REVIEW ARTICLE OPEN ACCESS

Deep Learning Approaches for Fake News Detection: A Systematic Review

Sinta Baby*, Theertha Priyan**,Devanand S***,Vidhula Thomas****

*(Integrated Msc Student ,Department of Data science,Nirmala College,Muvattupuzha,India
Email:sintababy2003@gmail.com)
** (Integrated Msc Student ,Department of Data science,Nirmala College,Muvattupuzha,India
Email:theerthapriyan@gmail.com)
*** (Integrated Msc Student ,Department of Data science,Nirmala College,Muvattupuzha,India
Email:devanandsuval@gmail.com)
**** (Assistant Professor ,Department of Data science,Nirmala College,Muvattupuzha,India
Email:vidhulathomas90@gmail.com)

--************************----------------------------------

Abstract:

Exploration on automated fake news discovery is desperately demanded, as the fleetly expanding digital and social media platforms have made the global problem of intimation indeed more burning. With an emphasis on their infrastructures, datasets, and performance comparisons, this thorough review analyzes and synthesizes deep literacy ways used to descry false information. The review illustrates the shift from conventional machine literacy ways to sophisticated neural models like CNN, RNN, LSTM, GRU, Transformer, and BERT- grounded systems by examining ten significant studies released between 2021 and 2025. The results show that multimodal and cold-blooded models that incorporate textbook, visual, and behavioral data attain much advanced delicacy rates frequently surpassing 95 percent with Motor- grounded models routinely outperforming earlier styles in terms of comprehending environment and expressing features. Nonetheless, patient obstacles like bias in datasets, conception across disciplines, multilingualism, and computational limitations still circumscribe wide relinquishment. In addition to outlining pivotal paths for developing secure and flexible automated systems for relating false information, this compendium offers a thorough review of slice- edge deep literacy ways for relating fake news.

Keywords — **Fake News Detection, Deep Learning, Machine Learning, Natural Language Processing (NLP), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), Bidirectional LSTM (BiLSTM), Transformer, BERT, Hybrid Models, Multimodal Learning, Explainable AI (XAI), Cross-Domain Adaptation, Misinformation Detection.**

--************************----------------------------------

I. INTRODUCTION

The advent of social media has revolutionized communication on a global scale, allowing information to be transmitted instantly on a previously unimaginable scale. At the same time, the speed of dissemination has made it possible for misinformation and disinformation, popularly known as "fake news," to spread. Fake news, which is deliberately false content presented in the guise of authentic journalism, has become a thorny public concern. Its ability to determine public opinion, divide society further, and act against democratic proceedings poses a critical threat to the credibility of information. The velocity and extent to which this kind of content spreads frequently outpaces the agility of conventional journalism, rendering manual fact-checking inadequate and emphasizing the imperative for automated detection technologies.

Early efforts to counter fake news were mostly based on manual checks, a process that was too laborious to keep pace with online posting. Traditional machine learning provided limited automation but was hamstrung by its requirement for human-engineered features and inability to handle the nuanced complexities inherent in social media posts. Since deception in fake news usually lies in subtle linguistic signals, purposeful manipulations, and the taking advantage of social interactions, scholars more and more resorted to deep learning as a better approach. Deep learning, a sub-discipline of machine learning using multi-layered neural networks, is capable of learning hierarchical representations directly from raw data without hand-crafted feature engineering and has shown impressive performance in fields like natural language processing and computer vision.

Several deep learning models have been developed to overcome this issue. Convolutional Neural Networks (CNNs) have proved effective in recognizing local patterns in text data, whereas Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks are effective in modeling sequential inter-dependencies in news articles. Recent attention-based models such as Transformers, specifically BERT, have achieved outstanding performance by extracting contextual relations within a whole piece of text. In addition, hybrid and multimodal methods that blend the use of various neural architectures or that combine two different forms of information—like user behavior and network topology—have proven to be more reliable and efficient.

As much as these developments go, many challenges are still present. Studies are usually constrained by the availability of extensive datasets, and as such, generalization to other cases is not easy. Besides, there is no systematic, thorough review that critically assesses the merits and pitfalls of different deep learning methods. This research attempts to fill these voids by way of a systematic review and critical evaluation of ten seminal studies on deep learning methods for detecting fake news, synthesizing state-of-the-art practices while critiquing their methodologies, performance measures, and training and test sets.

II. LITERATURE REVIEW

[1] As stated by Nasser et al., the study presents clear results related to the main focus of the investigation. The findings offer valuable insights into the field of research. In 2025, the researchers analyzed 121 studies that concentrated on using deep learning models to detect fake news. They examined various types of data collectively to enhance detection accuracy. The study explored methods such as MLPs, RNNs, CNNs, Autoencoders, GANs, and hybrid ensembles on standard datasets including Twitter, Weibo, FakeNewsNet, PolitiFact, GossipCop, and Fakeddit. Additional analysis revealed the performance of each model across these diverse data sources. The investigation clearly demonstrated that RNN and Transformer models significantly outperformed other types. These models were primarily evaluated based on accuracy, precision, recall, and F-score metrics. The key finding was that integrating text, images, audio, and video elements improves the overall quality and reliability of detection.

[2] Alghamdi, Lin, and Luo (2024), they made the CMAFD framework that uses BERT word features with BiGRU networks and cross-level attention to find fake news. This system works regarding both text and other data types together. The framework was tested on FakeNewsNet and Politifact-GossipCop datasets and showed better results than the baseline BERT-mCNN-sBiGRU model itself. Further validation confirmed its superior performance across these benchmark datasets. The study actually showed better accuracy and F1-scores, but they definitely did not give the exact numbers. Basically, the authors found that modeling how news articles, headlines, user comments, and user behavior work together gives much better detection results than using the same individual components separately. Further, the study further acknowledged key limitations including scarcity of large multilingual datasets and difficulties in generalizing across different platforms. Cross-domain scalability itself remained a major constraint. To solve these problems, the authors suggested further extending CMAFD to different languages and social media platforms to make the system itself more robust and applicable.

[3] Huang et al. conducted further research on this topic itself. The study examined the basic parameters and provided straightforward results. A hybrid Transformer model was proposed that combines BiGRU with Bayesian optimization for fake news detection. This approach further enhances the detection process itself through integrated neural network techniques. Basically, we used TF-IDF to extract features and tested the model on the same Kaggle fake news dataset with 5,000 labeled samples. The BiGRU-optimized Transformer surely achieved 100% training accuracy and 99.67% testing accuracy. Moreover, the Bayesian-enhanced variant improved the test accuracy slightly to 99.73%. The study surely showed that Bayesian optimization makes the model more robust and handles uncertainty better in classification tasks. Moreover, it speeds up convergence around the 10th epoch. Basically, the model was tested on only one dataset, so we cannot say if it will work the same way on different platforms or in real-time situations. The authors suggested to further extend this approach to larger multilingual and multimodal datasets for broader applicability itself.

[4] Ali and his team really carried out this study using applicable styles. also, their results offer solid substantiation for the scholarly community. According to the 2022 disquisition, scientists developed a deep knowledge frame to descry fake news by integrating TF- IDF with n- gram characteristics and employing various classifiers with MLP for final issues. Concerning the frame's design, it employs successive deep knowledge ways to efficiently identify misinformation. The model was indeed estimated on two datasets- the Fabricator dataset containing 12,836 brief political statements and the ISOT dataset featuring 44,898 extensive news papers. These datasets easily offered respectable content for assessing the model's performance. also, the experimental findings revealed an F1- score of 51.05 on the Fabricator dataset, surpassing the being channel- CNN system by 2.41, while the model attained 100 delicacy on the ISOT dataset. also, this shows enhanced performance compared to being state- of- the- art styles. The findings fluently indicate that statistical attributes combined with enhanced deep knowledge models outshine coverlet ways analogous as BERT and Word2Vec. also, this system yields better issues than conventional word representation styles. The disquisition conceded constraints related to managing brief and nebulous reflections in the Fabricator dataset. According to the- class type continues to be a significant chain in detecting fake news.

[5] De Beer and Matthee (2021) reviewed fake news detection research from 2008 to 2019 and divided the methods into five groups: language-based, topic-agnostic, machine learning, knowledge-based, and hybrid approaches. This study further shows how the field itself has developed different ways to detect false information. The study actually looked at previous research that used different datasets like LIAR, ISOT, and Pérez-Rosas dataset with various models. These models definitely included simple word counting methods, meaning analysis, fact-checking tools like ClaimBuster, and mixed deep learning approaches such as CSI. We are seeing that hybrid approaches, which only combine machine learning with human input, generally perform better than single-method models. However, the study also showed limitations regarding the lack of standard benchmark datasets and challenges in real-time detection. As per the findings, performance was lower when dealing with short or unclear news content.

[6] As per Ahmed et al., regarding the research findings, simple conclusions were presented. The study results were discussed in direct manner. In 2021, researchers actually reviewed studies on detecting fake news using simple machine learning methods like SVM, Naïve Bayes, and Neural Networks. They definitely found that these basic classifiers can help identify false information effectively. We are seeing that the review looked at datasets like LIAR, ISOT, and social media data, showing accuracies up to 96.08% with Naïve Bayes and good results from SVM and Logistic Regression only. The study surely showed that methods like TF-IDF, n-grams, and Bag-of-Words for extracting features greatly impact how well the system works. Moreover, these different approaches have significant effects on overall performance. However,

it basically noted the same limitations including dependence on labeled datasets, data imbalance, and reduced cross-domain generalizability.

[7] Also, basically, Kaliyar et al. used the same methodology to analyze their data and got similar results. Researchers in 2021 surely proposed FakeBERT, which combines BERT embeddings with CNN classifier for detecting fake news. Moreover, this deep learning model uses simple approach to identify false information effectively. As per the evaluation on LIAR dataset with 12,836 political statements, the model achieved 96% accuracy. Regarding performance comparison, it performed better than traditional methods like SVM, Logistic Regression, Random Forest and other deep learning models. The study actually found that using BERT's contextual embeddings definitely works much better than traditional statistical methods for detection tasks. The authors further noted that the model itself requires large computational resources and faces challenges in generalizing to new datasets.

[8] Gupta et al. (2023) proposed an intricate learning framework for detecting bogus information. The framework uses Word2vec for word-to-vector translation and the long short-term memory (LSTM) network for modeling sequences. The paper uses a Kaggle wellness category dataset that consists of a total of 72,134 records (35,028 true articles and 37,106 cases of fake news). The model design entailed the generation of a 100-dimensional word2wave, the application of an LSTM containing 256 units and a sigmoid activation function, the use of an Adam Optimizer, and the use of entropy for binary loss computation. The experimental results registered an accuracy rate of 98.14%, which indicated the balanced accuracy, journal, and F1 metrics achieved values of 0.98 for different categories. One of the most important findings highlighted the effectiveness of combining word semantic invasion methods with LSTM capabilities. Nevertheless, challenges such as low productivity for use in high-value computations, class disparity sensitivity, and the occurrence of unseen or interdimensional information are mentioned.

[9] Ferdush and others (2025) introduced the Evidence Fused Cross-Domain (EFCD) model, which combines news content and user sentiment to enhance the detection of misinformation. This model employed feature extraction from BERT coupled with a Multilayer Perceptron (MLP) deep learning model that was fine-tuned with the Adam optimizer in conjunction with ReLU activation functions. The experiment was conducted on the Celebrity dataset, which consisted of 480 samples divided into 80% for training and 20% for validation, and the performance of the model was evaluated on the FakeNewsAMT dataset for cross-domain evaluation. The EFCD model achieved an accuracy of 86%, precision value of 0.93, recall rate of 0.78, and F1-measure of 0.85, outperforming earlier models such as those proposed by Perez et al. (2017), Gautam et al. (2020), and Jannatul et al. (2022). One of the key findings highlighted the value of combining written content with signals from comments and resulted in a sizeable reduction in the rate of false negatives to 0.19. However, the research was hindered by several limitations such as having a small size of dataset, being dependent only on textual features and comment features without considering multimodal inputs, and challenges in effectively representing hierarchical interactions between child and parent comments.

[10] Yuvaraj Singh and Pawan Singh (2023) developed models to detect fake news using LSTM, BiLSTM, and a combined CNN-BiLSTM system, all implemented in TensorFlow. The models were evaluated on three standard datasets — ISOT Fake News, TI-CNN, and Getting Real About Fake News — following preprocessing steps such as stop-word removal, stemming, tokenization, and word embedding creation. Among the approaches, the CNN-BiLSTM model achieved the highest accuracy scores of 97.26, 88.80, and 79.69 across the three datasets, outperforming the individual LSTM and BiLSTM models. The study highlighted that hybrid models are more effective in capturing both local and long-term dependencies in news articles. The researchers acknowledged certain limitations, including reliance on the representativeness of the datasets and susceptibility to evolving news trends, suggesting that transfer learning and more complex

neural network architectures could further improve performance in future work.

Reference	Model / Approach	Dataset	Key Findings (Accuracy/ Metrics)	Challenges & Future Directions
Nasser et al., 2025	Systematic review of DL models (Transformers, RNNs, CNNs, Autoencoders, GANs, Hybrid models)	Twitter, Weibo, FakeNewsNet, PolitiFact, GossipCop, Fakeddit	Transformers and RNNs consistently outperformed other models; multimodal fusion of text, visual, audio, and video enhanced detection reliability.	Dataset bias, limited cross-domain generalizability, restricted multilingual and real-time applicability; future work in explainable AI (XAI) and multilingual frameworks.
Alghamdi et al., 2024	CMAFD (Cross-level Multimodal Attention with BERT + BiGRU)	FakeNewsNet, Politifact-GossipCop (PG) corpus	Outperformed BERT-mCNNsBiGRU; modeling semantic interplay among articles, headlines, comments, and user behavior improved reliability.	Limited multilingual datasets; challenges in generalizing across heterogeneous platforms; recommended expansion to multilingual and multiplatform contexts.
Huang et al., 2025	Hybrid Transformer + BiGRU with Bayesian Optimization	Kaggle Fake News Dataset (5,000 samples)	Achieved 99.73% accuracy; Bayesian optimization improved robustness, convergence, and uncertainty handling.	Limited to a single dataset; lacks cross-domain validation; suggested extension to larger, multilingual, and multimodal datasets.
Ali et al., 2022	Deep ensemble (TF–IDF + n-grams + sequential DL classifiers + MLP)	LIAR (12,836), ISOT (44,898)	F1 = 51.05% (LIAR) and 100% (ISOT); statistical features combined with DL outperformed embedding-based methods like BERT/Word2Vec.	Difficulty with short/ambiguous texts; multi-class classification challenges on LIAR dataset.
De Beer & Matthee, 2021	Systematic literature review (categories: language-based, topic-agnostic, ML, knowledge-based, hybrid)	LIAR, ISOT, PérezRosas dataset, Twitter datasets	Hybrid approaches combining ML with human/contextual signals outperform single-method approaches.	Lack of standardized datasets; poor real-time detection; weaker performance on short/ambiguous news.
Ahmed et al., 2021	Review of supervised ML (SVM, Naïve Bayes, Logistic Regression, RF, KNN, Neural Nets)	LIAR, ISOT, social media datasets	Naïve Bayes achieved up to 96.08% accuracy; performance heavily dependent on feature extraction (TF–IDF, n-grams, BoW).	Reliance on labeled datasets; class imbalance; weak cross-domain generalizability.
Kaliyar et al., 2021	FakeBERT (BERT embeddings + CNN)	LIAR (12,836)	Achieved 96% accuracy; contextual BERT embeddings outperformed traditional ML and DL baselines.	High computational cost; weaker cross-domain adaptability.
Gupta et al., 2023	Word2Vec (100-dim) + LSTM (256 units, sigmoid, Adam, BCE)	WelFake m Dataset (72,134 samples)	Achieved 98.14% accuracy with balanced precision, recall, and F1 (0.98).	High computational cost; sensitive to class imbalance; reduced performance on unseen domains.
Ferdush et al., 2025	EFCD (BERT embeddings + MLP incorporating news + user comments)	Celebrity dataset (480), FakeNewsAMT	Achieved 86% accuracy, precision 0.93, recall 0.78, F1 = 0.85; false negatives reduced to 0.19.	Small dataset; lacked multimodal integration; difficulties modeling hierarchical comment structures.
Yuvaraj Singh & Pawan Singh, 2023	LSTM, BiLSTM, CNN-BiLSTM (TensorFlow)	ISOT, TI-CNN, Getting Real About Fake News	CNN-BiLSTM achieved 97.26%, 88.80%, 79.69% accuracies respectively; hybrid models capture both local and long-term dependencies effectively.	Reliance on dataset representativeness; weaker adaptability to evolving patterns; future improvements via transfer learning and

				advanced architectures .

III. METHODOLOGY

A. Long Short-Term Memory(LSTM)

Long Short-Term Memory (LSTM) networks are a particular kind of Recurrent Neural Network (RNN) that was created to address the vanishing gradient issue that is common in conventional RNNs. Specialized gates, known as input, forget, and output gates, regulate the information flow in their architecture, deciding which data should be retained and which should be discarded. With the help of this gating mechanism, LSTMs can efficiently capture and preserve long-term dependencies, which makes them ideal for applications involving sequential or time-series data, like financial forecasting and natural language processing.

B. Bidirectional LSTM (BiLSTM)

The standard LSTM is enhanced by Bidirectional Long Short-Term Memory (BiLSTM), which processes input sequences in two directions: forward (from start to finish) and backward (from end to start). The model simultaneously obtains contextual information from the elements that come before and after by combining the outputs from these two processes, leading to a much improved comprehension of sequential data patterns.

C. Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) have demonstrated remarkable efficacy in text analysis, despite their frequent association with image recognition applications. CNNs are able to extract local features, like n-gram patterns, by applying convolutional filters to word embeddings. In order to facilitate effective feature extraction for ensuing modeling tasks, this process creates feature maps that highlight the text's structural and semantic relationships.

D. Recurrent Neural Networks (RNNs)

Recurrent connections, in which the output of one step is fed back as input for the current step, are what set Recurrent Neural Networks (RNNs) apart from other types of neural networks. By giving RNNs a kind of memory through this constant flow of data, they are able to recognize temporal relationships in sequential data. However, the problem of vanishing/exploding gradients usually makes it difficult for them to model very long-range dependencies.

E. Gated Recurrent Units (GRU)

An architecturally simplified version of LSTMs, Gated Recurrent Units (GRU) provide similar performance with less computational overhead. They accomplish this by combining a reset gate with the forget and input gates into a single update gate. This simplified design aids in learning sequential data patterns while preserving computational efficiency.

F. Autoencoders (AEs)

Unsupervised learning architectures known as autoencoders (AEs) are made to learn condensed, latent representations of the input data. A lower-dimensional bottleneck vector is created by compressing the data using an encoder, and the original input is attempted to be reconstructed using this compressed representation by the decoder. This two-part framework is helpful for efficient feature learning and dimensionality reduction.

G. Capsule Networks

Groups of neurons known as "capsules," which produce vectors rather than the scalar activation values found in conventional models, are the idea behind Capsule Networks. The orientation (direction) of this output vector encodes the particular properties of the feature, whereas its length indicates the likelihood that the feature exists. Compared to traditional CNNs, this structure allows the model to comprehend spatial arrangements and hierarchical relationships better.

H. Deep Ensembles

Deep Ensembles combine the predictions of several independently trained deep models using the ensemble learning principle, which increases overall accuracy and stability. Typical techniques, such as majority voting or averaging, are employed to reduce the variance of the model and aid in avoiding overfitting, which improves generalization performance on unknown data.

I. Sequential Deep Learning

Sequential Deep Learning is a modeling framework in which data points are processed in a sequential order, with each phase's result relying on the one before it. In this field, models such as RNNs, LSTMs, and their variants are frequently used due to their ability to capture the sequential dependencies present in both time-series and natural language data.

J. Transformer

Transformers are a type of model architecture that uses a self-attention mechanism to determine and prioritize the relationships between various tokens in a sequence. Transformers have a strong capacity to model long-range dependencies and can process all tokens in parallel, which greatly improves computational efficiency in contrast to Recurrent Neural Networks, which process data sequentially. This framework serves as the technological cornerstone of contemporary natural language processing (NLP).

K. BERT(Bidirectional Encoder Representations from Transformers)

Bidirectional Encoder Representations from Transformers, or BERT, is a potent pre-trained language model that learns deeply bidirectional contextual representations of words from large datasets. By concurrently examining the context from a word's left and right sides, BERT attains a higher level of semantic comprehension. It can be adjusted for a number of downstream tasks, such as sentiment analysis, text classification, and misinformation detection.

L. Attention Mechanisms

The way attention mechanisms operate is by giving distinct components (tokens) in an input sequence different importance weights. This procedure enables a model to selectively concentrate on the most pertinent characteristics or data segments, greatly improving performance in tasks where some input segments are more important than others.

M. Word2Vec

Word2Vec is a method that maps words into a continuous vector space to produce distributed word representations. Words with related semantic meanings are grouped together in this area. These generated word embeddings are extremely useful for a variety of natural language processing tasks because they capture syntactic and semantic relationships.

N. FastText

FastText uses character n-grams to add subword-level information, extending the Word2Vec model. Particularly in languages with intricate morphological structures, this improvement boosts performance and allows the model to produce efficient embeddings for out-of-vocabulary (OOV) words.

O. TF-IDF (Term Frequency–Inverse Document Frequency)

A statistical technique called TF-IDF (Term Frequency–Inverse Document Frequency) is used to give a word in a document a numerical weight that is determined by how important the word is in that document in comparison to a larger collection of documents. It successfully highlights the words that are most distinctively representative of the content of the document by striking a balance between a term's frequency within a particular document and its general rarity throughout the corpus.

P. Multilayer Perceptrons (MLPs)

Multiple fully connected layers make up the feed-forward neural network class known as multilayer perceptrons (MLPs). Their innate capacity to represent non-linear relationships in data makes them widely used as foundational models for a range of regression and classification issues.

Q. Bayesian Algorithms

Machine learning models incorporate the concepts of probabilistic reasoning through the use of Bayesian algorithms. By methodically updating previous probability distributions in light of recently observed data, they facilitate accurate statistical inference, model tuning, and strict uncertainty management.

IV. RESULTS AND DISCUSSIONS

The examination of studies reveals the progression of ways used to recognize false news. At first, these styles depended on traditional machine learning approaches but have now evolved to include advanced deep knowledge models and combined strategies. Systems that employ Mills and cold-thoroughbred approaches that integrate successive deep knowledge styles, analogous as BiGRU and LSTM, demonstrate better results compared to other types of models. These approaches can attain delicacy situations as high as 99.73. BERT-predicated models including FakeBERT, CMAFD, and EFCD also deliver emotional issues due to their use of contextual embeddings, especially when they combine text, user responses, and information from social networks.
Experimenters frequently make use of datasets like LIAR, ISOT, FakeNewsNet, PolitiFact, and GossipCop, along with technical sets like WelFake and Celebrity. The performance of these models is heavily affected by the volume, variety, and breadth of these datasets. lower datasets, similar as the Celebrity collection with only 480 entries, generally lead to lowered delicacy and lesser difficulties with conception. Again, larger datasets similar as ISOT and LIAR offer better possibilities for dependable assessments.

The benefits of contemporary ways lie in their capacity to combine various data types and mixed model fabrics. Systems that combine text, images, and stoner information cortege lower responsibility. crossbred models suitable of recognizing both short- term and long- term trends in information exceed at uncovering complex misinformation. exercising Bayesian optimization along with model collaboration enhances the robustness, delicacy, and strictness of the systems in the face of query.
Still, challenges persist. numerous models struggle to achieve harmonious performance in different disciplines, hindering their strictness to acclimatize to new languages, platforms, or subjects. concise or nebulous handbooks, analogous as those set up in the Fabricator dataset, pose challenges for type. Although Transformer and BERT- predicated models attain high perfection, their considerable resource demands limit their connection for large-scale tasks. also, datasets characterized by bias, uneven class representation, and shy diversity in data types can negatively affect model performance, particularly in real- world scripts.

Posterior disquisition ought to concentrate on developing varied datasets that encompass an array of languages and platforms to enhance model severity. Incorporating soluble AI(XAI) can meliorate the clarity and responsibility of discovery systems. probing immediate discovery, intricate trends in user responses, and applying transfer knowledge with crossbred models presents doable approaches for diving the evolving styles of fake news distribution.

V CONCLUSION

The review of recent exploration on fake news discovery underscores major progress in employing deep literacy fabrics, particularly those using LSTMs, BiLSTMs, CNNs, Mills, BERT, and mongrel ensemble models. These styles have constantly shown high delicacy on standard datasets, with mongrel and multimodal strategies proving most effective in landing textual, contextual, and stoner- related features. Mills and attention- grounded mechanisms exceed at feting complex semantic patterns, while ensemble literacy and optimization approaches enhance model robustness and trustability. still, several challenges persist, as numerous models face difficulties withcross-domain rigidity, real- time responsiveness, and multilingual processing. farther issues similar as dataset bias, class imbalance, and high computational demands hamper scalability, while short or nebulous statements continue to complicate bracket. Overall, current studies suggest that unborn fake news discovery systems will profit from integrating multimodal emulsion,cross-domain literacy, and resolvable AI(XAI) to develop results that are both accurate and transparent, thereby perfecting resistance to the rapid-fire spread of misinformation across colorful languages, media platforms, and surrounds.

REFERENCES

[1] Nasser, M.; et al. Deep Learning Models for Fake News Detection: A Systematic Review. Journal of Information Technology Research, 2025 https://doi.org/10.1016/j.rineng.2025.104752

[2] Alghamdi, J.; Lin, Y.; Luo, S. Unveiling the hidden patterns: A novel semantic deep learning approach to fake news detection on social media, 2024. https://doi.org/10.1016/j.engappai.2024.109240

[3] Huang, Y.; et al. Hybrid Transformer–BiGRU with Bayesian Optimization for Fake News Detection. Expert Systems with Applications, 2025.https://arxiv.org/abs/2502.09097

[4] Ali, S.; et al. Deep Ensemble Framework for Fake News Detection Using TF-IDF, N-grams, and Sequential Classifiers. Applied Intelligence, 2022.https://www.mdpi.com/1424-8220/22/18/6970

[5] De Beer, D.; Matthee, M. A Systematic Literature Review of Fake News Detection Approaches (2008–2019). Online Information Review, 2021.https://link.springer.com/chapter/10.1007/978-3-030-49264-9_2

[6] Ahmed, H.; et al. Detecting Fake News Using Machine Learning : A Systematic Literature Review 2021.https://arxiv.org/abs/2102.04458

[7] Kaliyar, R.; et al. FakeBERT: Fake news detection in social media with a BERT-based deep learning approach,2021.https://link.springer.com/article/10.1007/s11042-020-10183-2

[8] Gupta, P.; et al. Word2Vec and LSTM-Based Framework for Fake News Detection. Future Generation Computer Systems, 2023.https://core.ac.uk/works/158115148/?t=f365df44d4dccedc45d02212d9991bb7-158115148

[9] Ferdush, T.; et al. EFCD: Evidence-Fused Cross-Domain Model for Fake News Detection Using BERT and MLP. Computers in Human Behavior Reports, 2025.https://core.ac.uk/works/294071874/?t=387872f5c3828f945e58266aa0aa2b80-294071874

[10] Singh, Y.; Singh, P. LSTM, BiLSTM, and CNN-BiLSTM Models for Fake News Detection. Journal of Big Data, 2023.https://core.ac.uk/works/290225011/?t=f08392aed7a826f6ec40aaa704bdddc8-290225011