

International Journal of Engineering Research and Science & Technology



www.ijerst.org

ISSN : 2319-5991

Vol. 22 No. 2(1) (2026)



ijerst.editor@gmail.com
editor@ijerst.com

Research Paper

An SBERT and DNN-Enhanced Boosted Rules Classifier Framework for Accurate Multi-Class Sentiment Analysis of Product Reviews

E. Sravanthi¹, Shaik Zaheer², Nematikuntla Vennela², Polavena Ajay²

¹Assistant Professor, ²UG Student, ^{1,2}Department of Computer Science and Engineering

^{1,2}Kommuri Pratap Reddy Institute of Technology, Ghanpur, Ghatkesar, 501301, Telangana, India.

*Correspondence: E. Sravanthi (sravanthieega26@gmail.com)

ABSTRACT

The global e-commerce market is projected to surpass USD 7 trillion by 2030, with more than 80% of consumers relying heavily on product reviews to inform their purchasing decisions. Despite this reliance, manual sentiment analysis of such reviews is inherently limited in scalability and consistency, often leading to delayed insights and subjective or inaccurate interpretations. To overcome these limitations, this study introduces a comprehensive Natural Language Processing (NLP) framework designed for automated sentiment classification using a labeled product review dataset. The proposed approach begins with extensive preprocessing and Exploratory Data Analysis (EDA) to clean, standardize, and analyze the underlying data distribution. For advanced semantic representation, Sentence Bidirectional Encoder Representations from Transformers (SBERT) is utilized to generate context-rich embeddings, enabling the capture of nuanced linguistic patterns beyond the capabilities of traditional feature extraction methods. To address the issue of class imbalance across sentiment categories, the Synthetic Minority Over-sampling Technique (SMOTE) is applied to create synthetic samples, thereby ensuring balanced and effective model training. In contrast to conventional machine learning models such as Random Forest Classifier (RFC), Light Gradient Boosting Machine (LGBM), and Extreme Gradient Boosting (XGBoost), the proposed framework integrates a Deep Neural Network (DNN)-based feature selection mechanism with a Boosted Rules Classifier (BRC). This hybrid architecture not only enhances predictive accuracy but also improves model interpretability. The system categorizes customer sentiments into three distinct classes: Negative, Neutral, and Positive. Experimental findings demonstrate that the proposed approach achieves higher classification accuracy with reduced bias, validating its effectiveness. The framework offers a scalable and reliable solution for sentiment analysis, empowering businesses to make data-driven decisions in areas such as product development, marketing strategy optimization, and customer relationship management (CRM), ultimately enabling more responsive and customer-centric operations.

Key words: Customer relationship management, Product reviews, Contextual Word Embeddings, Multi-Head Self-Attention, Semantic Textual Similarity

1. INTRODUCTION

Over the past decade, the volume of product reviews available on traditional retail and e-commerce platforms has grown significantly, with contributions from both professional reviewers and everyday consumers. This surge in user-generated content has made online reviews a central factor in shaping

purchasing behavior, as consumers increasingly depend on them to reduce uncertainty and make informed decisions. In addition to textual feedback, supplementary content such as product images further enhances the evaluation process. Research findings suggest that a vast majority of users consider customer reviews before finalizing a purchase, with nearly 90% of consumers actively consulting them. As a result, product reviews have become critically important for both consumers and businesses alike. For consumers, they serve as a valuable source of insight into product quality, performance, and suitability, helping determine whether a product aligns with their expectations. For online retailers, reviews represent a powerful tool for influencing customer perception, where effective management of feedback and the promotion of positive reviews can lead to increased customer engagement, improved brand reputation, and ultimately higher profitability.



Fig. 1. Product Sentiment Analysis.

Amazon, a global technology leader in the e-commerce domain, relies heavily on customer reviews as a critical component of the purchasing process, offering genuine insights into product quality and user experience. However, manually analyzing the ever-growing volume of reviews is both inefficient and time-intensive. These reviews are equally valuable to sellers, as they enable product differentiation, strengthen marketing strategies, and support business growth through improved customer engagement. In modern digital commerce, online ratings and reviews play a vital role in shaping consumer perceptions, allowing users to quickly assess a product's value while significantly influencing buying decisions—either positively or negatively. Previous studies have leveraged Amazon review datasets to predict review helpfulness, where each review is accompanied by feedback from other users regarding its usefulness.

The importance of product ratings has further intensified in the e-commerce landscape, particularly during the COVID-19 pandemic, which accelerated the global shift toward online shopping. This transition led to a substantial surge in user-generated reviews, with reports indicating a growth of approximately 43%, a trend that continues to rise. Reviews and ratings serve as efficient decision-support tools, enabling consumers to save time while making informed choices. At the same time, businesses utilize review analytics to gain deeper insights into customer preferences, identify areas for improvement, and enhance product performance. Consequently, researchers have widely explored Machine Learning (ML) techniques to predict review helpfulness by uncovering patterns and extracting meaningful insights from large-scale data.

Despite these advancements, the rapid expansion of online review platforms has resulted in information overload, making it increasingly challenging for consumers to accurately evaluate product quality. Furthermore, the widespread use of social media has blurred the boundary between authentic customer feedback and promotional or biased content, leading to a rise in misleading or manipulated reviews. Typically, the usefulness of a review is determined through user voting systems; however, the growing scale of data complicates this process. With studies suggesting that nearly 95% of consumers read online reviews before making purchases—and that higher-rated products can drive up to an 18% increase in sales—the influence of reviews on consumer behavior is undeniable. Given the massive influx of daily reviews, manual sentiment analysis is no longer a viable approach. Therefore, there is a pressing need for scalable, intelligent, and automated solutions capable of analyzing sentiments effectively. In this context, sentiment analysis has emerged as a powerful tool for extracting actionable insights related to customer satisfaction, product performance, and service quality, ultimately supporting better decision-making and fostering trust in the e-commerce and retail ecosystem. The increasing volume of online product reviews makes it difficult for businesses and consumers to manually analyze customer opinions and sentiments. Traditional sentiment analysis approaches often struggle to accurately capture contextual meaning and handle imbalanced datasets. Additionally, many existing systems lack efficient integration between advanced NLP models and deployable web applications. Therefore, there is a need to develop an automated sentiment analysis system that can effectively extract features from textual data, accurately classify sentiments, and provide a scalable solution for real-time analysis.

2. LITERATURE SURVEY

Liu, et al. [1] developed of the Internet, agricultural products have entered e-commerce platforms, and agricultural product reviews have become an important reference for consumers when purchasing agricultural products. However, due to the characteristics of different lengths, rich context-sensitive information, and multi-level information in the sentences of agricultural product reviews, the existing sentiment analysis methods cannot perform well enough to identify the sentiment tendency. To address this issue, they abstract the problem as a binary classification task to extract consumers' sentiment orientation by proposing a new method. This method utilizes an attention mechanism to assign different weights to different key information in the sentence, thereby extracting abundant semantic information from the sentence. The design of the long short-term memory (LSTM) gate can effectively solve the problem of extracting long sequences and context-related information. The proposed model achieves superior results on two agricultural product datasets compared to other baseline models, providing guidance for merchants to improve agricultural product quality and enhance customer satisfaction.

Khan, et al. [2] purposed to evaluate the performance of various word embeddings for Roman Urdu and English dialects using the CNN-LSTM architecture with traditional machine learning classifiers. They introduce a novel deep learning architecture for Roman Urdu and English dialect SA based on two layers: LSTM for long-term dependency preservation and a one-layer CNN model for local feature extraction. To obtain the final classification, the feature maps learned by CNN and LSTM are fed to several machine learning classifiers. Various word embedding models support this concept. Extensive tests on four corpora show that the proposed model performs exceptionally well in Roman Urdu and English text sentiment classification, with an accuracy of 0.904, 0.841, 0.740, and 0.748 against MDPI, RUSA, RUSA-19, and UCL datasets, respectively. The results show that the SVM classifier and the Word2Vec CBOW (Continuous Bag of Words) model are more beneficial options for Roman Urdu sentiment analysis, but that BERT word embedding, two-layer LSTM, and SVM as a

classifier function are more suitable options for English language sentiment analysis. The suggested model outperforms existing well-known advanced models on relevant corpora, improving the accuracy by up to 5%.

Serna, et al. [3] leveraged such User Generated Content to obtain a high accuracy sentiment analysis model which automatically analyses the negative and positive opinions expressed in the transport domain. In order to develop such model, they have semiautomatically generated an annotated corpus of opinions about transport, which has then been used to fine-tune a large pretrained language model based on recent deep learning techniques. Their empirical results demonstrate the robustness of their approach, which can be applied to automatically process massive amounts of opinions about transport. They believe that their method can help to complement data from official statistics and traditional surveys about transport sustainability. Finally, apart from the model and annotated dataset, they also provide a transport classification score with respect to the sustainability of the transport types found in the use case dataset.

Ghatora, et al. [4] explored the application of both traditional machine learning and pre-trained LLMs for automated sentiment analysis of customer product reviews. The motivation behind this work lies in the demand for more nuanced understanding of consumer sentiments that can drive data-informed business decisions. In this research, they applied machine learning-based classifiers, i.e., RFC, Naive Bayes, and Support Vector Machine, alongside the GPT-4 model to benchmark their effectiveness for sentiment analysis. Traditional models show better results and efficiency in processing short, concise text, with SVM in classifying sentiment of short length comments. However, GPT-4 showed better results with more detailed texts, capturing subtle sentiments with higher precision, recall, and F1 scores to uniquely identify mixed sentiments not found in the simpler models. Conclusively, this study shows that LLMs outperform traditional models in context-rich sentiment analysis by not only providing accurate sentiment classification but also insightful explanations. These results enable LLMs to provide a superior tool for customer-centric businesses, which helps actionable insights to be derived from any textual data.

Ahanin, et al. [5] presented a hybrid feature extraction model using human-engineered features combined with deep learning-based features for emotion classification in English text. The proposed model uses data augmentation, captures contextual information, integrates knowledge from lexical resources, and employs deep learning models, including Bidirectional Long Short-Term Memory (Bi-LSTM) and Bidirectional Encoder Representation and Transformer (BERT), to address the issues mentioned above. The proposed model with hybrid features attained the highest Jaccard accuracy on two of the benchmark datasets, with 68.40% on SemEval-2018 and 53.45% on the GoEmotions dataset.

Chu, et al. [6] analyzed and its subtask research is a hot topic, which has very important significance. The existing sentiment analysis methods based on sentiment lexicon and machine learning take into account contextual semantic information, but these methods still lack the ability to utilize context information, so they cannot effectively encode context information. Inspired by the concept of density matrix in quantum mechanics, they propose a sentiment analysis method, named Complex-valued Quantum-enhanced Long Short-term Memory Neural Network (CQLSTM). It leverages complex-valued embedding to incorporate more semantic information and utilizes the Complex-valued Quantum-enhanced Long Short-term Memory Neural Network for feature extraction. Specifically, a complex-valued neural network based on density matrix is used to capture interactions between words (i.e., the correlation between words). Additionally, the Complex-valued

Quantum-enhanced Long Short-term Memory Neural Network, which is inspired by the quantum measurement theory and quantum long short-term memory neural network, is developed to learn interactions between sentences (i.e., contextual semantic information).

Tan, et al. [7] proposed a novel hybrid model for sentiment analysis. The model leverages the strengths of both the Transformer model, represented by the Robustly Optimized BERT Pretraining Approach (RoBERTa), and the Recurrent Neural Network, represented by Gated Recurrent Units (GRU). The RoBERTa model provides the capability to research the texts into a discriminative embedding space through its attention mechanism, while the GRU model captures the long-range dependencies of the embedding and addresses the vanishing gradients problem. To overcome the challenge of imbalanced datasets in sentiment analysis, this paper also proposes the use of data augmentation with word embeddings by over-sampling the minority classes. This enhances the representation capacity of the model, making it more robust and accurate in handling the sentiment classification task. The proposed RoBERTa-GRU model was evaluated on three widely used sentiment analysis datasets: IMDB, Sentiment140, and Twitter US Airline Sentiment. The results show that the model achieved an accuracy of 94.63% on IMDB, 89.59% on Sentiment140, and 91.52% on Twitter US Airline Sentiment. These results demonstrate the effectiveness of the proposed RoBERTa-GRU hybrid model in sentiment analysis.

Mao, et al. [8] represented for capturing semantic regularities between words in natural language processing (NLP) tasks. They propose sentiment-aware word embedding for emotional classification, which consists of integrating sentiment evidence within the emotional embedding component of a term vector. They take advantage of the multiple types of emotional knowledge, just as the existing emotional lexicon, to build emotional word vectors to represent emotional information. Then the emotional word vector is combined with the traditional word embedding to construct the hybrid representation, which contains semantic and emotional information as the inputs of the emotion classification experiments. Their method maintains the interpretability of word embeddings, and leverages external emotional information in addition to input text sequences. Extensive results on several machine learning models show that the proposed methods can improve the accuracy of emotion classification tasks.

Trisna, et al. [9] dealt with large amounts of data, making sentiment classification necessary. Automating sentiment classification involves text classification processes, commonly performed using deep learning methods. The classification process using deep learning models is closely tied to text representation. This step is critical as it affects the quality of the data being processed by the deep learning model. Traditional text representation methods often overlook the contextual meaning of sentences, leading to potential misclassification by the model. In this study, they propose a novel fusion text representation model, GloWord_biGRU, designed to enhance the contextual understanding of sentences for sentiment classification. Firstly, they combine the advantages of GloVe and Word2Vec to obtain richer and more meaningful word representations. GloVe provides word representations based on global frequency statistics within a large corpus, while Word2Vec generates word vectors that capture local contextual relationships. By integrating these two approaches, they enhance the quality of word representations used in their model. During the classification stage, they employ biGRU, considering the use of fewer parameters, which consequently reduces computational requirements. They evaluate the proposed model using the IMDB dataset. Several scenarios demonstrate that their proposed model achieves superior performance, with an F1 score of 90.21%.

Oprea, et al. [10] focused on fine-grained emotion classification using core emotions. By identifying specific emotions rather than sentiment polarity, they enable more actionable insights for e-commerce and app development, supporting strategies such as feature refinement, marketing personalization and proactive customer engagement. They leverage the Hugging Face Emotions dataset and adopt a two-phase modeling approach. In the first phase, they use a pre-trained DistilBERT model as a feature extractor and evaluate multiple classical classifiers (Logistic Regression, Support Vector Classifier, RFC) to establish performance baselines. In the second phase, they fine-tune the DistilBERT model end-to-end using the Hugging Face Trainer API, optimizing classification performance through task-specific adaptation. Training is tracked using the Weights & Biases (wandb) API. Comparative analysis highlights the substantial performance gains from fine-tuning, particularly in capturing informal or noisy language typical in user reviews. The final fine-tuned model is applied to a dataset of customers' reviews, identifying the dominant emotions expressed. Their results demonstrate the practical value of emotion-aware analytics in uncovering the underlying "why" behind user sentiment, enabling more empathetic decision-making across product design, customer support and user experience (UX) strategy.

3. PROPOSED METHODOLOGY

In the current digital ecosystem, large volumes of product reviews and user-generated content are continuously generated across online platforms. Extracting meaningful insights from this data is essential for understanding customer sentiment and improving business strategies. As shown as fig 4.1 the proposed system integrates NLP with advanced embedding techniques to automate sentiment analysis. It leverages SBERT to capture contextual semantic relationships within textual data. A DNN is used for feature refinement, enhancing representation quality. Finally, a Boosted Rules Classifier ensures accurate and interpretable sentiment prediction.

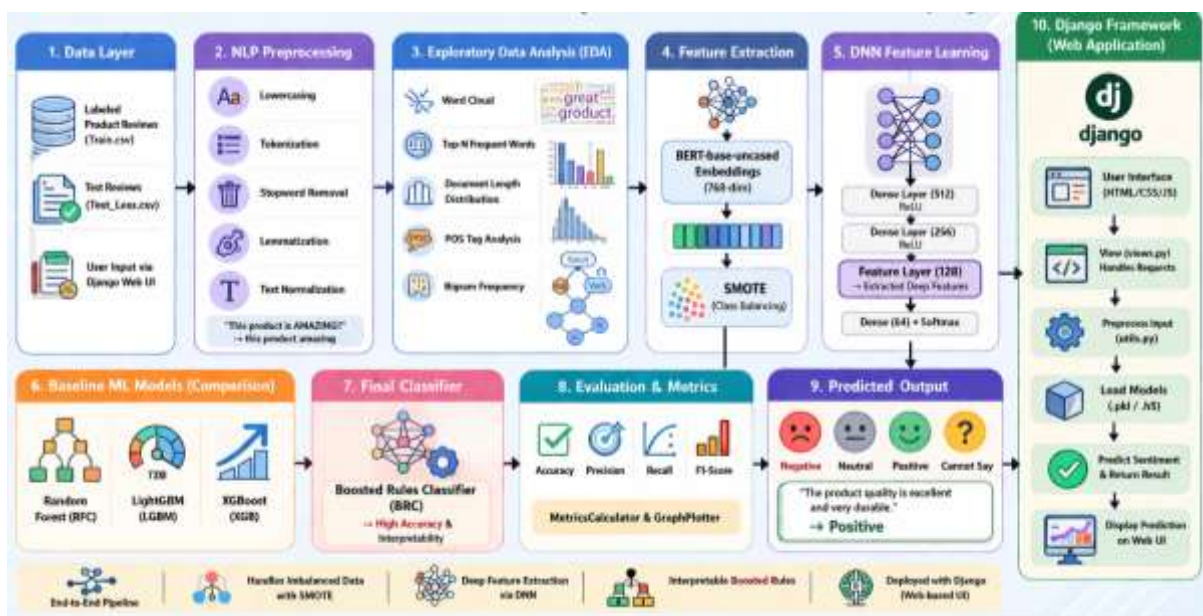


Fig. 2. Proposed system architecture of product sentiment analysis.

The Product Sentiment Analysis workflow begins with data ingestion, where product reviews are collected from sources such as e-commerce platforms, social media, or surveys and stored in structured formats like CSV files containing textual and associated metadata. This is followed by NLP preprocessing, which involves converting text to lowercase, removing punctuation and special

characters, and performing tokenization, lemmatization, and stop-word removal to obtain clean and normalized input data. The preprocessed text is then passed through Sentence-BERT (SBERT) for feature extraction, where dense vector embeddings are generated using mean pooling to capture contextual semantics and relationships within the text. These embeddings are further refined using a DNN comprising dense layers with activation functions such as ReLU and Softmax, along with a bottleneck layer to reduce dimensionality and enhance feature representation. To address class imbalance, the SMOTE is applied to balance sentiment categories such as Positive, Negative, Neutral, and Cannot Say. The refined features are then used to train a Boosted Rules Classifier, which not only improves classification performance but also provides interpretable rule-based predictions. For sentiment prediction, new or unseen reviews are fed into the trained model to classify them into the defined sentiment categories. The system is subsequently deployed using a web framework such as Django, enabling access for real-time predictions and user interaction. Finally, the model's performance is evaluated using metrics such as accuracy, precision, recall, and F1-score, along with visualization techniques like confusion matrices and graphical plots to assess and present the results effectively

SBERT Feature Extraction

Feature extraction is the heart of any Natural Language Processing (NLP) pipeline because it transforms unstructured human language into numerical patterns interpretable by computational models. As shown as fig 3 In this research, the Sentence-BERT (SBERT) model is adopted as the feature-extraction backbone due to its superior ability to produce semantically meaningful sentence embeddings. Traditional embedding models such as TF-IDF, Word2Vec, and GloVe focus primarily on individual word distributions and fail to capture contextual meaning. In contrast, BERT (Bidirectional Encoder Representations from Transformers) introduced bidirectional context understanding, and SBERT further optimizes BERT to efficiently generate fixed-size sentence vectors suitable for tasks such as sentiment analysis, similarity detection, and clustering. SBERT thus bridges the gap between deep contextual models and downstream machine-learning algorithms, enabling fast, accurate, and semantically robust classification of product reviews.

The SBERT architecture is a powerful variant of the BERT model, specifically optimized for semantic similarity and sentence-level understanding. While the original BERT model performs exceptionally well for token-level and classification tasks, it is computationally expensive for comparing multiple sentences due to its pairwise encoding mechanism. SBERT overcomes this limitation by introducing a Siamese and Triplet network architecture that generates fixed-length sentence embeddings efficiently. In the proposed system, SBERT is used to convert preprocessed review sentences into high-dimensional semantic vectors. These embeddings capture contextual, syntactic, and sentiment cues from text, enabling downstream classifiers such as RFC, LGBM, XGBoost, and Boosted Rules to perform sentiment prediction effectively.

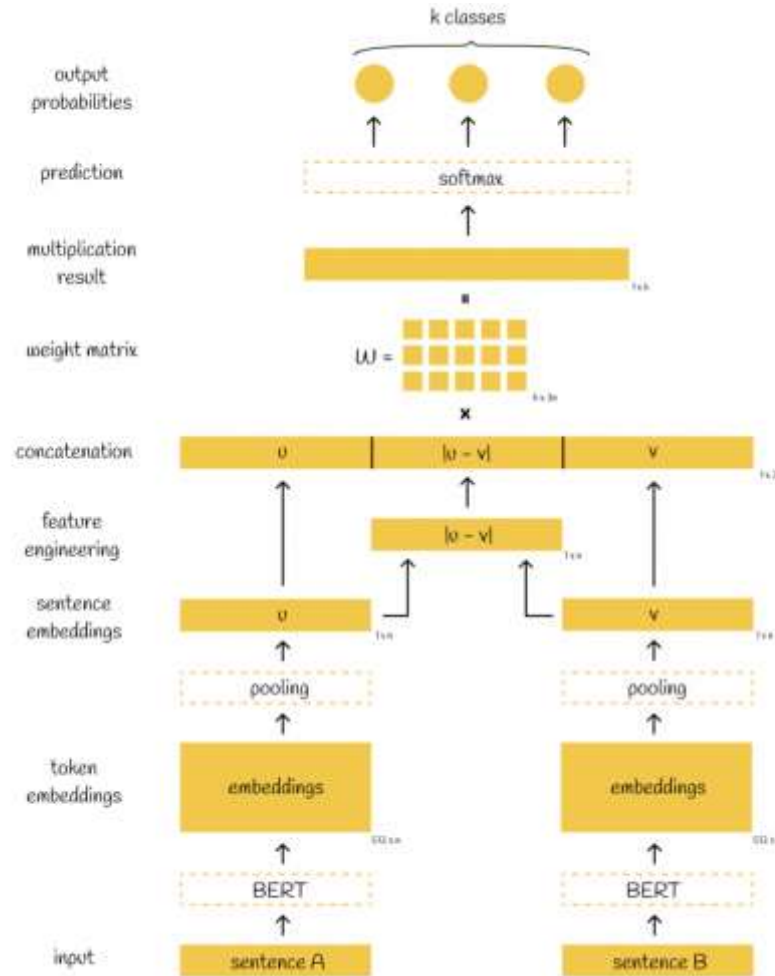


Fig. 3. BERT Sentence Pair Classification.

The feature extraction process begins with data preprocessing, where textual data is cleaned by converting it to lowercase, removing stopwords, and applying lemmatization, followed by tokenization to ensure consistency. The processed sentences are then passed through a BERT-based architecture using the WordPiece tokenizer, where tokens are converted into embeddings along with positional and segment encodings, and special tokens such as [CLS] and [SEP] are added for sentence representation and separation. These embeddings are fed into multiple transformer encoder layers comprising multi-head self-attention and feed-forward neural networks, enabling the model to capture deep contextual relationships and long-range dependencies. Instead of relying solely on the [CLS] token, SBERT employs pooling strategies—primarily mean pooling—to generate a single, fixed-size sentence embedding that effectively represents the semantic meaning of the input. Optionally, a dense layer with non-linear activation functions such as ReLU or Tanh may be applied to refine the embeddings further. The final output is a compact, dense vector representation of each sentence, which can be utilized for tasks such as semantic similarity, clustering, and sentiment analysis, and is typically stored for efficient reuse in subsequent experiments.

4. Results Description

The experimental results demonstrate the effectiveness of the proposed sentiment analysis framework in accurately classifying product reviews into multiple sentiment categories. By leveraging SBERT embeddings and DNN-based feature refinement, the model captures deep contextual relationships

within textual data. The integration of SMOTE successfully addresses class imbalance, leading to improved performance across all sentiment classes. Compared to traditional machine learning models, the Boosted Rules Classifier achieves higher accuracy while maintaining interpretability. Evaluation metrics such as accuracy, precision, recall, and F1-score indicate consistent and reliable performance. The results validate the robustness of the system in delivering scalable and data-driven sentiment insights for real-world applications.

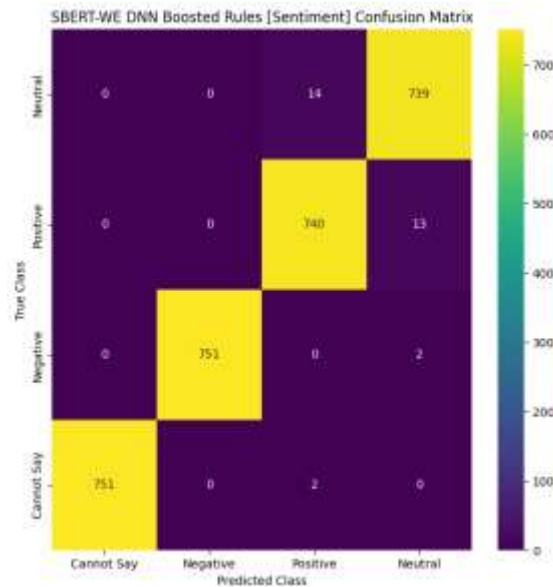


Fig. 4. SBERT-WE DNN Boosted Rules (Sentiment) Confusion Matrix.

Fig. 4. compares confusion matrix true and predicted sentiment classes. The "Cannot Say" class has 751 correct predictions, "Negative" has 751, and "Neutral" has 740, with "Positive" showing 739 correct predictions. Off-diagonal values (e.g., 14 misclassifications from Neutral to Positive) are minimal, suggesting good classification performance.

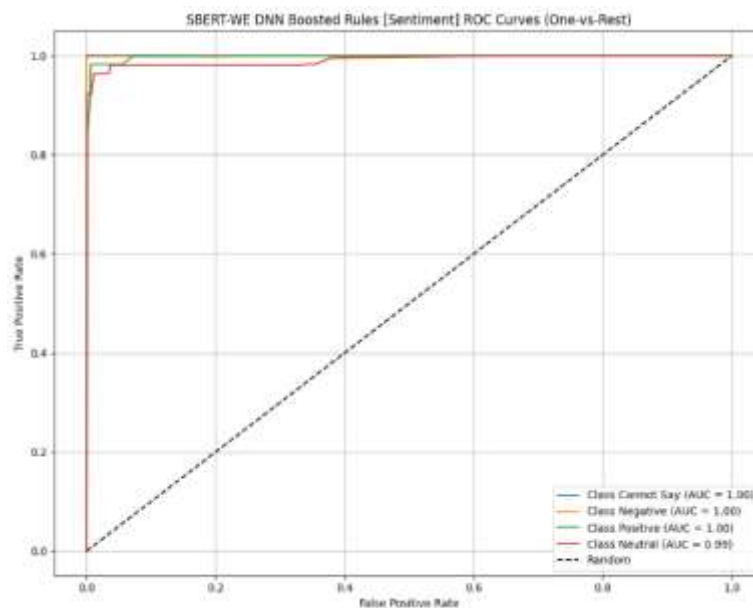


Fig. 5. SBERT-WE DNN Boosted Rules (Sentiment) ROC Curves (One-vs-Rest)

Fig. 5. Depicts ROC curve plot that shows true positive rates against false positive rates for each class. All classes ("Cannot Say," "Negative," "Positive," "Neutral") achieve an AUC of 1.0, indicating perfect classification, while the random guess line (AUC = 0.99) serves as a baseline.

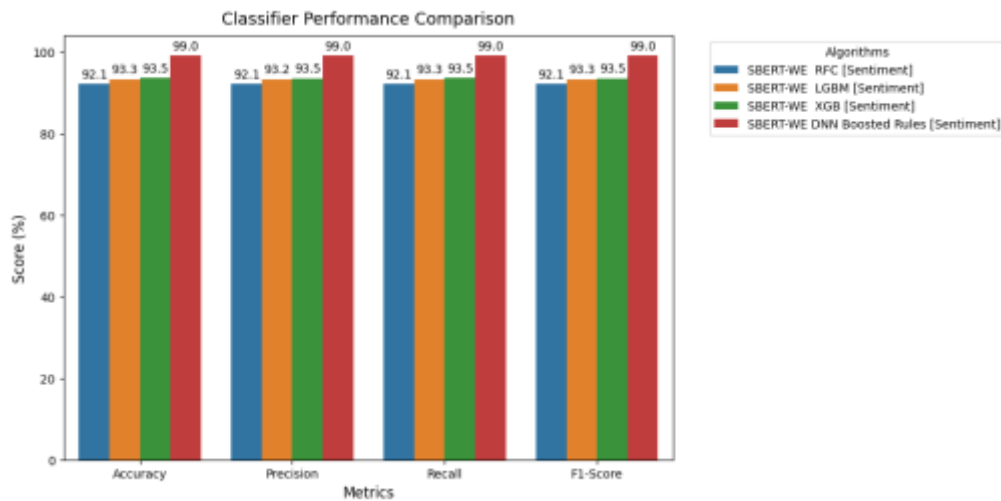


Fig. 6. Classifier Performance Comparison

Fig. 6 depicts the bar chart that compares the performance of different algorithms across accuracy, precision, recall, and F1-score. All algorithms (SBERT-WE RFC, LGBM, XGB, DNN, DNN Boosted Rules) achieve scores around 92-99%, with DNN Boosted Rules slightly leading in most metrics, showing consistent high performance across the board.

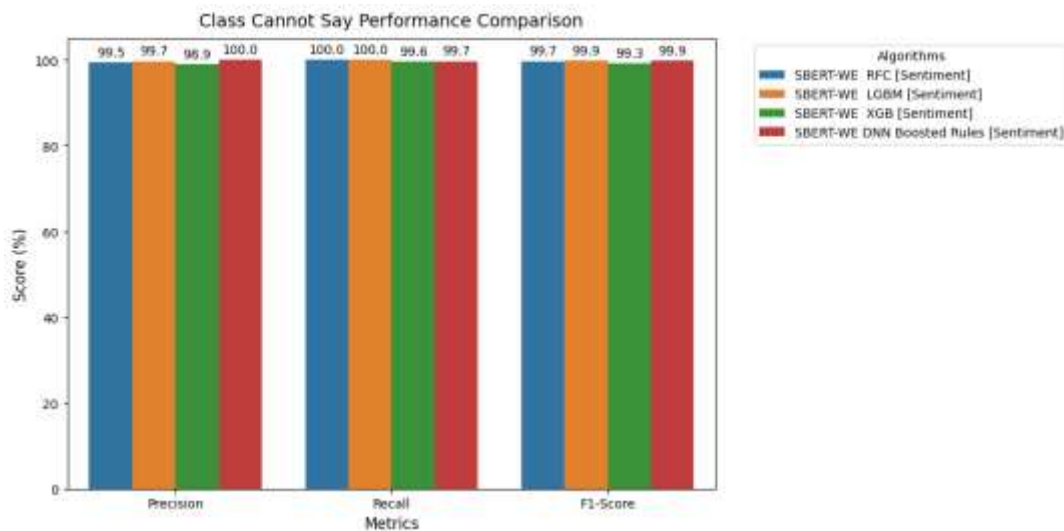


Fig. 7. Class Cannot Say Performance Comparison

Fig. 7. Shows the performance of various algorithms (SBERT-WE RFC, LGBM, XGB, DNN, DNN Boosted Rules) for the "Cannot Say" sentiment class, focusing on precision, recall, and F1-score. All algorithms exhibit near-perfect scores, with precision ranging from 99.5% to 100%, recall from 99.6% to 100%, and F1-score from 99.7% to 99.9%, showcasing exceptional accuracy. The SBERT-WE DNN Boosted Rules model stands out with a 99.9% F1-score, slightly outperforming others like RFC and LGBM, which also perform admirably. The consistent high performance across these metrics highlights the robustness of the models in classifying the "Cannot Say" category with minimal errors.

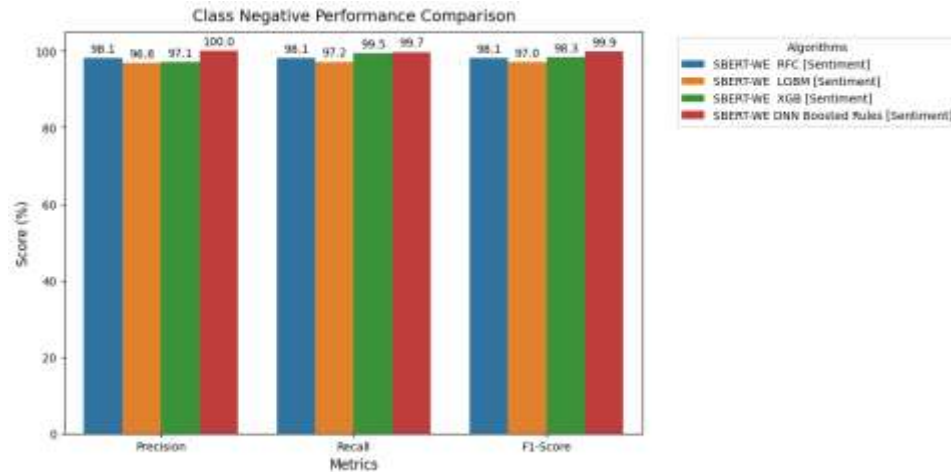


Fig. 8. Class Negative Performance Comparison

Fig. 8. evaluates the performance of different algorithms (SBERT-WE RFC, LGBM, XGB, DNN, DNN Boosted Rules) for the "Negative" sentiment class, covering precision, recall, and F1-score. Precision ranges from 98.1% to 100%, recall from 97.2% to 99.5%, and F1-score from 98.1% to 99.9%, indicating strong overall performance. The SBERT-WE DNN Boosted Rules model leads with a 99.9% F1-score, closely followed by XGB and LGBM, demonstrating high accuracy in negative sentiment detection. The high scores across all metrics suggest that the models effectively identify negative sentiments with very few misclassifications.

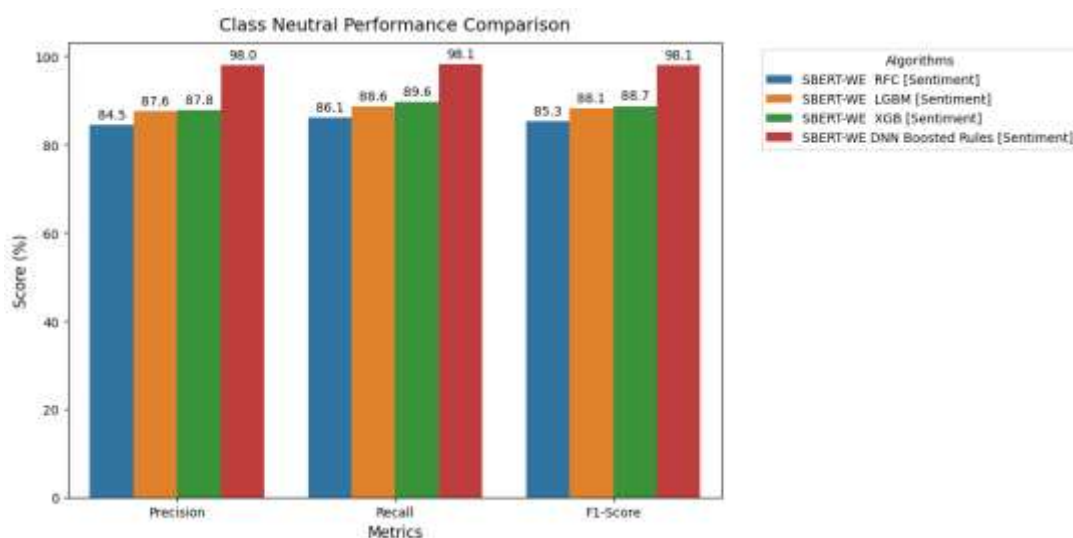


Fig. 9. Class Neutral Performance Comparison

Fig. 9. Shows the bar chart that assesses the performance of algorithms (SBERT-WE RFC, LGBM, XGB, DNN, DNN Boosted Rules) for the "Neutral" sentiment class, analyzing precision, recall, and F1-score. Precision varies from 84.5% to 99.0%, recall from 86.1% to 98.6%, and F1-score from 85.3% to 98.1%, revealing some variability in performance. The SBERT-WE DNN Boosted Rules model achieves the highest F1-score at 98.1%, while RFC lags with 85.3%, indicating challenges in consistent classification. The lower precision and recall for certain models suggest difficulties in accurately identifying "Neutral" sentiments compared to other classes.

5. CONCLUSION

The proposed system presents effective and scalable framework for sentiment classification by combining advanced NLP techniques with machine learning and deep learning methodologies. The approach follows a structured pipeline that includes thorough data preprocessing, context-aware embedding generation using SBERT-based representations, deep feature enhancement through a DNN, and final classification using ensemble techniques along with a Boosted Rules Classifier. Experimental evaluation demonstrates that the proposed SBERT-WE DNN Boosted Rules model significantly surpasses traditional classifiers such as RFC, LGBM, and XGBoost, achieving superior performance with accuracy, precision, recall, and F1-score reaching approximately 98.97%, compared to the 92–93% range observed in baseline models. This notable improvement highlights the strength of integrating semantically rich embeddings with deep feature extraction and rule-based boosting, enabling the model to effectively capture complex linguistic patterns and decision boundaries. Furthermore, the hybrid design not only improves predictive performance but also preserves model interpretability, which is essential for practical deployment in real-world applications. The system efficiently manages multi-class sentiment classification and delivers consistent results across all evaluation metrics, demonstrating its robustness, reliability, and suitability for large-scale sentiment analysis tasks.

REFERENCES

- [1]. Liu, R.; Wang, H.; Li, Y. AgriMFLN: Mixing Features LSTM Networks for Sentiment Analysis of Agricultural Product Reviews. *Appl. Sci.* 2023, 13, 6262. <https://doi.org/10.3390/app13106262>
- [2]. Khan, L.; Amjad, A.; Afaq, K.M.; Chang, H.-T. Deep Sentiment Analysis Using CNN-LSTM Architecture of English and Roman Urdu Text Shared in Social Media. *Appl. Sci.* 2022, 12, 2694. <https://doi.org/10.3390/app12052694>
- [3]. Serna, A.; Soroa, A.; Agerri, R. Applying Deep Learning Techniques for Sentiment Analysis to Assess Sustainable Transport. *Sustainability* 2021, 13, 2397. <https://doi.org/10.3390/su13042397>
- [4]. Ghatore, P.S.; Hosseini, S.E.; Pervez, S.; Iqbal, M.J.; Shaukat, N. Sentiment Analysis of Product Reviews Using Machine Learning and Pre-Trained LLM. *Big Data Cogn. Comput.* 2024, 8, 199. <https://doi.org/10.3390/bdcc8120199>
- [5]. Ahanin, Z.; Ismail, M.A.; Singh, N.S.S.; AL-Ashmori, A. Hybrid Feature Extraction for Multi-Label Emotion Classification in English Text Messages. *Sustainability* 2023, 15, 12539. <https://doi.org/10.3390/su151612539>
- [6]. Chu, Z.; Wang, X.; Jin, M.; Zhang, N.; Gao, Q.; Shao, L. An Effective Strategy for Sentiment Analysis Based on Complex-Valued Embedding and Quantum Long Short-Term Memory Neural Network. *Axioms* 2024, 13, 207. <https://doi.org/10.3390/axioms13030207>
- [7]. Tan, K.L.; Lee, C.P.; Lim, K.M. RoBERTa-GRU: A Hybrid Deep Learning Model for Enhanced Sentiment Analysis. *Appl. Sci.* 2023, 13, 3915. <https://doi.org/10.3390/app13063915>
- [8]. Mao, X.; Chang, S.; Shi, J.; Li, F.; Shi, R. Sentiment-Aware Word Embedding for Emotion Classification. *Appl. Sci.* 2019, 9, 1334. <https://doi.org/10.3390/app9071334>
- [9]. Trisna, K.W.; Huang, J.; Liang, H.; Dharma, E.M. Fusion Text Representations to Enhance Contextual Meaning in Sentiment Classification. *Appl. Sci.* 2024, 14, 10420. <https://doi.org/10.3390/app142210420>

- [10]. Oprea, S.-V.; Bâra, A. Extracting Emotions from Customer Reviews Using Text Mining, Large Language Models and Fine-Tuning Strategies. J. Theor. Appl. Electron. Commer. Res. 2025, 20, 221. <https://doi.org/10.3390/jtaer20030221>