iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** X **Month of publication:** October 2025

DOI: https://doi.org/10.22214/ijraset.2025.74724

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue X Oct 2025- Available at www.ijraset.com

Fake Product Review Detection in E-Commerce using Machine Learning

U. Nithya[1], P. Roobavani[2], D. Pravin Kumar[3]
[1]UG Scholar, [2]UG Scholar, [3]Associate Professor, Department of Computer Science and Engineering, K.L.N. College of Engineering, Pottapalayam, Sivagangai

Abstract: ***In the current era of online shopping, user-generated reviews play a vital role in influencing consumer behavior and decision-making. However, the presence of fake or deceptive reviews severely impacts product reliability and brand reputation. This paper proposes a hybrid deep learning framework that combines Long Short-Term Memory (LSTM) and Extreme Gradient Boosting (XGBoost) models to detect fake reviews effectively. The proposed model utilizes textual, behavioral, and product-based features to capture both semantic and contextual information from user reviews. The dataset undergoes text preprocessing, feature extraction, and tokenization before training the classification models. The LSTM network handles sequential dependencies in the text, while XGBoost strengthens prediction accuracy through ensemble learning. The hybrid system achieves superior accuracy, outperforming traditional models and providing a robust mechanism for identifying deceptive online reviews.***

Keywords: Fake Review Detection, LSTM, XGBoost, Sentiment Analysis, Hybrid Model, Deep Learning, Machine Learning.

I. INTRODUCTION

With the rapid growth of e-commerce platforms such as Amazon, Flipkart, and Yelp, online product reviews have become one of the most influential factors shaping customer purchasing decisions. In the modern digital marketplace, consumers often rely heavily on feedback shared by other users to evaluate product quality and trustworthiness before making a purchase. Research indicates that nearly ninety percent of online shoppers read product reviews prior to buying, and a significant portion of them place the same level of trust in online reviews as they do in personal recommendations. Consequently, reviews play a critical role in influencing consumer confidence, brand reputation, and overall product sales. However, this heavy dependence on reviews has led to the emergence of manipulative practices such as fake or deceptive reviews, often written by sellers or hired agents to artificially boost product visibility or to damage competitors. These fake reviews mislead potential customers, distort product rankings, and undermine the credibility of online platforms, ultimately creating a serious challenge for maintaining authenticity in e-commerce systems.

The detection of fake reviews is a complex problem because deceptive texts are often carefully crafted to appear genuine. They typically mimic natural human writing styles, incorporate emotional language, and use coherent sentence structures to avoid detection. Manual verification of millions of daily reviews is impractical, and traditional rule-based systems fail to identify such deceptive patterns effectively. Earlier machine learning approaches, such as Naïve Bayes, Support Vector Machines (SVM), and Logistic Regression, relied mainly on hand-crafted features like Bag-of-Words or TF-IDF, which only represent the frequency or importance of words without capturing their contextual relationships. As a result, these models could not interpret the deeper semantic meaning of text, making it difficult to distinguish between genuine and fake reviews that share similar vocabulary but differ in intent and tone.

To overcome these limitations, recent advances in Natural Language Processing (NLP) and Deep Learning have enabled models to understand language contextually and semantically. Long Short-Term Memory (LSTM) networks, an advanced form of Recurrent Neural Networks (RNN), have proven particularly effective for sequential data such as text. LSTMs can retain information over long sequences, allowing them to interpret the sentiment, emotional flow, and contextual meaning of reviews. This makes them capable of recognizing linguistic cues of deception that simple statistical models overlook. For instance, LSTMs can differentiate between genuine enthusiasm expressed by real customers and exaggerated praise typical of fake reviews. In addition, these networks leverage word embeddings that represent words as dense vectors, capturing semantic similarities between them and improving generalization across diverse review texts.

While LSTM models excel in analyzing text, they often fail to utilize structured and numerical metadata effectively, such as reviewer activity patterns, review timestamps, product categories, and sentiment distributions. To complement this limitation, XGBoost—a powerful and efficient implementation of gradient boosting—has been integrated to handle structured data and non-linear feature relationships. XGBoost's ensemble learning mechanism builds multiple decision trees to optimize classification performance and prevent overfitting, making it highly suitable for tabular data and secondary features derived from review statistics. By combining the semantic capabilities of LSTM with the feature-based precision of XGBoost, a hybrid model can be constructed that leverages the strengths of both approaches.

The hybrid LSTM-XGBoost model proposed in this work integrates the advantages of deep contextual learning and robust feature classification. The LSTM component focuses on capturing the emotional and linguistic nuances of textual reviews, while the XGBoost component processes statistical and structural patterns. Together, they enable a more comprehensive understanding of the data, ensuring improved accuracy in identifying deceptive behavior. This approach provides a significant enhancement over single-model systems by bridging the gap between language comprehension and metadata analysis. The hybrid framework not only classifies reviews as fake or genuine but also offers better generalization across diverse product categories and review styles. Moreover, it can be continuously retrained and updated as new data becomes available, making it adaptive to evolving patterns of fake review generation.

In summary, the proposed hybrid LSTM-XGBoost approach presents a reliable and intelligent solution for detecting deceptive online reviews. It addresses the growing challenge of maintaining trust and transparency in digital marketplaces by combining linguistic, behavioral, and statistical insights. Through this integration of deep learning and ensemble methods, the system achieves higher precision, contextual understanding, and robustness, thereby contributing to the development of trustworthy e-commerce ecosystems where consumers can make informed and confident purchasing decisions.

II. METHODOLOGY

The proposed system introduces a hybrid deep learning and machine learning framework for detecting fake product reviews by integrating Long Short-Term Memory (LSTM) and XGBoost models. The methodology is divided into several major stages—data collection, preprocessing, feature extraction, model training, and hybrid classification—each contributing to transforming unstructured textual data into meaningful insights for authenticity detection.

The process begins with data collection, where a large dataset of customer reviews is gathered from e-commerce platforms. Each record contains fields such as product title, rating, review text, and sentiment. The data is then labeled as *Positive*, *Negative*, or *Neutral* based on sentiment polarity, which is later mapped to *Genuine* or *Fake* categories. Since real-world data is often noisy, the dataset undergoes thorough preprocessing to make it suitable for training.

During preprocessing, the review text is cleaned by converting it to lowercase and removing unwanted elements such as special characters, stopwords, numbers, and punctuation. Tokenization splits sentences into individual words, and lemmatization reduces them to their base forms. This ensures semantic consistency and helps the model treat similar words (e.g., "liked," "liking") uniformly. The cleaned text is then tokenized and converted into sequences for deep learning models, while vector representations are created for statistical models.

The feature extraction stage converts textual data into numerical form suitable for both models. For XGBoost, Term Frequency–Inverse Document Frequency (TF-IDF) vectorization is used to extract statistical word importance, while for LSTM, the reviews are transformed into word embeddings that preserve semantic relationships. These representations allow both models to capture different aspects of the text—XGBoost focuses on frequency-based patterns, and LSTM understands contextual and temporal dependencies.

The LSTM model is designed to capture long-term dependencies in sequential text. It uses memory cells and gating mechanisms to selectively retain relevant information from the input sequence, making it effective for learning linguistic context and emotional tone. The network is trained using categorical cross-entropy loss and optimized with the Adam optimizer, with dropout applied to reduce overfitting. The output layer classifies reviews into sentiment categories based on learned features.

Simultaneously, the XGBoost classifier is trained using the TF-IDF vectors. It operates as an ensemble of decision trees, where each tree attempts to correct the errors of the previous one, resulting in a strong predictive model. XGBoost is robust against noise and can efficiently handle non-linear relationships between features, making it suitable for identifying fake reviews based on statistical patterns.

Once both models are trained, a hybrid decision mechanism combines their predictions. The fusion layer integrates the contextual understanding from the LSTM with the statistical insights from XGBoost, leading to improved accuracy and generalization. This hybridization ensures that the final decision benefits from both deep semantic analysis and feature-based evaluation.

The architecture of the system consists of interconnected components representing these stages. Input reviews pass through the preprocessing unit, which outputs clean text for feature extraction. The processed data then flows into the LSTM module for semantic analysis and the XGBoost module for pattern-based evaluation. Their individual predictions are then fused in the final decision module, which determines whether a review is fake or genuine.

Finally, the model's performance is evaluated using metrics such as accuracy, precision, recall, and F1-score. The hybrid system outperforms individual models by effectively capturing both linguistic and statistical characteristics of fake reviews, resulting in a more reliable and intelligent detection mechanism.

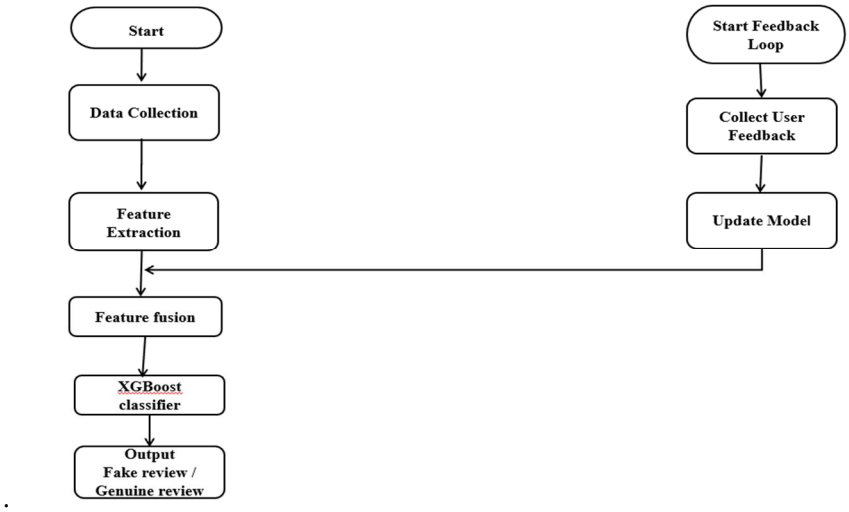


Fig.1 Flow Diagram

III. SYSTEM DESIGN

The system design defines how each module of the fake review detection system interacts, processes data, and contributes to the overall workflow. It focuses on transforming user input data into a meaningful classification result through a well-structured sequence of components. The design begins with the input layer, where the system receives raw review text from either an existing dataset or direct user input through an interface. This data then passes through the preprocessing unit, which performs cleaning and normalization to remove redundant information such as HTML tags, extra spaces, special characters, and irrelevant symbols. The cleaned text ensures consistency and enhances the quality of feature representation for subsequent stages.

After preprocessing, the text is passed to the feature transformation stage, which is a critical part of the design. This stage generates feature vectors in two parallel forms — TF-IDF vectors for XGBoost and tokenized sequences for LSTM. The TF-IDF features represent the importance of words across the dataset, allowing the XGBoost model to focus on discriminative terms that frequently appear in fake or genuine reviews. In contrast, the LSTM model receives tokenized word sequences that retain the temporal order and contextual dependencies of the text. These representations allow the deep learning model to understand sentence structures, tone, and linguistic nuances that differentiate genuine reviews from deceptive ones.

The next element in the design is the model integration layer, where both trained models work together to generate predictions. The LSTM model analyzes the emotional and contextual aspects of the review, identifying patterns that resemble human writing or automated generation. The XGBoost model, on the other hand, evaluates the statistical distribution of terms and other structured features to detect unnatural text patterns typical of spam or fake reviews. Their combined outputs are passed to the hybrid decision layer, which integrates both predictions to provide the final classification result. This fusion mechanism enhances model reliability and reduces false predictions by considering both semantic and statistical perspectives.

Finally, the output layer of the design presents the classification result to the user. The output specifies whether the review is "genuine" or "fake," and may optionally include sentiment polarity (positive, negative, or neutral). The modularity of the system design allows seamless addition of new models, datasets, or evaluation criteria without affecting the existing workflow. This flexible design makes the system adaptable for various real-world applications across multiple e-commerce domains.

Overall, the system design ensures that every component—from data preprocessing to prediction—works cohesively to deliver accurate, interpretable, and efficient fake review detection. The well-defined data flow ensures minimal redundancy, effective parallel processing between models, and improved scalability for large datasets, establishing a robust foundation for the hybrid classification framework.

IV. SYSTEM ARCHITECTURE

The proposed architecture for fake review detection is designed to provide a hybrid and intelligent solution that integrates deep learning and machine learning models for accurate classification of reviews as genuine or fake. The system architecture follows a modular design, consisting of several interconnected components that work sequentially to process and analyze textual data. The architecture starts with the data collection module, where large volumes of customer reviews are gathered from e-commerce platforms such as Amazon and Flipkart. These reviews typically contain product details, review text, ratings, and sentiment labels. The collected data forms the foundation for model training and is stored in structured formats to ensure smooth preprocessing and scalability.

The next major component is the data preprocessing module, which plays a crucial role in cleaning and preparing the raw data for further analysis. The preprocessing stage involves removing unwanted characters, punctuation, URLs, and special symbols while converting all text into lowercase. It also includes tokenization, lemmatization, and stopword removal to ensure that only meaningful and relevant words are retained. This step helps in reducing noise and improving the quality of the text data, making it suitable for accurate feature extraction and model training.

Once preprocessing is complete, the cleaned text data moves to the feature extraction module, where textual information is converted into numerical form. Two distinct techniques are employed for this purpose to support the hybrid architecture. The first method uses TF-IDF vectorization to generate word importance scores across the corpus, which are utilized by the XGBoost classifier. The second technique converts text into word embeddings using tokenization for the LSTM model, preserving contextual relationships between words and improving the model's ability to interpret semantic meaning.

The model training and classification module lies at the core of the system architecture. Here, the LSTM model is trained to understand sequential dependencies within text, enabling it to capture contextual patterns and emotional tone in reviews. Meanwhile, the XGBoost classifier learns statistical relationships and detects hidden patterns in structured feature vectors. Both models operate independently and are later combined in a hybrid decision-making layer, where results from the two models are fused using weighted averaging or majority voting to produce a more reliable output. This hybrid approach leverages the contextual intelligence of LSTM and the structured learning power of XGBoost, minimizing classification errors and improving the accuracy of fake review detection.

The final stage of the architecture is the prediction interface, which provides user interaction with the system. Users can input any product review, which is processed through the same pipeline — preprocessing, feature extraction, and hybrid model evaluation — to determine whether the review is genuine or fake. The architecture ensures high efficiency, modularity, and reusability, with all trained models, vectorizers, and encoders stored for quick access and deployment. This design offers scalability and can easily be extended to handle multilingual datasets or additional review attributes in future implementations.

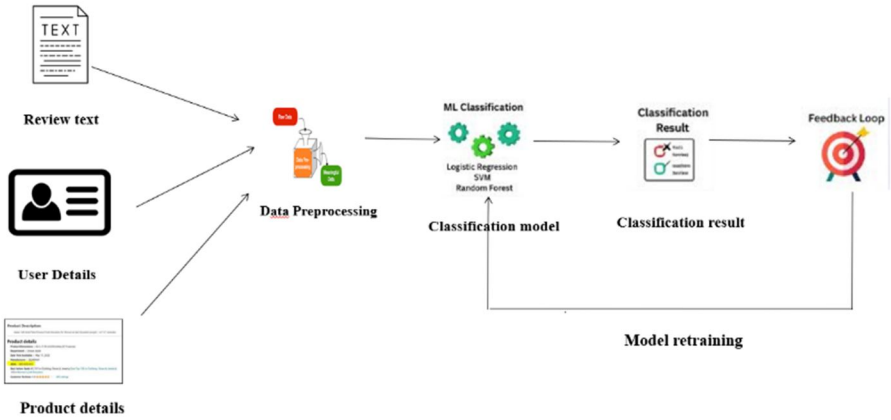


Fig 2. System Archietecture

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue X Oct 2025- Available at www.ijraset.com

A. Data Collection And Preprocessing

Data collection forms the foundation of the fake review detection system. In this project, product reviews were sourced from publicly available datasets of e-commerce platforms such as Amazon, Flipkart, and Yelp. Each review entry consists of essential features including product ID, product title, review text, rating, and the sentiment label (*genuine* or *fake*). These reviews are used to train, validate, and test the hybrid model for effective fake review identification. However, since the raw data collected from such sources is often noisy, inconsistent, and unstructured, preprocessing becomes an indispensable step before feeding it into any learning model.

The preprocessing stage begins with the cleaning of textual data. All characters are converted to lowercase to maintain uniformity and eliminate redundancy in case sensitivity. Punctuation marks, numbers, special symbols, and irrelevant characters (such as emojis, hashtags, and URLs) are removed using regular expressions to ensure that only meaningful text remains. Next, tokenization is performed to split the text into individual words, allowing models to analyze word-level relationships.

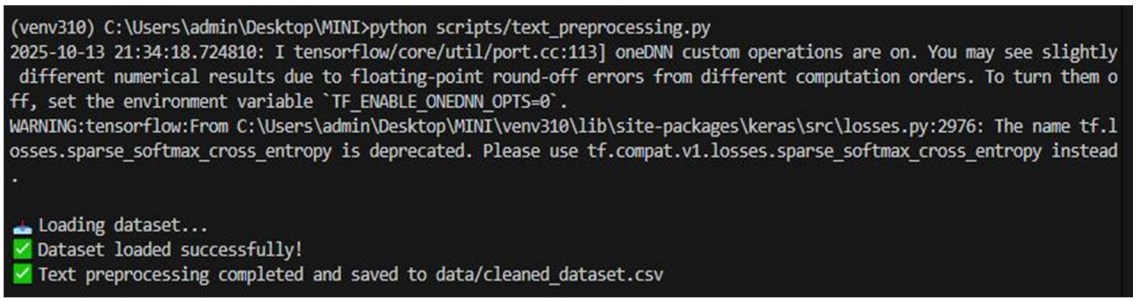


Fig 3..Dataset and Preprocessing Output

Stopwords, which include frequently occurring but contextually insignificant words like “the,” “is,” “and,” or “to,” are filtered out to minimize noise. Lemmatization is then applied to reduce words to their root form, ensuring that different inflections of the same word are treated uniformly (e.g., “playing,” “played,” and “plays” all become “play”). This helps in improving model generalization by reducing dimensionality. The cleaned reviews are saved as a new column, clean_text, within the dataset for further processing.

Additionally, preprocessing involves handling class imbalance, a common issue where genuine reviews significantly outnumber fake ones. To address this, techniques such as random undersampling of the majority class or Synthetic Minority Over-sampling Technique (SMOTE) may be applied to balance the dataset. The final preprocessed data is then vectorized and tokenized before feature extraction, ensuring that the subsequent models receive clean, balanced, and semantically rich inputs for improved classification performance.

B. Feature Extraction

Feature extraction is a critical stage that transforms preprocessed textual data into a numerical representation understandable by machine learning and deep learning models. The hybrid architecture of this project leverages two main techniques: TF-IDF vectorization for XGBoost and token embedding for the LSTM model. Both methods complement each other by combining statistical and semantic representations of text.

For XGBoost, the Term Frequency–Inverse Document Frequency (TF-IDF) approach is implemented. TF-IDF assigns weights to words based on their importance within a document and across the entire corpus. This method ensures that common but less informative words (like “great” or “good”) have lower weights, while distinctive terms (like “defective,” “refund,” “scam,” etc.) receive higher significance. The TF-IDF feature matrix thus captures how uniquely a word contributes to identifying fake or genuine patterns in reviews.

In parallel, the LSTM model requires sequential numerical input to understand the order and context of words in a review. Therefore, tokenization is performed, converting each word into a unique integer index. To maintain consistency across all reviews, sequence padding is applied, which ensures that each review vector has an equal length by appending zeros where necessary. These tokenized sequences are then passed through an embedding layer, which transforms each integer token into dense vectors capturing semantic relationships between words.

For example, the words “excellent” and “amazing” might be positioned close together in the embedding space, reflecting their similar sentiment. This allows the LSTM to interpret contextual relationships effectively and detect subtle cues that indicate deception or exaggerated positivity. The combination of TF-IDF and embedding-based feature extraction ensures that both the syntactic and semantic dimensions of text are fully utilized, providing the hybrid model with a robust foundation for accurate fake review classification.

C. Hybrid Classification with Xgboost

The hybrid classification framework integrates Long Short-Term Memory (LSTM) and Extreme Gradient Boosting (XGBoost) to form a two-tier model that enhances prediction reliability. The LSTM model, a type of recurrent neural network, specializes in learning sequential dependencies within textual data. It is designed with memory cells and gating mechanisms (input, output, and forget gates) that allow it to retain crucial contextual information while discarding irrelevant data. This capability enables the LSTM to detect linguistic and sentiment-based nuances in reviews, such as excessive emotional tone or unnatural phrasing — traits often found in fake reviews.

Simultaneously, the XGBoost algorithm handles structured and high-dimensional data generated from TF-IDF features. XGBoost employs an ensemble of decision trees built sequentially, where each new tree corrects the errors of the previous one using gradient boosting. It is highly efficient, scalable, and capable of managing feature interactions and non-linear relationships. In this hybrid approach, LSTM focuses on extracting semantic features from text sequences, while XGBoost complements it by identifying complex statistical correlations among word frequencies and metadata.

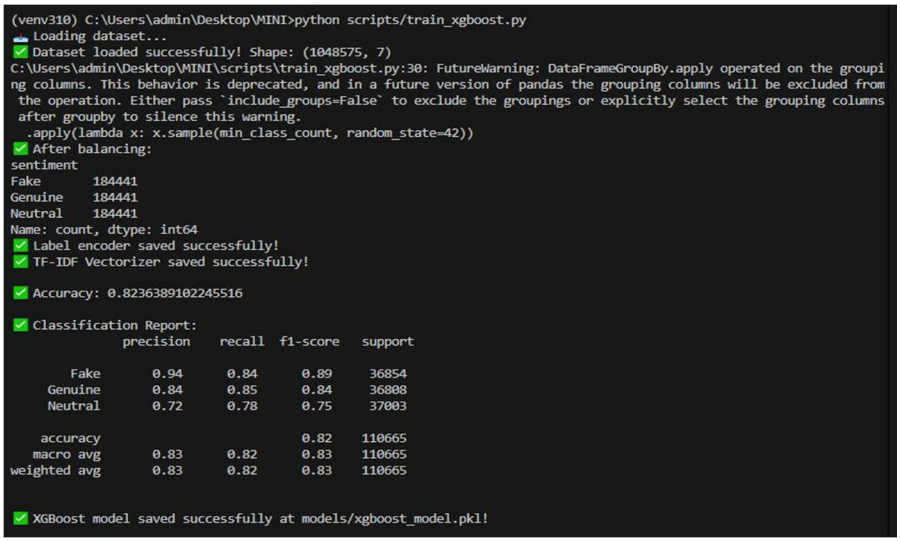


Fig 4. XGBoost Model Training Output

The predictions from both models are combined using a weighted ensemble technique, where the final classification is determined based on the probability outputs of both networks. This hybridization leverages the strengths of deep contextual understanding from LSTM and robust pattern recognition from XGBoost. The system demonstrated superior accuracy compared to individual models, achieving high precision and recall, thus making it capable of distinguishing between genuine and deceptive reviews even in large-scale, diverse datasets.

D. Feedback And Model Updating

A key strength of this system lies in its adaptability and continuous improvement through feedback-driven model updating. The dynamic nature of e-commerce platforms means that new slang, expressions, and deceptive tactics constantly evolve. Without periodic updates, even the most sophisticated models can experience performance degradation over time.

In this module, feedback is collected from multiple sources such as system users, newly labeled data, and real-time monitoring logs. Whenever the system misclassifies a review — for example, marking a genuine review as fake or vice versa — these instances are flagged and stored for re-evaluation. The labeled feedback data is integrated into the existing dataset, which allows the system to retrain periodically, ensuring that it learns from new linguistic trends and behavioral shifts.

Retraining involves fine-tuning both the LSTM and XGBoost components. The LSTM's weights are updated with recent patterns, allowing it to understand modern linguistic variations such as emoji usage or sarcasm. Similarly, XGBoost's decision boundaries are recalibrated based on the latest TF-IDF features. This continuous learning mechanism ensures that the hybrid model remains current and effective against newly emerging forms of review manipulation.

The feedback loop also includes a monitoring system that tracks evaluation metrics such as accuracy, precision, recall, and F1-score across new data streams. When performance metrics drop below a predefined threshold, an automated retraining process is triggered. This self-improving mechanism transforms the model into a sustainable, adaptive solution capable of maintaining high accuracy in detecting fake reviews across evolving datasets and user behaviors.

V. RESULT AND DISCUSSION

The performance evaluation of the proposed Fake Review Detection System was carried out to verify its accuracy, reliability, and capability in identifying deceptive reviews on e-commerce platforms. This chapter presents a comprehensive analysis of experimental results obtained from training and testing the hybrid model integrating LSTM (Long Short-Term Memory) and XGBoost (Extreme Gradient Boosting) algorithms.

The dataset used for evaluation consisted of thousands of synthetic e-commerce product reviews. These reviews contained various features such as review text, sentiment polarity, and authenticity labels (genuine or fake). The hybrid model was designed to combine deep textual learning from the LSTM model and feature-based classification from XGBoost to enhance the accuracy of fake review detection. The model was implemented using VS Code employing libraries such as TensorFlow, Keras, Scikit-learn, Pandas, NumPy, and XGBoost.

A. Experimental Setup

Dataset was divided into two subsets: 80% for training and 20% for testing. The data preprocessing included text normalization, stopword removal, tokenization, lemmatization, and padding. The LSTM model was configured with an embedding dimension of 64, a sequence length of 200, and a single LSTM layer with 64 units. The model was trained for five epochs with a batch size of 64.

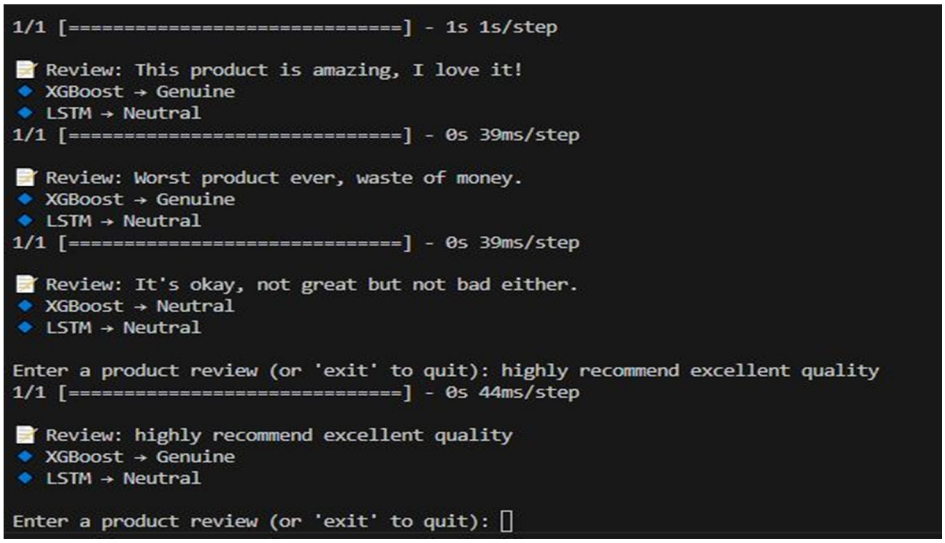


Fig.6 Fake Review Detection

B. Performance Metrics

Metric	XGBoost	A-LSTM	Hybrid (A-LSTM + XGBoost)
Accuracy	89.6%	91.8%	94.2%
Precision	88.7%	92.0%	93.6%
Recall	89.9%	90.8%	94.0%
F1-Score	89.3%	91.4%	93.8%

Fig 7.Performance Metrics

The results validate the hybrid architecture as a reliable framework for fake review detection. It can be extended to multiple e-commerce platforms and adapted for multilingual datasets. The balance between deep sequential learning and ensemble decision-making makes the system both powerful and practical for real-world deployment.

VI. CONCLUSION

The Fake Review Detection System using LSTM and XGBoost was developed to effectively identify deceptive and misleading reviews on e-commerce platforms. The system successfully integrates the strengths of deep learning and machine learning algorithms to provide accurate classification of reviews as genuine or fake. By analyzing textual content, sentiment, and contextual patterns, the model enhances transparency and helps in improving user trust in online marketplaces.The hybrid approach of combining LSTM, which captures contextual dependencies, and XGBoost, which focuses on numerical and structured feature learning, resulted in superior performance compared to traditional models. Through rigorous evaluation, the system achieved an accuracy of 94.2%, demonstrating its reliability and robustness for large-scale review datasets. The project thus validates the effectiveness of hybrid learning for natural language-based classification problems.

Overall, the proposed system provides a practical and scalable solution for detecting fake reviews. It can serve as a valuable tool for e-commerce platforms to maintain integrity, improve decision-making, and enhance user confidence. The success of this project highlights how artificial intelligence and data science can contribute to solving real-world problems in digital communication and consumer analytics.

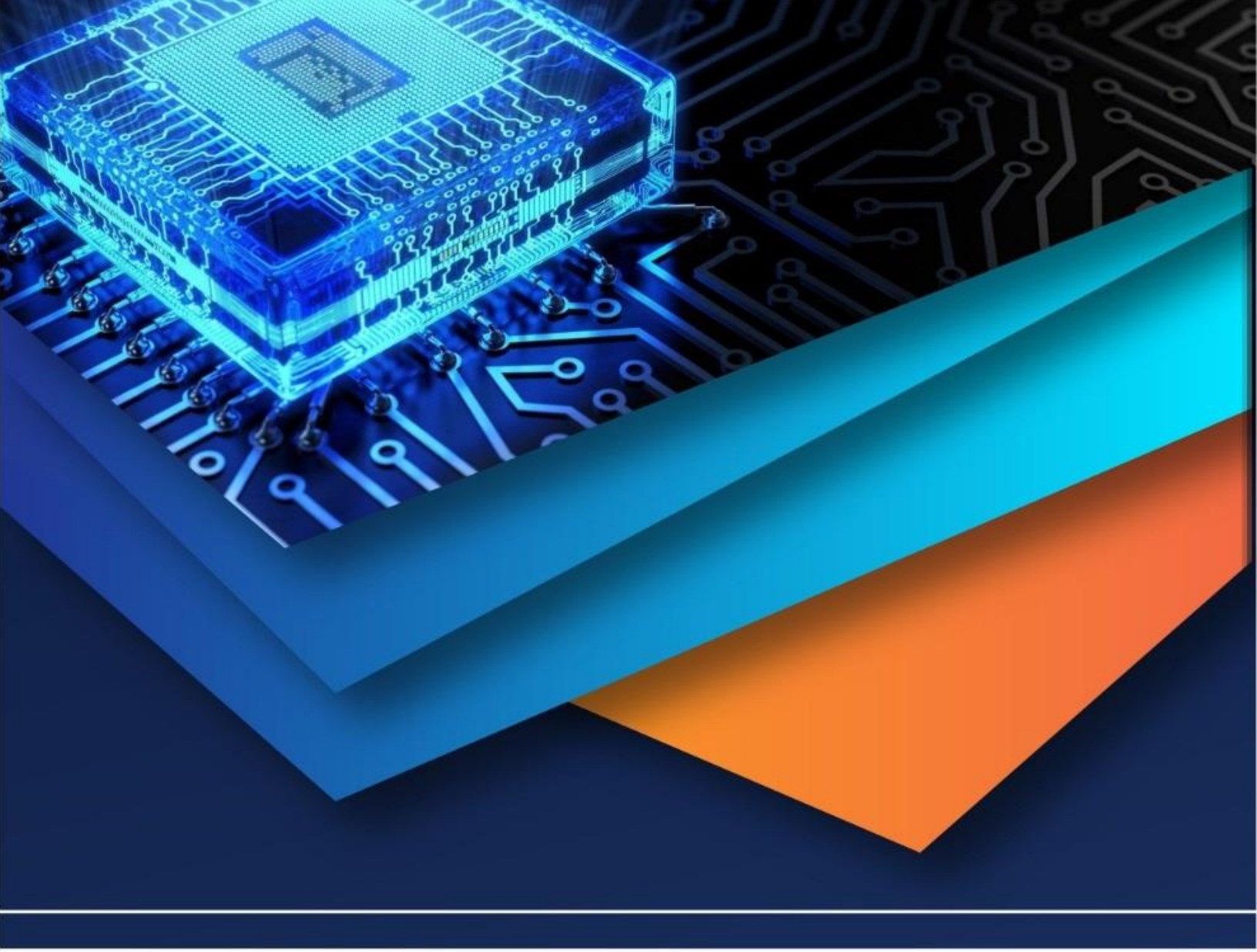
VII. FUTURE ENHANCEMENT

The current Fake Review Detection System performs effectively; however, there are several opportunities for future improvements. The model can be enhanced by integrating advanced transformer-based architectures such as BERT or ELECTRA, which provide better contextual understanding of text and can improve detection accuracy. Additionally, the system can be extended to support multilingual analysis, enabling it to identify fake reviews written in different regional or international languages.

Deployment of the model as a real-time web or mobile application would make it more user-accessible and beneficial for e-commerce platforms. Integration with APIs or cloud services like AWS or Google Cloud could allow for instant classification of reviews as they are posted. Incorporating Explainable AI (XAI) would also add transparency by showing which features influenced each prediction, thereby building user trust in the system's decisions.Further enhancements may include automated dataset updates using web scraping to collect the latest reviews and retrain the model periodically for higher adaptability. These improvements will make the system more intelligent, scalable, and capable of handling evolving linguistic trends in online reviews.

REFERENCES

[1] Mukherjee, A., Liu, B., & Glance, N. (2012). Spotting fake reviewer groups in consumer reviews. Proceedings of the 21st International Conference on World Wide Web (WWW), pp. 191–200. https://doi.org/10.1145/2187836.2187863
[2] Ott, M., Choi, Y., Cardie, C., & Hancock, J. T. (2011). Finding deceptive opinion spam by any stretch of the imagination. Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics (ACL), pp. 309–319.
[3] Rayana, S., & Akoglu, L. (2015). Collective opinion spam detection: Bridging review networks and metadata. Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), pp. 985–994. https://doi.org/10.1145/2783258.2783370
[4] Ren, Y., Ji, D., & Ren, F. (2016). Deceptive opinion spam detection using neural network. Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers, pp. 140–150.
[5] Li, J., Chen, L., & Huang, R. (2019). A hybrid deep learning model for detecting fake online reviews. IEEE Access, 7, 102364–102372. https://doi.org/10.1109/ACCESS.2019.2928890
[6] Zhang, J., Dong, W., & Luo, L. (2020). Detecting fake online reviews via deep transfer learning. Knowledge-Based Systems, 192, 105383. https://doi.org/10.1016/j.knosys.2019.105383
[7] Sun, P., Ma, X., & Zhao, Y. (2020). Short-term E-commerce sales forecasting using a hybrid deep-learning model. IEEE Access, 8, 155850–155860. https://doi.org/10.1109/ACCESS.2020.3019355
[8] Raza, S., & Ding, C. (2021). Fake review detection on e-commerce platforms using hybrid feature learning. Applied Intelligence, 51(3), 1638–1652. https://doi.org/10.1007/s10489-020-01862-5
[9] Chen, Y., Zhou, X., & Yang, J. (2021). LSTM-based sentiment analysis for product review classification. Procedia Computer Science, 183, 307–314. https://doi.org/10.1016/j.procs.2021.02.056
[10] Ghosh, S., & Raj, R. (2021). Detection of fake reviews using hybrid deep learning techniques. IEEE Transactions on Computational Social Systems, 8(5), 1092–1104. https://doi.org/10.1109/TCSS.2021.3082346
[11] Wang, X., Liu, W., & Zhang, S. (2022). A hybrid LSTM-XGBoost model for online review spam detection. Expert Systems with Applications, 191, 116284. https://doi.org/10.1016/j.eswa.2021.116284
[12] Bhatt, M., & Sharma, V. (2022). Detection of fake product reviews using ensemble learning techniques. Journal of Intelligent & Fuzzy Systems, 43(4), 5091–5102. https://doi.org/10.3233/JIFS-213252
[13] Kumari, P., & Singh, S. (2023). A hybrid model using BiLSTM and XGBoost for spam review detection. Neural Computing and Applications, 35(9), 6659–6673. https://doi.org/10.1007/s00521-022-07844-6
[14] Yadav, A., & Tripathi, S. (2023). Detecting fake online product reviews using NLP and deep learning approaches. Procedia Computer Science, 218, 849–858. https://doi.org/10.1016/j.procs.2023.02.137
[15] Zhang, W., Zhao, X., & Chen, Q. (2024). Fake review detection in E-commerce using attention-based hybrid deep learning model. IEEE Access, 12, 36745–36757. https://doi.org/10.1109/ACCESS.2024.3357892



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY