An Effective Ensemble Technique for Fake News Detection using Support Vector Machine, Random Forest, and Gradient Boosting Classifiers

M.F.F. Fusna Fareed and I.M. Kalith

Abstract The spread of fake news on social media, blogs, and websites has facilitated the dissemination of misinformation. Various studies have attempted to spot fake news using supervised and unsupervised learning approaches, but these methods often struggle with low accuracy due to poor feature selection methods, improper parameter tuning, and imbalanced datasets. This research aims to develop an efficient approach to automatically finding of fake news using machine learning techniques. The suggested methodology integrates three powerful machine learning models, such as SVM, Random Forest (RF), and Gradient Boosting (GB) coupled with an advanced text preprocessing technique, Count Vectorizer for Feature Extraction. These models are combined in a soft-voting ensemble, which enhances classification precision by leveraging the strengths of each classifier. The created model was trained and tested on a comprehensive selected dataset, achieving an accuracy of 91%, a recall of 94%, and an F1 score of 91%. This research highlights the effectiveness of using an ensemble approach for detecting fake news and its scalability for large datasets.

***Index Terms*— Fake News Detection, Gradient Boosting, Random Forest, Support Vector Machine**

I. Introduction

IN today's fast-evolving digital era, the issue of fake news has appeared as a noteworthy concern, posing threats to the accuracy and reliability of information. The term "fake news" refers to false or misleading information presented as legitimate news, which includes false narratives, misleading images, and deceptive headlines that can spread swiftly through social media platforms such as Facebook and Twitter, leading to reputational harm and societal damage. The ease of content creation and dissemination allows misinformation to proliferate, demanding urgent attention from researchers and policymakers alike. People can transmit bogus news to others, and everyone will start believing it if it is not identified early. Fake news can have an impact on people, groups, or political parties. US election of 2016, false information affected people's opinions and choices [1].

A. Research Problem

Fake news detection is difficult since it is not a simple task. Traditional detection methods, such as manual fact-checking, are time-consuming and labor-intensive, while rule-based systems struggle to keep up with the rapidly evolving tactics of misinformation. Additionally, statistical models often lack the interpretability and generalization required for effective fake news detection [2]. Even though researchers have introduced various machine learning-based classification models to address fake news, with notable progress in recent times. However, due to the complexity and evolving nature of fake news, single classification models often fall short in delivering consistently high accuracy. To tackle this, this research proposes using an ensemble approach for the automatic detection of fake news in news article platforms, which combines the strengths of multiple machine learning algorithms to improve detection capabilities.

B. Aim and Objectives of Research

This research aims to focus on detecting fake news in English-language textual articles, excluding other media formats. Using publicly available datasets for training and testing the machine learning models, an ensemble classification model will be developed for detecting fake news to improve accuracy compared to existing approaches.

C. Objectives

- To focus on solving the detection issues with machine learning algorithms.
- Assessing the performance of SVM, GB, and RF compared to existing approaches.
- To evaluate the performance of the proposed ensemble models using certain performance metrics and prove the efficiency of the proposed work by comparing it with other detection techniques' performance.

M.F.F. Fusna Fareed is a graduate from the Department of IT, Faculty of Arts and Culture, South Eastern University of Sri Lanka, Sri Lanka. (Email: fusnafareed@gmail.com)

I.M. Kalith is a Senior Lecturer at the Department of ICT, South Eastern University of Sri Lanka, Sri Lanka. (Email: imkalith@seu.ac.lk)

ISSN: 2961-5992

II. Literature Review

There is no commonly accepted definition of fake news [3]. Research communities often define fake news as ‘‘Fake news is a news article that is intentionally and verifiably false’’ [1]. According to the Cambridge Dictionary, fake news is inaccurate information that is sometimes sensational and created to be extensively shared or disseminated in order to make money or to advance or undermine a public figure, political movement, company, etc. Since many people attempt to mislead others by disseminating incorrect information on social media, it is imperative to identify false information early on [4]. It is essential to identify false news in order to prevent people or organizations from losing their reputation due to it. It's critical to conduct a comprehensive study on content-based false news identification utilising deep learning and machine learning [5].

The issue of how to manage fake news arises here because it is uncontrollable. The answer that has been suggested is machine learning. Machine learning can be used to detect fake news [6]. According to researchers, Machine Learning (ML) Classification Machine Learning (CML) enables software systems to provide more accurate outcomes without requiring direct reprogramming [2]. According to investigators' findings, spotting such bogus news is both a crucial and difficult endeavour. To accomplish this job, the researchers used the three machine learning methodologies. These consist of Naïve Bayes, neural networks, and SVM. The accuracy of the Naïve Bayes model was 96.08%. However, one study (Aphiwongsophon et al., 2018) found that the neural network and SVM algorithms had an accuracy of 90.90%.

It is possible to identify fake news automatically or manually. Either expert-based or crowd-sourced methods are used for manual fake news detection. Prior to the invention of automatic methods, the majority of processes were manual. When computing power and technology were lacking, this approach was widely used [3]. By acquiring fact checks, one can perform manual fact-checking, which enables a reader to evaluate work critically and take into account its integrity and relevance. Over 100 active fact-checking websites exist globally. However, these fact-checking websites usually need help linking to all the pages with the questionable claims since fake news spreads and changes so quickly, which lessens the potential impact of fact checks [7].

In machine learning, training the classifier is an essential task. This has a significant impact on how accurate these classifiers' outputs. A classifier needs to be properly trained using the right data set. Machine learning classifiers have been trained by several researchers to identify false news. The majority of the training data set is in an unbalanced form, which is the primary issue that arises when training these classifiers [4]. Mohan, et al., explain that supervised machine learning classifiers are employed to detect fake news. They have used the three distinct feature extraction models to train these classifiers. In reality, the classifiers are trained using these features. These models are the TF-IDF Model, N-Gram Model, and Bag of Words Model. These models use the training data set to extract features, which are subsequently used to train the classifier [5]. The evaluation of the ML models was applied in order to assess the models' performance and their generalizability to real world data. Accuracy, precision, recall, and F1-score were evaluation criteria used to gauge the ML models' performance. The evaluation was applied with three types of data. The first type included data that had not been subjected to feature selection; whereas the second and third types included data that was subject to Chi-squared feature selection and univariate feature selection, respectively [6].

A. Summary of Review

The reviewed literature highlights the growing significance of fake news detection and the vital role of machine learning in addressing this challenge. Traditional and modern approaches, including supervised and unsupervised methods, have demonstrated strengths but face limitations in handling the complexity of evolving fake news patterns. Feature engineering has improved detection performance; yet more advanced techniques are needed to process the nuanced language of fake news. However, due to the intricacy and developing nature of fake news, single classification models often fall short in delivering consistently high accuracy [8]. To address these limitations, ensemble methods have shown promise, offering a more robust solution by combining the strengths of multiple classifiers. However, there are some studies conducted under the ensemble approach there is a need to enhance the classification models’ efficiency in automatic fake news detection in news articles to get better output. Therefore, this study’s proposed ensemble method aims to leverage these findings, building on the existing literature to develop a more efficient, accurate, and adaptable approach to automatic fake news detection in news articles.

III. Methodology

The methodology explains the way the research has been handled with the knowledge of literature review. As mentioned in the literature review, the scope of this research spans many modern computer science fields. As mentioned in the above introduction section, the goal of this research is to find an efficient approach to detect fake news in English news articles. This methodology provides a comprehensive overview of implementation steps which have been carried out to develop this solution. Therefore, this section explains the data set, preprocessing data, and machine learning approaches that are used to accomplish this task.

As shown in Figure 1, the conceptual model explains the overall architecture of the fake news detection research approach. Initially, the collected data was annotated but not preprocessed, so the collected dataset was initially preprocessed by some techniques. Then, feature extraction was done using the CountVectorizer method. In the training stage, three supervised learning models have been used for the selected dataset. Furthermore, in the testing stage, it combined all three classifiers and predicted the result for selected data sets by using the Ensemble model and generating results.

A. Dataset

This study examines the potential for automatic detection of fake news in English texts. For this purpose, a dataset was chosen from the Kaggle website[11], which contained 7796

ISSN: 2961-5992

tweets, and it has binary classification labels as fake or real . That chosen dataset has been trained and tested against the selected machine learning algorithms to measure the performance of detecting the text automatically. The dataset was split into two parts for this purpose: a training set and other one is testing set, with the training set receiving 80% of the data and test set 20%. The dataset contains three attributes such as Title (Title of the News), Text (Text or Content of the News) and Label (Labelling the news as Fake or Real).

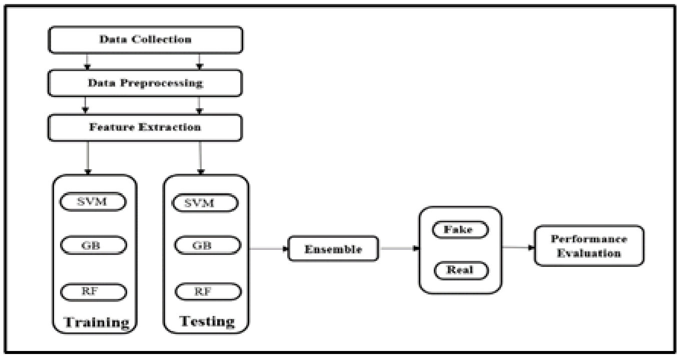


Fig. 1: Proposed model

B. Data Pre-Processing

This research leverages natural language processing and machine learning to train a model for detecting fake news in English-language content. The process involves collecting, preprocessing, and labeling data to ensure a noise-free context. Techniques like stemming reduce dictionary size and improve accuracy. A bag-of-words approach will label news as fake or real, supported by a testing dataset to evaluate the model. Following data collection, the data preparation process includes collecting, cleaning, filtering, and compiling the data into a single file or data table. To prepare for further tasks, the raw data must be initially cleaned and filtered, which comprises annotating the articles and posts in the dataset and creating the preparatory model for fake news identification. A list of the tasks completed during data preparation includes Lemmatization, Lower Case Conversion, Punctuation Removal, Special Characters and Numbers, Stop Word Removal, and Remove Username. Hashtag and URL and Stemming.

C. Proposed Text Preprocessing

To prepare the data for testing and training the model, English news material is pre-processed. The language and basic text handling techniques, such as filtering, lemmatisation, and removing punctuation and special characters, are used in its execution [5]. To prepare the data for testing and training the model, English news material is pre-processed. The language and basic text handling techniques, such as filtering, lemmatisation, and removing punctuation and special characters, are used in its execution [8], See figure 2.

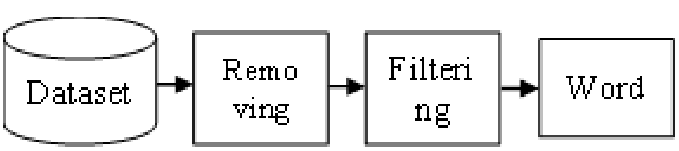


Fig. 2: Data Preprocessing

Special characters, emoticons, symbols, and punctuation are constantly present in news stories from websites or social media platforms. As a result, the text cleaning procedure is crucial to guaranteeing the analysis's accuracy.

D. Feature Extraction

The next stage of the data preprocessing is Feature extraction; it's typically used as an initial step to convert text into a numerical system for training machine learning models. For this purpose, the CountVectorizer technique was used. It's an effective technique for converting text into numerical feature vectors based on word frequencies. It's a type of bag-of-words model.

E. Training and Testing data

The next stage of the data pre-processing is to split the chosen dataset into two categories for the purpose of training and testing the dataset with the chosen machine learning algorithms. that are Training data: 80% and Testing data: 20%.

F. Classification using Ensemble Techniques:

In fake news analysis, classification is used to determine whether the given statement is true or false. The ensemble of classifiers combines the decisions of various weak classifiers to form one strong classifier. In that case, for this purpose Voting classifier was used.

G. Evaluation Metrics

A systematic investigation has been carried out since this research is quantitative in nature. After preprocessing and feature extraction methods to selected dataset, the results were used to develop a model that predicted the outcomes. Then, the model's output data will be examined using metric to determine the research's success. For the experiment, an evaluation measure was constructed and used in its entirety. NLP and ML were used in this study, hence the assessment metrics accuracy, precision, recall, and F1-score were chosen to evaluate the classification model's performance.

H. Classification Accuracy

Classification Accuracy is a measurement system that measures the degree of closeness of measurement between the original data and the detected data.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+FP+TN+FN}} \quad (1)$$

where,

TP – True Positive (actual, real news that is correctly identified as real news)
FN – False Negative (actual real news that is incorrectly identified as fake news)
TN – True negative (an actual fake news that is correctly identified as fake news)
FP – False Positive (actual fake news that is incorrectly identified as real news)

ISSN: 2961-5992

I. Precision Rate

Precision measures how accurate the model's positive predictions are. In other words, out of all the samples the model predicted as belonging to the positive class, precision tells us how many of those were correct. It focuses specifically on the model's performance in predicting the positive class accurately; by looking at the true positives compared to all instances labeled as positive by the model. The following formula for precision calculation (2).

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (2)$$

J. Recall Rate

Recall measures the model’s ability to identify all actual positive cases. In the context of fake news detection, recall tells us how many real news articles the model correctly identified out of all the real news articles in the dataset. The following formula for recall calculation (3).

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \quad (3)$$

K. F-Measure

The F-measure combines both precision and recall into a single score to give a balanced view of a model's performance. It’s useful when you need a balance between precision (how many predictions were correct) and recall (how many actual positives were identified). A high F- Measure means the model is good at identifying the relevant cases without excessive false alarms or missed detections. The following formula for F-measure calculation (4).

$$\text{F} - \text{Measure} = 2 \times \frac{\text{Precision} * \text{Recall}}{(\text{Precision+Recall})} \quad (4)$$

L. Proposed Classifiers

Three algorithms were employed in the first phase of this study: SVM, RF, and GB. These algorithms were considered the basis for ensemble learning.

1) Support Vector Machine (SVM)

Support Vector Machines are ML models that use hyperplanes for classification. The hyperplane is a decision boundary that separates different classes (e.g., real vs. fake news) in the feature space. Finding the optimum margin is SVM's learning objective. Hyperplanes are often a viable solution for linearly separable data. This situation is quite uncommon in non-artificial datasets. For the purpose of improving classification accuracy, the data is transformed into a higher dimension using a kernel function, which ultimately allows for linear separation [9].

2) Random Forest (RF)

A RF is an ensemble classifier, which is a classifier that keeps and blends a variety of decision tree classifiers. Usually, concept bagging with multiple feature sets is used for assembly. A more sophisticated version of the supervised learning model known as Decision Trees (DT) is called RF. The final forecast is based on the class that obtained the most votes, and RF is made up of many decision trees that each predict a class's conclusion independently. Because of the low connection between trees, random forests have a lower error rate than other models [10].

3) Gradient Boost (GB)

Gradient Boosting is a powerful machine learning algorithm widely used for classification and regression tasks, making it particularly useful in fields like fake news detection. This algorithm builds an ensemble of decision trees in a sequential manner, where each new tree corrects the errors made by previous ones. It begins by making initial predictions, often just the mean of the target variable for regression tasks. Then, it calculates the residuals, or errors, between the predicted and actual values. Each subsequent tree is trained to minimize these residuals, essentially "boosting" the model’s accuracy over each iteration [6].

IV. Results and Discussion

In this section, we present the results of the development of an efficient method for automatic detection of Fake news in news articles using an ensemble of Gradient Boosting, Random Forest, and Support Vector Machine (SVM) classifiers. The model was trained and tested using a dataset sourced from Kaggle, which contains labeled examples of both fake and real news. The key performance metrics evaluated include accuracy, precision, recall, F1-score, and the confusion matrix for each individual model, as well as the overall ensemble model, which was used to analyze the performance of the developed method.

A. Result Analysis

The results highlight the performance of individual classifiers, including Support Vector Machine (SVM), Gradient Boosting (GB), and Random Forest (RF), in predicting fake news and Ensemble model performance. Each achieved an accuracy, precision, recall, and F1-score of over 85%, showcasing their effectiveness (Table 01). However, the core contribution lies in the ensemble model, developed using a VotingClassifier with soft voting, which combines the strengths of these classifiers. The ensemble model, as shown in Table 01, achieved the highest accuracy (91%) among all tested models. Its precision (88%) matched Random Forest, while its recall (94%) significantly improved, effectively reducing false negatives. The F1-score (91%) demonstrated a balanced trade-off between precision and recall, minimizing false positives and negatives. By leveraging SVM's linear separability, Gradient Boosting's complex pattern modeling, and Random Forest's robustness, the ensemble model outperformed individual classifiers. This makes it a reliable choice for fake news detection, ensuring superior overall performance.

We plot a confusion matrix for each model after assessing its performance in accordance with Figure 03. Out of the four models, the ensemble model performs best, according to the confusion matrix. It has the highest True Positive (TP) and True

ISSN: 2961-5992

Negative (TN) rates, which means it correctly identifies the most fake news and real news. It also has the lowest False Positive (FP) and False Negative (FN) rates, which means it makes fewer mistakes. This suggests that the Ensemble model is the most reliable for detecting fake news. The number of misclassifications was lower than for any of the individual classifiers from the above table developed ensemble model has highly correct detection, rather than the other classifiers and also the developed ensemble method has lower incorrect detection when compared to other models. This confirms that the ensemble model is not only more accurate overall but also less prone to errors in classifying fake and real news. In the context of fake news detection, a model's ability to minimize false negatives (failing to detect fake news) is particularly important. The high recall of the ensemble model indicates that it excels in this regard, making it a suitable tool for applications where identifying fake news is critical.

TABLE I
SHOWS THE INDIVIDUAL CLASSIFIER AND ENSEMBLE MODEL PERFORMANCE

Classifier	Accuracy	Precision	Recall	F1 Score
SVM	0.87	0.85	0.88	0.86
(GB)	0.88	0.85	0.92	0.88
(RM)	0.89	0.88	0.90	0.89
Ensemble model	0.91	0.88	0.94	0.91

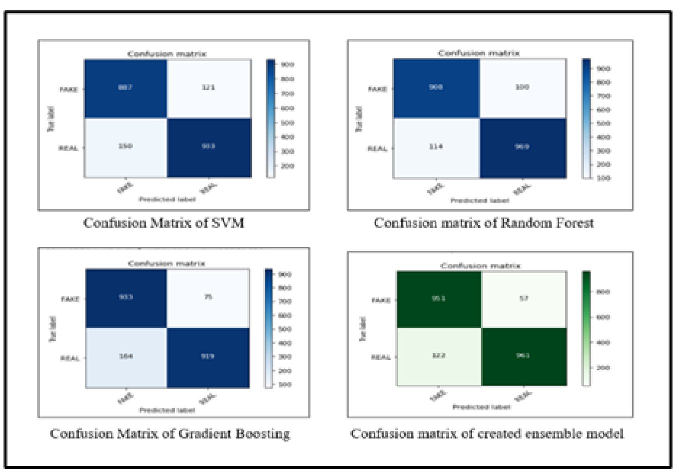
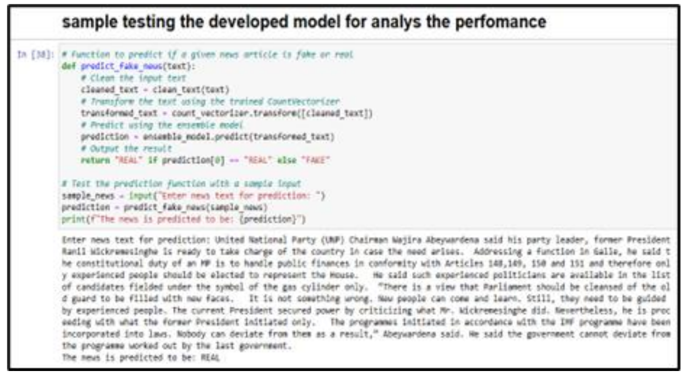


Fig. 3: Confusion Matrix analysis of selected classifiers and created Ensemble Model

The developed efficient model was tested with real-world recognized news articles and website data to evaluate whether the method is reliable or not. I got a reliable prediction from the method that defined real news as real news, and it detected actual fake news as fake news. Figures 04 and 05 prove the real and fake news testing.



sample testing the developed model for analys the perfomance

```
In [38]: # Function to predict if a given news article is fake or real
def predict_fake_news(text):
    # Clean the input text
    cleaned_text = clean_text(text)
    # Transform the text using the trained CountVectorizer
    transformed_text = count_vectorizer.transform([cleaned_text])
    # Predict using the ensemble model
    prediction = ensemble_model.predict(transformed_text)
    # Output the result
    return "REAL" if prediction[0] == "REAL" else "FAKE"

# Test the prediction function with a sample input
sample_news = input("Enter news text for prediction: ")
prediction = predict_fake_news(sample_news)
print(f"The news is predicted to be: {prediction}")

Enter news text for prediction: United National Party (UNP) Chairman Wajira Abeywardena said his party leader, former President Ranil Wickremesinghe is ready to take charge of the country in case the need arises.  Addressing a function in Galle, he said the constitutional duty of an MP is to handle public finances in conformity with Articles 148,149, 150 and 151 and therefore only experienced people should be elected to represent the House.   He said such experienced politicians are available in the list of candidates fielded under the symbol of the gas cylinder only.  "There is a view that Parliament should be cleansed of the old guard to be filled with new faces.   It is not something wrong. New people can come and learn. Still, they need to be guided by experienced people. The current President secured power by criticizing what Mr. Wickremesinghe did. Nevertheless, he is proceeding with what the former President initiated only.   The programmes initiated in accordance with the IMF programme have been incorporated into laws. Nobody can deviate from them as a result," Abeywardena said. He said the government cannot deviate from the programme worked out by the last government.
The news is predicted to be: REAL
```

Fig. 4: Testing the Real News

```
In [39]: # Function to predict if a given news article is fake or real
def predict_fake_news(text):
    # Clean the input text
    cleaned_text = clean_text(text)
    # Transform the text using the trained CountVectorizer
    transformed_text = count_vectorizer.transform([cleaned_text])
    # Predict using the ensemble model
    prediction = ensemble_model.predict(transformed_text)
    # Output the result
    return "REAL" if prediction[0] == "REAL" else "FAKE"

# Test the prediction function with a sample input
sample_news = input("Enter news text for prediction: ")
prediction = predict_fake_news(sample_news)
print(f"The news is predicted to be: {prediction}")

Enter news text for prediction: Singer-songwriter Beyoncé endorsed Vice President Kamala Harris for president at a campaign rally in Houston on Oct. 25. Social media posts have made the unfounded claim that Beyoncé was paid $10 million for the endorsement.
The news is predicted to be: FAKE
```

Fig. 5: Testing the Fake News

V. CONCLUSION

In conclusion, this research developed an effective method for automatically detecting fake news in English news articles by integrating Support Vector Machine (SVM), Gradient Boosting (GB), and Random Forest (RF) classifiers through an ensemble approach.

Using CountVectorizer for feature extraction, the model efficiently captured word patterns, achieving a high accuracy of 91%, a recall of 94%, and F1-score of 91%. The VotingClassifier's soft voting combined the strengths of individual algorithms, ensuring reliable performance with minimal classification errors. Confusion matrix analysis confirmed the model's robustness, making it a scalable and practical solution for real-world applications. The high recall rate underscores its capability to identify fake news effectively, a critical requirement for such systems. Future work will involve creating a user-friendly interface and exploring enhancements through larger datasets, metadata integration, and advanced techniques like deep learning and stacking. The model's applicability across multilingual and diverse fake news scenarios also offers exciting research opportunities.

REFERENCES

[1] C. Dewey, "How fake news spread through social media during the 2016 election," The Washington Post, Nov. 17, 2016.

[2] K. Shu, A. Sliva, S. Wang, J. Tang and H. Liu, "Fake news detection on social media: A data mining perspective," ACM SIGKDD Explorations Newsletter, vol. 19, no. 1, pp. 22–36, 2017.

[3] X. Wang, Y. Liu, J. Zhang and H. Chen, "Challenges in training machine learning classifiers with imbalanced datasets," Journal of Data Science and Machine Learning, vol. 12, no. 4, pp. 325–333, 2020.

ISSN: 2961-5992

[4] H. Kudarvalli and J. Fiaidhi, "Experiments on detecting fake news using machine learning algorithms," International Journal of Reliable Information Assurance, vol. 8, no. 1, pp. 15–26, 2020.

[5] N. Capuano, G. Fenza, V. Loia and F. D. Nota, "Content-based fake news detection with machine and deep learning: A systematic review," Neurocomputing, vol. 530, pp. 91–103, 2023.

[6] J. Y. Khan et al., "A benchmark study on machine learning methods for fake news detection," Computation and Language, preprint, 2019.

[7] S. Guha, "Related fact checks: A tool for combating fake news," arXiv preprint, arXiv:1711.00715, 2017.

[8] F. A. Alshuwaier and F. A. Alsulaiman, "Fake news detection using machine learning and deep learning algorithms: A comprehensive review and future perspectives," Computers, vol. 14, no. 9, p. 394, 2025.

[9] B. Gregorutti, B. Michel and P. Saint-Pierre, "Correlation and variable importance in random forests," Statistics and Computing, vol. 27, no. 3, pp. 659–678, 2017.

[10] M. M. Miron'czuk and J. A. Protasiewicz, "A recent overview of the state-of-the-art elements of text classification," Expert Systems with Applications, vol. 106, pp. 36–54, 2018.

[11] R. Kumar, "Fake News Dataset," Kaggle, 2023. [Online]. Available: https://www.kaggle.com/datasets/rajatkumar30/fake-news

ISSN: 2961-5992