Navigating News Authenticity

Prem Uppoor
Computer Science and Technology(AI)
Manipal Institute of Technology, Bengaluru

Khush Suvarna
Computer Science and Technology(AI)
Manipal Institute of Technology, Bengaluru

Sarthak Santra
Computer Science and Technology(AI)
Manipal Institute of Technology, Bengaluru

Krishna Khandelwal
Computer Science and Technology(AI)
Manipal Institute of Technology, Bengaluru

Aaryan Chetan
Computer Science and Technology(AI)
Manipal Institute of Technology, Bengaluru

***Abstract*—The proliferation of fake news has become a significant challenge in the digital age, affecting public opinion and societal trust. This literature review examines the current state of research on fake news detection technologies, highlighting key methodologies, advancements, and challenges in the field. It synthesizes studies on various approaches including machine learning algorithms, natural language processing techniques, and network analysis methods used to identify and mitigate the spread of misinformation. The review also addresses the limitations of existing systems, such as issues with data bias, scalability, and the evolving tactics employed by misinformation creators. By evaluating recent developments and identifying gaps in the current research, this review aims to provide a comprehensive overview of effective strategies for fake news detection and to suggest directions for future research in enhancing the accuracy and reliability of these systems.**

Keywords: Fake news, Misinformation, Fake news detection, Machine learning algorithms, Natural language processing, Network analysis, Data bias, Scalability, Misinformation creators , Accuracy, Reliability, Advancements, Challenges, Public opinion, Societal trust, Future research

I. Introduction

In the digital age, the spread of information is faster and more widespread than ever before. This unprecedented access to information has brought about numerous benefits, but it has also given rise to significant challenges, one of the most pressing being the proliferation of fake news. Fake news refers to false or misleading information presented as news, often created with the intent to deceive or manipulate public opinion. One research suggests that health-related fake news is more common only during a health crisis, whereas fake news related to religion and politics seems more prevalent, emerging from online media. Text & photo and text & video have three-fourths of the total share of fake news, and most of them are from online media: online media is the main source of fake news on social media as well. On the other hand, mainstream media mostly produces political fake news **[6].** During the covid-19 crisis, the spread of the Corona Virus Disease (Covid-19) across the globe had already left the world in a grim situation. A humongous number of fake news stories were being shared through different social media platforms during the wake of the pandemic **[7].** While the Internet and multimedia applications have made it easy to produce and spread media, they have also made it possible to distribute fake news to masses. With over 200 million active users in India and growing, WhatsApp's reach and features make it a top choice to spread fake news. This not only influences public opinion in India but has also sometimes created panic and incited to violence **[8]**. Fake news can take many forms, including fabricated stories, misleading headlines, and manipulated images or videos. Its effects can be far-reaching, impacting everything from individual beliefs to societal norms and democratic processes. Addressing fake news requires critical thinking skills, media literacy, and a concerted effort from both individuals and institutions to verify information and promote truthfulness in media. In this research, we focus on addressing the challenge of detecting fake news in a multilingual environment, specifically focusing on the Indian and global news landscape. By leveraging advanced natural language processing (NLP) techniques and machine learning models such as RoBERTa, Support Vector Machines (SVM), and Logistic Regression, we aim to develop a robust fake news detection system. Our approach employs ensemble learning, stacking the predictions from these models to improve accuracy and generalizability across different datasets. The goal of this paper is to highlight the efficacy of AI-driven models in combating misinformation and provide a scalable solution for news authenticity validation in multilingual settings. Through this project, we aim to enhance the tools available to both consumers and platforms in identifying and curbing the spread of fake news.

II. Literature Survey

Satwinder Singh, Anant Kumar, et al **[1]** described the ROC (receiver operating characteristic curve) stating how efficient the model can distinguish between fake and real news. Area

under ROC curve is degree of usefulness of a test in general greater the area more useful the test. In His work the model created a linear graph with area under the curve giving decent performance. The area under the curve will improve as more meaningful and structured data is available. Since data collections from reliable source is difficult and time-consuming task it limits the performance. Shiza Maham, Abdullah Tariq, et al **[2]** used ANN (Adversarial News Net) model framework which enhances fake news detection by utilizing adversarial training for increased model robustness. Emoticons are extracted to understand their significance in fake news, and this information improves model performance. This model was evaluated on four public datasets, ANN outperforms baseline methods, showing a 2.1% accuracy improvement over Random Forest and 2.4% over BERT after adversarial training. Tashko Pavlov, Georgina Mirceva **[3]** Their approach involved fine-tuning pre-trained BERT and RoBERTa models on real and fake news. By applying these models to tweet data, it demonstrates improved accuracy, recall, and F1 score compared to previous research using fine-tuned BERT models. The results shows that both models effectively detect fake news, but the fine-tuned BERT model outperformed the RoBERTa model in accuracy. Uma Sharma, Sidarth Saran et al **[4]** used a system consisting of three steps: static machine learning classifier for fake news detection, a dynamic component that checks the truth probability of user-provided text, and a URL authenticity checker against trusted and blacklisted sites. The implementation used Count Vectors and Tf-Idf vectors at both word and Ngram levels. Accuracy was recorded for all models, and K-fold cross-validation was employed to enhance model effectiveness. The best model, Logistic Regression, initially achieved 65% accuracy. After grid search parameter optimization, its accuracy improved to 80%. Therefore, the model can classify news articles or headlines with 80% accuracy. Noshin Nirvana Prachi, Md. Habibullah, et al **[5]** used various machine learning, deep learning, and NLP techniques, including logistic regression, decision tree, naive bayes, SVM, LSTM, and BERT. Feature vectors were generated using regex, tokenization, stop words, lemmatization, and TF-IDF. Model performance was evaluated through accuracy, precision, recall, F1 score, and ROC curve. Logistic regression, decision tree, naive bayes, and SVM achieved accuracies of 73.75%, 89.66%, 74.19%, and 76.65%, respectively, while LSTM reached 95%, and BERT obtained the highest accuracy of 98%.

III. Methodology

Data Collection and Integration: Collect datasets from multiple sources such as Fake News Dataset, LIAR Dataset and IFND Dataset and merging them into a single dataset.

Text Pre-processing:We did unified pre- processing which includes converting text to lower case and removing punctuations from the sentences using NLTK module. Then for training our model we used Roberta, Logistic Regression and Support Vector Machine (SVM). So, for that we did model specific pre-processing which includes:

1) Tokenization using RobertaTokenizer from transformers module for Roberta model.
2) Removing stop words and vectorizing the sentences using TF-IDF from Sci-Kit Learn module for SVM / Logistic Regression model.

Model Training : Using the pre-trained RoBERTa model, Logistic Regression and Support Vector Machine (SVM) we fit the pre-processed dataset to all these models and fine tune it by changing the hyperparameters, evaluating the model using validation dataset and getting the predictions from the models as outcomes.

Ensembling The Model : Using a meta-learner (such as Logistic Regression) to combine predictions from the multiple models by using the Stacking Method (Generate predictions from base models and using predictions as features to train the meta-learner) and fine tuning it.

.

Model Testing : Evaluating the final model on the test dataset and Determining performance metrics such as Accuracy, F1 score, precision, recall, etc. and finally deploying it.

IV. Discussion

In this study, three individual models—RoBERTa, Support Vector Machine (SVM), and Logistic Regression—were initially evaluated for fake news detection. RoBERTa achieved the highest individual performance with an accuracy of 94%, exhibiting balanced precision, recall, and F1-scores. Both SVM and Logistic Regression performed similarly, with 90% accuracy each, but had lower recall for real news (class "0") compared to fake news (class "1").

To further enhance performance, a stacking ensemble of these three models was implemented. This ensemble approach combined the strengths of each model, resulting in a significant boost in accuracy to 95%. The stacking method allowed for improved generalization by leveraging each model's strengths, which helped mitigate the weaknesses observed in the individual models.

V. Result of discussion

The stacking ensemble model outperformed all individual models, with an accuracy of 95%, demonstrating the effective-

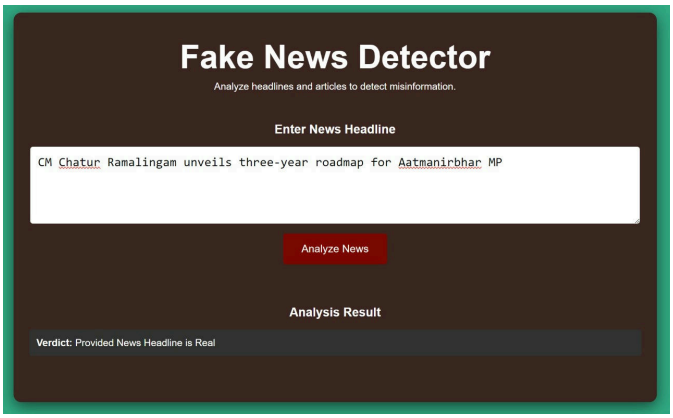

Fig. 1: Example of Real News

Fig. 2: Example of Fake News

ness of combining multiple classifiers to handle complex tasks like fake news detection. The ensemble showed an improvement in both precision and recall for both classes, especially in minimizing false negatives for real news. The RoBERTa model provided strong baseline performance, while SVM and Logistic Regression contributed valuable complementary patterns that were effectively captured in the ensemble, leading to enhanced predictive accuracy and robustness.

VI. Conclusion

The stacking ensemble method, combining RoBERTa, SVM, and Logistic Regression, proved to be the most effective approach, achieving a final accuracy of 95%. This result highlights the power of ensemble methods in complex classification tasks where no single model perfectly captures all patterns. Future work could involve experimenting with additional models in the ensemble or optimizing the stacking method to further enhance the model's performance in real-world applications of fake news detection.

VII. References:

[1] Kumar, Anant & Singh, Satwinder & Gurpreet, Mrs. (2019). Fake News Detection of Indian and United States Election Data using Machine Learning Algorithm. International Journal of Innovative Technology and Exploring Engineering. 8. 1559-1563. 10.35940/ijitee.K1829.0981119.

[2] Maham, S., Tariq, A., Khan, M.U.G. et al. ANN: adversarial news net for robust fake news classification. Sci Rep 14, 7897 (2024). https://doi.org/10.1038/s41598-024-56567-4

[3] T. Pavlov and G. Mirceva, "COVID-19 Fake News Detection by Using BERT and RoBERTa models," 2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO), Opatija, Croatia, 2022, pp. 312-316, Doi: 10.23919/MIPRO55190.2022.9803414.

[4] Sharma, Uma, Sidarth Saran, and Shankar M. Patil. "Fake news detection using machine learning algorithms." International Journal of creative research thoughts (IJCRT) 8.6 (2020): 509-518.

[5] Prachi, Noshin Nirvana, et al. "Detection of Fake News Using Machine Learning and Natural Language Processing Algorithms [J]." Journal of Advances in Information Technology 13.6 (2022).

[6] Md. S. Al-Zaman, "Social Media Fake News in India," Asian Journal for Public Opinion Research, vol. 9, no. 1, pp. 25–47, Feb. 2021.

[7] Raj, A. and Goswami, M.P., 2020. Is fake news spreading more rapidly than COVID-19 in India. Journal of Content, Community and Communication, 11(10), pp.208-220.

[8] Farooq, G. (2018). Politics of Fake News: How WhatsApp Became a Potent Propaganda Tool in India. Media Watch, 9(1), 106-117. https://doi.org/10.15655/mw_2018_v9i1_49279