iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: IV Month of publication: April 2025

DOI: https://doi.org/10.22214/ijraset.2025.68779

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

Sentiment Analysis of Customer Product Review Using Machine Learning

Abhishek Kamble[1], Shruti Chandane[2], Anuja Kamble[3], Urmila Pol[4], Parashuram S.Vadar[5], Tejashree T. Moharekar[6]
[1, 2, 3, 4]*Department of Computer Science, Shivaji University, Kolhapur*
[5, 6]*YCSRD, Shivaji University, Kolhapur,*

Abstract: Understanding customer opinions through sentiment analysis has become essential for businesses aiming to enhance customer satisfaction and refine their strategies. As digital platforms and e-commerce grow, consumers increasingly depend on online reviews for decision-making. This research delves into the analysis of customer product reviews using a range of sentiment analysis techniques, including lexicon-based approaches, traditional machine learning algorithms, and advanced deep learning models like Recurrent Neural Networks (RNN) and Transformer-based architectures such as BERT. The study emphasizes effective text preprocessing methods—such as tokenization, stemming, stopword elimination, and vectorization—to improve classification outcomes. Sentiment classification is assessed using metrics like accuracy, precision, recall, and F1-score to identify the most suitable model. Moreover, aspect-based sentiment analysis (ABSA) is employed to extract detailed opinions related to specific product features such as price, usability, durability, and customer service. The study also addresses challenges such as sarcasm detection, language diversity, and domain-specific interpretations. The insights from this research aim to support the development of intelligent sentiment analysis tools, enabling businesses to monitor feedback efficiently, tailor marketing strategies, and enhance customer experiences for long-term success.
Keywords: Sentiment analysis, e-commerce, Recurrent Neural Networks (RNN), BERT, Aspect-Based Sentiment Analysis (ABSA)

I. INTRODUCTION

In the era of digital transformation, customer opinions shared across online platforms have become a significant source of business intelligence. Reviews posted on e-commerce websites, social media, and other feedback channels reflect consumer experiences, expectations, and satisfaction levels.

Analyzing these reviews enables businesses to uncover insights into customer behaviour, highlight product strengths and weaknesses, and make informed decisions aimed at enhancing service quality and brand loyalty. The process of extracting meaningful information from large volumes of unstructured text data involves multiple stages, including data collection, text preprocessing, and structured analysis. One of the most effective techniques for this purpose is sentiment analysis, which seeks to determine the emotional tone conveyed in a given piece of text. It allows businesses to classify customer feedback into categories such as positive, negative, or neutral, and in some cases, even more nuanced sentiments.

With the advancement of machine learning (ML) and natural language processing (NLP), automated sentiment analysis has become increasingly reliable and scalable. ML models, when trained on labeled datasets of product reviews, can efficiently predict sentiments in real-time. Key preprocessing tasks like tokenization, stopword removal, and stemming are essential to prepare raw text for model training. Feature extraction techniques such as Bag of Words (BoW), TF-IDF, and word embeddings like Word2Vec or GloVe help convert textual data into numerical representations suitable for ML algorithms. Various classification models—including Logistic Regression, Support Vector Machines (SVM), Random Forest, Naive Bayes, and deep learning frameworks like Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs)—are employed to perform sentiment prediction. These models are evaluated based on standard performance metrics such as accuracy, precision, recall, and F1-score to ensure effective classification. Beyond simple sentiment tagging, modern approaches also explore aspect-based sentiment analysis (ABSA), offering more granular insights into specific product features like usability, price, or customer support. Applications of sentiment analysis span across customer service, product design, brand monitoring, and market research.

This research paper presents a comprehensive study of sentiment analysis techniques applied to customer product reviews using machine learning. It aims to identify the most efficient methodologies for understanding consumer sentiment and delivering actionable insights to support strategic business decisions.

II. REVIEW OF LITERATURE

Sentiment Analysis (SA), also known as opinion mining, involves the use of computational techniques to detect and classify emotions conveyed in textual content. When applied to customer product reviews, its primary objective is to assess the emotional tone of the feedback, commonly sorting it into categories such as positive, negative, or neutral. This technique has gained prominence in domains such as e-commerce, social media analytics, and customer experience evaluation, where companies strive to derive actionable insights from large volumes of unstructured text data [5].

This study conducts a comparative analysis of sentiment classification on customer reviews by employing both traditional machine learning algorithms and advanced deep learning models. The deep learning techniques explored include Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), while the conventional approaches encompass Logistic Regression, Random Forest, and Naïve Bayes classifiers. The analysis is based on a dataset of Amazon product reviews, with star ratings serving as an indicator of the underlying sentiment. The performance of each model is evaluated through detailed experiments, focusing on accuracy and overall effectiveness in sentiment detection. The findings offer critical insights into the advantages and limitations of each methodological approach, contributing to a deeper understanding of their applicability in real-world sentiment analysis task [2].

This study evaluates the performance of various machine learning classifiers—namely Random Forest, Naïve Bayes, and Support Vector Machine (SVM)—in comparison with the GPT-4 large language model (LLM) for sentiment analysis tasks. Traditional classifiers demonstrated higher efficiency when analyzing short and straightforward text inputs, with SVM particularly excelling in sentiment classification of brief comments. Conversely, GPT-4 achieved superior performance on longer and more descriptive reviews, effectively identifying nuanced sentiments. It outperformed traditional models in terms of precision, recall, and F1-score, especially when detecting mixed or complex emotional tones. The findings underscore the capability of LLMs like GPT-4 to surpass conventional methods in handling context-rich sentiment analysis, offering not only accurate classifications but also interpretative insights. Such capabilities make LLMs valuable tools for businesses aiming to extract actionable intelligence from textual customer feedback [4]. This study uses the PRISMA framework to systematically review literature on sentiment analysis. It examines key aspects such as application areas, evaluation methods, techniques, metrics, and tools. Product reviews were the most frequent application, with cross-validation widely used for model evaluation. SVM and LSTM emerged as the most accurate techniques. The F1-score was identified as the leading performance metric. Python dominated as the primary programming language. The review highlights both machine learning and deep learning applications in sentiment analysis. It also discusses current challenges and future directions, emphasizing its value in guiding business decisions [3].

This paper introduces an optimized machine learning approach, the Local Search Improvised Bat Algorithm-based Elman Neural Network (LSIBA-ENN), designed for sentiment analysis (SA) of online product reviews. The proposed methodology follows four key stages: data collection via web scraping, preprocessing, feature extraction/term weighting, and sentiment classification. Feature processing is enhanced using Log Term Frequency-based Modified Inverse Class Frequency (LTF-MICF) and Hybrid Mutation-based Earthworm Algorithm (HM-EWA). The processed data is classified using LSIBA-ENN into positive, negative, and neutral sentiments. Performance was evaluated using two benchmark datasets. Results demonstrate that LSIBA-ENN significantly outperforms existing models, particularly in recall metrics, where it achieved 87.79 using LTF-MICF, compared to lower scores with other feature weighting methods like W2V, TF, TF-IDF, and TF-DFS [10].

This study presents an optimized machine learning model, the Enhanced Golden Jackal Optimizer-based Long Short-Term Memory (EGJO-LSTM), for sentiment analysis of e-commerce product reviews. The proposed framework includes data collection, preprocessing, feature extraction and selection, followed by sentiment classification. Customer reviews are gathered using web scraping tools and then refined through preprocessing techniques. Log-term Frequency-based Modified Inverse Class Frequency (LF-MICF) and Improved Grey Wolf Optimizer (IGWO) are applied for feature weighting and selection. The processed data is subsequently classified using the EGJO-LSTM model into positive, negative, or neutral sentiments. Performance was assessed using an Amazon.com dataset, and the model was compared against leading machine learning methods using metrics like precision, accuracy, recall, and F1-score. Results indicate that EGJO-LSTM significantly outperforms traditional and hybrid models, showing improvements of 25% in precision and 32% in accuracy, particularly when used with the LF-MICF weighting approach [6].

This paper investigates the relationship between customer reviews on Amazon.com and their corresponding star ratings using a proposed framework. The study transforms the rating prediction task into a multi-class classification problem, where reviews are categorized into one of five classes based on their star ratings. The performance of various classifiers is evaluated and compared, with results indicating that Logistic Regression outperforms the other models. Additionally, the analysis reveals that factors such as review polarity and review length significantly influence the predicted star rating [8].

This study performs sentiment analysis on online reviews and uses the extracted sentiments as features to predict product ratings through various machine learning algorithms. The predictions were further analyzed using explainable AI (XAI) techniques to identify any biases in the models. In Study 1, algorithms such as k-NN, support vector machines, random forests, gradient boosting, and XGBoost were benchmarked, with random forests and XGBoost identified as the best for rating prediction. Study 2's global feature importance analysis highlighted "joy" sentiment and "negative emotional valence" as key predictors. XAI visualization methods, including local feature attributions and partial dependency plots, uncovered instance-level prediction errors. Study 3, focusing on classification benchmarking, found a high no-information rate of 64.4%, suggesting class imbalance as a contributing factor. In conclusion, while machine learning models performed well, caution is needed, as biases in the dataset may lead to skewed predictions. This study illustrates how XAI methods can uncover such biases [7].

This study analyzes unstructured Amazon product reviews, which are typically disorganized, using sentiment analysis to efficiently process large volumes of data. The proposed method combines an ensemble approach with Naïve Bayes, K-Nearest Neighbors (KNN), Long Short-Term Memory (LSTM), and the Natural Language Toolkit (NLTK). The dataset of customer reviews is first processed and grouped using the LSTM and KNN models, followed by classification with the Naïve Bayes model. The resulting model successfully classifies Amazon reviews into positive, negative, and neutral sentiments, achieving accurate sentiment analysis [1].

This paper introduces the concept of Review Reading Comprehension (RRC), which aims to transform customer reviews into a valuable knowledge source for answering user questions. To the best of our knowledge, RRC has not been previously explored. The study creates a new RRC dataset, Review RC, based on a benchmark for aspect-based sentiment analysis. Due to the limited training data available for both RRC and aspect-based sentiment analysis, the paper proposes a novel post-training method for the BERT language model to improve its fine-tuning performance for RRC tasks. The post-training approach is also tested on other review-based tasks, such as aspect extraction and aspect sentiment classification, with experimental results showing its effectiveness in enhancing model performance [9].

III. METHODOLOGY

Sentiment analysis of customer product reviews using machine learning (ML) entails a series of steps to automate the classification of customer opinions into positive, negative, or neutral categories. The typical methodology involves data collection, preprocessing, feature extraction, model selection, training, evaluation, and deployment. Each of these stages plays a crucial role in ensuring the accurate and efficient analysis of customer feedback.

The following outlines the key steps involved in the proposed methodology:

A. Data Collection

The initial step in sentiment analysis is to gather a substantial set of customer product reviews. These reviews are typically collected from various online sources, such as E-commerce platforms (Amazon, eBay, etc.), Social media networks (Twitter, Facebook), Product review websites (Yelp, TripAdvisor) and Direct customer feedback forms the dataset should consist of both the text of the reviews and their corresponding sentiment labels, which may be pre-labelled or manually annotated for supervised learning.

B. Data Preprocessing

The collected raw reviews are often noisy and require preprocessing to prepare them for machine learning. Key preprocessing steps include Text Cleaning in which Removing special characters, stopwords, and converting the text to lowercase to standardize it takes place. In Tokenization the text splits into individual tokens (e.g., words or phrases). Spelling Correction Corrects spelling mistakes if present reduces words to their base or root form, either by stemming (e.g., “running” to “run”) or lemmatization (e.g., “better” to “good”). For Handling Class Imbalance Techniques such as oversampling or under sampling can be used to address the imbalance in class distributions, common in many real-world datasets.

C. Feature Extraction

The next step is converting the preprocessed text into numerical features that can be input into machine learning models. Common feature extraction techniques include Bag of Words (BoW) which represents text as a frequency count of words, with each document being converted into a vector. TF-IDF (Term Frequency-Inverse Document Frequency) Measures the importance of a word in a document relative to the entire corpus, giving more weight to rare words. For Word Embeddings; Advanced methods like Word2Vec, GloVe, or FastText generate dense vector representations that capture semantic meanings of words. N-grams Captures sequences of 'n' words to provide more context than individual words.

D. Model Selection

After feature extraction, a suitable machine learning model is chosen to classify sentiment. Commonly used models include Naive Bayes is a probabilistic classifier based on Bayes' theorem that assumes feature independence. Support Vector Machine (SVM) is a powerful classifier that finds the hyperplane that best separates different classes in high-dimensional spaces. Logistic Regression model for predicting binary outcomes based on input features and Random Forest and Decision Trees; Ensemble methods that use multiple decision trees to improve classification performance.

E. Model Training

In this phase, the selected model is trained using the training dataset. Training involves finding patterns between the features and the target sentiment labels. Key steps during training include Cross-Validation for Evaluating the model's performance by splitting the data into subsets and training/testing on each subset. Hyperparameter Tuning for Optimizing model parameters, such as learning rate or the number of trees in a random forest, using techniques like grid search or random search.

F. Model Evaluation

Once the model is trained, it is evaluated on a separate test dataset to gauge its performance. Evaluation metrics include Accuracy means the proportion of correct predictions out of the total number of predictions. Precision, Recall, and F1-Score these are metrics that are particularly useful for imbalanced datasets. Confusion Matrix Visualizes the model's performance by showing the true positives, false positives, true negatives, and false negatives. ROC Curve and AUC measures the trade-off between the true positive rate and false positive rate.

G. Model Deployment

After evaluation, the model can be deployed for real-world use. Deployment steps include Model Serving were Hosting the model on a server or cloud platform to provide real-time predictions. API Integration for developing an API for the model to process incoming reviews and classify their sentiment and monitoring for Continuously tracking the model's performance to ensure it adapts to new trends and data shifts.

H. Continuous Improvement

To maintain model accuracy over time, ongoing improvements are necessary In Model Retraining, regularly retraining the model with updated data to capture new patterns. Active Learning for incorporating human feedback to correct uncertain predictions and enhance the model's learning and Model Drift Detection Monitors shifts in data distribution to detect and address any degradation in model performance.

The results of sentiment analysis of customer product reviews using machine learning (ML) methods reveal insights into both the effectiveness of the model and the customer sentiments expressed in the reviews. The analysis begins with the examination of sentiment distribution, categorizing reviews into positive, negative, and neutral sentiments. For example, after applying sentiment classification, the distribution could show that 60% of the reviews are positive, 30% are negative, and 10% are neutral. This provides a general understanding of how customers feel about the product.

A confusion matrix is an important tool for evaluating the model's performance. It compares the model's predicted sentiment labels with the actual sentiments. In a typical example, true positives (TP) for positive reviews could be as high as 85%, while false positives (FP) might be around 5%. True negatives (TN) could be 80%, and false negatives (FN) could be 10%. This matrix provides insights into how well the model is classifying each sentiment category and highlights areas for improvement.

Evaluation metrics, such as accuracy, precision, recall, and F1-score, are crucial for assessing the model’s overall performance. In this case, the model could achieve 85% accuracy, with a precision of 90% for positive sentiment, a recall of 80% for positive sentiment, and an F1-score of 85%. Additionally, the Area Under the ROC Curve (AUC) might show a value of 0.9, indicating strong performance in distinguishing between positive and negative sentiments.

The analysis might also include a comparison between different machine learning models. For instance, after testing models like Naive Bayes, Support Vector Machine (SVM), Logistic Regression, Random Forest, and Deep Learning (LSTM), the LSTM model might outperform others with a 91% accuracy. This suggests that LSTM is particularly effective at capturing long-range dependencies in the text, which is beneficial for complex sentiment expressions.

Insights from the sentiment analysis could include customer feedback on product features, highlighting strengths and weaknesses. Positive reviews may focus on aspects like product quality and ease of use, while negative reviews may raise concerns about delivery issues or product defects. These insights can be valuable for improving products or shaping marketing strategies.

Despite the promising results, several challenges remain. Ambiguity in text, such as sarcasm or mixed emotions, can hinder the model's ability to accurately classify sentiments. Additionally, the model may struggle with understanding the contextual meaning of words, as their sentiment can change depending on the context. An imbalanced dataset, where one sentiment category dominates, can also bias the model towards the majority class, leading to suboptimal performance for the minority class.

In conclusion, while machine learning-based sentiment analysis provides valuable insights into customer opinions, continuous model improvement and careful attention to data quality are necessary for maintaining accuracy and handling the complexities of natural language.

IV. CONCLUSION

In conclusion, sentiment analysis of customer product reviews using machine learning has become a valuable tool for businesses, providing automated insights into customer opinions. The methodology involves stages like data collection, preprocessing, model selection, and continuous improvement, enabling businesses to leverage feedback for better decision-making. Deep learning techniques have enhanced sentiment classification accuracy, allowing businesses to gain actionable insights from large datasets. Key benefits include improved customer insights, better decision-making, scalability, and real-time processing. Despite challenges such as ambiguity and domain-specific terminology, the need for continuous model improvement remains. Sentiment analysis empowers businesses to stay competitive, refine products, and enhance customer satisfaction in a dynamic market.

REFERENCES

[1] J, I., & Kumar G, N., "Amazon Product Reviews Sentiment Analysis using Machine Learning", JETIR, 234-239, 2022.

[2] Ashbaugh, L., & Zhang, Y., "A Comparative Study of Sentiment Analysis on Customer Reviews Using Machine Learning and Deep Learning", MDPI Computers, 1-16, 2024.

[3] Daza, A., Gonzalez ´ Rueda, N., Sanchez, M. S., Espíritu, W. R., & Quinones, M. C., "Sentiment Analysis on E-Commerce Product Reviews Using Machine Learning and Deep Learning Algorithms: A Bibliometric Analysis, Systematic Literature Review, Challenges and Future Works", International Journal of Information Management Data Insights, 1-20, 2024.

[4] Ghatora, P. S., Hosseini, S. E., Pervez, S., Iqbal, M. J., & Shaukat, N., "Sentiment Analysis of Product Reviews Using Machine Learning. Big Data Cognitive Computing" MDPI, 1-18, 2024.

[5] Pang, B., & Lee, L., "Opinion mining and sentiment analysis", 2008.

[6] Rasappan, P., Premkumar, M., Sinha, G., & Chandrasekara, K., "Transforming sentiment analysis for e-commerce product reviews: Hybrid deep learning model with an innovative term weighting and feature selection", Information Processing & Management, 2024.

[7] So, C., "What Emotions Make One or Five Stars? Understanding Ratings of Online Product Reviews by Sentiment Analysis and XAI", 1st International Conference on Artificial Intelligence in HCI, AI-HCI. Kopenhagen, Denmark, Springer, 2020.

[8] Taparia, A., & Bagla, T., "Sentiment Analysis: Predicting Product Reviews' Ratings using Online Customer Reviews", SSRN, 2020.

[9] Xu, H., Liu, B., Shu, L., & Yu, P. S., "BERT Post-Training for Review Reading Comprehension and Aspect-based Sentiment Analysis", arxiv logo, 2019.

[10] Zhao, H., Liu, Z., Yao, X., & Yang, Q., "A machine learning-based sentiment analysis of online product reviews with a novel term weighting and feature selection approach", Information Processing & Management, 2021.