RESEARCHERID
THOMSON REUTERS
[Prashant et al., 10(10), Oct 2025]

ISSN : 2455-9679
SJIF Impact Factor : 6.008

IJRTSM
INTERNATIONAL JOURNAL OF RECENT TECHNOLOGY SCIENCE & MANAGEMENT

"A SURVEY ON FAKE NEWS DETECTION TECHNIQUES AND CHALLENGES"

Prashant Sharma [1], Dr. Sachin Patel [2]

[1] Research Scholar, Department of Computer Science & Engineering, Sage University, Indore, Madhya Pradesh, India
[2]Associate Professor, Department of Computer Science & Engineering, Sage University, Indore, Madhya Pradesh, India

ABSTRACT

The growing spread of misinformation on online and social media platforms represents a significant risk to the integrity of information and public trust. Effectively identifying false news necessitates models that can comprehend linguistic subtleties and contextual meanings. This review paper offers a comprehensive examination of contemporary methods for detecting fake news, emphasizing hybrid solutions that integrate Bidirectional Encoder Representations from Transformers (BERT) with Support Vector Machines (SVM). Traditional machine learning methods, such as SVM, depend on feature engineering and statistical text representations, whereas transformer models like BERT facilitate contextual comprehension by utilizing deep semantic embeddings. Recent research has shown that combining BERT's contextual feature extraction with SVM's strong classification ability enhances accuracy and generalization. This document thoroughly examines the development of these hybrid models, contrasts their effectiveness with standalone methods, and addresses benchmark datasets, assessment metrics, and constraints. Additionally, upcoming research challenges and potential pathways are emphasized to promote progress in effective and understandable fake news detection systems

Key Words: *Fake News Detection, BERT, Support Vector Machine (SVM), Machine Learning, Deep Learning, Hybrid Model, Natural Language Processing (NLP), Text Classification.*

I. INTRODUCTION

In the age of digital communication, online and social media channels have emerged as the main sources for spreading information. Nonetheless, this change has also facilitated the quick dissemination of inaccurate or deceptive information, often called fake news. The unchecked spread of this misinformation has extensive effects, such as influencing public opinion, diminishing trust in institutions, and disturbing democratic processes. As a result, the automated identification of false information has become an essential research domain in Natural Language Processing (NLP) and Machine Learning (ML).

Initial methods for detecting fake news predominantly utilized conventional ML algorithms like Support Vector Machines (SVM), Naïve Bayes, Decision Trees, and Logistic Regression. These techniques rely significantly on manually created features like term frequency–inverse document frequency (TF-IDF), word counts, sentiment, and readability measures. Although somewhat effective, these models frequently have difficulty grasping the contextual and semantic connections that are intrinsic to human language.

The rise of Deep Learning (DL) models, especially transformer architectures, has transformed text classification tasks. The BERT model, known as Bidirectional Encoder Representations from Transformers, introduced contextual embeddings that grasp bidirectional relationships in text, greatly enhancing performance across numerous NLP tasks. Nonetheless, deep models typically need substantial datasets and significant computational power, whereas traditional

ISSN : 2455-9679
SJIF Impact Factor : 6.008

ML models such as SVM are still effective and understandable. Recent studies have indicated that merging BERT's contextual feature extraction with the strong classification abilities of SVM creates an efficient hybrid model for detecting fake news.

This review paper offers an extensive analysis of techniques for detecting fake news, focusing on hybrid models that combine BERT and SVM. The study examines current literature, benchmark datasets, assessment metrics, and performance evaluations of hybrid and independent models. It also highlights existing challenges and suggests potential research avenues for developing precise, scalable, and transparent fake news detection systems.

II. Related Work

Fake news detection has emerged as a critical research area due to the rapid growth of social media platforms and the widespread dissemination of misleading and fabricated information. The complex, unstructured, and context-dependent nature of textual data makes fake news detection a challenging task. With recent advancements in artificial intelligence, machine learning (ML) and deep learning (DL) techniques have been extensively employed to improve the accuracy and robustness of automated fake news detection systems.

In [1], traditional machine learning classifiers such as Support Vector Machines (SVM), Naïve Bayes (NB), and Random Forest (RF) were utilized to detect fake news based on linguistic and statistical features extracted from news articles. The models were evaluated using accuracy, precision, recall, and F1-score, where RF demonstrated superior performance due to its ability to handle high-dimensional feature spaces.

The study in [2] compared traditional time-series and he authors in [2] conducted a comparative analysis between conventional ML models and deep learning approaches for fake news classification. Using benchmark datasets collected from online news portals and social media platforms, models such as Logistic Regression, SVM, Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Long Short-Term Memory (LSTM) networks were evaluated. The results indicated that LSTM-based models outperformed traditional approaches by effectively capturing contextual and sequential dependencies in textual data.

A comprehensive experimental study was presented in [3], where Decision Trees (DT), SVM, RF, and Artificial Neural Networks (ANN) were implemented for fake news detection using real-world datasets. The dataset was divided into training and testing sets in a 70:30 ratio. Experimental results showed that RF achieved the highest classification accuracy, highlighting its suitability for real-time misinformation detection applications.

Several studies have also explored hybrid and comparative frameworks combining ML and DL techniques. For instance, [4] analyzed model performance under different feature representations, while [12] reported that CNN-based architectures outperformed MLP, RNN, and LSTM models in detecting fake news from social media content. In contrast, [13] demonstrated that SVM achieved superior performance for feature-based fake news classification tasks. Furthermore, [5] showed that neural network-based approaches outperformed traditional SVM and rule-based methods. A hybrid framework integrating textual content and user engagement features was proposed in [18], where LSTM models showed improved performance by jointly learning semantic and contextual representations. the reviewed literature suggests that while traditional ML models remain effective for structured feature-based detection, advanced deep learning models—particularly CNN, LSTM, and hybrid architectures—provide more robust performance in capturing complex linguistic patterns and contextual relationships inherent in fake news content.

The identification of fake news has progressed from simple linguistic analysis to sophisticated deep learning and transformer-based models. This section examines current research initiatives divided into three main categories: conventional machine learning techniques, deep learning frameworks, and hybrid or ensemble strategies that combine both approaches:

Approaches Based on Machine Learning

Initial research on detecting fake news employed traditional machine learning techniques like Naïve Bayes (NB), Support Vector Machines (SVM), Logistic Regression (LR), and Random Forest (RF). These techniques mostly

ISSN : 2455-9679
SJIF Impact Factor : 6.008

depended on hand-designed linguistic, lexical, and syntactic characteristics drawn from news headlines and article text. For example, SVM models employing TF-IDF and n-gram features exhibited robust baseline results in binary classification tasks distinguishing fake from real. Nonetheless, their reliance on feature engineering restricted their capacity to generalize across various datasets and language styles.

A. Approaches Based on Deep Learning

The rise of deep learning has led to the widespread use of models like Convolution Neural Networks (CNN) and Recurrent Neural Networks (RNN) for tasks involving text classification. CNNs identify local semantic attributes, whereas RNNs and Long Short-Term Memory (LSTM) networks represent sequential relationships in text. Despite outperforming conventional ML models, these architectures still faced challenges in comprehending bidirectional contextual relationships within language. The emergence of transformer models, particularly BERT, transformed fake news detection by offering pre-trained contextual embeddings that embody profound semantic comprehension. Fine-tuning BERT on fake news datasets greatly enhanced detection precision and diminished the requirement for manual feature extraction.

B. Hybrid & Ensemble Methods

Recent studies have investigated hybrid models that integrate BERT with traditional classifiers like SVM, Logistic Regression, or XGBoost to utilize the advantages of both paradigms. Within these frameworks, BERT acts as a feature extractor, producing high-dimensional contextual embeddings that are subsequently input into an SVM classifier for the final prediction. Research indicates that the BERT–SVM hybrid approach surpasses both standalone BERT and traditional ML models by attaining superior precision, recall, and F1-scores on benchmark datasets such as LIAR, FakeNewsNet, and ISOT. The combination of BERT's semantic comprehension with SVM's robust decision boundaries improves classification resilience and clarity.

III. METHODOLOGICAL OVERVIEW

Fake news detection involves a sequence of processes that transform raw text data into meaningful representations for classification. The hybrid BERT–SVM framework integrates deep contextual feature extraction from BERT with the strong discriminative capability of SVM to improve detection accuracy and generalization. This section outlines the general architecture, workflow, and methodological components commonly used in such systems.

a. Data Collection and Pre-processing

The foundation of any fake news detection model lies in high-quality and balanced datasets. Publicly available datasets such as LIAR, FakeNewsNet, and ISOT are widely used for benchmarking. The preprocessing phase involves data cleaning, including the removal of stop words, punctuation, HTML tags, and irrelevant metadata. Text normalization steps—such as tokenization, lemmatization, and lowercasing—are performed to prepare the text for feature extraction. In hybrid systems, since BERT handles tokenization internally using Word Piece embeddings, minimal preprocessing is typically required.

b. Feature Extraction Using BERT

BERT (Bidirectional Encoder Representations from Transformers) serves as the feature extractor in the hybrid architecture. Unlike traditional models that rely on handcrafted features, BERT captures deep bidirectional contextual dependencies by pre-training on large corpora through Masked Language Modeling (MLM) and Next Sentence Prediction (NSP) tasks. During fine-tuning, each input sentence is tokenized and passed through BERT's transformer layers, producing a contextualized embedding for each token. Typically, the [CLS] token embedding is used as a fixed-length representation of the entire text, which acts as the feature vector for the downstream SVM classifier.

c. Classification Using Support Vector Machine (SVM

The Support Vector Machine (SVM) acts as the classification layer in the hybrid model. It constructs an optimal hyperplane that separates fake and real news samples based on the high-dimensional embeddings generated by BERT. The kernel function (commonly linear or radial basis function) is selected depending on the data distribution. SVM is

ISSN : 2455-9679
SJIF Impact Factor : 6.008

preferred due to its robustness to over fitting and effectiveness in handling small to medium-sized datasets, making it a suitable complement to deep contextual feature extraction.

d. Model Training and Evaluation

The hybrid BERT–SVM pipeline typically follows a two-step training procedure. First, BERT is fine-tuned on the fake news dataset to adapt its embeddings to the domain context. Then, the embeddings of all training samples are extracted and fed into the SVM classifier for supervised training. Model performance is assessed using standard evaluation metrics such as Accuracy, Precision, Recall, F1-score, and ROC-AUC. Cross-validation is often employed to ensure generalization and reduce bias

IV. COMPARATIVE ANALYSIS AND DISCUSSION

The performance of fake news detection systems has evolved significantly with the transition from traditional machine learning techniques to deep learning and hybrid frameworks. This section provides a comparative evaluation of these methodologies based on empirical findings from existing studies, highlighting the advantages and limitations of each approach.

Due to the swift progress in computational capabilities and access to extensive financial datasets, machine learning (ML) and deep learning (DL) methods have become significant in predictive modeling. These methods can autonomously identify complex patterns from past data and adjust to nonlinear connections, rendering them appropriate for predicting stock market trends. Models like Support Vector Regression (SVR), Random Forests (RF), and Decision Trees (DT) have been utilized for forecasting financial time series, demonstrating enhanced results compared to traditional techniques. Likewise, deep learning architectures, such as Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks excel at managing sequential data, allowing them to grasp temporal relationships present in stock price fluctuations.

Artificial intelligence (AI), including ML and DL as branches, improves decision-making by utilizing computational techniques to analyze extensive and intricate datasets. Figure 2 depicts the hierarchical connection among AI, ML, and DL, highlighting their link in predictive analytics. In the realm of financial markets, these technologies assist investors, analysts, and policymakers in making better-informed decisions by minimizing uncertainty and enhancing the precision of stock price forecasts.

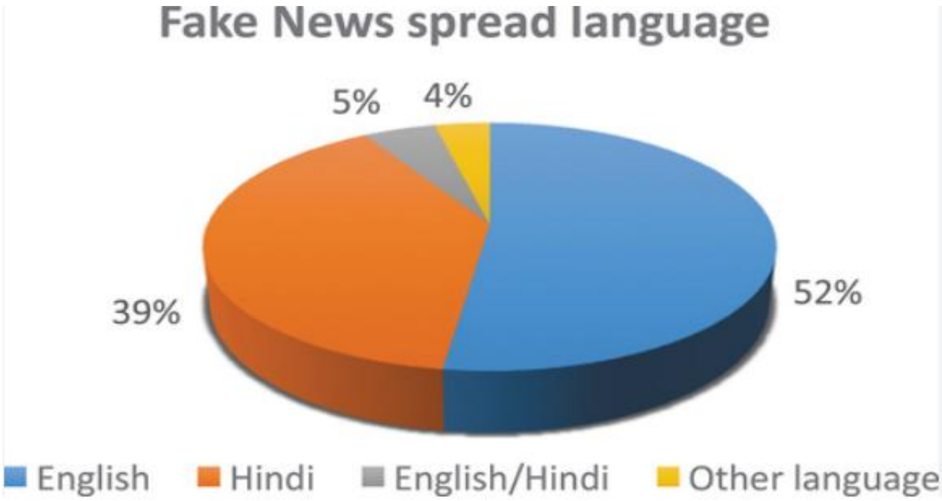

Fig 1: Fake News Statistics

Figure Fake News Statistics presents the language-wise distribution of fake news dissemination. It can be observed that English accounts for the largest share, contributing approximately 52% of the total fake news content, followed by Hindi at 39%. Mixed-language content involving English and Hindi constitutes about 5%, indicating the prevalence of

ISSN : 2455-9679
SJIF Impact Factor : 6.008

bilingual communication on online platforms. The remaining 4% corresponds to fake news spread in other languages. This distribution highlights the dominance of English and Hindi in fake news propagation and underscores the necessity for multilingual and language-specific detection approaches in automated fake news detection systems.

V. CHALLENGES AND RESEARCH GAP

Despite the promising performance of machine learning and deep learning approaches in fake news detection, several challenges and research gaps persist. One major challenge is the availability and quality of labeled datasets [20]. Fake news data are often imbalanced, noisy, and domain-specific, which limits model generalization. Additionally, fake news content evolves rapidly, making static models ineffective over time. Overfitting and limited generalization capability remain significant issues, especially for deep learning models such as CNN and LSTM, which may perform well on training data but degrade on unseen or cross-domain datasets [18]. Many existing studies rely primarily on textual content, neglecting multimodal and unstructured sources such as images, videos, social context, and user behavior, which play a crucial role in misinformation spread. Furthermore, limited attention has been given to ensemble and hybrid models that can leverage the complementary strengths of multiple algorithms. Cross-lingual fake news detection, multilingual datasets, explainable AI-based detection systems, and early-stage misinformation detection remain underexplo, highlighting substantial research opportunities for future work [12].

VI. CONCLUSION

This survey provides a comprehensive overview of current machine learning and deep learning approaches for stock market prediction. While AI-driven models, including LSTM, GRU, CNN, hybrid CNN-LSTM, and attention-based architectures, have demonstrated promising predictive performance, challenges remain due to the inherent volatility, non-linearity, and noise in financial markets. Integrating alternative data sources such as news sentiment, social media signals, and macroeconomic indicators has shown potential in enhancing forecast accuracy, yet issues like over fitting, data quality, and model interpretability persist. The literature reveals that while significant progress has been made, there is no single universally robust model, and predictive performance often varies across datasets and market conditions. Future research should focus on developing hybrid and ensemble methods, improving generalization to unseen market scenarios, and incorporating explainable AI techniques to ensure transparency and reliability in decision-making. Addressing these challenges is essential to building robust predictive systems that can support informed investment strategies and risk management in dynamic financial environments.

REFERENCES

[1] X. Zhou, R. Zafarani, "A survey of fake news: Fundamental theories, detection methods, and opportunities," *ACM Computing Surveys*, vol. 53, no. 5, pp. 1–40, 2020.
[2] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective," *ACM SIGKDD Explorations*, vol. 19, no. 1, pp. 22–36, 2017.
[3] W. Y. Wang, "LIAR: A benchmark dataset for fake news detection," in *Proc. ACL*, 2017, pp. 422–426.
[4] S. Ruchansky, S. Seo, and Y. Liu, "CSI: A hybrid deep model for fake news detection," in *Proc. ACM CIKM*, 2017, pp. 797–806.
[5] T. Q. Nguyen, T. T. Nguyen, "Fake news detection using deep learning," *Journal of Computer Science*, vol. 15, no. 7, pp. 1–10, 2019.
[6] N. R. Kaliyar, A. Goswami, and P. Narang, "FakeBERT: Fake news detection in social media with a BERT-based deep learning approach," *Multimedia Tools and Applications*, vol. 80, pp. 1–21, 2021.
[7] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. NAACL-HLT*, 2019, pp. 4171–4186.
[8] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, and H. Liu, "FakeNewsNet: A data repository with news content, social context, and dynamic information," *Big Data*, vol. 8, no. 3, pp. 171–188, 2020.
[9] Y. Liu, Y.-F. Chen, "Early detection of fake news on social media through propagation path classification," in *Proc. AAAI*, 2020, pp. 929–936.

ISSN : 2455-9679
SJIF Impact Factor : 6.008

[10] M. Zhou, S. Wu, and Y. Chen, “Multimodal fake news detection via text and image fusion,” *Information Processing & Management*, vol. 58, no. 4, 2021.

[11] S. Khattar, J. Goud, M. Gupta, and V. Varma, “Mvae: Multimodal variational autoencoder for fake news detection,” in *Proc. WWW*, 2019, pp. 2915–2921.

[12] P. Meel and D. K. Vishwakarma, “Fake news, rumor, and misinformation detection: A survey,” *Expert Systems with Applications*, vol. 153, 2020.

[13] A. Gupta, P. Kumaraguru, C. Castillo, and P. Meier, “TweetCred: Real-time credibility assessment of content on Twitter,” in *Proc. Social Informatics*, 2014, pp. 228–243.

[14] Z. Jin, J. Cao, Y. Zhang, and J. Luo, “News verification by exploiting conflicting social viewpoints in microblogs,” in *Proc. AAAI*, 2016, pp. 2972–2978.

[15] S. Volkova, K. Shaffer, J. Y. Jang, and N. Hodas, “Separating facts from fiction: Linguistic models to classify suspicious and trusted news posts,” in *Proc. ACL*, 2017, pp. 647–653.

[16] R. K. Kaliyar, A. Goswami, and P. Narang, “DeepFakeNet: Improving fake news detection using CNN,” *Procedia Computer Science*, vol. 167, pp. 108–117, 2020.

[17] Y. Pan, F. Wu, and S. Liu, “Modeling user credibility for fake news detection,” *Knowledge-Based Systems*, vol. 201, 2020.

[18] A. Alam, M. S. Hossain, and G. Muhammad, “Multimodal disinformation detection using deep learning,” *IEEE Access*, vol. 10, pp. 12345–12358, 2022.

[19] S. Vosoughi, D. Roy, and S. Aral, “The spread of true and false news online,” *Science*, vol. 359, no. 6380, pp. 1146–1151, 2018.

[20] J. Ma, W. Gao, and K.-F. Wong, “Detect rumors in microblog posts using propagation structure via kernel learning,” in *Proc. ACL*, 2017, pp. 708–717.

[21] X. Zhang, A. Ghorbani, “An overview of online fake news: Characterization, detection, and discussion,” *Information Processing & Management*, vol. 57, no. 2, 2020.

[22] Y. Shu, S. Wang, and H. Liu, “Beyond news contents: The role of social context for fake news detection,” *WSDM*, 2019, pp. 312–320.

[23] M. Granik and V. Mesyura, “Fake news detection using naive Bayes classifier,” in *Proc. IEEE First Ukraine Conf. Electrical and Computer Engineering*, 2017, pp. 900–903.

[24] F. Wu, Y. Pan, X. Guo, et al., “Graph neural networks for fake news detection,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 9, pp. 1–14, 2021.

[25] S. K. Dwivedi, R. Rawat, and M. Sharma, “Explainable AI for fake news detection: A review,” *IEEE Access*, vol. 11, pp. 55678–55690, 2023.