Bima Journal of Science and Technology, Vol. 9(2B) Jul, 2025 ISSN: 2536-6041

E-ISSN: 3115-4662 DOI: 10.64290/bima.v9i2B.1285

A Comparison of Algorithmic Techniques for the Identification of Fake News in Machine Learning for Twitter Misinformation Detection

Ugwunna C. O,[1]* Joshua J [2]., Alabi O. A[3]., Orji E. E[4]. and Emmanuel G. U[3].

[1]College of Science and Computing, Wigwe University Isiokpo, Rivers State Nigeria
[2]Department of Computer Science, Ahmadu Bello University, Zaria Nigeria
[3]Department of Computer Science, Federal University of Agriculture Abeokuta, Nigeria
[4]Department of Computer Science, Federal Polytechnic Ohodo Enugu State, Nigeria.

Corresponding Author: charles.ugwunna@wigweuniversity.edu.ng

ABSTRACT

Misinformation on Twitter is a common issue, and its ill effect is endangering individuals, communities, and society. Its detection through efficient means will minimize its damage. Hence, we conducted a comparative analysis using different machine learning models and neural networks for Twitter misinformation detection with the development environment established in Google Colab and Python as the underlying programming language. We used a set of 44,898 labeled tweets drawn from Kaggle and labeled as fake or true news, and with four significant features. We tested and trained four models: Support Vector Machine (SVM), Recurrent Neural Network (RNN), Naive Bayes, and Random Forest (RF). We tested model performance through several metrics. The accuracy was 0.99 for SVM (precision: 0.99, recall: 0.99, F1-score: 0.99) and Random Forest (accuracy: 0.99, precision: 0.93, recall: 0.95, F1-score: 0.94), which was the highest among all other models. Naive Bayes (accuracy: 0.94, precision: 0.93, recall: 0.95, F1-score: 0.94) and RNN (accuracy: 0.79, precision: 0.94, recall: 0.64, F1-score: 0.76) performed relatively low. This study has contributed to the creation of strong misinformation detection systems, maintaining the integrity of online information and preventing the dissemination of false information on Twitter. Our comparative analysis sheds light on the most effective machine-learning methods for detecting misinformation, opening the door to future research and applications.

Keywords: Misinformation, Fake news, Twitter, Machine-Learning and Detection

INTRODUCTION

Misinformation is the spread of inaccurate or incorrect information, as compared to disinformation that involves false deception (Woolley & Howard, 2016). Misinformation was originally defined as unreliable or not clearly sourced information with a spotlight on information quality. Current research, however, considers misinformation as an action of deception rather than inaccuracy because half-truths can be entailed (Fallis, 2015). According to Walton, Pointon, Barker, Turner, and Wilkinson (2022), the widespread circulation of misleading or inaccurate information—whether intentional or not—has emerged as a pressing global issue, especially magnified by the pandemic. This phenomenon impacts how people assess credibility and respond both mentally and physically. Their study emphasizes that such content goes by numerous terms, including disinformation, misinformation, fake news, and post-truth, all reflecting different facets of the same challenge.

Misinformation affects nearly every area of life, including social, political, and economic. It can destabilize nations, affect elections, and change public opinions (Wu, Morstatter, Carley, & Liu, 2019). Fact-checking news is consumed by 1,500 people in six times the speed, and fake news normally get shared by 70% when compared to authentic news. Fake news has been shown to spread further than the accurate information. As an example, during the 2016 American presidential election, a false story that Donald Trump was endorsed by Barack Obama garnered 480,000 engagements, while the actual event gathered slightly over 176,250 (West & Bergstrom, 2021). Social media users can fabricate disinformation to affect politics or accommodate personal ideology, while typical miscommunications keep spreading misinformation. Online platforms have become leading sources of information. Unlike mainstream media, social networks allow for unlimited dissemination of information, bypassing editorial gatekeepers who verify content (Scheufele & Krause, 2019). The younger generations rely more on social media than journalism for information, which results in the spread of misinformation (Starbird, Dailey, Mohamed, Lee, & Spiro, 2018). This spread of misinformation results in intellectual polarization, trust deficits within communities, and social polarization.

Over the years, Machine learning and Deep learning models have been used to detect Twitter misinformation. Apoorva, Goyal, Kumar, Singh, and Sharma (2022) conducted a study aimed at early detection of depression using linguistic cues from Twitter posts. They trained Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) models on a combined dataset sourced from Sentiment140 and the Twitter API. The models analyzed users' tweets to identify depressive markers, with LSTM outperforming RNN by achieving an accuracy of 96.21%. Their research highlights the potential of deep learning for timely mental health monitoring via social media.

Pavithra and Joseph (2019) examined how social media propagates information quickly, quite frequently fuelling rumours. They compared the performances of utilizing CNN, RNN, and RvNN for rumour detection on Twitter. CNN and RNN only considered the text in tweets. RvNN performed better by noting tweet-response pairs. Zhang, Ibrahim, and Fadzil (2024) reviewed deep learning methods for rumor detection in social media. They discussed the growing issue of false information and the benefits of DL models. The article categorized key features and reviewed CNN, RNN, and GNN models. It makes suggestions on how detection methods can be improved and gives indications on future work. Yadav and Gupta (2024) introduced an emotion-aware framework to detect fake news from multimodal data—images and text—on social media. Their method uses a vision transformer to filter out irrelevant visual content and improve classification. Experiments across five datasets demonstrated superior performance in accuracy, precision, recall, and F1 score. The study's ablation analysis confirmed each component's contribution, outperforming existing methods.

Someswara Rao, Raminaidu, Raju, and Sujatha (2024) addressed the challenge of fake news on social media by developing a lightweight RNN-based model using NLP techniques. Their approach aims to classify misleading content efficiently amid the growing flood of online data. The model is designed to improve detection accuracy while remaining resource efficient. Results show it effectively identifies false information, aiding in safer digital communication. Alghamdi, Lin, and Luo **(2022)** reviewed diverse ML and DL techniques for fake news detection. They evaluated classical, advanced, and transformer-based models across four real-world datasets. Their experiments showed contextual embeddings like BERTbase outperformed traditional methods. The study offers practical insights for enhancing detection accuracy using news content alone.

A recent study by Prabha, Malik, Kumari, Arya, Parihar, and Singh (2024) investigated the rise of fake news in the digital era, particularly through social media. The research evaluated the performance of four machine learning models—Naïve Bayes, Decision Tree, Random Forest, and Logistic Regression—applied to news verification tasks. Among these, the Decision Tree classifier demonstrated the highest accuracy at 99.56%. The study also emphasized ongoing security threats and challenges in authenticating online content. This study compares and ranks Neural Network and Machine Learning models for Twitter misinformation detection using the appropriate classification metrics. It decides the best option to prevent false information. and provides a performance comparison of all the models used.

Research GAP

a. The task of classifying fake news continues to face challenges due to the limited size and diversity of available datasets. This constraint often causes models to learn patterns that are too closely tied to the training data, limiting their ability to generalize to new or unseen content, as observed in the study by Someswara Rao et al. (2024).

b. Many deep learning approaches demand substantial computational resources, which translates to longer training durations and greater energy consumption, as highlighted by Alghamdi et al. (2022).

c. The intricate design of many model architectures presents difficulties in interpretation, making them less transparent and harder to explain. This raises concerns about their accountability and trustworthiness, as noted by Yadav & Gupta (2024)

Contributions

a. In this study, we conduct a comparative study of four widely used models—SVM, RF, Naïve Bayes and RNN for enhancing the detection of misinformation on Twitter.

b. Every model undergoes a critical and thorough evaluation process wherein it is tested against major performance metrics that include, but are not limited to, precision, recall, accuracy, and F1-score. The primary reason for this tough assessment is to ascertain and conclude the best approach out of the various models being tested.

c. Results indicate that SVM and RF outperform the rest of the models in accuracy as well as F1-score, confirming their potential to build stronger systems for misinformation detection.

MATERIALS AND METHODS

This paper provides a comparative analysis of certain machine learning and deep learning models for detecting misinformation on Twitter. These models include RF, SVM, and Naïve Bayes for machine learning and RNN for deep learning. The methodology proposed by the system has four key steps: data collection, preprocessing, train and test the model, and evaluate the performance. Figure 1 presents the framework of the proposed methodology.

Bima Journal of Science and Technology, Vol. 9(2B) Jul, 2025 ISSN: 2536-6041

E-ISSN: 3115-4662 DOI: 10.64290/bima.v9i2B.1285

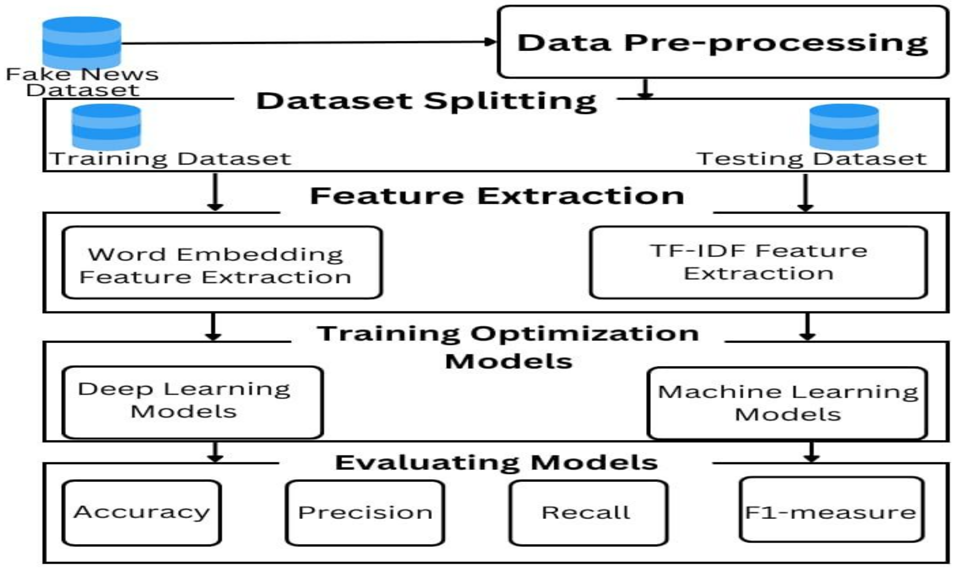


Figure 1: Architecture of the models

The model is a sequential dual-path architecture, as illustrated in Figure 1. Twitter data are uploaded to Google Colab and divided into training (70%) and test (30%) sets. Training data are fed to two parallel paths: RNN with word embedding feature and machine learning models with TF-IDF feature. The two steps intersect in the evaluation step, wherein the performance of the models is compared with the test dataset for choosing the optimum approach to Twitter misinformation detection.

Dataset Collection

44,898 tweets were obtained from Kaggle and marked as 'true news' (21,417 tweets) and 'fake news' (23,481 tweets) as depicted in table 1. Four principal features are present in the dataset: 'title,' 'text,' 'subject,' and 'date,' which are stored in CSV files. The data have been split into training (70%) for model learning and testing (30%) for evaluation.

Table 1: Summary of the dataset features and total size

Feature	Description
Total Tweets	44,898 tweets
True News	21,417 tweets
Fake News	23,481 tweets
Columns	Title, Text, Subject, Date, Label
Train/Test Split	70% Train (31,428 tweets), 30% Test (13,470 tweets)

Data Preprocessing

After processing data, the result of processing is saved in a new CSV file. There existed a single CSV file containing all the tweets—fake and actual—of both folders. "text" and "title" were combined after dropping the columns "date" and "subject" and saving them in the new CSV file. A new column indicating the category of tweet marked (Real & Fake) was added. Table 2 shows Dataset Structure and Table 3 show the Processed Dataset in this project.

Bima Journal of Science and Technology, Vol. 9(2B) Jul, 2025 ISSN: 2536-6041
E-ISSN: 3115-4662 DOI: 10.64290/bima.v9i2B.1285

Table 2: Dataset Structure

S/N	Title	Text	Subject	Date	Label
0	As U.S budget fight looms, Republic flip t…	WASHINGTON (Reuter)- The head of a conservation.…	politicsNews	December 31, 2017	Real
1	U.S. military to accept transgender recruits o…	WASHINGTON (Reuter) - Transgender people will of a conservation.…	politicsNews	December 29, 2017	Real

Data Preprocessing Steps

This stage is essential because it helps in removing and reducing the noises and unwanted parts from data before extracting any feature to improve the classification performance. Many processes, such as tokenization, normalization, removal of stop words, and light stemming wer performed during the data preprocessing stage. Figure 2 depicts the prepossessing steps.

Figure shows the preprocessing pipeline used to clean and structure raw text in advance of model training. As stated by Awajan et al., (2022), tokenization, noise filtering, stop word elimination, text normalization, and stemming are procedures involved. These reduce the dimensionality and improve feature extraction. The result is to enhanced model accuracy in detecting misinformation on Twitter.

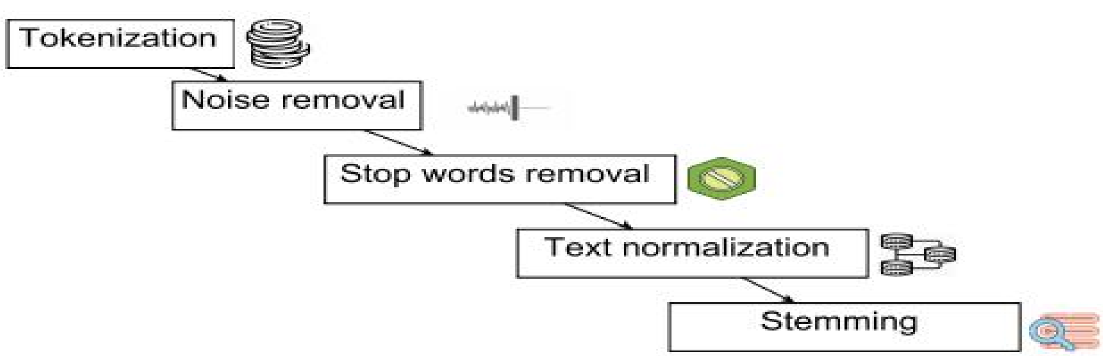


Figure 2: Preprocessing steps (Awajan, Alazab, Khurma, Alsaadeh, & Wedyan, 2022)

Table 3: Processed Dataset.

S/N	Text	Label
0	Breaking gop chairman Grassley enough demand the group…..	0
1	Failed gop candidate remembered hilarious mockery gallo…..	0
2	Mike penny new dc neighbor hilariously trolling sail…..	0
3	California ag pledge defends birth control insurance deman…..	1
4	Az rancher living u mexico border destroy nanc…	0

X = data[‘text’], y = data[‘label’]

E-ISSN: 3115-4662 DOI: 10.64290/bima.v9i2B.1285

Table 3 illustrates the processed dataset, which consists of two essential columns: "Text" (normalized text of tweets) and "Label" (denoting true or false). These essential elements enable the model's training and evaluation. This setup enables the efficient application of machine learning in recognizing misinformation.

Training and Testing Phase

The model is trained on 70% of labeled tweets indicating whether they spread misinformation. It is then tested on the remaining 30% to evaluate performance and generalization. Each algorithm predicts the likelihood of misinformation and outputs a score or label for comparison.

Model Workflow

The flowchart shown below represents the overall process involved in the proposed system (Figure 3).

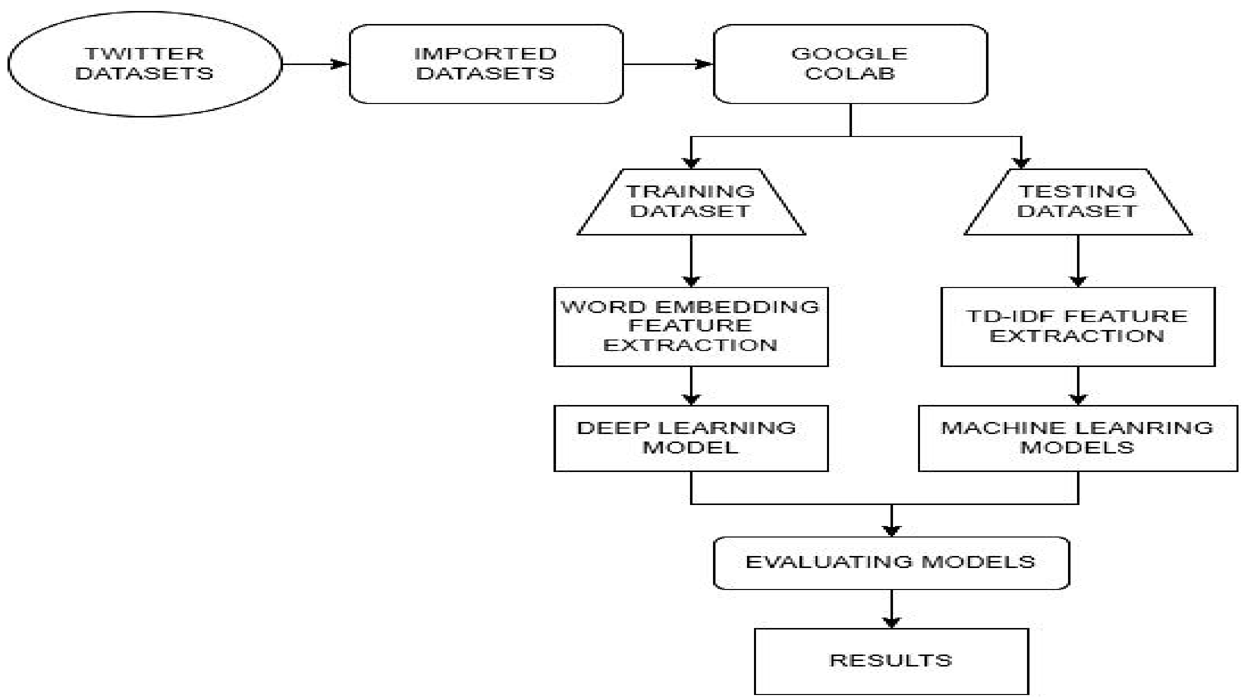


Figure 3: Workflow of the models.

Data from Twitter is gathered, uploaded to Google Colab, and divided into training and testing datasets. RNN model use word embeddings, while machine learning models use TF-IDF. Performance of models is evaluated based on accuracy, precision, and recall detecting misinformation.

Evaluation Metrics

Model performance was measured using classification performance evaluation metrics , which are derived from values of confusion matrix. Although accuracy provides a general notion, it might be deceptive while dealing with imbalanced datasets; therefore, additional metrics were also considered. Recall measures the performance of the model in detecting actual fake tweets, minimizing undetected disinformation. Accuracy ensures that genuine tweets are not flagged unnecessarily, preventing erroneous censorship. F1 score balances precision and

E-ISSN: 3115-4662 DOI: 10.64290/bima.v9i2B.1285

recall, and the confusion matrix determines misinformation detection enhancement.

RESULTS AND DISCUSSION

To determine the correct and incorrect number of false news, we use these measures to estimate the accuracy of a classification model, we will determine for True Positive, True Negative, False Positive, and False Negative.

RNN

The RNN model was trained using the dataset to determine news stories as real or fabricated.10 epochs were used to train the model, each epoch being one complete pass over the training set.

10 Epoch Reading Parameter of RNN

(Table 4)

Table 4: 10 Epochs Parameter of RNN

	Loss	Accuracy	Val_loss	Val_accuracy	Lr
0	0.698969	0.527181	0.664993	0.524837	0.01
1	0.649858	0.596819	0.616749	0.599391	0.01
2	0.569782	0.678273	0.538000	0.684239	0.001
3	0.495827	0.738304	0.467375	0.754855	0.001
4	0.447105	0.775421	0.419284	0.787345	0.001
5	0.404775	0.807465	0.456819	0.789135	0.001
6	0.413694	0.803228	0.408119	0.796384	0.001
7	0.383828	0.825069	0.900063	0.611564	0.001
8	0.438799	0.790548	0.404548	0.828336	0.001
9	0.455355	0.783566	0.416983	0.790656	0.001

Table 4 presents the training progress of the model across 10 epochs, demonstrating a consistent improvement in performance. As training progressed, the loss steadily decreased while accuracy increased, indicating effective learning. Similar trends were observed in the validation metrics, suggesting that the model generalised well to unseen data. After completing 10 epochs, the RNN model achieved a final training accuracy of approximately 0.7836 and a validation accuracy of about 0.7907, providing a solid basis for evaluating its effectiveness.

Plotted RNNs Graph of Accuracy & Val_Accuracy against Ten Epochs

The graph plotted for RNN on Accuracy & Val_accuracy against 10 epochs shows the training accuracy and validation accuracy of the RNN model over 10 epochs. Accuracy reflects how well the model performs on the training data, while validation accuracy (val_accuracy) indicates how effectively the model generalizes to new, unseen data.

As indicated by Figure 4, the model achieved 78.36% training accuracy and 79.07% validation accuracy at epoch 10. The conclusion is drawn as follows:

i. The slightly higher val_accuracy is an indication of good generalization without overfitting.
ii. Accuracy and val_accuracy plots are converging, which shows consistent improvement.
iii. Both plots are still on the rise, which is a sign that further training may further enhance performance.
iv. The model is learning nicely and is fitting well to unseen data.

E-ISSN: 3115-4662 DOI: 10.64290/bima.v9i2B.1285

v. Hyperparameter fine-tuning of learning rate, batch size, or network depth may improve performance.

vi. Figure 4 shows a well-balanced training process with room for further optimization.

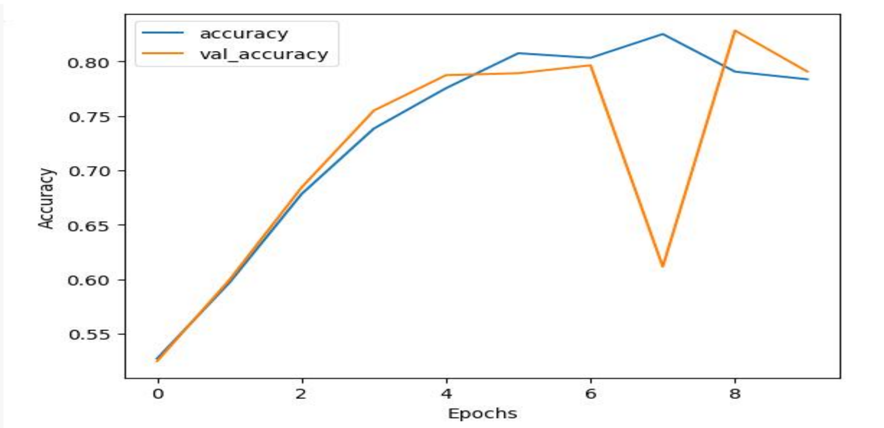


Figure 4: Graph of Accuracy against 10 Epochs.

Plotted RNNs Graph of Loss & Val_Loss against Ten Epochs

The plot plotted for RNN on loss & Val_loss vs 10 epochs shows the training loss and validation loss of the RNN model over 10 epochs (Figure 5). The discrepancy or difference that arises between the predictions done by the model and the true actual labels is quantified using a measure named Loss. In the same way, the difference that can be observed between the predictions done by the model and the true labels of the validation set is also quantified by employing a measure named Val_loss.

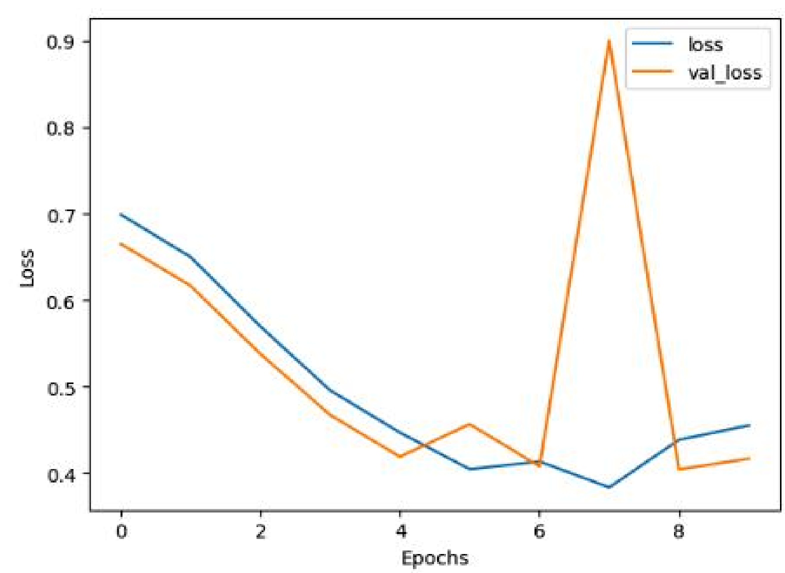


Figure 5: Graph of Loss against 10 epochs.

As can be seen in Figure 5, at epoch 10, the training loss is 0.455355, while the validation loss is slightly lower at 0.416986, which is a sign of good generalization. That val_loss is lower is a sign that the model is not overfitting and performing well on unseen data. The loss and val_loss curves are both converging, and the fact that they continue to decrease means that longer training might drive error even lower. Model learning is still good, and performance can possibly improve with more epochs. Hyperparameter tuning—like modifying learning rate or batch size—can optimize results further.

In general, the trend in Figure 5 demonstrates a model learning effectively without any indications of overfitting.

Classification Report of RNN

Table 5: Classification Report for RNN.

	Precision	Recall	F1-score	Support
0	0.94	0.64	0.76	5849
1	0.71	0.95	0.81	5324
accuracy			0.79	11173
Macro avg	0.82	0.85	0.79	11173
Weighted avg	0.83	0.79	0.79	11173

Table 5 is the classification report of the RNN model, a thorough evaluation of its performance in detecting misinformation on Twitter. According to the report, the RNN model achieves precision of 0.94 for class 0 (true information) and 0.71 for class 1 (misinformation), which implies that it is better at detecting true information. The recall scores show that the model is better at recognizing misinformation (0.95) than at recognizing actual information (0.64). The F1 scores, which are the trade-off between precision and recall, for class 0 and class 1 are 0.76 and 0.80, respectively, i.e., with slightly improved performance in recognizing misinformation. Overall, the RNN model is at 0.79 accuracy level, indicating how good it is at detecting Twitter misinformation, albeit not yet at getting perfect detection of genuine information.

Confusion Matrix

This approach uses a structured tabular format to evaluate the effectiveness of a classification model—specifically, a recurrent neural network (RNN). It captures both accurate and incorrect predictions by aligning the expected labels with those generated by the algorithm. This framework helps assess when adjustments or refinements to the model may be necessary.

As indicated in Figure 6, this study's observations are as follows:

- The diagonal elements (3766 and 5068) represent the number of correct classifications.

- The off-diagonal elements (2083 and 256) represent the number of incorrect classifications.

3766 is True Positive, and 5068 is True negative, thus, TP is the correct fake news.

Overall, the confusion matrix suggests that the RNN model is performing well, with a high number of correct classifications (3766 + 5068 = 8834) and a relatively low number of incorrect classifications (2083 + 256 = 2339). However, there is some room for improvement, particularly in reducing the number of incorrect classifications for both classes.

E-ISSN: 3115-4662 DOI: 10.64290/bima.v9i2B.1285

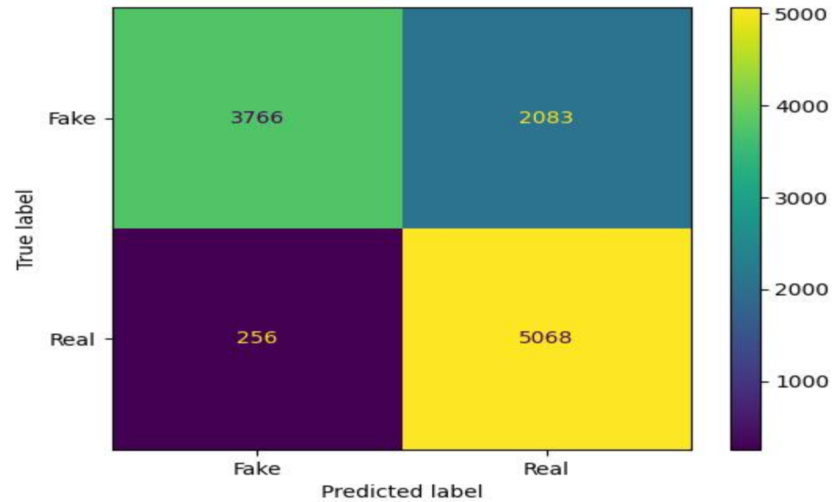


Figure 6: Confusion matrix of RNN.

Naive Bayes (NB)

This section explains the classification report and confusion matrix of Naive Bayes. The breakdown of the metrics is as follows:

Classification Report of NB

In Table 6 presented above, it is stated that 0.94 accuracy is noted, which implies that the model is performing well overall. Precision, recall, and F1-score measures are also noted to be high for both classes, indicating that the model is efficient to predict both classes.

A few observations in Table 6, which are:

- Class 1 has slightly more precision of 0.94 and recall compared to Class 0, and therefore the model is slightly more accurate at identifying Class 0 instances.

- The support values inform us that there are fewer instances of Class 1 compared to Class 0.

- The high values of F1-score inform us that the model is achieving an acceptable trade-off between precision and recall.

Overall, the Naive Bayes model is performing well on this dataset, with high precision, recall, and F1-score values for both classes.

Table 6: Classification Report for Naive Bayes

	Precision	Recall	F1-score	Support
0	0.93	0.95	0.94	5849
1	0.94	0.93	0.93	5324
accuracy			0.94	11173
Macro avg	0.94	0.94	0.94	11173
Weighted avg	0.94	0.94	0.94	11173

Confusion Matrix

Confusion matrix is a valuable tool that is utilized to calculate and identify how well a classification model is functioning—in this instance, the Naive Bayes classifier. The confusion matrix provides a comprehensive comparison between true labels, which are the actual categories of the data, and the predictions of the model, thereby illuminating both the correct classifications and incorrect ones. This detailed examination of information is essential in precisely calculating and understanding how well the model being analyzed is performing.

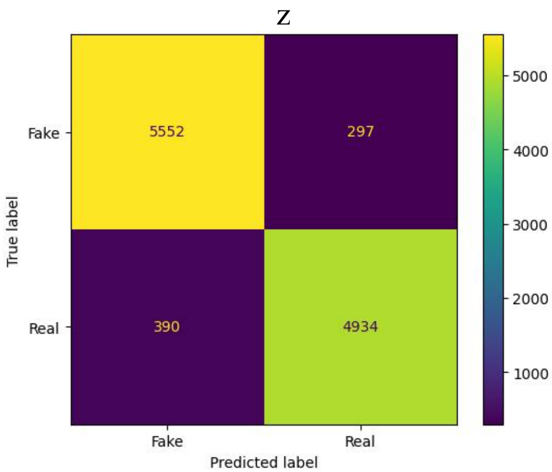


Figure 7: Confusion matrix of NB

Figure 7 shows the confusion matrix for NB model with 5552 true positive and 4939 true negative. Wrong identifications are 297 false positives and 390 false negatives. The model correctly predicted 10,491 and mis predicted 687. While performance is strong, removal of misclassifications would improve accuracy even more.

RANDOM FOREST (RF)

This section shows the classification report and confusion matrix of Random Forest. Here's a breakdown of the metrics:

Classification Report of RF

As shown in Table 7 above, the accuracy of 0.94 indicates that the model is performing well overall. The precision, recall, and F1-score metrics are also high for both classes, indicating that the model is effective in predicting both classes. Overall, the Random Forest classifier works excellent for Class 1 and good for Class 0 with some room for improvement in recall for Class 0 instances.

Table 7: Classification Report for Random Forest.

	Precision	**Recall**	**F1-score**	**Support**
0	0.99	0.99	0.99	5849
1	0.99	0.99	0.99	5324
accuracy			0.99	11173
Macro avg	0.99	0.99	0.99	11173
Weighted avg	0.99	0.99	0.99	11173

Some of the likely causes of this excellent performance in Table 7 are:

- The data may be comparatively easy or structured, and hence easy for the model to learn and generalize.

- Random Forest algorithm can possibly be best applied with this type of data, which is able to identify intricate patterns and relationships.

- The model may have been trained on enough data, which will make it able to learn and generalize well.

Confusion Matrix

The confusion matrix is a chart used to compare the accuracy of a classification model, in this case, a Random Forest. It plots crosswise the actual and predicted labels, both the correct and incorrect classifications. The metrics can then be used to identify the performance of the model and areas it needs to be enhanced.

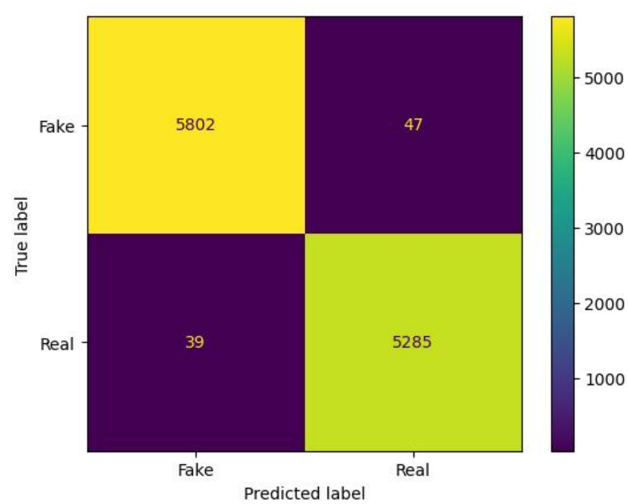


Figure 8: Confusion matrix of RF.

As clearly shown in Figure 8, the following observations were made from the study:

- The diagonal elements (5802 and 5285) represent the number of correct classifications.

- The off-diagonal elements (47 and 39) represent the number of incorrect classifications.

5802 is True Positive, and 5285 is True Negative, thus, TP is the correct fake news.

Overall, the confusion matrix suggests that the RF model is performing well, with a high number of correct classifications (5802 + 5285 = 11,087) and a relatively low number of incorrect classifications (47 + 39 = 86).

Support Vector Machine (SVM)

The Support Vector Machine (SVM) trained using Linear SVC () demonstrates exceptional performance on the binary classification problem

Classification Report of SVM

Table 8: Classification Report for SVM

	Precision	Recall	F1-score	Support
0	1.00	0.99	0.99	5849
1	0.99	0.99	0.99	5324
accuracy			0.99	11173
Macro avg	0.99	0.99	0.99	11173
Weighted avg	0.99	0.99	0.99	11173

Table 8 above shows that the Linear SVC () implementation of SVM is well-suited for this problem, and the model has effectively learned the decision boundary between the two classes.

Here are some observations given from the result in Table 8:

- The SVM model achieves near-perfect performance on both classes, indicating excellent generalization.

- The precision, recall, and F1-score values are extremely close to 1, indicating minimal errors.

- The support values indicate a slightly larger number of instances in Class 0 than in Class 1.

- The accuracy of 0.99 confirms the model's exceptional performance.

Confusion Matrix

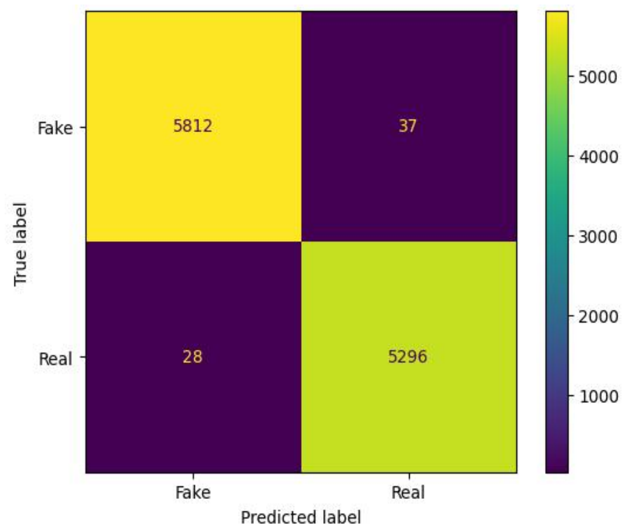


Figure 9: Confusion matrix of SVC.

The confusion matrix evaluates the SVM through the comparison of predicted and true labels, showing correct and incorrectly classified cases for improving performance. As evident from the confusion matrix in Figure. 9, the RF model performs extremely well with 5812 true positives (predicting fake news correctly) and 5296 true negatives. The off-diagonal cells—28 false positives and 37 false negatives—are relatively few misclassifications. In all, the model correctly classifies 11,087 times and incorrectly only 65 times, a circumstance that indicates high accuracy.

Model Comparisons Using a Graph

Figures 10 to 13 illustrate how RNN, Naive Bayes, RF, and SVM fare as far as accuracy, precision, recall, and F1-score are concerned. The plots emphasize how every algorithm performs in terms of important evaluation metrics.

Comparison by Accuracy

As depicted in Figure 10, which shows the comparison of the model's accuracy, which illustrates the following:

i. Random Forest and SVM consistently outperform RNN, while Naive Bayes has a surprisingly high accuracy.
ii. RNN may require additional tuning or data preprocessing to improve its performance.
iii. Naive Bayes has high accuracy despite being a simple yet effective model, but its performance may vary depending on the dataset.

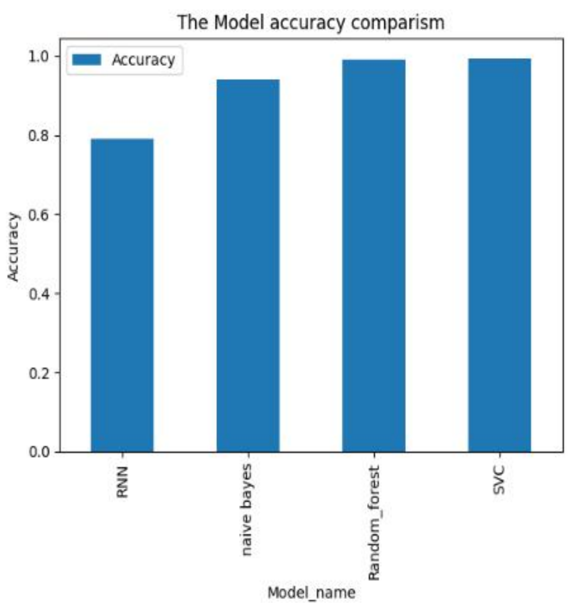
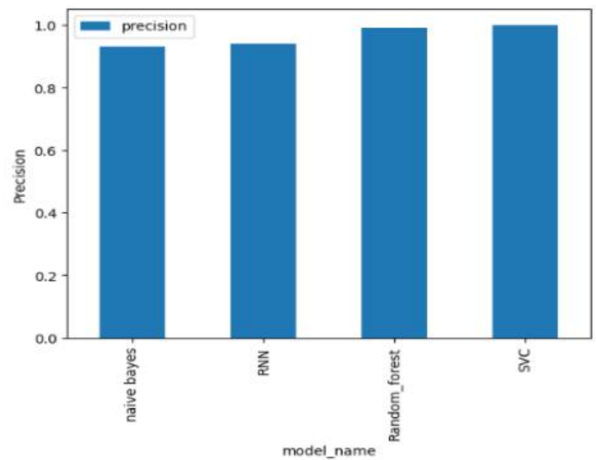


Figure 10: Comparison of Model Accuracy

Comparison by Precision

As depicted in Figure 11, this shows the comparison of the models' precision, which illustrates the following:

i. SVM has the highest precision (1.00), indicating, it has the lowest rate of false positives. It perfectly classified all instances as either positive or negative.
   - Naive Bayes and RNN have high precision values (0.96 and 0.94, respectively), indicating that they also have low rates of false positives.
   - Random Forest has a slightly lower precision (0.93), indicating a slightly higher rate of false positives compared to the other models.



Comparison by Recall

Recall: The count of true positives (accurately predicted instances) out of all actual positive instances in the data set. It estimates the model's recall, i.e., how accurately it can locate all relevant instances.

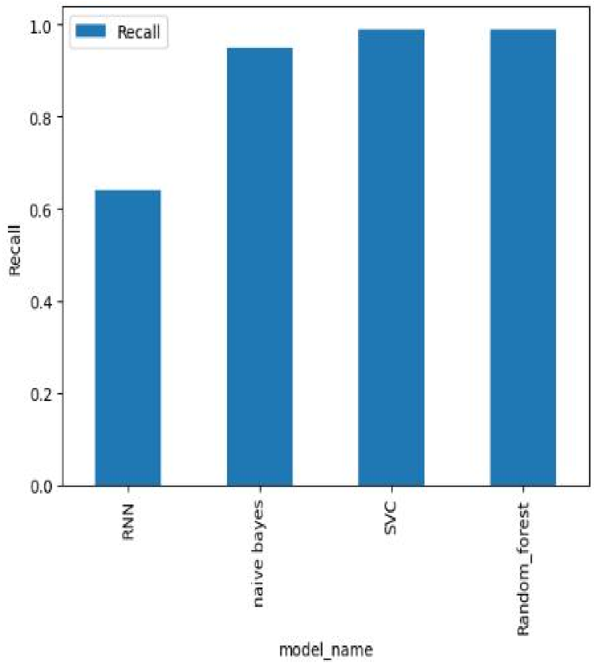


Figure 12: Comparison of Model Recall.

As indicated in Figure 12, this shows the comparison of the models' recall, which illustrates the following:

i. SVM has the highest recall (0.99), indicating that it detected almost all actual positive instances (only 1% were missed).
ii. Naive Bayes and Random Forest have high recall values (0.95), indicating that they also detected most actual positive instances (only 5% were missed).
iii. RNN has a significantly lower recall (0.64), indicating that it missed a substantial number of actual positive instances (36% were missed).

Comparison by F1 Score

As depicted in Figure 13, this shows the comparison of the models' F1 score, which illustrates the following:

i. SVM had the highest F1 score (0.99) with an excellent precision-recall balance.
ii. Naive Bayes and Random Forest were also satisfactory with F1 scores of 0.94.
i. RNN lagged with an F1 score of 0.76, with a worse balance and scope for improvement.

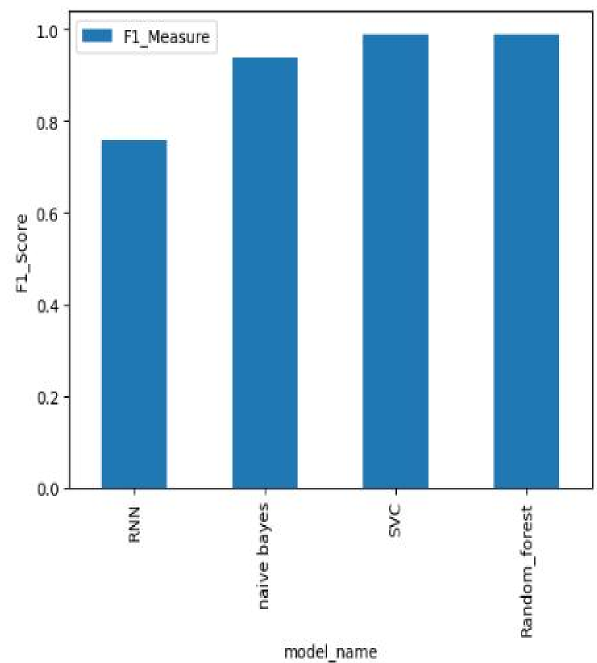


Figure 13: Comparison of Model F1 score.

Strengths and Weaknesses of the Models

Table 13: Strengths and weaknesses of the models

	Random Forest	SVM	RNN	Naive Bayes
Strengths	High Metric score.	High Metrics score.	Good performance on sequential data.	Simple, fast, easy to implement, and high accuracy.
Weaknesses	Computationally expensive.	Requires careful parameter tuning.	Struggles with Class 0 recall, lower precision for Class 1.	Limited by assumptions and simplicity.

Bima Journal of Science and Technology, Vol. 9(2B) Jul, 2025 ISSN: 2536-6041

E-ISSN: 3115-4662 DOI: 10.64290/bima.v9i2B.1285

As depicted in Table 13 above, shows the strengths and weaknesses of the various models used in this study.

DISCUSSION

The comparison of Tables 4 to 8 indicates considerable performance variation among the models tested. SVM, RF gave the most promising results overall. SVM presented very good performance with 0.994 accuracy and flawless precision (1.00). Its recall (0.99) and F1 score (0.99) point to high reliability in classification. Random Forest also gave a remarkable performance with 0.991 accuracy and an F1 score of 0.94. It properly traded off recall (0.95) and precision (0.93) and was a good model for misinformation detection.

Naive Bayes (NB), while simpler in structure, showed great potential with an accuracy of 0.934. It offered balanced precision and recall (both around 0.93–0.95), along with a decent F1 score. This renders NB particularly handy in quick, resource-limited environments where computational simplicity is a priority. Conversely, the Recurrent Neural Network (RNN) did not perform well as the other models. Despite having attained a high precision of 0.94, its recall was just 0.64, resulting in a lower F1 score of 0.76. This shows that the RNN failed to detect many misinformation tweets, thus narrowing its credibility in practical applications. The model's performance could be improved with more advanced neural network structures such as LSTM or GRU.

These findings reconfirm that traditional machine learning algorithms—specifically SVM and RF—are viable tools for identifying misinformation on Twitter, especially if they are trained on well-preprocessed and quality datasets. Neural networks, while promising, require additional polishing and optimization. Our study provides direction to follow-up research in exploring hybrid methods or refined neural network architectures to continue advancing misinformation detection techniques.

Limitations of the Models

i. Limited Recall for RNN: The RNN fails to detect a significant number of actual fake news instances, resulting in a low recall of 0.64.

ii Computational Complexity: RF and SVM models demand meticulous hyper-parameter tuning and consume significant computational resources, especially when handling large datasets.

iii Bias in Dataset: The dataset from Kaggle, whose primary interest is the true or fake labeled tweets, has biases that limit its generalization to other datasets.

iv. Class Imbalance: The slight imbalance in the distribution of true and fake tweets affects the performance of certain models.

v Model Interpretability: The complexity of models like RNN obscures their decision- making processes, reducing transparency and making interpretation difficult.

CONCLUSION

This research compares the performance of Naive Bayes, RF, SVM, and RNN for the detection of false news. Amongst them, SVM achieved maximum performance with precision 1.00, recall 0.99, and F1 measure of 0.99. Naive Bayes and RF too yielded good results with F1 values of nearly 0.94 each. RNN was behind with an F1 value of 0.76. These findings show that SVM is the most reliable model for detecting fake news. Naive Bayes and Random Forest are also reliable options, however. RNN needs a lot of improvement to be a strong rival. This study supplies valuable knowledge to build accurate fake news detection models and offers recommendations to future research. Future studies can be focused on hybrid or ensemble approaches, especially to improve deep learning models such as RNN.

REFERENCES

Alghamdi, J., Lin, Y., & Luo, S. (2022). A comparative study of machine learning and deep learning techniques for fake news detection. *Information*, *13*(12), 576.

Apoorva, A., Goyal, V., Kumar, A., Singh, R., & Sharma, S. (2022, December). Depression detection on Twitter using RNN and LSTM models. In International conference on advanced network technologies and intelligent computing (pp. 305-319). Cham: Springer Nature Switzerland.

Awajan, A., Alazab, M., Khurma, R. A., Alsaadeh, R., & Wedyan, M. (2022). Fake news detection and prevention using artificial intelligence techniques: A review of a decade of research. International Journal of Computer Information Systems and Industrial Management Applications, 14, 12–12.

Fallis, D. (2015). What is disinformation?. *Library trends*, *63*(3), 401-426.

Prabha, C., Malik, M., Kumari, S., Arya, N., Parihar, P., & Singh, J. (2024, March). Detection of fake news: A comparative analysis using machine learning. In *AIP Conference Proceedings* (Vol. 3072, No. 1). AIP Publishing.

Pavithra, C. P., & Joseph, S. (2019, August 15). *Deep learning approach for rumour detection in Twitter: A comparative analysis*. In *Proceedings of the International Conference on Systems, Energy & Environment (ICSEE) 2019*, GCE Kannur, Kerala, July 2019. SSRN.

Scheufele, D. A., & Krause, N. M. (2019). Science audiences, misinformation, and fake news. Proceedings of the National Academy of Sciences, 116(16), 7662–7669. https://doi.org/10.1073/pnas.1805871115

Someswara Rao, C., Raminaidu, C., Raju, K. B., & Sujatha, B. (2024). Effective Fake News Classification Based on Lightweight RNN with NLP. *Annals of Data Science*, *11*(6), 2141-2165.

Starbird, K., Dailey, D., Mohamed, O., Lee, G., & Spiro, E. S. (2018). Engage early, correct more: How journalists participate in false rumors online during crisis events. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (pp. 1–12). https://doi.org/10.1145/3173574.3173722

Walton, G., Pointon, M., Barker, J., Turner, M., & Wilkinson, A. J. (2022). Information discernment and the psychophysiological effects of misinformation. *Global Knowledge, Memory and Communication*, *71*(8/9), 873-898.

West, J. D., & Bergstrom, C. T. (2021). Misinformation in and about science. *Proceedings of the National Academy of Sciences, 118*(15), e1912444117. https://doi.org/10.1073/pnas.1912444117

Woolley, S. C., & Howard, P. N. (2016). Political communication, computational propaganda, and autonomous agents. *International Journal of Communication, 10*, 4882–4890.

Wu, L., Morstatter, F., Carley, K. M., & Liu, H. (2019). Misinformation in social media: definition, manipulation, and detection. *ACM SIGKDD explorations newsletter*, *21*(2), 80-90.

Yadav, A., & Gupta, A. (2024). An emotion-driven, transformer-based network for multimodal fake news detection. International Journal of Multimedia Information Retrieval, 13(1), 21–34. https://doi.org/10.1007/s13735-024-00281-3

Zhang, L., Ibrahim, S., & Fadzil, A. F. A. (2024). Rumor detection based on deep learning techniques: a systematic review. TELKOMNIKA (Telecommunication Computing Electronics and Control), 22(4), 869-886.