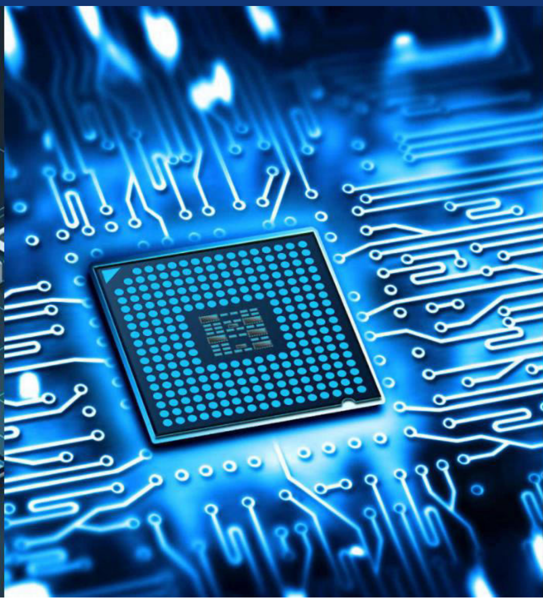
IJIRCCE

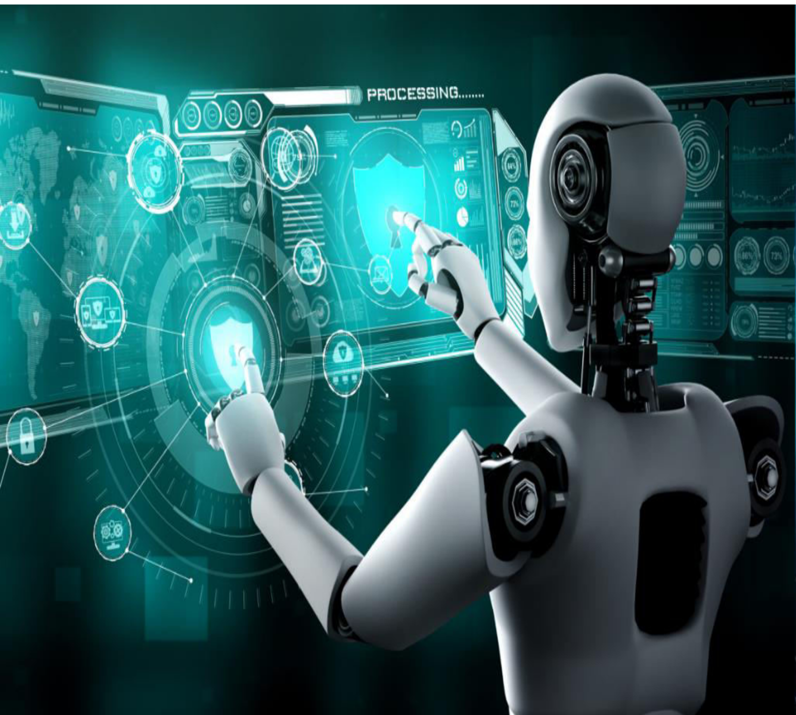
ISSN(O): 2320-9801

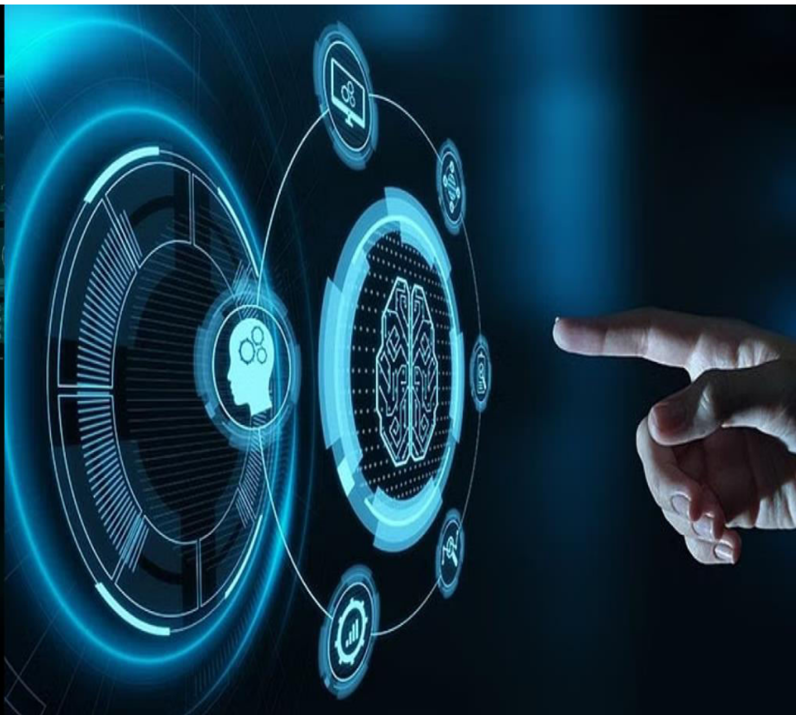
ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 9.927

Volume 14, Issue 8, August 2026

www.ijircce.com ijircce@gmail.com +91-9940572462 +91 63819 07438

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Context-Aware Sentiment Analysis on Product Reviews Using Hybrid Model

Dr. M. Arathi, Asripathi Nikhitha

Professor, Dept. of Computer Science and Engineering, JNTUH University College of Engineering, Science and Technology, Hyderabad, Telangana, India

PG Student, Dept. of Computer Science and Engineering, JNTUH University College of Engineering, Science and Technology, Hyderabad, Telangana, India

ABSTRACT: Sentiment analysis and sarcasm detection as become an important area in natural language processing (NLP) due to growth of e-commerce and social media platforms. Customers give feedback through reviews which helps to understand the contextual meaning and sentiment present in the text.The system integrates DistilBERT and CNN-LSTM hybrid model where CNN (Convolutional Neural Network) used to extract the local features, and LSTM (Long Short-Term Memory) learns sequential and long-term dependencies in the text, and DistilBERT is used to capture contextual embeddings in the text. The proposed system processes the reviews through preprocessing techniques including text cleaning, negation handling, and tokenization. Experimental results shows that the hybrid model achieves high accuracy compared to standalone and architecture architectures. The proposed model is applied in e-commerce platforms, recommendation systems and customer feedback analysis for better decision making.

KEYWORDS: Sentiment Analysis, DistilBERT, CNN, LSTM, Deep Learning, Natural Language Processing, Product Reviews.

I. INTRODUCTION

An intelligent framework for context-aware sentiment analysis of product reviews is designed to analyse sentiment in reviews collected from customers. The rapid growth of e-commerce applications, social media applications has large amount of user-generated textual data. Users express their opinions regarding the products, services and organizations in the form of comments, reviews due to this it has became a challenging for the businesses and researches to extract the meaningful insights from the long textual data. Traditional methods classified the text positive, negative or neutral, as human language is complex, context-dependent and it also consists mixed opinions the sentiment analysis has became a challenging task.

This project performs the context-aware sentiment analysis using natural language processing(NLP) and deep learning models for understanding the customers opinions. This system predicts whether the review is positive, negative, or neutral. This framework combines deep learning techniques with contextual, rule-based analysis to overcome the traditional methods. It uses the DistilBERT and CNN-LSTM (Convolutional neural network- Long short-term memory networks) hybrid model for sentiment analysis. DistilBERT is used to understand contextual meaning and semantic relationships among words within the sentences, while CNN-LSTM is used to extract the local features and sequential dependencies from reviews. The Attention mechanism is used to focus the most influential features. It does the multi-task learning by classifying the sentiment into positive, negative or neutral.

The system processes the user reviews for preprocessing techniques, including text cleaning, tokenization, balancing, normalizing, negation handling and aspect extraction. DistilBERT is used for contextual understanding of the words present within the text, and the embedding is passed to convolutional neural network to extract the local features and then to Long short-term memory to learn the sequential dependencies present in the text. The attention layer assigns the greater importance to the important words and the phrases to capture the relevant information. The implementation is used in e-commerce platforms, customer recommendations and Business decision analysis to understand the customer opinions.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

II. LITERATURE SURVEY

Sentiment analysis has become an important part of the research area in NLP (Natural Language Processing) as there is a rapid growth of online reviews, recommendation systems, social media and e-commerce applications. Traditional machine learning methods, such as Naïve Bayes, decision trees, and Support Vector Machines (SVM) initially used for sentiment classification of product reviews. Eventhough the results are good enough, they are not able to detect the sarcasm and perform the aspect-based analysis [1][2]. These approaches struggled in capturing the contextual understanding, complex linguistic patterns, and semantic relationships in the reviews. Godia et al. performed sentiment analysis on Flipkart product reviews using NLP techniques such as n-grams, TF-IDF, and SMOTE for handling class imbalance. Multiple machine learning algorithms, including Logistic Regression, Decision Tree, KNN, and Naïve Bayes [4].

AS Talaat proposed, transformer-based models enhanced the sentiment performance by introducing contextual embeddings. The models such as BERT, DistilBERT and RoBERTa improved the performance and contextual understanding by learning bidirectional language representations from huge datasets. Researchers combined the transformer-based models with BiLSTM and BiGRU models to increase the accuracy for both text-based and emoji-based reviews. Experimental results indicated that hybrid transformer-based models performed well compared with traditional models and standalone architectures [3]. Sheemu et al.developed a sentiment analysis framework for Amazon computer accessory reviews using both traditional machine learning models and deep learning techniques such as CNN and BERT, along with LIME for model interpretability [8].Sinha et al. compared sentiment classification on Amazon product reviews using TF-IDF and BERT-based language models to evaluate their effectiveness in opinion mining [9]. Mishra et al. proposed the hybrid model with encoder-based architectures with deep learning for sentiment analysis [10]. Rana et al. proposed the hybrid CNN-LSTM model integrated with the uncertain lexical terns for sentiment analysis [11].

Transformation in deep learning models has improved the sentiment analysis performance. As a result, researchers introduced the hybrid models combining the CNN (convolutional neural network), LSTM (Long short-term memory), BiLSTM (Bidirectional LSTM) and transformer-based models BERT, RoBERTa and DistilBERT. CNN models used to extract the local features from the text. The LSTM-based architectures capture the sequential and long-term dependencies. The combination of these models increases the performance as compared to the standalone architectures.

Paul et al. proposed sentiment analysis for e-commerce reviews using a CNN-LSTM hybrid model for classification. CNN is used for feature extraction, and LSTM captures the sequential learning and long-term dependencies in the reviews. Their model achieved high accuracy in sentiment analysis and demonstrated the hybrid deep learning techniques for sentiment analysis [5][7]. Similarly, other researchers compared CNN, LSTM, CNN-LSTM, and CNN-BiLSTM models, showing that the models improved the contextual understanding and sentiment accuracy [6][12].

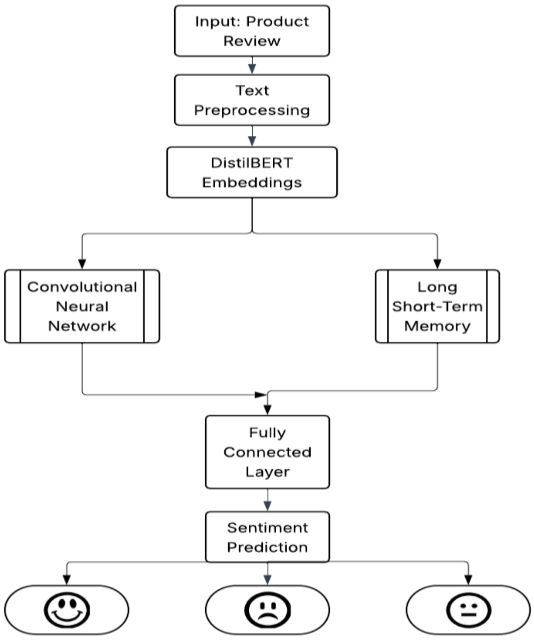
Ali Shaik et al. proposed a hybrid deep learning model combining BERT and LSTM for sentiment analysis of Google Play reviews, with SHAP integrated to provide explainable and interpretable predictions[13].

III. METHODOLOGY

The proposed system is an intelligent hybrid deep learning framework developed for sentiment analysis and sarcasm detection on product reviews. It is used to classify the customer feedback collected from e-commerce into positive, negative and neutral sentiments. The proposed model integrates the DistilBERT with the CNN-LSTM hybrid model to improve contextual understanding and improve the performance. DistilBERT captures the semantic embeddings between the words within the sentence and also identifies the context, CNN extracts the local textual features and important features, and LSTM learns the long-term and sequential dependencies.

© 2026 IJIRCCE | Volume 14, Issue 8, August 2026 DOI: 10.15680/IJIRCCE.2026.1408021
www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|
International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)
(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Figure I Proposed System Architecture



A. Input review text
The input is given in the form of a text which is a product review.

B. Preprocessing
Preprocessing is done for efficiency. The steps that are involved are cleaning the text and removing the stop words and converting the text into lowercase and normalizing the text. It removes the unwanted symbols and URLs.

C. DistilBERT embedding
DistilBERT is the transformer based model which reads the text from both the directions and tokenizes the text by using the tokeniser and generates the embeddings for each token and then it is passed to convolutional neural network. It helps in capturing the contextual embeddings. It identifies the semantic relationships and context of the words present within the text.

D. CNN feature extraction
Convolutional neural network extracts the local and important features.The input layer receives the input from the distilBERT. The input is the DistilBERT contextual embedding. This embedding contains the semantic relationships. This input is passed to CNN layer for feature extraction. The CNN layer applies the multiple filters and extracts the local features and important patterns. Each filter learns the different important patterns helps to classification during training as positive, negative or neutral and the output is feature map. Pooling layer is used to decrease the complexity and dimensionality. It is used in reducing the size of the feature map. Max pooling is used which selects the highest value from the small region in the feature map. Fully-connected layer is used to combine all the extracted features and learns higher-level relationships and output layer is used for classification.

E. LSTM sequential learning
Long short-term memory used to learn the long-term sequential dependencies and contextual dependencies in the review text. It contains three gates input gate, forget gate and output gate. They act as a smart filters.
Forget gate: it decides which data should be deleted from the long-term memory.
Input gate: It stores the new valuable information in the long-term memory
Output gate: It decides which data should be exposed as the current output.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

F. Fully connected classification

Fully connected classification layer is also called as dense layer acts as a final decision maker. It takes the extracted text from the preceeding layers and performs the multi class classification by classifying into positive, negative and neutral.

G. Sentiment prediction

It classifies the sentiment of the review into positive, negative or neutral based on the context.

IV. RESULTS AND DISCUSSION

We used Dell, which has a 12th Gen Intel(R) Core (TM) i5-1240P processor running at 1.70 GHz and 16GB RAM with 3200 MT/s, having a memory of 477 GB and having Windows on it as an operating system. We used Python as our programming language for implementing the hybrid CNN-LSTM and DistilBERT architecture for sentiment analysis and sarcasm detection, and imported NLP libraries PyTorch, Flask, NumPy, Transformers and Regular Expressions (re) for training. The dataset is the Amazon product reviews dataset collected from e-commerce reviews, which contains product-related information like reviewer ID, review text, ratings, summary helpful votes and review timestamps. The dataset is trained to perform sentiment analysis, which classifies into positive, negative or neutral.

Different deep learning models were implemented and evaluated for sentiment analysis on the product review dataset. The performance comparison was done using accuracy and F1-score.

Table I. Performance Comparison of Different Models

Model	Accuracy (%)	F1- Score (%)
Baseline Model	87.00	86.52
CNN-LSTM Model	89.00	88.74
DistiBERT and CNN-LSTM Hybrid model	92.98	92.76

Table I. The baseline model achieved an accuracy of 87%, and CNN-LSTM improved the accuracy to 89% by combining convolutional feature extraction and sequential learning. While DistilBERT and CNN-LSTM hybrid model improved the accuracy to 92.98 and 92.76 F1-score, which improved the contextual understanding and semantic relationships from product reviews.

Table II. Sentiment Classification Report

	Precision	Recall	F1- Score	Support
Negative (0)	0.9066	0.8823	0.8943	2201
Neutral (1)	0.9810	0.8973	0.9373	1441
Positive (2)	0.9265	0.9426	0.9426	5122
Macro avg	0.9380	0.9248	0.9248	8764
Weighted avg	0.9305	0.9296	0.9296	8764

Table II. The positive class achieved the highest recall value 95.94%, which identifies the positive reviews effectively. The neutral class achieved the highest precision value 98.10%, indicating it identifies mixed opinions and balanced reviews correctly.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Figure I. Sentiment Confusion Matrix

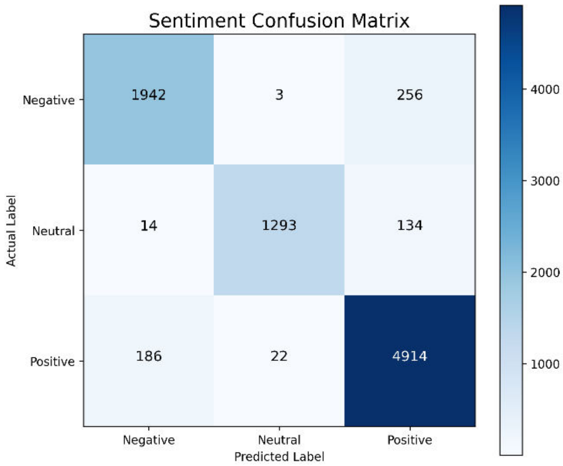


Figure I. The confusion matrix shows that most of the reviews were classified correctly with fewer misclassifications. The model shows better contextual understanding and balanced prediction.The experimental results show that the proposed hybrid DistilBERT and CNN-LSTM model performs effectively for both sentiment analysis and sarcasm detection, where DistilBERT generates the contextual embeddings, CNN for feature extraction and LSTM for sequential learning. The proposed system achieved: Sentiment analysis accuracy: 92.98%.

V. CONCLUSION

This paper, we presented a DistilBERT and CNN-LSTM hybrid model for sentiment analysis on product reviews. DistilBERT generates the contextual embeddings from tokens, which improves contextual understanding and semantic relationships. While CNN extracts the important local features, LSTM learns sequential dependencies. This hybrid model improved the performance compared to traditional methods and standalone architectures. Traditional methods perform simple text classification; the proposed model captures the contextual understanding, sequential dependencies present in review text. The system classifies the reviews into positive, negative or neutral sentiments and also detects the sarcasm. The model achieved 92.98% sentiment analysis accuracy indicating strong contextual understanding and improving the performance.

The project also has a strong foundation for future enhancements. This project help in enhancing in intelligence and real-world applicability. The model performs the sentiment analysis on product reviews by using DistilBERT and CNN-LSTM hybrid model. It can be extended to support multilingual sentiment analysis when the reviews is in multiple languages can be analyzed correctly. It can also be integrated with the advanced transformer models like BERT, RoBERTa, large language models (LLMs) for predicting correctly and increasing the model performance. It can be further enhanced to perform the multimodal sentiment analysis by combining the text with the emojis, video and audio.

REFERENCES

[1] Rahul, V. Raj and Monika. (2019). "Sentiment Analysis on Product Reviews", International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), Greater Noida, India, 2019, pp. 5-9, doi: 10.1109/ICCCIS48478.2019.8974527.

[2] S. Maurya and V. Pratap. (2022) "Sentiment Analysis on Amazon Product Reviews," International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON), Faridabad, India, 2022, pp. 236-240, doi: 10.1109/COM-IT-CON54601.2022.9850758.

[3] Talaat, A.S. (2023) Sentiment analysis classification system using hybrid BERT models. *J Big Data 10*, 110, https://doi.org/10.1186/s40537-023-00781-w

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

[4] A. Godia and L. K. Tiwari. (2024) "Sentiment Analysis and Classification of Product Reviews: A Comprehensive Study Using NLP and Machine Learning Techniques," 10th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2024, pp. 1247-1252, doi: 10.1109/ICACCS60874.2024.10717296.

[5] P. Paul, S. Acharya, B. Misra, S. Majumder, N. Dey and P. Pise. (2024). "Sentiment Analysis for E-Commerce Product Reviews Using CNN-LSTM," First International Conference for Women in Computing (InCoWoCo). Pune, India, 2024, pp. 1-7, doi: 10.1109/InCoWoCo64194.2024.10863425.

[6] Rana, Muhammad Rizwan Rashid and Nawaz, Asif and Ali, Tariq and Alattas, Ahmed Saleh and AbdElminaam, Diaa Salama. (2024). "Sentiment Analysis of Product Reviews Using Transformer Enhanced 1D-CNN and BiLSTM," ACM Trans. Netw. 24(3):112–131, doi: 10.2478/cait-2024-0028.

[7] P. Chinmayi, G. B. Regulwar, N. Anoop, K. G. Reddy and P. Sunitha. (2024). "Sentimental Analysis of Product Reviews with Combined CNN-LSTM Model using Deep Learning,"1st International Conference on Cognitive, Green and Ubiquitous Computing (IC-CGU), Bhubaneswar, India, pp. 1-6, doi: 10.1109/IC-CGU58078.2024.10530690.

[8] S. Y. Sheemu, M. A. A. Symum, A. Zaman, A. S. M. Asif, A. Shakil and R. Zannah. (2024). "Sentimental Analysis of Customer Product Reviews to Understand Customer Needs Using Machine Learning," 27th International Conference on Computer and Information Technology (ICCIT), Cox's Bazar, Bangladesh, pp. 1164-1169, doi: 10.1109/ICCIT64611.2024.11022600.

[9] H. Sinha. (2024). "A Sentiment Analysis of Product Review Data Through Large Language Models (LLMs)," International Conference on Intelligent Computing and Sustainable Innovations in Technology (IC-SIT), Bhubaneswar, India, 2024, pp. 1-6, doi: 10.1109/IC-SIT63503.2024.10862104.

[10] S. Mishra and R. Ranjan. (2024). "Exploring Sentiment Analysis: A Hybrid Approach using Deep Learning and Encoder,". 1st International Conference on Advances in Computing, Communication and Networking (ICAC2N) Greater Noida, India, pp. 394-397, doi: 10.1109/ICAC2N63387.2024.10895721.

[11] A. Rana and S. S. Yadav. (2025). "A Hybrid Approach for Sentiment Analysis in Online Product Ranking Using CNN and LSTM with Uncertain Lexical Term," 3rd International Conference on Communication, Security, and Artificial Intelligence (ICCSAI), Greater Noida, India, 2025, pp. 119-123, doi: 10.1109/ICCSAI64074.2025.11064059.

[12] V. K, A. Chabukswar, D. S, N. D S and S. Pahuja. (2025). "Sentiment Analysis on Product Reviews using Hybrid Deep Learning Models," International Conference on Advances in Next-Gen Computer Science (ICANCS), Bangalore, India, 2025, pp. 1-6, doi: 10.1109/ICANCS65819.2025.11376894.

[13] M. Ali Shaik and V. Vuppula. (2025). "Explainable Hybrid LSTM–BERT with SHAP for App-Review Sentiment," International Conference Sustainable Communication Networks and Application (ICSCN), Theni, India, 2025, pp. 949-954, doi: 10.1109/ICSCN67106.2025.11308579.