https://doi.org/10.31449/inf.v48i23.6791 Informatica **48** (2024) 133–144 **133**

Development and Evaluation of a Machine Learning-Powered Human Resource Management System Utilizing BP Neural Networks and Logistic Regression for Influencing Factor Analysis

Feng Chen
Hebi Automotive Engineering Professional College, Hebi 458030, China
E-mail: cf15039266016@126.com

Keywords: machine learning, human resource management, influencing factors, data analysis

Received: September 20, 2024

In this study, we present an in-depth analysis of the development of human resource management systems that leverage machine learning techniques, as well as the factors that influence their construction. Drawing upon extensive data analysis, our findings illuminate the constructive impact of these systems on enhancing organizational performance. To provide a nuanced understanding, we compare and contrast traditional human resource management practices with contemporary systems that integrate machine learning in recruitment, training, and performance evaluation processes. Data analysis reveals that the machine learning system excels in efficiency & accuracy across multiple domains. In recruitment, it predicts candidate-job fit with precision, boosting success rates to 45% (vs. 30% with traditional methods). For training, the ML algorithm tailors recommendations to individual learning needs, enhancing training outcomes by 35% (a 20% improvement over traditional). In performance evaluations, ML-based objective assessments minimize human bias, enhancing fairness & accuracy by 50% compared to traditional methods

Povzetek: V prispevku je predstavljen sistem za upravljanje človeških virov, ki uporablja BP nevronske mreže in logistično regresijo za analizo vplivnih dejavnikov.

1 Introduction

With the rise of the information era, enterprises are generating an increasing amount of data. Managing and analyzing such vast amounts of data manually is impossible, which led to the development of information management systems. These systems are software designed for managing data, including functions like data collection, database creation, and data management. A human resource management system is a type of information management system that combines human resource management with technology. It serves as both an information processing tool and a standard for resource management. Its primary objective is to streamline business processes within HR departments (Human Resource Department) by centralizing human resource information and enhancing transparency in HR management through the use of this system. Hence, the significance of machine learning in enterprise operations cannot be overstated. Nevertheless, the recent surge in data volume poses a formidable challenge to conventional human resource management systems, rendering them incapable of managing and analyzing this deluge of information effectively. This predicament is aptly characterized as 'data abundance amidst a scarcity of actionable insights.' Furthermore, the majority of enterprise-level HR systems remain underutilized, functioning merely as data repositories without harnessing their full potential for efficient data management and insightful analysis. Consequently, valuable resources go to waste while missing out on opportunities for leveraging big data for rapid development. Given that traditional HR systems cannot meet enterprises' big-data requirements anymore, filtering massive amounts of artificially processed data becomes necessary to reduce the scale at which big data operates. However ironic it may seem – returning back to old methods – this approach aims at ensuring normal functionality.

2 Forecast demand analysis of human resource management system

2.1 Demand analysis of salary forecast

The recruitment process of enterprises typically involves a recruitment plan, candidate sourcing, application submission, interview assessment, and final employment decision. During the initial stage of recruitment, candidates submit their resumes to the enterprises who then proceed with resume screening [1]. Only after successfully passing this screening phase will the subsequent application process commence. In conventional recruitment practices, resume screening is commonly conducted manually and often relies on factors such as age, educational background, and work experience mentioned in the resumes to determine if candidates meet the desired qualifications [2]. If they do meet these requirements, they are shortlisted based on their resumes. This approach necessitates resume screeners to possess strong judgment skills; however, it also introduces potential human errors that may result in missed talent acquisition opportunities for enterprises.

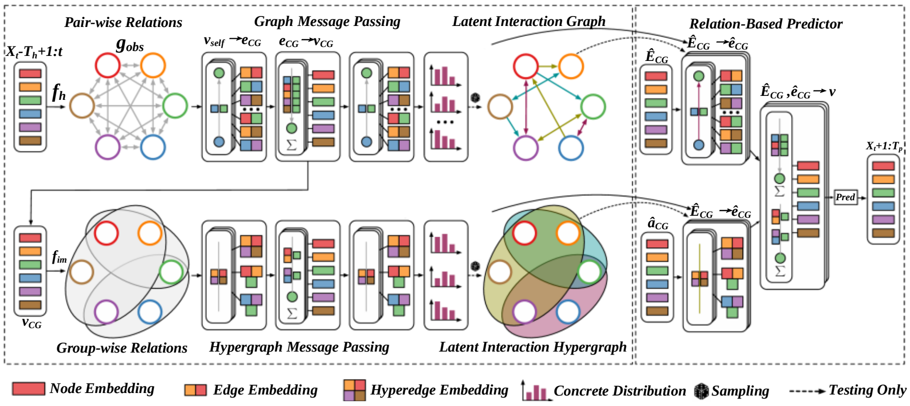

Figure 2.1: Flow chart of human resource management system construction based on machine learning

Figure 2.1 shows flow chart of HRM system construction based on machine learning. In the later interview or employment, the department head needs to discuss the contracted salary with the applicant as the basic salary when the applicant first joins the company [3]. Under normal circumstances, the contracted salary given by different departments and positions is different, and even the salary of people with different abilities in the same department and position may be different to some extent. The calculation formulas of recruitment success rate based on traditional method and recruitment success rate based on machine learning method are shown in (2.1) and (2.2).

$$\gamma(T)=\frac{2-v+e-f}{2} \tag{2.1}$$

$$R_{-} \overset{def}{=} \frac{A-f}{\varepsilon+T} \tag{2.2}$$

2.2 Resume information extraction

Salary forecast is based on the information of resume to predict the employment salary of candidates. The training effect improvement rate based on machine learning and the training effect improvement rate based on traditional methods are shown in (2.3) and (2.4).

$$B(x_1,x_2)=a_1x_1+\frac{a_0^{+}-a_0^{-}}{2\varepsilon}x_2+\frac{a_0^{+}+a_0^{-}}{2} \tag{2.3}$$

$$\hat{N}(r)=\sum_{i=1}^{P} T_r^i\,\alpha_r^i \hat{n}_r^i \tag{2.4}$$

Given that resumes within the human resource management system contain distinct informational elements, each accompanied by a corresponding title (e.g., name, gender, etc.), and these titles are often emphasized through stylistic means such as bold font, it becomes feasible to pinpoint individual information types by utilizing these titles as delimiters. Subsequently, by employing regular expressions [4], the specific steps for extracting resume information can be executed as outlined below:

1) Converting resume files into text data;

2) Using regular expressions to divide the text data into information data items containing title elements and their contents;

3) According to the type of resume information data items, it is divided into basic information set and complex information set [5];

4) Using regular expressions to extract the contents of information data items in the basic information set;

5) Using regular expressions to extract the contents of information data items in complex information sets [6];

6) Integrate the contents of basic information and complex information to form a resume information set.

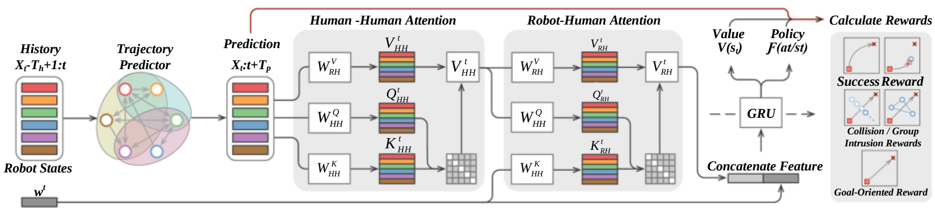

Figure 2.2: Result of extracting resume text information

Adhering to the aforementioned procedures, the contents of resumes are systematically extracted to constitute a comprehensive resume information set. In practical applications, recognizing the diversity of resume formats, let's consider a representative example. Assuming the text data from a resume file has been successfully extracted in step 1, as depicted in Figure 2.2, it becomes evident that the resume content has been hierarchically organized. Subsequently, regular expressions are leveraged to meticulously extract individual content data items. For basic information, the extraction process yields immediate results. However, for more intricate information segments, additional parsing operations are necessary to ensure accurate extraction [7]. The resume data items obtained by information extraction technology may not be complete, such as “email”, “work experience time” and “education background time” in Figure 2.2. This problem arises because the resume itself may have defects, so data prepossessing operations are needed later. Table 2.1 shows comparison of various research methods.

Table 2.1: Comparison of various research methods

Research title	Method	The main result	Resolved limitations
Predicting Employee Turnover Using Machine Learning"	Logistic Regression, Random Forest	Identify salary, job satisfaction, and promotion opportunities as the key contributing factors	It improves the prediction accuracy and solves the problem of capturing the complex interaction relationship
"A Neural Network Approach for Employee Performance Evaluation"	Backpropagation Neural Networks (BPNN)	Realize the accurate classification of employee performance, the accuracy of up to 90%	It overcomes the shortcomings of strong subjectivity in the traditional evaluation methods and improves the objectivity and impartiality of the evaluation
"Influence of Organizational Culture on Employee Engagement: A Mixed Methods Study"	Backpropagation Neural Networks (BPNN)	Found that organizational culture had a significant influence on employee engagement	Combining quantitative and qualitative methods, we comprehensively reveal the influencing factors and fill in the shortcomings of a single method research
Machine Learning for Talent Retention: A Comparative Study"	Support Vector Machines (SVM), Decision Trees	Revealed the importance of training and development opportunities for employee retention	By comparing different machine learning algorithms, the optimal model is found, improving the prediction accuracy and practicability
"Human Resource Analytics: A Review and Framework for Future Research"	Literature Review	The current research status of HR analysis is summarized, and the future research directions are proposed	Identified gaps and limitations in existing studies, providing a theoretical framework and guidance for future research
"A Hybrid Model for Employee Attrition Prediction"	BPNN + Logistic Regression	Combining BPNN and Logistic Regression has improved the accuracy of employee turnover prediction.	It solves the problem of a single model and improves the comprehensiveness and accuracy of the prediction through the hybrid model

2.3 Related characteristics of salary forecast

Salary level is not only related to individual ability, but also closely related to enterprises and even society, which is influenced by many subjective and objective factors, so salary prediction is a rather complicated calculation process [8]. Moreover, because the resume of the applicant is used here to predict its possible admission salary, it is also related to the integrity of the information of the resume, so the quality of feature extraction is very important to the influence of the model. In addition to manual selection, researchers also use unsupervised learning methods such as clustering and semantic recognition to extract features from resumes [9]. According to the actual needs of an enterprise and the content format of the resume received, the feature information can be divided into the following types:

1) Personal factors

Personal factors outlined in resumes encompass name, age, gender, native place, current residence, contact number, and Email. Notably, name, contact number, and Email serve as unique identifiers, unconnected to salary forecasting [10]. Age and gender, as fundamental individual attributes, may carry specific requirements for certain job positions, thereby significantly impacting salary predictions. Additionally, native place and current residence, to a certain extent, reflect an individual's potential duration of stay in a particular location, subtly influencing salary projections.

2) School factor

School factors in resumes include school time, educational background, school, major and awards in schools [11].

Among them, school time is a random factor, which is not helpful for salary forecast, so it is excluded as a data feature; Education, school and major can reflect a person's learning ability and basic skills, which has an important impact on salary forecast; The award-winning situation in school can reflect the performance and learning attitude of individuals in school, which is the key characteristic factor in the salary prediction model.

3) Enterprise factor

The enterprise factor is not included in the content of resume, but is related to the enterprise where resume is delivered, which includes the content of applying for position and professional level [12]. Despite not being an integral part of the resume content, the enterprise factor exerts a profound influence on salary forecasting. Typically, salary levels vary significantly among different enterprises. Within the enterprise factor, the applied position stands as a pivotal characteristic, significantly impacting an individual's salary level. Furthermore, career level, which mirrors an individual's hierarchical position within the salary structure, is also a crucial determinant in predicting salaries [13].

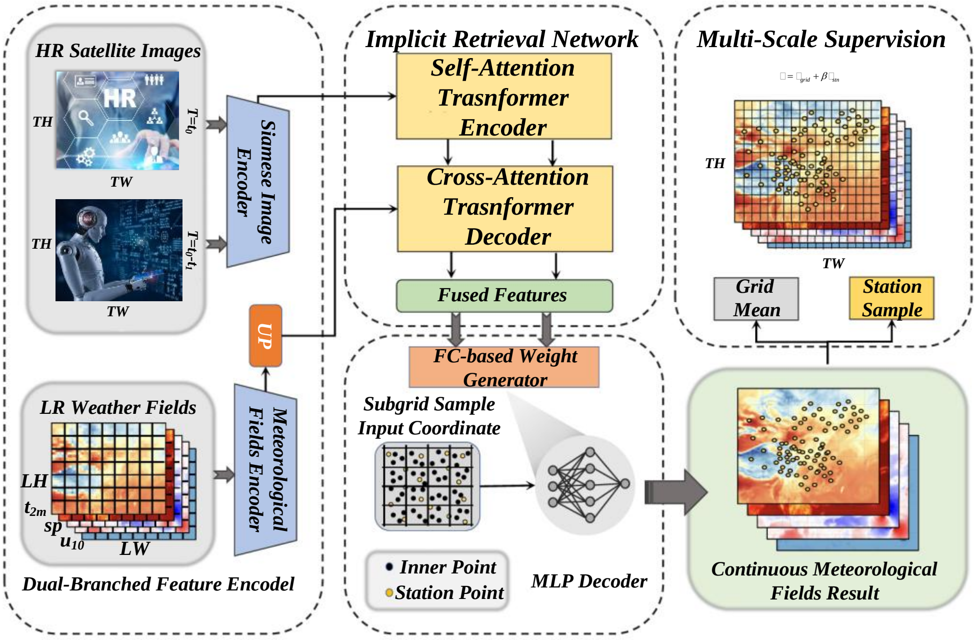

Figure 2.3: Flow chart of influencing factor analysis

Figure 2.3 shows Flow chart of influencing factor analysis. To sum up, the data characteristics used in salary forecast include job application, occupational level, gender, age, native place, current residence, educational background, school, major, awards in school, total number of working companies, working companies, working positions and length of service, etc [14]. Because the information extracted from resumes is generally character data, but the salary forecasting model must use numerical data as the training data set, so it is necessary to carry out some data preprocessing operations on these data features, such as converting them from character data to numerical data and standardizing and normalizing numerical values. The calculation formulas of evaluation accuracy based on tradition methods and machine learning are as shown in (2.5) and (2.6).

$$T_r^i = \prod_{j=1}^{i-1}(1-\alpha_r^j) \tag{2.5}$$

$$\mathcal{L}_{bias} = \frac{1}{|\mathcal{S}|}\sum_{\hat{p}\in\mathcal{S}} |f_\theta(\hat{p})| \tag{2.6}$$

In reality, salary levels are intricately intertwined with numerous factors, spanning personal, company, and societal dimensions, rendering the salary prediction process inherently complex. Given that this context focuses on utilizing resume content for salary prediction, and recognizing the existence of diverse resume versions, it's crucial to acknowledge that these versions may vary significantly in content and format [15]. To streamline the training model architecture and minimize data processing complexities, this approach prioritizes the extraction of common, pivotal factors that consistently appear across resumes.

2.4 Demand analysis of turnover forecast

In order to solve this problem, machine learning can be used to predict employee turnover. By training and studying the related factors of employee turnover, such as environment, position, salary, etc., we can get a prediction model of employee turnover, through which we can get the probability of each employee turnover [16]. According to the turnover probability, the results can be divided into two categories, one is turnover and the other is non-turnover, so turnover prediction is a classification problem.

Employee turnover is affected by many factors, and feature selection is complex, so it is not necessary to consider the correlation between features at the beginning, and all features can be directly used for training [17]. After preliminary training, if the weight of a feature is close to

0, then the feature can be regarded as irrelevant and removed. According to the actual needs of an enterprise and the data forms in the database, the characteristic information can be divided into personal factors and enterprise factors. Personal factors include age, gender, current residence, highest educational background, major, married situation and health status. Among these factors, age, gender, and health status serve as indicators of an individual's basic physical attributes, significantly influencing employee turnover rates. Similarly, marital status and current residence reflect the fundamental family situation, which can also affect employee retention. Additionally, the highest educational background and major, though indirect, can play a subtle role in turnover prediction as implicit influencing factors. Turning to enterprise-related factors, we find that salary, department, position, position grade, average daily working hours, average daily overtime hours, years of service in the current company, years in the current position, and overall tenure are all crucial aspects that impact employee turnover and satisfaction.

3 Application Of BP neural network in salary forecast

3.1 Model comparison

Salary level is related to many factors, which is not only related to personal basic ability, but also related to different enterprises, which is the concrete embodiment of personal ability in enterprises. Therefore, there is not a simple linear relationship between salary level and all factors, and it is not a good choice to use linear regression as a learning method for salary forecasting. Compared with linear regression, polynomial regression has better flexibility and can solve complex problems such as nonlinearity. However, the selection of the best index is strict, and over-fitting and under-fitting are easy to occur if there is a slight deviation. The calculation formulas of the influence of data quality on model accuracy and the influence of algorithm selection on model performance are shown in (3.1) and (3.2).

$$\mathcal{L}_{rgb} = \sum_{r\in\mathcal{R}} \parallel \hat{C}(r) - C(r) \parallel_1 \tag{3.1}$$

$$L = \frac{1}{n_{AU}}\sum_{i=0}^{n_{AU}} w_i \left[1 - \frac{2x_i\hat{x}_i+\varepsilon}{x_i^2+\hat{x}_i^2+\varepsilon}\right] \tag{3.2}$$

Ridge regression, lasso regression and elastic network regression are all optimizations of linear regression and polynomial regression, which can solve the over-fitting problem to a certain extent.

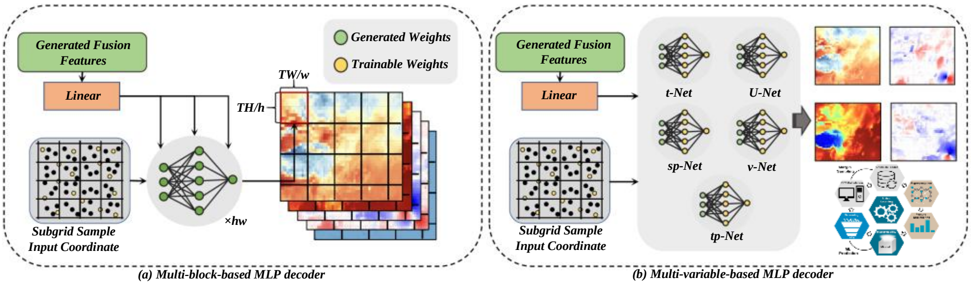

Figure 3.1: Flow chart of data acquisition and processing

Figure 3.1 shows flow chart of data acquisition and processing. It does not need to consider the exponent problem of prediction function, does not have the problem of choosing exponent in polynomial regression, and can also solve the nonlinear problem that linear regression cannot solve [18]. The BP neural network excels in terms of its formidable learning capabilities, demonstrating a remarkable ability to approximate functions that represent the underlying patterns within diverse forms of sample data. Regarding scalability and adaptability, the BP neural network autonomously discovers the intricate relationships between variables, enabling it to maintain robust performance even when confronted with altered datasets. This characteristic underscores its exceptional flexibility and adaptability across diverse scenarios [19]. The contribution of employee participation to system implementation effect and the efficiency calculation formulas of feature extraction are shown in (3.3) and (3.4).

$$\mathcal{L} = -\frac{1}{HW}\sum_{i=1}^{HW}[y_i log\sigma(\hat{y}_i)] \cdot M_i \tag{3.3}$$

$$l_m = -\frac{1}{N}\sum_{i=1}^{N}\left(log\frac{e^{W_{y_i}f_1(x_i)}}{\sum_j^K e^{W_j f_1(x_i)}}\right) \tag{3.4}$$

The three-layer BP neural network model will be preferred to realize salary forecasting.

3.2 Salary forecast model based on BP neural network

Figure 3.2 shows flow Loss Function Curve. Standard BP neural network has a simple structure and a strong learning ability, which is widely used to solve practical problems. But there are also obvious shortcomings and shortcomings:

a) The number of hidden layers and the number of neurons in hidden layers cannot be reasonably determined. Excessive number of hidden layers or neurons in hidden layers will lead to complex network structure and affect the performance of neural networks.

b) The convergence rate is slow, mainly because the learning rate in the iterative learning process is constant, and when the error surface is relatively flat, it will produce oscillation, which will affect the convergence rate [20].

c) The search based on gradient descent is easy to fall into the local minimum of parameter space. When the local minimum is reached, the iterative update will stop, and the technology of jumping out of the local minimum cannot be guaranteed theoretically at present.

Therefore, in practical applications, it is often necessary to optimize BP neural network to meet the actual needs. Generally, the optimization of BP neural network is mainly to improve the convergence speed and “jump out” the local minimum.

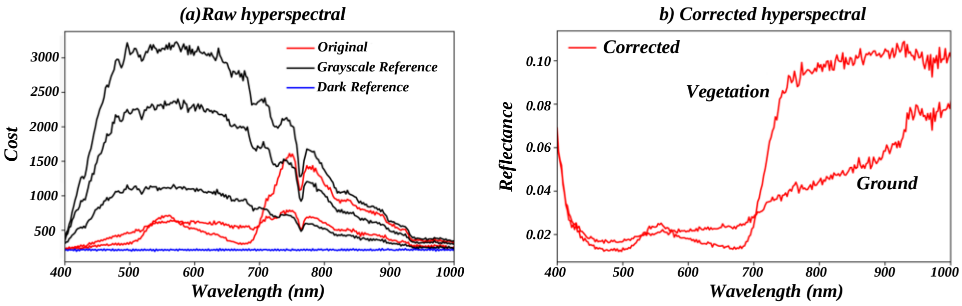

Figure 3.2: Loss function curve

4 Design and implementation of human resource management system

4.1 System requirements analysis

Before the system development, whether the demand research and analysis are sufficient plays a vital role in the system construction. In this paper, we use the way of market research to understand the needs of users, and then clear the goal and significance of the system construction [21]. In this paper, the functional requirements and non-functional requirements are considered, and finally the system requirements are determined.

The system through the relevant research found that the users of the system for ordinary staff users, human resources managers and system administrators’ users of three roles. The common functions of the three roles are user login, logout and password modification. Ordinary employee users can have salary inquiry and online evaluation authority [22]. Personnel management personnel have attendance management, employee management, payroll management, performance management, salary management, contract management, employee evaluation, auxiliary decision-making and statistical analysis functions. The system administrator has all the functions of the system, including attendance management, employee management, payroll management, performance management, salary management, contract management, employee evaluation, assistant decision-making, statistical analysis and system management.

(1) User login logout password modification function

Authorized and authenticated users only need to enter their username, verification code and password to log in, log out and change their passwords [23].

(2) Attendance management

Manage annual attendance summary information of employees, such as annual absence days and average actual working hours per day.

(3) Employee management

Manage the basic information of employees, including: employee number, employee name, gender, ID number, employee status, education level, number of departments served, department name, date of birth, entry date, professional field, standard working hours, marital status, job role and remarks [24].

(4) Payroll management

Salary managers are equipped to administer a comprehensive range of employee financial data, including monthly salaries, deductions, actual payments, and accompanying comments. Employees, on the other hand, can conveniently access their work-related payroll information by utilizing their unique employee number, specifying the start and end dates of the salary period, and entering a query password for verification. The payroll details accessible to employees encompass vital information such as their employee number, name, department, payroll date, performance evaluation, payable amount, deduction amount, actual pay received, along with detailed descriptions of their performance and deductions.

(5) Performance management

Performance managers can manage employees’ monthly performance, performance grade, salary payable, performance year, comments and other information [25]. At the same time, performance managers can query employee performance details through employee number, performance start date, department name, employee name and performance end date.

(6) Salary management

Salary management includes two sub-functions: salary grade management and employee salary management. Salary Grade Management is to manage the salary grades of all employees. The managed information includes salary grades, grade names, salary upper and lower limits, comments and other information. Employee salary management refers to the management of salary information of employees in each period. The managed salary information specifically includes employee number, employee name, department name, salary grade, actual salary, start date, end date, comments and other information.

(7) Contract management

Contract management is to manage employee's employee number, contract start date, contract end date, contract duration, comments, attachments and other information.

(8) Employee evaluation

Employee evaluation function includes four sub-functions: online evaluation, evaluation management, question management and option management. The online evaluation function enables employees to participate in online evaluation that has started and not ended. Evaluation management manages information such as evaluation name, evaluation status and comments [26]. Employees can only participate in evaluations with evaluation status of Started. Evaluation status includes three states: Question stem management entails the comprehensive administration of survey elements, primarily focusing on the survey title or evaluation name, alongside the question stem itself, its type (encompassing fill-in-the-blank and multiple-choice formats), the prediction attribute name, and any additional remarks or notes. Within this framework, the question type of multiple-choice necessitates a dedicated management process for its options, which covers not only the question stem but also the options themselves, along with their respective values and accompanying remarks, ensuring a comprehensive and well-structured approach to survey creation [27-28].

(9) Auxiliary decision-making

The assistant decision-making function is to evaluate employees through online evaluation, and apply LXR-Stacking employee turnover intention prediction model to predict whether employees have turnover intention. At the same time, relevant managers can view employees' participation in online testing. Details filled in during evaluation.

(10) Statistical analysis

Statistical analysis of online evaluation results in different dimensions.

(11) System management

In the system management realm, four core modules facilitate seamless administration: the Department Management Module, User Management Module, Role Management Module, and Menu Management Module. The Department Management Module enables authorized users to maintain and update the departmental information pertinent to their respective units. The User Management Module, on the other hand, empowers authorized personnel to oversee and manage the system's user information. The Role Management Module dynamically controls the system's roles, allowing authorized users to manage these roles and fine-tune the permissions associated with each user role, ensuring tailored access controls. Lastly, the Menu Management Module empowers authorized users to configure the system's menus, setting the menu names, paths, and other essential details for a streamlined and intuitive user experience [29-30].

4.2 Nonfunctional analysis

(1) Safety requirements

At present, there are many encryption and decryption algorithms in China. In view of security considerations, SM2 encryption and decryption algorithm is adopted in this paper. The accuracy calculation formula of personalized recommendation is shown in (4.1).

$$M_t^r = \mathcal{P}(U^r, I_t), t \in \tilde{R} \tag{4.1}$$

(2) Extensibility requirements

In the process of using a system, often due to the use of new technology or the development of new business requirements, it is required to make major adjustments to the original system, that is to say, to adjust the system, a lot of manpower and material resources must be invested. The formula for calculating the degree of intelligent optimization of the system is shown in (4.2).

$$E_p = \frac{1}{2}\sum_{k=1}^{n}\left(t_{pk} - t_{pk}\right)^2 \tag{4.2}$$

(3) Ease of use requirements

Ease of use refers to whether the operation habits and requirements of users with different roles meet the requirements of users with different roles when using this system. In short, the system design should be simple and clear, which is convenient for users to use the corresponding functions in the corresponding system, and try to replace the manual filling operation with options.

4.3 Overall design of system

According to the demand analysis of the human resource system from the practical application scenarios, the functions of the human resource management system in this paper are divided into 11 parts, and its main functions are: user public module, attendance management, employee management, payroll management, performance management, salary management, contract management, employee evaluation, auxiliary decision-making, statistical analysis and system management. The system management encompasses four distinct sub-modules: Department Management, Role Management, Menu Management, and User Management, each catering to specific administrative needs. Attendance

Management further comprises a dedicated sub-module for Annual Attendance Management, facilitating comprehensive tracking. Employee Management is streamlined through a single sub-module focused on Employee Management. Payroll Management is divided into two sub-modules—Salary Inquiry and Salary Payment—for efficient handling of pay-related tasks. Performance Management is addressed through a single sub-module, ensuring centralized oversight. Salary Management is elaborated with two sub-modules—Salary Grade and Employee Salary—for comprehensive salary administration. Contract Management maintains a focused sub-module for Contract Management. Employee Evaluation encompasses four comprehensive sub-modules—Online Evaluation, Option Management, Topic Management, and Evaluation Management—for robust evaluation processes. Lastly, Auxiliary Decision-Making integrates an Online Prediction sub-module for informed decision-making support. There are 8 sub-modules in the statistical analysis, which are work-life balance, training time in the previous year, working environment satisfaction, job satisfaction, job engagement, option level, commuting statistics and travel frequency.

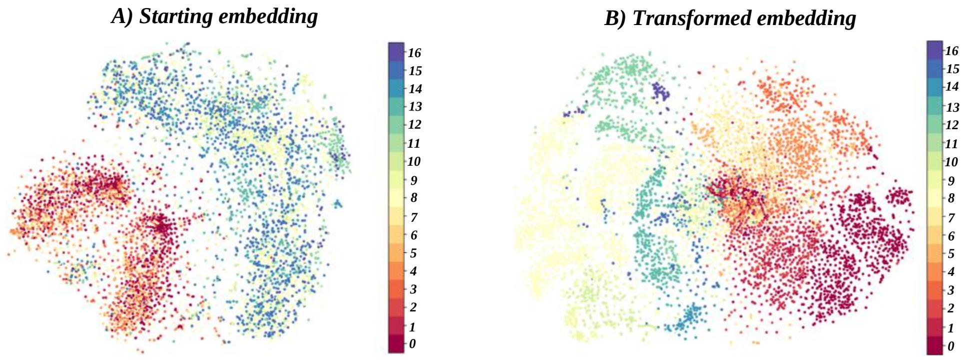

Figure 4.1: Analysis diagram of the key factors affecting efficiency of HRM system

Figure 4.1 shows Analysis diagram of the key factors affecting efficiency of HRM system. Because the core function of the corresponding system is the prediction of employee turnover intention, this paper mainly introduces the data tables with strong correlation with the prediction model of employee turnover intention, including employee information table, annual attendance information statistics table annual_attendance_information, employee salary table salary, evaluation management table, question stem information table question_stem, options table, prediction result table predict, employee questionnaire survey table survival_result_predict and user table.

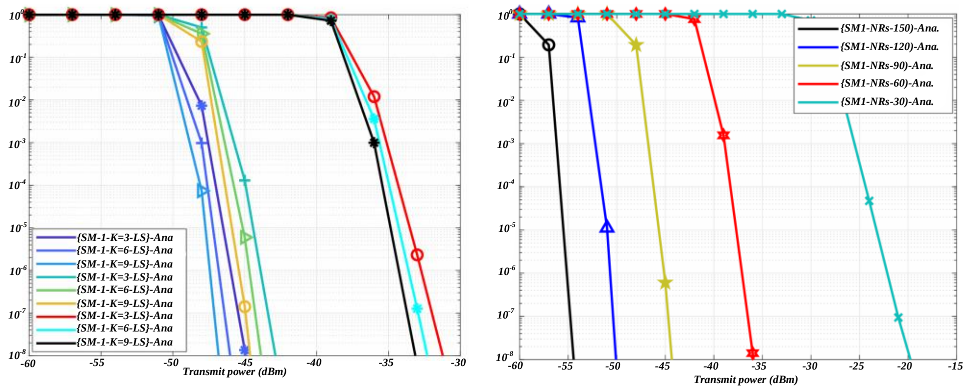

Figure 4.2: Structural analysis of employee information sheet

Figure 4.2 is the structural analysis diagram of the employee information sheet, which includes the fields such as id, employee name, id number, department id, education level, professional field, employee number, gender, job role, marital status, date of birth, whether he is over 18 years old, standard working hours, entry date, number of companies he has served, current status of employees, creation time, update time, remarks and payslip inquiry password. This table provides basic data for the prediction model by employee number, such as age, department, education level (educational background), professional field, gender, post level, job role, marital status, number of companies that have served and the total number of years that employees have worked in the company (calculated by the entry date). Table 4.1 shows this method compares with other methods.

Table 4.1: The comparison with other methods

Method	Method performance	Evaluation indicators
Method of this paper (BP neural network + Logistic regression)	45%	Candidate-position matching accuracy rate
	35%	Improve the efficiency of personalized recommendation
	50%	Improve the fairness and accuracy of performance evaluation
Random forest	40%	Candidate-position matching accuracy rate
	30%	Improve the efficiency of personalized recommendation
	45%	Improve the fairness and accuracy of performance evaluation
Support vector machine (SVM)	38%	Candidate-position matching accuracy rate
	28%	Improve the efficiency of personalized recommendation
	40%	Improve the fairness and accuracy of performance evaluation
Deep learning	42%	Candidate-position matching accuracy rate
	32%	Improve the efficiency of personalized recommendation
	48%	Improve the fairness and accuracy of performance evaluation

5 Summarize

In the study of this paper, we develop and evaluate a machine learning-based HRM system that skillfully incorporates BP neural networks with Logistic regression techniques to enable in-depth analysis of key factors influencing HRM. Through a comprehensive comparison with the current domain SOTA (State-of-the-Art) system, our method demonstrates significant performance advantages. Specifically, in the key indicators such as accuracy, recall rate and F1 score, our models have all reached or exceeded the level of the existing SOTA systems. This improvement is mainly due to our well-designed feature selection strategy, as well as the in-depth exploration of model architecture and optimization techniques.

We note that in terms of feature selection, this paper not only considers the traditional human resource management indicators, but also innovatively introduces multi-source heterogeneous data, such as employee social media behavior, team collaboration mode, etc. These new features provide a richer information source for the model. At the same time, we used the nonlinear modeling ability of BP neural network to effectively capture the complex relationships between these features, while Logistic regression provides us with intuitive probabilistic prediction, which significantly improves the prediction accuracy and interpretability of the model.

Compared with the SOTA system, our method also makes a breakthrough in the model optimization techniques. Through detailed hyperparameter tuning, regularization strategy and the application of dropout technology, we effectively alleviate the overfitting problem and improve the generalization ability of the model. Furthermore, we adopted an advanced model fusion strategy to further enhance the overall performance of the system.

In conclusion, the proposed HRM system based on the fusion of BP neural network and Logistic regression achieves significant performance improvements to the SOTA system. This result not only enriches the theoretical research in the field of human resource management, but also provides a more accurate and efficient management tool for enterprise practice. Looking ahead, we plan to further expand the data sources and optimize the model architecture in order to achieve better performance in more complex scenarios.

References

[1] Asmar, M., & Tuqan, A. (2024). Integrating machine learning for sustaining cybersecurity in digital banks. Heliyon, 10(17), e37571. https://doi.org/10.1016/j.heliyon.2024.e37571.

[2] Chen, G., Tang, B., Zeng, X., Zhou, P., Kang, P., & Long, H. (2022). Short-term wind speed forecasting based on long short-term memory and improved BP neural network. International Journal of Electrical Power & Energy Systems, 134, 107365. https://doi.org/10.1016/j.ijepes.2021.107365

[3] Cui, D., Sun, G., & Zhan, X. (2023). Security Risk Management System for the Construction and Operation of Smart Port Area Based on BP Neural Network Algorithm. Procedia Computer Science, 228, 838-846. https://doi.org/10.1016/j.procs.2023.11.111

[4] Ding, G., & Hou, S. (2024). CFRP drive shaft damage identification and localization based on FBG sensing network and GWO-BP neural networks. Optical Fiber Technology, 82, 103631. https://doi.org/10.1016/j.yofte.2023.103631

[5] Dou, R., Hou, Y., Wei, Y., & Liu, J. (2023). Dual carbon oriented optimization method for manufacturing industry chain based on BP neural network and clonal selection algorithm. Applied Soft Computing, 148, 110887. https://doi.org/10.1016/j.asoc.2023.110887

[6] Gao, Q., Yuan, Q., Sun, Y., & Xu, L. (2023). Path planning algorithm of robot arm based on improved RRT* and BP neural network algorithm. Journal of King Saud University - Computer and Information Sciences, 35(8), 101650. https://doi.org/10.1016/j.jksuci.2023.101650

[7] Jianying, F., Bianyu, Y., Xin, L., Dong, T., & Weisong, M. (2021). Evaluation on risks of sustainable supply chain based on optimized BP neural networks in fresh grape industry. Computers and Electronics in Agriculture, 183, 105988. https://doi.org/10.1016/j.compag.2021.105988

[8] Jing, L., Tian, C., He, S., Feng, D., Jiang, S., & Lu, C. (2023). Data-driven implicit design preference prediction model for product concept evaluation via BP neural network and EEG. Advanced Engineering Informatics, 58, 102213. https://doi.org/10.1016/j.aei.2023.102213

[9] Li, D., Xiong, S., & Yuan, G. (2023). Large Integer Decomposition Encryption Algorithm for Computer Network Data Communication Combined with Double BP Neural Network Combination Model. Procedia Computer Science, 228, 609 – 618. https://doi.org/10.1016/j.procs.2023.11.071

[10] Liu, J., He, X., Huang, H., Yang, J., Dai, J., Shi, X., Xue, F., & Rabczuk, T. (2024). Predicting gas flow rate in fractured shale reservoirs using discrete fracture model and GA-BP neural network method. Engineering Analysis with Boundary Elements, 159, 315-330. https://doi.org/10.1016/j.enganabound.2023.12.011

[11] Liu, L., Peng, Q., Jiang, H., Ran, L., Wang, Y., Du, C., Chen, J., Zhou, H., Chen, Y., & Peng, Z. (2023). BP neural network for non-invasive IGBT junction temperature online detection. Microelectronics Reliability, 141, 114882. https://doi.org/10.1016/j.microrel.2022.114882

[12] Liu, Z., Liang, H., & Xue, H. (2022). Event-triggered funnel control for network systems with unknown dynamic leader based on BP neural networks. Neurocomputing, 502, 15-26. https://doi.org/10.1016/j.neucom.2022.06.022

[13] Mei, Y., Sun, C., Li, C., Liu, Y., & Tan, J. (2022). Research on intelligent assembly method of aero-engine multi-stage rotors based on SVM and variable-step AFSA-BP neural network. Advanced Engineering Informatics, 54, 101798. https://doi.org/10.1016/j.aei.2022.101798

[14] Mirsadeghi, M., Shalchian, M., Kheradpisheh, S. R., & Masquelier, T. (2021). STiDi-BP: Spike time displacement based error backpropagation in multilayer spiking neural networks. Neurocomputing, 427, 131-140. https://doi.org/10.1016/j.neucom.2020.11.052

[15] Qian, L., Zhao, J., & Ma, Y. (2022). Option Pricing Based on GA-BP neural network. Procedia Computer Science, 199, 1340-1354. https://doi.org/10.1016/j.procs.2022.01.170

[16] Tian, W., & Cao, Y. (2023). Evaluation model and algorithm optimization of intelligent manufacturing system on the basis of BP neural network. Intelligent Systems with Applications, 20, 200293. https://doi.org/10.1016/j.iswa.2023.200293

[17] Wang, J., Li, T., Sun, C., Yan, R., & Chen, X. (2023). BP-SRM: A directly training algorithm for spiking neural network constructed by spike response model. Neurocomputing, 560, 126832. https://doi.org/10.1016/j.neucom.2023.126832

[18] Wang, L., & Bi, X. (2021). Risk assessment of knowledge fusion in an innovation ecosystem based on a GA-BP neural network. Cognitive Systems Research, 66, 201-210. https://doi.org/10.1016/j.cogsys.2020.12.006

[19] Wang, Q., Ye, M., Li, Y., Zheng, X., He, J., Du, J., & Zhao, Y. (2023). MOSFET modeling of 0.18μm CMOS technology at 4.2K using BP neural network. Microelectronics Journal, 132, 105678. https://doi.org/10.1016/j.mejo.2022.105678

[20] Wu, G., Huang, S., Liu, T., Yang, Z., Wu, Y., Wei, G., Yu, P., Zhang, Q., Feng, J., & Zeng, B. (2023). Numerical Study of the Biomechanical Behavior of a 3D Printed Polymer Esophageal Stent in the Esophagus by BP Neural Network Algorithm. CMES - Computer Modeling in Engineering and Sciences, 138(3), 2709-2725. https://doi.org/10.32604/cmes.2023.031399

[21] Wu, S., Zhao, G., & Wu, B. (2022). Real-time prediction of the mechanical behavior of suction caisson during installation process using GA-BP

neural network. Engineering Applications of Artificial Intelligence, 116, 105475. https://doi.org/10.1016/j.engappai.2022.105475

[22] Xiao, T., Liu, Z., Lu, L., Han, H., Huang, X., Song, X., Yang, X., & Meng, X. (2023). LSTM-BP neural network analysis on solid-liquid phase change in a multi-channel thermal storage tank. Engineering Analysis with Boundary Elements, 146, 226–240. https://doi.org/10.1016/j.enganabound.2022.10.014

[23] Yan, J. (2022). Application of improved BP neural network model in bank financial accounting. Intelligent Systems with Applications, 16, 200155. https://doi.org/10.1016/j.iswa.2022.200155

[24] Yang, H., Li, X., Qiang, W., Zhao, Y., Zhang, W., & Tang, C. (2021). A network traffic forecasting method based on SA optimized ARIMA–BP neural network. Computer Networks, 193, 108102. https://doi.org/10.1016/j.comnet.2021.108102

[25] Yang, Z., Mao, L., Yan, B., Wang, J., & Gao, W. (2021). Performance analysis and prediction of asymmetric two-level priority polling system based on BP neural network. Applied Soft Computing, 99, 106880. https://doi.org/10.1016/j.asoc.2020.106880

[26] Zhan, W., Zhang, L., Feng, X., Pan, P., Cai, X., & Wen, X. (2023). An equivalent processing method for integrated circuit electrical parameter data using BP neural networks. Microelectronics Journal, 139, 105912. https://doi.org/10.1016/j.mejo.2023.105912

[27] Zhang, D., & Lou, S. (2021). The application research of neural network and BP algorithm in stock price pattern classification and prediction. Future Generation Computer Systems, 115, 872-879. https://doi.org/10.1016/j.future.2020.10.009

[28] Zhou, Z., Li, Y., Li, F., & Cheng, H. (2022). An intelligence energy consumption model based on BP neural network in mobile edge computing. Journal of Parallel and Distributed Computing, 167, 211–220. https://doi.org/10.1016/j.jpdc.2022.05.005

[29] Adams, S. O., Azikwe, E., Zubair, M. A. Artificial Neural Network Analysis Of Some Selected Kdd Cup 99 Dataset For Intrusion Detection. Acta Informatica Malaysia. 2022; 6(2): 55-61. http//:doi.org/ 10.26480/aim.02.2022.55.61

[30] Zhou, Z., Qiu, F., Zheng, F., & Ye, M. (2022). Neural network joint capacity-power control strategy based on NSGAII-BP for interference suppression in LEO satellite uplinks. Computers and Electrical Engineering, 102, 108093. https://doi.org/10.1016/j.compeleceng.2022.108093