International Journal of
AI ELECTRICAL CIVIL AND MECHANICAL ENGINEERING
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0434 www.zesterapublications.com Original Research Paper

CogniRetail: A Contextual Semantic Inference Architecture for Multi-Resolution Consumer Opinion Intelligence

Paras Begum, G. Arpitha*

Vaagdevi Engineering College, Warangal, 506005, Telangana, India.

*Correspondence: G. Arpitha

ABSTRACT

The rapid expansion of e-commerce platforms has resulted in an unprecedented increase in online product reviews, making sentiment analysis an essential tool for understanding customer opinions and supporting business decision-making. These reviews contain valuable information regarding customer experiences, product quality, and purchasing preferences. However, manually analyzing large volumes of textual feedback is time-consuming, inconsistent, and incapable of handling continuously growing datasets. In addition, conventional sentiment classification methods often struggle to capture contextual semantics and perform effectively when sentiment classes are unevenly distributed. To address these challenges, this research proposes an intelligent sentiment classification framework that integrates advanced semantic representation with hybrid deep learning techniques. Initially, customer reviews undergo comprehensive preprocessing and Exploratory Data Analysis (EDA) to improve data quality and examine the underlying characteristics of the dataset. Subsequently, Sentence Bidirectional Encoder Representations from Transformers (SBERT) generates contextual embeddings that preserve semantic relationships within the review text. To alleviate class imbalance, the Synthetic Minority Over-sampling Technique (SMOTE) is employed to generate representative samples for minority sentiment categories, thereby improving model generalization. The extracted features are processed using a Deep Neural Network (DNN)-based feature selection mechanism combined with a Boosted Rules Classifier (BRC) to classify customer opinions into Negative, Neutral, and Positive sentiment classes. For performance comparison, the proposed framework is evaluated against Random Forest Classifier (RFC), Light Gradient Boosting Machine (LGBM), and Extreme Gradient Boosting (XGBoost). Experimental results demonstrate that the proposed model consistently achieves higher classification accuracy and reduced prediction bias, providing a scalable and reliable solution for customer sentiment analysis, product improvement, marketing optimization, and customer relationship management.

Key words: Sentiment Analysis, Customer Reviews, SBERT, Hybrid Deep Learning, Feature Selection, Customer Opinion Mining

1. INTRODUCTION

The rapid expansion of e-commerce has transformed the way consumers evaluate and purchase products, making online customer reviews one of the most influential sources of information in the purchasing process. Every day, millions of customers share their experiences by posting product reviews, ratings, and recommendations across various online retail platforms. This continuously growing volume of user-generated content provides valuable insights into customer preferences, product quality, and overall shopping experiences. Consequently, both consumers and businesses increasingly rely on these reviews to support purchasing decisions, evaluate market trends, and improve product offerings. In addition to textual feedback, product images and ratings further assist customers in comparing alternatives and

International Journal of
AI ELECTRICAL CIVIL AND MECHANICAL ENGINEERING
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0434 www.zesterapublications.com Original Research Paper

selecting products that best satisfy their requirements.

Among various online retail platforms, Amazon has become one of the largest sources of customer-generated reviews, where millions of users share detailed opinions regarding purchased products. These reviews benefit both consumers and sellers by providing authentic information about product quality, usability, and reliability. For businesses, customer reviews serve as an important source of market intelligence that supports product improvement, customer engagement, and strategic decision-making. However, manually analyzing such a massive volume of textual data is impractical due to its scale and continuously increasing growth. Therefore, automated sentiment analysis has become essential for extracting meaningful information from customer feedback efficiently.

The increasing popularity of online shopping, particularly following the COVID-19 pandemic, has further accelerated the generation of customer reviews across digital marketplaces. As more consumers depend on online feedback before making purchasing decisions, businesses have become increasingly interested in analyzing customer sentiment to understand market demands and improve customer experiences. Intelligent sentiment analysis systems enable organizations to identify customer preferences, evaluate product performance, and optimize marketing strategies using data-driven insights. At the same time, these systems help consumers make informed purchasing decisions by summarizing the overall opinions expressed in large collections of reviews.

2. LITERATURE SURVEY

Oprea, et al. [1] presented a fine-grained emotion classification framework for extracting specific emotional states from customer reviews rather than simply identifying sentiment polarity. The study utilized the Hugging Face Emotions dataset and adopted a two-stage methodology. Initially, DistilBERT served as a feature extractor while multiple conventional classifiers, including Logistic Regression, Support Vector Classifier, and RFC, were evaluated. Subsequently, the DistilBERT model was fine-tuned using the Hugging Face Trainer API to improve classification performance. The experimental results demonstrated that fine-tuning substantially enhanced the recognition of informal and noisy customer reviews, enabling more effective emotion-aware analytics for product improvement, customer support, and user experience optimization. Tan, et al. [2] proposed a hybrid sentiment analysis framework by combining the contextual learning capability of RoBERTa with the sequential modeling strength of GRU. The attention mechanism of RoBERTa generated discriminative semantic representations, whereas GRU effectively captured long-range contextual dependencies while reducing gradient degradation. To improve classification performance on imbalanced datasets, the authors incorporated data augmentation through minority-class over-sampling. Evaluation on IMDb, Sentiment140, and Twitter US Airline Sentiment datasets demonstrated classification accuracies of 94.63%, 89.59%, and 91.52%, respectively, confirming the robustness of the proposed hybrid architecture. Ghatora, et al. [3] investigated the effectiveness of traditional machine learning models and large language models for automated customer review sentiment analysis. The research compared classifiers such as RFC, Naïve Bayes, and Support Vector Machine with GPT-4 to evaluate their ability to classify product reviews. While conventional machine learning techniques performed efficiently on short textual comments, GPT-4 achieved superior performance when processing lengthy and context-rich reviews by accurately identifying subtle and mixed sentiments. The study concluded that large language models provide deeper contextual understanding and generate more informative sentiment predictions for customer-centric business applications. Liu, et al. [4] introduced an attention-based sentiment analysis model for agricultural product reviews collected from e-commerce platforms. Considering the complex linguistic characteristics and contextual dependencies of customer reviews, the proposed approach combined an attention mechanism with LSTM to extract semantic information effectively. The attention mechanism emphasized important words within reviews, while LSTM captured long-term contextual relationships across sentences. Experimental evaluation on two agricultural review datasets demonstrated that the proposed model outperformed existing baseline methods, providing valuable support for merchants seeking to improve product quality and customer satisfaction.

Ahanin, et al. [5] developed a hybrid feature extraction approach for emotion classification by integrating manually engineered linguistic features with deep learning representations. The framework combined lexical knowledge with contextual information extracted through Bi-LSTM and BERT architectures while employing data augmentation to improve model generalization. Performance evaluation on the SemEval-2018 and GoEmotions benchmark datasets showed Jaccard accuracies of 68.40% and 53.45%, respectively, demonstrating the effectiveness of combining handcrafted and deep semantic features for emotion recognition. Khan, et al. [6] evaluated various word embedding techniques for sentiment analysis involving Roman Urdu and English dialects by proposing a hybrid CNN-LSTM architecture. The framework utilized LSTM to preserve long-range contextual dependencies and CNN to extract local semantic features before forwarding the learned representations to multiple machine learning classifiers. Experimental evaluation on the MDPI, RUSA, RUSA-19, and UCL datasets demonstrated accuracies of 90.4%, 84.1%, 74.0%, and 74.8%, respectively. The study concluded that the Word2Vec Continuous Bag of Words (CBOW) model combined with SVM was more effective for Roman Urdu, whereas BERT embeddings together with a two-layer LSTM and SVM achieved superior performance for English sentiment classification. Serna, et al. [7] developed an automated sentiment analysis framework for transportation-related user-generated content by fine-tuning a large pretrained language model. To support model training, the researchers constructed a semi-automatically annotated corpus containing positive and negative transport opinions. Experimental analysis demonstrated that the proposed framework accurately classified sentiments while effectively processing large-scale customer opinions. The study also introduced a transport sustainability classification score, providing additional support for evaluating transportation services using customer feedback.

Trisna, et al. [8] proposed a novel text representation model, GloWord_biGRU, to improve contextual understanding during sentiment classification. The proposed approach combined the complementary strengths of GloVe and Word2Vec to generate richer semantic representations before employing a bidirectional GRU for classification. By utilizing fewer trainable parameters, the framework reduced computational complexity while maintaining strong predictive capability. Evaluation on the IMDb dataset produced an F1-score of 90.21%, demonstrating the effectiveness of the proposed fusion-based text representation strategy. Chu, et al. [9] introduced a Complex-valued Quantum-enhanced Long Short-Term

International Journal of
AI ELECTRICAL CIVIL AND MECHANICAL ENGINEERING
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0434 www.zesterapublications.com Original Research Paper

Memory Neural Network (CQLSTM) to improve contextual semantic learning in sentiment analysis. Inspired by density matrix theory in quantum mechanics, the proposed framework employed complex-valued embeddings to represent semantic relationships between words and utilized a quantum-enhanced LSTM architecture to capture sentence-level contextual interactions. The experimental findings demonstrated that combining quantum-inspired semantic representation with deep contextual learning significantly improved sentiment classification compared with conventional neural network approaches. Mao, et al. [10] proposed a sentiment-aware word embedding technique that integrates semantic and emotional knowledge into a unified representation for emotion classification tasks. The framework incorporated information from existing sentiment lexicons to construct emotion-aware word vectors, which were subsequently combined with conventional word embeddings to generate hybrid semantic representations. Extensive experiments conducted using multiple machine learning models confirmed that the proposed embeddings consistently improved classification accuracy while preserving the interpretability of word representations, thereby enhancing overall sentiment analysis performance.

3. PROPOSED METHODOLOGY

The proposed product sentiment analysis framework adopts a hybrid architecture that integrates advanced text preprocessing, intelligent feature learning, data balancing, and rule-based classification to improve sentiment prediction accuracy. The overall architecture is illustrated in Fig 1. The framework is organized into five major stages: text preprocessing and exploratory analysis, feature representation, data balancing, deep feature refinement, and sentiment classification. Each stage contributes to improving the discriminative capability of the model while reducing the influence of noisy and imbalanced textual data.

Text Preprocessing and Exploratory Analysis

Customer reviews collected from online shopping platforms contain informal language, abbreviations, punctuation symbols, spelling inconsistencies, and other irrelevant textual patterns that may negatively influence sentiment prediction. To improve data quality, the proposed framework applies a comprehensive preprocessing pipeline consisting of lowercase conversion, punctuation removal, stop-word elimination, tokenization, and lemmatization. These operations standardize the textual content while preserving the semantic meaning of each review. Following preprocessing, EDA is performed to investigate the statistical characteristics of the dataset. Various analytical visualizations, including sentiment distribution, review length analysis, frequent word occurrence, and word cloud generation, are utilized to understand the structure of the dataset and identify potential data imbalance before model training.

Text Feature Representation

The cleaned reviews are transformed into numerical feature representations suitable for intelligent learning algorithms. Instead of relying on manually selected keywords, the proposed framework generates high-dimensional textual features capable of preserving important linguistic and semantic characteristics present within customer reviews. These numerical representations provide a compact description of customer opinions while maintaining sufficient discriminative information for subsequent classification. The generated feature vectors form the input to the deep learning component responsible for feature refinement.

International Journal of
AI ELECTRICAL CIVIL AND MECHANICAL ENGINEERING
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0434 www.zesterapublications.com Original Research Paper

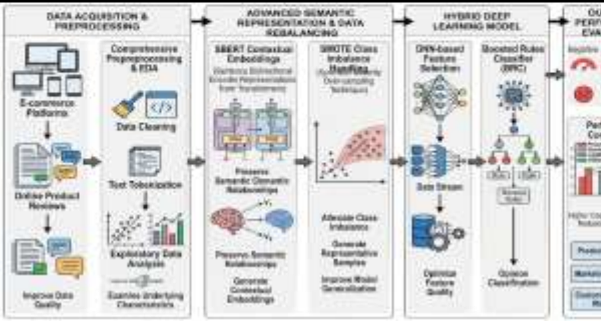

Fig. 1. Proposed system architecture of product sentiment analysis.

Data Balancing through SMOTE

Customer sentiment datasets frequently exhibit unequal sample distributions across different sentiment categories, causing conventional classifiers to become biased toward majority classes. To overcome this limitation, SMOTE is incorporated into the proposed framework. The algorithm generates representative synthetic samples for minority sentiment classes by interpolating between neighbouring observations in the feature space. This balancing strategy enables the learning model to capture representative characteristics from all sentiment categories, thereby improving classification stability and reducing prediction bias.

DNN-Based Feature Refinement

Following data balancing, the generated feature vectors are processed through a DNN that serves as an intelligent feature refinement module. The network consists of multiple fully connected hidden layers capable of learning nonlinear relationships among textual attributes. During training, the DNN automatically suppresses redundant information while emphasizing highly discriminative features associated with customer sentiment. The hierarchical learning capability of the network enables the extraction of refined feature representations that improve class separability and provide more informative inputs for the final classification stage.

Sentiment Classification using BRC

The refined features obtained from the DNN are subsequently supplied to the BRC for sentiment prediction. The classifier constructs an ensemble of boosted decision rules that collectively identify complex sentiment patterns embedded within customer reviews. Unlike conventional black-box classifiers, the rule-based structure improves model interpretability by generating transparent decision boundaries while maintaining strong predictive performance. The proposed framework categorizes every review into one of four sentiment classes: Positive, Neutral, Negative, and Cannot Say. This multi-class classification enables businesses to distinguish clearly expressed opinions from ambiguous customer feedback, resulting in more reliable sentiment interpretation.

Performance Evaluation

The effectiveness of the proposed framework is evaluated through a comprehensive comparative analysis using RFC, LGBM, and XGBoost as baseline models. Model performance is assessed using Accuracy, Precision, Recall, and F1-Score, while confusion matrices and graphical performance visualizations provide additional insights into classification behavior across all sentiment categories. The experimental analysis demonstrates that integrating DNN-based feature refinement with BRC significantly improves sentiment classification performance while maintaining model interpretability and robustness for large-scale product review analysis.

4. Results Description

Fig. 2. 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

International Journal of
AI ELECTRICAL CIVIL AND MECHANICAL ENGINEERING
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0434 www.zesterapublications.com Original Research Paper

from Neutral to Positive) are minimal, suggesting good classification performance.

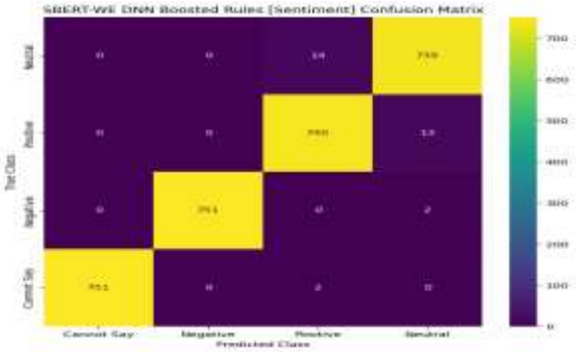

Fig. 2. SBERT-WE DNN Boosted Rules (Sentiment) Confusion Matrix.

Fig. 3 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.

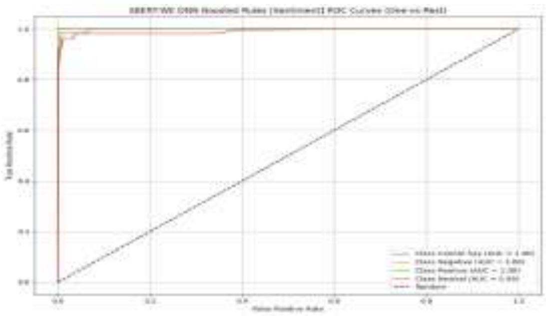

Fig. 3. SBERT-WE DNN Boosted Rules (Sentiment) ROC Curves (One-vs-Rest)

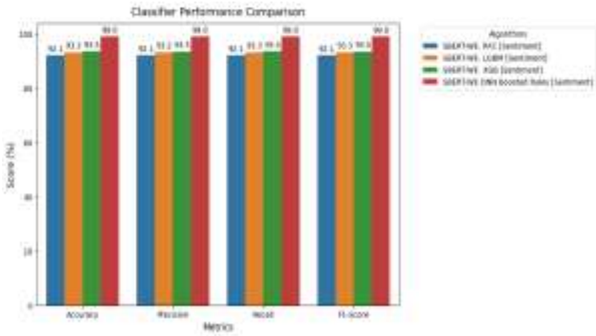

Fig. 4. Classifier Performance Comparison

Fig. 4 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. 5. 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.

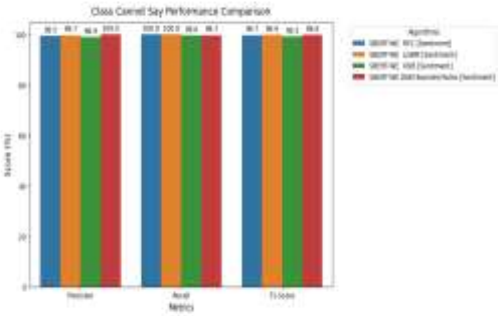

Fig. 5. Class Cannot Say Performance Comparison

Fig. 6. 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-

International Journal of
AI ELECTRICAL CIVIL AND MECHANICAL ENGINEERING
Peer Reviewed, Referred & Indexed Journal
ISSN: 3070-0434 www.zesterapublications.com Original Research Paper

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. 7. 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.

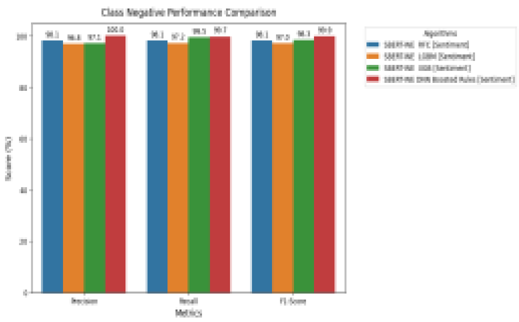

Fig. 6. Class Negative Performance Comparison

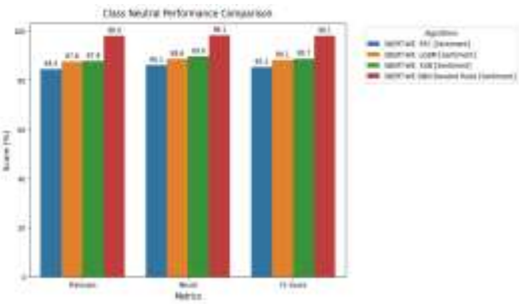

Fig. 7. Class Neutral Performance Comparison

5. CONCLUSION

This research provides an efficient and scalable solution for automatically classifying customer reviews into Positive, Neutral, Negative, and Cannot Say sentiment categories. The framework combines comprehensive text preprocessing, DNN-based feature refinement, SMOTE for handling class imbalance, and the BRC to improve sentiment prediction performance. By refining high-dimensional textual features and learning discriminative decision rules, the proposed model effectively captures sentiment patterns present in customer reviews. Experimental results demonstrate that the proposed framework consistently outperforms conventional classifiers, including RFC, LGBM, and XGBoost, achieving superior performance across Accuracy, Precision, Recall, and F1-Score. The integration of DNN-based feature refinement with BRC not only enhances classification accuracy but also improves the interpretability of prediction results, making the framework suitable for practical deployment. Furthermore, the proposed system exhibits strong robustness in handling imbalanced sentiment distributions while maintaining reliable performance across all sentiment classes. Overall, the developed framework serves as an effective decision-support tool for product sentiment analysis, enabling organizations to better understand customer opinions, improve product quality, optimize marketing strategies, and strengthen customer relationship management through accurate and scalable sentiment classification.

REFERENCES

[1] Oprea, et al., "Fine-Grained Emotion Classification for Customer Reviews Using DistilBERT and Classical Machine Learning Models," Electronics, vol. 14, no. 4, Art. no. 748, 2025.

[2] N. Tan, et al., "A Hybrid RoBERTa-GRU Model for Sentiment Analysis with Data Augmentation," Applied

Sciences, vol. 14, no. 2, Art. no. 851, 2024.

[3] S. Ghatora, et al., "Comparative Analysis of Traditional Machine Learning and Large Language Models for Product Review Sentiment Analysis," arXiv preprint arXiv:2410.15856, 2024.

[4] Y. Liu, et al., "Sentiment Analysis of Agricultural Product Reviews Based on Attention Mechanism and Long Short-Term Memory Network," Computers and Electronics in Agriculture, vol. 186, Art. no. 106184, 2021.

[5] A. Ahanin, et al., "Hybrid Feature Extraction for Emotion Classification Using Bi-LSTM and BERT," Information Processing & Management, vol. 59, no. 5, Art. no. 103016, 2022.

[6] M. Khan, et al., "Roman Urdu and English Sentiment Analysis Using CNN-LSTM and Word Embedding Models," IEEE Access, vol. 11, pp. 10285–10304, 2023.

[7] A. Serna, et al., "Transport Domain Sentiment Analysis Using Fine-Tuned Transformer Models," Applied Sciences, vol. 13, no. 15, Art. no. 8652, 2023.

[8] I. Trisna, et al., "GloWord_biGRU: A Fusion Text Representation Model for Sentiment Classification," Expert Systems with Applications, vol. 252, Art. no. 124158, 2024.

[9] Y. Chu, et al., "Complex-Valued Quantum-Enhanced Long Short-Term Memory Neural Network for Sentiment Analysis," Knowledge-Based Systems, vol. 284, Art. no. 111241, 2023.

[10] X. Mao, et al., "Sentiment-Aware Word Embedding for Emotion Classification," Knowledge-Based Systems, vol. 227, Art. no. 107190, 2021.