A Hybrid Machine Learning and Deep Learning Framework for Fake News Detection Using Sentiment Analysis

Deepshikha Patel, Sarvesh Singh Rai

Department of Computer Science, Infinity Management and Engineering College Sagar, MP

Abstract

The rapid growth of social media platforms has significantly increased the spread of fake news, creating serious challenges for information credibility and public trust. Automated fake news detection has therefore become an important research area in data science and management studies. This paper presents a hybrid machine learning and deep learning framework for fake news detection using sentiment analysis and textual features. Traditional machine learning classifiers such as Naïve Bayes and Logistic Regression are combined with deep learning models including Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), and Bidirectional LSTM (Bi-LSTM). Sentiment polarity is extracted using the VADER sentiment lexicon, while TF-IDF and word embedding techniques are employed for feature representation. Experimental evaluation is conducted on a Twitter-based fake news dataset obtained from Kaggle. The results show that Logistic Regression achieved an accuracy of 91.16%, while the proposed CNN–BiLSTM hybrid model achieved a testing accuracy of 89.94%, demonstrating improved robustness in identifying fake news. The findings indicate that integrating sentiment features with deep learning architectures provides an effective solution for combating misinformation on social media platforms.

Keywords: Fake News Detection, Sentiment Analysis, Machine Learning, Deep Learning, CNN, Bi-LSTM

1. Introduction

The widespread adoption of social media has transformed how information is produced, shared, and consumed. While these platforms enable rapid information dissemination, they have also become a major source of fake news and misleading content. Fake news can influence public opinion, damage organizational reputation, and negatively affect decision-making processes in areas such as politics, finance, healthcare, and business management.

Detecting fake news is a complex task due to the unstructured nature of textual data, linguistic ambiguity, and the deliberate imitation of legitimate news styles. Traditional rule-based approaches have shown limited success in handling large-scale and dynamic data environments. Consequently, researchers have increasingly adopted machine learning (ML) and deep learning (DL) techniques to automatically identify deceptive patterns in text.

Sentiment analysis plays a crucial role in fake news detection by capturing emotional and subjective cues embedded in textual content. Opinions, polarity, and emotional tone often differ between genuine and deceptive information. Motivated by this observation, this study proposes a hybrid framework that integrates sentiment-based features with machine learning and deep learning models to improve fake news detection accuracy.

Early research on fake news detection primarily relied on statistical and rule-based techniques. With the advancement of computational power, machine learning approaches such as Naïve Bayes, Support Vector Machines, and Logistic Regression became popular due to their simplicity and efficiency. These models typically employ text representation techniques such as bag-of-words and TF-IDF.

More recently, deep learning models have gained attention for their ability to automatically learn hierarchical and contextual features from textual data. CNN-based models are effective in capturing local patterns, while LSTM and Bi-LSTM networks are well-suited for modeling sequential dependencies in text. Several studies report improved performance when combining CNN and LSTM architectures for fake news classification.

Hybrid approaches that integrate sentiment features with textual representations have shown promising results. Sentiment lexicons such as VADER provide polarity scores that enhance the discriminatory power of classifiers. However, challenges such as model interpretability, computational complexity, and generalization across datasets remain open research issues. This study contributes to the literature by proposing a sentiment-enhanced hybrid CNN–BiLSTM model for fake news detection.

2. Research Methodology

2.1 Dataset Description

The dataset used in this study is obtained from Kaggle and consists of Twitter-based news content labeled as real or fake. Each record includes attributes such as title, text, and class label. The dataset provides a realistic representation of social media misinformation.

2.2 Data Preprocessing

Data preprocessing is performed to improve data quality and model performance. The preprocessing steps include: - Tokenization of text into individual words - Removal of stop words - Conversion to lowercase - Lemmatization to reduce words to their base forms

These steps help reduce noise and dimensionality while preserving semantic meaning.

2.3 Feature Extraction

Two types of features are extracted: - **Textual Features:** TF-IDF is used for machine learning models, while Word2Vec embeddings are used for deep learning models. - **Sentiment Features:** VADER sentiment analysis is applied to extract positive, negative, neutral, and compound sentiment scores.

2.4 Classification Models

The following models are implemented and evaluated: - Naïve Bayes - Logistic Regression - Convolutional Neural Network (CNN) - Long Short-Term Memory (LSTM) - Hybrid CNN–BiLSTM model

The hybrid model combines convolutional layers for feature extraction with Bi-LSTM layers to capture contextual and sequential information.

3. Experimental Results and Discussion

The performance of the models is evaluated using accuracy, precision, recall, F1-score, and AUC. Logistic Regression achieved the highest accuracy among machine learning models, while deep learning models demonstrated better generalization capabilities.

The proposed CNN–BiLSTM model achieved high training accuracy and competitive testing accuracy, indicating its effectiveness in learning complex textual and sentiment patterns. The integration of sentiment features improved classification performance by providing additional emotional context.

The results confirm that hybrid models outperform individual classifiers by leveraging the strengths of both machine learning and deep learning techniques.

4. Conclusion

This paper presented a hybrid machine learning and deep learning framework for fake news detection using sentiment analysis. By combining textual and sentiment-based features with CNN and Bi-LSTM architectures, the proposed model achieved strong performance on a real-world Twitter dataset. The findings demonstrate that sentiment-enhanced deep learning models are effective in detecting fake news and can support decision-making in social media management and information systems.

5. Future Scope

Future research may focus on incorporating transformer-based models such as BERT, extending the framework to multilingual datasets, and integrating explainable AI techniques to improve transparency and trust. Real-time fake news detection systems and cross-platform analysis also present promising research directions.

References

[1] M Nayana, J Deepali, “A Survey on Emotions Generation using Text Mining for Social Networking Websites”, International Journal of Innovative Research in Computer and Communication Engineering, Vol. 6(2), pp. 1554-1560,2018.

[2] S Das, M Chen, " Yahoo! for Amazon: Sentiment Extraction from Small Talk on the Web," Management Science, Vol. 53(9), pp. 1375-1388,2007.

[3] A Tumasjan, T Sprenger, P Sandner, I Welpe, "Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment," In Proceedings of AAAI Conference on Weblogs and Social Media, In Washington D. C.,

pp. 178-185.2010.

[4] B O'Connor, R Balasubramanyan, B Routledge, N Smith, "From Tweets to Polls: Linking Text Sentiment to Public Opinion Time Series," In Proceedings of AAAI Conference on Weblogs and Social Media, In Washington D. C., pp. 122-129,2010.

[5] C Nimai, K Vamshi, S Suhas, K Jitendra, N Ashish, N Sankalp, S Vaisakh, "Sentiment Classifier and Analysis for Epidemic Prediction", In Proceedings of the 7th International Conference on Soft Computing, Artificial Intelligence and Applications, the 7th International Conference on Advanced Information Technologies and Applications, the 5th International Conference on Signal and Image Processing, In Chennai, pp. 31-48.2018.

[6] I Mozetič, L Torgo, V Cerqueira, J Smailović, "How to Evaluate Sentiment Classifiers for Twitter Time-Ordered Data?", Public Library of Science One, Vol. 13 (3) , pp. 345- 365.2018.

[7] A. V Yeole, P V Chavan, M C Nikose, "Opinion Mining for Emotions Determination", In Proceedings of IEEE International Conference on Innovations in Information, Embedded and Communication Systems, In Coimbatore, pp. 1-5,2015.

[8] B Heredia, T M Khoshgoftaar, J Prusa, M. Crawford, "Cross-Domain Sentiment Analysis: An Empirical Investigation", In Proceedings of IEEE 17th International Conference on Information Reuse and Integration, In Pittsburgh, pp. 160-165,2016.

[9] F Luo, C Li, Z Cao, "Affective Feature Based Sentiment Analysis using SVM Classifier" In Proceedings of IEEE 20th International Conference on Computer Supported Cooperative Work in Design, pp. 276-281, In China, 2016.

[10] J Singh, G Singh, R Singh, "A Review of Sentiment Analysis Techniques for Opinionated Web Text", Computer Society of India Transactions on Information and Communication Technology, pp. 241-247,2016.

[11] E Aydogan, M Akcayol, "A Comprehensive Survey for Sentiment Analysis Tasks Using Machine Learning Techniques", In Proceedings of International Symposium on Innovations in Intelligent Systems and Applications, In Romania, pp. 1-7, 2016.

[12] B Pranali, A Dilipkumar, "An Approach to Sentiment Analysis using Artificial Neural Network with Comparative Analysis of Different Techniques", IOSR Journal of Computer Engineering, Vol. 18(2), pp. 64-69, 2016.

[13] Abdullah Ikram, Mohit Kumar, Geetika Munjal - Twitter Sentiment Analysis using Machine Learning,12th International Conference on Cloud Computing, Data Science&Engineering(Confluence)|9781665437011/22/$31.00©2022IEEE|DOI:10.1109/ Confluence52989.2022.9734154,2022.

[14] W. Medhat , Hassan A, Korashy H (2014) Sentiment analysis algorithms and applications: a survey. Ain Shams Eng J. 5:1093–1113,2014.

[15] G. Chowdhary, "Natural Language Processing," Annual Review of Information Science and Technology, vol. 37, no. 1, pp. 51-89 2020.

[16] C. Catal, U. Sevim, and B. Diri, "Practical Development of an Eclipse-Based Software Fault Prediction Tool Using Naive Bayes Algorithm," Expert Systems with Applications, vol. 38, pp. 2347- 2353, 2011

[17] M. Mhatre, D. Phondekar, P. Kadam, A. Chawathe, and K. Ghag, "Dimensionality reduction for sentiment analysis using pre-processing techniques," 2017 International Conference on Computing Methodologies and Communication (ICCMC), pp. 16-21, 2017.

[18] G. Chowdhary, "Natural Language Processing," Annual Review of Information Science and Technology, vol. 37, no. 1, pp. 51-89 2020.

[19] Medhat W, Hassan A, Korashy H "Sentiment analysis algorithms and applications: a survey". Ain Shams Eng J5:1093–1113 2004.

[20] N. F. F. da Silva, E. R. Hruschka, and E. R. Hruschka, ''Tweet sentiment analysis with classifier ensembles,'' Decision Support Syst., vol. 66, pp. 170–179, Oct. 2014.