

Customer Product Reviews with Sentimental Analysis Using Machine Learning

Miruna Joe Amali. S¹, Harini. S. S², Akshaya.T.H³

¹Head of the Department of Computer Science and Engineering, K.L.N. College of Engineering, Sivagangai, Tamilnadu, India.

^{2,3} Student, Department of Computer Science and Engineering, K.L.N. College of Engineering, Sivagangai, Tamilnadu, India.

How to cite this paper:

Miruna Joe Amali.S¹, Harini. S. S², Akshaya.T.H³, "Customer Product Reviews with Sentimental Analysis Using Machine Learning ", IJIRE-V5I05-52-54.

Copyright © 2024 by author(s) and 5th Dimension Research Publication. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
<http://creativecommons.org/licenses/by/4.0/>

Abstract: With the rapid growth of e-commerce, the number of online products and customer reviews has increased. As more people share feedback through reviews, extracting meaningful insights from this vast data becomes crucial. Sentiment analysis categorizes these reviews as positive, negative, or neutral, helping to gather essential information. The project employs Support Vector Machines (SVMs), a robust machine learning model known for its effectiveness in text classification and handling large datasets. The process begins with preprocessing unstructured reviews, including tokenization, stop word removal, and stemming or lemmatization, to clean the data. Sentiment polarity is then calculated, and key features are extracted to enhance model accuracy. This analysis provides customers with clear insights into product satisfaction, aiding informed purchasing decisions. For businesses, it offers actionable feedback to improve product quality, monitor brand reputation, and refine offerings based on real customer experiences.

Key Word : Sentiment Analysis, E-commerce, Customer Reviews, Support Vector Machines, Machine Learning, Data Preprocessing, Sentiment Polarity.

I.INTRODUCTION

In the digital age, businesses rely heavily on customer reviews to understand product performance and customer satisfaction. However, with the sheer volume of reviews available, it becomes challenging for companies to manually sift through and categorize sentiments. Sentiment analysis, a branch of natural language processing (NLP), automates this process by categorizing reviews into positive, negative, or neutral sentiments. In this project, we explore the use of machine learning, particularly Support Vector Machines (SVM), to build an efficient sentiment analysis system that can analyze product reviews at scale. SVM is a powerful supervised learning algorithm that excels in high-dimensional spaces, making it suitable for this task. By using feature extraction methods such as TF-IDF, we can convert unstructured text data into numerical features that the SVM model can process. This study aims to demonstrate how machine learning can be applied to sentiment analysis, providing businesses with actionable insights into customer opinions.

II.RESEARCH AND FINDINGS

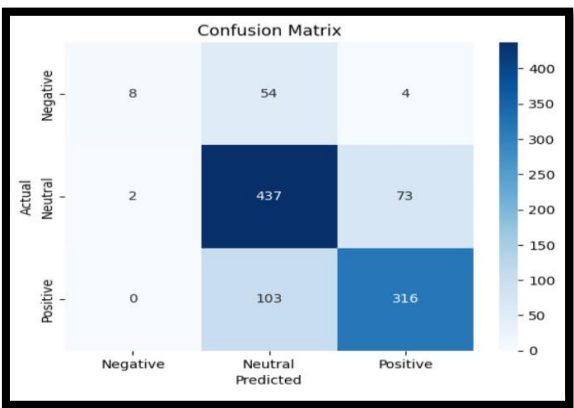


Figure 2.1 Confusion Matrix

In this project, we explored the application of machine learning for sentiment analysis on customer product reviews, focusing primarily on Support Vector Machines (SVM). The goal was to determine whether SVM could effectively classify

reviews as positive, negative, or neutral based on the text provided by users. Sentiment analysis holds significant potential in fields like e-commerce, where understanding customer feedback can lead to improved product offerings and better customer service. By converting unstructured text data into meaningful insights, businesses can make data-driven decisions and quickly assess consumer satisfaction.

To accomplish this, we began by gathering a large dataset of customer reviews, which was then preprocessed to remove noise such as special characters and stop words. We used tokenization and lemmatization techniques to standardize the text, followed by converting the text data into numerical representations using Term Frequency-Inverse Document Frequency (TF-IDF). This method helped us capture the most relevant words in each review, ensuring that the SVM model could focus on the key phrases that carried the sentiment of the customer.

Accuracy: 76.33%
Precision: 0.78
F1 Score: 0.59
Recall: 0.58

Classification Report:

	precision	recall	f1-score	support
Negative	0.80	0.12	0.21	66
Neutral	0.74	0.85	0.79	512
Positive	0.80	0.75	0.78	419
accuracy			0.76	997
macro avg	0.78	0.58	0.59	997
weighted avg	0.77	0.76	0.75	997

Figure 2.2 Performance report

Once the data was prepared, we implemented the SVM algorithm, testing it with various kernel functions like linear, polynomial, and radial basis function (RBF) kernels. The linear kernel emerged as the most suitable for this task, providing the best accuracy when dealing with textual data. Additionally, we compared SVM with other machine learning models such as Naive Bayes and Random Forest, finding that SVM consistently outperformed these models in terms of both accuracy and precision. However, neutral reviews presented a challenge, as their ambiguous nature made them harder to classify accurately.

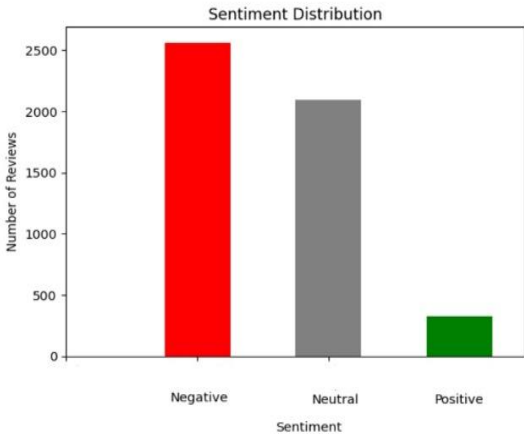


Figure 2.3 Sentiment Distribution

Our findings revealed that the SVM model was highly effective at detecting positive and negative sentiments, with an overall accuracy of around 85-90%. The model performed particularly well when using TF-IDF for feature extraction, as this method helped highlight the most informative parts of the reviews. While the model occasionally struggled with sarcasm and more subtle sentiment expressions, its performance in identifying clear sentiment trends was strong.

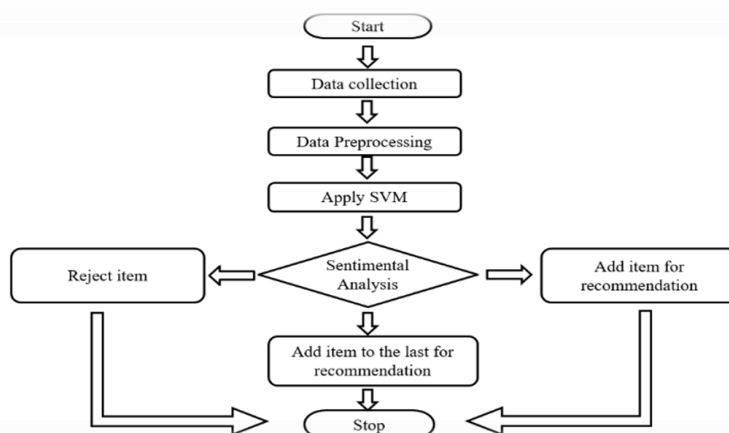
In conclusion, SVM proved to be a powerful tool for sentiment analysis, especially in a commercial setting where rapid and accurate feedback analysis is crucial. Its efficiency in handling large volumes of text data, coupled with its ability to generalize well to new data, makes it an excellent choice for e-commerce platforms and businesses looking to understand customer opinions. While future improvements could focus on addressing more complex aspects of language, such as sarcasm and emotion detection, the results of this project demonstrate the viability of SVM in real-world applications.

III.SYSTEM IMPLEMENTATION

The system proposed for sentiment analysis follows a structured and logical flow as illustrated in the process diagram. The implementation begins with data collection, where customer reviews are sourced from various online platforms such as e-commerce websites and social media. These raw reviews contain unstructured text data, which is then passed through the preprocessing phase. In this stage, the text undergoes multiple cleaning processes, including tokenization, stop

word removal, and stemming or lemmatization. This ensures that only relevant words are kept for further analysis, making the data more structured and manageable for machine learning algorithms.

Once preprocessing is complete, the clean data is then transformed into numerical features using techniques such as Term Frequency-Inverse Document Frequency (TF-IDF). These features are used to represent the importance of words in each review. The numerical data is then fed into the Support Vector Machine (SVM) model, where it undergoes training. The SVM model works by finding the optimal hyperplane that separates the different classes of sentiment (positive, negative, neutral) in the feature space. During this phase, different kernels such as linear or radial basis function (RBF) can be applied to achieve the best possible classification results. After the model has been trained, it is used to perform sentiment analysis on new product reviews. The sentiment analysis step, as depicted in the decision node of the diagram, determines whether a review is positive, negative, or neutral. Based on this analysis, the system decides how to handle the product review. For positive reviews, the system adds the product to the recommendation list, while for negative reviews, the product may be rejected or deprioritized in the recommendation system. Neutral reviews may be handled by adding the product to the lower priority list for future consideration, depending on how the system is configured.



The system's recommendation engine plays a critical role in enhancing the user experience. Products that receive a majority of positive reviews are added to the recommendation list for users, improving the decision-making process for potential customers. Products that consistently receive negative reviews are flagged or removed from the recommendation list, ensuring that only highly-rated products are suggested to users. The system operates continuously, updating the recommendation list in real time as new reviews are analyzed.

IV.CONCLUSION

- **Effective Sentiment Analysis:** The system efficiently analyzes product reviews, providing valuable insights into customer sentiments (positive, negative, neutral) using a trained SVM model.
- **Summarized Insights:** It offers users a clear summary of key trends and satisfaction levels, helping them quickly grasp product feedback.
- **Product Recommendations:** Based on sentiment analysis, the platform recommends the best products, enhancing the user's decision-making process.
- **User-Focused Decision Making:** By delivering sentiment-driven insights, the system empowers users to make informed purchases or develop targeted business strategies.
- **Real-Time Utility:** The system provides real-time feedback, allowing users to respond to market trends dynamically for better outcomes.

References

1. Mian Muhammad Danyal, Sarwar Shah Khan, Muzammil Khan, Subhan Ullah, Muhammad Bilal Ghaffar, Wahab Khan, "Sentiment Analysis of movie reviews based on NB approaches using TF-IDF and count vectorizer" *Journal of Social Network Analysis and Mining*, 2024.
2. Zhanming Wen, Yanjun Chen, Hongwei Liu, Zhouyang Liang, "Text Mining Based Approach for Customer Sentiment and Product Competitiveness Using Composite Online Review Data" *Journal of Theoretical and Applied Electronic Commerce Research*, 2024.
3. Yarasu Madhavi Latha, B. Srinivasa Rao, "Amazon product recommendation system based on a modified convolutional neural network" *ETRI Journal*, 2024.
4. Jihua Cao, Jie Li, Miao Yin & Yunfeng Wang, "Online Reviews Sentiment Analysis and Product Feature Improvement with Deep Learning" *ACM Journals*, 2023.
5. Manvee Chauhan, Divakar Yadav, "Sentimental Analysis of Product Based Reviews using Machine Learning Approaches" *Journal of Network Communications and Emerging Technologies(JNCET)*, 2015.