Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 17s (2025)
https://www.jneonatalsurg.com

OPEN ACCESS

Hybrid transformer-based deep learning model analysis for Review classification in E-commerce

M.Geetha[1], Dr.N.Vimala[2]

[1]Research scholar, Computer Science, LRG Govt Arts College for Women, Tiruppur-641604, Bharathiyar University, Email ID: mp.geetha88@gmail.com

[2]Assistant Professor & HEAD, Computer Science, Puratchi Thalaivi Amma Govt Arts & science college, Palladam-641664, Email ID: vimalananjappan@gmail.com

Cite this paper as: M.Geetha, Dr. N.Vimala, (2025) Hybrid transformer-based deep learning model analysis for Review classification in E-commerce. *Journal of Neonatal Surgery*, 14 (17s), 872-884.

ABSTRACT

As e-commerce platforms continue to grow, understanding customer sentiment has become essential for improving user experiences, refining product recommendations, and supporting decision-making processes. Sentiment classification models play a vital role in analyzing extensive customer reviews. However, accurately categorizing sentiments, particularly neutral ones, remains a significant challenge due to the overlapping nature of textual expressions. This proposed work evaluates the performance of five deep learning models, LSTM, GRU, Bi-LSTM, Bi-GRU, and a hybrid model called ResBERT with Bi-GRU for sentiment classification tasks. A major contribution of this research is the investigation of a Hybrid transformer-based deep learning model designed to improve classification accuracy, particularly in handling delicate or complex sentiments. Experimental results demonstrate that the Hybrid transformer-based deep learning model consistently outperforms the others, achieving the highest accuracy. These findings underscore the potential of the proposed model in advancing sentiment analysis, offering valuable insights for e-commerce platforms seeking to better interpret customer feedback and refine their business strategies.

Keywords: *Sentiment Classification, E-commerce Reviews, Deep Learning Models, Transformer-Based Architectures.*

1. INTRODUCTION

With the rapid expansion of e-commerce platforms, millions of customers now express their experiences and opinions through online reviews. These reviews serve as a critical resource for businesses to enhance user experience, improve product recommendations, and refine marketing strategies. However, analysing vast amounts of unstructured text data presents a significant challenge, requiring automated sentiment classification models to efficiently interpret user feedback.[1]

BERT represents a groundbreaking advancement in natural language processing, introducing pre-trained, bidirectional, and fully unsupervised language representations. Trained on a vast corpus that includes the entirety of English Wikipedia, BERT was developed by Google AI and is available as an open-source model. Its fine-tuning capabilities allow it to surpass traditional word embedding techniques, setting new standards across a wide range of NLP tasks. Meanwhile, BiGRU (Bidirectional Gated Recurrent Unit) serves as an effective feature extraction technique for sentiment analysis. As a variant of recurrent neural networks (RNNs), BiGRU processes data in both forward and backward directions, enabling it to capture richer contextual information. Its versatility makes it well-suited for various NLP applications, including sentiment classification.[11]

Deep learning models such as Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Bidirectional LSTM (BiLSTM), and Bidirectional GRU (BiGRU) have been widely used for sentiment classification. While these models achieve competitive performance, they often struggle with accurately classifying neutral sentiments due to overlapping textual expressions and contextual ambiguities. Recent advancements in hybrid transformer-based deep learning models have shown exceptional ability in capturing long-range dependencies and contextual relationships within textual data.

2. LITERATURE REVIEW

Kaur et al. (2023)designed a deep learning based consumer sentiment analysis model, leveraging NLP techniques, LSTM networks to analyse the increasing volume of textual data across digital platforms. The research methodology included pre-processing, feature extraction, and sentiment classification, where NLP techniques were utilized to refine textual data before feature extraction. The proposed model, which combines the Hybrid Feature Vector (HFV) with a LSTM classifier, significantly enhances sentiment analysis performance compared to traditional individual methodologies. The theoretical framework and experimental results, while promising, may not reflect real-world complexities and dynamics of consumer sentiment in live environments, which could differ from controlled research settings. Future work suggests integrating transformer based models such as BERT and RoBERT for enhanced contextual understanding, exploring multimodal sentiment analysis, and improving real-time sentiment prediction.[2]

Zhao et al. (2023)the authors introduce an innovative approach to public opinion sentiment analysis by leveraging multi-model fusion with transfer learning. This method aims to maximize the utility of the limited labeled data available by integrating various models' advantages to enhance the learning of sentiment features.The methodology described involves the use of the ERNIE model for generating dynamic word vector representations and a combination of TextCNN and BiGRU to extract both local and overall features from the text. The attention mechanism helps in highlighting important sentiment-related features, thereby enhancing overall model performance. Future work includes incorporating multimodal sentiment analysis (text, image, and audio), exploring domain adaptive transformers like BERT and RoBERT, and enhancing real-time sentiment analysis models for large-scale social media data processing.[3]

Kardakis et al. (2021) employed attention based models built on RNN, specifically LSTM and GRU, to enhance sentiment analysis using benchmark datasets like IMDB and SST-2. A comparative analysis revealed that integrating attention mechanisms improved classification accuracy by up to 3.5%, with models achieving a maximum accuracy of 88.2% on movie review sentiment classification tasks. These improvements highlight the effectiveness of attention based techniques in refining feature extraction and enhancing the recognition of opinions and emotions. Future work includes incorporating transformer based architectures for more advanced context modelling, exploring hybrid deep learning frameworks, and optimizing real-time sentiment analysis for large scale text data processing.[4]

Jain et al. (2023)The paper emphasizes that semantics (meaning) and sentiments (emotions) are integral to daily communication. They help convey messages with the intended tone, making accurate interpretation crucial for understanding the true meaning behind expressions. This involves teaching computers to analyze sentiments using data collected from audio, video, and text sources. The automation is designed to improve both the efficiency and accuracy of sentiment interpretation. These technologies are capable of processing and analyzing vast amounts of data, which is essential for accurately identifying and understanding sentiments.The paper mentions various ML techniques applied in sentiment classification, including Support Vector Machines (SVMs), Bayesian Networks (BNs), Decision Trees (DTs), Convolutional Neural Networks (CNNs), and K-Means Clustering. These methods are employed to classify different parts of a message into corresponding emotions, contributing to the overall sentiment analysis process.The model developed for analyzing textual data achieved impressive metrics, including an F1-score of 91%, a recall of 0.91, and a precision of 0.91. The results demonstrate that the model effectively classifies sentiments into positive, negative, and neutral categories. Future research may aim to further enhance emotion recognition accuracy by leveraging more advanced machine learning and deep learning techniques. This could include exploring alternative neural network architectures or incorporating additional data sources to improve the model’s ability to interpret complex emotional expressions.[7]

Wu et al. (2022) introduced an enhanced BERT-based model for multimodal sentiment analysis, featuring three key innovations: a Hierarchical Multi-Head Self-Attention module for improved hierarchical feature representation, a Gate Channel module that replaces BERT’s Feed Forward layer to enhance information filtering, and a tensor fusion mechanism based on self-attention to effectively integrate multimodal features while maintaining inter-modal dependencies. Experiments on the CMU-MOSI dataset demonstrated a 0.44% increase in accuracy and a 0.46% improvement in F1-score compared to the original BERT model with custom fusion. Future research directions include incorporating reinforcement learning for adaptive sentiment modeling, leveraging multimodal transformers for more effective feature fusion, and enhancing real-time sentiment analysis for broader NLP and emotion recognition applications. [12]

3. METHODOLOGY

The sentiment classification methodology follows a well-structured pipeline, starting with Dataset Collection, where customer reviews from e-commerce platforms (specifically, **Kaggle's Candes air coolers dataset**) are gathered. In the Text Pre-processing phase, the text undergoes tokenization, stop-word removal, stemming, lemmatization, and normalization to clean and standardize the data. The processed text is then transformed through Feature Extraction (TF-IDF) to convert words into numerical representations. During the Classification stage, deep learning models such as LSTM, GRU, BI-LSTM, and BI-GRU are used to analyze sentiments. A Hybrid transformer-based deep learning model, combining ResBERT with Bi-GRU, is introduced to improve classification accuracy by leveraging both transformer-based and sequential learning techniques. Finally, in the Evaluation step, model performance is assessed using Accuracy, Precision, Recall, and F1-Score to identify the most effective sentiment classification model. This pipeline ensures a robust and optimized approach for sentiment analysis in the e-commerce domain.

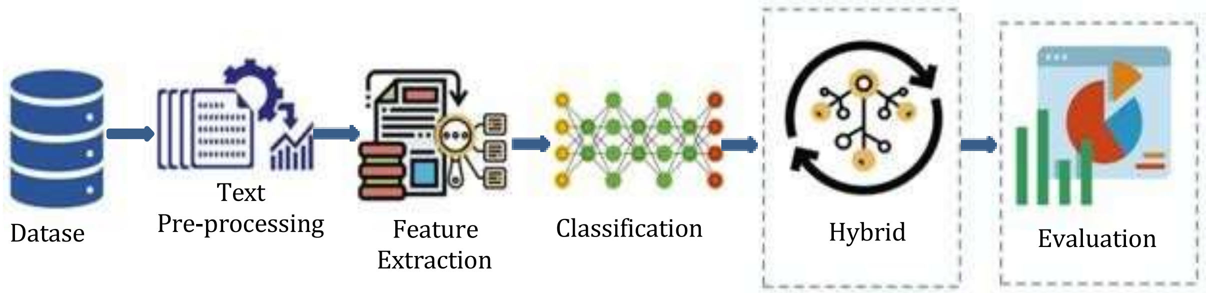

Fig.1. Hybrid Transformer-based Deep Learning Model for Sentiment Classification Workflow

3.1 Dataset Description

This dataset consists of **2,05,052** product reviews for **Candes air coolers** collected from an e- commerce platform and retrieved from Kaggle. It includes user ratings, textual reviews, sentiment classifications, and product details. The dataset is essential for sentiment analysis, opinion mining, and enhancing product recommendation systems.

Table.1. Dataset Attributes

Column Name	Description
product_name	Name and model of the air cooler.
product_price	Price of the product in INR.
Rate	Customer rating (1 to 5).
Review Summary	A short summary of the customer review.
Review	Detailed textual feedback from the customer.
Sentiment	Sentiment label (positive, negative, neutral) derived from the review.

The above table describes the dataset attributes, including product details, customer ratings, review summaries, detailed feedback, and sentiment classification.

3.2 Text Pre-processing

Pre-processing converts unstructured text into a structured format suitable for NLP. Key steps include lowercasing, stop word removal, tokenization, stemming, and lemmatization. Punctuation, special characters, HTML tags are removed, while spelling correction and emoji handling enhance accuracy. Text normalization standardizes content by removing extra spaces, accents, and converting numbers.

3.3 Feature Extraction

Feature extraction is crucial for distinguishing between positive, negative, and neutral sentiments in product reviews. A key method used is Term Frequency Inverse Document Frequency (TF-IDF), which assigns weights to words based on their significance in a review relative to the entire dataset.By reducing noise and emphasizing key terms, TF-IDF enhances feature selection.

3.4 Classification

Opinion mining, or sentiment analysis, is crucial for extracting valuable insights from product reviews, helping both businesses and consumers make informed decisions. Traditional machine learning models struggle with capturing contextual dependencies in text, leading to suboptimal classification. Advanced deep learning techniques such as BI-LSTM, Bidirectional Gated Recurrent Units (Bi-GRU) have significantly improved sentiment analysis by effectively modeling sequential data and long range dependencies. BI-LSTM enhances sentiment classification by processing information in both forward and backward directions. BI-GRU, an extension of GRU, further refines sentiment analysis by leveraging bidirectional processing and attention mechanisms to emphasize sentiment bearing words. Additionally, transformer based models like ResBERT, when integrated with Bi-GRU and residual connections (ResBERT with Bi-GRU), enhance feature representation and classification accuracy by preserving contextual embedding while mitigating gradient vanishing issues. This work explores the effectiveness of these Hybrid transformer-based deep learning model approaches, demonstrating how bidirectional architectures, attention mechanisms, and transfer learning contribute to improved sentiment classification in product reviews.

3.4 1.Bidirectional Long Short-Term Memory (BI-LSTM):

LSTM networks are essential in opinion mining for product reviews, efficiently managing sequential data and long-term dependencies. Unlike RNNs, they prevent the vanishing gradient problem by retaining key information, improving sentiment classification. Uni- LSTM processes data sequentially, while BI-LSTM captures both past and future contexts, enhancing its ability to extract deeper insights for sentiment analysis.

3.4.2 Bidirectional Gated Recurrent Unit (BI-GRU)

BI-GRU in sentiment analysis captures both past and future circumstances via bidirectional processing. This augments its capacity to discern intricate semantic links, hence enhancing sentiment categorization in product evaluations. Performance is enhanced using attention techniques that prioritize relevant features and the application of transfer learning from pre-trained models. The BI-GRU allows the model to capture contextual dependencies from both previous and future states. A GRU consists of two main gates:

- **Update gate (yt):** Controls the amount of the previous hidden state to retain.
- **Reset gate (rt):** Regulates how much of the past information should be discarded when calculating the new candidate state.

Table.2. Component and Description of BI-GRU

Component	Description
Hidden State h_t	$h_t = (1 - y_t).\ h_{t-1} + \tilde{h}_t$, h_t at time step t is computed as a linear interpolation between the previous hidden state h_{t-1} and the candidate state $\tilde{h}_t$
Candidate State $\tilde{h}_t$	$\tilde{h}_t = \tan h\ (W_h.\ x_t + r_t.\ (U_h\ .\ h_{t-1}) + b_h$ Where x_t is the input at time t, W_h and U_h are weight matrices, r_t is the reset gate, and b_h is the bias term.

When the reset gate r_t is zero, the model effectively discards past information, ensuring a dynamic and adaptive memory mechanism. The bidirectional nature of the Bi-GRU further enhances context understanding by concatenating hidden states from both forward and backward passes, leading to more comprehensive sentiment analysis.

3.4.3 ResBERT with Bi- GRU

ResBERT with Bi-GRU integrates attention mechanisms and transfer learning, utilizing ResBERT for contextual embedding, Bi-GRU for sequential dependencies, and residual connections for enhanced feature representation. This approach effectively captures sentiment nuances in product reviews, improving classification accuracy and interpretability.

Table.3. Component of ResBERT with Bi- GRU

Component	Function	Computation
ResBERT Embedding	Captures contextual meaning of words based on surrounding context.	xi = ResBERT (wi), where wi the input is token, and xi is the corresponding word embedding
Bi-GRU Layer	Captures sequential dependencies in both the forward and backward directions.	hi = [hFi‖hBi]
Residual Connection	Enhances gradient flow and retains meaningful information.	hres = hi + xi, where hi represent the Bi-GRU output and xi is the original ResBERT embedding.
Attention Mechanism	Focuses on sentiment bearing words by assigning higher weights to influential tokens.	ai = softmax (Whi), where W is a learnable parameter matrix.

Context Vector	Aggregates weighted hidden states to represent overall sentiment information.	C = $\sum_{i=1}^{T}$ aihi, where T is the sequence length.

The final classification step uses the context vector from attention weighted hidden states, processed through ResBERT embedding, Bi-GRU, residual connections, and attention mechanisms. A fully connected layer with a sigmoid or softmax activation function is used to predict the sentiment label. enhancing accuracy and interpretability by focusing on influential words and preserving rich contextual information.

Algorithm: ResBERT with Bi-GRU for Sentiment Analysis

Input: Reviews dataset RD

Output: Sentiment class (positive, negative, neutral)

1**. Preprocessing and Zero-Shot Classification**

- For each review R in RD :

Calculate polarity scores (Pos_s, Neu_s, Neg_s) using zero-shot BERT classification.

- If Pos_s > Neg_s and Pos_s > Neu_s :

Assign label = positive.

- Else if Neg_s > Pos_s and Neg_s > Neu_s :

Assign label = negative.

- Else: Assign label = neutral.

2. **Tokenization and Embedding Extraction**

- For each R in preprocessed RD :

Perform WordPiece tokenization to generate token IDs and attention masks.

Extract word embedding vectors Vec using the ResBERT model.

3. **Data Splitting**

- Divide the dataset into training (T_{train}) and testing (T_{text}) sets.

4. Model Training

- For each R in T_{train} :

Pass embedding through a Bidirectional GRU (Bi-GRU) layer to capture sequential dependencies $h_i = [h_{Fi}][h_{Bi}]$

- Apply residual connections between **ResBERT** embedding and Bi-GRU outputs:

$hres = hi + xi$

- Use an attention mechanism to compute attention scores:

$a_i = softmax\,(Wh_i)$

- Aggregate weighted hidden states to form the context vector.

  $\sum_{i=1}^{T} a_i h_i$

- Pass the context vector through a dropout layer (dropout = 0.2).
- Flatten the output and pass it through a dense layer.
- Use a sigmoid activation to calculate probabilities for sentiment labels.

5. Model Testing

- For each R in T_{test}:

Classify R into sentiment classes using the trained model.

Output the predicted sentiment (positive, negative, and neutral).

4. RESULT AND DISCUSSION

This section presents the evaluation of sentiment classification models based on feature analysis and performance metrics. The proposed Hybrid transformer-based deep learning model (ResBERT + Bi-GRU) is compared with LSTM, GRU, BI-LSTM, and BI-GRU to determine its effectiveness in sentiment classification. Performance is assessed using key evaluation metrics, including accuracy, precision, recall, F1-score, and confusion matrix analysis.

- **4.1 Experimentation details:**
- The experiment was conducted on a macOS machine with a 64-bit architecture. The machine features an Intel 2.6GHz 8-core i7 CPU, 16GB 2400MHz DDR4 memory, and a Radeon Pro 560X 4GB GPU. The programs used throughout the experiment were developed using Python 3.8 and executed within the Anaconda environment. This hardware and software configuration provided a robust and efficient platform for running the experiments and analyzing the results.

4.2.Feature analysis

Feature analysis is essential for understanding data patterns and improving model performance in sentiment classification. This process examines key attributes such as word frequency, review length, sentiment distribution, and common phrases to extract meaningful insights. Identifying prevalent words and sentiment-specific terms helps refine text preprocessing, improve classification accuracy, and address potential biases in the dataset. It enhances feature selection, optimizing machine learning models for reliable and interpretable sentiment predictions.

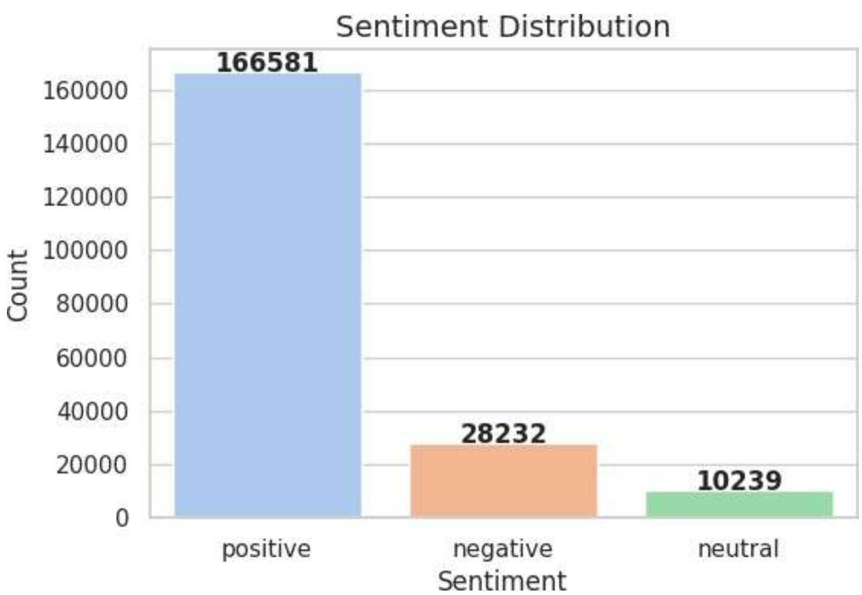

Fig.2. Sentiment Distribution of the Dataset

The figure presents the sentiment distribution of the dataset, highlighting the dominance of positive reviews (1,66,581), followed by negative reviews (28,232) and neutral reviews (10,239). This imbalance suggests a strong inclination toward positive sentiments, which may impact model training and classification performance. Addressing this disparity through resampling or weighted loss functions can improve model robustness in sentiment analysis.

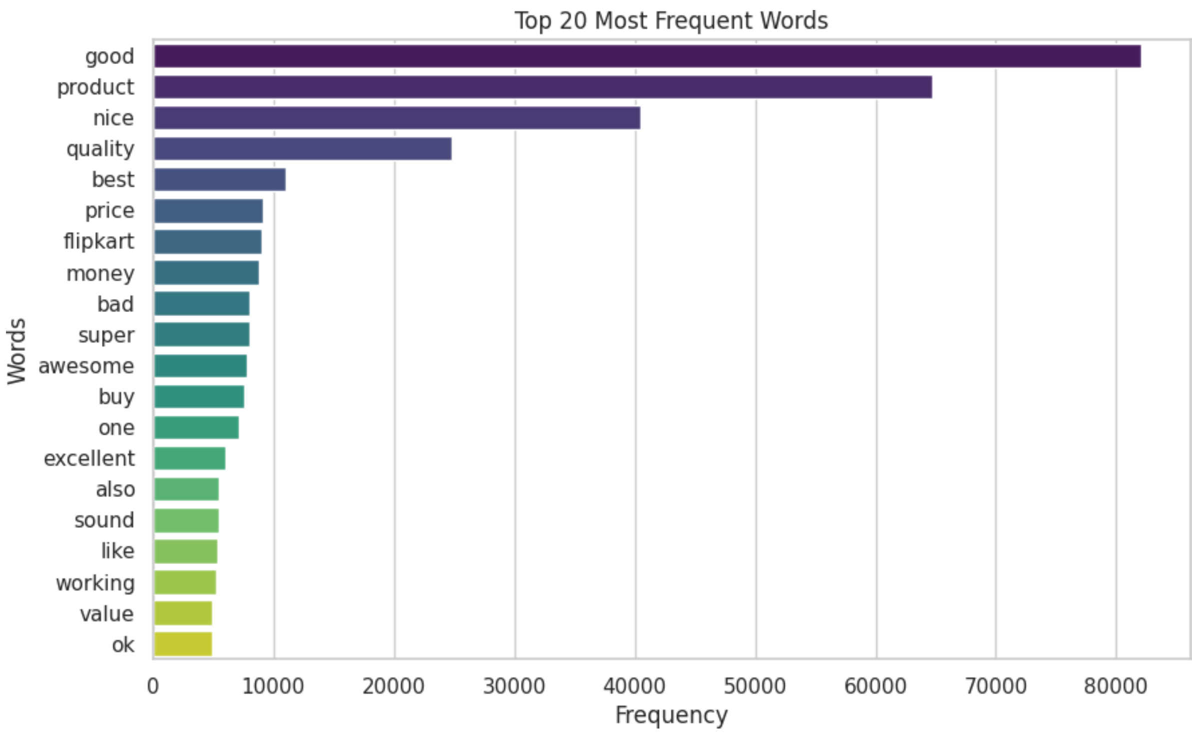

Fig.3. Top 20 Most Common Words in Customer Reviews

The figure illustrates the top 20 most common words in the dataset, with "good," "product," and "nice" being the most commonly used terms. These high-frequency words indicate a strong emphasis on product quality and user satisfaction. The presence of terms like "bad" and "money" suggests a mix of both positive and negative sentiments, highlighting key aspects influencing customer reviews.

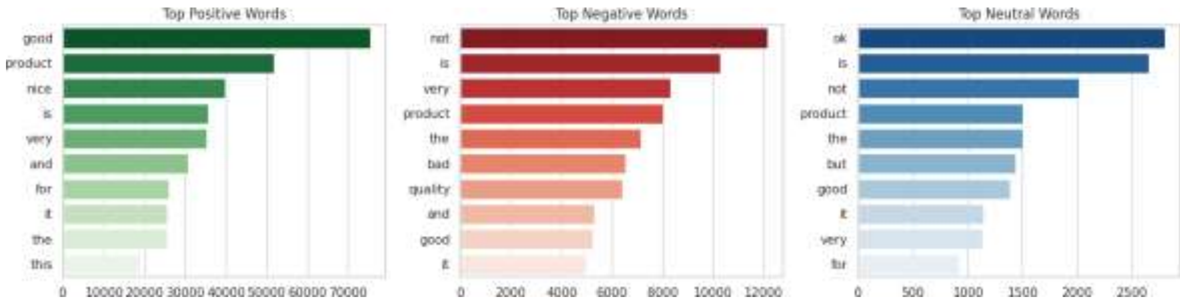

Fig.4. Most Frequent Words across Sentiment Categories

The figure presents the most frequent words across different sentiment categories: positive, negative, and neutral. In positive sentiment, words like "good," "product," and "nice" dominate, reflecting customer satisfaction. Negative sentiment is characterized by words such as "not," "bad," and "quality," indicating dissatisfaction or criticism. Neutral sentiment includes words like "ok," "is," and "not," suggesting mixed or ambiguous feedback. This analysis provides insights into customer perception and common themes in product reviews.

Fig.5. Word cloud in Product Reviews

The figure presents word cloud for positive, negative, and neutral reviews, visualizing the most commonly used words in each sentiment category. Positive reviews prominently feature words like "good," "nice," "product," and "quality," indicating customer satisfaction. Negative reviews highlight terms such as "bad," "don't buy," "worst," and "waste money," reflecting dissatisfaction and product issues. Neutral reviews contain words like "ok," "quality," "average," and "working," suggesting mixed or moderate opinions. This analysis provides insights into key themes and sentiments expressed in customer feedback.

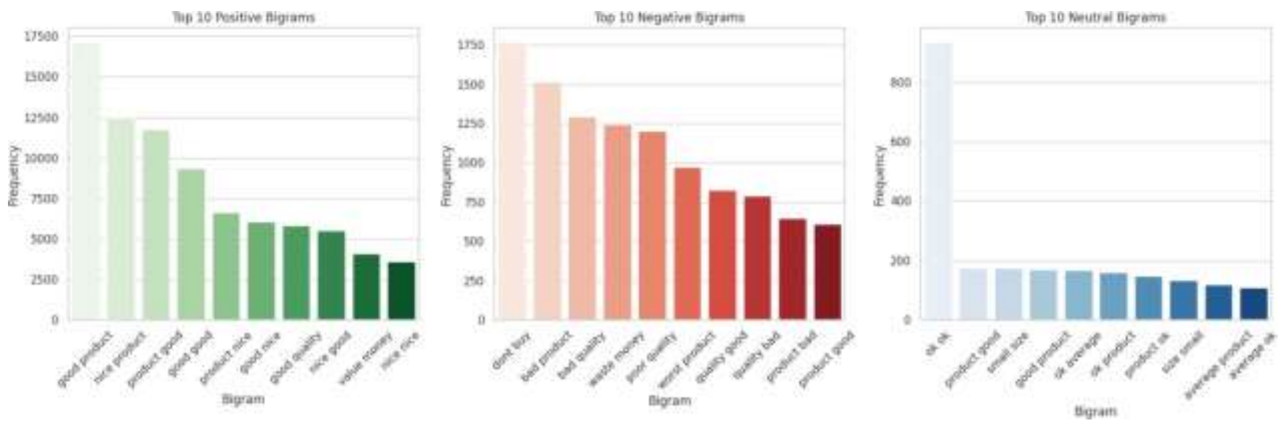

Fig.6. Top 10 Bigrams across Positive, Negative, and Neutral Reviews

Feature analysis offers valuable insights into data distribution, sentiment patterns, and key linguistic trends, improving the model's accuracy and robustness in sentiment classification.

4.3. Evaluation Metrics

This standard employs assessment metrics like as accuracy, precision, recall and F1-score and the confusion matrix to analyse model performance. The following measures are elaborated upon below.

Table.4. Evaluation Metrics

Metric	**Definition**	**Formula**
Accuracy	Calculates the ratio of correct predictions made by the model.	$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$
Precision	Assesses the model's ability to make accurate positive predictions.	$Precision = \frac{TP}{TP + FN}$
Recall	Evaluates the model's ability to accurately identify all relevant positive instances.	$\text{Recall}= \frac{\text{TP}}{\text{TP+FN}}$
F1-Score	The harmonic mean of precision and recall, providing a balance between the two metrics.	$\text{F-score}=2 \text{ X } \frac{\text{Precision X Recall}}{\text{Precision Recall}}$

Table.5. Precision Comparison of Sentiment Classification Models

Category	**LSTM**	**GRU**	**BI-LSTM**	**BI-GRU**	**ResBERT with Bi- GRU**
negative	0.90	0.88	0.90	0.89	**0.91**
neutral	0.79	0.78	0.74	0.74	**0.85**
positive	0.96	0.97	0.97	0.98	**0.99**
Avg	0.88	0.87	0.87	0.87	**0.91**

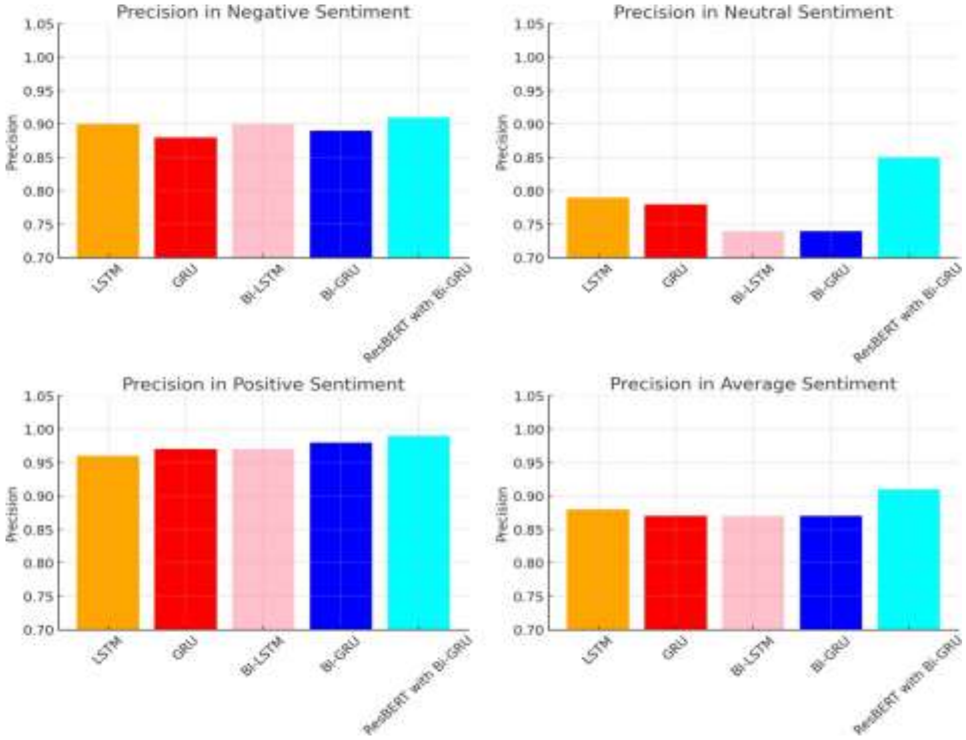

Fig.7. Precision Comparison Across Sentiment Categories for Different Models

Table.6. Recall Comparison of Sentiment Classification Models

Category	LSTM	GRU	BI-LSTM	BI-GRU	ResBERT with Bi- GRU
negative	0.88	0.87	0.87	0.89	**0.91**
neutral	0.46	0.51	0.51	0.60	**0.68**
positive	0.99	0.97	0.99	0.98	**0.99**
Avg	0.78	0.79	0.79	0.82	**0.86**

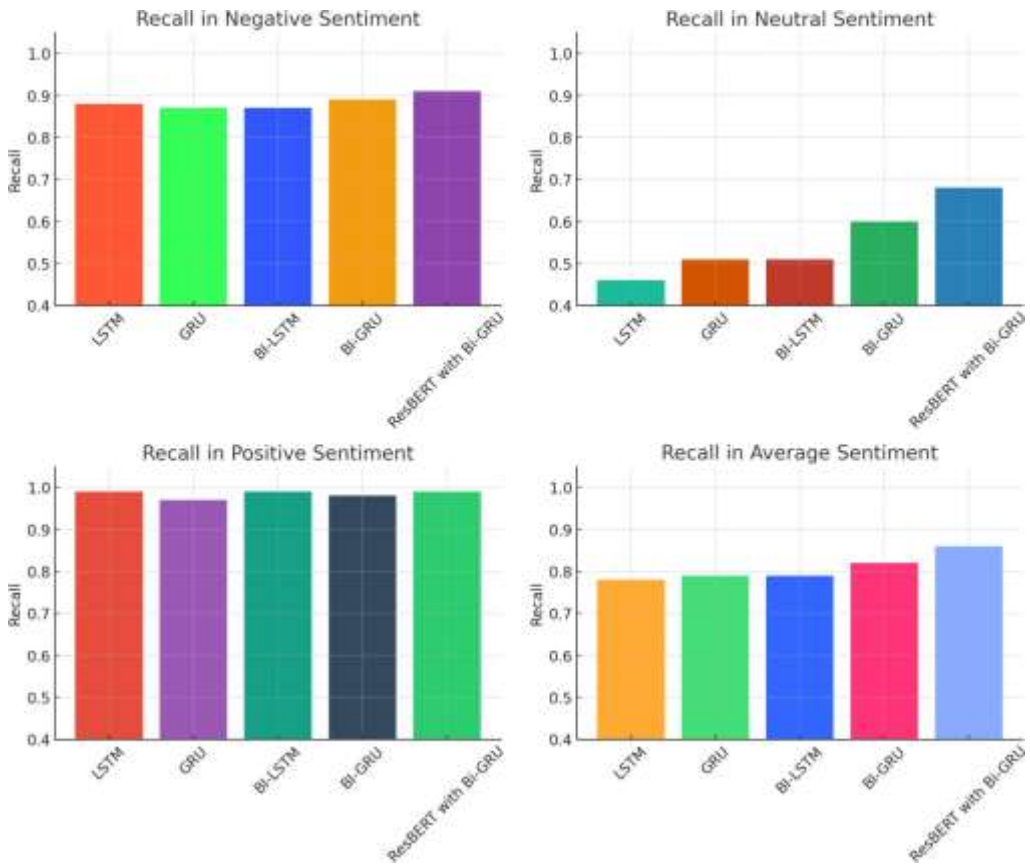

Fig.8. Recall Comparison Across Sentiment Categories for Different Models

Table.7. F1-Score Comparison of Sentiment Classification Models

Category	LSTM	GRU	BI-LSTM	BI-GRU	ResBERT with Bi- GRU
negative	0.88	0.87	0.87	0.89	**0.91**
neutral	0.46	0.51	0.51	0.60	**0.68**
positive	0.99	0.97	0.99	0.98	**0.99**
Avg	0.78	0.79	0.79	0.82	**0.86**

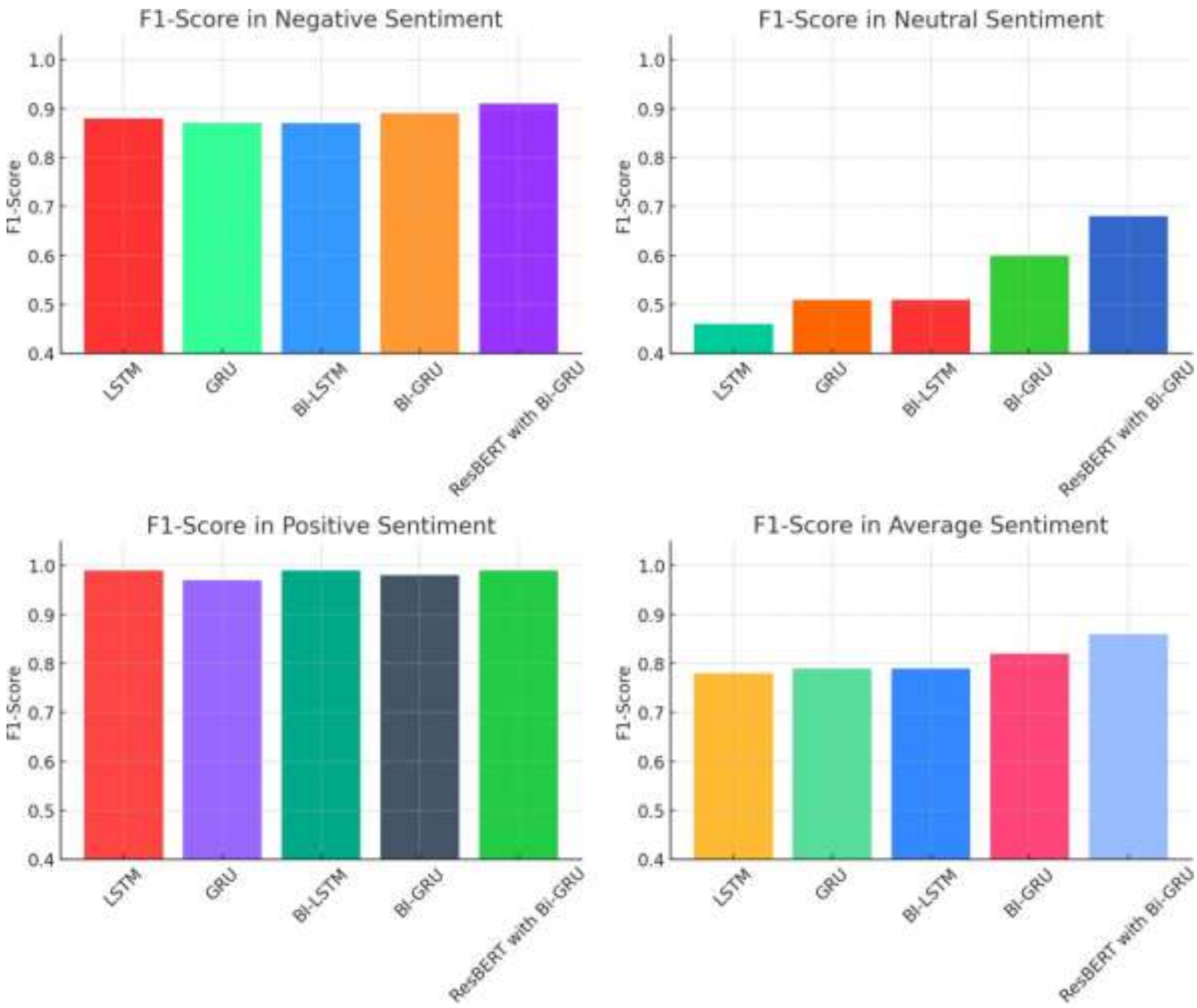

Fig.9. F1-score Comparison Across Sentiment Categories for Different Models

Table.8. Accuracy and Loss Comparison of Sentiment Classification Models

Model	**Accuracy**	**Training Loss**	**Validation Loss**
LSTM	0.95	0.1451	0.1476
GRU	0.94	0.1483	0.1512
BiLSTM	0.95	0.1340	0.1491
BiGRU	0.95	0.1206	0.1559
ResBERT + BiGRU	**0.96**	**0.0983**	**0.1125**

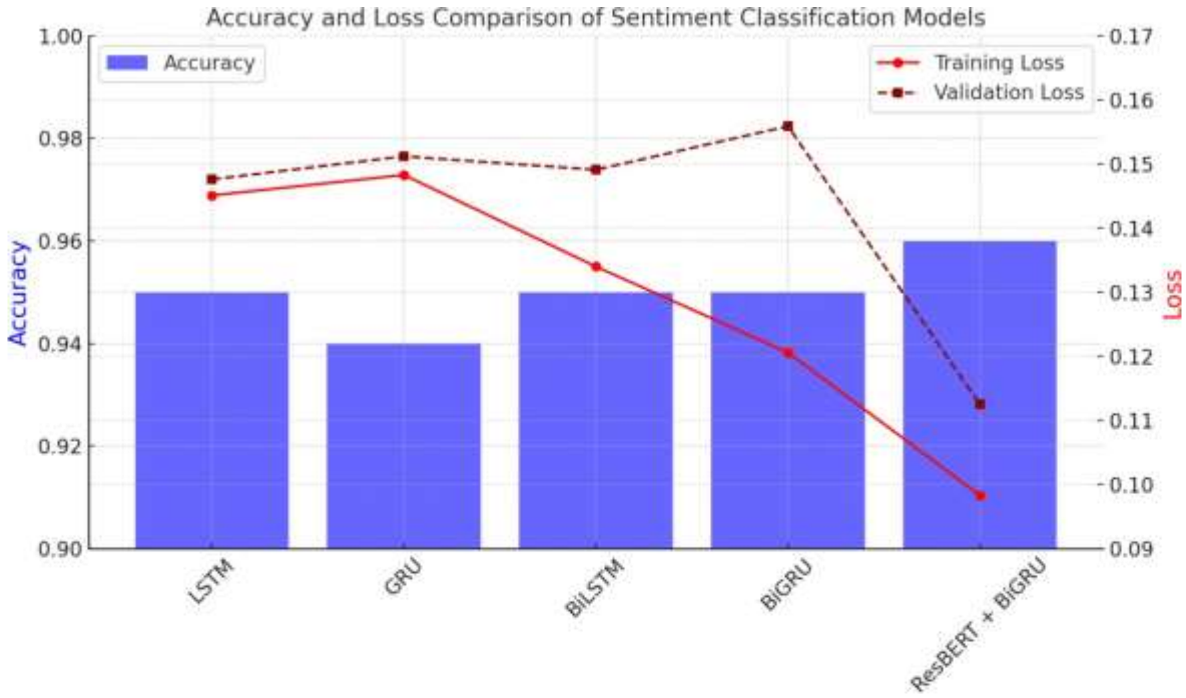

Fig.10. Accuracy and Loss Comparison of Sentiment Classification Models

A confusion matrix offers a detailed analysis of model predictions by comparing the predicted and actual class labels, aiding in the evaluation of classification performance across various categories.

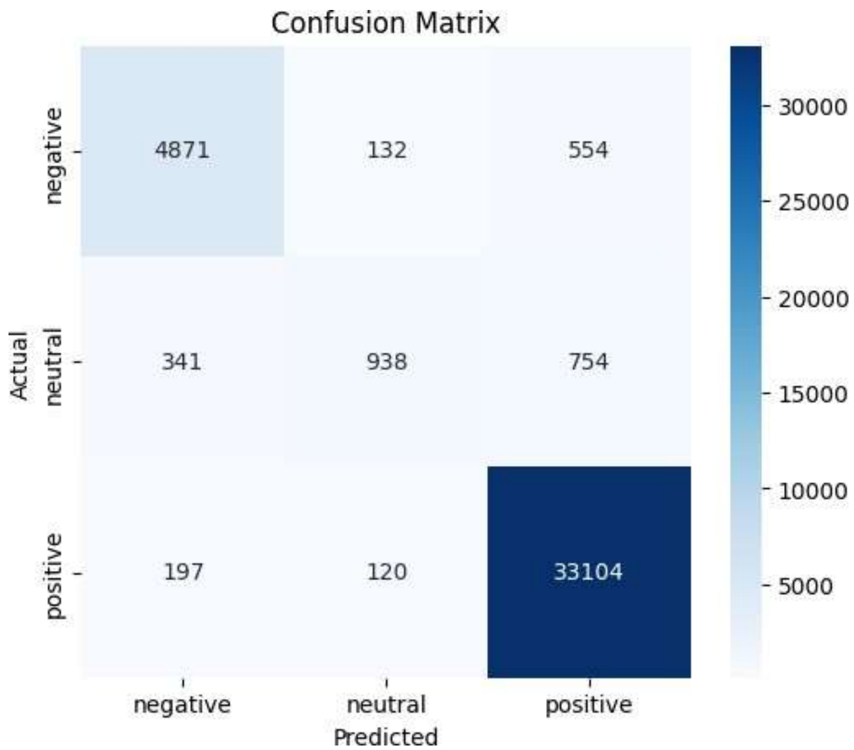

Fig.11. Confusion Matrix for ResBERT + Bi-GRU (Proposed Model)

Above figure presents the confusion matrix for the proposed ResBERT + BiGRU model, demonstrating strong classification performance across all sentiment classes. While it correctly classifies most instances, Neutral sentiment remains the most challenging category, with some misclassifications into Negative and Positive. The confusion matrix demonstrates the classification performance of a sentiment analysis model across negative, neutral, and positive sentiment categories.

5. CONCLUSION

This work investigates sentiment classification in e-commerce product reviews using deep learning models, including LSTM, GRU, BI-LSTM, BI-GRU, and the proposed Hybrid transformer-based deep learning model (ResBERT + Bi-GRU). The results show that the Hybrid transformer-based deep learning model outperforms all other models, achieving the highest accuracy of 96%, proving to be an effective solution for sentiment analysis. The research highlights the challenges in neutral sentiment classification, where misclassifications were more frequent compared to positive and negative sentiments. Feature analysis, including word frequency and bigram analysis, provides insights into key themes influencing customer reviews. The confusion matrix analysis further confirms the model's robustness, especially in accurately detecting positive sentiments with high precision and recall. These findings emphasize the effectiveness Hybrid transformer-based deep learning model in improving sentiment classification performance. Future work can focus on handling class imbalance using data augmentation or advanced resampling techniques to further improve the model's performance in sentiment analysis tasks.

REFERENCES

[1] H. He, G. Zhou, and S. Zhao, “Exploring E-Commerce Product Experience Based on Fusion Sentiment Analysis Method,” IEEE Access, vol. 10, pp. 1–12, Oct. 2022.

[2] G. Kaur and A. Sharma, "A deep learning-based model using hybrid feature extraction approach for consumer sentiment analysis," Journal of Big Data, vol. 10, p. 5, 2023.

[3] Z. Zhao, W. Liu, and K. Wang, "Research on sentiment analysis method of opinion mining based on multi-model fusion transfer learning," Journal of Big Data, vol. 10,p. 155, 2023.

[4] Kardakis, S.; Perikos, I.; Grivokostopoulou, F.; Hatzilygeroudis, I. Examining Attention Mechanisms in Deep Learning Models for Sentiment Analysis. Appl. Sci. 2021, 11, 3883.

[5] Li, H.; Lu, Y.; Zhu, H. Multi-Modal Sentiment Analysis Based on Image and Text Fusion Based on Cross-Attention Mechanism. Electronics 2024, 13, 2069.

[6] K. Qiu, Y. Zhang, J. Zhao, S. Zhang, Q. Wang, and F. Chen, "A Multimodal Sentiment Analysis Approach Based on a Joint Chained Interactive Attention Mechanism," Electronics, vol. 13, no. 1922, 2024.

[7] Jain, R., Rai, R.S., Jain, S. et al. Real time sentiment analysis of natural language using multimedia input. Multimed Tools Appl 82, 41021–41036 (2023).

[8] Z. Zhao, H. Zhu, Z. Xue, Z. Liu, J. Tian, M. C. H. Chua, and M. Liu, "An image- text consistency driven multimodal sentiment analysis approach for social media," Information Processing & Management, vol. 56, no. 6, 2019, Art. no. 102097.

[9] S. Ghaffarian, J. Valente, M. van der Voort, and B. Tekinerdogan, "Effect of Attention Mechanism in Deep Learning-Based Remote Sensing Image Processing: A Systematic Literature Review," Remote Sens., vol. 13, no. 2965, 2021.

[10] M. S. Md Suhaimin, M. H. Ahmad Hijazi, E. G. Moung, P. N. E. Nohuddin, S. Chua, and F. Coenen, "Social media sentiment analysis and opinion mining in public security: Taxonomy, trend analysis, issues and future directions," Journal of King Saud University - Computer and Information Sciences, vol. 35, no. 10, p. 101776, 2023.

[11] Rana, M. R. R., Nawaz, A., Rehman, S. U., Abid, M. A., Garayevi, M., & Kajanová, J. (2025). BERT-BiGRU-Senti-GCN: An Advanced NLP Framework for Analyzing Customer Sentiments in E-Commerce. *International Journal of Computational Intelligence Systems*, *18*(1), 21.

[12] S. Sari and M. Kalender, "Sentiment Analysis and Opinion Mining Using Deep Learning for the Reviews on Google Play," Department of Computer Engineering, Yeditepe University, Istanbul, Turkey, Feb. 2021.

[13] Wu, J.; Zhu, T.; Zheng, X.; Wang, C. Multi-Modal Sentiment Analysis Based on Interactive Attention Mechanism. Appl. Sci. 2022, 12, 8174.

[14] .https://www.kaggle.com/datasets/niraliivaghani/flipkart-product-customer-reviews- dataset/data

[15] F. Xing, H. Ashraf, U. Ihsan, N. A. Khan and R. Bahuguna, "Social Media Text Sentiment Analysis: Exploration Of Machine Learning Methods," 2024 IEEE 1stKarachi Section Humanitarian Technology Conference (KHI-HTC), Tandojam, Pakistan, 2024, pp. 1-6.

[16] A. Koufakou, "Deep learning for opinion mining and topic classification of course reviews," Education and Information Technologies, vol. 29, pp. 2973–2997, 2024.