Journal of Computer Science

Research Article

Implementing a Hybrid Parallel Framework Utilizing Machine and Deep Learning for Rapid Rumor Detection in Social Media

Neetu Rani[1], Amit Kumar Bhardwaj[2], Shaily Jain[3], Chander Prabha[3], Prakash Srivastava[4], Mohammad Zubair Khan[5], Jeehaan Algaraady[6], Mohammad Mahyoob Albuhairy[7] and Abdulaziz Alblwi[8]

[1]*Department of Computer Science and Engineering, Chandigarh University, Mohali,* India
[2]*LM Thapar School of Management, Thapar Institute of Engineering and Technology, Patiala, India*
[3]*Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India*
[4]*Department of Computer Science and Engineering, Graphic Era (Deemed to be University), Dehradun, India*
[5]*Faculty of Computer and Information Systems, Islamic University of Madinah, Madinah 42351, Saudi Arabia*
[6]*Languages and Translation College, Taiz University, Taiz, Yemen*
[7]*Energy, Industry, and Advanced Technologies Research Center, Taibah University, Madinah, Saudi Arabia*
[8]*Department of Computer Science and Information, Applied College, Taibah University, Madinah, 42353, Saudi Arabia*

Article history
Received: 15-12-2024
Revised: 25-05-2025
Accepted: 17-07-2025

Corresponding Authors:
Mohammed Zubair Khan
Faculty of Computer and Information Systems, Islamic University of Madinah, Madinah 42351, Saudi Arabia
Email: zubair.762001@gmail.com
And:
Chander Prabha
Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India
Email: prabhanice@gmail.com

Abstract: Numerous businesses have faced significant repercussions due to the widespread dissemination of false information and rumors across social media platforms. The impact of fake news extends to tarnishing public perception, damaging corporate reputations, disrupting communities, undermining governmental integrity, exposing companies to risks, and posing a grave threat to social cohesion. This research article delves into the endeavours of prominent researchers focused on utilizing machine learning for rumour detection. Additionally, it explores a newly proposed framework wherein several established methods viz. Adaboost, Hard Voting, Gradient Boosting, and Random Forest; and a novel hybrid deep learning model CNN + BiLSTM + BiGRU operate simultaneously to identify rumours in a parallel environment. Utilizing time-series vector representations of Twitter, Facebook and FakeNewsNet datasets, this study suggests an ensemble approach for rumor detection. The proposed model demonstrates better accuracy, f1-score, recall, and efficiency compared to existing models and minimizes time consumption due to parallel computational capabilities.

Keywords: Rumor Detection, Social Media, Machine Learning, Parallel Computing

Introduction

In today's age of powerful social media, a famous quote gains even more relevance: "A lie travels around the globe while the truth is getting ready for the journey". This holds particularly true for companies that fear the potential damage of false news on their reputation. Therefore, closely monitoring online postings related to their business becomes essential to mitigate the risks. Social media websites have evolved into critical contexts for sharing data worldwide. Popular microblogging sites like WhatsApp, Facebook, Instagram, Twitter, and Sina Weibo dominate the transmission of information (Can and Alatas, 2019). Though social networking emerged in the 1960s, its popularity has risen in recent years as a primary mode of communication, fueled by the proliferation of novel gadgets and internet-based learning. Online media permit information to spread quicker than ever, but ensuring its integrity remains a significant challenge. Malicious rumors and misinformation can swiftly permeate social networks through various communication channels (Chen *et al*., 2021). Surprisingly, fake news spreads much more rapidly than real news, as confirmed by a recent study by MIT researchers, (Guadagno and Guttieri, 2021). For businesses, rumors, especially negative ones, can seriously impact their reputation. In the context of stock markets, they have historically affected share prices in response to rumors about business announcements, earnings expectations, and undervalued stocks (Yahya, 2022).

The power of social media to sway opinions and influence events was evident during the 2016 United States constitutional elections, where politicians and their supporters extensively used social media platforms to promote their agendas. Various stories, amounting to 529, revolving around Donald Trump and Hillary Clinton spread like wildfire through social media, significantly influencing the election outcome (Azeez and Jimoh, 2023). Numerous instances illustrate the detrimental effects of rumors on companies and individuals alike. For example, in 2013, the news agency was hacked to spread false information about White House on their social media account, causing social anxieties and an immediate drop in the stock market (Nasery *et al*., 2023). In 2014, 92 rumors about the "Malaysia Airlines Flight MH370" were spread on Sina Weibo, affecting people's emotions and hindering their understanding of the actual condition (Alsaleh, 2022). Brands are not immune to false info either; Nestlé India faced declining sales in 2015 due to rumors about one of its flagship products containing harmful substances (Gandhi, 2022).

Rumors about semiconductor production shortages affected iPhone production and other electronic industries, damaging company reputations and lowering stock prices (Alkhodair *et al*. 2020). The study focused on the impact of stock market rumors shows that they have a strong effect on the price of stock especially in regions such as Istanbul where rumors of earnings expectations by investors and foreign investors have more pronounced influence (Bamiatzi *et al*., 2016). In a different case study, Sahara Bank also met its crises because of rumor on the disappointment between investors who aspire to higher rates of returns (Sandhu and Saluja, 2023). Moreover, inaccurate rumors may have tremendous effects on businesses, and the example of the event in 2019 with a tweet about a Tesla autonomous car crashing into a robot car at CES convention leading to the decrease in the Tesla stock value shows that rumors may have serious consequences (Strauss and Smith, 2019). There are studies that have demonstrated how stock market rumors control price rates and decrease market performance as with the case of the Indonesian market (Wirama *et al*., 2017).

Rumors have a great capacity to impact certain industries and cause severe consequences to the businesses. As an example, Chinese workers and pipeline projects had unsubstantiated allegations in the oil industry in Sudan that resulted in tense relationships and economic losses. In the same context, a research paper noted the devastating consequences that the hens representing the poultry farming industry faced when an information campaign against hens being a source of coronavirus lead to a disastrous 80 percent decline in the sales of chicken meat and affected legions of small farmers. The following are the examples of the rumors meant negatively to businesses in reality.

United Airlines in 2017 suffered with a PR crisis after a rumor spread online that a doctor was either injured or killed after he was forced to leave a United Airlines overbooked flight, consequently, decreasing the stock value of the company by a large margin.

Starbucks (2018): An incident results in detention of two African American men in Philadelphia at a Coffee shop for waiting on a friend resulted in accusations of racial prejudice, after a viral video of the event. This caused a boycott across the country and a temporary closure of Starbucks stores to racial bias training.

Tesla (2019): Presumptions that the company killed someone when using Autopilot damaged the stock price of the company momentarily. Tesla responded to these allegations by stating their focus on the safety practices.

Apple (2020): Rumour surrounding the release of a new model that was cheaper than the previous one made people stop buying existing models in anticipation of the new one which later did not sell as expected.

Johnson and Johnson (1982): There were panic sales and a false alarm on cyanide contained in Tylenol. This resulted in the innovation of tamper-evident packaging and the increased safety precautions of the industry.

Coca-Cola (2003): A rumor that coca cola products had dangerous pesticides circulated in India and sales plummeted as people believed that the brand was not doing well and that the rumors were true despite the fact that no pesticides were detected.

Nirav Modi (2018): The case of Nirav Modi, a jeweller, as a result of allegedly defrauding banks generated a loss of consumer confidence, sales stagnated, and his branches subsequently fell.

WhatsApp Child Kidnapping Rumors (2018): Misinformation that had spread through WhatsApp, centered on child kidnapping across India led to mob violence that hampered business operations.

Snapdeal and Aamir Khan Controversy (2015): Snapdeal brand ambassador actor Aamir Khan drew a lot of criticism on himself due to his controversial public statement, resulting in a backlash against the company, which led the latter to abandon their relationships with the former.

Baba Ramdev and Patanjali Rumours: Since its foundation, Patanjali Ayurved Limited led by Baba Ramdev has over the years had to deal with various rumors regarding misleading advertisements and quality of products it offers, a factor that has hurt its reputation and sales.

These cases are vivid examples of how crucial it is that companies are expected to respond to rumors quickly by providing clear, precise information that will help them prevent losing long-term credibility. False information is a very dangerous subject that needs to be detected and countered to defend a business reputation and position in the market.

This is because, to effectively detect rumors in the social media in real-time, this research paper concentrates on exploiting temporal features of Twitter and Facebook data. To train a classification model, the proposed research extracts tweet creation timestamps immediately to be used as such functional characteristics. On this basis an enhanced deep learning collaborative-based manifold time series study model is generated to become effective on the recognition of rumors in social media. Conversations are converted into time-series vectors, each indicating the total number of reactions throughout the discussion. These time-series vectors are then fed into the deep-learning models, allowing for efficient rumor identification based on the temporal properties of the data. The main points of our proposed approach are:

- Computational cost of the suggested approach decreases substantially since it does not require the analysis of tweet content or user social activities. Instead, it only requires the timestamps of tweets to extract features. The obtained feature set is also composed entirely of numeric values, which is extremely well-suited with classification models
- The suggested ensemble approach improves the performance of classification models by using a majority-voting framework inside the ensemble to use the unique advantages of various networks. Overall performance and accuracy are enhanced by this method

The Fake News datasets were castoff to authenticate the proposed method: (Bisaillon, 2018). The performance outcomes from this validation illustrate the efficacy of the proposed method. The proposed technique, allowing for fast and efficient identification of rumors in social media, analyses the temporal aspects of Twitter and Facebook data via an ensemble-based classification model.

Related Work

This section delves into the abundant tools, approaches, and several machines and deep learning algorithms researchers have presented to address the challenge of detecting rumors.

Machine Learning Approach to Rumour Detection

Machine learning methods involve training machines to analyze data and produce more accurate results proficiently. These algorithms understand patterns in the data and extract relevant information from accessible datasets. The primary goal of these methods is to augment computer programs' ability to access and learn from information without human intervention. They rely on previous data to make predictions for future outcomes. The aim of machine learning is to come up with automatic learning methods devoid of human interference. However, without prior acquaintance, this framework could not detect rumors in trending news.

Hamidian and Diab (2016) presented a J48 classified and trained a model on the WEKA platform, achieving an f-score of 82%. However, the proficiency of the pre-processing task was limited due to constraints with the WEKA tools. Zhao *et al*. (2015) introduced a phrase investigation-based technique for detecting rumors, which clustered similar phrases with an accuracy of 52%. However, the work suffered from slow response and required manual labeling.

Zubiaga *et al*. (2019) utilized CRF for rumor recognition based on contents and socials feature, achieving 60% f1-score on tested dataset. Vijeev *et al*. (2018) has utilized the NLP algorithm to identify rumors based on data and the characteristics of the users and where the Chi-square approach has been used to vigor the best features. SVM, NB and RF classes have been tested on the PHEME dataset with the highest and an accuracy of 74.6% with an algorithm.

Chen (2021) detect rumors from Weibo microblogs based on people behavior by means of an un-supervised model with the slot time of seven as well as achieved an accuracy of 71%. Suissa *et al*. (2022) proposed an auto-detection methodology to find rumors, employing contents, users, and features for training several classifiers like SVM, DT, Bayes network, and J48. The J48 achieved 88.9% precision on a handmade feature. Twitter was monitored to spot events. Aljamal *et al*. (2025) also developed a method for detecting long-standing rumors, but it could not able to find rumors from the news without prior data.

Deep Learning Methods for Rumour Detection

Deep learning methods, a subset of machine learning, have become a prominent research topic due to their outstanding performance in various applications, such as NLP and text mining. Deep learning models have the ability with the purpose of discovering concealed features in text and images. Among the widely used artificial neural network standards include Recurrent Neural Network (RNN) and the Convolutional Neural Network (CNN) (Chauhan and Singh, 2018).

Rathakrishnan and Sathiyanarayanan (2023) explored deep learning methods to detect rumors and

DOI: 10.3844/jcssp.2025.3005.3018

the causes of the same on comments in the tweet. The system separates topics and classifies them into the following ranks deny, support, query, and comment. Comments considered as negative are categorized as aggressive, viciousness, misogyny, and hatred mongering based on Improved Deep Learning Neural Network (IDLNN). The improved ANISPIMF can perform better on the COVID-19 dataset with 0.6, 0.7 and 1% gain in precision, recall and accuracy respectively compared to current methodologies.

Tan *et al*. (2023) made a comparison of approaches to recognition of rumors on three aspects: Feature selection, model structure, and research methods. categorizes the methods into contents, socials and proliferation structure features. The study also contrasts the deep learning models such as CNN, RNN, GNN and Transformer on the basis of the model. It examines 30 works into seven rumor detection techniques (propagation trees, adversarial learning, cross-domain methods, multi-task learning, unsupervised and semi-supervised methods, and based knowledge graphs).

Ma *et al*. (2018) introduced a tanh-RNN, GRU, and LSTM-based model for rumor detection, using a transmission tree as input rooted from the initial receptive post. The model was tested on Twitter and Sina Weibo datasets, yielding experimental results. Nguyen *et al.* (2017) united CNN and LSTM for rumor discovery, utilizing CNN for expression extraction and LSTM for tweet representation. They evaluated their model on Snopes and Urban Legends data, achieving an accuracy of 82%.

Alkhodair *et al*. (2020) detected rumors from news on Twitter by combining Word2Vec and LSTM-RNN methods. It tested on synthetic and social characteristics of the PHEME dataset and achieved an 80 percent accuracy. Asghar *et al*. (2021) created a model of detection, which was composed by BiLSTM and CNN together. Moreover, it was experimented with approx. The PHEME dataset had 6000 tweets concerning news and their accuracy was 86 percent and a Chi-Square statistical test confirmed that their work was effective. Nonetheless, the model only paid attention to text-based characteristics and English terms.

Liu *et al*. (2024) presented the BiGRU-CNN model that allowed them to classify the Chinese tweets into multiple categories achieving 79 percent accuracy. Having used the CNN algorithm with the dataset of movie reviews, Sankar *et al*. (2020) managed to achieve 82 percent of accuracy estimating opinions of people regarding the selected movies, implementing words-based encryption. Zeng *et al*. (2021) applied LSTM-based rumor detection methodology, achieving an accuracy of 85 percent on Shango dataset which was collected on China Science Communication.

Li *et al*. (2022) proposed a rumour tracking integration model (RL-ERT) on a deep reinforcement learning framework, in which many elements are integrated with various elements using the weight adjustment strategy network and some social factors are used to enhance the model performance. A similar model was proposed by Sridhar and Sanagavarapu (2021) who used data taken off the Kaggle online resource to detect rumors with an impressive accuracy of 97%. The BiGRU model to make early detection of rumors was proposed by researchers Yang *et al*. (2022), which achieved an 88 percent/91 percent accuracy on the Twitter dataset and Sina Weibo dataset, respectively.

Cen and Li (2022) proposed BiLSTM network system to detect fake information on social media sites and identify them automatically. The model was able to decipher semantic information of an input text data and reconstructed three types of social information, that is, people information, communication content, and Weibo context. The method of vectorization was by using two-word vectors, and classification was provided through a Soft Max deposit. The model reached 94.9, 94.1, 94.4 and 93.9 in accuracy, recall, precision and F1 values respectively which are better than that of existing approaches. Also, scientists in (Sadiq *et al*., 2021) used deep learning in reinforcement learning models to detect rumors and assessed the effectiveness of the proposed research on the PHEME and RumourEval() datasets. The models gave a 95 percent accuracy of RumourEval() and 94 percent of the PHEME dataset. Furthermore.

Table 1 compares the various existing techniques of rumor detection. Examining Table 1 reveals that earlier research (Cen and Li, 2022; Li *et al*., 2022; Gao *et al*., 2020) concentrated on manual data labeling, tackled issues of slow response time, and faced high computation costs, respectively.

Consequently, the need arises for a model capable of autonomously identifying rumors in their early stages, ensuring heightened accuracy, reducing computation time, and automatically classifying data into distinct categories. Moreover, an essential requirement is a hybrid model for automatic rumor detection with enhanced accuracy. Despite the utilization of advanced machine and deep learning techniques by previous researchers, our understanding indicates that integrating these deep learning techniques has not been explored before to alleviate their potential drawbacks.

Table 1: Existing rumour detection studies comparison

Author and Year	Platform	Algorithm/ Technique	Features	Results (Accuracy)	Drawbacks
Cen and Li (2022)	Weibo	Multi-BiLSTM	Content, user and social features	94%	Manual crawling and labelling
Sadiq *et al*. (2021)	Twitter	CNN-BiLSTM	Content, network, Twitter-specific features	92%	Only single feature TF-IDF was used
Li *et al*. (2022)	Twitter	Clustering method	Content-features	90% Precision	Response time was slow
Vijeev *et al*. (2018)	PHEME	RF, SVM, NB	Content, user features	74.6%	Few features were used
Ma *et al*. (2018)	Twitter	RNN	Word2Vec	73%	Require Improvement in accuracy
Asghar *et al*. (2021)	PHEME	BiLSTM-CNN	Word2Vector and FastText	86.12%	Require Improvement
Gao *et al*. (2020)	PHEME and Twitter	LSTM	Twitter-based features as well as metadata	68%	High Cost of computation
Ajao *et al*. (2018)	PHEME	LSTM- CNN	LSTM, LSTM with dropout, LSTM-CNN	8	Performance needs improvement

To address these limitations, our proposed approach introduces a hybrid model combining CNN + BiLSTM + BiGRU for rumor detection. The proposed work's main point is classifying tweets into rumors or non-rumors through binary classification, utilizing a time series method to differentiate between the two categories. The algorithms implemented in our research operate both serially and in parallel to augment efficiency. The CNN + BiLSTM + BiGRU hybrid model is justified due to its complementary strengths in feature extraction and temporal modeling, improved accuracy and generalization across diverse datasets and reduced overfitting and improved efficiency because of BiGRU's lightweight design.

Methods

The proposed framework in methodology contains two main components: Pre-processing and ensemble model. During the first step, raw Twitter and Facebook data are converted into the requisite format and then further fed into an ensemble learning model for the final prediction of whether a tweet is a rumor or non-rumor.

Figure 1 depicts the structure of our suggested paradigm. The framework accepts conversations for the input phase, where every conversation contains source-tweet and its related activities. Every tweet and text from Facebook respectively during the data pre-processing stage is processed and determined the value of its creation timestamp. After parsing every tweet, time-series data for various intervals is created and cleaned.

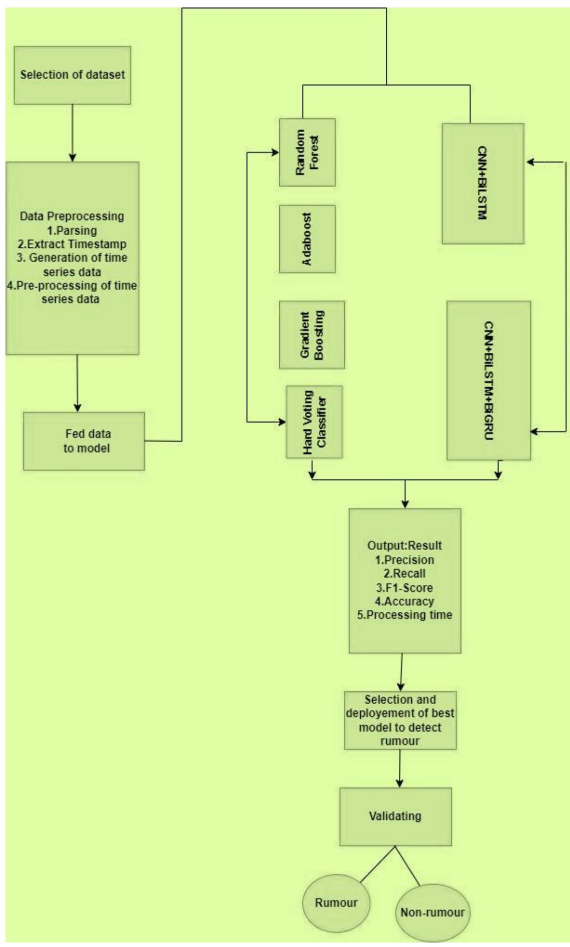

Fig. 1: Proposed Methodology Framework

Further, for data pre-processing, data sparsity is performed, data is normalized, and its duplicity is removed. The data pre-processing step used Darts library for data cleaning which helps in detecting anomalies in time series. The ensemble model is then fed with this cleaned data. The proposed research contains advanced machine learning and deep learning algorithms where every classifier produces results. Finally, we used the majority-voting approach to decide whether each prediction made by those base learners should be classified as a rumor or a non-rumor by adding up all the prediction outcomes and choosing 0 for the non-rumor and 1 for the rumor.

The input phase is an initial module that extracts datasets from different platforms for experimental purposes. The dataset collected from the web platform is processed in the second phase, pre-processing. The information is converted into a machine-understandable format in the third phase via various feature extraction methods. Once they have been extracted in the fourth section, several more advanced machine learning techniques are employed and compared against prior research which employed a similarly structured dataset. A new framework is created for rumor identification in the fifth module, and it further employs deep learning approaches, including CNN + BiLSTM + BiGRU.

This proposed work uses binary classification to classify tweets as rumor and non-rumor. Methodology used to design a new model is discussed in detail:

Input Phase

The first phase in the proposed research work is the input phase; here, the dataset is extracted for experimental purposes from various platforms. The datasets used are open source and are publicly accessible from online media (Bisaillon, 2018).

The features of the datasets used in this research are shown in Table 2. The Id field explains the unique identifier, the title shows what the article is about, the author identifies who tweeted about it, the date field indicates when it was published online, the text field shows the information that was written in the article, and the label value specifies whether it is a rumor or not. The values lie between 0 and 1, where 1 denotes reliable data and 0 denotes unreliable data.

Table 2: Information about datasets

Twitter Dataset		Facebook Dataset		FakeNewsNet Dataset	
Field	Detail	Field	Detail	Field	Detail
Id	The unique number of news article	Title	Article Heading	Title	Full news article body including images
Title	Article label	Text	Text involved in the article	Text	Tweet-level reactions, metadata, and social graph data
Author	Tweet Author	Subject	Theme of the article	Subject	Politics / Entertainment / Economy
Text	Information includes in the article	Date	Date whenever document dispatched	Date	2018-09-12

Creation of Time-Series Content

Temporal traits of twitter and Facebook data are applied in the proposed research work to come up with timely detection of rumors online media. Twitter chats result in time series data, with a tweet list making up a conversation. Further, time series vectors are created and provided as input for model creation. We converted every conversation of the dataset into a time series vector by providing different time intervals. Afterward, the successful conversion of all conversations signifies that every vector signifies the entire discussion.

The subsequent Fig. 2 demonstrates the process of creating a time series vector for each conversation.

Pre-Processing

The subsequent phase in the data preparation pipeline focuses on reducing the sparsity inherent in time-series data. In this context, sparsity arises because the vector length of each data sample is defined based on the longest conversation associated with the corresponding time point TTT. This results in shorter conversations being padded, leading to high-dimensional yet sparsely populated vectors that can negatively impact model training and performance. Social media text data, which forms the basis of this study, is often unstructured, informal, and noisy. These involve use of irrelevant symbols, emojis, overuse of punctuations marks, slangs or grammatic inconsistencies, URLs, mention of the user, and unrelated hashtags which do not add any semantics to the message. This noise has the potential to mask the real signal of the data, as well as derail subsequent Natural Language Processing (NLP) tasks.

Accordingly, data cleaning that will eliminate undesirable factors (Sarker, 2017). The sharper and reliable text classification as well as time-series forecasting models have been noted to reflect effectively on pre-processing of data (Xu *et al*., 2024). The methods commonly used in this aspect are lowercasing, stop word elimination, stemming or lemmatization, normalization of tokens and deletion of non-alphanumeric characters.

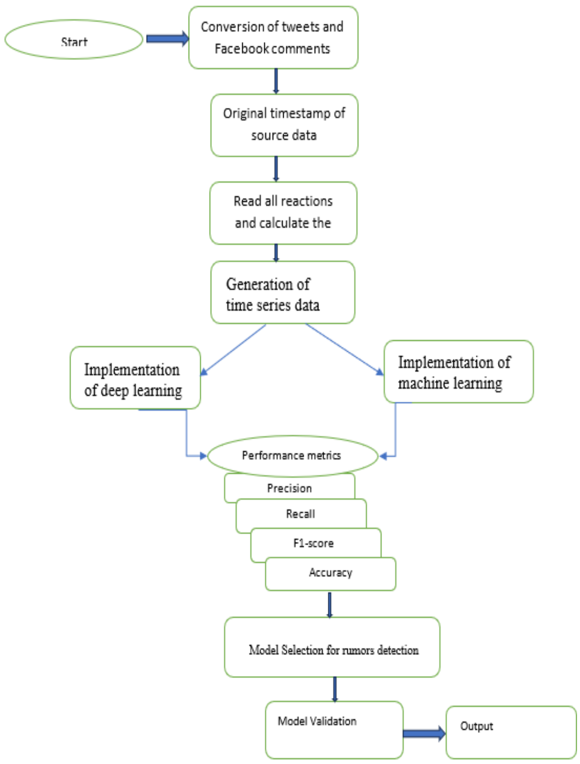

Fig. 2: Process of creating a time series vector for each conversation

Also, the frequency and time of the posts are the two attributes that are kept unchanged to retain temporal dynamics required in time-series analysis. In such cases, the proposed methodology will use the Darts library which is a complete Python framework dedicated to time-series modeling. The workflow has the advantage of preserving the temporal nature of conversations around social media whilst at the same time removing noise thus enriching the quality of the information available to support the model.

Feature Extraction

The feature extraction step is the most important after the data preprocessing process is performed. It includes transforming cleaned textual data into numerical form. In particular, feature extraction is the process of transforming the preprocessed sentences that were originally expressed as arrays of integers into dense, meaningful vectors. All these numerical vectors summarize the semantics and syntax phrases presented in the text. Embedding techniques are used to undergo this transformation. Embedding is a process of mapping words or token into a continuous vector space and the semantically similar words are positioned near to one another. This enables models to learn about linguistic context and n-grams between words, which is vital in activities like classification, sentiment analysis and fake news detection (Fazil *et al.*, 2021; Kumar *et al*., 2025). Such word embeddings can be generated in different ways. TF-IDF (Term Frequency-Inverse Document Frequency) and Count Vectorizer are typical examples of traditional machine learning pipelines. Rather, text is translated into sparse vector representations according to word frequency and word significance in the corpora.

On the contrary, in case of deep learning models, more sophisticated embedding methods are used. Among them is GloVe (Global Vectors for Word Representation) which uses global word co-existence data of a huge corpus to produce condensed word vectors. It has a more profound expression of textual semantics which means deep learning algorithms can learn intricate language patterns better. In the given work, the classic methods of feature extraction (TF-IDF, CountVectorizer) are employed along with the deep learning-related word embedding (GloVe). This mixed method provides the fuller comparison and assessment of various techniques of representation, which allow improving the model performance in a variety of conditions of the experiment.

Process Creation of Time-Series Vector

First step involves converting raw tweet data into a usable format, likely involving preprocessing like cleaning and tokenization. Then locate the timestamp of the original tweet and being analyzed. After that we gather all reactions (like replies, retweets) to the source tweet and note the timestamps associated with them to determine the frequency of reactions for specific time intervals, which is key for time-series data analysis.

We create time series dataset from the processed reaction data to track changes and patterns over time. We choose some machine learning or deep learning model suitable for analyzing the time-series data and detecting rumors and apply the selected model to the data. Lastly assess the output of the model with the help of such indicators as precision, recall, F1-score, and accuracy and give the results or predictions whether this tweet is a rumor. This flowchart outlines a comprehensive process for analyzing tweet data to detect and classify rumors.

Results and Discussion

The feature extraction and dataset cleaning process are done after which different advanced machine and deep learning techniques are used. Later, the dataset required the deep learning methods (CNN, BiLSTM, BiGRU). In this study, the results of the proposed hybrid model were tested on accuracy as the main measure of performance. To validate the suggested model, the findings gained due to the confusion matrix will be compared with another classification effectiveness.

Experimental Results

Table 3 illustrates the results of various advanced machine-learning algorithms on twitter dataset. Firstly, we run various algorithms on Twitter and Facebook datasets in serial mode by allocating different timestamps to them. Then, algorithms are run in parallel mode to demonstrate their results.

The above results demonstrate various performance metrics of different algorithms and the processing time given to every machine learning algorithm. The implementation results and evaluation of metrices on Datasets of Twitter and Facebook are presented in Figs. 3 and 4, respectively. On the accuracy, precision, recall, and f1-score, the results indicate that the hard voting classifier is better with respect to T = 20 ns. Gradient boosting does well at 97 percent in terms of precision. When exploring the proposed research, the majority voting classifier had an accuracy, precision, recall as well as an F1 score amounting to 96 per cent, which was similar to the performance of individual models incorporated in the ensemble. When running algorithms in parallel, the overall processing time would equal the processing time of the slowest algorithm, which is 20 ns for the Hard voting classifier. Regarding the performance metrics running algorithms in parallel does not inherently affect their values. Each algorithm's performance metrics would remain the same as if they were run individually.

Gradient Boosting is the algorithm that provides good results for the majority voting classifier. It has the highest performance metrics among all the algorithms. With an accuracy of 93%, precision of 94%, recall of 95%, and F1-score of 95%, Gradient Boosting outperforms Random Forest, AdaBoost, and the Hard Voting classifier. The results on the Facebook dataset metrics, as shown in Table 4. The results show that the Gradient Boosting algorithm achieves the highest accuracy, 96%, in the Twitter and 93% in the Facebook datasets.

Therefore, when using a majority voting classifier, the Gradient Boosting algorithm would be considered better in this case. When running algorithms in parallel, the overall processing time would equal the processing time of the slowest algorithm, which is 16 ns for the Hard voting classifier. If we run both the algorithms applied to the in parallel, Twitter and Facebook datasets lead to the time of the overall processes being equal to the maximum computer processing time. The processing time of Facebook data in the case can go to a maximum of 16 ns.

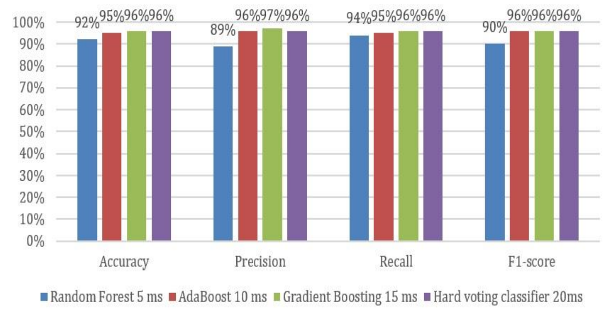

Fig. 3: Results on the Twitter dataset

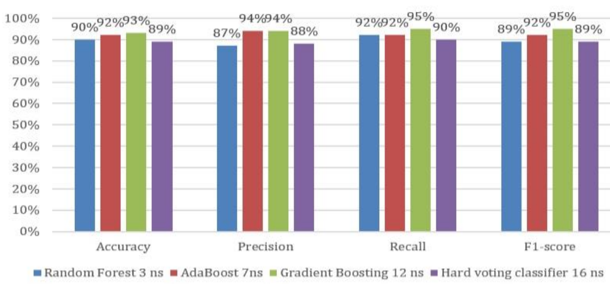

Fig. 4: Results on the Facebook dataset

Table 3: Results on the Twitter dataset using ML approaches

Approach	Processing Execution Time	Accuracy	Precision	Recall	F1-score
Random Forest	5 ms	92%	89%	94%	90%
AdaBoost	10 ms	95%	96%	95%	96%
Gradient Boosting	15 ms	96%	97%	96%	96%
Hard voting classifier	20ms	96%	96%	96%	96%

Table 4: Results on the Facebook dataset using ML approaches

Approach	Processing Execution Time	Accuracy	Precision	Recall	F1-score
Random Forest	3 ms	90%	87%	92%	89%
AdaBoost	7ms	92%	94%	92%	92%
Gradient Boosting	12 ms	93%	94%	95%	95%
Hard voting classifier	16 ms	89%	88%	90%	89%

Proposed Hybrid Framework (Cnn + Bilstm + Bigru)

Figure 1 illustrates the suggested framework. Some libraries and tools are utilized for model development, including Keras and Tensor Flow. Comments are transformed into segments before the model is trained. The trained word vector was then provided as an input for the model after the comments had been turned into a word vector using custom embedding. The sentence's integer vector is changed into a dense vector throughout the feature extraction procedure. The text is transformed into numbers via embedding. The dataset was then subjected to the application of CNN + BiLSTM + BiGRU. Since our problem is binary, the dataset was finally subjected to sigmoid activation function analysis. This function's range of values is -1 to 1 and 0 to 1. The shape of the sigmoid function looks like a curve, as shown in Fig. 5. The graph shown in Figs. 6 and 7 shows a steady reduction in training loss and increased accuracy, which specifies that the model studies well from the training data. The validation loss decreases initially and accuracy increases then stabilize, showing that the model is generalizing well on unseen data with minimal overfitting. The gap between the two curves is small, indicating balanced performance.

Various standards are used in the compilation of the model to determine the performance of the recommended work. Such variables, such as Adam optimizer to update network weights over and over again, depend on training data. There is a further application of a binary-cross-entropy method that compares the predictions that are empirical and foregone. Lastly, we used the predict method in Python, with which we can predict including the labels of the data values using the training model. The success of our proposed method is shown by its ability to discriminate rumors and non-rumors. Figure 8 shows the confusion matrix of the classification. The confusion matrix will have positive, that is, the truth and negative, or fake news.

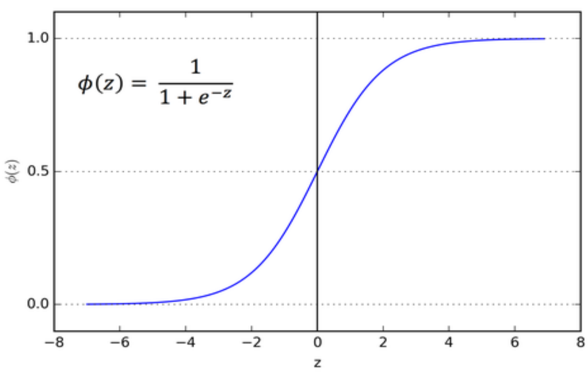

Fig. 5: Sigmoid activation function

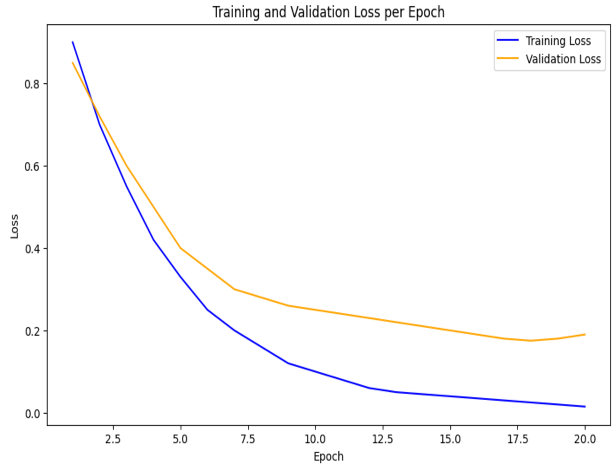

Fig. 6: Training and Validation Loss per epoch

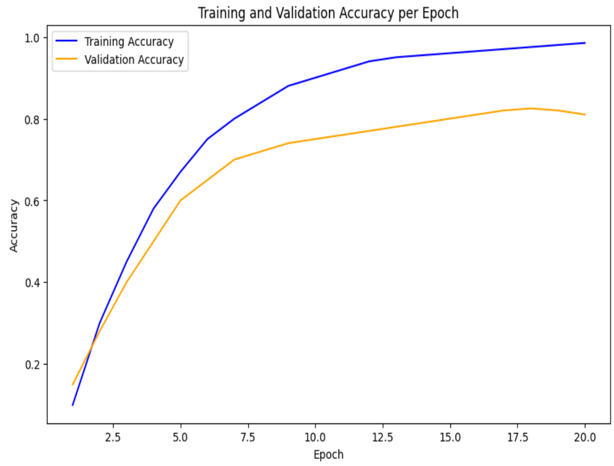

Fig. 7: Training and Validation Accuracy per epoch

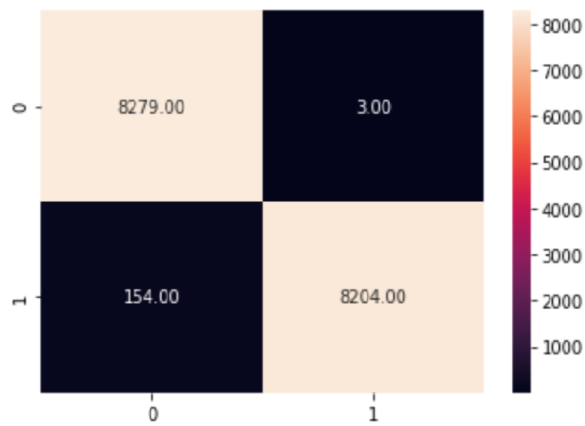

Fig. 8: Confusion Matrix

Proposed Hybrid Framework Results

To assess the effectiveness of the proposed hybrid model, we compared its performance against a range of traditional and deep learning-based classifiers that have been previously evaluated on the same datasets (Fazil *et al.*, 2021; Elaoud *et al.*, 2020). The models considered for comparison include traditional machine learning algorithms such as Logistic Regression, Decision Trees, AdaBoost, and ensemble models like the Voting Classifier, as well as deep learning approaches including CNN, BiLSTM, and SVM-based models as shown in Tables 5-7. The proposed CNN + BiLSTM + BiGRU model achieves the highest scores across all evaluation metrics Accuracy, Precision, Recall, and F1-score—on both Facebook and Twitter datasets. In traditional models such as Logistic Regression (F1-score: 91%) or even relatively strong learners like Decision Trees (F1-score: 94%), the hybrid model shows a clear margin of improvement, ranging from 3 to 8 percentage points in F1-score. Interestingly, standalone deep learning models like Wang-CNN and Wang-BiLSTM perform significantly worse. Figures 9-12 graphically represent the comparative analysis of accuracy, precision, recall and F1-score respectively. This suggests that neither CNN nor BiLSTM alone is sufficient to model the complexity and temporal dependencies present in the social media text. The hybrid model's superior performance indicates that stacking convolutional layers with bidirectional recurrent units (BiLSTM + BiGRU) provides a more powerful feature extraction and sequence modeling capability. Each component in the hybrid model contributes to overall effectiveness: CNN layers are responsible for capturing local n-gram features and spatial patterns in text. BiLSTM captures long-range dependencies and context in both forward and backward directions. BiGRU acts as a lightweight and computationally efficient alternative to LSTM, further reinforcing temporal modeling. The combination leverages the complementary strengths of each component, resulting in a model that is both deep (complex) and generalizable. The model maintains consistently high performance across two very different social media platforms: Facebook, known for longer, well-structured posts. Twitter, which features short, noisy, and informal language.

This robustness indicates that the model is not overfitted to a specific text structure or style and is capable of generalizing well across different types of user-generated content. While some traditional models (e.g., Logistic Regression or Decision Trees) perform reasonably well and may be less resource-intensive, they lack the contextual understanding needed to detect nuanced misinformation or semantic subtleties. In contrast, the proposed model's high F1-score demonstrates that it can reliably distinguish between true and false information, making it suitable for deployment in practical misinformation detection systems. The evaluation of the F1-score specifies that the proposed algorithm has the highest F1-score of 99, followed by decision tree (94), Adaboost (92), Logistic regression (91), and so on. Similar trends can be seen when comparing recall, accuracy, and precision results. The suggested approach outperforms previous algorithms in terms of performance. The suggested hybrid model has significantly better outcomes on all the evaluation metrics than all the preceding models, which proves that multimodal deep learning architectures stimulate the solution to complex natural language processes such as detecting misinformation. The consistency of the higher precision and recall of the model points to the possibility of its real-time application in online social networking or algorithms that scan facts.

Table 5: Results on the Twitter dataset using DL approaches

Approach	Processing Execution Time	Accuracy	Precision	Recall	F1-score
CNN+BiLSTM	18 Sec	97%	97%	96%	98%
CNN+BiLSTM +BiGRU	20 Sec	99%	98%	98%	99%

Table 6: Results on the Facebook dataset using DL approaches

Approach	Processing Execution Time	Accuracy	Precision	Recall	F1-score
CNN + BiLSTM	27 Sec	97%	96%	96%	96%
CNN + BiLSTM + BiGRU	30 Sec	98%	98%	97%	97%

Table 7: Comparison with previous research

Approach	Accuracy	Precision	Recall	F1-score
Logistic regression (LR)	91 %	92 %	90 %	91 %
Voting classifier (RF, LR, KNN)	88 %	88 %	89 %	88 %
Decision trees	94 %	94 %	95%	94 %
AdaBoost	92 %	92 %	93%	92 %
Perez-LSVM	79 %	79 %	81 %	80 %
Wang-CNN	66 %	65 %	71 %	67 %
Wang-BiLSTM	52 %	43 %	59 %	44 %
Proposed CNN + BiLSTM + BiGRU (Facebook dataset)	99 %	98 %	98 %	99 %
Proposed CNN + BiLSTM + BiGRU (Twitter dataset)	98%	98%	97 %	97 %
Proposed CNN + BiLSTM + BiGRU (FakeNewsNet Multimodal dataset)	94%	93%	94%	96%

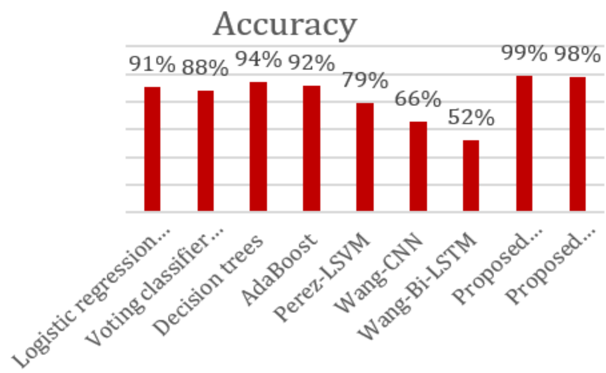

Fig. 9: Comparison of Accuracy with previous work

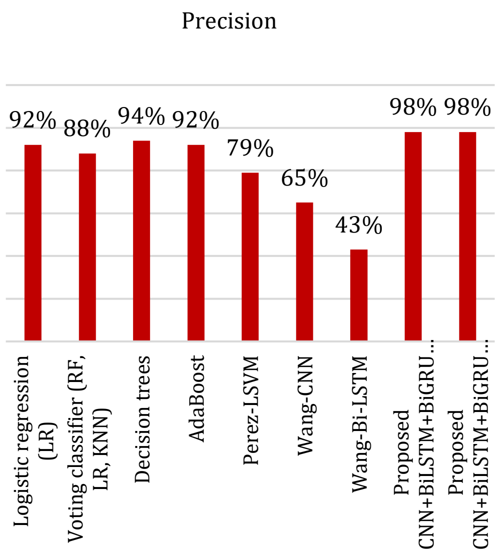

Fig. 10: Comparison of Precision with previous work

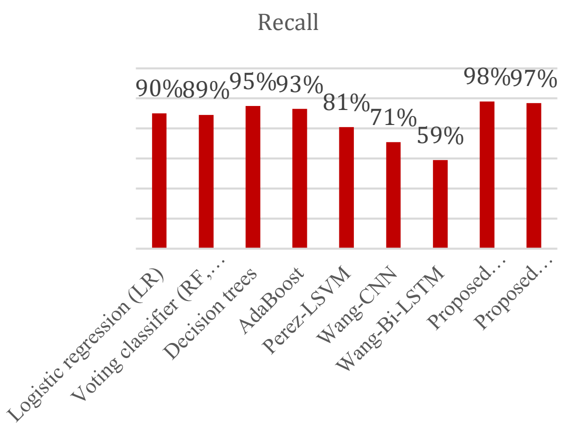

Fig. 11: Comparison of Recall with previous work

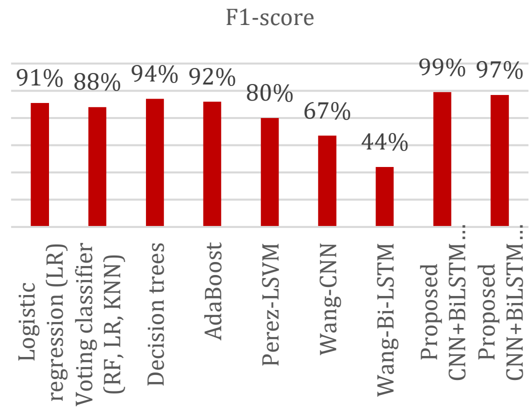

Fig. 12: Comparison of F1-score

Conclusion

The proposed research uses an ensemble framework to quickly identify rumors on social media platforms. The data pre-processing techniques create a time-series dataset, which reduces feature complexity and lowers the computation time during the training period. It converts Twitter and Facebook conversations among time-series vectors via timestamp creations, which are extracted and processed without any wait time. In this study, various advanced machine and deep learning methods are applied to mitigate the problem of rumor detection. The CNN, BiLSTM, and BiGRU algorithms applied in this research were integrated into a novel hybrid deep learning framework for textual data. This framework takes a unique approach that makes use of customized embedding. This hybrid model's results showed an accuracy of 99%, higher than the maximum accuracy obtained in earlier studies. This suggested framework can function as a decision engine in a recommender system, offering businesses helpful assistance in identifying rumors that might influence our society. The current study has certain disadvantages along with numerous advantages

Our study exclusively employed text-based characteristics for rumor categorization. However, more powerful findings may be obtained by enclosing more types of characteristics. The experiment only utilized English text. Other elements, such as pictures and contextual information, can be taken along with text-based features to obtain more accurate results.

Future research will be conducted on the text data, considering the language viewpoint. Other deep-learning approaches for rumor identification will be explored.

Acknowledgment

Thank you to the publisher for their support in the publication of this research article. We are grateful for the resources and platform provided by the publisher, which have enabled us to share our findings with a wider audience. We appreciate the efforts of the editorial team in reviewing and editing our work, and we are thankful for the opportunity to contribute to the field of research through this publication.

Funding Information

The authors have not received any financial support or funding to report.

Author's Contributions

All authors contributed equally to the conceptualization, investigation, writing, original draft preparation, review, and editing of the manuscript.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

Data will be available from the corresponding author upon reasonable request.

References

Alkhodair, S. A., Ding, S. H. H., Fung, B. C. M., & Liu, J. (2020). Detecting breaking news rumors of emerging topics in social media. *Information Processing & Management*, *57*(2), 102018. https://doi.org/10.1016/j.ipm.2019.02.016

Azeez, I. A. A., & Jimoh, A. A. (2023). The Impact of Islamophobia on Society, Attitudes, Policies and The Youths. *Eduvest - Journal of Universal Studies*, *3*(7), 1322–1339. https://doi.org/10.59188/eduvest.v3i7.873

Asghar, M. Z., Habib, A., Habib, A., Khan, A., Ali, R., & Khattak, A. (2021). Exploring deep neural networks for rumor detection. *Journal of Ambient Intelligence and Humanized Computing*, *12*(4), 4315–4333. https://doi.org/10.1007/s12652-019-01527-4

AlJamal, M., Alquran, R., Alsarhan, A., Aljaidi, M., Al-Jamal, W. Q., & Alkoradees, A. F. (2025). Optimized Novel Text Embedding Approach for Fake News Detection on Twitter X: Integrating Social Context, Temporal Dynamics, and Enhanced Interpretability. *International Journal of Computational Intelligence Systems*, *18*(1) https://doi.org/10.1007/s44196-024-00730-2

Alsaleh, H. (2022). *An evidence-based digital nudging in support of health misinformation assessment on social media sites*.

Ajao, O., Bhowmik, D., & Zargari, S. (2018). Fake News Identification on Twitter with Hybrid CNN and RNN Models. *Proceedings of the 9th International Conference on Social Media and Society*. SMSociety' 18: International Conference on Social Media and Society, Copenhagen Denmark. https://doi.org/10.1145/3217804.3217917

Bamiatzi, V., Bozos, K., & Lambertides, N. (2016). Mapping the trading behavior of the middle class in emerging markets: Evidence from the Istanbul Stock Exchange. *International Business Review*, *25*(3), 679–690. https://doi.org/10.1016/j.ibusrev.2015.02.004

Bisaillon, C. (2018). *Fake and real news dataset [Data set*. https://www.kaggle.com/datasets/clmentbisaillon/fake-and-real-news-dataset

Can, U., & Alatas, B. (2019). A new direction in social network analysis: Online social network analysis problems and applications. *Physica A: Statistical Mechanics and Its Applications*, *535*, 122372. https://doi.org/10.1016/j.physa.2019.122372

Cen, J., & Li, Y. (2022). A Rumor Detection Method from Social Network Based on Deep Learning in Big Data Environment. *Computational Intelligence and Neuroscience*, *2022*, 1–8. https://doi.org/10.1155/2022/1354233

Chauhan, N. K., & Singh, K. (2018). A Review on Conventional Machine Learning vs Deep Learning. *2018 International Conference on Computing, Power and Communication Technologies (GUCON)*. 2018 International Conference on Computing, Power and Communication Technologies (GUCON), Greater Noida, Uttar Pradesh, India. https://doi.org/10.1109/gucon.2018.8675097

Chen, K., Luo, Y., Hu, A., Zhao, J., & Zhang, L. (2021). Characteristics of Misinformation Spreading on Social Media During the COVID-19 Outbreak in China: A Descriptive Analysis. *Risk Management and Healthcare Policy*, *Volume 14*, 1869–1879. https://doi.org/10.2147/rmhp.s312327

Chen, Y.-S. (2021). *Virtual Communities and Social Networking in Franchising*. 251–264. https://doi.org/10.4018/978-1-7998-9020-1.ch012

Elaoud, A., Barhoumi, W., Zagrouba, E., & Agrebi, B. (2020). Skeleton-based comparison of throwing motion for handball players. *Journal of Ambient Intelligence and Humanized Computing*, *11*(1), 419–431. https://doi.org/10.1007/s12652-019-01301-6

Fazil, M., Sah, A. K., & Abulaish, M. (2021). DeepSBD: A Deep Neural Network Model With Attention Mechanism for SocialBot Detection. *IEEE Transactions on Information Forensics and Security*, *16*, 4211–4223. https://doi.org/10.1109/tifs.2021.3102498

Gandhi, P. R. (2022). *Marketing and consumer perception of Maggi noodles*. http://dspace.dtu.ac.in:8080/jspui/handle/repository/19492

Gao, J., Han, S., Song, X., & Ciravegna, F. (2020). RP-DNN: A tweet level propagation context based deep neural networks for early rumor detection in social media. *ArXiv:2002.12683*.

Guadagno, R. E., & Guttieri, K. (2021). *Fake News and Information Warfare*. 218–242. https://doi.org/10.4018/978-1-7998-7291-7.ch013

Hamidian, S., & Diab, M. (2016). Rumor Identification and Belief Investigation on Twitter. *Proceedings of the 7th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*. Proceedings of the 7th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, San Diego, California. https://doi.org/10.18653/v1/w16-0403

Kumar, M., Khan, L., & Chang, H.-T. (2025). Evolving techniques in sentiment analysis: a comprehensive review. *PeerJ Computer Science*, *11*, e2592. https://doi.org/10.7717/peerj-cs.2592

Li, G., Dong, M., Ming, L., Luo, C., Yu, H., Hu, X., & Zheng, B. (2022). Deep reinforcement learning based ensemble model for rumor tracking. *Information Systems*, *103*, 101772. https://doi.org/10.1016/j.is.2021.101772

Liu, J., Sun, Y., Zhang, Y., & Lu, C. (2024). Research on Online Review Information Classification Based on Multimodal Deep Learning. *Applied Sciences*, *14*(9), 3801. https://doi.org/10.3390/app14093801

Ma, J., Gao, W., & Wong, K.-F. (2018). Rumor Detection on Twitter with Tree-structured Recursive Neural Networks. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Melbourne, Australia. https://doi.org/10.18653/v1/p18-1184

Nguyen, T. N., Li, C., & Niederée, C. (2017). On Early-Stage Debunking Rumors on Twitter: Leveraging the Wisdom of Weak Learners. *In International Conference on Social Informatics*, *10540*, 141–158. https://doi.org/10.1007/978-3-319-67256-4_13

Nasery, M., Turel, O., & Yuan, Y. (2023). Combating Fake News on Social Media: A Framework, Review, and Future Opportunities. *Communications of the Association for Information Systems*, *53*(1), 833–876. https://doi.org/10.17705/1cais.05335

Rathakrishnan, A., & Sathiyanarayanan, R. (2023). Rumor Detection on Social Media Using Deep Learning Algorithms with Fuzzy Inference System for Healthcare Analytics System Using COVID-19 Dataset. *International Journal of Computational Intelligence and Applications*, *22*(01). https://doi.org/10.1142/s1469026823410080

Sandhu, N., & Saluja, S. (2023). Fraud Triangle as an Audit Tool. *Management and Labour Studies*, *48*(3), 418–443. https://doi.org/10.1177/0258042x231160970

Sarker, A. (2017). A customizable pipeline for social media text normalization. *Social Network Analysis and Mining*, *7*(1). https://doi.org/10.1007/s13278-017-0464-z

Sridhar, S., & Sanagavarapu, S. (2021). Fake News Detection and Analysis using Multitask Learning with BiLSTM CapsNet model. *2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*. 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India. https://doi.org/10.1109/confluence51648.2021.9377080

Strauss, N., & Smith, C. H. (2019). Buying on rumors: how financial news flows affect the share price of Tesla. *Corporate Communications: An International Journal*, *24*(4), 593–607. https://doi.org/10.1108/ccij-09-2018-0091

Suissa, O., Elmalech, A., & Zhitomirsky-Geffet, M. (2022). Text analysis using deep neural networks in digital humanities and information science. *Journal of the Association for Information Science and Technology*, *73*(2), 268–287. https://doi.org/10.1002/asi.24544

Sadiq, S., Mehmood, A., Ullah, S., Ahmad, M., Choi, G. S., & On, B.-W. (2021). Aggression detection through deep neural model on Twitter. *Future Generation Computer Systems*, *114*, 120–129. https://doi.org/10.1016/j.future.2020.07.050

Sankar, H., Subramaniyaswamy, V., Vijayakumar, V., Arun Kumar, S., Logesh, R., & Umamakeswari, A. (2020). Intelligent sentiment analysis approach using edge computing-based deep learning technique. *Software: Practice and Experience*, *50*(5), 645–657. https://doi.org/10.1002/spe.2687

Tan, L., Wang, G., Jia, F., & Lian, X. (2023). Research status of deep learning methods for rumor detection. *Multimedia Tools and Applications*, *82*(2), 2941–2982. https://doi.org/10.1007/s11042-022-12800-8

Vijeev, A., Mahapatra, A., Shyamkrishna, A., & Murthy, S. (2018). A Hybrid Approach to Rumour Detection in Microblogging Platforms. *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 337–342. https://doi.org/10.1109/icacci.2018.8554371

Wirama, D. G., Wiksuana, I. G. B., Zuraidah, M. S., & Kazemian, S. (2017). Price manipulation by dissemination of rumours: Evidence from the Indonesian stock market. *International Journal of Economics and Financial Issues*, *7*(1), 429–434. https://doi.org/10.2139/ssrn.2803432

Xu, X., Xu, Z., Ling, Z., Jind, Z., & Du, S. (2024). Comprehensive implementation of TextCNN for enhanced collaboration between natural language processing and system recommendation. *International Conference on Image, Signal Processing, and Pattern Recognition (ISPP 2024)*. International Conference on Image, Signal Processing, and Pattern Recognition (ISPP 2024), Guangzhou, China. https://doi.org/10.1117/12.3033511

Yahya, N. C. (2022). Impact of investors' sentiment on share prices: Potential investing strategies during the COVID-19 outbreak. *Journal of Entrepreneurship, Business and Economics*, *10*(2), 18–43. https://doi.org/10.59188/jebe.v10i2.172

Yang, Z., Chen, X., Wang, H., Wang, W., Miao, Z., & Jiang, T. (2022). A New Joint Approach with Temporal and Profile Information for Social Bot Detection. *Security and Communication Networks*, *2022*, 1–14. https://doi.org/10.1155/2022/9119388

Zhao, Z., Resnick, P., & Mei, Q. (2015). Enquiring Minds. *Proceedings of the 24th International Conference on World Wide Web*. WWW '15: 24th International World Wide Web Conference, Florence Italy. https://doi.org/10.1145/2736277.2741637

Zubiaga, A., Aker, A., Bontcheva, K., Liakata, M., & Procter, R. (2019). *Detection and Resolution of Rumours in Social Media*. ACM Computing Surveys. https://doi.org/10.1145/3161603

Zeng, H., Wang, R., Huang, Y., Cui, X., & Jiang, Q. (2021). Scientific Rumors Detection in Short Online Texts. *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 1233–1240. https://doi.org/10.1109/smc52423.2021.9659056