IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⸎ Peer-reviewed & Refereed journal ⸎ Vol. 14, Issue 2, February 2025

DOI: 10.17148/IJARCCE.2025.14260

Sentimental Analysis on Product Review in E-Commerce Platform using Machine Learning and Deep Learning

Mr. Satish Kumar Parasa[1], Chintala Guna Vardhan[2], Dodda Diswanth[3], Chavatapalli Prathap[4], Banavath Pavan Naik[5]

Assistant Professor, CSE (Cybersecurity, IOT Including Blockchain Technology),
Vasireddy Venkatadri Institute of Technology, Guntur, India[1]

Student, CSE (Cybersecurity, IOT Including Blockchain Technology),
Vasireddy Venkatadri Institute of Technology, Guntur, India[2]

Student, CSE (Cybersecurity, IOT Including Blockchain Technology),
Vasireddy Venkatadri Institute of Technology, Guntur, India[3]

Student, CSE (Cybersecurity, IOT Including Blockchain Technology),
Vasireddy Venkatadri Institute of Technology, Guntur, India[4]

Student, CSE (Cybersecurity, IOT Including Blockchain Technology),
Vasireddy Venkatadri Institute of Technology, Guntur, India[5]

Abstract: This paper focuses on developing a comprehensive sentiment analysis system for customer reviews, combining traditional machine learning and advanced deep learning techniques. The system classifies reviews into positive, negative, or neutral categories through robust text preprocessing, feature extraction, and model training. Traditional classifiers like Random Forest, Naive Bayes, Logistic Regression, and SVM are utilized alongside advanced NLP models such as VADER for quick analysis of short reviews and BERT for an in- depth understanding of longer, context-rich reviews. The system employs ensemble methods to enhance accuracy and consistency in sentiment classification. It evaluates performance through metrics such as accuracy, precision, recall, and F1-score to ensure reliability and scalability. A user- friendly Flask-based web application enables seamless dataset uploads, real-time analysis, sentiment visualization, and downloadable results. The project aims to provide an efficient and accurate sentiment analysis solution adaptable to diverse e-commerce platforms and customer feedback scenarios

Keywords: Sentiment Analysis; NLP; VADER; BERT; Customer Reviews; TF-IDF.

I. INTRODUCTION

Customer sentiment is a critical factor influencing the success of businesses, particularly in the dynamic landscape of e-commerce. Reviews and feedback shared by customers serve as valuable insights into product quality, service satisfaction, and areas requiring improvement. However, the sheer volume and diversity of online reviews present a challenge for manual analysis, necessitating the use of automated sentiment analysis systems. These systems can process large datasets efficiently, enabling businesses to identify trends, understand customer needs, and improve their offerings. This project aims to develop a sentiment analysis system that classifies customer reviews into positive, negative, or neutral categories using a combination of machine learning and deep learning techniques. The system begins by preprocessing the text data through cleaning, tokenization, and feature extraction, ensuring that irrelevant noise is removed while preserving essential information. By employing traditional machine learning classifiers such as Random Forest, Naïve Bayes, Logistic Regression, and SVM, the system establishes a solid foundation for accurate predictions.
In addition to traditional methods, advanced natural language processing (NLP) models are integrated to handle the complexities of customer reviews. Shorter reviews are processed using VADER, a rule-based sentiment analysis tool known for its speed and simplicity. Longer reviews, which often contain more nuanced and context-rich language, are analyzed using BERT, a state-of-the-art deep learning model. This dual-model approach ensures both efficiency and depth in sentiment classification, catering to the varying nature of review lengths. To further enhance accuracy and consistency, the system incorporates ensemble techniques that combine outputs from multiple models. By leveraging the

DOI: 10.17148/IJARCCE.2025.14260

strengths of each method, the ensemble approach reduces discrepancies and improves the reliability of sentiment predictions. Performance evaluation is conducted using metrics such as accuracy, precision, recall, and F1-score, ensuring the system meets high-quality standards. To provide users with a seamless experience, a Flask-based web application has been developed. This intuitive interface allows users to upload datasets, process sentiment analysis, visualize results, and download the sentiment-labeled data for further use. The application is designed to be user-friendly, ensuring that users with minimal technical expertise can operate it with ease.

This project stands out by addressing the diverse needs of businesses, offering a flexible and scalable sentiment analysis solution. By combining traditional machine learning with advanced deep learning models, it provides a robust framework for understanding customer sentiment, enabling businesses to make informed decisions and improve their customer engagement strategies.

II. LITERATURE SURVEY

Sentiment analysis, often referred to as opinion mining, has been an area of extensive research in recent years. Its applications span across domains such as customer feedback analysis, social media monitoring, and brand reputation management. Various methodologies, ranging from traditional machine learning techniques to advanced deep learning models, have been proposed and implemented to tackle sentiment classification tasks. Early approaches to sentiment analysis relied heavily on machine learning algorithms such as Naive Bayes, Support Vector Machines (SVM), and Logistic Regression. These models require careful feature engineering to extract meaningful insights from text data, often leveraging techniques like bag-of-words and TF-IDF for text vectorization[1] [2].

While these methods are computationally efficient and effective for structured datasets, they struggle with capturing the semantic meaning and context of words, which is crucial for accurate sentiment classification. The advent of rule-based systems, such as the VADER (Valence Aware Dictionary and sEntiment Reasoner) sentiment analysis tool, introduced a lightweight and efficient method for analyzing textual data. VADER uses a lexicon of pre-defined words with associated sentiment scores, making it particularly effective for short and informal text, such as social media posts or product reviews [3].

However, its reliance on pre-defined rules limits its ability to generalize to more complex and context-rich text. Recent advances in deep learning have revolutionized sentiment analysis, with models like BERT (Bidirectional Encoder Representations from Transformers) leading the way. BERT's ability to understand the bidirectional context of words allows it to capture nuanced sentiment in text, making it highly effective for analyzing long and complex reviews [4].

However, deep learning models like BERT are computationally intensive and require significant resources, posing challenges for real- time applications. Hybrid and ensemble methods have emerged to address the limitations of individual approaches. Combining the strengths of rule-based systems, traditional machine learning, and deep learning models enhances overall performance and robustness [5].

These ensemble techniques are particularly useful in scenarios where datasets contain a mix of short and long reviews with varying levels of complexity. Existing sentiment analysis systems have also integrated user- friendly web interfaces to enhance accessibility. Flask-based applications are commonly used to enable seamless interaction with the system, allowing users to upload datasets, process text, and visualize results [6].

Such interfaces make advanced sentiment analysis techniques accessible to businesses without requiring deep technical expertise. Despite significant progress, challenges remain in the field of sentiment analysis. Issues such as handling sarcasm, ambiguous language, and domain-specific jargon continue to pose difficulties [7].

Moreover, ensuring scalability and maintaining high accuracy across diverse datasets are ongoing areas of research. This project builds upon these advancements by integrating traditional machine learning classifiers, VADER, and BERT into a unified system. By leveraging an ensemble approach and providing a Flask-based interface, it addresses the limitations of existing methods while offering a robust and scalable solution for sentiment analysis. Existing sentiment analysis systems have also integrated user- friendly web interfaces to enhance accessibility. Flask-based applications are commonly used to enable seamless interaction with the system, allowing users to upload datasets, process text, and visualize results [8].

Such interfaces make advanced sentiment analysis techniques accessible to businesses without requiring deep technical expertise. Web-based systems have also enabled real-time sentiment analysis, facilitating instant feedback on customer reviews, social media posts, and other textual data streams. This is particularly valuable for businesses seeking to respond rapidly to changing customer sentiments. Despite significant progress, challenges remain in the field of sentiment analysis. Issues such as handling sarcasm, ambiguous language, and domain-specific jargon continue to pose difficulties [9].

Additionally, the variability in sentence structure, use of informal language, and multilingual data further complicates the task of sentiment classification. Moreover, ensuring scalability and maintaining high accuracy across diverse datasets

Impact Factor 8.102 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 14, Issue 2, February 2025

DOI: 10.17148/IJARCCE.2025.14260

are ongoing areas of research. Multi- modal sentiment analysis, which incorporates not just text but also voice, images, and video, is another frontier that has garnered increasing attention as it has the potential to provide a more holistic understanding of sentiment in complex data sources.

III. PROPOSED METHODOLOGY

The proposed methodology for sentiment analysis in this project is designed to effectively classify customer reviews into three sentiment categories: positive, negative, and neutral. The system utilizes a combination of traditional machine learning classifiers, rule-based systems, and advanced deep learning models. The following steps outline the key components of the methodology.

3.1 Data Collection and Preprocessing

The first step in the methodology is the collection of a dataset of customer reviews. The dataset must contain reviews labeled with their corresponding sentiment (positive, negative, or neutral). This can be gathered from various sources, such as e-commerce platforms, product review websites, or social media platforms. The dataset is then preprocessed to clean and prepare it for analysis.

Preprocessing Steps Include:

3.1.1 Text Normalization**:** All text is converted to Lowercase to ensure uniformity

3.1.2 Removing Noise: Punctuation, special characters, and non-alphanumeric symbols are removed from the text using regular expressions.

3.1.3 Tokenization: The text is broken down into individual words or tokens, which are easier to process and analyse.

3.1.4 Stop word Removal: Common but unimportant words (e.g., "the", "is", "and") are removed to reduce the dimensionality of the text data.

3.1.5 Lemmatization: Words are reduced to their base form (e.g., "running" becomes "run"), improving the model’s ability to understand different variations of the same word.

3.1.6 Negation Handling: Words that indicate negation (e.g., "not", "never") are processed to ensure they are appropriately reflected in the sentiment interpretation.

3.2 Feature Extraction

The data is preprocessed, feature extraction techniques are applied to transform the raw text into numerical representations suitable for machine learning models.

3.2.1 TF-IDF Vectorization: The Term Frequency-Inverse Document Frequency (TF-IDF) method is used to convert the text data into a numerical format, considering both the frequency of words in individual documents and their rarity across the entire dataset. This helps highlight important words in reviews while reducing the impact of less meaningful ones.

3.2.2 Sentiment Scores (VADER): For short reviews, sentiment scores are extracted using VADER, a rule- based sentiment analysis tool. VADER provides a sentiment score for each review based on a predefined lexicon, classifying reviews as positive, negative, or neutral based on the aggregate score.

3.2.3 Word Frequency Features: In addition to TF-IDF scores, word frequencies are calculated to capture the most common terms in the dataset, which can provide additional insight into the sentiment expressed in reviews.

3.3 Model Selection and Training

The system employs a hybrid approach, utilizing both traditional machine learning classifiers and deep learning models for sentiment classification.

3.3.1 Random Forest Classifier: Instead of relying on a single decision path, Random Forest constructs numerous decision trees using randomized subsets of your data and features. Each tree casts its vote for a class, and the final prediction is the most frequent outcome. This ensemble approach minimizes overfitting and boosts predictive power by combining the outputs of diverse trees. It's adept at handling complex datasets with many features and provides insights into feature importance. While a robust method, its computational demands can increase with large datasets, and its internal workings are less transparent than a single decision tree.

3.3.2 Naive Bayes Classifier: Naive Bayes applies Bayes' theorem, assuming features are independent, to estimate the probability of a class given the observed features. This probabilistic approach is particularly efficient for text-based classification. Despite the simplifying "naive" assumption, it often delivers strong performance, especially with high-dimensional text data. It's a straightforward, rapid method suitable for tasks like spam filtering or sentiment analysis. However, the independence assumption can limit its accuracy when features are strongly correlated.

3.3.3 Logistic Regression: Logistic Regression models the likelihood of a binary or multi-class outcome through a linear combination of input features, transformed by the sigmoid function. It's a straightforward, computationally efficient, and easily interpretable method, providing information on feature importance via its coefficients. It's well-suited for data that

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁝ Peer-reviewed & Refereed journal ⁝ Vol. 14, Issue 2, February 2025

DOI: 10.17148/IJARCCE.2025.14260

can be separated by a linear boundary, and it yields probability estimates. However, it relies on the assumption of a linear relationship and may struggle to capture intricate, non-linear patterns.

3.3.4 Support Vector Machines (SVM): SVM aims to define the optimal dividing line, known as a hyperplane, that maximizes the separation between different classes. It excels in high-dimensional spaces, such as those encountered in text analysis, and employs kernel functions to handle non-linear data arrangements. SVM focuses on establishing a strong separation boundary using support vectors, the data points nearest to the hyperplane. While a potent tool, its computational cost can be significant for extensive datasets, and the selection of the appropriate kernel and parameters is vital for achieving optimal results.

3.3.5 Deep Learning Model (BERT): For longer reviews, BERT (Bidirectional Encoder Representations from Transformers) is utilized. BERT is a state-of-the-art model that understands the context of words in both directions (left-to-right and right-to-left) by using a transformer-based architecture. Fine-tuned for sentiment analysis, BERT captures more nuanced sentiment in long and complex reviews, overcoming the limitations of traditional models.

3.3.6 Ensemble Approach: The outputs of the various models are combined to improve overall prediction accuracy. If different models predict conflicting results, a neutral label is assigned, ensuring robustness across varying review types.

3.4 Sentiment Classification

The classification process is determined by the length and complexity of the review. Short reviews are processed using the VADER sentiment analysis tool, while longer reviews are classified using the BERT model for a more contextual understanding.

3.4.1 Short Reviews: Reviews with fewer than 120 characters are passed through the VADER sentiment analysis tool to determine sentiment. VADER evaluates the polarity scores of the reviews, assigning them a label of positive, negative, or neutral based on the sentiment intensity.

3.4.2 Long Reviews: Reviews exceeding 120 characters are processed by BERT to understand the full context and more accurately classify sentiment. The BERT model outputs a sentiment label (positive, negative, or neutral) with a confidence score.

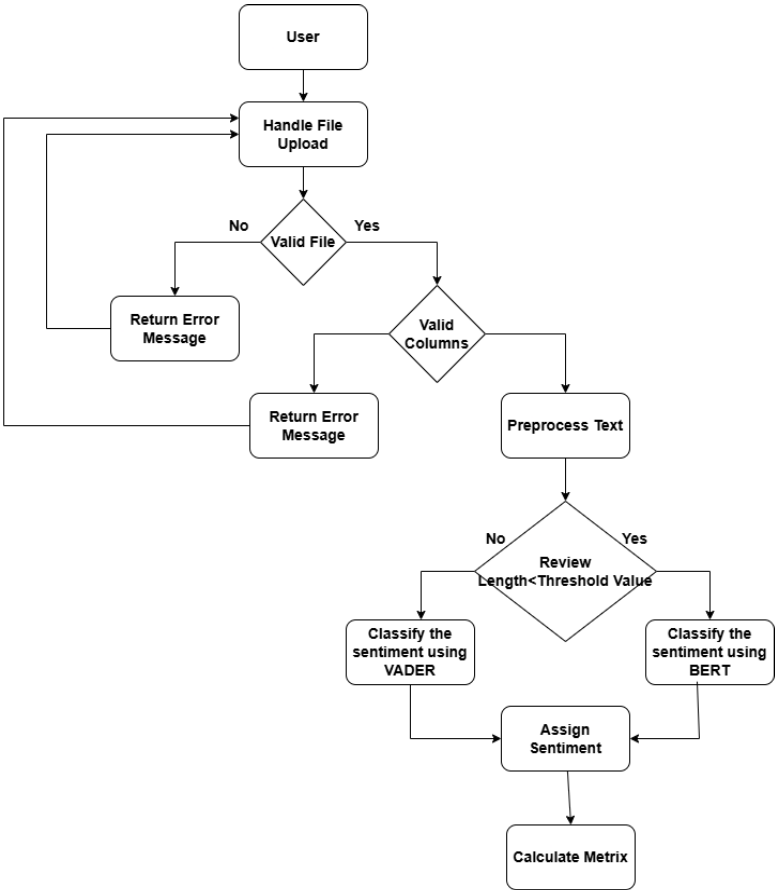

Fig.1: Flow Chart

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 14, Issue 2, February 2025
DOI: 10.17148/IJARCCE.2025.14260

Table1: Project Key Features

Key Features	Description
Sentiment Classification	Classifies reviews into positive, negative, or neutral.
Techniques Used Text	Combines machine learning and deep learning models.
Preprocessing	Involves tokenization, stop word removal, and lemmatization.
Ensemble Approach	Uses multiple models for improved accuracy.
Web Interface	Provides a Flask-based user interface.
Evaluation Metrics	Measures performance with accuracy, precision, recall, and F1-score.
Scalability	Handles short and long reviews effectively.

This project focuses on developing an efficient sentiment analysis system that classifies customer reviews into positive, negative, or neutral categories. It combines machine learning algorithms and advanced natural language processing models like VADER and BERT for effective sentiment detection. The system provides a web interface for easy file upload and sentiment analysis, delivering valuable insights into customer feedback. The model's performance is evaluated with accuracy, precision, recall, and F1-score, ensuring robust and reliable results for diverse review datasets.

IV. RESULTS AND DISCUSSIONS

The sentiment analysis model demonstrated a high level of accuracy in classifying customer reviews into positive, negative, and neutral categories. The model attained 96% accuracy, indicating its ability to correctly classify the instances. Furthermore, it exhibited high precision at 96.22%, signifying a low rate of false positives. The recall of 96% demonstrates the model's effectiveness in identifying true positives. Finally, the F1-score of 96.01%.

a) Accuracy
The model achieved an accuracy of 96%, meaning that 96% of the predictions made by the model were correct. This indicates a strong overall performance in sentiment classification.

b) Precision
With a precision of 96.22%, the model was able to correctly identify positive, negative, and neutral sentiments in 96.22% of the cases where it made a positive prediction. This suggests that the model is effective in minimizing false positives and accurately identifying the intended sentiment when it classifies a review as positive or negative.

c) Recall
The recall of 96% reflects the model's ability to correctly identify true sentiment cases out of all possible positive, negative, and neutral instances in the dataset. A recall rate of 96% indicates that the model is good at capturing most of the correct sentiments in the reviews, although there is still room for improvement in minimizing false negatives.

d)F1-Score
The F1-score of 96% combines both precision and recall into a single metric, giving a balanced view of the model's performance. The F1-score close to precision and recall indicates that the model is not heavily biased toward any specific class and performs well in all three categories (positive, negative, neutral).

e) Model Evaluation
The combination of these performance metrics indicates that the model is both accurate and robust in handling a diverse range of customer reviews. However, there are areas for further improvement, particularly in addressing any edge cases or more complex sentences that may result in slight misclassifications. Future iterations of the model could explore fine-tuning the deep learning models, enhancing the feature engineering process, and incorporating additional data for improved results. Overall, the sentiment analysis model provides an effective solution for classifying customer reviews with high accuracy and balanced performance across different metrics
The uploaded image represents the "Upload CSV" page, where users can upload their product review datasets in CSV format for sentiment analysis. This page provides an easy- to-use interface, allowing users to select a file and submit it for processing. Once uploaded, the system will analyze the reviews and return sentiment classification results, offering valuable insights into customer feedback.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 2, February 2025
DOI: 10.17148/IJARCCE.2025.14260

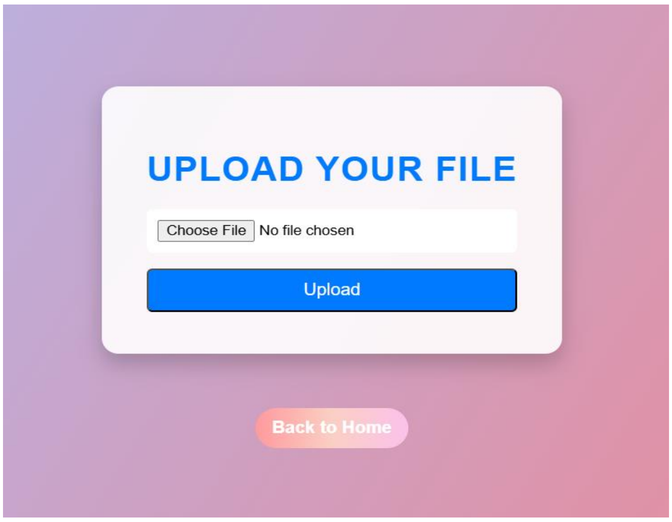

Fig.2: Upload CSV File

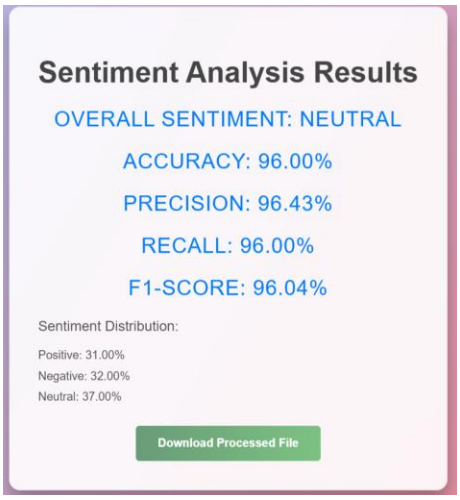

Fig.3: Sentiment Analysis Results

The sentiment analysis model has detected the sentiment of the provided review as Neutral. This classification indicates that the model found no strong positive or negative emotions in the text. The review likely contains factual or balanced content that does not lean towards an opinionated sentiment. The neutral classification reflects the model's ability to identify reviews that are neither overly positive nor negative, allowing businesses to gain insights from a variety of customer feedback.

V. CONCLUSION

This project has successfully developed a comprehensive sentiment analysis system that classifies customer reviews into positive, negative, or neutral categories. By integrating traditional machine learning classifiers such as Random Forest, Naive Bayes, Logistic Regression, and SVM, with advanced NLP models like VADER and BERT, the system is well-equipped to handle both short and long reviews. The ensemble approach ensures that the strengths of each model are leveraged, resulting in a highly accurate and reliable classification system. With an overall accuracy of 96%, the model demonstrates strong performance across key metrics such as precision, recall, and F1-score, ensuring that both positive and negative sentiments are correctly identified. The system's ability to process various forms of review text, including informal language and domain-specific terminology, further enhances its robustness. The Flask- based web interface offers an intuitive user experience, making it easy for businesses to upload datasets, analyze sentiments, and download results. This user-friendly design ensures that even non-technical users can utilize the tool effectively, gaining valuable insights into customer opinions and sentiments.

Despite its success, the project acknowledges areas for improvement, such as handling more complex nuances like sarcasm and ambiguous language. Future enhancements could involve refining the feature extraction process, incorporating additional data sources, and further fine- tuning deep learning models to improve performance. Overall, this sentiment analysis system provides a powerful tool for businesses to monitor customer feedback, manage brand reputation, and make data-driven decisions.

REFERENCES

[1]. Asogor, E. A., & Garba, E. J. (2024). Predicting Consumer Behaviour from E-Commerce Platform Product Reviews: A Machine Learning Approach to Sentiment Analysis. International Journal of Development Mathematics (IJDM), 1(3), 209-217.
[2]. Savci, P., & Das, B. (2023). Prediction of the customers' interests using sentiment analysis in e-commerce data for comparison of Arabic, English, and Turkish languages. Journal of King Saud University-Computer and Information Sciences, 35(3), 227-237.
[3]. Loukili, M., Messaoudi, F., & El Ghazi, M. (2023). Sentiment analysis of product reviews for e-commerce recommendation based on machine learning. International Journal of Advances in Soft Computing & Its Applications, 15(1).

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 14, Issue 2, February 2025
DOI: 10.17148/IJARCCE.2025.14260

[4]. Dieksona, Z. A., Prakosoa, M. R. B., Qalby, M. S., Putraa, M. S. A. F. S., Achmada, S., & Sutoyoa, R. (2023). "Sentiment Analysis of Customer Reviews: A Case Study of Traveloka." *Procedia Computer Science*, 216, 682-690.
[5]. Mohammed, M. A. (2022). Sentiment Based Product Recommendation System for E Commerce Using Machine Learning Approaches. International Journal of Innovative Research in Computer Science & Technology, 10(6), 120-137
[6]. J. C. Gope, T. Tabassum, M. M. Mabrur, K. Yu and M. Arifuzzaman, "Sentiment Analysis of Amazon Product Reviews Using Machine Learning and Deep Learning Models," 2022 International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE), Gazipur, Bangladesh, 2022, pp. 1-6, doi: 10.1109/ICAEEE54957.2022.9836420.
[7]. Mohammed, M. A. (2022). Sentiment Based Product Recommendation System for E Commerce Using Machine Learning Approaches. International Journal of Innovative Research in Computer Science & Technology, 10(6), 120-137.
[8]. Shaba, M., Roland, A., Simon, J., Misra, S., Ayeni, F. (2022). A Real-Time Sentimental Analysis on E-Commerce Sites in Nigeria Using Machine Learning. In: Abraham, A., et al. Hybrid Intelligent Systems. HIS
[9]. 2021. Lecture Notes in Networks and Systems, vol 420. Springer, Cham. https://doi.org/10.1007/978-3-030-96305-7_42
[10]. Basani, Y., Sibuea, H. V., Sianipar, S. I. P., & Samosir, J. P. (2019, March). Application of sentiment analysis on product review e-commerce. In Journal of Physics: Conference Series (Vol. 1175, No. 1, p. 012103). IOP Publishing.
[11]. Noor, F., Bakhtyar, M., & Baber, J. (2019). Sentiment analysis in e-commerce using svm on roman urdu text. In Emerging Technologies in Computing: Second International Conference, iCETiC 2019, London, UK, August 19–20, 2019, Proceedings 2 (pp. 213-222). Springer International Publishing.