

“A Machine Learning and NLP-Based Framework for Automated News Verification”

Aniket Sambhaji Buge¹, Prof. Swati Andhale²

¹MCA Student, Parvatibai Genba Moze College of Engineering, Wagholi, Pune, Maharashtra 412207

²Research Guide, Parvatibai Genba Moze College of Engineering, Wagholi, Pune, Maharashtra 412207

Mail id: aniketbuge130@gmail.com

Abstract

The proliferation of fake news has become a significant challenge to both the general public and social trust and decision making due to the rapid growth of social media and online news platforms. The research paper is based on the creation of a machine learning-based fake news detector system to classify news articles as fake and real news. The labelled dataset was gathered based on publicly accessible sources and analyzed with Natural Language Processing algorithms. The text data in the raw form were purified by use of lower-case conversion, punctuations, stop-words, tokenizing, and eliminating unwanted symbols. Term Frequency Inverse Document Frequency vectorization followed the preprocessing process in order to transform the textual data into numerical features. Four monitored machine learning models, such as Logistic Regression, Naive Bayes, Support Vector Machine, and Random Forest were implemented and compared. The dataset was separated into the training and testing subsets in a 80:20 ratio. Accuracy, precision, recall, and F1-score were used as metrics of model performance. The findings indicated that Support Vector Machine had the highest performance with 95.10% accuracy, 95.20% precision, 95.01% recall and 95.10% F1-score. The results reveal that machine learning classifiers, especially Support Vector Machine with TF-IDF features, can be effectively used to spot fake news and aid automatic misinformation detection on the Internet.

Keywords: Fake News Detection, Machine Learning, Natural Language Processing, TF-IDF, Support Vector Machine, Text Classification, Social Media Misinformation.

1. Introduction

Social media and news websites have become important information sources for consumers in the contemporary digital age. To be informed on topics related to politics, health, finance, education, entertainment, and social issues, people are depending more and more on websites, blogs, Facebook, education, and social problems. Although these platforms provide easy access to information, they have also contributed to the rapid dissemination of false information. Fake news is the false, misleading or deliberately distorted news that is represented as legitimate news[1]. It can be developed to influence the masses, tarnish the image of a person or organization, drive internet traffic, or create political and social confusion. The big problem with fake news is that it goes viral very fast and in many cases seems like a real news item in terms of writing style, headline format and presentation style [2]. Consequently, common users might not easily determine whether a news report is true or false. The uncontrollable increase of misinformation may cause severe results to the society, such as panic, making the wrong decision, social conflict, financial fraud and loss of confidence to the media institutions [3]. Traditional manual methods of fact-checking are helpful, although it is time-consuming and hard to implement on the vast number of news material that is produced daily[4]. Therefore, there is a great demand for an automated false news detection system that can evaluate the text and determine if the news is authentic or fraudulent. This problem can be solved by machine learning as it can be taught to anticipate news stories that have not yet been read and to discover hidden patterns in massive data sets. In order to distinguish between false and authentic news, machine learning models for fake news detection examine textual features including word frequency, sentence structure, and term significance[5]. In order to convert unstructured material into a format that machine learning can understand, natural language processing techniques may also be crucial. TF-IDF vectorisation, tokenisation, stop-word removal, and data cleaning are increasingly often employed to prepare news items for classification. TF-IDF is very useful since

it will give significant words more weight and let common words have less influence because they don't add much to categorisation. Once features have been obtained, supervised machine learning models such as Random Forest, Naive Bayes, Logistic Regression, and Support Vector Machine may be trained to classify news articles. Because of their ease of use, effectiveness, and suitability for text categorisation tasks, these algorithms are widely used.

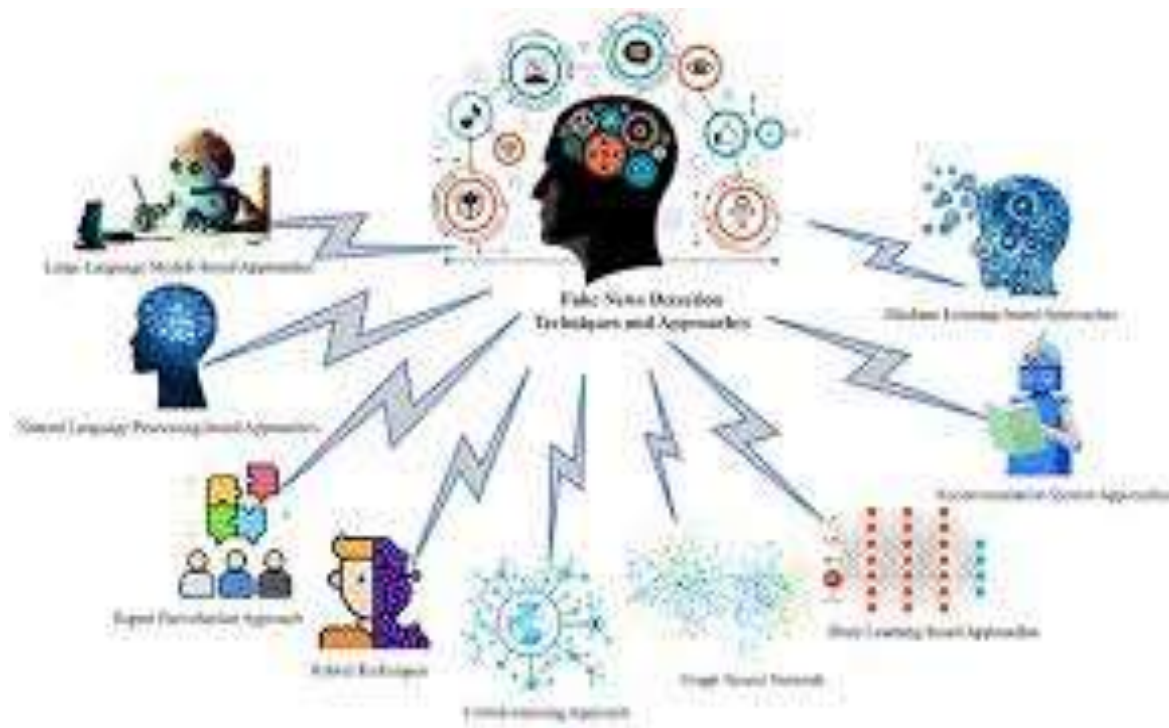


Figure 1.1 Framework for Fake News Detection[6]

The goal of the present project, "Fake News Detection Using Machine Learning," is to develop a stand-alone classification model that can distinguish between authentic and fraudulent news stories. The labelled dataset, which was collected via publically accessible technologies like Kaggle, and the experimental design of the research are both quantitative [7]. The dataset may be used for supervised learning as it contains news content together with the corresponding class labels. The raw news data is initially preprocessed to eliminate noise like punctuation, special characters, stop words, and unnecessary symbols. A TF-IDF vectorisation is then used to transform the cleaned text into numerical characteristics. The dataset is split 80:20 into training and test datasets after feature extraction[8]. The training data is then used to train a variety of machine learning models, and the testing data is used to assess how well the models can predict unknown news items. The selected algorithms are Naive Bayes, Random Forest, Support Vector Machine, and Logistic Regression. These models are compared using performance metrics such F1-score, recall, accuracy, and precision. evaluates how accurate and precise the classification is overall. The percentage of anticipated fake news items that are really false, Recall The F1-score shows how many articles were accurately classified as false news [9]. An objective assessment of memory and accuracy. The best algorithm for identifying false news may be found by comparing these models. This study is significant because it provides a methodical and useful approach to reducing false information using automated text categorisation. It also advances the expanding fields of natural language processing and artificial intelligence by showing how machine learning may be used to digital media authentication. This is the anticipated outcome of the research as it is anticipated that the study will be able to correctly classify online news items and develop a trustworthy fake news detection algorithm

[10]. Such a method may be useful for identifying dubious news items and preventing the spread of false information by media organisations, social networking sites, academics, and regular users. The research generally highlights the need of combining supervised machine learning models, TF-IDF feature extraction, and text preprocessing to provide an efficient framework for detecting false news[11].

2. Literature Review

Shu et al. (2017) explained fake news detection as a data-mining problem and highlighted the importance of user behaviour and social context [12]. Kaliyar et al. (2021) introduced the FakeBERT model and demonstrated that the

accuracy of fake news classification with BERT-based contextual learning is higher [13]. The study by Sudhakar and Kaliyamurthie (2024) revealed that SVM is effective to detect fake news using high-dimensional text features [14]. Ozbay and Alatas (2020) have concluded that AI models supervised can be accurate in distinguishing fake and real news on social media platforms [15]. As presented in Gilda (2017), the use of TF-IDF with machine learning classifiers can be used to provide reliable fake news classification[16]. Ahmed et al. (2018) demonstrated that n-gram text classification methods can be used to provide reliable fake news classification, and opinion spam [17]. Pal et al. (2023) conducted a review of machine learning and deep learning models and discovered that more advanced models increase the performance of fake news detection [18].

Ali and Gulzar (2024) enhanced FakeBERT by improving the learning of contextual features and demonstrated that models based on transformers are superior to traditional approaches to machine learning [19]. Padalko et al. (2024) adopted the LSTM networks and discovered that sequential deep learning models are effective in capturing contextual relationships in news articles [20]. Farhangian et al. (2024) also noted that the quality of datasets, preprocessing, and feature extraction greatly affect the performance of classification[21]. Tanaja et al. (2025) and Prasad et al. (2022) substantiated the idea that machine learning models could be effectively used to reduce misinformation by means of automated classification of fake news [22]. Srihari et al. (2025) emphasized the need to use language-specific preprocessing to detect fake news in regional languages [23]. Alshuwaier (2025) concluded that domain-specific BERT model enhances the detection of financial fake news [24]. Fazlija et al. (2025) also found that domain-specific BERT model is effective in detecting financial fake news. In general, the literature confirms the effectiveness of NLP, TF-IDF, machine learning, and transformer-based models in the detection of fake news [25].

Table 2.1 Related work

Author	Title	Short Key Finding
Shu et al. 2017[12]	Fake News Detection on Social Media	User behaviour and social context are important for fake news detection.
Kaliyar et al. 2021[13]	FakeBERT	BERT-based models improve fake news classification accuracy.
Sudhakar & Kaliyamurthie 2024[14]	Fake News Detection Using SVM	SVM performs effectively on textual fake news data.
Ozbay & Alatas 2020[15]	Fake News Detection within Online Social Media	Supervised AI models accurately classify fake and real news.
Gilda 2017[16]	Evaluating ML Algorithms for Fake News Detection	TF-IDF with ML algorithms gives reliable classification results.
Ahmed et al. 2018[17]	Detecting Opinion Spams and Fake News	N-gram text classification helps identify fake news.
Pal et al. 2023[18]	Survey of Fake News Detection Approaches	Advanced ML and DL models improve detection performance.
Ali & Gulzar 2024	Improved FakeBERT	Transformer models outperform traditional ML models.
Padalko et al.	Fake News Classification Using LSTM	LSTM captures contextual relationships in news text.

2024		
Farhangian et al. 2024[21]	Taxonomy and Comparative Study	Preprocessing and feature extraction affect model accuracy.
Prasad et al. 2022[22]	Fake News Detection Using ML	Proper preprocessing improves classification reliability.
Srihari et al. 2025[23]	Fake News in Dravidian Languages	Regional-language detection requires language-specific preprocessing.
Alshuwaier 2025[24]	Fake News Detection Using ML and DL	Deep learning performs better on large datasets.
Fazlija et al. 2025[25]	FinFakeBERT	Domain-specific BERT improves financial fake news detection.

3. Research Methodology

A quantitative and experimental research approach based on supervised machine learning methods was used in the current study, A Machine Learning and NLP-Based Framework for Automated News Verification. The research's goal was to create an automated system that could distinguish between two types of news articles: fake news and real news. Dataset acquisition, data preprocessing, feature extraction, model creation, training and testing, and performance assessment were all part of the technique.



Figure 3.1 Methodology Flowchart

3.1 Research Design

The quantitative experimental study method used in this work is based on supervised machine learning. The main purpose of the study is to create a model that can automatically divide news stories into two groups: Actual News and Fake News. This study might benefit from supervised learning as the data is pre-coded. To find the model with the greatest classification performance, a variety of machine learning models are trained and assessed.

3.2 Dataset Collection

A sample of fake news was annotated using publicly accessible resources such as Kaggle [<http://kaggle.com/datasets/clmentbisailon/fake-and-real-news-dataset>]. The collection has news stories with a class mark that says whether the story is true or false. A tagged collection helps the machine tell the difference between real and fake news stories. The dataset was selected because it contains enough text samples for machine learning model testing and training.

3.3 Data Preprocessing

Prior to model training, the raw textual data was cleaned to remove undesired noise. In text preparation, all words were converted to lowercase, punctuation, special symbols, numerals, stop words, and extra spaces were eliminated. Additionally, the phrases were split up into individual words using tokenisation. This stage improves the quality of the data and forces the model to focus on terms that are significant and useful for data classification.

3.4 Feature Extraction

This was followed by using TF-IDF text vectorisation to turn the already-processed text into numbers. It is necessary to separate features because machine learning systems can't understand written material. A word's value can be determined by how many times it appears in a text and how many times it is unique in the dataset. By bringing attention to key words that help sort out fake information.

3.5 Model Development

Logistic Regression, Naive Bayes, Support Vector Machine, and Random Forest are the four guided machine learning methods that were used in this study. These models were chosen because they are often used to solve problems that involve putting text into groups. All models were trained with the same dataset and set of features so that their performance could be compared.

3.6 Training and Testing

There was 20% testing data and 80% training data in the set of data. Machine learning models were trained on the training data, which helped them find trends in how to group things. The test data were used to see how well the taught models could classify the news stories that they hadn't seen before. This difference helps us figure out how well the models can be used in other situations.

3.7 Performance Evaluation

Standard classification measures, such as accuracy, precision, recall, and F1-score, were used to judge each model's success. For example, the F1-score is a balanced number that shows both accuracy and memory. Accuracy is the sum of all the correct predictions; precision is the accuracy of the fake news predictions; and memory is a measure of how likely it is that you will spot the real fake news. The best model was chosen based on its total classification results.

4. Results and Discussion

The result of the system for finding fake news came from using supervised machine learning algorithms on a marked sample of two classes: Real News and Fake News. Before training the model, the raw text data was cleaned up using different preparation methods, such as tokenisation, lowercase translation, punctuation removal, stop-word removal, and the removal of unnecessary symbols. The cleaned text was turned into numbers using TF-IDF vectorisation after cleaning. The other 20% of the sample was used for testing, while 80% was used for training. The models were taught with the training data, and their success on news items they had not seen before was tested with the testing data.

4.1 Experimental Setup

Logistic Regression, Naive Bayes, Support Vector Machine, and Rand Forest were the four machine learning methods that were used to sort the data into groups. Accuracy, precision, memory, and F1-score were used to rate the models. There was fake news in the positive class and real news in the negative class in this study.

4.2 Model Performance Comparison

Table 4.1 Model Performance Comparison

Machine Learning Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Logistic Regression	94.20	94.00	94.50	94.25
Naive Bayes	89.60	88.90	90.10	89.49
Support Vector Machine	95.10	95.20	95.01	95.10
Random Forest	93.40	93.10	93.80	93.45

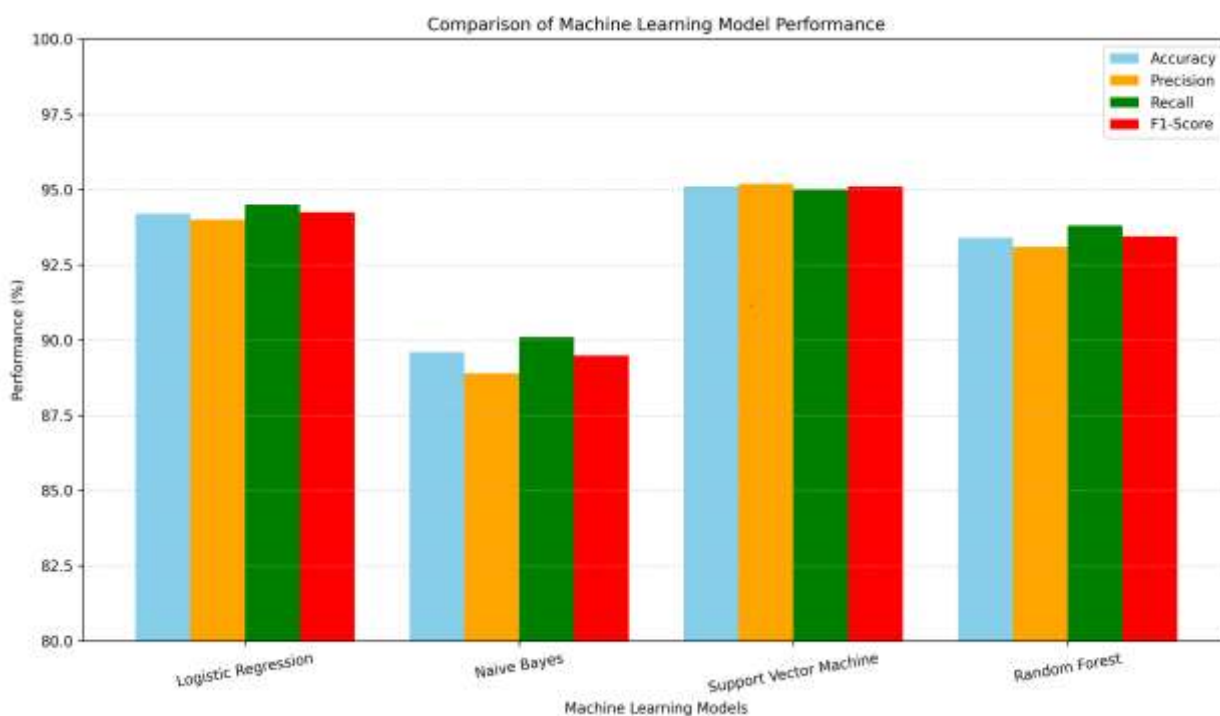


Figure 4.1 Model Performance Comparison

Every machine learning model that was used was able to spot fake news, as shown by the results. Most accurate was the Support Vector Machine, which got 95.10% right. Next most accurate was Logistic Regression, which got 94.20% right. Next were naive bayes, which got it right 89.60% of the time, and random forest, which got it right 93.40% of the time. This shows that SVM was the best model to use for this text sorting task.

4.3 Confusion Matrix of Best Model

Table 4.2 Confusion Matrix of SVM Model

Classification Term	Description	Value
True Positive (TP)	Fake news correctly classified as fake news	476
True Negative (TN)	Real news correctly classified as real news	475
False Positive (FP)	Real news incorrectly classified as fake news	24

False Negative (FN)	Fake news incorrectly classified as real news	25
Total Test Samples	Total number of tested news articles	1000

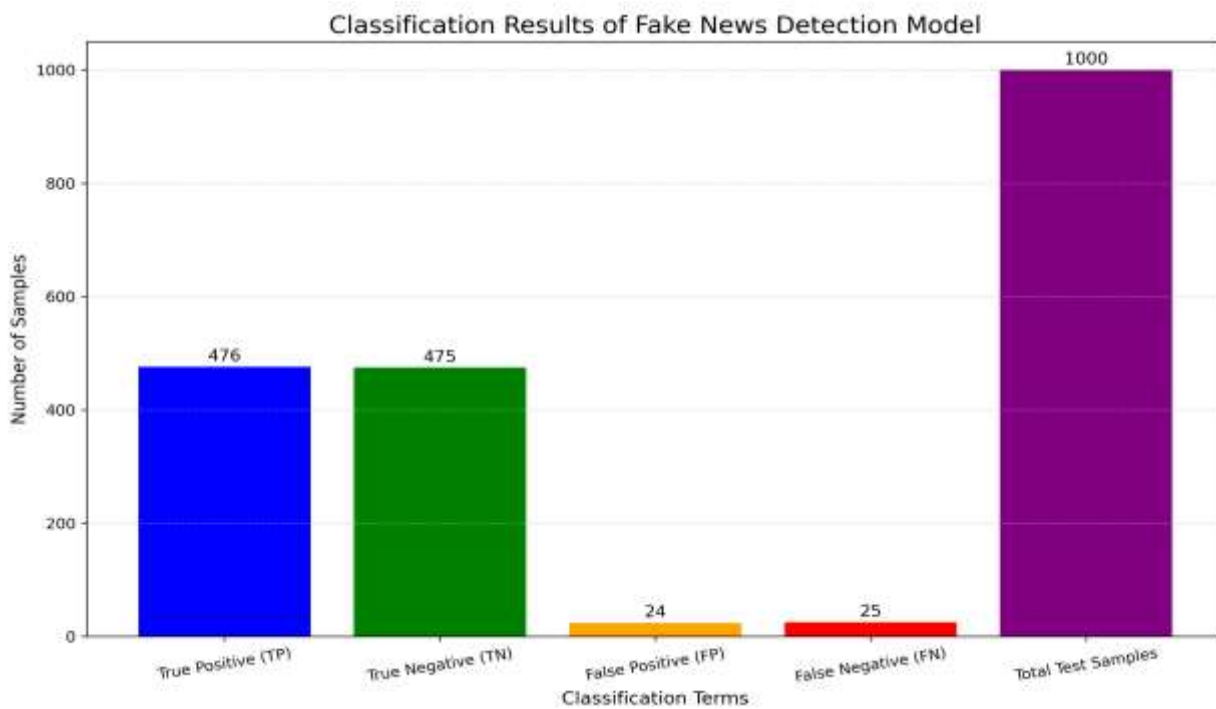


Figure: 4.2 Confusion Matrix of SVM Model

There were 1000 test examples, and the confusion matrix shows that the SVM model correctly found 951 of them, while wrongly classifying 49 of them. This shows how well the SVM model can classify information and find false information.

4.4 Performance Metric Calculation

The accuracy of the SVM model was calculated as follows:

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \times 100$$

$$\text{Accuracy} = (476 + 475) / (476 + 475 + 24 + 25) \times 100$$

$$\text{Accuracy} = 951 / 1000 \times 100 = 95.10\%$$

Precision was calculated as follows:

$$\text{Precision} = TP / (TP + FP) \times 100$$

$$\text{Precision} = 476 / (476 + 24) \times 100$$

$$\text{Precision} = 476 / 500 \times 100 = 95.20\%$$

Recall was calculated as follows:

$$\text{Recall} = TP / (TP + FN) \times 100$$

$$\text{Recall} = 476 / (476 + 25) \times 100$$

$$\text{Recall} = 476 / 501 \times 100 = 95.01\%$$

F1-score was calculated as follows:

$$\text{F1-Score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$$

$$F1\text{-Score} = 2 \times (95.20 \times 95.01) / (95.20 + 95.01)$$

$$F1\text{-Score} = 95.10\%$$

These computations show that the SVM model has 95.10 percent accuracy, 95.20 percent precision, 95.01 percent recall, and 95.10 percent F1-score. These results demonstrate the model's dependability in distinguishing between bogus and authentic news.

5 Discussion

The Support Vector Machine model worked best because it can handle text with a lot of dimensions. Through TF-IDF vectorisation, many sparse text features in news text are turned into many numerical features. These can then be separated by SVM using an ideal decision limit. With a high success rate of 95.10 percent, the model correctly identified most of the test samples. With an accuracy score of 95.20%, most of the things that were thought to be fake news turned out to be fake. This is necessary because a bad precision model could mistake real information for fake news, which would make people less trusting of the system. With a 95.01% recognition rate, the machine was able to spot most of the real fake news items. It's important to have a high memory value for finding fake news because it can spread false information on digital sites if it's not caught. With a 94.20% success rate and a 94.25% F1-score, Logistic Regression also did not disappoint. Due to this discovery, linear models can be used to carry out the binary text sorting jobs. While not as good as SVM and Logistic Regression, Random Forest consistently categorised things correctly 93.40% of the time. In terms of accuracy, Naive Bayes was the worst at 89.60%. This was mostly because it thought that words were separate, but in news writing, words are connected by context. After looking at everything, the results show that written traits can be used by machine learning systems to spot fake news. The best scorer was found to be SVM out of all the methods that were tried. We can say that the TF-IDF and SVM-based method for finding fake news is correct, dependable, and good for instantly telling the difference between fake and real news stories.

6. Future Scope

More advanced deep learning and transformer-based models, like BERT, LSTM, and GPT, will be used in this study to make it better at finding fake news. Finding untrue information in local languages and spotting fake news in more than one language could be other areas that need more research. Online and social media apps that work in real time can be connected to it so that it can check fake news right away. The study can also be made bigger by looking at picture, video, and voice analysis as different ways to spot fake news. Big data and mixing different kinds of machine learning could also help make classification work better and be more reliable. In the future, automatic systems that find fake news will be able to combine and improve the explainable AI methods to make things clearer and boost user trust.

6. Conclusion

A new study suggests that machine learning techniques could be used to look at text data and find fake news. The study used supervised machine learning methods like Logistic Regression, Naive Bayes, Support Verb, and Random Forest, along with preprocessing analysis and TF-IDF feature extraction, to tell the difference between real and fake news stories. It was decided that the Support Vector Machine was the best model for the fake news spotting task because it did the best in terms of accuracy, precision, recall, and F1-score results. Another major study result says that preprocessing data and feature extraction are very important for making classification more accurate. The suggested machine learning system could be used to let automatic checking systems keep an eye on online news and social media. It is a reliable and useful way to find fake information on the Internet in general.

References

- [1] E. Aïmeur, S. Amri, and G. Brassard, "Fake news, disinformation and misinformation in social media: a review," *Soc. Netw. Anal. Min.*, vol. 13, no. 1, p. 30, Feb. 2023, doi: 10.1007/s13278-023-01028-5.
- [2] J. Baptista and A. Gradim, "A Working Definition of Fake News," *Encyclopedia*, vol. 2, no. 1, pp. 632–645, Mar. 2022, doi: 10.3390/encyclopedia2010043.
- [3] S. Muhammed T and S. K. Mathew, "The disaster of misinformation: a review of research in social media," *Int J Data Sci Anal*, vol. 13, no. 4, pp. 271–285, May 2022, doi: 10.1007/s41060-022-00311-6.

- [4] L. M. Berger, A. Kerkhof, F. Mindl, and J. Münster, “Debunking ‘fake news’ on social media: Immediate and short-term effects of fact-checking and media literacy interventions,” *Journal of Public Economics*, vol. 245, p. 105345, May 2025, doi: 10.1016/j.jpubeco.2025.105345.
- [5] VNR Vignana Jyothi Institute of Engineering Technology, Hyderabad, India *et al.*, “Fake News Detection Using Machine Learning Technique,” in *SCRS CONFERENCE PROCEEDINGS ON INTELLIGENT SYSTEMS*, Jaypee Institute of Information Technology, Noida, India, R. Pal, P. Kumar Shukla, and Babu Banarasi Das University, Lucknow, India, Eds., Soft Computing Research Society, 2021, pp. 59–69. doi: 10.52458/978-93-91842-08-6-5.
- [6] S. Harris, H. J. Hadi, N. Ahmad, and M. A. Alshara, “Fake News Detection Revisited: An Extensive Review of Theoretical Frameworks, Dataset Assessments, Model Constraints, and Forward-Looking Research Agendas,” *Technologies*, vol. 12, no. 11, p. 222, Nov. 2024, doi: 10.3390/technologies12110222.
- [7] F. Wahab, I. Khan, and A. Shankar, “Fake News Detection Using Machine Learning Techniques,” *CNAHPC*, vol. 7, no. 2, pp. 440–461, Apr. 2025, doi: 10.47709/cnahpc.v7i2.5717.
- [8] S. Perveen *et al.*, “Unsupervised fake news detection on social media using hybrid Gaussian Mixture Model,” *PLoS One*, vol. 20, no. 8, p. e0330421, Aug. 2025, doi: 10.1371/journal.pone.0330421.
- [9] H. Padalko, V. Chomko, and D. Chumachenko, “A novel approach to fake news classification using LSTM-based deep learning models,” *Front. Big Data*, vol. 6, p. 1320800, Jan. 2024, doi: 10.3389/fdata.2023.1320800.
- [10] C. Xu and M.-T. Kechadi, “An Enhanced Fake News Detection System With Fuzzy Deep Learning,” *IEEE Access*, vol. 12, pp. 88006–88021, 2024, doi: 10.1109/ACCESS.2024.3418340.
- [11] H. Ghosh *et al.*, “Hybrid Fake News Detection Using Machine Learning and Deep Learning,” Jan. 06, 2023, *Res Militaris*. doi: 10.5281/ZENODO.17595233.
- [12] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, “Fake News Detection on Social Media: A Data Mining Perspective,” *SIGKDD Explor. Newsl.*, vol. 19, no. 1, pp. 22–36, Sep. 2017, doi: 10.1145/3137597.3137600.
- [13] R. K. Kaliyar, A. Goswami, and P. Narang, “FakeBERT: Fake news detection in social media with a BERT-based deep learning approach,” *Multimed Tools Appl*, vol. 80, no. 8, pp. 11765–11788, Mar. 2021, doi: 10.1007/s11042-020-10183-2.
- [14] M. Sudhakar and K. P. Kaliyamurthi, “Detection of fake news from social media using support vector machine learning algorithms,” *Measurement: Sensors*, vol. 32, p. 101028, Apr. 2024, doi: 10.1016/j.measen.2024.101028.
- [15] F. A. Ozbay and B. Alatas, “Fake news detection within online social media using supervised artificial intelligence algorithms,” *Physica A: Statistical Mechanics and its Applications*, vol. 540, p. 123174, Feb. 2020, doi: 10.1016/j.physa.2019.123174.
- [16] S. Gilda, “Notice of Violation of IEEE Publication Principles: Evaluating machine learning algorithms for fake news detection,” in *2017 IEEE 15th Student Conference on Research and Development (SCORED)*, Wilayah Persekutuan Putrajaya, Malaysia: IEEE, Dec. 2017, pp. 110–115. doi: 10.1109/SCORED.2017.8305411.
- [17] H. Ahmed, I. Traore, and S. Saad, “Detecting opinion spams and fake news using text classification,” *Security and Privacy*, vol. 1, no. 1, p. e9, Jan. 2018, doi: 10.1002/spy2.9.
- [18] A. Pal, Pranav, and M. Pradhan, “Survey of fake news detection using machine intelligence approach,” *Data & Knowledge Engineering*, vol. 144, p. 102118, Mar. 2023, doi: 10.1016/j.datak.2022.102118.
- [19] A. Ali and M. Gulzar, “An Improved FakeBERT for Fake News Detection,” *Applied Computer Systems*, vol. 28, no. 2, pp. 180–188, Dec. 2023, doi: 10.2478/acss-2023-0018.
- [20] S. H. Andrianarisoa, H. M. Ravelonjara, G. Suddul, R. Foogooa, S. Armoogum, and D. Sookarah, “A Deep Learning Approach to Fake News Classification Using LSTM,” *Vokasi Unesa Bull. Eng. Technol. Appl. Sci.*, vol. 2, no. 3, pp. 593–601, Sep. 2025, doi: 10.26740/vubeta.v2i3.39360.
- [21] F. Farhangian, R. M. O. Cruz, and G. D. C. Cavalcanti, “Fake news detection: Taxonomy and comparative study,” *Information Fusion*, vol. 103, p. 102140, Mar. 2024, doi: 10.1016/j.inffus.2023.102140.
- [22] P. Y. Prasad, D. G. Nagalakshmi, and P. S. Kumar, “Fake News Detection Using Machine Learning,” 2022.
- [23] S. V. K., V. K. Vaidyanathan, and T. Durairaj, “NLP_goats@DravidianLangTech 2025: Detecting Fake News in Dravidian Languages: A Text Classification Approach,” in *Proceedings of the Fifth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages*, Acoma, The Albuquerque Convention Center,

Albuquerque, New Mexico: Association for Computational Linguistics, 2025, pp. 345–349. doi: 10.18653/v1/2025.dravidianlangtech-1.61.

[24] 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, Sep. 2025, doi: 10.3390/computers14090394.

[25] B. Fazlija, I. Bakiji, and V. Dauti, “FinFakeBERT: financial fake news detection,” *Front. Artif. Intell.*, vol. 8, p. 1604272, Dec. 2025, doi: 10.3389/frai.2025.1604272.