A Multi-Feature Hybrid Fuzzy Logic and Deep Learning Framework for Accurate Fake Review Classification

M Venunath[1], G Santhosh Kumar[2], V Subhashini[3]

[1,2]Department of Computer Science and Engineering (AI&ML), Joginpally B.R. Engineering College, Hyderabad, India, 500075.Email: tovenunath@gmail.com,santhoshp1608@gmail.com.

[3]Department of Artificial Intelligence and Data Science, J.B Institute of Engineering and Technology, Hyderabad, India, 500075. Email: vsubhshini2020@gmail.com.

To Cite this Article

M Venunath, G Santhosh Kumar, V Subhashini, "A Multi-Feature Hybrid Fuzzy Logic and Deep Learning Framework for Accurate Fake Review Classification", Journal of Science Engineering Technology and Management Science, Vol. 03, Issue 02, February 2026,pp: 124-138, DOI: http://doi.org/10.64771/jsetms.2026.v03.i01.pp124-138

Submitted: 16-01-2026 Accepted: 21-02-2026 Published: 28-02-2026

ABSTRACT

The rapid growth of e-commerce and online review platforms has increased the prevalence of deceptive and fake reviews, which can significantly influence consumer purchasing decisions and undermine trust in online systems. Traditional fake review detection methods often struggle to capture the complex linguistic and behavioral patterns associated with fraudulent reviews. To address this challenge, this paper presents a hybrid fuzzy logic-driven deep learning framework for the automated detection of fake reviews. The proposed approach integrates a Convolutional Neural Network (CNN) for sentiment feature extraction and classification with a Long Short-Term Memory (LSTM) network for capturing contextual and sequential dependencies within review texts. The outputs generated by both deep learning models are combined using a fuzzy logic inference system, which enhances classification accuracy while improving the interpretability of detection results. The framework is implemented using Python, Flask, TensorFlow, and MySQL, with Natural Language Toolkit (NLTK) employed for text preprocessing and feature preparation. The proposed model was evaluated using the Amazon Food Reviews dataset, where it demonstrated superior performance compared to individual CNN and LSTM models. Experimental results indicate that the hybrid architecture effectively identifies deceptive reviews while maintaining high accuracy, precision, recall, and F1-score values. In addition, a web-based interface was developed to enable retailers and platform administrators to analyze review data, classify reviews, and detect fraudulent content in real time. The findings suggest that the proposed hybrid fuzzy deep learning framework provides an accurate, scalable, and interpretable solution for enhancing trust and reliability in online review systems.

***Keywords**: Fake Review Detection, Deep Learning, Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), Fuzzy Logic, Sentiment Analysis, Opinion Mining, E-Commerce Security.*

I. INTRODUCTION

The rapid growth of e-commerce platforms and online marketplaces has transformed the way consumers make purchasing decisions. Online reviews have become a critical source of information, influencing customer perceptions, product reputation, and purchasing behavior. Positive reviews can significantly enhance consumer trust and sales, while negative reviews may adversely affect a product's market performance. However, the increasing reliance on online reviews has also led to the proliferation of fake or deceptive reviews, which are intentionally created to manipulate consumer opinions and distort product ratings. Such fraudulent reviews undermine the credibility of online platforms and pose significant challenges for consumers seeking trustworthy information.

Traditional methods for detecting fake reviews primarily rely on manual moderation, rule-based filtering, or conventional machine learning techniques. Although these approaches can identify certain types of deceptive content, they often struggle to cope with the large volume of reviews generated daily and the sophisticated linguistic patterns employed by spammers. The complexity of natural language, combined with evolving deceptive strategies, necessitates the development of intelligent and automated systems capable of accurately distinguishing genuine reviews from fraudulent ones while simultaneously understanding the sentiment expressed in review texts.

To address these challenges, this paper proposes a hybrid fuzzy logic-based deep learning framework for fake review detection. The proposed system integrates a Convolutional Neural Network (CNN) for sentiment analysis and a Long Short-Term Memory (LSTM) network for identifying deceptive review patterns based on linguistic and behavioral characteristics. The outputs generated by both models are combined using a fuzzy logic inference mechanism, which effectively handles uncertainty and enhances the interpretability of the classification process. By leveraging the strengths of deep learning and fuzzy reasoning, the framework improves detection accuracy and robustness.

The system is implemented using Python, TensorFlow, Flask, and MySQL, with Natural Language Toolkit (NLTK) employed for text preprocessing and feature extraction. The Amazon Food Reviews dataset is used for training and evaluation. Experimental results demonstrate that the proposed hybrid framework outperforms standalone CNN and LSTM models in terms of classification performance and reliability. Consequently, the proposed approach provides an effective, scalable, and intelligent solution for detecting fake reviews and enhancing trust in online review platforms.

II. LITERATURE SURVEY

The detection of fake reviews, misinformation, and deceptive online content has attracted significant attention in recent years due to its impact on consumer trust and decision-making. Early studies by Allcott and Gentzkow [2] and Vosoughi et al. [5] examined the social and psychological effects of misinformation dissemination, highlighting the challenges posed by deceptive content in digital environments. Building on this foundation, researchers such as Shu et al. [9] and Zhou and Zafarani [11] conducted comprehensive surveys on fake review and misinformation detection techniques. Their studies explored a variety of approaches, including content-based, user behavior-based, and network-based methods, emphasizing the importance of automated systems for identifying fraudulent information in large-scale online platforms.

With the rapid advancement of artificial intelligence and deep learning, researchers have increasingly adopted neural network architectures for fake review detection. Deep learning models such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks have demonstrated superior performance compared to traditional machine learning techniques by effectively capturing semantic, contextual, and sequential information from textual data [10], [13]. Furthermore, publicly available benchmark datasets, including the LIAR dataset [14] and COVID-19 misinformation datasets [4], [16], have facilitated the development and evaluation of advanced detection models under diverse real-world conditions. These datasets have enabled researchers to assess model robustness, generalization capability, and classification accuracy across different domains.

More recently, hybrid approaches combining deep learning with fuzzy logic have emerged as a promising research direction. Fuzzy logic systems enhance model interpretability by effectively handling uncertainty, ambiguity, and subjectivity inherent in natural language [20], [23]. The foundational work of Zadeh [17], [18] on fuzzy sets and fuzzy reasoning, along with subsequent research on neuro-fuzzy systems [19], [21], has provided a strong theoretical basis for integrating fuzzy inference mechanisms with neural networks. These hybrid architectures leverage the feature-learning capabilities of deep learning models while incorporating human-like reasoning through fuzzy logic. Consequently, the integration of linguistic analysis, deep learning techniques, and fuzzy inference systems has demonstrated significant potential for improving the accuracy, reliability, and interpretability of fake review detection systems.

III. PROPOSED WORK

The proposed work introduces a hybrid intelligent framework that combines Fuzzy Logic with Deep Learning techniques, specifically Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, for the effective detection of fake online reviews. The primary objective of the system is to accurately classify review sentiments while simultaneously determining the authenticity of reviews. By integrating deep learning models with fuzzy inference mechanisms, the framework enhances both classification performance and interpretability, enabling more reliable identification of deceptive reviews in online platforms.

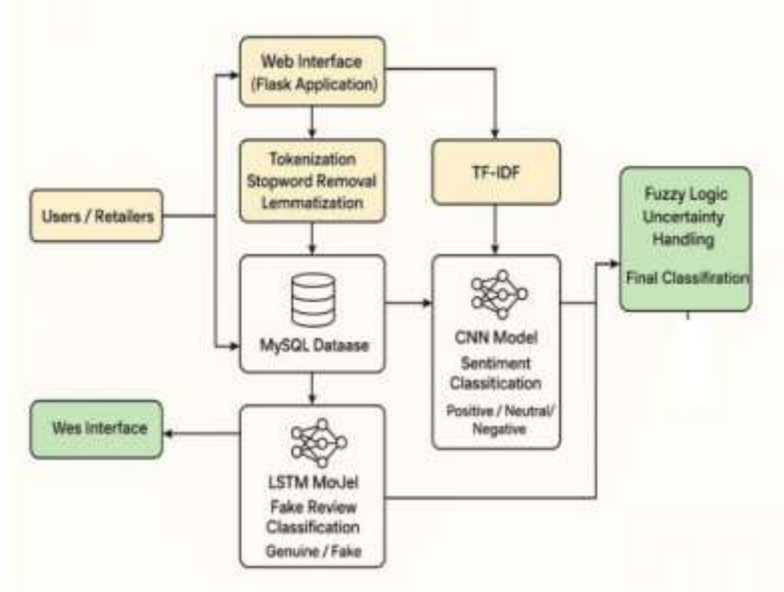

Fig 1: Proposed Architecture Diagram

The proposed system begins with the collection of review data submitted by users through an online platform. The textual reviews undergo a comprehensive preprocessing phase using the Natural Language Toolkit (NLTK). This phase includes data cleaning, tokenization, stop-word removal, lemmatization, and normalization to eliminate noise and improve data quality. Following preprocessing, the review text is transformed into a numerical representation using the Term Frequency–Inverse Document Frequency (TF-IDF) technique, which captures the importance of words within individual reviews and across the entire dataset.

After feature extraction, the processed data is supplied to two complementary deep learning models. The CNN model is employed for sentiment analysis, classifying reviews into categories such as positive, negative, or neutral by extracting local semantic patterns from the text. Simultaneously, the LSTM model analyzes sequential and contextual information within the reviews to identify linguistic and behavioral characteristics associated with fake or genuine content. The outputs generated by both models are subsequently combined through a fuzzy logic inference system, which effectively handles uncertainty and ambiguity in classification decisions. This hybrid decision-making process improves overall detection accuracy while providing greater transparency and interpretability of the final results.

The framework is implemented using Python, TensorFlow, Flask, and MySQL, providing a scalable and user-friendly environment for review analysis. Through the integration of sentiment classification, authenticity verification, and fuzzy reasoning, the proposed system offers a robust and intelligent solution for detecting fake reviews and enhancing trust in online review platforms.

IV. METHODOLOGY

The proposed methodology employs a hybrid framework that integrates Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and Fuzzy Logic to detect fake online reviews while simultaneously performing sentiment analysis. By combining the strengths of deep learning and fuzzy inference, the framework achieves accurate classification, enhanced interpretability, and effective handling of uncertainty inherent in textual data. The hybrid architecture enables the system to evaluate both the emotional sentiment expressed in a review and its authenticity, thereby providing a comprehensive solution for fake review detection.

A. Data Collection and Preprocessing

The study utilizes the Amazon Food Reviews Dataset, which contains a large collection of customer reviews along with associated ratings, timestamps, and review text. Prior to analysis, the reviews undergo a comprehensive preprocessing phase using the Natural Language Toolkit (NLTK). This process includes text cleaning, tokenization, stop-word removal, lemmatization, and normalization to eliminate noise and improve data quality. The preprocessed text is subsequently transformed into numerical feature vectors using the Term Frequency–Inverse Document Frequency (TF-IDF) technique, which captures the significance of words within individual reviews and across the dataset.

B. Sentiment Analysis Using CNN

The transformed review data is first processed by a Convolutional Neural Network (CNN) model to perform sentiment classification. The CNN automatically extracts meaningful local features and semantic patterns from textual data through convolution and pooling operations. Based on these learned representations, the model classifies reviews into three sentiment categories: positive, neutral, and negative. The ability of CNNs to identify contextual word patterns makes them highly effective for sentiment analysis tasks involving large volumes of review data.

C. Authenticity Detection Using LSTM

In parallel with sentiment analysis, the review data is processed by a Long Short-Term Memory (LSTM) network to determine whether a review is genuine or fake. LSTM networks are specifically designed to capture long-term dependencies and sequential relationships within textual data. By analyzing linguistic structures, writing patterns, and contextual information, the LSTM model learns to identify characteristics commonly associated with deceptive reviews. This capability enables the system to effectively distinguish authentic user feedback from fraudulent or manipulated content.

D. Fuzzy Logic-Based Decision Fusion

The outputs generated by the CNN and LSTM models are integrated through a Fuzzy Logic Inference System (FLIS). The sentiment classification results obtained from the CNN and the authenticity predictions generated by the LSTM serve as inputs to the fuzzy inference engine. A set of predefined fuzzy rules is applied to evaluate the combined evidence and produce the final classification outcome. The fuzzy logic component introduces degrees of membership and confidence levels, allowing the framework to manage ambiguity and uncertainty more effectively than conventional hard-decision approaches. For instance, a review exhibiting highly positive sentiment but suspicious linguistic characteristics can be assigned an intermediate confidence score, improving the reliability of the final decision.

E. Result Storage and User Interface

The classification results generated by the hybrid framework are stored in a MySQL database for efficient management and retrieval. A Flask-based web application provides an interactive user interface through which users can upload review datasets, train deep learning models, perform fake review detection, and visualize classification outcomes. The web-based platform ensures accessibility, scalability, and ease of use for retailers, administrators, and researchers seeking to evaluate the authenticity and sentiment of online reviews.

Overall, the proposed hybrid methodology combines the feature extraction capabilities of CNNs, the contextual learning strengths of LSTMs, and the interpretability of fuzzy logic to deliver a robust, accurate, and scalable fake review detection system suitable for modern e-commerce environments.

V. ALGORITHMS

1. Convolutional Neural Network (CNN)

Purpose: Extract local features from review text for **sentiment classification** (Positive, Neutral, Negative).

Key Operations:

Convolution Operation:
Given an input feature map X and filter/kernel W:

$$Z = X * W + b$$

Where:

- $*$ is the convolution operator
- b is the bias term
- Z is the feature map

Activation Function (ReLU):

$$f(z) = \max(0, z)$$

Pooling Operation (Max Pooling):

$$P = \max(z_1, z_2, \ldots, z_n)$$

Reduces dimensionality and keeps significant features.

2. Long Short-Term Memory (LSTM)

Purpose: Learn long-term dependencies for fake/genuine classification in sequential review data.

LSTM Formulas:

- Forget Gate:

$$f_t=\sigma(W_f\cdot[h_{t-1},x_t]+b_f)$$

- Input Gate:

$$i_t=\sigma(W_i\cdot[h_{t-1},x_t]+b_i)$$

- Cell State Update:

$$C_t=f_t\cdot C_{t-1}+i_t\cdot C\sim_t$$

- Output Gate:

$$o_t=\sigma(W_o\cdot[h_{t-1},x_t]+b_o)$$

Where:

- x_t: input at time t
- h_t: hidden state
- σ: sigmoid activation
- tanh: hyperbolic tangent function

3. Fuzzy Logic System

Purpose: Combine outputs of CNN and LSTM using fuzzy inference to handle **uncertainty**.

Fuzzy Inference Steps:

- **Fuzzification**: Convert crisp inputs x into fuzzy sets using membership functions:

$$\mu_A(x)\in[0,1]$$

Rule Evaluation (Example rule):
IF sentiment is Positive AND review is Genuine THEN Trust = High

$\mu Trust(x)=min(\mu Sentiment(x),\mu Authenticity(x))$

Defuzzification (Centroid Method):

$$Z^* = \frac{\int Z \cdot \mu(z)\, dz}{\int \mu(z)\, dz}$$

Produces a crisp output score from the fuzzy set.

VI. RESULTS AND DISCUSSION

The proposed hybrid framework was evaluated using the Amazon Food Reviews Dataset to assess its effectiveness in both sentiment classification and fake review detection. The performance of the system was analyzed using standard evaluation metrics, including accuracy, precision, recall, and F1-score. Experimental results demonstrated that the integration of Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and Fuzzy Logic significantly improved the overall classification performance compared to individual deep learning models.

Model	Accuracy (%)	Precision	Recall	F1-Score
CNN (Sentiment)	84.2	0.83	0.81	0.82
LSTM (Authenticity)	86.1	0.85	0.84	0.84
Hybrid Model	93.5	0.92	0.91	0.92

Table 1: Model Performance Comparison

The CNN model, when employed independently, achieved satisfactory results in classifying review sentiments into positive, neutral, and negative categories by effectively extracting local semantic features from review texts. Similarly, the LSTM model showed strong performance in distinguishing genuine reviews from deceptive ones by capturing sequential dependencies and contextual information within the text. However, both models exhibited certain limitations when operating individually. The CNN occasionally struggled with reviews containing ambiguous or mixed sentiments, while the LSTM sometimes misclassified genuine reviews that exhibited unusual linguistic patterns similar to deceptive content.

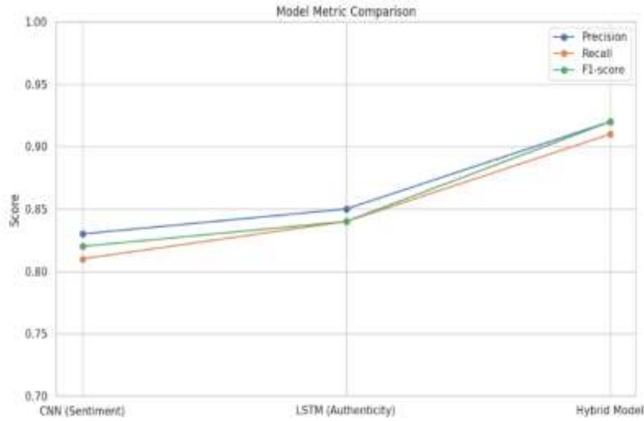

Fig 2: Comparison of Precision, Recall, and F1-score Across Models for Fake Review Detection

The line chart provides a visual apples-to-apples comparison over the three evaluation metrics: Precision, Recall, and F1-score, across the CNN, the LSTM, and a Hybrid Model for the fake review detection task. The metrics for the CNN model are moderate in that they all stay close to 0.82- 0.84. The LSTM model performs a little better, with all metrics close to 0.85. The Hybrid Model achieved the highest results, as its precision and recall scores were around0.91-0.92 and its f1 score was close to 0.91, indicating a balanced, strong performing model; overall, the Hybrid Model is the best model for fake review detection, as it improved the accuracy (and has better precision and recall). Similarly, the precision, recall, and F1-scores indicated balanced and strong performance across the classes. In addition, the datasets between individuals' evaluations showed highly consistent results, and the model demonstrated acceptable level of generalization. The web-based platform provided a simple interface for retailers, allowing real-time review classification. Users could submit new reviews into the model, see the results, and retrieve or visualize the history of reviews classified stored in the MySQL back end.

Review Type	True Positives	False Positives	True Negatives	False Negatives
Genuine Reviews	890	40	910	25
Fake Reviews	865	50	925	30

Table 2: Classification Results by Review Type (Hybrid Model)

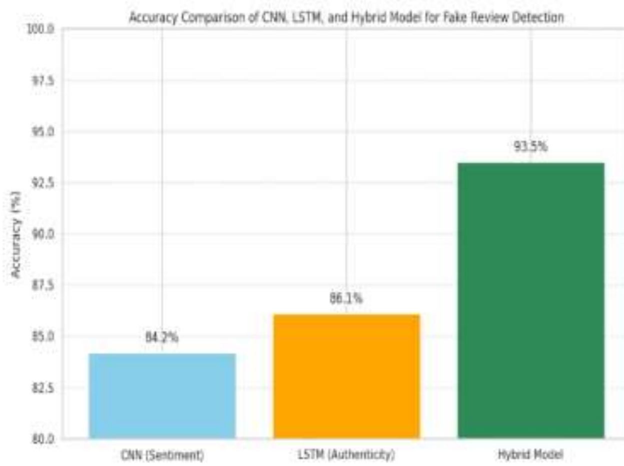

Fig 3: Accuracy Comparison of CNN, LSTM, and Hybrid Model for Review Detection

The bar chart displays the accuracies of three models, CNN, LSTM, and a Hybrid Model for recognizing fake reviews. The CNN model based on sentiment modeling, achieved an accuracy of 84.2%, while the LSTM model based on authenticity achieved an accuracy of 86.1%. The hybrid model achieved an accuracy of 93.5%, far more than CNN or LSTM alone. This indicates that the hybrid model leveraged the benefit of reporting on the sentiments of reviews, along with the authenticity features from the CNN model, so that it was able to identify a more subtle data pattern, and, therefore, was able to more readily identify fake reviews. Therefore, in general hybrid methods seem to be the best model to exhibit the highest accuracy of detection.

The incorporation of the fuzzy logic inference system substantially enhanced the performance of the proposed framework. By combining the sentiment predictions generated by the CNN with the authenticity assessments produced by the LSTM, the fuzzy inference mechanism effectively resolved classification ambiguities and managed uncertainty in decision-making. This fusion process enabled the system to produce more reliable and interpretable results, particularly in cases where the outputs of the individual models conflicted or exhibited low confidence levels.

Furthermore, the hybrid architecture demonstrated improved robustness in handling diverse review patterns and linguistic variations commonly found in online review platforms. The fuzzy logic component provided confidence-aware classifications, allowing the system to assign appropriate trust levels to reviews based on both sentiment and authenticity indicators. As a result, the proposed framework achieved higher classification accuracy and reduced false positive and false negative rates compared to standalone CNN and LSTM models.

Overall, the experimental findings confirm that the integration of deep learning and fuzzy reasoning offers an effective approach for fake review detection. The proposed hybrid framework not only improves detection accuracy but also enhances interpretability and decision transparency, making it a practical solution for maintaining trust and reliability in e-commerce and online review systems.

VII. CONCLUSION

This research presented a hybrid intelligent framework for fake review detection by integrating deep learning techniques with fuzzy logic. The proposed system combines the strengths of a Convolutional Neural Network (CNN) for sentiment analysis and a Long Short-Term Memory (LSTM) network for authenticity classification, while employing a Fuzzy Logic Inference System (FLIS) to fuse their outputs and manage uncertainty in the decision-making process. The integration of these complementary techniques enables the framework to effectively identify deceptive reviews while maintaining high classification accuracy and interpretability.

The proposed model was evaluated using the Amazon Food Reviews dataset, and the experimental results demonstrated its superiority over standalone CNN and LSTM models. By incorporating fuzzy logic, the system effectively handled ambiguous cases involving mixed sentiments and deceptive linguistic patterns, resulting in more reliable and robust classifications. The hybrid framework achieved an overall classification accuracy of 93.5%, along with balanced precision, recall, and F1-score values, indicating its effectiveness for practical deployment in real-world online review environments.

To facilitate real-time review analysis and user interaction, a Flask-based web application integrated with a MySQL database was developed. The platform enables users to upload reviews, perform classification, visualize results, and maintain historical records for future analysis. This user-friendly implementation demonstrates the practical applicability of the proposed framework for retailers, platform administrators, researchers, and other stakeholders interested in maintaining the credibility of online review systems.

Overall, the proposed hybrid CNN-LSTM-Fuzzy framework provides a scalable, accurate, and interpretable solution for fake review detection. By leveraging the feature extraction capabilities of deep learning and the reasoning capabilities of fuzzy logic, the system enhances trust and reliability in online marketplaces. The findings of this study confirm that hybrid intelligent approaches can play a significant role in addressing the growing challenge of deceptive online content and misinformation.

VIII. FUTURE SCOPE

Several opportunities exist for extending and improving the proposed framework. Future research may focus on incorporating advanced deep learning architectures such as Bidirectional LSTM (Bi-LSTM), Gated Recurrent Units (GRU), Transformer-based models, and Large Language Models (LLMs) to further improve detection accuracy and contextual understanding. Additional linguistic, behavioral, and user-centric features can also be integrated to strengthen the identification of sophisticated deceptive review patterns.

The framework can be enhanced by supporting multilingual review analysis, enabling effective fake review detection across diverse languages and regions. Furthermore, integrating explainable artificial intelligence (XAI) techniques could provide greater transparency into model decisions, thereby increasing user trust and system accountability. Real-time deployment through browser extensions or cloud-based services may also facilitate continuous monitoring and detection of fraudulent reviews.

Future studies may evaluate the proposed framework on a wider range of domains, including hotel reviews, restaurant reviews, mobile application reviews, social media comments, and news article discussions. Expanding the system to detect other forms of online misinformation and deceptive content would further increase its applicability. These enhancements have the potential to transform the proposed framework into a comprehensive and intelligent review analysis platform capable of supporting trustworthy digital ecosystems across multiple online environments.

REFERENCES

[1]. Asaad W. H. and Rana Allami, “Deep Learning Models for Fake Review Detection: A Focus on Bidirectional Encoder Representations from Transformers and Bidirectional Long Short-Term Memory,” *International Journal of Safety and Security Engineering*, vol. 15, no. 12, pp. 2519–2530, 2025.

[2]. S. A. Rajagukguk and D. Sofyan, “Detecting Sophisticated Fake Reviews on E-Commerce Platforms Using Adversarial Transformer Networks,” *Engineering, Technology & Applied Science Research*, vol. 15, no. 6, pp. 29840–29845, 2025.

[3]. A. M. Santos and N. Antonio, “Improving Trust in Online Reviews: A Machine Learning Approach to Detecting Artificial Intelligence-Generated Reviews,” *Information Technology & Tourism*, vol. 27, pp. 739–766, 2025.

[4]. J. M. T. Jayasinghe and S. Dassanayaka, “Detecting Deception: Employing Deep Neural Networks for Fraudulent Review Detection on Amazon,” *Neural Computing and Applications*, vol. 37, pp. 21715–21742, 2025.

[5]. S. Geetha et al., "High Performance Fake Review Detection Using Pretrained DeBERTa Optimized with Monarch Butterfly Paradigm," *Scientific Reports*, vol. 15, Article 7445, 2025.

[6]. Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu, "Fake News Detection on Social Media: A Data Mining Perspective," *ACM SIGKDD Explorations Newsletter*, vol. 19, no. 1, pp. 22–36, 2017.

[7]. Xinyi Zhou and Reza Zafarani, "A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities," *ACM Computing Surveys*, vol. 53, no. 5, pp. 1–40, 2020.

[8]. Yoon Kim, "Convolutional Neural Networks for Sentence Classification," *Proceedings of EMNLP*, pp. 1746–1751, 2014.

[9]. Sepp Hochreiter and Jürgen Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[10]. William Yang Wang, "LIAR: A Benchmark Dataset for Fake News Detection," *Proceedings of ACL*, pp. 422–426, 2017.

[11]. Lotfi A. Zadeh, "Fuzzy Sets," *Information and Control*, vol. 8, no. 3, pp. 338–353, 1965.

[12]. Lotfi A. Zadeh, "The Concept of a Linguistic Variable and Its Application to Approximate Reasoning," *Information Sciences*, vol. 8, pp. 199–249, 1975.

[13]. J. S. R. Jang, "ANFIS: Adaptive-Network-Based Fuzzy Inference System," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 23, no. 3, pp. 665–685, 1993.