Fake News Detection using Machine Learning Algorithms

*Note: Sub-titles are not captured in Xplore and should not be used

Adithya Challa, Rahul Srikanth
Department of Computer Science & Engineering (Data Science)
School of Engineering
Dayananda Sagar Univerity
Bengaluru, India
adithya.challa1@gmail.com,
rahul.n.srikanth@gmail.com

Sri Sai Aravindan, Suhas S
Department of Computer Science & Engineering (AI & ML)
School of Engineering
Dayananda Sagar Univerity
Bengaluru, India
aravinda6547@gmail.com,
suhas.cse20@gmail.com

Tina Babu
Department of Computer Science & Engineering
School of Engineering
Dayananda Sagar Univerity
Bengaluru, India
b.tina-cse@dsu.edu.in

***Abstract*—The prevalence of fake news poses a significant challenge to the credibility of news and information sources, particularly in the context of Indian politics. In response to this issue, we propose a machine learning-based approach to detect fake news by analyzing news titles through vectorization and tokenization with a pre-defined dataset of news articles labeled as true or fake. Our goal is to develop a model that can accurately classify news articles based on their textual content, distinguishing between genuine news and fabricated propaganda.**

To evaluate the performance of our approach, we conduct experiments on a benchmark dataset of news articles. The results demonstrate that our proposed method achieves high accuracy in detecting fake news, outperforming several state-of-the-art methods in the literature. Our model can effectively help detect fake news and protect the credibility of news and information sources in the context of Indian politics.This approach can be extended to other domains and languages, providing a robust and effective mechanism for detecting and mitigating the impact of fake news.

***Index Terms*—Fake News, Self-Learning, Response Generation, Natural Language Processing, Context Free Grammar, Stochastic Gradient Decent**

I. Introduction

[1]The term ”fake news” carries divergent meanings for different individuals. Essentially, it refers to false news stories lacking verifiable facts, sources, or quotes. Such stories can be deliberate propaganda designed to mislead readers or created as ”click bait” for economic incentives tied to the number of clicks. The rapid and effortless sharing of information on social media platforms has led to a proliferation of fake news in recent years. This research paper explores the feasibility of detecting fake news based solely on textual information using traditional machine-learning techniques. To develop effective detection models, it is crucial to grasp the concept and characteristics of fake news. This involves two key aspects: understanding the nature of fake news (characterization) and subsequently building detection models. Starting with a thorough characterization is fundamental to successful fake news detection.

A. Fake News Characterization

[2]The definition of fake news comprises two essential components: authenticity and intent. Authenticity refers to the presence of false information within the content that can be verified as such. It implies that conspiracy theories are not categorized as fake news since their veracity is often challenging to ascertain. The second aspect, intent, pertains to the purposeful creation of false information with the aim of deceiving the reader.

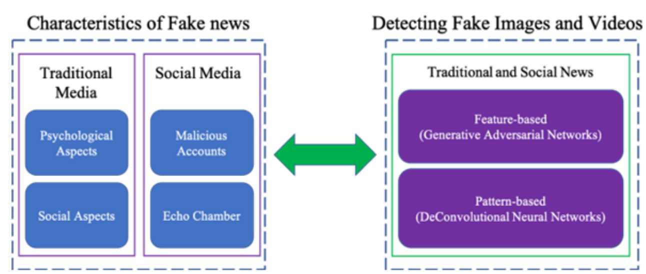

Fig. 1. Process of Characterization

B. Fundamental Theories

[3] Insights from fundamental theories on human cognition and behavior, developed in disciplines such as social sciences and economics, hold immense value in the analysis of fake news. These theories present opportunities for qualitative and quantitative investigations of large-scale fake news data. Moreover, they can aid in constructing well-justified and interpretable models for detecting and addressing fake news, which have been scarce thus far. Through an extensive review of literature spanning various disciplines, we have identified notable theories that hold potential for studying fake

news. Table 2 provides these theories, accompanied by brief descriptions, focusing on their relevance to either (1) the news itself or (2) the individuals responsible for its dissemination.

1) News-related theories: [4] News-related theories reveal the possible characteristics of fake news content compared to true news content. For instance, theories have implied that fake news potentially differs from the truth in terms of, e.g., writing style and quality, quantity such as word counts, and sentiments expressed. It should be noted that these theories, developed by forensic psychology, target deceptive statements or testimonies but not fake news, though these are similar concepts. Thus, one research opportunity is to verify whether these attributes are statistically distinguishable among disinformation, fake news, and the truth using big fake news data.

2) User-related theories: [5] User-related theories investigate the characteristics of users involved in fake news activities, e.g., posting, forwarding, liking, and commenting. Fake news, unlike information such as fake reviews, can "attract" both malicious and normal users. Malicious users spread fake news often intentionally and are driven by benefits. Some normal users (which we denote as vulnerable normal users) can frequently and unintentionally spread fake news without recognizing the falsehood. Such vulnerability psychologically stems from (i) social impacts and (ii) self-impact

II. Literature Review

Nguyen Vo student of Ho Chi Minh City University of Technology (HCMUT) Cambodia did his research on fake news detection and implemented it. He used Bi-directional GRU with Attention mechanism in his project fake news detection; Yang et al. originally proposed this mechanism. He also used some Deep learning algorithms and tried to implement other deep learning models such that AutoEncoders, GAN, CNN [6], [7].

Samir Bajaj of Stanford University published a research paper on fake news detection. He detects fake news with the help of the NLP perspective and implements some other deep learning algorithms. He took an authentic data set from the Signal Media News dataset. [8]

Facebook and WhatsApp are also working on fake news detection as they wrote in an article. They have been working for the last few years, and it is currently under the alpha phase.

Three students of Vivekananda Education Society's Institute of Technology, Mumbai published their research paper on fake news detection. They wrote in their research paper; the social media age started in the 20th century. Eventually, the web usage is increasing, the posts are increasing, the number of articles is increasing. They used various techniques and tools to detect fake news like NLP techniques, machine learning, and artificial intelligence.

A student named Avinash Shakya from ABES Engineering College, Lucknow published his research paper on fake news detection. He wrote in his research paper that most smartphone users prefer to read the news via social media over the internet. Though the news websites publishing the news provide the source of the authentication. There is no suitable way to authenticate the news on social media like WhatsApp, Twitter, Facebook, and other microblogs and social media websites. They provided a strategy of a mix of Naive Bayes classifier, Support vector machines, and semantic investigation. This three-section strategy is a blend between machine 9 learning calculations that subdivide into managed learning procedures and characteristic language-preparing techniques. They got an accuracy of 93.50 percent using this method.

Taking inspiration from these studies, we have worked on improving the accuracy of fake news detection using traditional machine-learning techniques. By analyzing the tokens of words in news titles, we have applied vectorization to build our model for classification. Our goal is to build a model that can accurately classify a given article as either true or fake, and thereby contribute to the fight against the spread of fake news.

III. Methodology

A. Proposed Framework

In our proposed framework, We are expanding on the current literature by introducing ensemble techniques with various linguistic feature sets to classify news articles from multiple domains as true or fake along with Linguistic Inquiry and Word Count (LIWC) feature.

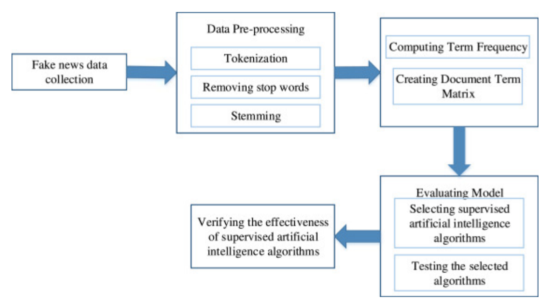

Fig. 2. Flowchart of Fake News detection

There are numerous reputed websites that post legitimate news content which is used for fact checking. In addition, there are open repositories which are maintained by researchers to keep an up-to-date list of currently available datasets and hyperlinks to potential fact-checking sites that may help in countering false news spread. However, we selected three datasets for our experiments which contain news from multiple domains (such as politics, entertainment, technology, and sports) and contain a mix of both truthful and fake articles, and merged the three datasets into a large dataset. The datasets are available online and are extracted from Kaggle. [9]

B. Algorithms

[10] We used the following learning algorithms in conjunction with our proposed methodology to evaluate the performance of fake news detection classifiers.

1) Naïve Bayes: Naive Bayes is a probabilistic classifier that draws inspiration from the Bayes theorem, operating under a straightforward assumption of attribute conditional independence. The classification process involves computing the maximum posterior probability, specifically P(Ci—X), while employing the aforementioned assumption within Bayes' theorem. By making this assumption, the computational complexity is significantly reduced, as only the class distribution needs to be taken into account. Despite the fact that this assumption does not hold true in most scenarios due to attribute dependence, Naive Bayes has demonstrated impressive performance.

$$p(\mathbf{x} \mid C_k) = \prod_{i=1}^{n} p_{ki}^{x_i}(1-p_{ki})^{(1-x_i)} \tag{1}$$

2) Logistic Regression: Logistic regression derives its name from the central function employed in the method, known as the logistic function or sigmoid function. Originally formulated by statisticians, the logistic function was devised to explain the characteristics of population growth in ecological studies. This function exhibits a rapid rise and eventually levels off at the carrying capacity of the environment. Represented by an S-shaped curve, it has the ability to transform any real-valued number into a value ranging between 0 and 1, while never reaching the exact boundaries of 0 or 1.

$$\frac{1}{1+e^{-value}} \tag{2}$$

The input values (x) are linearly combined using weights or coefficients (represented as the Greek capital letter Beta) to make predictions for an output value (y). Unlike linear regression, logistic regression models the output value as a binary value (0 or 1) rather than a numeric value. Here is an example of a logistic regression equation:

$$y = \frac{e^{b_0+b_1*x}}{1+e^{b_0+b_1*x}} \tag{3}$$

3) Support Vector Machine (SVM): Support Vector Machine (SVM) is a highly popular algorithm in Supervised Learning that is utilized for both Classification and Regression problems. However, its primary application lies in Classification tasks within the field of Machine Learning. The main objective of the SVM algorithm is to establish an optimal line or decision boundary capable of effectively segregating n-dimensional space into distinct classes. This enables accurate categorization of new data points in the future. The optimal decision boundary, known as a hyperplane, is created by selecting the most extreme points or vectors that assist in its formation. These exceptional cases are referred to as support vectors, thus giving rise to the name Support Vector Machine for the algorithm. [11]

$$\begin{aligned} b = \mathbf{w}^{\mathsf{T}}\varphi(\mathbf{x}_i) - y_i &= \left[\sum_{j=1}^{n} c_j y_j \varphi(\mathbf{x}_j)\cdot\varphi(\mathbf{x}_i)\right] - y_i \\ &= \left[\sum_{j=1}^{n} c_j y_j k(\mathbf{x}_j,\mathbf{x}_i)\right] - y_i. \end{aligned} \tag{4}$$

4) Decision Tree Learning: Decision Trees are a form of Supervised Machine Learning, where the input and corresponding output are defined in the training data. In this algorithm, data is repeatedly divided based on specific parameters. The structure of a decision tree consists of two fundamental elements: decision nodes and leaves. The leaves represent the final outcomes or decisions, while the decision nodes denote the points where the data is split based on certain criteria. By recursively traversing the decision nodes, a decision tree can provide a clear and interpretable flow of decisions leading to the final outcomes. [12]

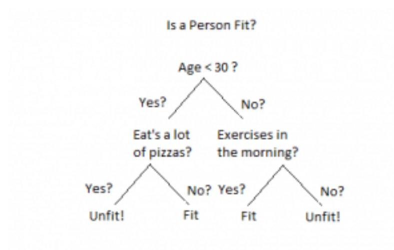

Fig. 3. Decision Tree Learning

5) Random Forest: Random Forest (RF) is an advanced variant of the Decision Tree (DT) algorithm and falls under the category of supervised learning models. RF comprises a large ensemble of individual decision trees that independently make predictions for a target class. The final prediction is determined by aggregating the class that receives the majority of votes from the individual trees. One key advantage of RF is its low error rate compared to other models, owing to the low correlation between the trees.

To train our random forest model, we experimented with different parameters. A grid search technique was employed, varying the number of estimators, to identify the optimal model capable of making highly accurate predictions.

When it comes to determining splits in the decision tree, various algorithms exist for regression or classification problems. In the case of classification problems, we employed the Gini index as a cost function to assess and determine the best split in the dataset. [13]

$$Gini = 1 - \sum_{i=1}^{n}(p_i)^2 \tag{5}$$

C. Datasets

The datasets utilized in this study are openly accessible and freely available on the internet. They encompass a range of news articles, comprising both fake and truthful content, sourced from diverse domains. Truthful news articles reflect accurate depictions of real-world events, whereas fake news websites propagate claims that are not substantiated by factual evidence. For this research, three distinct datasets were employed. The combined dataset comprises 18 collections of articles from these three datasets, which are subsequently referred to as ”True” and ”Fake.” Ultimately, these two final datasets are merged to create a comprehensive and extensive dataset, referred to as ”data.” [14]

D. Performance Metrics

To evaluate the performance of the algorithms, I used confusion matrix. A confusion matrix is a tabular representation of a classification model performance on the test set, which consists of four parameters: true positive, false positive, true negative, and false negative.

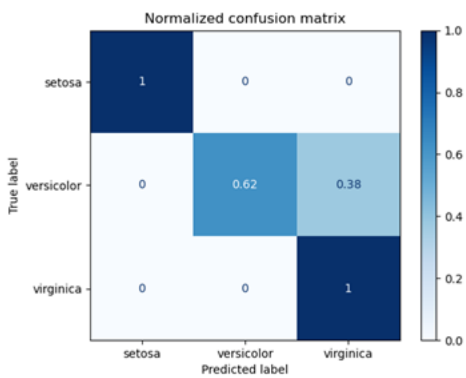

Fig. 4. Confusion Matrix representation

IV. Result Analysis

A. Results

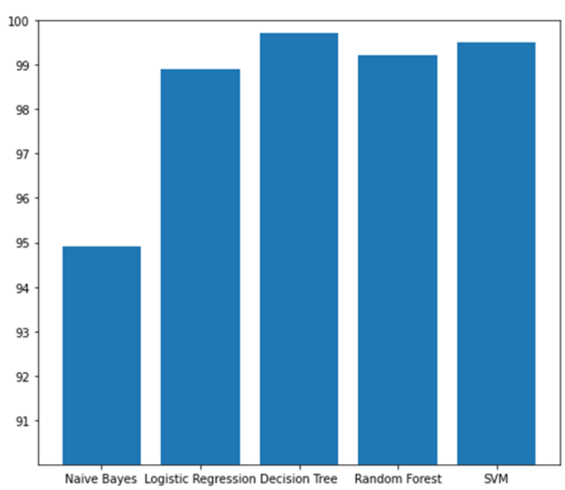

Fig. 5. Percentage analysis of all Machine Learning models

Our study successfully implemented five different classification algorithms to detect fake news. Our dataset consisted of 20,800 news articles, with an equal number of fake and real articles. We trained and tested the classifiers on this dataset using a 70:30 training-test split. We evaluated the performance of each algorithm using the confusion matrix and calculated the accuracy, precision, recall, and F1 score for each algorithm.

TABLE I
Accuracy measures for various Machine Learning algorithms

Sl. No	**Results**	
	Classifier Type	***Accuracy***
1	Naive Bayes	94.91 %
2	Logistic Regression	98.91 %
3	Support Vector Machine(SVM)	99.52 %
4	Decision Tree	99.91 %
5	Random Forest	99.22 %

The results showed that the Decision Tree algorithm had the highest accuracy of 99.73 percent, followed closely by the Support Vector Machine (SVM) with 99.52 percent. Random Forest achieved an accuracy of 99.22 percent, while Logistic Regression achieved an accuracy of 98.91 percent. The Naïve Bayes algorithm had the lowest accuracy at 94.91 percent. These results indicate that Decision Tree and SVM are highly effective algorithms for detecting fake news.

The confusion matrix helped us to understand the performance of each algorithm in more detail. For example, Decision Tree had 166 false negatives and 49 false positives, indicating that it correctly classified most of the real news articles but missed some of the fake news articles. On the other hand, SVM had 155 false negatives and 80 false positives, indicating that it correctly classified most of the fake news articles but missed some of the real news articles. Random Forest had 188 false negatives and 16 false positives, indicating that it correctly classified most of the real news articles but missed some of the fake news articles. Logistic Regression had 212 false negatives and 21 false positives, indicating that it correctly classified most of the real news articles but missed some of the fake news articles. Naïve Bayes had 838 false negatives and 3 false positives, indicating that it struggled to classify both real and fake news articles accurately.

Overall, our study highlights the effectiveness of using machine learning algorithms for fake news detection. The high accuracy achieved by Decision Tree and SVM indicates that these algorithms can be used to build reliable and efficient fake news detection systems. However, the limitations of the Naïve Bayes algorithm suggest that not all machine learning algorithms are equally effective in this task. Further research can explore the use of ensemble methods and deep learning algorithms for fake news detection to improve the accuracy and efficiency of the detection systems.

B. Confusion Matrices

The Confusion Matrix typically consists of four main categories:

1. True Positives (TP): These are the cases where the model correctly identifies a piece of news as fake. It indicates that

the system has successfully detected and classified the news article as intended, aligning with the actual ground truth.

2. True Negatives (TN): These cases represent the instances where the model correctly identifies a news article as genuine or non-fake. It signifies that the system has accurately recognized authentic news and distinguished it from fake news.

3. False Positives (FP): Also known as Type I errors, false positives occur when the model incorrectly classifies a genuine news article as fake. This suggests that the system has made an erroneous prediction, potentially flagging legitimate news as deceptive or misleading.

4. False Negatives (FN): False negatives, or Type II errors, happen when the model fails to detect fake news and mistakenly classifies it as genuine. It implies that the system has missed identifying deceptive content, allowing it to circulate as authentic information.

Below are the Confusion Matrices for the five machine-learning algorithms evaluated:

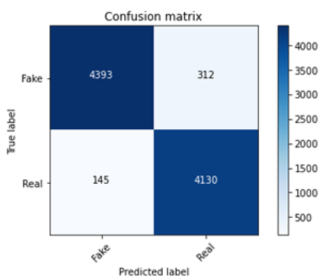

Fig. 6. Naive Bayes Confusion Matrix

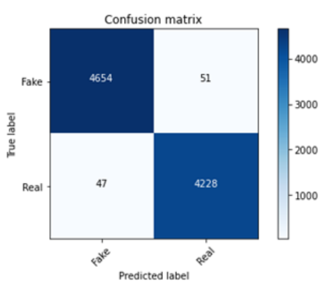

Fig. 7. Logistic Regression Confusion Matrix

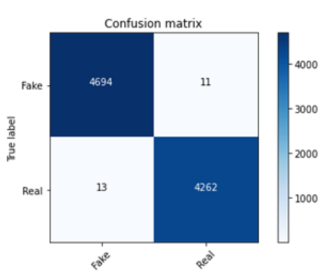

Fig. 8. SVM Confusion Matrix

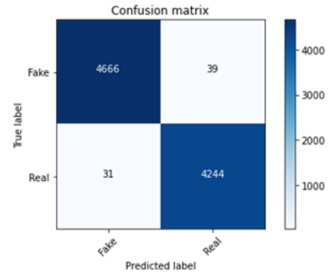

Fig. 9. Decision Tree Confusion Matrix

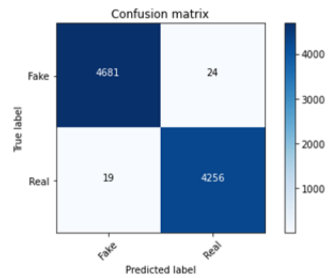

Fig. 10. Random Forest Confusion Matrix

V. Conclusion

The task of classifying news manually requires in-depth knowledge of the domain and expertise to identify anomalies in the text. In this research, we discussed the problem of classifying fake news articles using machine learning models and ensemble techniques. The data we used in our work is collected from the KAGGLE and contains news articles from various domains to cover most of the news rather than specifically classifying political news. The primary aim of the research is to identify patterns in text that differentiate fake articles from true news.

The learning models were trained and parameter-tuned to obtain optimal accuracy. Some models have achieved comparatively higher accuracy than others. We used multiple performance metrics to compare the results for each algorithm. The ensemble learners have shown an overall better score on all performance metrics as compared to the individual learners.

Fake news detection has many open issues that require the attention of researchers. For instance, in order to reduce the spread of fake news, identifying key elements involved in the spread of news is an important step. Graph theory and machine learning techniques can be employed to identify the key sources involved in the spread of fake news. Likewise, real-time fake news identification in videos can be another possible future direction.

Finally, this application is the only one that would be necessary in a larger toolbox that could function as a highly accurate fake news classifier. Other tools that would need to be built may include a fact detector and a stance detector. In order to combine all of these "routines," there would need to be some type of model that combines all of the tools and learns how to weight each of them in its final decision [15].

REFERENCES

[1] M. Zobaer Hossain, M. Ashraful Rahman, M. Saiful Islam, and S. Kar, "Banfakenews: A dataset for detecting fake news in bangla," *arXiv e-prints*, pp. arXiv–2004, 2020.

[2] N. Revathy, P. Hemashree, and T. Vigneshwaran, "Meme analysis through visit using filters fakers," *International Journal of Advances in Engineering & Technology*, vol. 10, no. 2, p. 192, 2017.

[3] R. Bekesh, L. Chyrun, P. Kravets, A. Demchuk, Y. Matseliukh, T. Batiuk, I. Peleshchak, R. Bigun, and I. Maiba, "Structural modeling of technical text analysis and synthesis processes." in *COLINS*, 2020, pp. 562–589.

[4] M. Javad Shafiee, P. Siva, P. Fieguth, and A. Wong, "Domain adaptation and transfer learning in stochasticnets," *arXiv e-prints*, pp. arXiv–1512, 2015.

[5] M. Gulzar Hussain, M. Rashidul Hasan, M. Rahman, J. Protim, and S. Al Hasan, "Detection of bangla fake news using mnb and svm classifier," *arXiv e-prints*, pp. arXiv–2005, 2020.

[6] T. Babu, T. Singh, D. Gupta, and S. Hameed, "Optimized cancer detection on various magnified histopathological colon imagesbased on dwt features and fcm clustering," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 30, no. 1, pp. 1–17, 2022.

[7] R. Haarika, T. Babu, and R. R. Nair, "Insect classification framework based on a novel fusion of high-level and shallow features," *Procedia Computer Science*, vol. 218, pp. 338–347, 2023, international Conference on Machine Learning and Data Engineering. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1877050923000169

[8] M. Horta Ribeiro, P. H. Calais, V. A. Almeida, and W. Meira Jr, "" everything i disagree with is# fakenews": Correlating political polarization and spread of misinformation," *arXiv e-prints*, pp. arXiv–1706, 2017.

[9] S. H. Kong, L. M. Tan, K. H. Gan, and N. H. Samsudin, "Fake news detection using deep learning," in *2020 IEEE 10th Symposium on Computer Applications Industrial Electronics (ISCAIE)*, 2020, pp. 102–107.

[10] J. Sekajugo, G. R. Kagoro, L. Jacobs, C. Kabaseke, E. Namara, O. Dewitte, and M. Kervyn, "Accuracy and completeness of a near real-time citizen science-based multi-disaster inventory in the rwenzori mountains, uganda," in *EGU General Assembly Conference Abstracts*, 2021, pp. EGU21–14 282.

[11] D. F. e. a. Umair M, Zubair M, "A multi-layer holistic approach for cursive text recognition," *Applied Sciences 2022*, vol. 12, no. 12652, 2022.

[12] R. K. Kaliyar, K. Fitwe, R. P, and A. Goswami, "Classification of hoax/non-hoax news articles on social media using an effective deep neural network," in *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, 2021, pp. 935–941.

[13] N. Hoy and T. Koulouri, "A systematic review on the detection of fake news articles," 2021.

[14] C. Chappuis, V. Zermatten, S. Lobry, B. Le Saux, and D. Tuia, "Prompt-rsvqa: Prompting visual context to a language model for remote sensing visual question answering," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2022, pp. 1372–1381.

[15] D. D. Srinivasa Rao, N. Rajasekhar, D. Sowmya, D. Archana, T. Hareesha, and S. Sravya, *Fake News Detection Using Machine Learning Technique*, 01 2021.