Deciphering COVID-19 Narratives: A Comparative Study of ML Models (RF, MNB, GB, LR, SVM) and DL Models (CNN, Bi-LSTM) for News Article Classification

Kana Das[1], Md. Asadullah[2], Md. Murad Hossain1[*], Annita Siddeka Tanni[1], Shahidul Islam[3], Masudul Islam[4], Mst Sharmin Akter Sumy[5]
[1]Department of Statistics, Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Gopalganj-8100, Bangladesh
[2]Bangladesh Rice Research Institute, Gazipur-1701, Bangladesh
[3]MIS Department, Friendship, Bangladesh
[4]Statistics Discipline, Khulna University, Khulna, Bangladesh; masudul_stat@ku.ac.bd;
[5]Department of Bioinformatics and Biostatistics, University of Louisville, 485 E.
Gray St, Louisville, 40202, KY, USA
E-mail: 12kanadas43@gmail.com, asadullahstat@gmail.com, murad.stat@bsmrstu.edu.bd, annitatanni28@gmail.com, jushahidms@gmail.com, masudul_stat@ku.ac.bd , mstsharminakter.sumy@louisville.edu
[*]Corresponding author

Keywords: Machine Learning (ML), Natural Language Processing (NLP), Roc Curve, GloVe, and FastText

Received: June 24, 2024

The COVID-19 pandemic has provided an unprecedented amount of information in news outlets, which include scientific, health-related, political, economic, and social narratives. This study compares the effectiveness of machine learning and deep learning algorithms for classifying text data, with a certain emphasis on how well the former handle COVID-19 news narratives. The study dataset contains news articles regarding COVID-19. To achieve the primary purpose of this research is to classify COVID-19 related news, we integrate multiple datasets. The analysis reveals machine learning models exhibit superior performance in text data classification. In particular, the Random Forest model reaches a 98% accuracy rate. In contrast, with regards to deep learning models, the Bidirectional Long Short-Term Memory model with FastText integration turns out to be the best option due to its exceptional accuracy. Exploratory data techniques such as topic modeling and word cloud approaches are incorporated to uncover hidden patterns in the data. Pre-trained (e.g., deep learning) and non-pre-trained ML models are implemented highlighting the versatility of ML in text classification tasks. The specific purpose to compare to the deep learning and machine learning algorithm to classification of the new article. Notably, a predictive model employing Bi-LSTM with the FastText pre-trained model achieved an impressive 94% accuracy in classifying COVID-19 news reports.

Povzetek: Primerjalna analiza modelov ML in DL za klasifikacijo novic COVID-19 razkrije RF kot najbolj natančno (98 %), medtem ko Bi-LSTM s FastText (94 %) odlikuje kontekstualno razumevanje, kar izboljšuje učinkovitost klasifikacije besedila.

1 Background of the study

Natural language processing has made extensive use of text classification to divide texts into different classes. In the context of COVID-19 news articles, text classification is crucial for identifying and categorizing the vast amount of information generated daily. In text categorization tasks, machine learning algorithms have demonstrated encouraging results, especially when applied to COVID-19 news items. Empirical evaluation has been conducted to determine the efficacy of several machine learning techniques for text classification of COVID-19 news items.
Jin et al. (2024) investigated the use of AI techniques, such as natural language processing and machine learning, to improve text classification [1]. Their findings highlight the potential of technologies to enhance accuracy and efficiency in text processing, supporting information retrieval and decision-making despite operational challenges. Didi et al. (2022) undertook studied on the categorization of tweets related to COVID-19 using machine learning methods and analysed public sentiment about the pandemic through their novel hybrid feature extraction method that combines syntactic elements with semantic aspects for more accurate text data representation and enhanced classification. Their research built upon previous work exploring Twitter's potential in understanding public opinion during the pandemic, focusing on sentiment analysis using ML models as naïve Bayes and Logistic Regression [2]. The authors demonstrated with a focus on diagnostics and predictive modelling. It emphasized deep learning's impact on healthcare applications, especially

Convolutional Neural Networks while addressing challenges like data scarcity and diversity [3]. Abdeen et al. (2021) introduced NeoNet, a cutting-edge machine learning algorithm created to categorize news stories and medical papers about COVID-19 according to their degree of veracity. Leveraging advanced Term Frequency-Inverse Document Frequency bigram features, NeoNet enables accurate prediction of document relevance and accuracy, with the goal of widespread adoption across various social media platforms [4]. In his evaluation of the COVID-19 pandemic's fake news prevalence and its effects, Abhishek Koirala discussed earlier studies on the identification of fake news and brings up the establishment of the Liar Dataset. The study addressed new challenges posed by COVID-19-related news, experimenting with deep learning models to improve classification accuracy but notes inconsistencies in the dataset that may hinder this process [5].

The methodologies and efficacy of neural network and neuro-fuzzy algorithms in detecting Covid-19-related disinformation on social media are thoroughly examined by Ravichandran and Keikhosrokiani (2023) [6]. This study highlights the role of NF and NN methodologies in assessing their strengths, and limitations, and providing recommendations for future research. Chughtai et al. (2021) paper emphasizes the role feature selection has in model performance and uses benchmark datasets to demonstrate the efficacy of their SVM classifier using MMR algorithm, emphasizing its improvement in F1-score and role in detecting misinformation [7]. Arbane et al. (2023) stress the importance of using advanced machine learning algorithms for automatic sentiment analysis to gain insights from social media data during the COVID-19 pandemic [8]. Previous studies, including those by Mansoor et al. (2020) and Samuel et al. (2020) used various machine-learning models to evaluate public sentiments at different pandemic stages, revealing details such as the impact of lockdowns on emotional responses [9,10]. Additionally, other researchers incorporated deep learning methods to examine sentiments [11]. Using machine learning classifiers [11], researchers investigated sentiments expressed on Twitter during the pandemic using various machine learning classifiers and predicts sentiments in two datasets collected before and after lockdown periods. The results suggest changes in public sentiment during and after lockdowns, providing insights into public opinion dynamics in response to the pandemic and associated restrictions [12].

Dangi et al. (2022) examines the nature of COVID 19 news coverage across the United Kingdom, India, Japan, and South Korea using topic modeling and sentiment analysis with top2vec and RoBERTa to deeply analyse a large amount of news data. It introduces an extensive approach that utilizes the top2vec algorithm to identify underlying topics in news articles, and RoBERTa for categorizing sentiments [11]. They focused on sarcasm detection using lexical and word embedding approaches while others developed systems for identifying fake news by analysing sentiment and named entity features from Twitter data [12]. Madani (2021) discussed about the MVEDL ensemble deep learning model, which is intended to categorize tweets about COVID-19 as informative or not. They evaluated on the "COVID-19 English labelled tweets" dataset, transformer models like RoBERTa, BERTweet, and CT-BERT that demonstrated a strong performance with an accuracy of 91.75% and an F1-score of 91.14% [14]. Malla and Alphonse (2021) investigate how well several transformer-based models—including BERT, RoBERTa, ALBERT, and DistilBERT perform in text classification for uses like sentiment analysis and the identification of false news. Their findings suggest that these models can significantly improve accuracy in classifying challenging data types found on social media platforms [15].

Qasim et al. (2022) explores the application of domain-specific BERT-based models, such as BioBERT and CovBERT. They highlight the superiority of CovBERT in handling vocabulary deficiencies in scientific summaries, achieving up to 94% accuracy compared to its predecessors [16]. They also examine the transition from machine learning to deep learning models and highlight the potential of the Cov-Dat-20 dataset in assisting epidemiologists in addressing the challenges posed by COVID-19. Khadhraoui et al. (2022) explores sentiment analysis of COVID-19-related tweets, covering research in different languages like Nepali and review underscores the constraints of current methodologies. They support a backdrop for the authors' exploration of hybrid feature extraction methods to enhance the classification of Nepali COVID-19 tweets [17]. Shahi et al. (2022) integrates Bidirectional Long Short-Term Memory neural networks with local interpretable model-agnostic explanations, aiming for an effective model and proposes experimentation with different RNN-based models, machine learning techniques, explains ability methods such as LIME and SHAP using datasets like the Constraint 2021 COVID-19 fake news dataset and the WNUT COVID-19 tweet dataset [18].

The challenge of accurately classifying biomedical papers related to the COVID-19 pandemic into relevant categories addressed by Ahmed et al. (2022). They look into numerous machine learning approaches such as document representation strategies, neural networks and random forests, focusing on a subset of the Lit Covid corpus and employing pre-processing and feature engineering methods such as TF-IDF, BOW, and BERT embedding [19]. Rabby and Berka (2023) discusses a range of studies related to the detection of fake news and spam, including exploration of review spam across millions of Amazon reviews, identification of spam in Arabic texts, comparison of different machine learning models using n-gram analysis to flag false information and focus on online articles for detecting fake news with a high accuracy rate [20]. Etaiwi (2022) also build an ensemble model utilizing fusion vector multiplication on a COVID-19 English fake news dataset, with a 98.88% accuracy and an F1-score of 98.93% for achieving high performance in their model evaluation [21]. Malla and Alphonse (2022) discuss the

challenge of classifying Covid-19 misinformation on social media conducting a systematic review of studies from 2018 to 2021 that use various machine learning techniques of misinformation classification, evaluates their efficiency, strengths, and limitations. They propose a novel hybrid ANFIS-DNN model to enhance accuracy and effectiveness in this domain [22].
Ravichandran and Keikhosrokiani (2023) present a novel collection of 24 Information Society (IS) laws, exploring their connections with electronics and artificial intelligence (AI) [23]. These laws highlight exponential growth in areas such as processing, storage, and communication, while addressing their societal and economic impacts. The researchers also discuss how AI advancements contribute to electronics and broader IS progress, emphasizing their interdependence and potential future developments [24]. Janko et al. (2021) investigate factors influencing the early spread of COVID-19 across countries, focusing on the period before countermeasures [25]. By analyzing a diverse dataset with statistical methods and machine learning (ML) feature selection, they identify key factors like culture, development, and travel. They also use a novel rule discovery algorithm to explore factor interconnections, cautioning against overreliance on ML alone. The best model, using a decision tree classifier, predicts infection classes with about 80% accuracy. The researcher of [26] presents SentiTextRank, an emotional variant of TextRank, for extractive summarization and classifying sentences into eight emotional categories from SenticNet, SentiTextRank using both single and multi-document summarization tasks. The work of authors in [27] investigate the emotional component of successful medical web pages related to spine pathology, hypothesizing that they would exhibit distinct emotional patterns. Using sentiment analysis and machine learning, the study retrieves high classification accuracy, with disgust emerging as a key emotion. The findings suggest that digital content impacts patients' biopsychosocial ecosystems, influencing chronic pain and health behaviours, raising ethical concerns for health information providers.
The comprehensive review of existing literature on text classification in the context of COVID-19 indicates several significant areas where further research is needed. The application of natural language processing, deep learning, and machine learning methods for categorizing documents, and analyzing sentiment, and fake news identification has advanced significantly, yet there are still unaddressed issues. To begin with, despite numerous studies focusing on fake news detection and sentiment analysis, the lack of standardized datasets and inconsistent labelling practices poses challenges for model generalization and benchmarking. Additionally, many studies have predominantly focused on English-language datasets without considering the requirement for multilingual models to combat misinformation across diverse cultural and linguistic contexts. While some research acknowledges the

Table 1: Comparison among the related articles.

Autor and Title References	Dataset	Algorithm	Result
[2]	Tweets dataset	TF-IDF, Word2Vec, Glove, and FastText, SVM.	Fast Text with TF-IDF performed better
[4]	COVID-19 News Articles Open Research dataset	TF-IDF, NN, SVM, and RF, NeoNet.	NeoNet better perform
[5]	COVID-19 News articles dataset	LR, Embedded LSTM, LSTM, Bidirectional LSTM.	Embedded LSTM Hybrid models better performer
[6]	COVID-19 News dataset between July 2018 and May 2021.	Neuro-fuzzy, NN and specially ANFIS.	Hybrid ANFIS-DNN better performer
[8]	COVID-19 related tweets and comments dataset.	LSTM, Bi-LSTM,	Bi-LSTM model is superior over LSTM.
[10]	Twitter datasets	LR, RF, MNB, SVM, and DT	Decision Tree Classifier better performer.
[11]	COVID-19 articles dataset	top2vec and RoBERTa	RoBERTa
[13]	COVID-19 English labeled tweets dataset	RoBERTA, CT-BERT, and BERTweet	Ensemble Deep Learning (MVEDL) model
[14]	COVID-19 fake news dataset and COVID-19 English tweet dataset	BERT-base, BERT-large, RoBERTa-base, RoBERTa-large, DistilBERT, XLM-RoBERTa-base, ALBERT-	RoBERTa-base model achieved the highest accuracy in COVID-19 fake news dataset, Bart-large, BERT-base are the respective winners of other datasets

		base-v2, Electra-small, and BART-large	
[15]	Dataset of biomedical articles on COVID-19	CovBERT and BERT	CovBERT outperform
[16]	Nepali-COVID-19 tweets dataset	FastText + TF-IDF, LR, SVM, NB, KNN, DT, RF, Extreme Tree classifier, AdaBoost, and MLP	FastText with TF-IDF, SVM + RBF is the best performing classifier.
[17]	Covid-19 fake news dataset	BiLSTM, LSTM, GRU, RNN, CNN, SVM, DT	BiLSTM model high classification accuracy
[18]	COVID-19 Open Research dataset	RF, Logistic regression, KNN, DT Multi-layer Perceptron, Neural Network (BERT), BOW	Random Forest and Neural Network (BERT)
[20]	COVID-19 fake news dataset	BERT, BERTweet, AlBERT, CT-BERT, RoBERTa and DistlBERT	RoBERTa
[21]	Covid-19 misinformation related papers dataset	Neuro-Fuzzy, Neural Network	ANFIS-DNN model

importance of explain ability in classification models; limited research exists on implementing and evaluating reasonable AI specifically tailored for COVID-19-related text classification. The following Table 1 highlights the current state of research on text classification using machine learning algorithms for COVID-19 news articles with comparison among them.

Table 1 shows most of the research for text data analysis are not using feature extraction technique. Some of the literature use FastText or TF-IDF technique for feature extraction technique. But they skip some powerful feature extraction technique like BERT or Glove technique. In this research, we utilize FastText, TF-IDF and BERT techniques and compare them.

2 Methods and materials

The three segments of the text classification approach are features extraction, text pre-processing, and dataset description are considered. In the Algorithm Selection phase, the innovative deep learning algorithms have been merged. How text pre-processing was done for machine learning is explained in the Text Classification Approach section. The mathematical description of the process is used to extract private data from the dataset and how text data can be transformed into a numeric form shown in the third phase. Lastly, the algorithms that have been shortlisted in this type of study are discussed in the fourth phase. Learning rate for Adam optimizer: 0.001 is considered as starting point. Batch sizes (e.g**.**, 16, 32, or 64) typically work well for text data analysis, especially in tasks like text classification or sentiment analysis. We use 32 for our text data analysis. Start with 10 epochs use early stopping to avoid overfitting.

2.1 Dataset description

The dataset from Kapoor et.al (2020) [34] and Lipenkova et.al (2021) [35] contained news articles regarding COVID-19 since the primary purpose of this research is to categorize news related to COVID-19. In this research we use news article text data. Two Datasets of news articles are extracted from www.inshorts.com and then labeled based on relation to COVID as well as the sentiment. They have been assembled from different repositories and reformatting for a similar distribution. After combining the two datasets, the sample size of our dataset is 14012 in total. Balancing imbalanced data for classification tasks in machine learning (ML) is crucial because imbalanced datasets can lead to biased models that favour the majority class and fail to detect the minority class effectively. The imbalance data use SMOTE (Synthetic Minority Over-Sampling Technique) interpolating between existing samples of the minority class. For analysis text data we pre-process the data by removing punctuation**,** removing numbers, removing special characters and symbols, removing URLs, emails, and mentions, removing stop words, tokenization, text normalization**,** vectorization of text, term frequency - inverse document frequency. The features column of Table 1, and the feature’s value of Table 2 is the header-wise first content information of the dataset. This dataset consists of six attributes which are depicted below.

Table 2: Attribute of the proposed research dataset

Features	Feature's Value
Headline	Headline of the article
Sentiment	1 if the article is positive, 0 otherwise
Covid	1 if the article is related to COVID, 0 otherwise
Description	Description of the news
Image	Image URL
Source	Source URL

2.2 Features extraction algorithm

Two new classifications, pre-trained model configuration, and non-pre-trained model setup have been incorporated into the features extraction. This process comprises two parts: the non-pre-trained model setup and the pre-trained model configuration. The unique word embedding technique Text to sequence, Fast Text, and Glove is demonstrated by the PTMS. However, the NPTMS provided an explanation of the typical features extraction method: Inverse document frequency paired with term frequency (TF-IDF).

A) **Pre-trained model structure**

FastText: FastText is a library developed by Facebook's AI research (FAIR) lab, designed for efficient text classification and representation learning. It is particularly useful for text classification, word representation (word embeddings), and language modeling tasks. FastText improves traditional **Word2Vec** by representing words as **subword-level units** (i.e., n-grams), making it more effective for handling rare or out-of-vocabulary (OOV) words. Additionally, FastText supports effective training and inference, making it a popular tool for text classification tasks. It is an efficient and powerful tool for text classification, leveraging subword information to handle out-of-vocabulary (OOV) words and offering better performance in languages with rich morphology. FastText is a novel technique for word embedding represents words as bags of n-gram characters. This approach addresses the issue of morphology neglect, in other words embedding representations by capturing subword information explained [28]**.** Consider the term "introduce" with n equal to 3, FastText generates three-gram characters shown in the following representation:

$< in$, int, ntr, tro, rod, odu, duc, uce, ce $>$

We are considering a word w that is correlated using an n-gram dictionary with a size of G as a way to represent the vector for each n-gram g. In this case, the acquired scoring function defined in Spirovski et al. (2018) [28] is:

$$s(w,c) = \sum_{g \in gw} z_g^T v_c \quad (1)$$

where $g_w \in \{1,2,\ldots,G\}$

Global Vectors (Glove): GloVe is another popular method for **word embedding** that is widely used in text analysis. GloVe was developed by Stanford researchers and designed to capture **global statistical information** about a corpus, unlike methods such as **Word2Vec**, which focus on local context windows. The key idea behind GloVe is the **co-occurrence matrix** of words in a corpus contains valuable information about the relationships between words. GloVe uses the co-occurrence data to generate dense word vector. Global Vectors (GloVe), a potent word embedding method has been applied to text classification [29]. This strategy bears a strong resemblance to Word2Vec, which provides a high-dimensional vector of each word and trains it across an extensive corpus using surrounding terms. Pre-trained word embedding are widely used, based on 50 dimensions for word presentation in Wikipedia 2014 and Gigaword 5, as well as 400,000 vocabularies introduced as the corpus [36]. Unlike the traditional word embeddings such as Word2Vec, which can't generate vectors for words not seen in training data, FastText can generate embeddings for any word by breaking it into subword units (n-grams). FastText is particularly useful for languages with complex morphology (e.g., Turkish, Finnish) or rare words as it captures meaningful subword features. Pre-trained FastText models are available for many languages, which can be used directly for feature extraction in downstream tasks, saving time and computational resources.

B) Non-pre-trained model structure

Term Frequency-Inverse Document Frequency (TF-IDF)

In order to reduce the influence of frequently occurring words in the dataset, inverse document frequency is a method that should be combined with term frequency [30]. Terms in the document that have a high or low frequency are given a higher weight by IDF. TF-IDF combines term frequency and inverse document frequency. Equation 2 mathematically represents a term's weight that is used in this study.

$$\text{TF-IDF}.W(d,t) = TF(d,t) * \log\left(\frac{N}{df(t)}\right) \quad (2)$$

In this scenario, $df(t)$ represents the number of documents in the corpus containing the word t, and N is the total document count. According to Tokunaga and Makoto the initial factor in equation 2 enhances recall, while the second term improves word embedding accuracy [37]. TF-IDF is a simple and computationally efficient method for text feature extraction. Unlike GloVe or FastText, TF-IDF doesn't require a pre-trained model. It can be computed directly from the text corpus. The resulting features (weights) are easy to interpret since they are based on word frequency and document distribution.

2.3 Deep learning algorithms

CNN and RNN are the main types of deep learning architectures used for text classification. Hierarchical machine learning or deep learning involves a series of algorithms performed in sequential order.

2.3.1 Bidirectional long short-term-memory (Bi-LSTM)

Bi-LSTM input sequences can be in both directions with two neuron sub-layers. This orientation is to generate a complete
input context. There are also backward hidden sequences, namely $\overleftarrow{h}$, $\overrightarrow{h}$. From this configuration, we can compute the output sequence y: two neuron sub-layers can be used in both directions for Bi-LSTM input sequences. The goal of this viewpoint is to produce an entire input context. Backward hidden sequences are also present, denoted as (h) $^{\leftarrow}$ and (h) $^{\rightarrow}$. We can calculate the output sequence based on this arrangement y:

$$\overrightarrow{h}_t = \mathcal{H}\left(W_{x\overrightarrow{h}}x_t + W_{\overrightarrow{h}}\overrightarrow{h}\overrightarrow{h}_{t-1} + b_{\overrightarrow{h}}\right) \quad (3)$$

$$\overleftarrow{h}_t = \mathcal{H}\left(W_{x\underset{h}{h}}x_t + W_{\leftarrow\leftarrow h}\overleftarrow{h}_{t+1} + b_{\overleftarrow{h}}\right) \quad (4)$$

$$y_t = W_{\overrightarrow{h}y}\overrightarrow{h}_t + W_{\overleftarrow{h}y}\overleftarrow{h}_t + b_y \quad (5)$$

It is an advanced architecture of the Long Short-Term Memory (LSTM) network, a type of recurrent neural network (RNN) is designed to learn sequential dependencies in data. The key feature of a Bi-LSTM is its ability to consider both past (backward) and future (forward) context while processing sequences. This makes it especially effective for tasks where the entire context of a sequence is critical, such as natural language processing (NLP), speech recognition, and time-series prediction.

2.3.2 Convolutional neural networks (CNN)

A widely used deep learning structure for categorizing hierarchical documents is the convolutional neural network defined [31]. While initially constructed for the processing of images, CNNs have proven to be useful for text classification as well explained [32]. CNN's use pooling to minimize the output's size from one layer to the next in the network to reduce the complexity of the computation. To minimize outcomes while retaining essential features, various pooling techniques are used [33]. The process of choosing the highest value in the pooling window is referred to as max pooling, which is a commonly employed technique. The feature maps are transformed into a single column before transmitting the pooled output from stacked feature maps to the next layer. In general, both the weights and the feature detector filters are modified during the back-propagation phase of a convolutional neural network. The number of channels is a potential issue that emerges when using CNN for text classification (size of the feature space). In general, the program has few channels (e.g., just 3 RGB channels) and can be very broad for text classification applications, resulting in very high dimensionality. The CNN based text classification architecture includes word embedding as input layer 1D convolution layers, 1D pooling layer, completely connected layers, and finally, the output layer [33].

2.4 ML model performance measure

Precision: The ratio of the model's accurate true positive estimate to the total positive estimate (including both correct and incorrect classifications). It is expressed as:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (6)$$

Recall / Sensitivity: The predictive ratio shows a positive correlation and is expressed as:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (7)$$

F1 score: This provides a more accurate estimate than the accuracy metric for the misclassified instances; it is calculated as the harmonic mean of Precision and Recall. In mathematical terms, it can be expressed as

$$\text{F1} = 2 \cdot \frac{\text{Precision.Recall}}{\text{Precision} + \text{Recall}} \quad (8)$$

Accuracy: The sum of all the precisely forecasted events. It is presented as:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (9)$$

2.5 Proposed algorithms architecture

The proposed algorithms provided a description of the algorithms utilized for COVID-19 related news classification. The findings indicate the effectiveness of machine learning algorithms in analyzing the text data used in this research. Both machine learning and deep learning methods are applied to categorize the text, as detailed in our proposed architecture shown in Figure 1.

3 Results analysis

The classification result is gathered using the deep learning (DL) approach and presented by the empirical consequence. The model assesses and analyzes precisely and the best model that emerged.

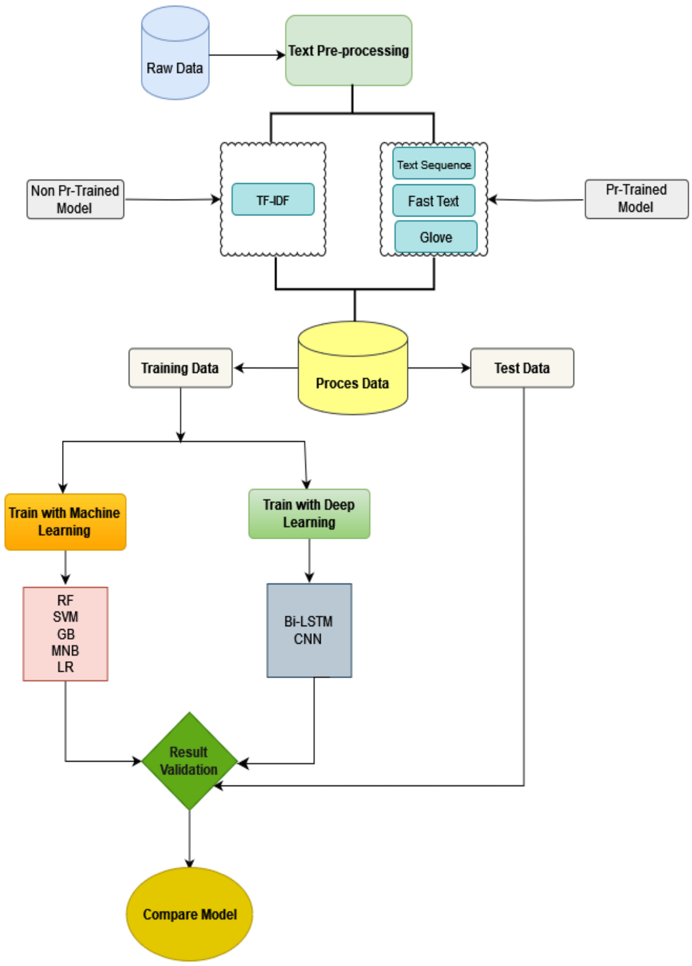

Figure 1: Architecture of the proposed algorithm

	Table 3: Classification result for deep learning algorithms							
Algorithm		**In case of "0"**			**In case of "1"**			
	Features Extraction Technique	**Precision**	**Recall**	**F1-score**	**Precision**	**Recall**	**F1-score**	**Accuracy**
CNN	Text to sequence	.93	.90	.92	.90	.93	.92	.92
BI-LSTM	Text to sequence	.93	.90	.92	.90	.93	.92	.92
BI-LSTM	Fast Text	.95	.92	.94	.93	.95	.94	.94
BI-LSTM	Glove	.93	.90	.92	.90	.99	.93	.92

Table 4: Classification result for machine learning algorithms (Baseline model)

Algorithm	In case of "0"				In case of "1"			
	Features Extraction Technique (FET)	**Precision**	**Recall**	**F1-score**	**Precision**	**Recall**	**F1-score**	**Accuracy**
RF	TF-IDF	.92	.85	.90	.88	.94	.9	.94
MNB	TF-IDF	.75	.86	.82	.87	.73	.8	.81
GB	TF-IDF	.88	.67	.78	.74	.91	.82	.8
LR	TF-IDF	.81	.82	.83	.84	.82	.83	.83
SVM	TF-IDF	.93	.93	.94	.95	.94	.94	.92

Table 5: Classification result for machine learning algorithms with (TF-IDF)

Algorithm	In case of "0"				In case of "1"			
	Features Extraction Technique (FET)	**Precision**	**Recall**	**F1-score**	**Precision**	**Recall**	**F1-score**	**Accuracy**
RF	TF-IDF	.97	.91	.94	.92	.98	.94	.98
MNB	TF-IDF	.80	.92	.86	.91	.77	.84	.85
GB	TF-IDF	.93	.73	.82	.78	.95	.86	.84
LR	TF-IDF	.86	.88	.87	.88	.86	.87	.87
SVM	TF-IDF	.98	.99	.98	.99	.98	.98	.96

The accuracy of the classification algorithms is evaluated based on a set of metrics for each class. These metrics involve accuracy, recall, and f1-score, computed using true and false positives along with false negatives. In Table 3, the classification results for deep learning algorithms and a cutting-edge natural language processing technique are presented based on feature extraction methodology. Additionally, Table 4 and Table 5 exhibit the classification outcomes from machine learning methods without any feature extraction technique and utilizing TF-IDF (term frequency inverse document frequency) approach.

3.1 Deep learning model result

Model assessment includes receiver operating characteristics, area under the curve, and confusion matrix. The algorithm that provides the best accuracy is used as the foundation for the model assessment. Using the Bi-LSTM and the Fast Text method, we were able to obtain a good score in Figure 9. The model assessment illustrates how well our suggested model works for the specific task.

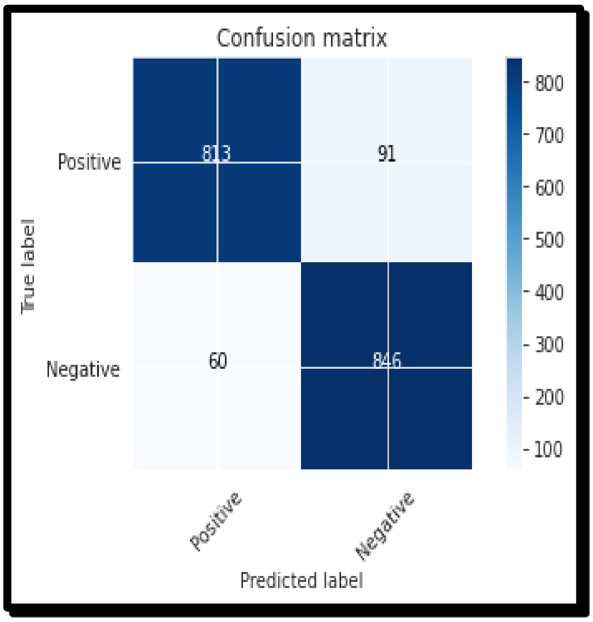

Figure-2: Confusion Matrix of Bi-LSTM using Glove

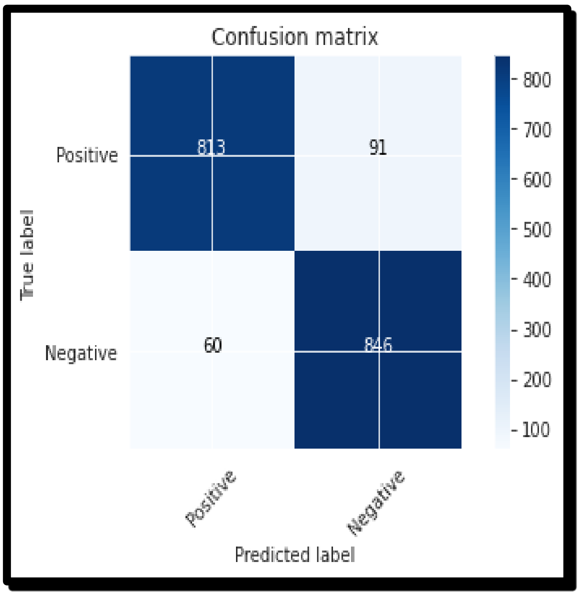

Figure-4: Confusion Matrix of CNN with text sequence

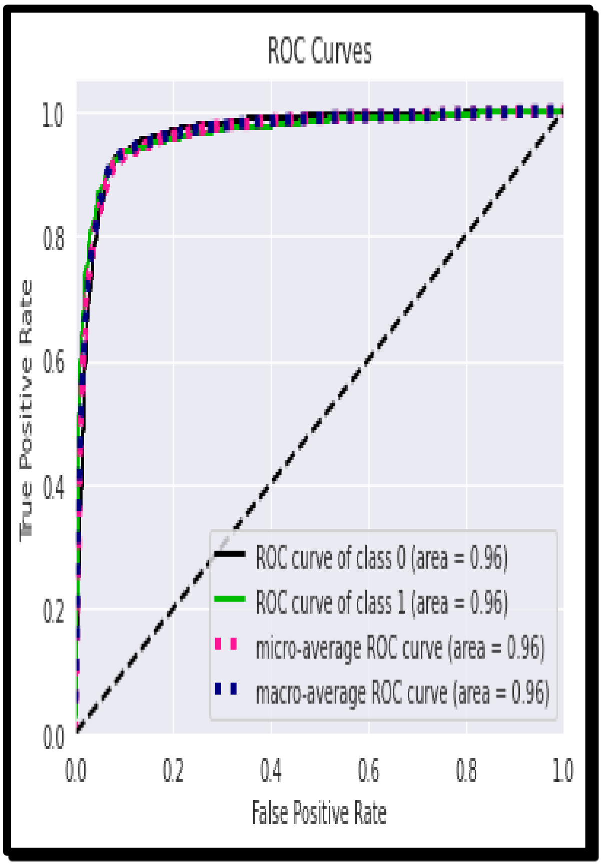

Figure-3: ROC-AUC curve of Bi-LSTM using Glove

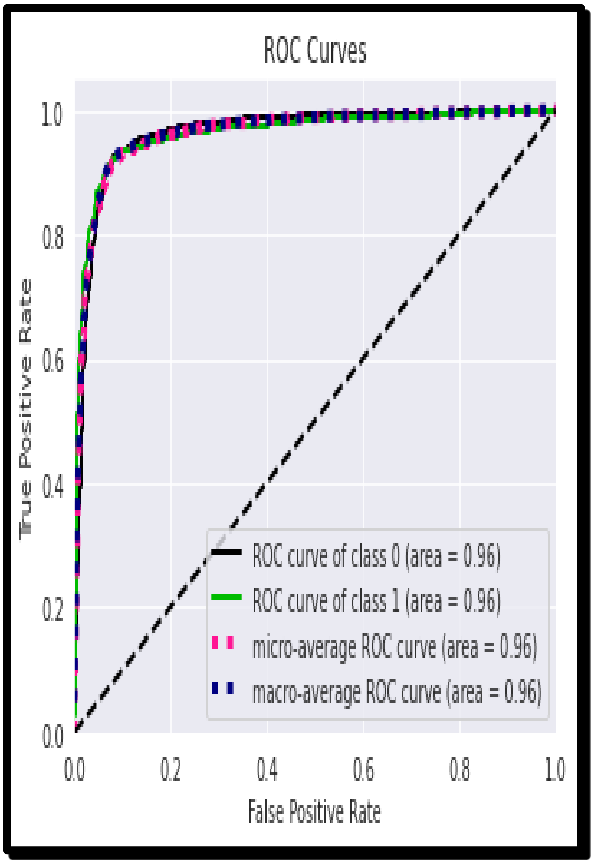

Figure-5: ROC-AUC curve of CNN with Text Sequence

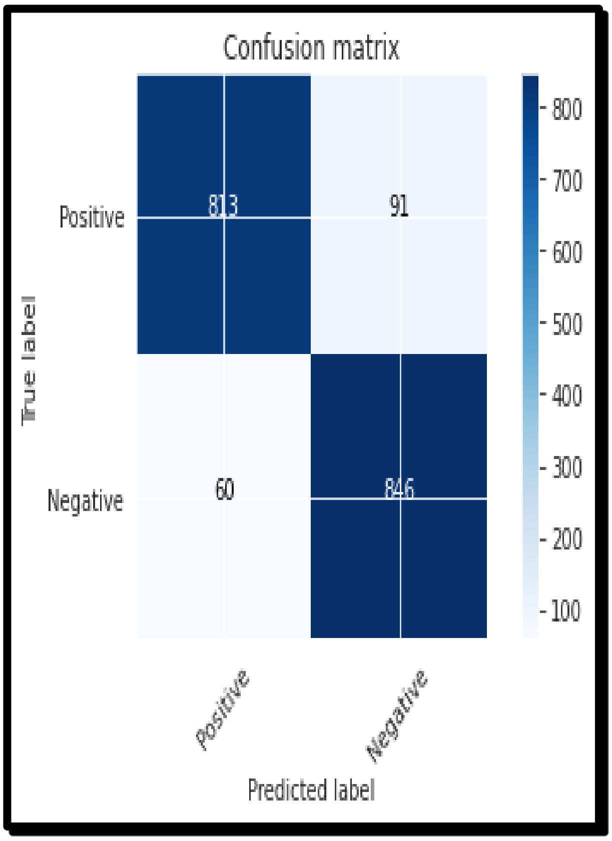

Figure-6: Confusion Matrix of Bi-LSTM with Text

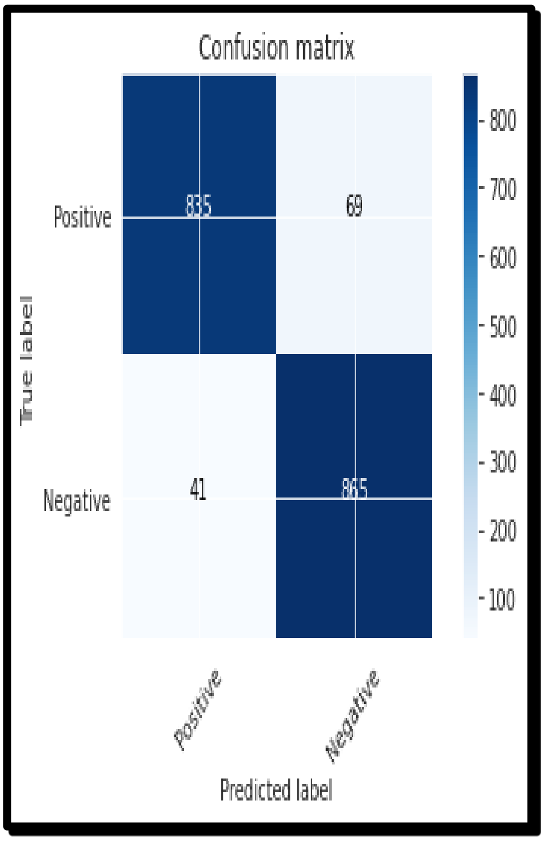

Figure-8: Confusion Matrix of Bi-LSTM using FastText

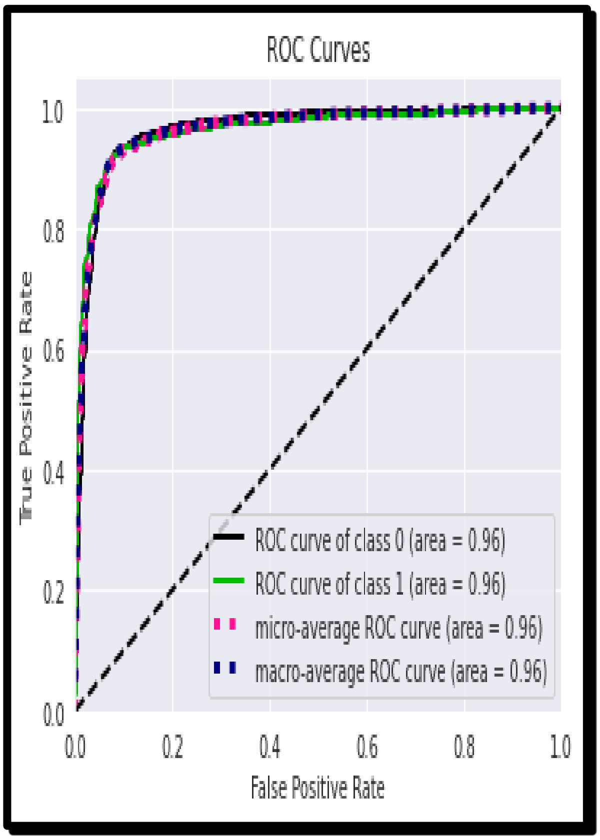

Figure-7: ROC-AUC curve f Bi-LSTM Text Sequence

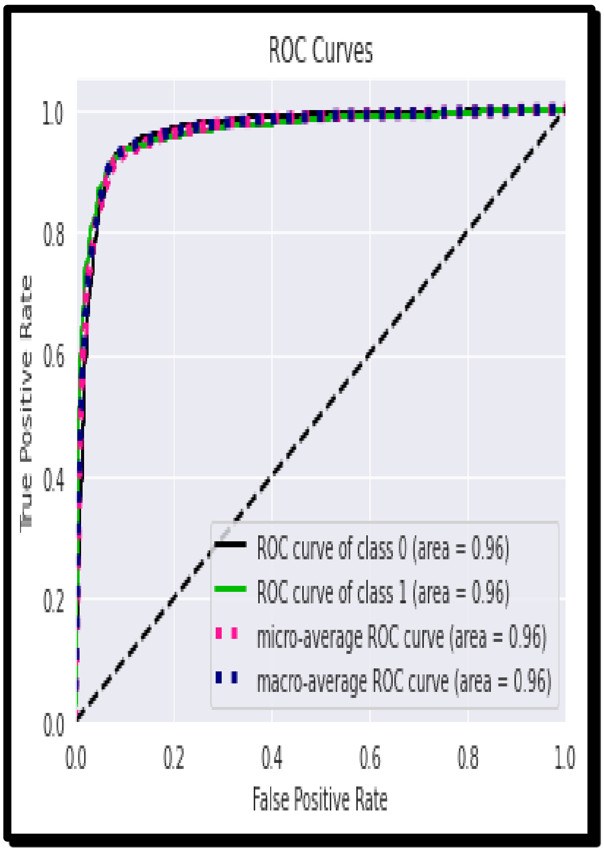

Figure-9: ROC-AUC curve of Bi-LSTM using FastText

3.2 ML model assessment

ML model evaluation is conducted using various approaches, including ROC, AUC, Accuracy of Models, and Confusion Matrix. Notably, the assessment in this section was performed utilizing the algorithm that yields the highest accuracy across all ML methods. We found a satisfactory score by using the SVM method. In this section shows tradional model assessment how well performs compare to the deep learning method in text classification. We only show the best model in machine learning method.

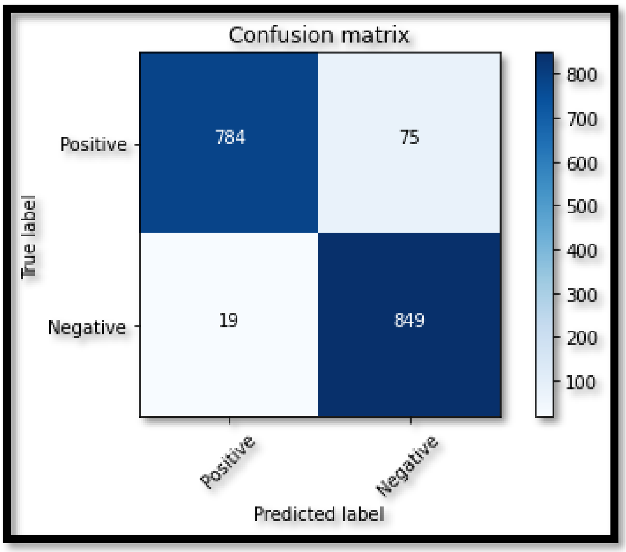

Figure-10: Confusion Matrix of RF

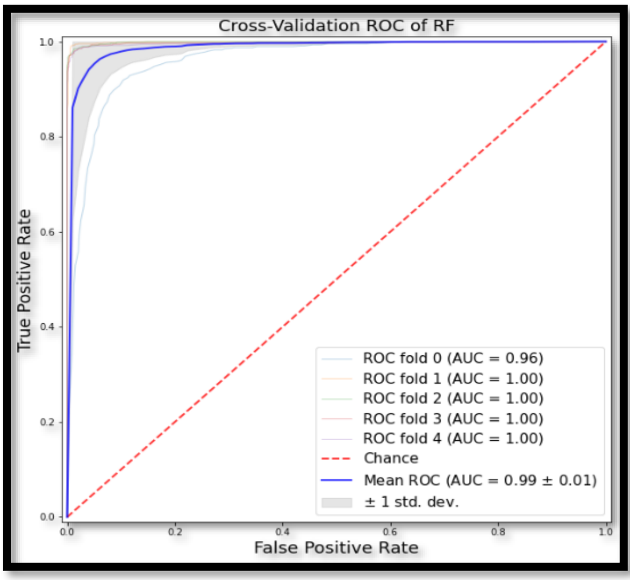

Figure-11: ROC-AUC curve of RF

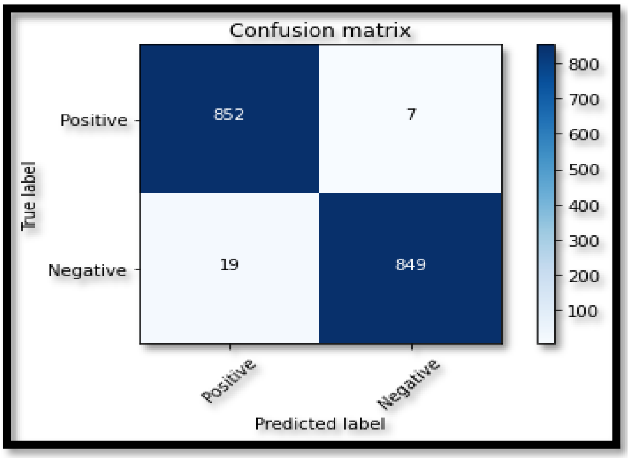

Figure-12: Confusion Matrix of SVM

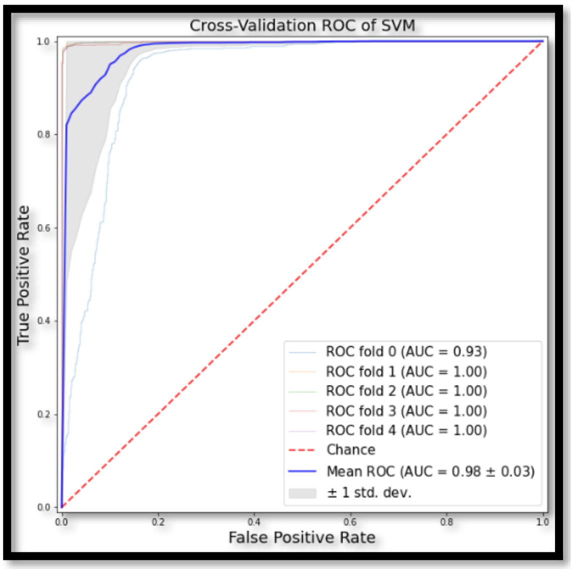

Figure-13: ROC-AUC curve of SVM

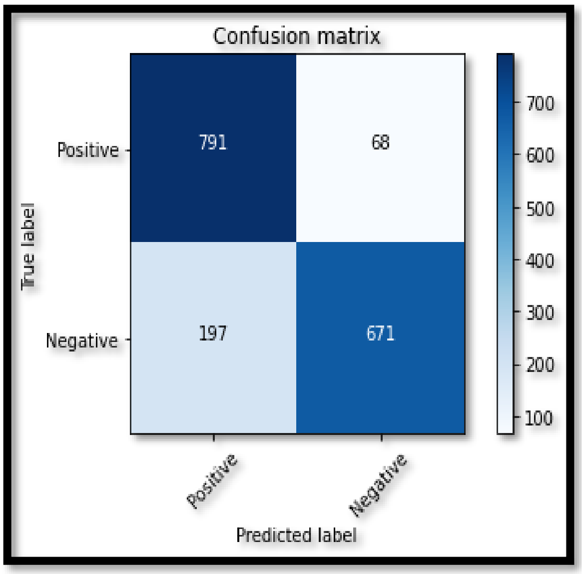

Figure-14: Confusion Matrix of MNB

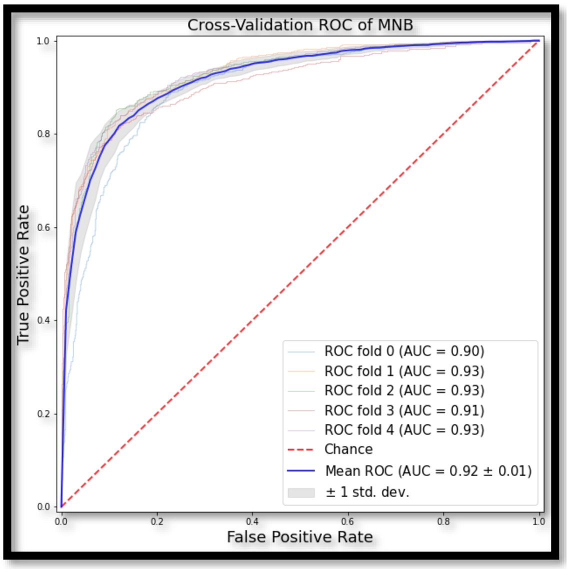

Figure-15: ROC-AUC curve of MNB

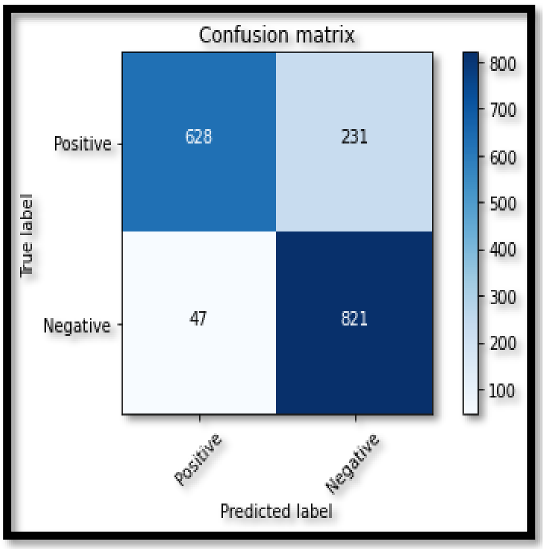

Figure-16: Confusion Matrix of GB

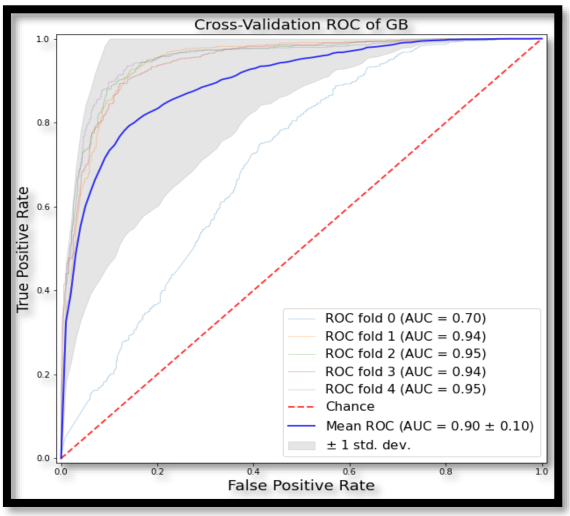

Figure-17: ROC-AUC curve of GB

Estimating binary classification problems often relies on the receiver operator characteristic curve, that plots true positive rate against the false positive rate at different threshold levels. The ROC serves as a probability graph that effectively distinguishes between signal and noise. A key metric derived from the ROC curve is crucial for evaluating a classifier's ability to differentiate between two groups: an AUC of 1 indicating accurate discrimination between positive and negative class points; while an AUC of 0 implies all negatives are predicted as positives and vice versa. For instance, as shown in Figure 11, the RF classifier model achieves an AUC value of approximately 0.96, demonstrating its precise discrimination capability between positive and negative class points.

3.3 Exploratory data analysis

Exploratory data analysis (EDA) is an important technique for examining the dataset and understanding its fundamental characteristics. The EDA provides valuable insights into the dataset combining topic modeling. While EDA uncovers meaningful patterns and observations, the topic modeling approach illustrates the hidden semantic structure of text and figured out the most dominant word in each sentence. In this study, text data was analyzed using these approaches to explore new things from the dataset.

3.3.1 Topic modelling approach

As defined by Hornik and Grün (2011), the topic Modeling approach is a systematic method for classifying items that are present in a written document and extracting hidden patterns from a text corpus [38]. This technique is widely applied for tasks such as feature selection, document clustering, and information extraction from unorganized data. Text in a document can be categorized into distinct topics using the Latent Dirichlet Allocation (LDA), which is an illustration of a topic modeling. The coherence score for latent Dirichlet allocation (LDA) is displayed in Figure 18 below, which is based on the topic count.

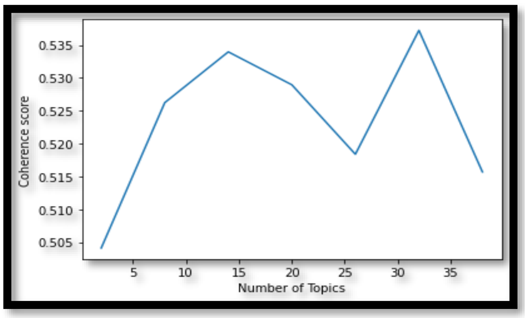

Figure-18: Coherence score for latent Dirichlet allocation (LDA)

We use the "corpora. Dictionary ()" function to create a dictionary of documents. The "Dictionary ()" function iterates through each document, giving each unique token a unique id and counting them. Word Cloud Visualization (WCV) has been developed, showing the dataset's most frequently occurring words in the word cloud. The word's visual representation is an important way to distinguish news categories and derive insightful intuitions. Nevertheless, Figure 19 provides a detailed illustration of Word Cloud, highlighting the dominant words and their importance.

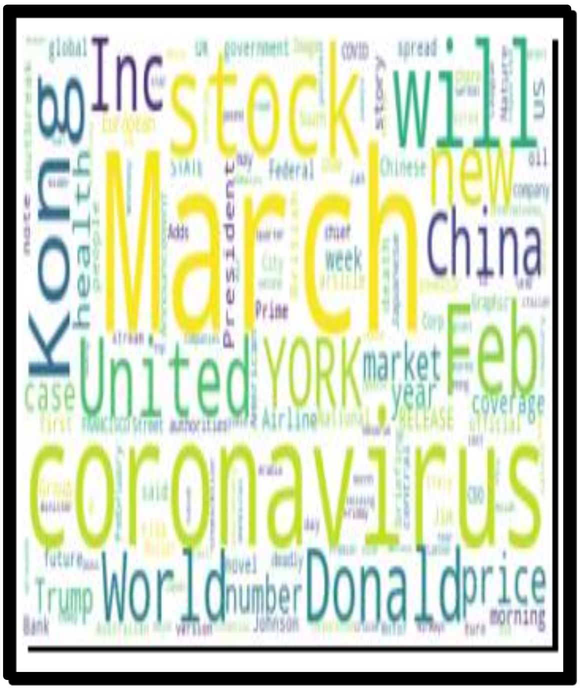

Figure-19: Word Cloud for News classification

4 Conclusion

To protect people's safety in the framework of COVID-19, it is essential to implement measures that make sure the spreading of precise information online. This study employs a predictive model-based utilize the fast-test pre-trained bidirectional long short-term memory (Bi LSTM) with an accuracy of 94% in classifying COVID-19 news. Additionally, the LDA technique is applied to reveal themes and detect hidden patterns within the dataset. The study reveals the deep learning method has lower accuracy compared to the machine learning method for text classification. Although the RF algorithm has achieved 98% accuracy, all method accuracy is lower compared to all deep learning method accuracy. However, Bi-LSTM with fast-test method perform better among all deep learning methods for text classification. Our results indicate that the Bi-LSTM + FastText model outperforms several traditional text classification models, such as SVM and RF, in terms of accuracy and robustness. This performance is consistent with findings in SOTA literature [39], where Bi-LSTM has been shown to capture long-term dependencies in text data more effectively than traditional models. The success of the Bi-LSTM can be attributed to its ability to process information from both the past and future context. This feature is particularly advantageous in sentiment analysis tasks where context is crucial. FastText is well known for its efficient handling of subword information, and also contributes to the performance by capturing richer semantic representations of words, even when out-of-vocabulary terms appear. This method embeddings overcome the limitations of traditional models (such as word2vec or GloVe) face when dealing with out-of-vocabulary (OOV) words. Different models rely deeply on predefined features (e.g., bag-of-words, TF-IDF). Our approach leverages word embeddings from FastText, allows the model to learn features directly from data. The most of research only perform the Bi-LSTM or other machine learning model without embedding or feature extraction technique or hyper parameter technique. In this paper, we incorporate all of these techniques for the classification of text analysis. Our findings show that embedding these techniques all together enhance performance. However, the difference in feature extraction methods could explain the differences in performance between our results and models that rely on more manual feature engineering. Variations in hyper parameter tuning may also explain performance differences. While our study provides a basic hyper parameter grid search, and a more extensive search or fine-tuning of layer sizes, learning rates, and regularization strategies may improve further performance. Overall, the machine learning model is the best classifier for text data classification.

5 Limitations

In this paper, a new machine learning and deep learning methods for classifying COVID-19 related electronics news were provided. While our analysis of the associated experimental data to demonstrate the suggested solution. Inaccurate news classification yielded a decent result for our trials. Nevertheless, this study has several shortcomings such as the dataset that used in this research is modest size in comparison to COVID 19 related text data. Our dataset contains lacks noun, verb, adjective, and other information since it does not use Part of Speech (POS) tags. During text pre-processing, we employed the lemmatization strategy, which takes a lot of time and results in poorly tuned hyperparameters on the BI-LSTM model. Despite these limitations, this study highlights the effectiveness of combining machine learning and deep learning methods for text classification and provides a foundation for future research to address these challenges.

References

[1] Jin, H., & Zhang, L. (2024, September). Text classification and its application based on natural language processing and machine learning. In *AIP*

Conference Proceedings (Vol. 3131, No. 1). AIP Publishing. https://doi.org/10.1063/5.0230295

[2] Didi, Y., Walha, A., & Wali, A. (2022). COVID-19 tweets classification based on a hybrid word embedding method. *Big Data and Cognitive Computing*, *6*(2), 58. https://doi.org/10.3390/bdcc6020058

[3] Tiwari, S., Chanak, P., & Singh, S. K. (2022). A review of the machine learning algorithms for COVID-19 case analysis. *IEEE Transactions on Artificial Intelligence*.https://doi.org/10.1109/TAI.2022.3142241

[4] Abdeen, M. A., Hamed, A. A., & Wu, X. (2021). Fighting the COVID-19 Infodemic in News Articles and False Publications: The NeoNet Text Classifier, a Supervised Machine Learning Algorithm. *Applied Sciences*, *11*(16),7265. https://doi.org/10.3390/app11167265

[5] Koirala, A. (2020). COVID-19 fake news classification with deep learning. *Preprint*, *4*. https://www.researchgate.net/profile/Abhishek-Koirala/publication/344966237_COVID-19_Fake_News_Classification_with_Deep_Learning/links/5f9b6ba5299bf1b53e5130b8/COVID-19-Fake-News-Classification-with-Deep-Learning.pdf

[6] Ravichandran, B. D., & Keikhosrokiani, P. (2023). Classification of Covid-19 misinformation on social media based on neuro-fuzzy and neural network: A systematic review. *Neural Computing and Applications*, *35*(1),699-717. https://doi.org/10.1007/s00521-022-07797-y(0123456789().,-volV) (0123456789().,-volV)

[7] Chughtai, M. A., Hou, J., Long, H., Li, Q., & Ismail, M. (2021, November). Design of a predictor for COVID-19 misinformation prediction. In *2021 International Conference on Innovative Computing (ICIC)* (pp.1-7). IEEE. https://doi.org/10.1109/ICIC53490.2021.9693057

[8] Arbane, M., Benlamri, R., Brik, Y., & Alahmar, A. D. (2023). Social media-based COVID-19 sentiment classification model using Bi-LSTM. *Expert Systems with Applications*, *212*, 118710. https://doi.org/10.1016/j.eswa.2022.118710

[9] Mansoor, M., Ur Rehman, Z., Shaheen, M., Khan, M. A., & Habib, M. (2020). Deep learning based semantic similarity detection using text data. *Information Technology and Control*, 49(4), 495-510. https://doi.org/10.5755/j01.itc.49.4.27118

[10] Samuel, J., Ali, G. M. N., Rahman, M. M., Esawi, E., & Samuel, Y. (2020). Covid-19 public sentiment insights and machine learning for tweets classification. *Information*, 11(6),314. https://doi.org/10.3390/info11060314

[11] Hossain, M. M., Asadullah, M., Tamanna, S., Tazwar, M. A. S., Alam, M. M., Hossain, M. A., Islam, M. & Sumy, M. S. A. (2024). Automated Machine Learning Algorithms for Predicting Anxiety and Depression in Bangladeshi University Students. *Journal of Information Systems Research and Practice*, 2(3), 16-31. https://mojc.um.edu.my/index.php/JISRP/article/view/54235

[12] Dangi, D., Dixit, D. K., & Bhagat, A. (2022). Sentiment analysis of COVID-19 social media data through machine learning. *Multimedia Tools and Applications*, *81*(29),42261-42283. https://doi.org/10.1007/s11042-022-13492-w:

[13] Ghasiya, P., & Okamura, K. (2021). Investigating COVID-19 news across four nations: A topic modeling and sentiment analysis approach. *Ieee Access*, *9*, 36645-36656. 10.1109/ACCESS.2021.3062875

[14] Madani, Y., Erritali, M., & Bouikhalene, B. (2021). Fake News Detection Approach Using Parallel Predictive Models and Spark to Avoid Misinformation Related to Covid-19 Epidemic. In *Intelligent Systems in Big Data, Semantic Web and Machine Learning* (pp. 179-195). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-72588-4_13

[15] Malla, S., & Alphonse, P. J. A. (2021). COVID-19 outbreak: An ensemble pre-trained deep learning model for detecting informative tweets. *Applied Soft Computing*, *107*,107495. https://doi.org/10.1016/j.asoc.2021.107495

[16] Qasim, R., Bangyal, W. H., Alqarni, M. A., & Ali Almazroi, A. (2022). A fine-tuned BERT-based transfer learning approach for text classification. *Journal of healthcare engineering*, *2022*. https://doi.org/10.1155/2022/3498123

[17] Khadhraoui, M., Bellaaj, H., Ammar, M. B., Hamam, H., & Jmaiel, M. (2022). Survey of BERT-based models for scientific text classification: COVID-19 case study. *Applied Sciences*, *12*(6), 2891. https://doi.org/10.3390/app12062891

[18] Shahi, T. B., Sitaula, C., & Paudel, N. (2022). A hybrid feature extraction method for Nepali COVID-19-related tweets classification. *Computational Intelligence and Neuroscience*, *2022*. https://doi.org/10.1155/2022/5681574

[19] Ahmed, M., Hossain, M. S., Islam, R. U., & Andersson, K. (2022). Explainable Text Classification Model for COVID-19 Fake News Detection. *Journal of Internet Services and Information Security (JISIS)*, *12*(2), 51-69. DOI:10.22667/JISIS.2022.05.31.051

[20] Rabby, G., & Berka, P. (2023). Multi-class classification of COVID-19 documents using machine learning algorithms. *Journal of Intelligent Information Systems*, *60*(2), 571-591. https://doi.org/10.1007/s10844-022-00768-8

[21] Etaiwi, H. A. (2022). Empirical Evaluation of Machine Learning Classification Algorithms for Detecting COVID-19 Fake News. *International Journal of Advances in Soft Computing & Its Applications*, *14*(1). DOI: 10.15849/IJASCA.220328.04

[22] Malla, S., & Alphonse, P. J. A. (2022). Fake or real news about COVID-19? Pretrained transformer model to detect potential misleading news. *The European Physical Journal Special Topics*, *231*(18), 3347-3356. https://doi.org/10.1140/epjs/s11734-022-00436-6

[23] Ravichandran, B. D., & Keikhosrokiani, P. (2023). Classification of Covid-19 misinformation on social media based on neuro-fuzzy and neural network: A systematic review. *Neural Computing and Applications*, *35*(1), 699-717. https://doi.org/10.1007/s00521-022-07797-y

[24] Gams, M., & Kolenik, T. (2021). Relations between electronics, artificial intelligence and information society through information society rules. *Electronics*, 10(4), 514. https://doi.org/10.3390/electronics10040514

[25] Janko, V., Slapničar, G., Dovgan, E., Reščič, N., Kolenik, T., Gjoreski, M., ... & Luštrek, M. (2021). Machine learning for analyzing non-countermeasure factors affecting early spread of COVID-19. *International Journal of Environmental Research and Public Health*, 18(13), 6750. https://doi.org/10.3390/ijerph18136750

[26] Hossain, M. M., Anselma, L., & Mazzei, A. (2023). Exploring sentiments in summarization: SentiTextRank, an Emotional Variant of TextRank. In CEUR WORKSHOP PROCEEDINGS (Vol. 3596, pp. 1-5). CEUR-WS. https://hdl.handle.net/2318/1950757

[27] Caldo, D., Bologna, S., Conte, L., Amin, M. S., Anselma, L., Basile, V., ... & De Nunzio, G. (2023). Machine learning algorithms distinguish discrete digital emotional fingerprints for web pages related to back pain. *Scientific Reports*, 13(1), 4654. https://doi.org/10.1038/s41598-023-31741-2

[28] Spirovski, K., Stevanoska, E., Kulakov, A., Popeska, Z., & Velinov, G. (2018, June). Comparison of different model's performances in task of document classification. In *Proceedings of the 8th international conference on web intelligence, mining and semantics* (pp. 1-12). https://doi.org/10.1145/3227609.322766

[29] Pennington, J., Socher, R., & Manning, C. D. (2014, October). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 1532-1543). https://aclanthology.org/D14-1162.pdf

[30] Sparck Jones, K. (1972). A statistical interpretation of term specificity and its application in retrieval. *Journal of documentation*, *28*(1), 11-21. https://doi.org/10.1108/eb026526

[31] Jaderberg, M., Simonyan, K., Vedaldi, A., & Zisserman, A. (2016). Reading text in the wild with convolutional neural networks. *International journal of computer vision*, *116*, 1-20. https://doi.org/10.1007/s11263-015-0823-z

[32] LeCun, Y., Touresky, D., Hinton, G., & Sejnowski, T. (1988, June). A theoretical framework for back-propagation. In *Proceedings of the 1988 connectionist models summer school* (Vol. 1, pp. 21-28). https://www.researchgate.net/profile/Yann-Lecun/publication/2360531_A_Theoretical_Framework_for_Back-Propagation/links/0deec519dfa297eac1000000/A-Theoretical-Framework-for-Back-Propagation.pdf

[33] Scherer, D., Müller, A., & Behnke, S. (2010, September). Evaluation of pooling operations in convolutional architectures for object recognition. In *International conference on artificial neural networks* (pp. 92-101). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-15825-4_10

[34] Wang, W., Zhu, D., Wang, X., Hu, Y., Qiu, Y., Wang, C., ... & Scherer, S. (2020, October). Tartanair: A dataset to push the limits of visual slam. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 4909-4916). IEEE. 10.1109/IROS45743.2020.9341801

[35] Lu, G., Businger, M., Dollfus, C., Wozniak, T., Fleck, M., Heroth, T., ... & Lipenkova, J. (2023). Agenda-setting for COVID-19: A study of large-scale economic news coverage using natural language processing. *International Journal of Data Science and Analytics*, *15*(3), 291-312. https://doi.org/10.1007/s41060-022-00364-7

[36] ADEMI, A. (2016). EVALUATION OF THE MODELS USED TO CREATE VECTOR SPACE REPRESENTATION OF WORDS. *SCIENCE, INNOVATION NEW Technology*, 31.

[37] Iwayama, M., & Tokunaga, T. Department of Computer Science Tokyo Institute of Technology. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=e10595e5fa4c94df2ceb8867327ad2fa0825c089

[38] Grün, B., & Hornik, K. (2011). topicmodels: An R package for fitting topic models. *Journal of statistical software*, *40*, 1-30. https://www.jstatsoft.org/article/view/v040i13

[39] Klaiber, M. (2021). A fundamental overview of sota-ensemble learning methods for deep learning: a systematic literature review. *Science in Information Technology Letters*, *2*(2), 1-14. https://pubs2.ascee.org/index.php/sitech/article/view/549