Advanced Fake News Detection Using Hybrid CNN-BiLSTM with Class Balancing

Siddenthi Uday Kumar
Information Technology
IIIT Allahabad
Prayagraj, India
siddenthiuday@gmail.com

Kushal Bansal
Information Technology
IIIT Allahabad
Prayagraj, India
siddenthiuday@gmail.com

Aditya Khairnar
Information Technology
IIIT Allahabad
Prayagraj, India
siddenthiuday@gmail.com

Jatavath Siddhartha Nayak
Information Technology
IIIT Allahabad
Prayagraj, India
siddenthiuday@gmail.com

Sushmitha Pothuraju
Information Technology
IIIT Allahabad
Prayagraj, India
siddenthiuday@gmail.com

***Abstract*—The spread of fake news in today's digital age has significant social and political consequences. This report presents the "Fake News Detection Using Machine Learning" project, which leverages artificial intelligence (AI) and machine learning (ML) techniques to create a system capable of classifying news articles as real or fake. The project builds on extensive literature review, exploring traditional and modern methodologies, including neural network architectures such as CNN, LSTM, and BiLSTM. Initially, experiments with a CNN + LSTM model achieved low accuracy (30%). Improvements were made by transitioning to a BiLSTM model, applying data preprocessing, and addressing class imbalance using SMOTE. With the adoption of a new dataset, the model's accuracy reached 98%, demonstrating significant progress. This report highlights the importance of model choice, preprocessing, and class balancing in achieving effective fake news detection, with future work focusing on model generalization and further optimization.**

***Index Terms*—component, formatting, style, styling, insert**

I. Introduction

In today's digital age, the spread of fake news has become a significant issue, influencing public trust, decision-making, and even societal norms. With the vast amount of information shared online, it has become increasingly challenging to distinguish between real and fake news. Fake news articles often closely mimic legitimate reporting, using persuasive language and realistic formatting to mislead readers, making simple detection methods ineffective.

This project aims to develop a reliable system that can classify news articles as either fake or real, addressing the growing concerns around misinformation. Traditional methods for detecting fake news often fall short due to the complexity and subtle nature of fake news content. To overcome these challenges, we employ advanced machine learning techniques that focus on understanding the text rather than relying solely on superficial features.

The system leverages pre-trained GloVe word embeddings to effectively capture the context and meaning of words in news articles. For the classification process, we utilize a hybrid model that combines Convolutional Neural Networks (CNNs) to extract meaningful patterns from the text and Bidirectional Long Short-Term Memory (BiLSTM) networks to analyze the sequential flow of words. This combination enables the system to handle the complexities of natural language while maintaining a high degree of accuracy.

To ensure simplicity and scalability, the model processes raw text data without relying on additional metadata, such as the credibility of the source or external features like publication date. Instead, the focus is on the linguistic and contextual patterns within the news articles themselves.

The results of this project contribute to the ongoing fight against misinformation by providing a tool that can identify fake news with high accuracy. By empowering users to access reliable information and reducing the influence of false narratives, this system supports informed decision-making and fosters a more trustworthy digital environment. Ultimately, this work represents a step forward in addressing the challenges posed by the spread of fake news in the modern world.

II. Literature Review

The foundation of the "Fake News Detection" project is built upon a thorough review of existing literature in the fields of artificial intelligence and machine learning, particularly focusing on their applications in identifying misinforma- tion

A. Standard ML techniques for fake news detection

Initially, we studied a research paper"FAKE NEWS DETECTION USING ML." [**?**] that detailed various data preprocessing techniques and classifiers used in fake news detection. Key methodologies explored including TF-IDF (Term Frequency-Inverse Document Frequency), Multinomial Naive Bayes (NB), Passive Aggressive Classifier, and Support Vector Classifier (SVC).

These methods primarily emphasize non-deep learning approaches, providing a baseline for traditional techniques in this domain. Additionally, we familiarized ourselves with performance metrics such as accuracy, precision, recall, and F1 score, essential for evaluating model effectiveness.

B. Survey Paper Analysis

To further enhance our understanding, we analyzed another research paper ”Fake news detection on social media: A data mining perspective.” [**?**] that examined a broader array of neural network methods, including Feedforward Neural Networks (FNN), Graph Neural Networks (GNN), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN) such as LSTM [**?**], BiLSTM, and GRU, alongside BERT. This literature highlighted the importance of transfer learning techniques in improving model accuracy and addressed the challenges faced in implementing these approaches within the context of fake news detection.

C. Class Balancing Problem

Moreover, we recognized the critical issue of class imbalance in datasets used for fake news detection. Various sampling methods, such as oversampling, undersampling, and SMOTE (Synthetic Minority Over-sampling Technique), were explored to tackle this imbalance effectively.

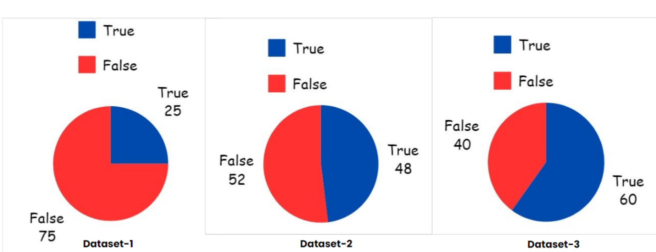
III. Methodology

A. Literature Review

We conducted a comprehensive review of existing literature, focusing on data preprocessing techniques and classifiers applicable to fake news detection. This included traditional methods such as TF-IDF, Multinomial Naive Bayes, and Support Vector Classifier, as well as modern neural network architectures like CNN, RNN, and BERT. We also learned about transfer learning techniques, which can enhance model performance by leveraging pre-trained models for improved accuracy in fake news detection.

B. Dataset Exploration

We started by using a publicly available Kaggle dataset, which we fed into an initial CNN + LSTM model. This model produced an accuracy of approximately 40%, indicating room for improvement in our approach.



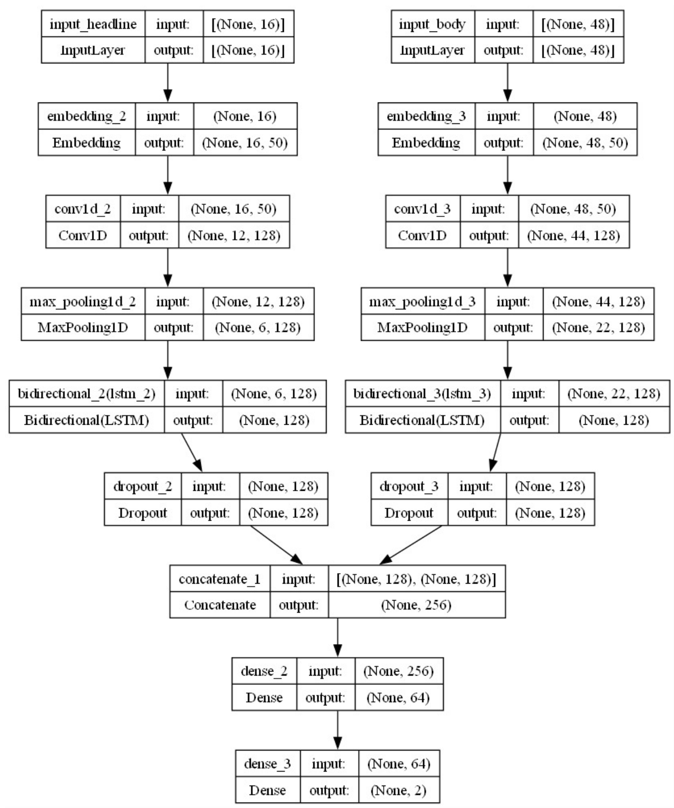
C. Class Imbalance Addressing

After analyzing the dataset, we discovered a significant class imbalance issue. To mitigate this, we applied the SMOTE (Synthetic Minority Over-sampling Technique) algorithm, which allowed us to oversample the minority class effectively. However, we later shifted to a new dataset where the class imbalance was less severe, enabling more accurate model training.

D. Model Improvement

To enhance model performance, we moved to a hybrid model combining BiLSTM and CNN. This approach improved the accuracy to 90%, demonstrating substantial progress. We also applied transfer learning techniques, using pre-trained models such as BERT, DistilBERT, and XLNet, which boosted the accuracy further to 96%.

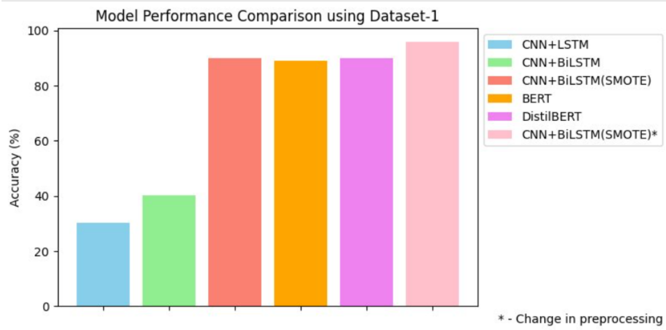
E. Model Architecture



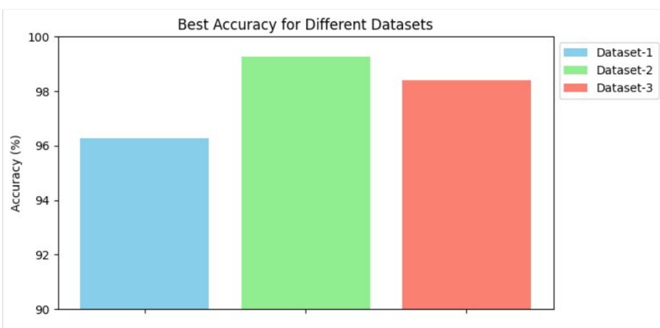
IV. Results and Discussions

A. Model Comparison

The models used include CNN, CNN + BiLSTM, BERT, DistilBERT, and CNN + BiLSTM (SMOTE). The CNN + BiLSTM (SMOTE) model outperformed others in terms of accuracy and generalization across datasets. A significant improvement was observed when combining SMOTE with CNN + BiLSTM, particularly on imbalanced datasets, which led to better performance metrics.



Datasets	Model	Accuracy	Precision	Recall	F1-Score
Dataset-1	CNN + LSTM	30.18%	29.19%	31.20%	30.10%
	CNN + BILSTM	40.16%	38.80%	32.00%	35.00%
	CNN+BILSTM+SMOTE	90.14%	95.08%	91.54%	93.28%
	BERT	89.23%	91.27%	88.41%	89.82%
	DistilBERT	90.15%	95.08%	91.54%	93.28%
	CNN+BILSTM+SMOTE (changes in pre-processing)	96.36%	95.69%	97.34%	96.13%
Dataset-2	CNN+BILSTM	99.27%	99.15%	99.46%	99.30%
Dataset-3	CNN+BILSTM+SMOTE (changes in pre-processing)	98.40%	98.22	98.60%	98.40%



B. Performance Metrics

For Dataset 1, the model achieved an accuracy of 96.36%, with significant improvements in precision, recall, and F1-score due to preprocessing and SMOTE. For Dataset 2, the model achieved an accuracy of 99.27%, while Dataset 3 showed an accuracy of 98.4%. The F1-score consistently demonstrated that models performed well on both real and fake news classifications.

C. Data Preprocessing Insights

Preprocessing steps such as tokenization, stopword removal, and handling of special characters played a crucial role. A separate treatment for headlines and body text improved feature extraction and classification accuracy. SMOTE was highly effective in balancing the class distribution and mitigating bias toward the majority class.

D. Challenges

Data imbalance made it difficult for models to generalize well across different classes, which was addressed with SMOTE and careful training data preparation. Extensive text preprocessing was needed to extract meaningful features for effective classification.

E. Model Insights

The hybrid CNN + BiLSTM (with SMOTE) leveraged both spatial features through CNN and temporal dependencies through BiLSTM for better performance. BERT and DistilBERT, despite their pre-trained nature, slightly lagged behind the hybrid model due to dataset-specific fine-tuning limitations.

V. Summary and Conclusions

A. Key Findings

The hybrid model CNN + BiLSTM with SMOTE and improved preprocessing emerged as the best model, achieving 96.36% accuracy, 95.69% precision, and 97.34% recall, with an F1-score of 96.13%. Data preprocessing and class balancing were crucial in achieving high accuracy and mitigating overfitting or bias.

B. Contributions

This project highlights the effectiveness of hybrid models and the importance of addressing data imbalance in fake news detection. The insights can be applied to other text classification tasks where class imbalance is a significant concern.

C. Future Scope

The future scope of this project includes real-time testing on live news streams to assess performance in dynamic environments. Integrating APIs will enable continuous news data retrieval. The model will also incorporate continuous learning to adapt to evolving fake news patterns. Finally, deployment in a scalable, low-latency environment will support efficient real-time use.

D. Conclusion

This study confirms that combining traditional machine learning techniques with modern architectures and proper data handling leads to significant improvements in fake news detection. Addressing challenges like data imbalance and leveraging hybrid architectures like CNN + BiLSTM ensures a more robust and accurate classification model.

VI. Appendix

A. Data Resources

- **Kaggle Dataset**:The initial dataset used for the project was sourced from Kaggle, comprising labeled news articles for classification as fake or real.
- **Additional Datasets**: Subsequent experiments involved other publicly available datasets that provided a broader range of news articles to test the model's effectiveness.

B. Data Preprocessing Steps

- **Text Cleaning**:Removal of special characters, numbers, and punctuation to ensure uniform and clean text data.
- **Tokenization**: Splitting text into individual words or tokens for further analysis.
- **Vectorization**: Utilizing TF-IDF to convert textual data into numerical form for processing by machine learning models.

Model Architectures

- **CNN + LSTM**:The initial model combining Convolutional Neural Networks (CNN) with Long Short-Term Memory (LSTM) networks to capture both spatial and temporal features.
- **BiLSTM**: The upgraded model using Bidirectional LSTM to leverage context from both directions in the text data.

C. Performance Metrics

- **Accuracy**:The proportion of correctly predicted instances among the total instances.
- **Precision**: The ratio of true positive predictions to the total predicted positives.
- **Precision**:The ratio of true positive predictions to the total actual positives.
- **F1 Score**:The harmonic mean of precision and recall, providing a single metric for model performance.

D. Class Imbalance Techniques

- **SMOTE (Synthetic Minority Over-sampling Technique):**: A method used to create synthetic samples of the minority class to balance the dataset.

REFERENCES

[1] Agrawal, Srishti, Vaishali Arora, Ruchika Arora, Pronika Chawla, and Madhumita Kathuria. ”FAKE NEWS DETECTION USING ML.” International Research Journal of Engineering and Technology (IRJET)., V olume: 0 7 (2020).

[2] Shu, Kai, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. ”Fake news detection on social media: A data mining perspective.” ACM SIGKDD explorations newsletter 19, no. 1 (2017): 22-36.

[3] Alnabhan, Mohammad Q., and Paula Branco. ”Fake News Detection Using Deep Learning: A Systematic Literature Review.” IEEE Access (2024).

[4] Gifu, Daniela. ”An intelligent system for detecting fake news.” Procedia Computer Science 221 (2023): 1058-1065.

[5] Alnabhan, Mohammad Q., and Paula Branco. ”Fake News Detection Using Deep Learning: A Systematic Literature Review.” IEEE Access (2024).

[6] Pal, Aishika, and Moumita Pradhan. ”Survey of fake news detection using machine intelligence approach.” Data & Knowledge Engineering 144 (2023): 102118.

[7] Asish, K. Raghavendra, Adarsh Gupta, Arpit Kumar, Alex Mason, Murali Krishna Enduri, and Satish Anamalamudi. ”A tool for fake news detection using machine learning techniques.” In 2022 2nd International Conference on Intelligent Technologies (CONIT), pp. 1-6. IEEE, 2022.

[8] Aljabri, Malak, Dorieh M. Alomari, and Menna Aboulnour. ”Fake news detection using machine learning models.” In 2022 14th International Conference on Computational Intelligence and Communication Networks (CICN), pp. 473-477. IEEE, 2022.

[Dataset-1] Nayak, Prabhanjan. ”Fake News Challenge.” Kaggle. April 10, 2022.

[Dataset-2] Emine, Bozkus. ”Fake News Detection Datasets.” Kaggle. May 23, 2022.

[Dataset-3] clmentbisaillon. ”fake-and-real-news-dataset.” Kaggle. April 19, 2024.