IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, April 2025
Impact Factor: 7.67

Leveraging Analysis of Sentiment for Fake News Identification in Cybersecurity: A Machine Learning Approach

Mani Gopalsamy
Independent Researcher
manigopalsamy14@gmail.com

Abstract: *A crucial difficulty in the rapidly changing world of digital communication is maintaining information integrity. This study investigates how machine learning methods can increase the accuracy of cybersecurity sentiment analysis classification. News reports of incorrect identification. It can differentiate between authentic and fraudulent news by analyzing sentiment and emotional trends in user interactions and news content. The emotions of surprise, disdain, and fear are often evoked by false news, whereas anticipation and trust are associated with accurate news. Outperforming traditional classifiers, the Bi-LSTM model obtains high accuracy, AUC, and F1 score. The findings show how effective sentiment-based feature integration is in identifying fake news, providing a viable strategy for reducing disinformation. Multimodal data and explainable AI algorithms can be used in future research to improve real-time detection. Consequently, it has the best accuracy of 96.89% in cybersecurity sentiment false news identification.*

Keywords: component, Fake News, Sentiment Analysis, Cybersecurity, Machine Learning, Social Media

I. INTRODUCTION

A major factor in false news is the way feeling is expressed. When social media users see information that they find provocative but over which they feel less in control, they are more likely to leave comments [1]. In contrast, people who feel more in control are more likely to share a message. To demonstrate that sentiment-related activity was adequate to differentiate between social bot and human accounts, Dickerson et al. combined a variety of sentiment metrics [2][3]. Headlines should pique readers' interest and emotionally connect with them in order to improve the dissemination of news. In order to trick readers, publishers deliberately employ combinations of emotional valence or polarity (positive and negative) and arousal (strong and weak), as a significant percentage of the false news audience simply reads the headlines. In order to assess if a news piece is reliable or should be regarded as false news, SA offers vital information about its content [4][5].

Fake news has an impact on people's everyday lives by influencing their ideas, manipulating their thoughts and emotions, and sometimes causing them to make poor choices. In many areas of society, including politics [6][7], the economy, society, health, technology, and sports, False information spreading on social media has a negative effect. According to a 2016 survey, 23% of US citizens unintentionally or intentionally disseminated bogus news. The likelihood of fake news spreading is 70% higher than that of real news, per a poll. Another poll found that many people, regardless of gender, age, or educational attainment, find it difficult to distinguish between fake and true news [8].

Social media is the main platform where false pictures are disseminated. Fake photographs are photos that have been altered to contradict the information they portray. False images circulated on social media cause societal discord and misinformation [9]. False information regarding the pandemic, such as the misleading or notable drop in cases caused by COVID-19, was disseminated by false news, although the second wave was actually wreaking havoc on the world, particularly the US. Despite the fact that rumors might occasionally be true, depending on the source's intentions, false news is always a misinformation effort [10]. These are spread by terrorists, political activists, hired posters, and bots, false information that is extensively disseminated on social media platforms worldwide, including by media

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, April 2025
Impact Factor: 7.67

organizations, individuals, and state-sponsored trolls. Their motivations can range from financial gain to harm and disgrace, disorder, opinion manipulation, or just advancing personal beliefs [11]

A. Contribution of the Study

The main contribution of the work is the combination of sentiment analysis and machine learning approaches to improve the identification of fake news. More specifically, the study helps by:

- **Utilizing the Fakeddit Dataset:** Refines this dataset by removing non-binary labels, duplicates, and incomplete entries, creating a balanced set of 22,788 records (with a majority of real news posts) that preserves the authentic sentiment patterns in the data.
- **Sentiment and Emotion Analysis:** Identifies eight emotions, including fear, rage, and trust, by combining the NRCLex tool with sentiment and emotion analysis. The research divides emotions into "Novelty" and "Expectation," offering important insights into how consumers react emotionally to both bogus and legitimate news.
- **Effective Model Training:** The study guarantees the best possible use of the dataset in training and testing sets in order to train and validate the model. The ability to identify long-term patterns in sequential data, identify contextual patterns in news content, and improve the effectiveness of false news classification are all made possible by the use of Bi-LSTM models.
- **Ensemble Learning Approaches:** Introduces a machine learning ensemble approach, comparing the performance of bagging and boosting techniques for fake news detection. Specifically, it evaluates the effectiveness of CatBoost (a boosting-based algorithm) and other models, such as LR, showcasing their strong performance about categorization reliability and accuracy.
- **Comprehensive Performance Evaluation:** The importance of evaluating the model's performance based on several metrics, including accuracy, F1-Score, and Area Under the Curve, or AUC. These metrics provide a trustworthy assessment of the model's classification accuracy for news articles, particularly when dealing with imbalanced datasets.

B. Structure of the Paper

This is how the rest of the paper is structured. Section II gives a summary of the research on sentiment in misleading news. ML for Cybersecurity Detection. Methods and methodology are presented in Section III, whilst Section IV deals with the analysis and discussion of the results. The study's findings and future directions are presented in Section V.

II. LITERATURE REVIEW

This section presents the prior research on using machine learning techniques to use sentiment analysis for the identification of bogus news in cybersecurity. We conclude from the literature evaluations above that sentiment has been the subject of a great deal of investigation. The detection of fake news is shown in Table I.

Bhutani et al. (2019) Proposed false news has become a major problem that is wreaking havoc all across the world. The creation of the most accurate algorithm will thus be revolutionary and significantly affect both the existing state of politics and prevailing societal concerns. Due to its instant accessibility, affordability, and speed just a click away people turn to social media and online news articles as their main sources of information. To improve the accuracy, they have thus proposed a unique method for identifying bogus news that considers sentiment. It also examines the efficacy of the proposed method using three different data sets. The results demonstrate the effectiveness of the suggested solution. Additionally, comparisons with other approaches used in this study are also provided [12].

Rozi, Arianto and Mahdyan (2023) aim to construct a system that uses sentiment analysis to identify false information. Their goal is to create a fake news detection system for the Indonesian language that uses sentiment analysis, even though earlier studies have successfully used the VADER (Valence Aware Dictionary and Sentiment Reasoner) approach for English, with Random Forest achieving a noteworthy 78% accuracy. It makes news and any other

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, April 2025
Impact Factor: 7.67

information sources easily accessible and shareable. However, this convenience also comes with the risk of encountering false as well as fake information [13].

Pillai et al. (2024) proposed for classifying fake news into actual and fraudulent news based on sentiment analysis. The social media and smartphones popularity have enhanced exponentially. By the electronic media, fake news has rising quick with new information which are hugely untrustworthy. Search engines like Google are incapable of fraudulent news because of their limitations with restricted keywords. The Optimized Convolution Neural Network (OPCNN) is which accommodates varying complexity levels by adjusting its architecture in Fake News Detection (FND). Features are extracted from pre-processed pictures using Principal Component Analysis (PCA), which lowers the data dimension with many linked variables and recalls the significant change in actual data. The ISOT dataset is pre-processed through four various techniques. The recall, accuracy, f1 score and precision with the ISOT dataset are considered to evaluate OPCNN performance [14].

Ritu (2023) provides an effective binary classification method for detecting false news that makes use of attention-based Recurrent Neural Network (RNN) models. Their combined model, with high accuracy and understanding, is rigorously tested on an online dataset and regularly produces high-quality performance measures. Spreading false, offensive, and misleading material under the guise of legitimate news is known as fake news. Its main goal is to damage people's reputations while making money off of ads that include inaccurate information [15].

Matsumoto, Yoshida and Muneyasu (2021) provide. A network-based method for spotting bogus news As Transformer Network (GTN) learns effective node representations, it could discover beneficial connections between nodes in the original network. The dissemination of false information has become a global issue that undermines public confidence. According to recent studies, false news and legitimate news propagate on social media in various ways. Thus, detection methods based on propagation have attracted a lot of attention. These techniques concurrently learn propagation patterns and user preferences by using GNN to build graphs with users as nodes and news-sharing chains as connections. However, it might be difficult to determine the relationships between disconnected nodes when attempting to extract user preferences from the network [16].

Chabukswar, Shenoy and Venugopal (2023) the internet is essential for the spread of information through public networks and websites, which is why fake news is created for political and economic gain in order to mislead and captivate readers. To stop misleading information from spreading, deep learning-based detection techniques are being researched for natural language processing applications. The DL model suggested in this research combines LSTM and Bi-directional LSTM with one-hot encoding representation to classify bogus news. The model is successfully validated by the body, headline, and label of the collection of political news items from 240 websites (URLs). Kera's built-in DL libraries are used by the TensorFlow framework, which has a TensorFlow Kaggle repository community [17].

Table 1: Comparative table for literature review Sentiment for Fake News Identification in Cybersecurity

Reference	Methodology	Dataset	Performance	Limitations & Future Work
Bhutani et al. (2019)	Sentiment-based Fake News Detection: Adding sentiment as an additional component improves classification accuracy	Three distinct datasets derived from internet news items and social media	demonstrates more accuracy in identifying bogus news than conventional methods	Requires further testing with real-time news streams and evaluation with large-scale datasets
Rozi, Arianto and Mahdyan (2023)	Sentiment analysis for detecting false news in Indonesian using the Random Forest classifier in conjunction with VADER (Valence Aware Dictionary and Sentiment Reasoner)	Indonesian Language News Dataset	Achieved 78% accuracy with Random Forest	Needs expansion to multilingual fake news detection and evaluation on larger datasets

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 3, April 2025 Impact Factor: 7.67

Pillai et al. (2024)	Optimized Convolutional Neural Network (OPCNN) with Principal Component Analysis (PCA) for feature extraction	ISOT dataset, processed using four different pre-processing techniques	High recall, accuracy, F1-score, and precision, showing effective deep learning-based classification	Further enhancements needed in real-time fake news detection and generalization across different datasets
Ritu (2023)	Attention-based recurrent neural network (RNN) model for the binary categorisation of misleading data	Internet-sourced datasets from social media platforms	Achieved superior accuracy in fake news classification compared to standard RNN models	Needs evaluation on diverse datasets, including multilingual and multimodal news sources
Matsumoto, Yoshida and Muneyasu (2021)	The Graph Transformer Network (GTN) detects connections between bogus news sources and propagation networks by learning effective node representations.	Social Media Graphs	Successfully extracts propagation patterns and improves detection accuracy by analyzing user interactions.	Struggles with learning relationships between unconnected nodes, requiring improvements in graph-based learning
Chabukswar, Shenoy and Venugopal (2023)	A deep learning method for classifying false news that combines one-hot encoding with LSTM (long short-term memory) + Bi-LSTM (bidirectional LSTM).	A dataset comprising 240 websites, headlines, and categories for political news articles	Achieved strong performance using TensorFlow and Keras deep learning libraries, demonstrating effective classification in political news	Needs to be expanded beyond political domains to general fake news detection, including health-related misinformation and financial fraud

A. Research Gap

The literature on leveraging sentiment analysis for fake news detection in cybersecurity reveals several research gaps. Firstly, there is a need for more cross-linguistic and multi-lingual approaches, as most studies focus on English, with limited exploration of other languages. Furthermore, little research has been done on combining sentiment analysis with multimodal data—like pictures and videos—for the purpose of detecting false news. Advanced feature extraction techniques, real-time detection systems, and explainable AI models also require further investigation, as current methods often rely on black-box models. Moreover, the generalization of models across diverse domains and addressing biases in sentiment analysis tools are crucial areas for improvement, especially in politically sensitive or context-specific content. Filling up these gaps might greatly improve the precision and usefulness of algorithms for detecting false news.

III. METHODOLOGY

Sentiment analysis-based fake news identification involves a systematic process that starts with gathering datasets of both authentic and fraudulent news stories. To standardize inputs, the data is pre-processed using text cleaning methods such stemming, tokenization, and stop-word removal. Oversampling and SMOTE are two balancing strategies used to address class imbalance. After that, the text is converted into numerical form using encoding techniques like TF-IDF or Word2Vec.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, April 2025
Impact Factor: 7.67

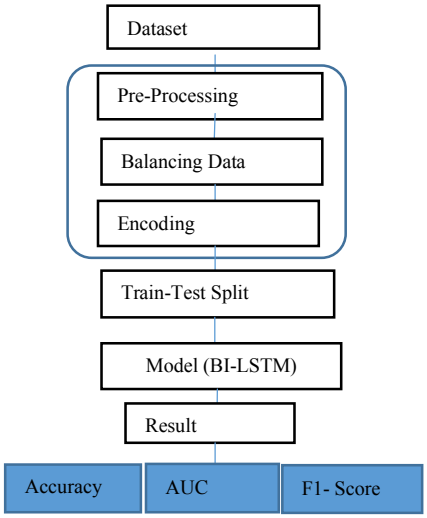

Figure 1: Methodology Flow Diagram for Cyber-Attack Detection

To make sure the model is trained on one portion and assessed on another, after that, Training and testing sets are created from the dataset. Because it can process information both forward and backward, a Bi-LSTM model is employed for categorization, successfully capturing contextual connections in text. After being instructed, the model produces findings that show if a news piece is authentic or fraudulent. The final assessment metrics are used Based on sentiment patterns shown in Figure 1, the model's accuracy, AUC (Area Under the Curve), and F1-Score are used to assess its performance and ensure that it is reliable in detecting false news.

The stages that make up a data flow diagram are explained in detail below:

A. Data Collection

The Fakeddit dataset which comprises text and image components obtained from Reddit social media platform and is available at GitHub as of February 22, 2022. During March 19, 2008 to October 24, 2019 Nakamura and Levy built this collection which comprises Reddit posts. The posts covered multiple domains and were equipped with various features, including image data and user information, reaching over one million entries. Due to the way social media functions the collected data contains multiple errors and incomplete values. The number of comments attached to a single post range from multiple to zero. Researcher teams provided three distinct labels for each post which enables three classification types including binary (two-way), ternary (three-way), and six-category formats.

B. Data Pre-processing

This work uses the Faked It dataset to create a binary classification model for identifying false and authentic news. To enhance data quality, all non-binary labels, duplicate entries, and records missing titles or comments were removed, resulting in a refined dataset of 22,788 records, 8,553 labeled as fake news and 14,235 as real news. Unlike traditional data balancing techniques, the dataset's original distribution was preserved to maintain the authenticity of sentiment and emotion patterns in comments. Sentiment and emotion analysis was conducted using the NRCLex tool, which classifies text into Eight feelings: surprise, grief, pleasure, disgust, wrath, fear, anticipation, and trust. Prior research suggests that fake news comments typically express fear, disgust, and surprise, whereas real news comments are more likely to reflect trust, joy, anticipation, and sadness.

- To facilitate analysis, emotions were grouped into two primary categories: "Novelty" (fear, disgust, surprise) and "Expectation" (trust, joy, anticipation, sadness), while neutral comments were categorized separately. Emotion values were normalized to a scale of 0 to 1 to minimize magnitude dependence and optimize model weight adjustments.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, April 2025
Impact Factor: 7.67

- The findings reinforced previous research, showing that fake news comments predominantly belonged to the "Novelty" category, whereas real news comments fell under the "Expectation" group.
- By integrating sentiment and emotion analysis, this study offers valuable insights into how fake and real news impact audience emotions, contributing to a better comprehension of user interaction and emotional reaction trends in the identification of fake news.

C. Training and Testing Split

Validation and testing encompass 20% of the pre-processed dataset, with the remaining 80% dedicated to training the machine learning models. This strategic division ensures an 80:20 ratio, optimizing the utilization of data for robust model training and effective validation.

D. Classification of Model

The methodology illustrated in the flowchart represents the method of identifying false information by machine learning and sentiment analysis, specifically with a Bi-LSTM model. Below is a step-by-step explanation of each stage:

1. Bi-LSTM

An extension of LSTMs, bi-LSTM models have shown promise as efficient SOC battery estimate tools. The two primary components of bi-LSTM models are one LSTM that processes input features from the past in a forward manner and another LSTM that processes input characteristics from the future in a backward way [18][19]. Bi-LSTM models can improve prediction performance thanks to these features, especially when dealing with lengthy sequential data [20]. Recently, there has been a noticeable increase in interest in DL-based techniques, particularly those that use RNNs like LSTM for the identification of bogus news [21][22]. An RNN type called LSTM is suitable for processing text data with intricate and protracted relationships, like the news items depicted in Figure 2, as it can identify temporal connections in sequential data. Published a study proposing an LSTM-based model with a 94% accuracy rate in Equation (1).

$$h\vec{t}\left(i.e.ht = \left[h\vec{t}; h\overleftarrow{t}\right]\right) \quad (1)$$

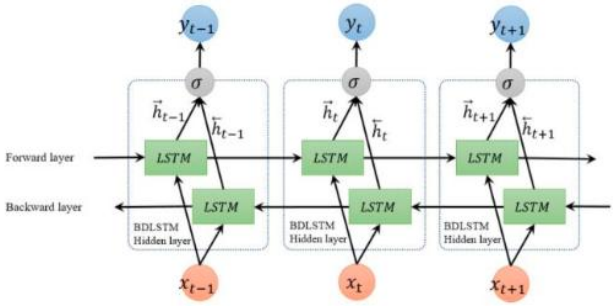

Figure 2: The architecture of Bi-LSTM model

2. Cat Boost

This paper suggests an ensemble machine learning method for detecting bogus news, with a view to comparatively evaluating the performance difference between bagging and boosting ensemble learners, selecting one of the best algorithms for each category for Bagging-based and CatBoost for Boosting based ensembles[23]. one on news stories about global politics, and the other on posts and news stories on COVID-19, and the results show improved performance compared to many existing results on the detection of fraudulent news. The study's findings help develop more effective methods for spotting misleading material, which are essential for maintaining the integrity of online information.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, April 2025
Impact Factor: 7.67

3. Logistic Regression

To predict text categorization, use logistic regression. The following list of current research describes how this machine-learning algorithm is used. Past studies on detecting false news are displayed below. When developing a model to identify false news, logistic regression provides a high degree of accuracy. The predictive capacity of logistic regression in probability values has demonstrated that it is quite successful in resolving binary categories. The LR detection model has a range of 79.0% to 89.0% and can handle both short and lengthy input texts with excellent accuracy, according to the data in the table and Figure 3. It has been demonstrated that TF-IDF's characteristics make it a useful tool for text preparation tasks[24].

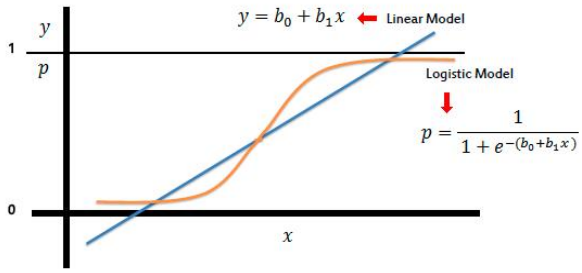

Figure 3: Logistic Regression

E. Performance Matrix Model Evaluation

In order to select the evaluation metric and analyze the model appropriately, it is necessary to understand how each metric is measured. The objective was to assess the performance of ML algorithms by examining each of these performance metrics, such as accuracy score, AUC, and F1-Score.

1. Accuracy

The accuracy of the model is defined as the proportion of examples it correctly classifies and the total error in class prediction. The model's performance across classes is summed up by this metric. Skewed data, however, might give false impressions of performance. The majority class might be correctly predicted by a classifier, but it might incorrectly categorize instances of the minority class. The accuracy is computed as displayed in Equation (2).

$$Acuracy = \frac{TP+TN}{TP+TN+FN+FP} \quad (2)$$

Where,

TP (True Positive): The number of times the model accurately forecasted the positive class.

TN (True Negative): The frequency with which the model accurately predicted the negative class.

FP (False Positive): The frequency of inaccurate model predictions for the positive class.

FN (False Negative): The frequency of inaccurate model predictions for the negative class.

2. AUC (Area Under the Curve)

In classification models, AUC is a crucial performance indicator that assesses the models' capacity to discriminate across different classes. This data comes from the ROC curve, which compares the true positive rate and false positive rate at different threshold levels. AUC = 0.5 indicates random guessing, whereas AUC = 1.0 indicates perfect classification. A higher AUC value indicates stronger model performance. AUCs above 0.9 are often regarded as exceptional, those between 0.8 and 0.9 as good, and those below 0.7 as bad. A high AUC in fake news identification indicates that the model successfully distinguishes between false and authentic news.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, April 2025
Impact Factor: 7.67

3. F1-Score

Accuracy and recall weighted harmonic mean is the F1-score, often known as the F-measure. This metric is best suited for usage when the dataset is significantly unbalanced. A more thorough evaluation is possible when a wider viewpoint is used. F1-score is calculated using the Equation (3).

$$F1 - score = 2 * \frac{precision * recall}{precision + recall} \quad (3)$$

IV. RESULTS AND DISCUSSION

The simulated sentiment analysis results for cybersecurity ML methods for identifying fake news are discussed. Outcomes, dataset description, performance metrics, and classifier statistics are all included in this part, which also provides the outcomes of the dataset evaluation that was done for this study.

A. Experiment Results

In this part, the outcomes of using the Bi-LSTM model for ML-based cybersecurity fake news detection on a large dataset are shown.

Bi-LSTM model performance matrices for Fake news detection in cybersecurity

Model	Accuracy	AUC	F1-score
Bi-LSTM	96.89	90.16	91.78

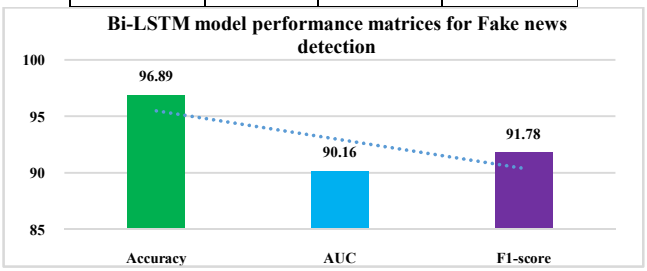

Figure 4: Bi-LSTM model performance matrices for Fake news detection

Table II and Figure 4 display the Bi-LSTM model's performance characteristics for cybersecurity fake news detection. The model's efficacy is demonstrated by its high classification accuracy of 96.89%. The model's ability to differentiate between bogus and true news is demonstrated by its 90.16% AUC score. Additionally, the F1-score of 91.78% highlights a balanced performance between accuracy and recall, demonstrating the model's effectiveness in detecting incorrect information.

Figure 5 shows an ML model's training and validation results, most likely for cybersecurity false news identification. The AUC (Area Under the Curve) score trend across five epochs is displayed in the top graph, where both training and validation AUC improve, though the validation AUC fluctuates before rising sharply. The loss trend can be seen in the bottom graph, where training and validation loss both sharply decline following the first epoch before leveling out at lower values. The increasing AUC and decreasing loss suggest a model improvement, but the widening gap between training and validation AUC indicates potential overfitting.

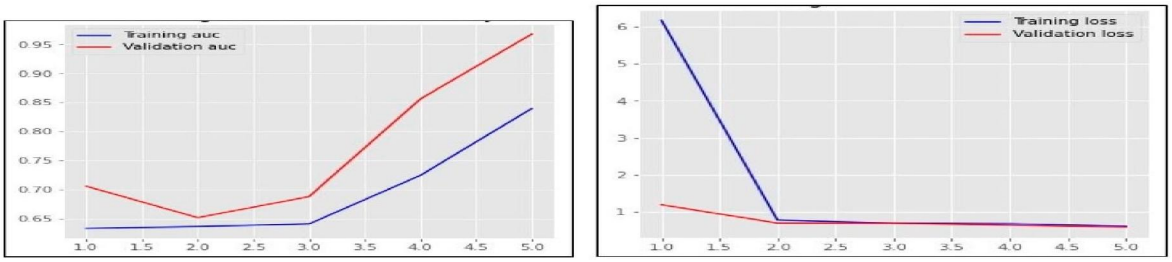

Figure 5: Training and Validation Performance of the Bi-LSTM Model for Fake News Detection

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, April 2025
Impact Factor: 7.67

B. Comparative Analysis and Discussion

A comparison of several cybersecurity false news detecting systems. Compares and assesses several ML and DL models for predicting the identification of false information based on performance measures.

Comparison between various models for fake news detection

Models	Bi-LSTM	CatBoost[25]	LR[24]
Accuracy	91.78	81	74
F1-Score	95	81	78
AUC	94.51	90.51	79

Table III presents a comparative study of many models—LSTM, CatBoost, and LR—for identifying false information. At 91.78% accuracy, 95 F1 scores, and 94.51 AUC, the LSTM model is the most effective, proving its better ability to identify false information. CatBoost exhibits a modest level of performance, as seen by its 81% accuracy, 81 F1-score, and 90.51 AUC. LR records the lowest accuracy (74%), F1-score (78), and AUC (79), suggesting limited effectiveness in contrast to the previous models.

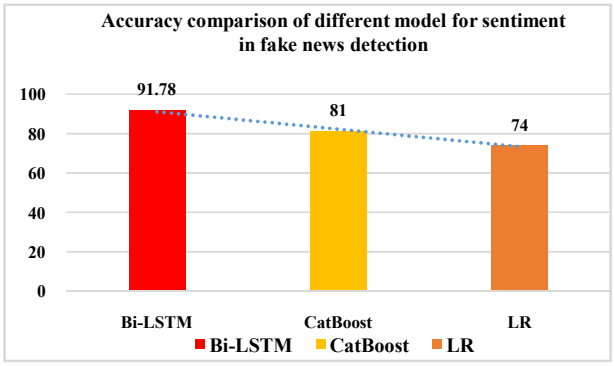

Figure 6: Accuracy comparison for different models for sentiment fake news detection

Three ML and DL models—LSTM, CATBOOST, and LR—are compared in Figure 6 for their accuracy in identifying fake news in cybersecurity. With an accuracy of 91.78%, LSTM was the most accurate. At 0.74%, LR's accuracy was the lowest. Examines the precision of three ML and DL methods.

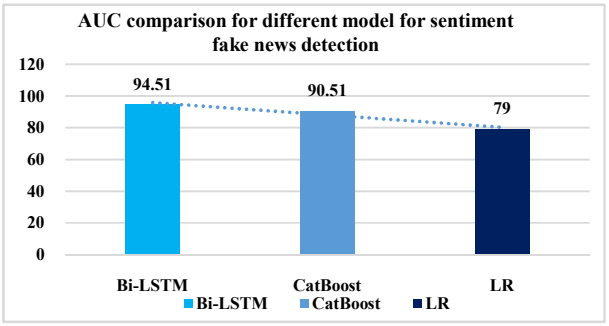

Figure 7: AUC comparison for different models for sentiment fake news detection.

Figure 7 compares AUC of three ML and DL Models, LSTM, CATBOOST and LR, in Predicting Fake news. LSTM achieved the Highest AUC score at 94% LR had very low AUC scores at 79%. This implies that the best model for detecting fake news is LSTM.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, April 2025
Impact Factor: 7.67

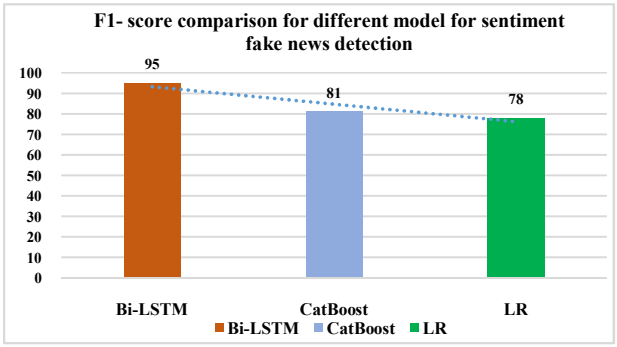

Figure 8: F1-score comparison for different models for sentiment fake news detection.

To forecast false news, Figure 8 compares the F1-score of three ML and DL models: LSTM, CATBOOST, and LR. With an F1-score of 95%, LSTM received the highest score, while LR had the lowest at 81%. As a result, the best model for spotting fake news is LSTM.

V. CONCLUSION AND FUTURE SCOPE

The cybersecurity domain's use of sentiment analysis to identify bogus news showcases its potential to improve classification accuracy. Using cutting-edge machine learning techniques, especially Bi-LSTM models, succeeded in surpassing traditional methods in terms of accuracy and performance metrics. Their findings indicate that sentiment and emotion patterns significantly influence the identification of fake news, as fake news tends to evoke feelings of surprise, contempt, and dread, whereas genuine news is more closely linked to joy, trust, and expectation. The superiority of Bi-LSTM for this job was further confirmed by comparing several machine learning models, achieving the highest accuracy, AUC, and F1 score.

Future research can focus on refining sentiment-based fake news detection models by incorporating multimodal data, such as images and videos, along with text-based sentiment analysis. Additionally, integrating explainable AI techniques can help in understanding how sentiment influences classification decisions, enhancing trust in automated systems. As fake news continues to evolve, adaptive and real-time detection approaches using sentiment and emotion analysis will be crucial in mitigating misinformation, especially in critical domains like cybersecurity, politics, and public health. This study examines how DL and ML might help IDSs detect threats.

REFERENCES

[1] M. I. Khan, A. Arif, and A. R. A. Khan, "AI-Driven Threat Detection: A Brief Overview of AI Techniques in Cybersecurity," *BIN Bull. Informatics*, vol. 2, no. 2, pp. 248–261, 2024.

[2] S. Pandya, "Comparative Analysis of Large Language Models and Traditional Methods for Sentiment Analysis of Tweets Dataset," *Int. J. Innov. Sci. Res. Technol.*, vol. 9, no. 12, pp. 1647–1657, 2024, doi: https://doi.org/10.5281/zenodo.14575886.

[3] J. Kumar Chaudhary, S. Tyagi, H. Prapan Sharma, S. Vaseem Akram, D. R. Sisodia, and D. Kapila, "Machine Learning Model-Based Financial Market Sentiment Prediction and Application," in *2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, IEEE, May 2023, pp. 1456–1459. doi: 10.1109/ICACITE57410.2023.10183344.

[4] M. Alonso Pardo, D. Vilares, C. Gómez-Rodríguez, and J. Vilares, "Sentiment Analysis for Fake News Detection," *Electronics*, vol. 10, p. 1348, 2021, doi: 10.3390/electronics10111348.

[5] J. M. Mahak Shah, Akaash Vishal Hazarika, Meetu Malhotra, Sachin C Patil, "Bridging Emotions and Architecture: Sentiment Analysis in Modern Distributed Systems," *arXiv Prepr. arXiv2503.18260*, 2025, doi: https://doi.org/10.48550/arXiv.2503.18260.

[6] S. S. S. Neeli, "Critical Cybersecurity Strategies for Database Protection against Cyber Attacks," *J. Artif. Intell. Mach. Learn. Data Sci.*, vol. 1, no. 1, p. 5, 2023.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, April 2025
Impact Factor: 7.67

[7] V. P. Shashikant Tank Kumar Mahendrabhai Shukla, Nimeshkumar Patel, "AI Based Cyber Security Data Analytic Device," pp. 414425–001, 2024.
[8] S. K. Hamed, M. J. Ab Aziz, and M. R. Yaakub, "Fake News Detection Model on Social Media by Leveraging Sentiment Analysis of News Content and Emotion Analysis of Users' Comments," *Sensors*, 2023, doi: 10.3390/s23041748.
[9] S. Chatterjee, "Risk Management in Advanced Persistent Threats (APTs) for Critical Infrastructure in the Utility Industry," *Int. J. Multidiscip. Res.*, vol. 3, no. 4, pp. 1–10, 2021.
[10] S. Duary, P. Choudhury, S. Mishra, V. Sharma, D. D. Rao, and A. Paul Aderemi, "Cybersecurity Threats Detection in Intelligent Networks using Predictive Analytics Approaches," in *2024 4th International Conference on Innovative Practices in Technology and Management (ICIPTM)*, IEEE, Feb. 2024, pp. 1–5. doi: 10.1109/ICIPTM59628.2024.10563348.
[11] A. Bhardwaj, S. Bharany, and S. K. Kim, "Fake social media news and distorted campaign detection framework using sentiment analysis & machine learning," *Heliyon*, vol. 10, no. 16, p. e36049, 2024, doi: 10.1016/j.heliyon.2024.e36049.
[12] B. Bhutani, N. Rastogi, P. Sehgal, and A. Purwar, "Fake News Detection Using Sentiment Analysis," in *2019 12th International Conference on Contemporary Computing, IC3 2019*, 2019. doi: 10.1109/IC3.2019.8844880.
[13] I. F. Rozi, R. Arianto, and H. H. Mahdyan, "Fake News Detection Using Sentiment Analysis Approach in Indonesian Language," in *2023 International Conference on Advanced Mechatronics, Intelligent Manufacture and Industrial Automation, ICAMIMIA 2023 - Proceedings*, 2023. doi: 10.1109/ICAMIMIA60881.2023.10427891.
[14] S. E. V. S. Pillai, K. Rajashekaran, R. A. Reddy, P. K. Pareek, and D. Sontakke, "Optimized Convolution Neural Network Based Fake News Detection Using Sentiment Analysis," *3rd IEEE Int. Conf. Distrib. Comput. Electr. Circuits Electron. ICDCECE 2024*, pp. 2024–2025, 2024, doi: 10.1109/ICDCECE60827.2024.10549302.
[15] Ritu, "Preserving Information Integrity: A Novel Machine Learning Approach for Fake News Detection," in *3rd IEEE International Conference on Mobile Networks and Wireless Communications, ICMNWC 2023*, 2023. doi: 10.1109/ICMNWC60182.2023.10435836.
[16] H. Matsumoto, S. Yoshida, and M. Muneyasu, "Propagation-Based Fake News Detection Using Graph Neural Networks with Transformer," in *2021 IEEE 10th Global Conference on Consumer Electronics, GCCE 2021*, 2021. doi: 10.1109/GCCE53005.2021.9621803.
[17] A. Chabukswar, P. D. Shenoy, and K. R. Venugopal, "Fake News Detection Using Optimized Deep Learning Model Through Effective Feature Extraction," in *2023 International Conference on Recent Advances in Information Technology for Sustainable Development, ICRAIS 2023 - Proceedings*, 2023. doi: 10.1109/ICRAIS59684.2023.10367082.
[18] M. S. Akaash Vishal Hazarika, "Blockchain-based Distributed AI Models: Trust in AI model sharing," *Int. J. Sci. Res. Arch.*, vol. 13, no. 2, pp. 3493–3498, 2024.
[19] S. Mathur and S. Gupta, "Classification and Detection of Automated Facial Mask to COVID-19 based on Deep CNN Model," in *2023 IEEE 7th Conference on Information and Communication Technology (CICT)*, IEEE, Dec. 2023, pp. 1–6. doi: 10.1109/CICT59886.2023.10455699.
[20] S. Nokhwal, P. Chilakalapudi, P. Donekal, S. Nokhwal, S. Pahune, and A. Chaudhary, "Accelerating Neural Network Training: A Brief Review," *ACM Int. Conf. Proceeding Ser.*, pp. 31–35, 2024, doi: 10.1145/3665065.3665071.
[21] S. Masarath, V. Waghmare, S. Kumar, R. Joshitta, and D. Rao, "Storage Matched Systems for Single-click Photo Recognitions using CNN," *2023 Int. Conf. Commun. Secur. Artif. Intell.*, pp. 1–7, 2024.
[22] N. Patel, "Enhanced Network Security: Real-Time Malicious Traffic Detection in SD-WAN Using LSTM-GRU Hybrid Model," *Int. Conf. Commun. Electron. Syst. - ICCES-2024*, 2024, doi: https://doi.org/10.47205/jdss.2024(5-IV)50.
[23] S. Nokhwal, S. Nokhwal, S. Pahune, and A. Chaudhary, "Quantum Generative Adversarial Networks: Bridging Classical and Quantum Realms," in *2024 8th International Conference on Intelligent Systems, Metaheuristics & Swarm Intelligence (ISMSI)*, New York, NY, USA, NY, USA: ACM, Apr. 2024, pp. 105–109. doi: 10.1145/3665065.3665082.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, April 2025
Impact Factor: 7.67

[24] M. Mokhtar, Y. Jusoh, N. Admodisastro, N. Che Pa, and A. Amruddin, "Fakebuster: Fake News Detection System Using Logistic Regression Technique In Machine Learning," *Int. J. Eng. Adv. Technol.*, vol. 9, pp. 2407–2410, 2019, doi: 10.35940/ijeat.A2633.109119.

[25] A. S. Adeyemi, S. O. A. P. D, K. H. B. P. D, O. N. Saliu, M. Olabisi, and N. T. Toye, "Fake News Detection using Ensemble Methods : an Empirical Evaluation of Bagging and Boosting Algorithms," vol. 6, no. 06, pp. 317–325, 2024, doi: 10.35629/5252-0606317325.