment Considering Exceptional Motion for 360° VR Video[J]. IEEE Transactions on Image Processing, 2019, 28(4):1646-1660.

[6] Iyer V R , Sheedy S P , Gunderson T M , et al. Procedure-Related Pain During Image-Guided Percutaneous Biopsies: A Retrospective Study of Prevalence and Predictive Factors[J]. American Journal of Roentgenology, 2019, 213(4):1-7.

[7] Chen H H , Singh V R , Luo Y . Speckle-based volume holographic microscopy for optically sectioned multi-plane fluorescent imaging[J]. Optics Express, 2015, 23(6):7075-7084.

[8] Ncetan K , Celik I O , Obeid A , et al. VR-Caps: A Virtual Environment for Capsule Endoscopy[J]. Medical Image Analysis, 2021, 70(7):101-110.

[9] Evelyn S , Apolo S F , Ignacio V R , et al. Image-Guided BrachyAblation (IGBA) en hepatocarcinoma. Descripción de la técnica y reporte del primer caso en Chile[J]. Revista medica de Chile, 2019, 147(6):808-812.

[10] Gevaert O , Mitchell L A , Achrol A S , et al. Errata: Glioblastoma multiforme: Exploratory radiogenomic analysis by using quantitative image features (Radiology (2014) 273, 1, (168-174) DOI: 10.1148/radiol.14131731)[J]. Radiology, 2015, 276(1):56-63.

[11] Ross A S , Bruno M J , Kozarek R A , et al. Novel single-use duodenoscope compared with 3 models of reusable duodenoscopes for ERCP: a randomized bench-model comparison[J]. Gastrointestinal Endoscopy, 2019, 91(2):521-530.

[12] Aguilera V R , K Apaza Ⅲ Edoya, Pereira B , et al. Clinical study with short implants ‐ Relation among insertion torque, osseointegration and bone loss[J]. Clinical Oral Implants Research, 2019, 30(19):457-464.

[13] Thies J , Zollhfer M , Stamminger M , et al. FaceVR: Real-Time Gaze-Aware Facial Reenactment in Virtual Reality[J]. ACM Transactions on Graphics, 2018, 37(2):1-15.

[14] Kozlov A A , Abdullaev S D , Flid V R , et al. Algorithm and criterion of quality for assessing the packing of polymer microspheres[J]. Russian Journal of Physical Chemistry A, 2016, 90(9):1835-1838.

[15] Zhen, Wang, Qian, et al. The image variations in mastoid segment of facial nerve and sinus tympani in congenital aural atresia by HRCT and 3D VR CT[J]. International Journal of Pediatric Otorhinolaryngology, 2015,3(1):59-67

[16] Chen, W., Liu, X., Qiao, L., Wang, J. and Zhao, Y., 2020. Construction of virtual reality-interactive classroom based on deep learning algorithm. Wireless Communications and Mobile Computing, 2020, pp.1-9.

[17] Wu, M.Y., Ting, P.W., Tang, Y.H., Chou, E.T. and Fu, L.C., 2020. Hand pose estimation in object-interaction based on deep learning for virtual reality applications. Journal of Visual Communication and Image Representation, 70, p.102802.

[18] Zhang, H. and Min, X., 2020. Optimization and simulation of garden image visual effect based on Particle Swarm and wavelet threshold. IEEE Access, 8, pp.154390-154403.

[19] Liu, J. and Chen, Y., 2021, March. Research on scene fusion and interaction method based on virtual reality technology. In Journal of Physics: Conference Series (Vol. 1827, No. 1, p. 012010). IOP Publishing.

[20] Lin, Q., 2020. Application and development of virtual reality technology in artificial intelligence deep learning. In IOP Conference Series: Materials Science and Engineering (Vol. 740, No. 1, p. 012151). IOP Publishing.

[21] T. Fukuda, M. Novak, H. Fujii, Y. Pencreach, and L. C. Fu, “Virtual reality rendering methods for training deep learning, analysing landscapes, and preventing virtual reality sickness,” International Journal of Architectural Computing, vol. 1, no. 2, Article ID 147807712095754, 2020.

[22] Zhu, Z. and Du, Y., 2021. Research on interior design optimization based on virtual reality technology. In Journal of Physics: Conference Series (Vol. 1746, No. 1, p. 012063). IOP Publishing.

[23] Geslin, E., Bartheye, O.O., Schmidt, C., Tcha-Tokey, K., Kulsuwan, T., Keziz, S. and Belouin, T., 2020. Bernardo autonomous emotional agents increase perception of VR stimuli. Network and Communication Technologies;, 5(1), pp.11-25.

[24] Li, L., 2021. Visual information enhancement method of multimedia human-computer interaction interface based on virtual reality technology. International Journal of Information and Communication Technology, 19(2), pp.127-142.

[25] Yang, S., 2021, February. Intelligent Improvement Measures for the Broadcasting and Hosting Major based on Multimedia and Virtual Reality. In 2021

Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV) (pp. 447-451). IEEE.

[26] Mao, W., 2022. Video analysis of intelligent teaching based on machine learning and virtual reality technology. *Neural Computing and Applications*, *34*(9), pp.6603-6614.

[27] Cui, L., Zhang, Z., Wang, J. and Meng, Z., 2022. Film Effect Optimization by Deep Learning and Virtual Reality Technology in New Media Environment. Computational Intelligence and Neuroscience, 2022.

Internet of Things – A Model for Data Analytics of KPI Platform in Continuous Process Industry

Jeeva Jose[1], Vijo Mathew[2]
[1]Department of Computer Applications BPC College, Piravom, Kerala, India
[2]Department of Strategies and IoT AIDEAS Engineering Limited, Banglore, India
E-mail: vijojeeva@yahoo.co.in, vijo.mathew@aideasengineering.com

Keywords: internet of things, data analytics, KPI, continuous process, DCS, PLC, SCADA, cement, industrial IoT

Received: November 15, 2021

Internet of Things (IoT) is gaining momentum now a days to real time operational environment. The related technologies of IoT is converging to the main stream of industrial applications and replacing the conventional models of data acquisition, analysis, visualization and control in continuous manufacturing process industries. In this paper, we are proposing an IoT based model platform for acquiring various data that is generated in a continuous process manufacturing plant. This includes data from mobile devices and ERP systems as well. This is analyzed using machine learning and artificial intelligence technologies which leads to visualization of Key Performance Indicators (KPIs). It can be displayed on plant level as well as head office level in static and mobile devices. Control instructions can also be given from static devices as well as from mobile devices. Along with proposed platform concept, a prototype is also developed for cement manufacturing plant which is a core engineering continuous process manufacturing industry. The general KPIs in cement plants are explained and the KPIs generated in visualizing devices by the prototype platform are also provided in this paper.

Povzetek: Članek predlaga model IoT platforme za analitiko ključnih kazalnikov uspešnosti (KPI) v industriji kontinuiranih procesov, ki vključuje integracijo podatkov iz mobilnih naprav in ERP sistemov, uporabo strojnega učenja in AI za vizualizacijo KPI-jev v proizvodnji cementa.

1 Introduction

In continuous process industry [1], raw material moves from the beginning of the process and advances through each production step before converting to a final product. Once the process is initiated, the parameters such as pressure, temperature, speed, humidity etc. need to be controlled within the limits. The sensors can collect the data, compare that with requirements and take corrective actions wherever required. Cement manufacturing is an example of continuous manufacturing process industry.

Professionals working in continuous process manufacturing plants are expected to monitor performance of various machines and process parameters continuously. This should also be controlled in real time basis. The man power required for this activity is very high. In addition to this, there are possibilities of human error while monitoring manually. Presently, most of the continuous manufacturing process plants are reasonably automated. Their operations are with Programmable Logic Controllers (PLC) [2] or Distributed Control Systems (DCS) [3] and monitoring can be done from the control room. A PLC [4], is a ruggedized computer used for industrial automation. These controllers can automate a specific process, machine function, or even an entire production line. DCS [5] is a computerized control system for a process or plant that consists of a large number of control loops, in which autonomous controllers are distributed throughout the system with a central operator supervisory control.

Even though some level of autonomous control operations system is implemented in some manufacturing facility, the human experts need to be physically deployed in all areas of operation. If data collection, analysis, display and control can be done without human intervention, it will ensure less error in operations and activities can be done in a faster pace. The service of professionals who are presently involved in data collection, processing, analyzing and controlling activities can be utilized in other important focus areas like development of process and control, that meets future product, customer and environmental requirements. Presently engineers and managers are having access to smart phones and have reliable Internet connectivity in most of the places where plants are located. If they can get process information on their mobile phone, the need to be present in the control room all the time can be avoided. This will improve the flexibility of these personnel and hence it will result in improving open thinking and productivity. A platform that can acquire data from DCS or PLC [6] in real time, with capability to analyze and visualize on static as well as mobile devices with alerts for manual interventions as needed, can support industry to meet this requirement. As the sensors, wireless connectivity, computing and visualizing capabilities are in the developed phase, an Internet of Things (IoT) [7] based platform will be the right choice for meeting this requirement. IoT refers to a system of interrelated, Internet-connected things that are able to collect and transfer data over a network without human

intervention. The things can be sensor, actuator or any equipment connected each other and to the Internet normally wireless and sometimes wired. The Industrial Internet of Things (IIoT) [8] refers to the extension and use of the IoT in industrial sectors and applications. This can be either connected to the Internet or work as an independent industrial network. An example for IIoT is the smart electrical grid which is interconnected with power generation, transmission and distribution with sensors, control system and actuators. IIoT needs to follow the components and communication standards required for that particular industry in which, it is implemented.

Platform [9] [10] is a digital hub which integrates the inputs from sensors, analyze the data and provides output for visualization or actions. In addition to automated sensor data, the inputs can be provided by manual intervention based on the policies and requirements. The development of IoT platform with capability of data acquisition, analysis and visualization in static and mobile devices will reduce human efforts, improve speed and will support for taking the right manual decisions when required. In an IoT enabled factory, there are many individual components like sensors, actuators etc. These may be interdependent components of a production line and will be aware of each other's activity in real time. So, the entire manufacturing process will become more efficient as well as much easy to monitor and manage with the platform. Data analytics [11] [12] [13] is the process of systematically applying statistical and/or logical techniques to describe and illustrate, condense, and evaluate data. In IIoT, the data collected by various sensors are processed, some process happens at the sensor end itself which is known as edge processing [14]. This is transferred to platform in which detailed analysis happen and the output is given for human visualization and/or for actuators to take actions. Many software tools such as, Python, R Programming, Hadoop etc. are used for analysis. For visualization software such as Tableau, Power BI etc. are used for Human Machine Interface (HMI) [15]. Predictive analytics capability on the platform will be able to predict possible breakdown scenarios well in advance and will help to take corrective actions.

2 Related work

PLC can be programmed for effective operation of the process with productivity, accuracy, precision and efficiency [16]. Before the introduction of PLC, the relay logic and contactor logics (RLC) were used [2] which include human intervention and resulted in errors. The introduction of microprocessors, microcontrollers, PLCs, Supervisory Control & Data Acquisition (SCADA) [17] [18] and DCS [19] have improved the control of manufacturing operations. These systems reduced human intervention and increased the flexibility in the process control. By automation, the working of a process or repetitive works can be done efficiently by proper controls within acceptable range. DCS made IoT implementation practically feasible. The communication from DCS to processor can be via Message Queuing Telemetry Transport (MQTT) protocol [20] [21]. For a robust system, the security enhancements should be compatible with MQTT Application Programming Interfaces (API) [22]. Open Platform Communications United Architecture (OPCUA) protocol [23] is another protocol which is getting wider acceptability in the industry. IIoT receives very large amount of data from sensors and other sources. IIoT search engines [24] are also presently available. Big data analytics can be used for analysis of these data. Predictive and prescriptive analytics [25] can be done by adding this to the operational processes. The sensor driven data analytics which is used for decision making will improve and optimize the process industry. An analytical platform [26] can support the collection, storage, processing and visualization of data. Such a platform will be able to connect to the existing plant environment and use the data gathered to build predictive functions to optimize the production process.

3 Background

Continuous manufacturing process industries like cement, steel, paper, sugar, petrochemicals, fertilizers etc. have a matured manufacturing process. In this industry, once capital equipment in the manufacturing facility is installed, it is expected to provide continuous service for next 30-40 years. Not much of the technical upgradations or changes are possible in this life span. During the earlier days, all the process in continuous manufacturing industry were sensed, measured and required changes were done manually. Later, mechanical automation for sensing temperature, pressure, volume and suitable automatic systems were introduced [27]. An example of this is automatic coal fire reduction when steam pressure reaches required value. With the wide use of electricity in industries, electro-mechanical sensing and automation systems were introduced. Electric switch cut-off with a thermostat when it reaches the preset heat is an example of this application. These systems were of unidirectional, which means that it does not have the capability to adjust the process, based on the feedback from output or other variable parameters. More over this control system hardware need to be custom developed as per the individual manufacturing plant or industry requirements. The introduction of PLC brought great flexibility by providing the option of using standard programmable controller irrespective of manufacturing plant or industry. The era of DCS brought a revolution by allowing standard computers to monitor and control manufacturing in process industries [28]. This helped to get real time data to the centralized control rooms and these control rooms can take remote actions by providing inputs to the actuators. Various technological improvements like change of wired sensor system to wireless, development of various industrial communication standards, high computational & storage capabilities, display options and control capabilities brought an IoT revolution to continuous process manufacturing industry.

3.1 State of the art

The new generation of sensors and actuators are small, energy efficient, accurate, reliable and identifiable electronically. The identification systems like beacons, Radio Frequency Identification (RFID) [29] [30], Near Field Communication (NFC) [31] [32] etc. helped for easy and accurate sensing. The development of industrial wireless communication standards as well as computation and control systems, initiated Industry 4.0, which is the digital factory concept. With the introduction of Industry 4.0 [33], manufacturing plants started real time sensing of data with sensors installed in various equipment as well as throughout the environment. This system has created an environment called Cyber Physical System (CPS) [34]. By connecting this system to Internet, IIoT came into existence. Presently IIoT is getting implemented in many industries with very less or controlled exposure to communication through Internet. Dependability and standardization are essential to the adoption of Wireless Sensor Networks (WSN) [35] in industrial applications. Communication standards such as ZigBee [36], Wireless HART [37], ISA100.11a [38] and WIA-PA [39] are well accepted presently. The development of technology for computing at the sensing point itself and transfer of data to central control room for supervisory and management analysis as per the required Key Performance Indicators (KPIs) [40] paved the way for the revolution of IIoT. Key Performance Indicator (KPI) is a quantifiable measure of performance over time for a specific objective. KPIs provide milestones to measure progress that help people across the organization to take right decisions. Most of the industry and organizations monitor and compare their performance based on the KPIs set up for that particular segment. KPIs are important for monitoring the performance and to identify opportunities for improvement of the industry. KPIs can be defined for individual equipment, sub processes as well as for the whole plant. Performances related to energy, raw material, final product, process control, operation, maintenance, etc. can be monitored by KPI. Benchmarking KPIs with similar equipment and plants is one method of setting industrial segment KPI standards. The outputs received as KPIs, are displayed at plant levels as well as at the head office. The KPIs from other plants also reach the head office for analysis at that level and comparison. The corrective and control instructions [41] can also be given from head office or plant level to supervisory or to the actuator level.

4 Problem identification

Covid-19 the pandemic, restricted employees and professionals in travelling to factories and offices as well as for conducting physical meetings. In this situation, information flow from continuous manufacturing plants to supervisory and management team became important for taking right decisions and running the operations smooth. The present infrastructure of PLC, DCS or IoT enabled manufacturing industries are having data visualization and process control facility available only in static devices located in plant control rooms or at offices. In this situation, to continue the manufacturing process seamlessly, there is a need of integrating mobile devices to the existing control system infrastructure for accessing the continuous process data and other operational information. The process control facility needs to be provided with authorized mobile devices and it should be capable of operating from anywhere in the world. To achieve this, the right connectivity methods matching present available infrastructure as well as ensuring security needs to be developed. The integration of existing IIoT to mobile devices meeting the security requirements is a challenge identified by continuous process manufacturing organizations.

5 Proposed solution

The solution that we propose to the identified problem is the development of industrial platform which can access data from wireless sensors, mobile devices, DCSs, PLCs, ERP and text files. In the proposed platform, data could be analyzed as per the KPI requirements. The machine learning and artificial intelligence algorithms [42] [43] need to be incorporated for taking autonomous regular or corrective actions. The platform can also provide predictive analysis outputs that can be utilized for advance actions. The analysis output, meeting the KPIs formats should be displayed in mobile devices as well as in static devices as per the requirement. It should also be able to provide control instructions from mobile devices.

5.1 Automation landscape

In a continuous process industry, the data is collected from sensors and actuators to take actions based on the inputs from PLC, Proportional Integral Derivative (PID) controller, DCS or Supervisory Control and Data Acquisition (SCADA). A PID controller is an instrument used in industrial control applications to regulate temperature, flow, pressure, speed and other process variables. PID controllers use a control loop feedback mechanism to control process variables and are the most accurate and stable controller. A SCADA [44] is an automation control system that is used in industries such as energy, oil and gas, water, power, and many more. This system can be a centralized one to monitor and control individual sites and all connected sites. Manufacturing Execution Systems (MES) are software solutions that ensure quality and efficiency. This is built into the manufacturing process and are proactively as well as systematically enforced. Enterprise Resource Planning (ERP) is a software system that utilizes a centralized database that contains all the necessary data in one location. Information Technology (IT) automation is the process of creating software and systems to replace repeatable processes and reduce manual intervention. With IT automation, software is used to take care of repeat instructions, process, or policies to save time and free up IT staff for some other strategic work. Operational technology involves hardware and software that detects or causes a change, through the direct monitoring and/or

control of industrial equipment, assets, process and events. Figure 1 shows the convergence zone of operation /automation and information technology. The operation / automation technology involves sensors, actuators, PLC, PID, personnel computers and SCADA. ERP and MES combines to form the information technology area. The proposed platform will be in the convergence zone. Various operational technology channels are explained in Table 1 and information technology channels are described in Table 2.

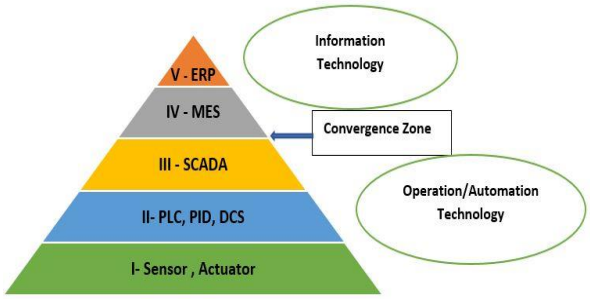


Figure 1: Convergence zone of operation/automation and information technology.

Table 1: Operational technology channels.

Operational Technology Channel	Description
OPC (Open Platform Communications)	Handles OPC connections using either OPC Unified Architecture (UA) specifications or OPC Data Access (DA) specifications. UA security is secured using certificates. DA security permissions can be applied using DCOM settings.
OPC Server	Acts as an OPC UA server. It can be accessed by a classic OPC client using a COM wrapper.
XML	Connects via a local or remote XML file.
CSV	Connects via a local or remote CSV file.
Webservice	Supports SOAP and REST communication and provides SOAP/REST host services. It runs as a server sending and receiving XML messages.
MQTT	Supports the ISO standard (ISO/IEC PRF 20922) protocol. ATS Bus supports encryption between the MQTT channel and the MQTT broker using X509 certificates.
RFID	Uses the Octane SDK to communicate with Impinj Speedway readers. The channel connects to the reader using a raw TCP/IP socket. These TCP/IP connections are not secured using certificates.
MTConnect	Supports communication with MTConnect agents that exchange information with CNC machines.
Socket	A bidirectional (client/server) TCP/IP communication channel. It can be used to process CSV, text or binary data. As a server the channel binds to a port. As a client the channel connects to a host name and port. It does not provide data encryption.
Serial Port	A bidirectional (client/server) RS-232 communication channel. It supports CSV, text and binary data payloads. COM ports can be virtual or physical.
Database	Communicates with Microsoft SQL Server and Oracle databases.

Table 2: Information technology channels.

Information Technology Channel	Description
XML	Connects via a local or remote XML file.
ActiveMQ	Connects via Apache ActiveMQ messaging service. Apache ActiveMQ is an open-source messaging and integrations patterns server. Encryption is not supported on this channel.
Webservice Server	Supports WCF and REST communication and provides WCF/REST host services. It runs as a server sending and receiving XML messages.
Webservice Client	Exchanges information with REST, SOAP and HTTP based web services.
Extension	Required when other IT channels don't have the functionality required to communicate with a customer's software. It read and write to a plug-in (.NET assembly) using a standard interface. It may or may not have secure communications depending on how it's used.

5.2 Line diagram

The line diagram of IoT based KPI platform for the continuous process manufacturing industry having multiple plant facilities is shown in Figure 2. The proposed platform will be installed in each plant as well as in head office. The data from each manufacturing plant will be transmitted to the plant level KPI platform from DCS through MQTT/OPC/Modbus channel. The data from the ERP will also be transferred similarly. Each plant will be connected to head office KPI platform through the Internet. Firewall will be placed at the point where each plant is connected to Internet as well as where the head office is connected to Internet.

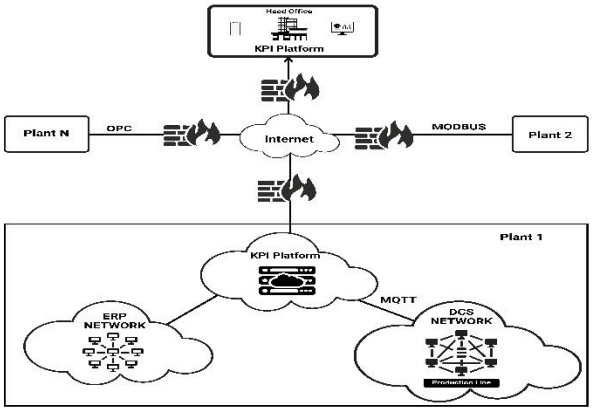


Figure 2: Line diagram of IoT based KPI platform.

5.3 Platform architecture

In the proposed IIoT platform, the operation/automation and the Information Technology will converge. Figure 3 shows the architecture of proposed KPI platform.

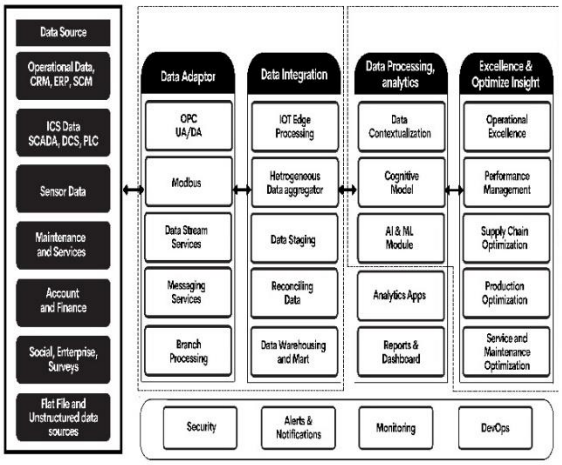


Figure 3: Architecture of KPI platform.

The proposed architecture has modules for acquiring inputs from various data sources. These sources can be sensor data, Industrial Control Systems (ICS), ERP, mobile applications etc. It can accept manual input data which comes as flat file as well as social media data which will be in the unstructured format. The data adaptor can be OPC, Modbus, MQTT etc. The data integration module integrates the data and will be made available for analysis. The artificial intelligence and machine learning applications are incorporated in data processing and analytics module. The output of this will be made available to dashboards. The security, monitoring, notifications, development, quality and operation modules will be common to all modules.

6 Implementation in cement manufacturing

Cement manufacturing [45] is highly automated continuous manufacturing process industry. The main stages of cement manufacturing are lime stone crushing, raw material handling, raw mill, kiln, coal mill and cement mill. The process needs to be monitored and controlled from starting point to final product end. Figure 4 shows the process of cement manufacturing.

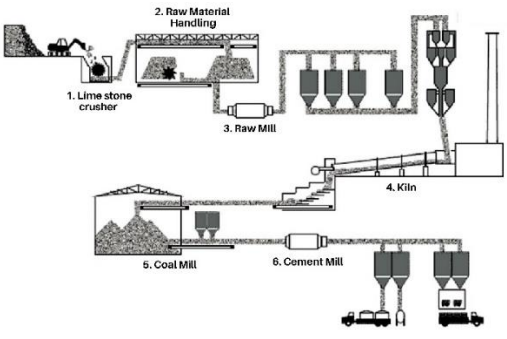


Figure 4: Process of cement manufacturing.

The identified KPIs [46] [47] normally using in cement manufacturing industry are provided. Table 3 explains the KPI for critical process parameters [48]. Table 4 shows the KPIs related to environment. Table 5 shows the material stock KPI. Table 6 explains the KPI for quality control parameters. These KPIs will be generated by the platform based on the inputs from IoT sensors.

Table 3: KPI for critical process parameters.

No.	Process	Parameter	Unit of Measurement
1	Lime Stone Crusher	Apron Feeder Speed	Rotations/Minute
		Crusher Motor Load	Kilowatt
		Limestone to Stacker	Tons/Hour
2	Raw Material Handling	Limestone Reclaimer	Tons/Hour
		Raw Mill Additive Reclaimer	Tons/Hour
		Raw coal reclaimer	Tons/Hour
		Cement Mill Additive Reclaimer	Tons/Hour
3	Raw Mill	Limestone Weigh Feeder	Tons/Hour
		Bauxide Weigh Feeder	Tons/Hour
		Hammetite Weigh Feeder	Tons/Hour
		Raw mill Total Feed	Tons/Hour
		Raw mill Motor Load	Kilowatt
		Raw Mill Differential Pressure	Millimeter Water Gauge
		Raw Fan Motor Load	Kilowatt
		Raw Mill Fan Speed	%
		Raw Mill Fan Flow	m3/Hour
		Bag House/ESP Fan Load	Kilowatt
		Bag House/ESP Fan Speed	%
		Bag House/ESP Fan Flow	m3/Hour
		Bag House/ESP Differential Pressure	Millimeter Water Gauge
		Classifier Speed	%
4	Kiln	Pre heater Fan Motor Load	Kilowatt
		Pre heater Fan Speed	%
		Pre heater Fan Flow	m3/Hour
		PH I/L O2	%
		PH I/L CO	%
		Calciner O2	%
		Calciner CO	%
		Calciner NOX	PPM
		Kiln I/L O2	%
		Kiln I/L CO	%
		Kiln I/L NOX	PPM
		Kiln Firing Coal	Tons/Hour
		Calciner Firing Coal	Tons/Hour
		Calciner Temperature	Degree Centigrade
		Kiln Feed	Tons/Hour
		Kiln motor Load	Kilowatt

		Kiln Speed	Rotations/Minute
		Kiln I/L Temperature	Degree Centigrade
		Burning Zone Temperature	Degree Centigrade
		Tertiary Air Temperature	Degree Centigrade
		Secondary air Temperature	Degree Centigrade
		Kiln Hood Draft	Millimeter Water Gauge
		Cooler Compartment Pressure	Millimeter Water Gauge
		Cooler Grate Speed	Rotations/Minute
		Clinker Temperature	Degree Centigrade
		Cooler ESP Fan Load	KW
		Cooler ESP Fan Speed	%
		Cooler ESP Fan Flow	M3/Hour
5	Coal Mill	Raw Coal Weigh Feeder	Tons/Hour
		Coal mill Motor Load	Kilowatt
		Coal Mill Differential Pressure	Millimeter Water Gauge
		Coal Mill Fan Motor Load	Kilowatt
		Coal Mill Fan Speed	%
		Coal Mill Fan Flow	M3/Hour
		Bag House Fan Load	Kilowatt
		Bag House Fan Speed	%
		Bag House Fan Flow	m3/Hour
		Bag House Differential Pressure	Millimeter Water Gauge
		Bag House I/L O2	%
		Bag House I/L CO	%
		Fine Coal Silo CO	%
		Bag House I/L Temperature	Degree Centigrade
		Classifier Speed	%
6	Cement Mill	Clinker Weigh Feeder	Tons/Hour
		Gypsum Weigh Feeder	Tons/Hour
		Puzzolana Weigh Feeder	Tons/Hour
		Cement mill Total Feed	Tons/Hour
		Cement mill Motor Load	Kilowatt
		Cement Mill Differential Pressure	Millimeter Water Gauge
		Cement Mill Fan Motor Load	Kilowatt
		Cement Mill Fan Speed	%
		Cement Mill Fan Flow	m3/Hour
		Bag House Fan Load	Kilowatt
		Bag House Fan Speed	%
		Bag House Fan Flow	m3/Hour
		Bag House Differential Pressure	Millimeter Water Gauge
		Classifier Speed	%

Cement is a commonly used construction material that requires large number of resources to manufacture and the manufacturing process have significant environmental impact [46]. The cement industries are facing challenges to implement sustainable manufacturing into their products and processes. Cement manufacturing is an intensive consumer of natural raw materials, fossil fuels, energy, and a major source of multiple pollutants. Thus, evaluating the sustainable manufacturing in this industry has become a necessity [49]. To meet the environmental requirements, the parameters related to manufacturing operations need to be monitored and is included as one of the KPIs.

Table 4: KPIs related to environment.

No.	Parameter	Unit of Measurement
1	Kiln Stack Emission	mg/Nm3
2	Coal Stack Emission	mg/Nm3
3	Cooler Stack Emission	mg/Nm3
4	Cement Stack Emission	mg/Nm3
5	Ambient Air Quality	Index
6	Water Consumption	m3/hr.
7	Waste water	m3/hr.

The information of raw material stock, material in process and finished goods availability is very important for business operations and planning. The availability of various chemicals and consumables using in manufacturing process also need to be monitored for optimum production to take place.

Table 5: Material stock KPI.

No.	Description	Unit of Measurement
1	Limestone Stock Pile	Ton
2	Raw mill Additives	Ton
3	Raw Meal Silo	Ton
4	Raw Coal Stock Pile	Ton
5	Fine Coal Silo	Ton
6	Clinker Stock Pile	Ton
7	Cement Mill Additives Gypsum	Ton
8	Cement Mill Additives Fly Ash	Ton
9	Cement Mill Performance Improver	Ton
10	Grinding Aid	Ton
11	Cement Silo	Ton
12	Water Reservoir	Litre
13	Diesel Stock	Litre

Table 6: KPI for quality control parameters.

No.	Parameter
1	Cao
2	LSF
3	Liter weight
4	Free Lime
5	C3S
6	C2S
7	Blain (OPC)
8	Blain (PPC)
9	Cement Particle Size

For monitoring KPIs, Data Acquisition Module (DAM) is installed on each site. It collects data from equipment in real time from various sensors. The platform is installed in the server available in customer premises. The data from each site is sent to platform server over Internet. Platform server processes the data with intelligence and presents it to different types of users like support team, managers, top management etc. Access control is in place so that each user sees what is relevant to user. Figure 5 shows the proposed architecture for deployment. This platform is developed based on line diagram of IoT based KPI platform shown in Figure 2 and architecture of KPI platform shown in Figure 3.

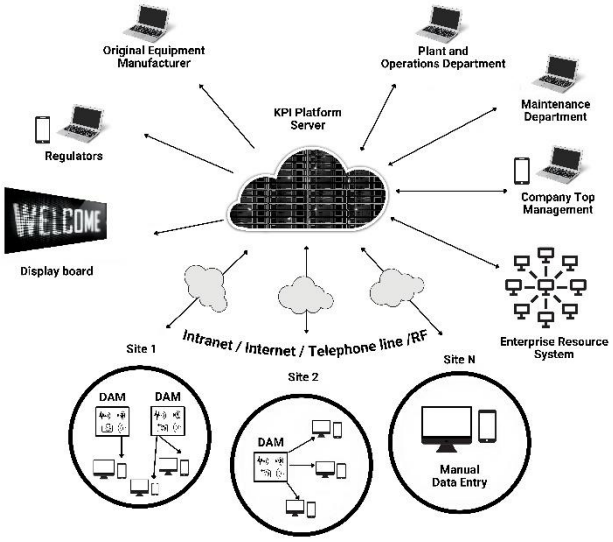


Figure 5: Proposed architecture for deployment.

The proof-of-concept platform is developed and the testing is done on a simulated environment. Few of the KPI reports generated in a mobile device are provided. Figure 6 shows the process parameter KPIs generated in visualizing device as output from platform. Environmental KPIs are shown in Figure 7. The material stock KPIs are provided in Figure 8. Quality control KPIs are shown in Figure 9. Production KPI is in Figure 10. Fuel consumption KPI is shown in Figure 11 and the power consumption is shown in Figure 12. Consolidation of data of all plants is also possible for head office application. Comparison of KPI between units within a plant or between other plants of similar size is also possible.

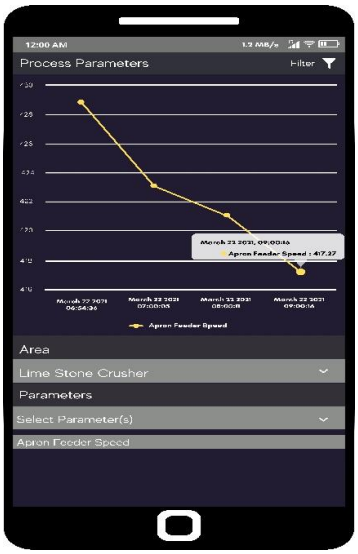


Figure 6: Process parameter KPIs.

Figure 7: Environmental KPIs.

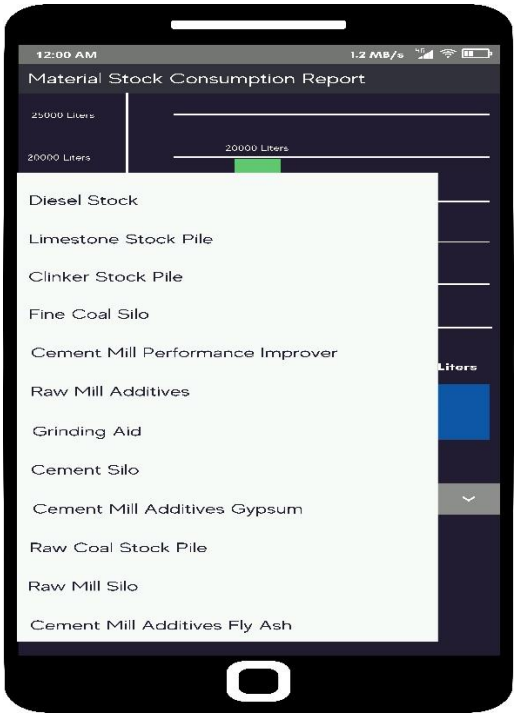


Figure 8: Material stock KPIs.

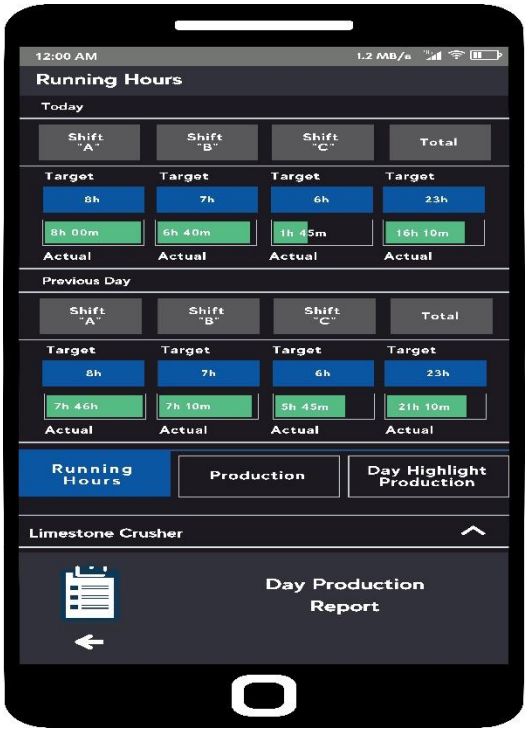


Figure 10: Production KPIs.

Figure 9: Quality control KPIs.

Figure 11: Fuel consumption KPIs.

Figure 12: Power consumption KPIs.

7 Conclusion

The developed platform is the solution for integrating mobile devices to the IoT based automation and control system of a continuous process industry. This platform is implemented at the convergence area of operations/automation and Information Technology. The platform is able to acquire various types of data, analyze the data collected and provide the required outputs to the static and mobile devices. The prototype platform developed is implemented in one of the cement manufacturing industries at the plant server and at the head office server as well. The KPIs required for this cement manufacturing plant is identified and deployed in this platform. This developmental model can be extended to steel, petrochemicals, sugar, paper, fertilizer, food, pharmaceutical industry etc. As a future work, the platform can be installed in the cloud which can be accessed by plants as well as head office. With the acceptance and popularity in industry with IoT based KPI platform, it can be developed in the cloud and provide Platform as a Service (PaaS) to customers.

References

[1.] C. R. Sekhar, P. Hema, and C. E. Reddy (2018). Equipment Effectiveness Improvement in a Continuous Process Industry. International Journal of Research and Analytical Reviews (IJRAR), vol. 5, pp. 134-142.

[2.] M. G. Hudedmani, R. M. Umayal, S. K. Kabberalli and R. Hittalamani (2017). Programmable Logic Controller (PLC) in Automation. Advanced Journal of Graduate Research, vol. 2, pp. 37-45. https://doi.org/10.21467/ajgr.2.1.37-45

[3.] R. Kirubashankar, K. Krishnamurthy and J. Indra (2009). Remote monitoring system for distributed control of industrial plant process. Journal of Scientific & Industrial Research, vol. 68, pp.858-860.

[4.] M. M. Lashin (2014). Different Applications of Programmable Logic Controller (PLC). International Journal of Computer Science, Engineering and Information Technology (IJCSEIT), vol. 4, pp.27-32. https://doi.org/10.5121/ijcseit.2014.4103

[5.] M. Dahm and A. Mathur (1990). Automation in the food processing industry: distributed control systems. Butterworth & Co. (Publishers) Ltd. Pp. 32-35.

[6.] K. Stouffer, J. Falco and K. Kent (2006). Guide to Supervisory Control and Data Acquisition (SCADA) and Industrial Control Systems Security. NIST Special Publication 800-82, Intelligent Systems Division, Gaithersburg.

[7.] S. Li, L. D. Xu and S. Zhao (2015). The internet of things: a survey. Inf Syst Front, Springer, vol. 17, pp. 243-259. https://doi.org/10.1007/s10796-014-9492-7

[8.] Sadeghi, C. Wachsmann, M. Waidner (2015). Security and Privacy Challenges in Industrial Internet of Things. DAC ’15, ACM, San Francisco, CA, USA, pp. 7-11. http://dx.doi.org/10.1145/2744769.2747942

[9.] M. Short and F. Abugchem (2017). A Microcontroller-Based Adaptive Model Predictive Control Platform for Process Control Applications. Electronics, vol. 6, pp.1-17. https://doi.org/10.3390/electronics6040088

[10.] J. C. Kabugoa, S. L. Jounelaa, R. Schiemannb and C. Binder (2020): Industry 4.0 based process data analytics platform: A waste-to-energy plant case study. International Journal of Electrical Power and Energy Systems, vol. 115. https://doi.org/10.1016/j.ijepes.2019.105508

[11.] Z. Ge, Z. Song, S. X. Ding and B. Huang (2017). Data Mining and Analytics in the Process Industry: The Role of Machine Learning. IEEE Access, vol. 5, pp. 20590-20616. https://doi.org/10.1109/ACCESS.2017.2756872

[12.] E. Goldin, D. Feldman, G. Georgoulas, M. Castano and G. Nikolakopoulos (2017). Cloud Computing for Big Data Analytics in the Process Control Industry. Proceedings of 25th Mediterranean Conference on Control and Automation (MED), IEEE, Valletta, Malta. https://doi.org/10.1109/MED.2017.7984310

[13.] M. H. Rehman, I. Yaqoob, K. Salah, M. Imran, P. P Jayaraman and C. Perera (2019). The Role of Big Data Analytics in Industrial Internet of Things. Future Generation Computer Systems, vol. 99, pp. 247-259. https://doi.org/10.1016/j.future.2019.04.020

[14.] L.V. Zhihan, L. Qiao, S. Verma and Kavita (2021). AI-enabled IoT-Edge Data Analytics for Connected Living. ACM Transactions on Internet Technology, vol. 21, pp. 1-20. https://doi.org/10.1145/3421510

[15.] D. Reguera-Bakhache, I. Garitano, R. Uribeetxeberria, C. Cernuda and U. Zurutuza (2020). Data-Driven Industrial Human-Machine Interface Temporal Adaptation for Process Optimization. Proceedings of 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), IEEE, Vienna, Austria. https://doi.org/10.1109/ETFA46521.2020.9211930

[16.] G.Prashanta (2020). Case Study: Industrial Automation Using PLC. Iconic Research and Engineering Journals, vol. 4, pp. 52-55.

[17.] S. Thalmann, J. Mangler, T. Schreck, C. Huemer, M. Streit, F. Pauker, G. Weichhart, S. Schulte, C. Kittl, C. Pollak, M. Vukovic, G. Kappel, M. Gashi, S. Rinderle-Ma, J. Suschnigg, N. Jekic, S. Lindstaedt (2018). Data Analytics for Industrial Process Improvement A Vision Paper. Proceedings of the IEEE 20th Conference on Business Informatics, IEEE, Vienna, Austria, pp. 92-96. https://doi.org/10.1109/CBI.2018.10051

[18.] S. D. Anton, D. Fraunholz, C. Lipps, F. Pohl, M. Zimmermann and H. D. Schotten (2017). Two decades of SCADA exploitation: A brief history. Proceedings of the 2017 IEEE Conference on Application, Information and Network Security (AINS), IEEE, Miri, Malaysia, pp. 98-104. https://doi.org/10.1109/AINS.2017.8270432

[19.] P. Samuel, V. R. Alexandru, M. Alexandru, Z. B. Constantin (2020). Architectural Issues in Implementing a Distributed Control System for an Industry 4.0 Prototype. Proceedings of 15th International Conference on Development and Application Systems, IEEE, Suceava, Romania, pp. 56-59. https://doi.org/10.1109/DAS49615.2020.9108924

[20.] M. B. Yassein, M. Q. Shatnawi, S. Aljwarneh and R. Al-Hatmi (2017). Internet of Things: Survey and open issues of MQTT protocol. Proceedings of the International Conference on Engineering & MIS (ICEMIS), Monastir, Tunisia, pp. 1-6. https://doi.org/10.1109/ICEMIS.2017.8273112

[21.] T. Yokotani, S. Ohno, H. Mukai and K. Ishibashi (2021). IoT Platform with Distributed Brokers on MQTT. International Journal of Future Computer and Communication, vol. 10, pp. 7-12. https://doi.org/10.18178/ijfcc.2021.10.1.572

[22.] H. Chien, Y. Chen, G. Qiu, J. F. Liao, R. Hung, P. Lin, X. Kou, M. Chiang and C. Su (2020). A MQTT-API-compatible IoT security-enhanced platform. Int. J. Sensor Networks, vol. 32, pp. 54-68. https://dx.doi.org/10.1504/IJSNET.2020.104463

[23.] H. Bauer, S. Hoppner, C. Iatrou, Z. Charania, S. Hartmann, S. Rehman, A.Dixius, G. Ellguth, D. Walter, J. Uhlig, F. Neumarker, M. Berthel, M. Stolba, F.Kelber, L. Urbas and C. Mayr (2021). Hardware Implementation of an OPC UA Server for Industrial Field Devices. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 29. https://doi.ieeecomputersociety.org/10.1109/TVLSI.2021.3117401

[24.] M. Younan, E. H. Houssein, M. Elhoseny, A. A. Ali (2019). Challenges and recommended technologies for the industrial internet of things: A comprehensive review. Measurement, vol. 151 http://dx.doi.org/10.1016/j.measurement.2019.107198

[25.] N. Mehdiyev, A. Emrich, B. Stahmer, P. Fettke and P. Loos (2017). iPRODICT – Intelligent Process Prediction based on Big Data Analytics. Proceedings of Business Process Management (BPM-17) Industry Track, Barcelona, Spain.

[26.] Ahmed, S. Obermeier, S. Sudhakaran and V. Roussev (2017). Programmable Logic Controller Forensics. IEEE Security & Privacy, vol. 15, pp. 18-24. https://doi.ieeecomputersociety.org/10.1109/MSP.2017.4251102

[27.] C. Rameback (2003). Process automation systems-history and future. Proceedings of 2003 IEEE Conference on Emerging Technologies and Factory Automation, IEEE, Lisbon, Portugal. https://doi.org/10.1109/ETFA.2003.1247680

[28.] D. R. Milivojevic, V. Despotovic, V. Tasic and M. Pavlov (2010). Process Control Program as an Element of Distributed Control System. Information Technology and Control, vol. 39. pp. 152-158.

[29.] J. M. Sardroud (2012). Influence of RFID technology on automated management of construction materials and components. Scientia Iranica, vol. 19(3), pp. 381-392. http://dx.doi.org/10.1016/j.scient.2012.02.023

[30.] A. Akbari, S. Mirshahi and M. Hashemipour (2015). Application of RFID System for the Process Control of Distributed Manufacturing System. Proceedings of Canadian Conference on Electrical and Computer Engineering, IEEE,

Halifax, NS, Canada. https://doi.org/10.1109/CCECE.2015.7129325

[31.] L. G. Kurmi, S. D. Patil and M. L. Yadav (2014). NFC Based Library Automation using Smart Phone. International Journal of Engineering Research & Technology (IJERT), vol. 3. pp. 1648-1651.

[32.] C. Lesjak, T. Ruprechter, H. Bock, J. Haid and

[33.] E.Brenner (2014). Facilitating a Secured Status Data Acquisition from Industrial Equipment via NFC. International Journal of Internet Technology and Secured Transactions, vol. 3(3), pp.288 – 299. http://dx.doi.org/10.20533/jitst.2046.3723.2014.0037

[34.] T. Kurfess, C. Saldana, K. Saleeby and M. Parto-Dezfouli (2020). A Review of Modern Communication Technologies for Digital Manufacturing Processes in Industry 4.0. Journal of Manufacturing Science and Engineering, vol. 142. pp. http://doi.org/10.1115/1.4048206

[35.] B. Dafflon, N. Moalla and Y. Ouzrout (2021). The challenges, approaches, and used techniques of CPS for manufacturing in Industry 4.0: a literature review. The International Journal of Advanced Manufacturing Technology, vol. 113, pp.2395–2412. http://dx.doi.org/10.20533/jitst.2046.3723.2014.0037

[36.] D. Raposo, A. Rodrigues, S. Sinche, J. S. Silva and F. Boavida (2018). Industrial IoT Monitoring: Technologies and Architecture Proposal. Sensors, vol. 18(10), pp. 1-32. http://dx.doi.org/10.3390/s18103568

[37.] S.S. Mahmood and P.Sharma (2019). Industrial Automation using Zigbee Communication Protocol. International Journal of Recent Technology and Engineering (IJRTE), vol.8, pp. 7240-7243. http://dx.doi.org/10.35940/ijrte.C6294.098319

[38.] P. A. M. Devan, F. A. Hussin, R. Ibrahim, K. Bingi and F. A. Khanday (2021). A Survey on the Application of WirelessHART for Industrial Process Monitoring and Control. Sensors, vol. 21(15), pp. 1-26. https://doi.org/10.3390/s21154951

[39.] T. Hasegawa, H. Hayashi, T. Kitai, H. Sasajima (2011). Industrial Wireless Standardization - Scope and Implementation of ISA SP100 Standard. Proceedings of SICE Annual Conference, IEEE, Tokyo, Japan, pp. 2059-2064.

[40.] Y. N. Valadao, G. Kunzel, I. Muller and C. E. Pereira (2018). Industrial Wireless Automation: Overview and Evolution of WIA-PA. Proceedings of the International Federation of Automatic Control, Energy Procedia, Elsevier, pp. 175-180. https://doi.org/10.1016/j.ifacol.2018.06.257

[41.] C. F. Lindberga, S.T. Tan, J.Y. Yan, F. Starfelt (2015). Key performance indicators improve industrial performance. Proceedings of the 7th International Conference on Applied Energy – ICAE2015, Energy Procedia, Elsevier, pp. 1785-1790. https://doi.org/10.1016/j.egypro.2015.07.474

[42.] B. Galloway and G. P. Hancke (2012). Introduction to Industrial Control Networks. IEEE Communications Surveys & Tutorials, vol.15, pp. 860 – 880. https://doi.org/10.1109/SURV.2012.071812.00124

[43.] L. Cattaneo, L. Fumagalli, M. Macchi and E. Negri (2018). Clarifying Data Analytics Concepts for Industrial Engineering. Proceedings of the International Federation of Automatic Control, Elsevier, pp. 820-825. https://doi.org/10.1016/j.ifacol.2018.08.440

A. K. Y. Benhamidouche (2021). Prediction of Cement Fineness Using Machine Learning Approaches. PhD. Thesis, Faculty of Technology, University Mohamed Boudiaf - M'sila, People's Democratic Republic of Algeria.

[44.] D. N. Huntzinger and T. D. Eatmon (2008). A life-cycle assessment of Portland cement manufacturing: comparing the traditional process with alternative technologies. Journal of Cleaner Production, vol. 17. pp. 668-675. https://doi.org/10.1016/j.jclepro.2008.04.007

[45.] A. Rahman, M.G. Rasul, M.M.K. Khan and S. Sharma (2013). Impact of alternative fuels on the cement manufacturing plant performance: an overview. Proceedings of the 5th BSME International Conference on Thermal Engineering, Elsevier, pp. 393-400. https://doi.org/10.1016/j.proeng.2013.03.138

[46.] E. Amrina and A. L. Vilsi (2015). Key Performance Indicators for Sustainable Manufacturing Evaluation in Cement Industry. Proceedings of the 12th Global Conference on Sustainable Manufacturing, Elsevier, pp. 19-23. https://doi.org/10.1016/j.procir.2014.07.173

[47.] J. P. John (2020). Parametric Studies of Cement Production Processes. Journal of Energy, vol. 2020, pp. 1-17. https://doi.org/10.1155/2020/4289043

[48.] R. Feiz, J. Ammenberg, L.Baas, M. Eklund, A. Helgstrand and R. Marshall (2015). Improving the CO2 performance of cement, part I: Utilizing life-cycle assessment and key performance indicators to assess development within the cement industry. Journal of Cleaner Production, vol.98, pp.272-

281.
http://dx.doi.org/10.1016/j.jclepro.2014.01.083

[49.] A. K. Mishra and A. Jha (2019). Quality Assessment of Sarbottam Cement of Nepal. International Journal of Operations Management and Services, vol. 9. pp. 1-22.

[50.] N.A. Madlool, R. Saidur, M.S. Hossain, N.A. and Rahim (2011). A critical review on energy use and savings in the cement industries. Renewable and Sustainable Energy Reviews, vol. 15. pp. 2042-2060. https://doi.org/10.1016/j.rser.2011.01.005

Generating Lyrics using Constrained Random Walks on a Word Network

Žiga Babnik, Jasmina Pegan, Domen Kos and Lovro Šubelj
Faculty of Computer and Information Science, Večna pot 113, 1000 Ljubljana, Slovenia
E-mail: zb1996@student.uni-lj.si, jp2634@student.uni-lj.si, dk6314@student.uni-lj.si, lovro.subelj@fri.uni-lj.si

Student paper

Keywords: lyrics, lyrics generation, Markov models, networks, network analysis, poetry generation, semantics

Received: November 16, 2020

In the paper we present an approach for automatic lyrics generation. From the American National Corpus of written texts we build a Word Network, which encodes word sequences. Lyrics are then generated by performing a constrained random walk over the Word Network. *The constraints include the structure of the generated sentence, the rhythm of the lines of the stanza or the rhymes of the stanza itself. Lyrics are generated using each constraint individually and also using all three constraints at the same time. We tested the single constraint strategies using a toy example, while the results of the joint strategy were subject to human review. While the given properties of the toy example, were kept in the results, replicating the toy example perfectly proved a difficult task. The results of the questionnaire showed that lack of a deeper meaning and strange capitalization were the main reasons that our results did not appear as though they were written by a human.*

Povzetek: Avtomatsko generiranje besedil pesmi bazira na uporabi omejenih naključnih sprehodov po besednem omrežju Word Network, vzpostavljenem iz Ameriškega nacionalnega korpusa. Besedila se generirajo z upoštevanjem strukture stavka, ritma in rim.

1 Introduction

Natural Language Processing (NLP) is becoming a very popular research field, with many researchers working on it. Many methods for speech recognition, understanding of language and generation of text are being developed. In this paper we concentrate on the subtask of NLP which is text generation. More specifically we address the problem of generating lyrics that resemble real lyrics in some way.

Since the development of deep neural networks most of the state of the art approaches for text generation are based on extracting features with deep learning. Our approach is to use NLP tools and methods to extract them manually and pack them into one or more networks, all containing some information about real lyrics. The main idea is to use a big data set of existing songs and possibly other texts and construct the needed networks out of them. With some constrained random walk through this networks we then generate lyrics for new songs. The nodes in the main network, we call the *Word Network*, are words and the edges are relations between them. We focus on building many strategies where each strategy ensures one property of real texts is satisfied, such as rhythm, rhymes and sentence structure, all of which play a major role in lyrics. By combining these individual strategies we want to create a system that generates lyrics which mimic real lyrics in many different aspects.

In section 2 we overview papers relevant to our research. Firstly we present an overview of the field in section 2.1, after which we take a closer look at three most relevant papers. In section 2.2 we present a paper [9] that introduces the *PoeTryMe* poetry generation system, in section 2.3 we present the paper [8] that introduces the *Tra-la-Lyrics* song lyrics generation system and in section 2.4 we present a paper [6] that introduces a Markov Constraint based system for lyrics generation.

In section 3 we present the data used to build the necessary networks and generate new lyrics as well as the *Word Network*, which is the central data structure of our system.

In section 4 we present our general approach for the implementation of a lyrics generation system. Firstly in section 4.1 we present how the lyrics structure was generated, which is then used in different text generation methods presented in section 4.2. Finally in section 4.3 we present how the generated text is reorganized using the structural information to produce the final generated lyrics.

In section 5 we present the results of the different generation strategies. Firstly in section 5.1 we present an example of generated lyrics for each of the developed strategies. In section 5.2 evaluation of the results using a toy example is presented, finally in section 5.3 evaluation using public review is presented. In section 6 we present the main results of our paper as well as propose our interpretation of them. Finally in section 7 we overview the paper.

2 Related work

2.1 A Survey on intelligent poetry generation

The authors of the paper *A Survey on Intelligent Poetry Generation: Languages, Features, Techniques, Reutilisation and Evaluation* [7] made an overview of the intelligent generation of poetry area. In the paper they discuss many topics, mainly surrounding different types of poetry, the structure of poems and how to recognise them. They also discuss the most common formulated features and how to design a generator which takes into account the features based on the language of the poem. Another important mention are so called *Content features* which depend on the grammatical correctness and meaningfulness of the text and how to achieve them.

In the second part of the paper they discuss artificial intelligence techniques for poem generation. One of the interesting approaches is to use genetic algorithms where the population is represented with initial drafts and in each iteration the most promising texts are kept. New poem are generated using mutations and crossover operations which are evaluated by some fitness function. Another approach is to present this as a constraint optimization problem where constraints are represented with the number of lines, syllables per line, number of rhymes, etc. The algorithm should generate a poem such that it optimizes those constraints. Standard machine learning methods were also used. A Support Vector Machine (SVM) [11] model was trained on a poetry corpus and used to predict the next word or syllable. They also used language models to generate poetry texts which were represented with Markov models and some Deep Neural Networks (DNNs) which includes Recurrent Neural Networks (RNNs) [10].

In the last part they also discuss the evaluation of such texts where most of the reliable evaluation is still performed by humans. They discuss some metrics which are mostly used for classification of the poem type by measuring occurrence of different properties in a generated poem.

2.2 PoeTryMe

In the paper *PoeTryMe: a versatile platform for poetry generation* [9] an automatized poetry generation system for Portuguese poetry is presented. It uses a set of seed words to describe the general context of the goal lyrics, and a poem template for structure and rhythm. PoeTryMe supports syllable-based rhythm with no regards to stress patterns. Also grammar and word relations represented by relational triples $(node_1, relation\_type, node_2)$ can be user-defined.

The paper categorizes poetry generation techniques into four categories: template-based where a sentence is generated in accordance to the template, generate-and-test where n sentences are generated and the best is chosen, evolutionary where n poems are generated, then the best few are selected and crossed repeatedly and case-based reasoning approach that uses adaptation of existing songs. Implementation of the algorithm PoeTryMe uses three different strategies to generate lines: basic which is categorized as template-based, generate-and-test and an evolutionary approach. The system is modular, it consists of a sentence generator, grammar processor, relations manager, contextualizer, syllables utility, sentiment processor and a generation strategy [8] already described.

Three generated poems are presented as the results. The authors confirm that following multiple properties of poetry such as meaningfulness, grammatical correctness and poeticness at the same time is hard. PoeTryMe generates grammatically correct sentences which are somehow related to a given keywords and at the same time conforming to given structure. Only the evolutionary approach has rhymes with high probability.

2.3 Tra-la-Lyrics 2.0

In the paper *Tra-la-Lyrics 2.0: Automatic Generation of Song Lyrics on a Semantic Domain* [8] a system for automatic generation of lyrics is presented. Tra-la-Lyrics 2.0 generates text with rhymes on a semantic domain with a given rhythm, based on input music. Its predecessor Tra-la-Lyrics generated rhymed rhythmicized text based on stressed syllables with no regards to semantics. The 2.0 version integrates the previous approach with PoeTryMe to achieve generation of meaningful lyrics on a given topic with rhythm and rhymes.

Tra-la-Lyrics has two rhyming strategies: Rhythm+Rhymes (RR) and Generative Grammar (GG). The RR strategy prefers rhymes at specific parts of the song. In addition to that, GG sets morphological constraints. As lyrics are often repetitive, both strategies also include a repetition parameter.

The implementation of Tra-la-Lyrics 2.0 was derived from PoeTryMe by changing the algorithm to accept a song as an input and by creating a new generation strategy which considers also the rhythm.

Results are again presented in the form of generated lyrics. The results of Tra-la-Lyrics and Tra-la-Lyrics 2.0 are evaluated empirically and numerically on a number of points such as rhythm, rhymes, semantics and meaningfulness. On the average Tra-la-Lyrics 2.0 outperforms its precedent, but although it shows improvement in meaningfulness, it is still far from perfect.

2.4 Markov constraints for generating lyrics

In the paper [6] the authors used Markov models to generate lyrics in the style of existing authors. Since the Markov chains are not suitable to satisfy the non-local properties of poems such as structural constraints, the authors developed a more advanced framework. Using so called Constrained Markov Processes (CMP) they generated texts that were consistent with the corpus. The idea is to represent the prob-

lem as the constraint satisfaction problem. A Markov probabilistic model is then built in two steps. They presented two different constraints. The first one is replacing the transition probability in the standard Markov model. It is called *Markov constraint* and beside the transition probability also holds a constraint variable. The other type is called *Control constraint* which needs to be satisfied in some specific state. The *Markov constraints* on each transition are then set so that they satisfy *Control constraint*. Using these techniques they were able to keep structural properties of the poems such as rhyme and rhythm.

They also demonstrated the methods and evaluate them. The evaluation was again performed manually by 12 volunteers.

3 Data

Our approach is based on a constrained random walk over the so-called *Word Network*. We have to ensure that the *Word Network* is large enough, so that we will be able to perform the constrained random walks on it. In order for the *Word Network* to be large enough it needs to be constructed from a large data set, we chose the *Open American National Corpus* data set [3] which contains over 6000 texts from different domains, totaling around 11 million words.

Since our approach tries to generate text that mimics lyrics by some property, we also need a data set that includes lyrics, from which we will be able to extract these properties. We chose the *Song Lyrics* data set [5] available on the Kaggle platform. The data set includes lyrics from 49 different artists such as Adele, The Beatles, Bob Marley and countless others, gained from free online lyrics hosting websites using a Python script. For each artist a single text file is available that contains lyrics from several songs of the artist. Since the data structures built from this data set are specific to certain sub-tasks of our approach, we will introduce them later on.

3.0.1 Word network

The *Word Network* is a directed network and represents the dependencies between single words in the lyrics, the nodes in the network represent individual words, while the links show if two words appear in the lyrics one after another. To build such a network we first tokenize each sentence of the texts. We than construct a list of all word tuples, such that the first word in the tuple is always followed by the second word in the tuple in the lyrics. To build the network we than iterate over all such tuples adding individual words as nodes in the network, where each word node gets the following attributes: the Part-of-Speech tag (POS tag) of the word and a list of all possible phonemes of the word. After adding both words from the tuple into the network we than do the following, if a link already exists between the words in the network we increase the weight of the link by one, on the other hand if a link does not exist we simply add it with weight equal to one.

Table 1 presents some basic statistics of the *Word Network*, while Figure 1 presents the indegree and outdegree distributions of the network.

Statistic	Result
Number of nodes (n)	60115
Number of links (m)	2357451
Average degree ($\overline{k}$)	39
Density (ρ)	0.00065
Number of nodes in LCC	60111
Average clustering coefficient ($\overline{C}$)	0.467

Table 1: Basic statistics of the *Word Network*

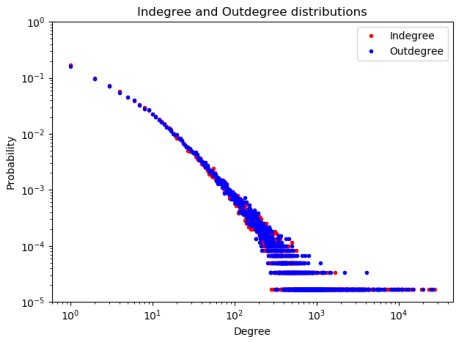


Figure 1: Indegree and outdegree distributions

From the statistics we see that the *Word Network* is quite dense, which is a desirable property for our approach. To calculate the number of nodes in the largest connected component, we first turned the *Word Network* into an undirected network, we see that most of the words are within the largest connected component. Finally we see that its degree distributions roughly follow a power-law distribution.

4 Methods

Our approach consists of three stages. In the first stage the general structure of the lyrics is generated, here we obtain the following: how different stanzas such as the chorus and verse follow each other and also how many lines are contained in each of them. This information is fed to the second stage which generates lines for each stanza in the lyrics structure. The third stage than collects the lines and stacks them according to the lyrics structure, while also adding details such as capitalization and commas. Figure 2 shows visually how our approach is structured at the highest level.

4.1 Generating lyrics structure

The approach used to generate lyrics structure uses a simple network called the *Structure Network*. The *Structure Network* is a directed network, which contains the four most basic blocks of lyrics: *intro*, *verse*, *chorus* and *bridge*. Figure 3 shows how these nodes are connected and the

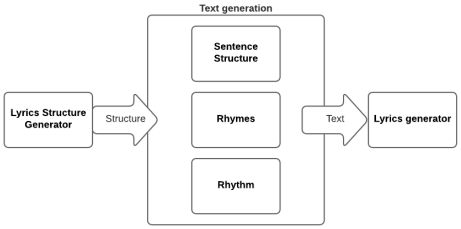


Figure 2: Visualization of approach pipeline

weights of each connection. The *Structure Network* was handcrafted and represents only a rough approximation of how a song can be structured.

To generate the lyrics structure a random walk starting from the *intro* node was performed. The walk was stopped once more than five steps were performed and the current observed node was not *verse*, meaning we did not want our lyrics to end with a *verse*.

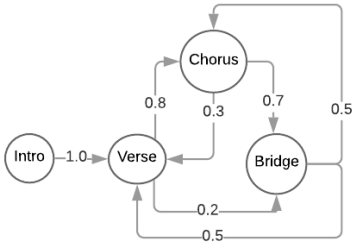


Figure 3: Visualization of the Structure Network

After the lyrics structure was generated we also generated the number of lines each part should contain. This was done simply by randomly selecting a number in a given interval. The interval was defined as $[3, 6]$ for *verse*, *chorus* and *bridge*, while for *intro* it was defined as $[2, 4]$.

4.2 Generating text with certain properties

In the following section we present approaches for generating texts with certain properties. These properties include proper sentence structure, rhymes and rhythm. We introduce strategies for generating lyrics that take only one of these properties into consideration and also a joint strategy which takes all three into consideration.

4.2.1 Generating text with proper sentence structure

We propose a model that takes into account the sentence structure present in individual lines of lyrics. Alongside the *Word Network* this strategy also uses the *Part-of-speech tag network* or *POS-tag network*.

The *POS-tag network* is a directed network that contains information about the sequences of line structures we can observe in lyrics. Firstly each line in the lyrics is represented as a sequence of Part-of-Speech (POS) tags, that tell us the structure of the line. The POS tag sequences of lines are then added to the network as nodes, we create a link between POS tag sequences X and Y, if it holds that in the lyrics POS tag sequence Y comes directly after POS tag sequence X. By performing a random walk over such a network we not only guarantee the proper structure of each individual line, but also proper ordering of lines.

To generate text for a given part, we first generate a sequence of line structures using the *POS-tag network*. This is done using simple random walks over the network, where the number of steps equals the length of the given part. The walk generates all the needed constraints for this strategy.

Once the constraint in the form of POS-tag structure of individual lines has been generated, we perform a constrained random walk on the *Word Network*. For each line a seperate constrained random walk is performed. The first word of a line is chosen randomly among all words with the proper POS-tag, for each successor we search among all neighbors of the current word, that again have the proper POS-tag. If such a neighbor does not exist, we start the walk for the current line from the beginning.

4.2.2 Generating text based on rhyme scheme

We first defined three types of rhymes. Since the *Word Network* is constructed from random texts and not lyrics we do not expect to find many neighbourhood words that correspond to a perfect rhyme. That is why we allow three types of rhymes. The first one is a perfect rhyme which is defined as the rhyme where the stressed vowels and any succeeding consonants are identical e.g. *believe* and *conceive* [4]. The second rhyme is called assonance or a vowel rhyme. It is a rhyme in which the same vowel sounds are used with different consonants in the stressed syllables of the rhyming words e.g. *pentient* and *reticence* [1]. The third rhyme is called consonant rhyme and is the repetition of consonants or consonant patterns especially at the ends of words e.g. *bell* and *ball* [2].

Our strategy generates words in such order that they follow the rhyme scheme we chose. By defining the number of words in a line and the rhyme scheme, e.g ”ABBA”, we then generate our lyrics by randomly choosing the node in the *Word Network*. As in a random walk we chose a successor by taking into account the weights of each edge. After reaching the last word in the line we chose the next node only from the successor that do not violate the rhyme scheme. In this step we chose the successor both uniformly at random and by weighting each rhyme. Most of the stop words have high weights in the *Word Network* and they are usually short and by definition most of them rhyme. That is why they were chosen in most of the cases. The strategy generated more natural rhymes when choosing them randomly.

4.2.3 Generating text with rhythm

We propose a model that generates lyrics based on a given rhythm. The model uses the *Word Network* and an additional network storing rhythm data. The *Rhythm network* is

a weighted directed network that represents rhythms found in the lyrics. It consists of three nodes: -1 represents start or ending of a line, 0 stands for an unstressed syllable and 1 indicates a stressed syllable. The weights of edges were decided by calculating the normalized number of transitions between the corresponding syllables. We can use this network to generate a random rhythm by starting at node -1 and choosing a random neighbor taking into account corresponding edge weights. When we reach -1 again, the generated line of rhythm is completed. The generated line of rhythm is then corrected to include more repetitiveness which is expected from rhythm to feel natural. The first n syllable stresses are chosen as baseline rhythm and are then propagated throughout the rhythm line with 70% probability. Number of syllables n is a randomly chosen number between 2 and 4.

The rhythm-based model has two variants: one is given a rhythm and the other generates the rhythm for each verse. Each variant has two sub-variants, one uses random walk strategy and the other uses POS tag strategy.

First, the rhythm is acquired in form of a string where 0 stands for an unstressed syllable and 1 for a stressed syllable. If the rhythm is given, it is expanded to match the line length. Otherwise, a random line rhythm is generated from the rhythm network for each verse of generated text. We start with a random node from the *Word Network* or the *POS-tag network*, depending on the variant, and expand the line with successors that match the required rhythm.

4.2.4 Generating text with multiple constraints

Our last strategy combined all the above constraints: structure, rhythm and rhyme. Each line was generated such that it took into account the structure, rhythm and the rhyme. Although the network is relatively large, it sometimes happened that none of the successors would satisfy all the constraints. In that case we performed a random jump and started generating the current line again. This was repeated until the generated text satisfied all the constraints.

4.3 Constructing lyrics from generated text

In the final phase of our approach we combine the generated text, with the structure of the lyrics to produce proper lyrics. Firstly we reorder the text according to the lyrics structure, so that the parts properly follow each other. Each part also gets an annotator in the form of `[<part name>]`. The ordered and annotated text is capitalized using a simple POS-tag heuristic, where we simply check the tag of each word. If the tag of the word is `NN` or `NNS`, or if the word is the first in line, we capitalize it. Commas are also added to each line, as well as a line separator between each part.

The text generators described previously generates text for each unique part only once, meaning that if the lyrics contain for instance more than one *chorus* all would contain the same text. To avoid exact repetition five to ten words in each part were chosen at random to be replaced. The whole text of the part along with the list of words to replace were then sent back to the generator. The generator then replaced the words according to the same constraints the text was generated in the beginning.

5 Results

In the following section we present the results of each strategy described in section 4.2. For each strategy we present lyrics generated by that strategy.

In order to evaluate how well the strategies work, we devised two different evaluation approaches. The first approach using a toy example was applied to all strategies where only one property of lyrics was being considered, the second approach using public review was applied to the strategy which considered all three properties.

5.1 Lyrics generated by strategies

In the following section we present lyrics generated by each of the strategies developed.

5.1.1 Lyrics generated by the sentence structure strategy

The lyrics generated based on the sentence structure strategy are presented below.

[intro]
Benjamin upward Half as described a Variety,
Carriages looser Sugar as,
Nov Angstroms o Ja Et Banditry,
Newsgroups has bacall Whitney.

[verse]
Okay like newsweek explain what you awoke,
Handhold Inspire Confidence Nudge between Gore Camp s play it s,
Think I this Observation shows helplessly he co their Interfaces,
Mine Everybody harsher Children,
Professor tour Hitler S to have,
Concludes it s my E.

[chorus]
Tina would have their Security Capital Sneeze Bob Bookies won T,
Hottest Tech Approach for nettlesome Human,
Versa Whereas outfit Blame for directional the Apparatus,
Lorge Jr Call Attention Let s h,
Craigie the Membrane regular blazes the Funding Mix.

[bridge]
Postage in Yakovlev for Discounting,
Sloth less stringent E ether Inhibitory,
Boris a Tampax described i attend academic Literature Briefly back mr Morris I think they.

[verse] Method like newsweek explain what you awoke,
Handhold Inspire Confidence Ricky between Gore Camp s play it s,
Think I this Detainee chides liter he co their Interfaces,
Mine Encyclopaedia Hanover Campsites,
Professor tour Hitler S to have,
Concludes it arsenal my E.

[chorus]
Tina could have their Security oilseed Danforth Bob Bookies won T,
Freshener tech Airport for nettlesome Human,
Teacher whereas Outfit Blame for directional Argentina Apparatus,
Strove Jr Call Attention Let s h,
Craigie the Neville Sari blazes the Funding Mix.

While the generated sentences might follow the correct POS tag structure it is clear that this constraint is not strong

enough to enable the generation of lyrics that would to some extent resemble real lyrics. Generation of meaningful lyrics is not one of our main goals, but rather that individual lines and smaller building blocks could resemble real parts of lyrics. This generation strategy is not informative enough to enable the generation of such lyrics.

5.1.2 Lyrics generated by the strategy based on the rhyme scheme

In the example below we can observe the lyrics generated by our strategy with a predefined rhyme scheme which is ”ABA” for *intro*, *verse* and *chorus*, while for the *bridge* ”ABABAB” was used.

[intro]
Shrug Democrats have and when in,
Voided the strengthening their biggest Risk,
Tetrads were controlled Trials since when.

[verse]
Plumage of ethical Story,
Stoneburner and the to,
Somehow forced the Study.

[bridge]
Randy s no Consensus that federal Reserve,
Lighting Shows which the Elk they be,
Summits to any other Poet who was,
Coelho but if you find Anybody who,
Dystrophy Pages for personal favorite Newspaper but,
Retried unpopular Gingrich in faces some Corroboration.

[chorus]
Lousiness impregnable the Incorrect,
Kolb and other racial,
Berri for Performance Checked.

[bridge]
Randy s no Consensus that federal Reserve,
Lighting Shows which the Elk they be,
Summits to all other Poet who preserve,
Coelho if you find Anybody than me,
Dystrophy Pages for personal favorite Newspaper Reserve,
Retried unpopular Gingrich in faces some be.

We can observe that some rhymes are more natural like ”be-me” and ”preserve-reserve”, while the others satisfy the definition but are not so natural to read, e.g. ”in-when”.

5.1.3 Lyrics generated by the rhythm-based strategy

We tested our rhythm-based strategy. First we look at the variant of the algorithm which accepts rhythm as an input. When given the rhythm ”011” of Prešeren’s *Povodni mož*, the random walk sub-version returns the following example.

[intro]
Duked out on the pre Birth Control of a Round Golf,
To find a Campaign seems and a Union Rules in,
The Lung lymphocytes a Tractor it is so says,
The good Seafood in S a new Car are not the.

[verse]
Kasparov was Things on the Times you about the,
Subject Matter when just Delights in e big Hot,
And more Years make are Fans will find both good or wade.

[chorus]
Embodiment of Error R a relaxed Dole,
Campaign had to turn out the Sum up to be those,
Of south will its Roots to its Entries and but and,
A M Rosenthal who was true but impressed by,
The Bush s alleged new Account for the six Lines,
Though a new Environment Act at the Field flat.

[bridge]
Nationals leaving its Building overlooks this,
Has the brooklyn Heights or a large Part of Spread his,
It a Ratio of the new Products that we,
Punished the Schedule slipped across Species female,
In the black and it a Policewoman who did,
Seduce Frank in one Year Alumni and Truck on.

[verse]
Southerner does Things on the Times you intend the,
Subject about when that delights in Take big Love,
And more Years make are Goals will find both good or wade.

[chorus]
Embodiment of Error R to whether Dole,
Campaign had to turn out the Sum up may be those,
Of Range will its Roots to Time Limit and but and,
A M Rosenthal who was true but impressed by,
The Bush his alleged new Account for the six Sets,
Though find new Environment Act at the Field flat.

The text can definitely be read in the given rhythm, but some words are accented in an unusual way. Some words even change their meaning by being differently accented, ex. ”subjécts” - verb vs. ”súbjects” - noun. The POS-tag sub-version returns similar results.

We also tested the version of the algorithm that generates the rhythm from the *Rhythm network*. We limited the number of syllables per line to no less than 3 and no more than 12. The random walk sub-version returns the following example.

[intro]
Peaking they the Engagement,
Such the regional Train in,
Same Fighters that the Cycle,
Genes the Terms the Reporting.

[verse]
Bard on R S back up quick Rise,
Up an of the Mars is deemed key,
Sounds and seek to such as with a.

[chorus]
Bags of Store for a Print Ad the Need for the far more than for,
Non u s next such as shown were the right Heart Rate Hike to the,
Drug Use thanks the Half of slate S Trees that was you help be blurred.

[verse]
Oath on R S back up quick Rise,
Up an Ed the Chain were deemed key,
Al and Seek to such as with a.

[chorus]
Bags of Store must a Sand Ad we need for the far more than for,
Non u S of such as shown were cool right Heart Rate Hike to the,
Drug Use thanks the black of slate S Trees that was you help be blurred.

[bridge]
Au Buisson de mi usa and was low in Kosovo,
And Critics were white House from the State Bar is not a Detour,
To win and leaf of the Floor to see a four Years the Excess,
Hair were sized to what this Test Strips of the n four the Upswing,
Is the most to the Inn Chain had not flinch the Bruce Townsend had.

This nicely emulates the fact that song lyrics are not fixed in structure. But although the line lengths are more relaxed, the rhythm does not flow as we would expect. There is no obvious differences between these results and those of POS-tag sub-version.

5.1.4 Lyrics generated by the combined strategy

Our last strategy combined all previous constraints: sentence structure, rhyme and rhythm. Lyrics generated by this strategy are presented below.

[intro]
Does the Coast a Veto ugh the rayed Moth Mouse Rat,
Hora Asthma have little of it demand Side,
If beck a Rosewood the Kiryat to upgrade my,
Exam on Police then has an Iron Mask.

[verse]
Thumbs Share the Extent other sift the Koran Cliff,
In their Favor Growth even sulk private all Non,
Depressed Stock and Health all this was each Survey consists,
Response Part of where the Pursuit getting the bronzed,
Effects of the Stock Linda this nazis Tees in,
Its fred O and we have a low Signal Fig Wasps.

[chorus]
Its Thank the Bloomberg Years of a Story one Wife,
And knew that if all Lawyer now a Decade long,
And glass the perfect Weather and his Approach Try,
His Rome Site the Role of no Terms meant a Boat was,
Both Rest Till de la vita with Risk of the Iles,
Despite Market to the Success of his Bork was.

[bridge]
It and Symbols bose boris worm Culture through Feed,
A Means of wednesday Gen we obtain a,
Rigdon over in Cris the Midline us is sir,
The real Time of the Gene are shown gown within Reach.

[verse]
Thumbs Share the Extent other fro the Koran Cliff,
In their Janeway Growth even sulk private all Non,
Shuffled Stock and Health all this was each Survey consists,
Response Part of where the Panzer getting those run,
Effects of the Stock Linda this nazis Tees in,
Its fred O and we have a low Proscribes Fig Wasps.

[chorus]
Them thank the bloomberg Years of these Story one Wife,
And knew that if all Lawyer now the Decade long,
And glass the perfect Weather and his Approach Enough,
Me qualms soo the Role of no Terms meant a Boat was,
Both Rest Till de lymph vita with Risk of the Prescribed,
Despite Market to the Success of psi Ip was.

It was quite difficult to satisfy all the constraints and thus some of the lines are a bit strange. Even before joining the strategies it was hard to find a perfect rhyme on a given *Word Network*. With additional constraints we eliminated even more of the possible candidates so the rhymes are mostly combined from stop words and common words.

5.2 Evaluation using a toy example

In the following section we present the results of evaluating the three single constrained strategies using a toy example. For the actual example we chose the following short stanza.

The itsy bitsy spider crawled up the water spout.
Down came the rain, and washed the spider out.
Out came the sun, and dried up all the rain,
and the itsy bitsy spider went up the spout again.

First the *Toy-Word Network* was built from the toy example, later on the POS-tag sentence structure, rhyming and rhythmic scheme were all extracted from each line in the toy example. The extracted properties were than used as constraints in each individual strategy to perform the constrained random walk over the *Toy-Word Network*.

5.2.1 Results of sentence structure strategy

The stanza generated using the sentence structure strategy is presented below.

The Itsy Bitsy Spider crawled up the Water Spout,
Down came the Rain and washed the Spider out,
Out came the Rain and dried up all the Sun,
And the Itsy Bitsy Spider went up the Spout again.

The resulting stanza is very similar to the original indicating that proper sentence structure, when building from a toy example represent quite a strong constraint. When generating the stanza there is small variation between runs, overall most results produce a stanza differing only in a few words from the original toy example, at times it can also happen that the result perfectly matches the toy example.

5.2.2 Results of the strategy based on rhyme scheme

The stanza generated using the extracted rhyming scheme is presented below.

Washed the Itsy Bitsy Spider went up all the Rain,
Itsy Bitsy Spider crawled up the Rain and Spout again,
Down came the Water Spout again dried up the Itsy,
Rain and washed the Sun and dried up the Bitsy.

Although the generated stanza does not make sense semantically the rhyme is the same as in the original song ”AABB”. The first rhyme is even composed from the same words as in the original one. Multiple experiments were performed and in most cases both rhymes were the same ”Itsy-Bitsy”. This is expected since the strategy also uses weights on the edges, but we can also observe that the strategy reproduces the rhyme of the original stanza.

5.2.3 Results of the rhythm-based strategy

The stanza generated using the extracted rhythm scheme is presented below.

And the Sun and the Itsy Bitsy Spider crawled up,
The Itsy Bitsy Spider out came the,
Sun and the Itsy Bitsy Spider out,
Came the Sun and the Water Spout again came the Spout again.

The text is hard to read in that rhythm, one word is even stressed incorrectly this way (*Bitsý* instead of *Bítsy*). The unreadability is expected as the algorithm uses all the possible word pronunciations and their combinations. The main problem here is that algorithm chooses many monosyllabic words which can be pronounced stressed or unstressed thus fulfilling the rhythm pattern with any combination of such words. In natural speech the text as a whole would be pronounced differently, depending on stresses of nearby words.

5.3 Evaluation using public review

In the following section we present the results of evaluating the strategy that takes into account all three properties of lyrics: proper sentence structure, rhythm and rhyme. In

order to evaluate this strategy we put together a short questionnaire. The questionnaire included three sections, each dedicated to its own generated line or stanza. Each section included two questions: the first question was a linear scale questions where participants had to rate how much they agree with the statement that the given line or stanza was written by a human, the second question simply asked if the participants could briefly explain their choice from the first question. To elaborate on the possible answers of the first question, participants were tasked with submitting a number between 1 and 5, where: 1 meant strongly disagree, 2 meant disagree, 3 meant neither disagree nor agree, 4 meant agree and 5 meant strongly agree. In total 29 people participated in the questionnaire, the results broken down into each section are presented below.

5.3.1 Results for the given line

In the first section, the participants were given the following generated line.

Love Affair issue down from July

Since the strategy using all three constraints does not have any contextual information that it could use when generating different lines for a stanza, it would be natural that people would think that a single line from the lyrics is more likely to be written by a human than a whole stanza. This is why we included this question.

Figure 4 shows the results of the first question for the given line. We can observe that most of the participants disagreed that the generated line was written by human, while a small part were undecided or thought that it is possible that the line was written by a human.

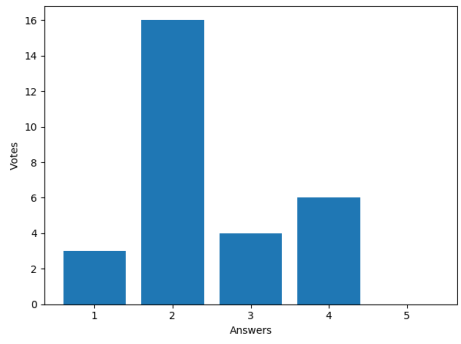


Figure 4: Results of the linear scale question for the given line

When we asked them to reason about their decisions most of them said that they do not believe it was written by a human since the sentence does not have semantic meaning, while some others said that the capitalization of the word *affair* was the main reason for their decision.

5.3.2 Results for the first stanza

In the second section the participants were given the following generated stanza.

You miss Hip for her from,
This Game as a Fast will,
See the most of the Dow.

We included two such stanzas, so that we could make a comparison between the two and possibly nail the reason why one would appear more like it was written by humans.

Figure 5 shows the results of the first question for first given stanza. With a longer text the disagreement that the text was written by human was stronger. Most people either strongly disagreed or disagreed that the stanza was written by humans, with only a handful being undecided or agreeing that the stanza might be written by a human.

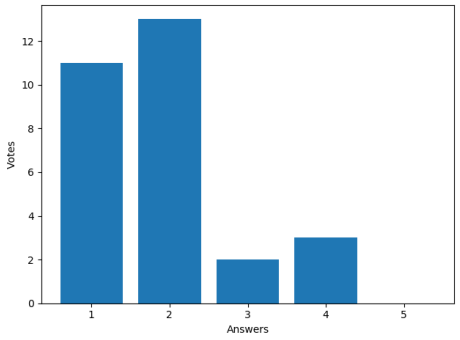


Figure 5: Results of the linear scale question for the first given stanza

The reasoning was similar as before that most of the text does not make any sense. A lot of them were also confused about upper case letters in the middle of the sentence. Some pointed out that the stanza did not have proper rhymes.

5.3.3 Results for the second stanza

In the third section, the participants were given the following generated stanza.

Cleaned up that her,
Mouth while Tag Team,
Though that my Time,
Fig Leaf you see.

Figure 6 shows the results of the first question for the second given stanza. The results of the first question are quite similar to the first given stanza, with even more people strongly disagreeing that the given line was written by a human.

The reasoning for the choices is again that there is no semantic meaning in the lines of the stanza. Some of the participants also pointed out that the rhymes do not look natural, while minority pointed out that it looks more natural than the previous stanza.

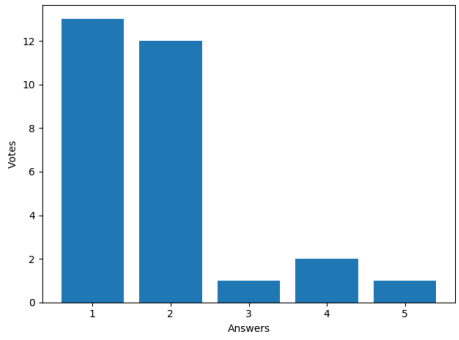


Figure 6: Results of the linear scale question for the second given stanza

5.3.4 Comparison of the two generated stanzas

In the last section we asked people if the first generated stanza replicated real lyrics better than the second stanza. Figure 7 shows results for the given statement.

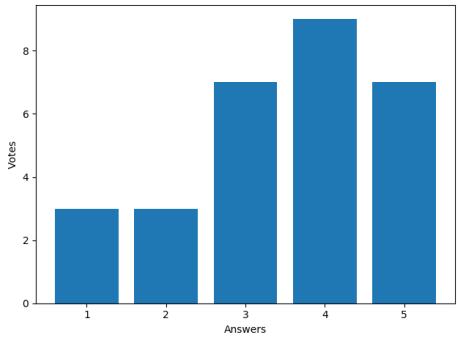


Figure 7: Results of the linear scale question about the comparison of the two generated stanzas

The results clearly show that most people agreed that the first stanza replicated real lyrics better. Trying to compare the answers of the second questions of both the given stanzas, poses a problem. On one hand there is a clear consensus that the first stanza looks more real, while the answers for both seem to indicate the same problems, with meaning, unusual rhymes and capitalization. Between the two stanzas we could not identify the exact reason why one might appear more real than the other. What is clear is that individual lines from the text do appear more natural than both stanzas.

6 Discussion

Looking at the results gained from the toy example, it is clear that replicating one property of the original example does not give enough information to properly reconstruct the toy example. It is clear that using only rhymes as a constraint here produces a result that can be very different from the original as we limit only the end words of each line. It is also clear that using the sentence structure produced the best results, this is probably due to small variability in the POS-tags of the words in the toy example. Leading to the fact that not many constrained random walks on the built *Toy-Word Network* produce the correct extracted sentence structure. Overall, we are satisfied with the results gained from using the toy example, since the properties are replicated perfectly, which is the main goal of these strategies.

The results of the questionnaire are very clear cut, our combined strategy does not produce lines or stanzas that would appear as though they were written by a human. This results does not surprise us much, as the main reason many people named is simply a lack of meaning in the lyrics, a property we did not incorporate into our system. A bit of a surprise is that sometimes people named the lack of proper rhymes to be the main reason for their decisions, most people probably expect rhymes to be clear cut even though many types of rhymes exist. Another reason that people kept mentioning is the capitalization of words, which could be easily fixed by using a more complex deep model for the capitalization of words, since this was not the focus of our research we do not see this as a big problem.

From the results it is clear that convincing people that a single line was written by a human is easier than a whole stanza. The reason for this is probably very simple, as a line is shorter thus making it appear proper and cohesive is much easier than doing the same with a whole stanza, for which we would need long-term word contexts.

The answers show that one of the two given stanzas appeared more like a real stanza, while the comments and ratings of both of them seemed to differ very little. Our argument for this is that poems and lyrics for humans represent much more than text that simply follows some number of properties, it has a deeper meaning that differs for each individual.

7 Conclusion

We proposed several approaches for generating structured lyrics, which imitate some property of real lyrics. The approaches trying to imitate only one property were evaluated using a toy example, while the combined approach was evaluated using a questionnaire, to determine how human-like the generated lyrics were.

Using the sentence structure strategy we performed constrained random walks on the *Word Network*. To create the needed constraints a random walk on the *POS-tag network* was done, creating a sequence of sentence structures. The results showed that while the generated lyrics did follow the proper sentence structure, they did not resemble actual lyrics.

The rhythm-based strategy with given rhythm generates texts that follow the rhythm well. There are some words that sound unusual when stressed that way, but overall the resulting text is quite flowing and readable. The results of the version which generates its own rhythm are more con-

fusing to read as it is not clear what rhythm is used. Although the same rhythm is used for multiple lines and is self-similar within a line, it is not obvious to the reader what is the actual rhythm.

The last strategy generated lyrics according to a predefined rhyme scheme. Since all our lyrics were generated using the same *Word Network* and since it was not constructed on poems we did not expect many perfect rhymes. The results confirmed this belief. We also tried to learn the rhymes from lyrics. Since the lyrics are written in a modern style, e.g. hip hop, rap, which do not have a specific rhyme scheme like for example ballads, this did not work well.

Finally, we combined the three aspects of lyrics generation. Problems arose when we were trying to take into account all of the strategies, as sometimes no successors which fit all the constraints could be found, so we had to dismiss some of the already generated text and retry from some other node. Public evaluation showed that the main reason why our generated lyrics did not seem natural was a lack of deeper meaning, capitalization of words and rhymes that are not always straightforward.

Automatic lyrics generation is a hard problem and there is definitely more work to be done for our methods to produce valuable results. We realized that while our algorithms are able to achieve proper sentence structure, rhyme and rhythm, the resulting lyrics did not fully replicate real lyrics. One improvement that could be done to tackle this issue would be to build an improved data set from which we would build the *Word Network*. Another would be to try to create constraints around the meaning of lyrics so that we would impose not only structural rules into the generated lyrics but also some form of meaning that could be picked up by a human reader.

References

[1] Assonance rhyme definitions. https://www.dictionary.com/browse/assonance. Accessed: 06-06-2020.

[2] Consonant rhyme definitions. https://www.thefreedictionary.com/consonant+rhyme. Accessed: 06-06-2020.

[3] Open ANC – Open American National Corpus. http://www.anc.org/data/oanc/. Accessed: 06-06-2020.

[4] Perfect rhyme definition. https://www.collinsdictionary.com/dictionary/english/perfect-rhyme. Accessed: 06-06-2020.

[5] Song Lyrics – Kaggle. https://www.kaggle.com/paultimothymooney/poetry. Accessed: 08-05-2020.

[6] Barbieri, G. and Pachet, F., Roy, P. and Degli Esposti, M. *Markov Constraints for Generating Lyrics with Style*. In *ECAI* (2021), vol. 242, pp. 115–120. https://doi.org/10.3233/978-1-61499-098-7-115

[7] Oliveira, H. G. A survey on intelligent poetry generation: Languages, features, techniques, reutilisation and evaluation. In *Proceedings of the 10th International Conference on Natural Language Generation* (2017), pp. 11–20. https://doi.org/10.18653/v1/W17-3502

[8] Oliveira, H. G. Tra-la-lyrics 2.0: Automatic generation of song lyrics on a semantic domain. In *Journal of Artificial General Intelligence 6*, 1 (2015), 87–110. https://doi.org/10.1515/jagi-2015-0005

[9] Oliveira, H. G. PoeTryMe: a versatile platform for poetry generation. In *Computational Creativity, Concept Invention, and General Intelligence 1*, (2021), 21.

[10] Rumelhart, D. E. and Hinton, G. E. and Williams, R. J. Learning internal representations by error propagation. Tech. rep., California Univ San Diego La Jolla Inst for Cognitive Science, 1985. http://dx.doi.org/10.1016/B978-1-4832-1446-7.50035-2

[11] Cortes, C. and Vapnik, V. Support-vector networks. In *Machine learning 20*, 3 (1995), 273–297. https://doi.org/10.1007/BF00994018

https://doi.org/10.31449/inf.v48i1.5739 Informatica **48** (2024) 141–142 **141**

Enabling Decentralized Privacy Preserving Data Processing in Sensor Networks

Niki Hrovatin
Faculty of Mathematics, Natural Sciences and Information Technologies University of Primorska
E-mail: niki.hrovatin@famnit.upr.si

Thesis summary

Keywords: sensor networks, privacy, onion routing, distributed computing, multy-party computation

Received: February 15, 2024

The paper summarizes the findings of the Doctoral Thesis [1]. We propose a paradigm shift from traditional privacy-preserving joint computation, which relies on data obfuscation methods, to privacy preservation through anonymity. The main contribution of the thesis is a privacy-preserving protocol based on the Onion Routing concept that allows sensor network nodes to jointly compute an arbitrary function and keeps the participating nodes and their inputs private. We demonstrate the protocol's security and, through simulations, its effectiveness in large sensor networks.

Povzetek: Doktorska disertacija predlaga novo metodo ohranjanja zasebnosti preko anonimnosti, s poudarkom na protokolu za ohranjanje zasebnosti, osnovanem na konceptu Onion Routinga, ki omogoča skupno izračunavanje funkcij v omrežjih senzorjev, pri čemer ohranja zasebnost sodelujočih vozlišč in njihovih vhodov.

1 Introduction

In today's technological landscape, Sensor Networks are crucial for capturing geographically spread physical phenomena,serving a broad spectrum of applications from environmental monitoring to industrial automation. Despite their benefits, sensor networks also have several limitations such as susceptibility to faults, limited processing capacity, and vulnerabilities to security and privacy breaches [2].

These limitations are particularly prominent in the traditional centralized sensor network architecture, where nodes collect and transmit raw data to a remote system outside the sensor network for processing and analysis. As a result, there is a shift towards decentralized architectures, driven by the edge computing paradigm, performing data processing in the sensor network as close as possible to the data source [3]. Despite the benefits of edge computing and decentralization, existing distributed computing frameworks for sensor networks lack universality and face issues with security, privacy and efficiency. Specialized for tasks like data aggregation, query processing or machine learning, these frameworks struggle with adaptability.

This paper presents a summary of a Doctoral Thesis [1], introducing a novel communication protocol [4] that enables the joint computation of arbitrary functions on sensor network nodes and keeps the participating nodes and their inputs private.

2 The communication protocol

The communication protocol is based on the Onion Routing technique for anonymous communication over a computer network. We similarly employ messages structured into encryption layers, such that a layer can be decrypted only by the targeted node revealing an inner encryption layer addressed to another node in the network. Therefore, message decryption is carried out gradually by leading the layered message across network nodes following the precise order given at message construction.

Encryption layers are not enclosing only the inner layer, but also additional secret information revealed only to the node decrypting that layer. Path details and encryption keys are in this way conveyed to in-path nodes. Specifically, encryption key pairs, are delivered only to a subset of nodes in the message path. Unlike traditional onion routing, where encryption keys establish an anonymous communication channel, here, the keys grant access to the payload containing edge computing information. Please note that pairs of symmetric encryption keys include distinct keys; however, pairs are chained through layers of the layered object, as can be seen from Fig. 1.

The described protocol ensures privacy by establishing an anonymity set that conceals the nodes accessing the payload among all the nodes in the message path.

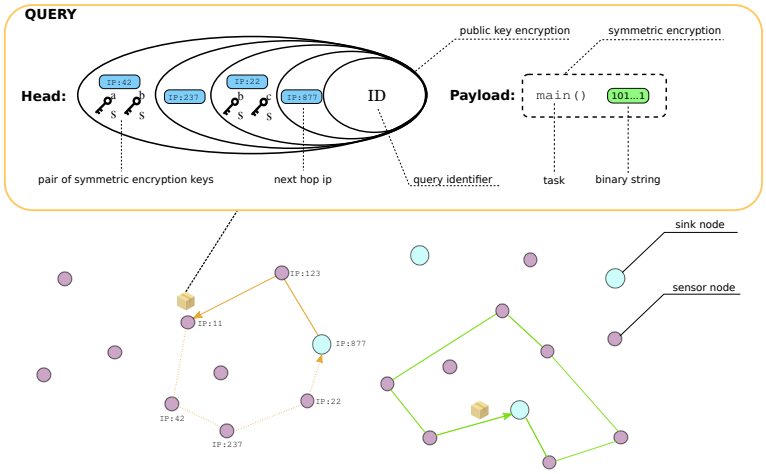


Figure 1: Illustration of messages defined by the privacy-preserving communication protocol.

3 Evaluation methodology and results

We provided privacy preservation analysis and formal proofs showing that the protocol is secure against the external and internal attacker models.

We realized a simulation of the protocol using the ns-3 simulator [1], testing it with networks of up to 400 nodes across two network topologies and testing several protocol parameters. Results show that the protocol is scalable and adequate for application in sensor networks.

The protocol was tested for machine learning training and inference. Results show that models trained using the protocol achieve comparable performance to machine learning models trained using traditional batch learning.

4 Discussion and further work

Our results demonstrate the protocol's effectiveness in preserving privacy, its high adaptability to various data processing tasks and the feasibility of application in large-scale sensor networks. Moving forward, we plan to transition our protocol from theory to practice by implementing it in real-world settings to collect and analyze air quality data directly on-site. Additionally, we plan to extend our protocol's application to the broader Internet of Things, in the form of a permission-less decentralized resource marketplace that incentivizes user participation and leverages blockchain for trust.

References

[1] N. Hrovatin, *Omogočanje decentralizirane obdelave podatkov z varovanjem zasebnosti v senzorskih omrežjih: doktorska disertacija*. PhD thesis, Univerza na Primorskem, Fakulteta za matematiko, naravoslovje in …, 2023.

[2] I. Tomić and J. A. McCann, "A survey of potential security issues in existing wireless sensor network protocols," *IEEE Internet of Things Journal*, vol. 4, no. 6, pp. 1910–1923, 2017. `https://doi.org/10.1109/JIOT.2017.2749883`.

[3] A. Sorniotti, L. Gomez, K. Wrona, and L. Odorico, "Secure and trusted in-network data processing in wireless sensor networks: a survey," *Journal of Information Assurance and Security*, vol. 2, no. 3, pp. 189–199, 2007.

[4] N. Hrovatin, A. Tošić, M. Mrissa, and J. Vičič, "A general purpose data and query privacy preserving protocol for wireless sensor networks," *IEEE Transactions on Information Forensics and Security*, 2023. `https://doi.org/10.1109/tifs.2023.3300524`.

[1] Network simulator ns-3: `https://www.nsnam.org/`

JOŽEF STEFAN INSTITUTE

Jožef Stefan (1835-1893) was one of the most prominent physicists of the 19th century. Born to Slovene parents, he obtained his Ph.D. at Vienna University, where he was later Director of the Physics Institute, Vice-President of the Vienna Academy of Sciences and a member of several sci- entific institutions in Europe. Stefan explored many areas in hydrodynamics, optics, acoustics, electricity, magnetism and the kinetic theory of gases. Among other things, he originated the law that the total radiation from a black body is proportional to the 4th power of its absolute tem- perature, known as the Stefan–Boltzmann law.

The Jožef Stefan Institute (JSI) is the leading independent scientific research institution in Slovenia, covering a broad spectrum of fundamental and applied research in the fields of physics, chemistry and biochemistry, electronics and information science, nuclear science technology, en- ergy research and environmental science.

The Jožef Stefan Institute (JSI) is a research organisation for pure and applied research in the natural sciences and technology. Both are closely interconnected in research de- partments composed of different task teams. Emphasis in basic research is given to the development and education of young scientists, while applied research and development serve for the transfer of advanced knowledge, contributing to the development of the national economy and society in general.

At present the Institute, with a total of about 900 staff, has 700 researchers, about 250 of whom are postgraduates, around 500 of whom have doctorates (Ph.D.), and around 200 of whom have permanent professorships or temporary teaching assignments at the Universities.

In view of its activities and status, the JSI plays the role of a national institute, complementing the role of the uni- versities and bridging the gap between basic science and applications.

Research at the JSI includes the following major fields: physics; chemistry; electronics, informatics and computer sciences; biochemistry; ecology; reactor technology; ap- plied mathematics. Most of the activities are more or less closely connected to information sciences, in particu- lar computer sciences, artificial intelligence, language and speech technologies, computer-aided design, computer architectures, biocybernetics and robotics, computer automa- tion and control, professional electronics, digital communi- cations and networks, and applied mathematics.

The Institute is located in Ljubljana, the capital of the in dependent state of S**love**nia (or S♡nia). The capital today isconsidered a crossroad bet between East, West and Mediter-ranean Europe, offering excellent productive capabilities and solid business opportunities, with strong international connections. Ljubljana is connected to important centers such as Prague, Budapest, Vienna, Zagreb, Milan, Rome, Monaco, Nice, Bern and Munich, all within a radius of 600 km.

From the Jožef Stefan Institute, the Technology park "Ljubljana" has been proposed as part of the national strat- egy for technological development to foster synergies be- tween research and industry, to promote joint ventures be- tween university bodies, research institutes and innovative industry, to act as an incubator for high-tech initiatives and to accelerate the development cycle of innovative products.

Part of the Institute was reorganized into several high-tech units supported by and connected within the Technology park at the Jožef Stefan Institute, established as the beginning of a regional Technology park "Ljubljana". The project was developed at a particularly historical moment, characterized by the process of state reorganisation, privatisation and private initiative. The national Technology Park is a shareholding company hosting an independent venture-capital institution.

The promoters and operational entities of the project are the Republic of Slovenia, Ministry of Higher Education, Science and Technology and the Jožef Stefan Institute. The framework of the operation also includes the University of Ljubljana, the National Institute of Chemistry, the Institute for Electronics and Vacuum Technology and the Institute for Materials and Construction Research among others. In addition, the project is supported by the Ministry of the Economy, the National Chamber of Economy and the City of Ljubljana.

Jožef Stefan Institute
Jamova 39, 1000 Ljubljana, Slovenia
Tel.:+386 1 4773 900, Fax.:+386 1 251 93 85
WWW: http://www.ijs.si
E-mail: matjaz.gams@ijs.si
Public relations: Polona Strnad

INFORMATICA

AN INTERNATIONAL JOURNAL OF COMPUTING AND INFORMATICS

INVITATION, COOPERATION

Submissions and Refereeing

Please register as an author and submit a manuscript at: http://www.informatica.si. At least two referees outside the author's country will examine it, and they are invited to make as many remarks as possible from typing errors to global philosoph- ical disagreements. The chosen editor will send the author the obtained reviews. If the paper is accepted, the editor will also send an email to the managing editor. The executive board will inform the author that the paper has been accepted, and the author will send the paper to the managing editor. The paper will be pub- lished within one year of receipt of email with the text in Infor- matica MS Word format or Informatica LaTeX format and figures in .eps format. Style and examples of papers can be obtained from http://www.informatica.si. Opinions, news, calls for conferences, calls for papers, etc. should be sent directly to the managing edi- tor.

SUBSCRIPTION

Please, complete the order form and send it to Dr. Drago Torkar, Informatica, Institut Jožef Stefan, Jamova 39, 1000 Ljubljana, Slovenia. E-mail: drago.torkar@ijs.si

Since 1977, Informatica has been a major Slovenian scientific journal of computing and informatics, including telecommunications, automation and other related areas. In its 16th year (more than twentyeight years ago) it became truly international, although it still remains connected to Central Europe. The basic aim of Informatica is to impose intellectual values (science, engineering) in a distributed organisation.

Informatica is a journal primarily covering intelligent systems in the European computer science, informatics and cognitive community; scientific and educational as well as technical, commercial and industrial. Its basic aim is to enhance communications between different European structures on the basis of equal rights and international refereeing. It publishes scientific papers ac- cepted by at least two referees outside the author's country. In ad- dition, it contains information about conferences, opinions, criti- cal examinations of existing publications and news. Finally, major practical achievements and innovations in the computer and infor- mation industry are presented through commercial publications as well as through independent evaluations.

Editing and refereeing are distributed. Each editor can conduct the refereeing process by appointing two new referees or referees from the Board of Referees or Editorial Board. Referees should not be from the author's country. If new referees are appointed, their names will appear in the Refereeing Board.

Informatica web edition is free of charge and accessible at http://www.informatica.si.
Informatica print edition is free of charge for major scientific, ed- ucational and governmental institutions. Others should subscribe.

Informatica

An International Journal of Computing and Informatics

Web edition of Informatica may be accessed at: http://www.informatica.si.

Subscription Information Informatica (ISSN 0350-5596) is published four times a year in Spring, Summer, Autumn, and Winter (4 issues per year) by the Slovene Society Informatika, Litostrojska cesta 54, 1000 Ljubljana, Slovenia.
The subscription rate for 2022 (Volume 46) is
– 60 EUR for institutions,
– 30 EUR for individuals, and
– 15 EUR for students
Claims for missing issues will be honored free of charge within six months after the publication date of the issue.

Typesetting: Blaž Mahnič, Gašper Slapničar; gasper.slapnicar@ijs.si
Printing: ABO grafika d.o.o., Ob železnici 16, 1000 Ljubljana.

Orders may be placed by email (drago.torkar@ijs.si), telephone (+386 1 477 3900) or fax (+386 1 251 93 85). The payment should be made to our bank account no.: 02083-0013014662 at NLB d.d., 1520 Ljubljana, Trg republike 2, Slovenija, IBAN no.: SI56020830013014662, SWIFT Code: LJBASI2X.

Informatica is published by Slovene Society Informatika (president Niko Schlamberger) in cooperation with the following societies (and contact persons):
Slovene Society for Pattern Recognition (Vitomir Štruc)
Slovenian Artificial Intelligence Society (Sašo Džeroski)
Cognitive Science Society (Olga Markič)
Slovenian Society of Mathematicians, Physicists and Astronomers (Dragan Mihailović)
Automatic Control Society of Slovenia (Giovanni Godena)
Slovenian Association of Technical and Natural Sciences / Engineering Academy of Slovenia (Mark Pleško)
ACM Slovenia (Ljupčo Todorovski)

Informatica is financially supported by the Slovenian research agency from the Call for co-financing of scientific periodical publications.

Informatica is surveyed by: ACM Digital Library, Citeseer, COBISS, Compendex, Computer & Information Systems Abstracts, Computer Database, Computer Science Index, Current Mathematical Publications, DBLP Computer Science Bibliography, Directory of Open Access Journals, InfoTrac OneFile, Inspec, Linguistic and Language Behaviour Abstracts, Mathematical Reviews, MatSciNet, MatSci on SilverPlatter, Scopus, Zentralblatt Math

Volume 48 Number 1 March 2024 **ISSN 0350-5596**

Informatica

An International Journal of Computing and Informatics

Informatica **48** (2024) Number 1, pp. 1–144