

SENTIMENT ANALYSIS SYSTEM USING PYTHON

Jayashree Janani Murugesan¹, Devadharshini S², Loganarayanan R³, Dr. R. Devi⁴

^{1,2,3}Final Year BCA – Data Science (UG Students) Department of Applied Computing & Emerging Technologies School of Computing Sciences Vels Institute of Science, Technology and Advanced Studies (VISTAS) Pallavaram, Chennai – 600117, Tamil Nadu, India.

⁴Guide: M.C.A., M.Phil., Ph.D. Professor & Head Department of Applied Computing & Emerging Technologies School of Computing Sciences Vels Institute of Science, Technology and Advanced Studies (VISTAS) Chennai – 600117, Tamil Nadu, India.

DOI: <https://www.doi.org/10.58257/IJPREMS52947>

ABSTRACT

Sentiment analysis is one of the most important applications of Natural Language Processing (NLP) used to identify opinions, emotions, and attitudes from textual data. In the modern digital world, customers continuously provide feedback through e-commerce websites, online shopping applications, service review portals, and social media platforms. These reviews contain valuable information regarding customer satisfaction, product quality, service efficiency, and overall brand reputation. Manual analysis of such large volumes of customer reviews is highly time-consuming, inconsistent, and inefficient. This project presents a Python-based Sentiment Analysis System developed to automatically classify customer reviews into positive, negative, and neutral sentiments. The system follows a complete analytical workflow including data collection, preprocessing, feature extraction, model development, evaluation, and visualization. Text preprocessing methods such as lowercasing, punctuation removal, stop-word elimination, tokenization, stemming, lemmatization, and TF-IDF vectorization are applied to improve classification accuracy and model efficiency. Machine learning algorithms are used for sentiment prediction and performance evaluation is carried out using Accuracy, Precision, Recall, F1-Score, and Confusion Matrix. Visualization tools such as bar charts, word clouds, and sentiment distribution graphs improve result interpretation. The proposed system helps organizations improve customer satisfaction, service quality, and strategic business decision-making through intelligent customer feedback analysis.

Keywords: Sentiment Analysis, Natural Language Processing, Machine Learning, Python, Customer Reviews, Opinion Mining, Text Classification, TF-IDF.

1. INTRODUCTION

In today's digital environment, customers frequently express their opinions regarding products and services through online reviews, social media platforms, surveys, feedback forms, and discussion forums. These reviews provide valuable business intelligence and help organizations understand customer satisfaction, market expectations, and product performance. Businesses use this information to improve services, strengthen customer relationships, and enhance brand reputation.

With the rapid growth of digital platforms, organizations receive thousands of customer reviews every day. Manual analysis of such large-scale feedback is difficult, expensive, and often inaccurate due to human limitations. Traditional review analysis methods fail to provide fast, scalable, and consistent insights when handling massive textual datasets.

Sentiment analysis provides an automated solution by identifying whether customer feedback expresses positive, negative, or neutral emotions. Using Natural Language Processing and Machine Learning techniques, text can be processed efficiently and converted into meaningful business intelligence. This project aims to develop a practical and reliable Sentiment Analysis System using Python for effective customer feedback analysis.

2. OBJECTIVES OF THE STUDY

The major objectives of the proposed system are listed below:

- To collect customer review datasets from publicly available online sources
- To preprocess raw textual data for improved analytical quality
- To perform tokenization, stop-word removal, stemming, and lemmatization
- To apply TF-IDF vectorization for feature extraction
- To classify reviews into positive, negative, and neutral sentiments
- To evaluate model performance using standard classification metrics
- To generate bar charts, word clouds, and confusion matrix reports
- To improve customer satisfaction analysis using automated systems
- To reduce manual effort in review monitoring and sentiment identification



www.ijprems.com
editor@ijprems.com

**INTERNATIONAL JOURNAL OF PROGRESSIVE
RESEARCH IN ENGINEERING MANAGEMENT
AND SCIENCE (IJPREMS)**

(Int Peer Reviewed Journal)

Vol. 06, Issue 04, April 2026, pp : 2409-2411

**e-ISSN :
2583-1062**

**Impact
Factor :
7.001**

-
- To create a scalable and practical Python-based sentiment analysis platform

3. LITERATURE SURVEY

Liu (2015) explained the importance of opinion mining and established the foundation of modern sentiment analysis by focusing on extracting sentiments and emotions from textual content. His work showed how businesses can use customer opinions for better strategic planning and service quality improvement. Jurafsky and Martin (2023) discussed advanced Natural Language Processing techniques including text classification, semantic understanding, and machine learning-based sentiment prediction. Their research provides strong theoretical support for NLP-driven business systems. Bird, Klein, and Loper (2009) introduced practical implementations of NLP using Python and demonstrated preprocessing techniques such as tokenization, stop-word removal, stemming, and lemmatization using the NLTK library. Raschka and Mirjalili (2020) explained machine learning algorithms suitable for classification tasks and emphasized feature extraction methods such as TF-IDF vectorization. Existing systems mainly focus on basic classification accuracy but often lack visualization support and complete workflow integration. This project improves these limitations by combining preprocessing, classification, evaluation, and business reporting into a unified solution.

4. PROBLEM DEFINITION

Organizations receive massive amounts of customer feedback from online platforms, e-commerce websites, and service review systems. Traditional manual analysis of reviews is inefficient and unable to handle large datasets effectively. Businesses often struggle to identify customer dissatisfaction quickly, leading to delayed service improvement and reduced customer retention. Many existing systems provide only basic review monitoring and lack automated sentiment classification, real-time interpretation, and visual analytical support. Without proper sentiment analysis, organizations cannot respond effectively to customer dissatisfaction or identify opportunities for product improvement. Therefore, there is a strong need for an intelligent and scalable sentiment analysis system that can automatically process textual reviews, classify emotional tone accurately, and provide meaningful insights for faster and better decision-making.

5. METHODOLOGY

Data Collection

Customer review datasets are collected from publicly available sources such as Amazon product reviews, Flipkart customer feedback, restaurant review systems, hotel review portals, and open-source sentiment datasets. These datasets contain textual opinions related to products and services across different business domains.

Data Preprocessing

Preprocessing is one of the most important stages in sentiment analysis because raw text often contains punctuation, symbols, repeated words, unnecessary stop words, and inconsistent formatting. The following preprocessing steps are applied:

- Lowercasing all text
 - Removing punctuation and special characters
 - Stop-word removal
 - Tokenization
 - Stemming
 - Lemmatization
 - Removal of duplicate entries
 - Handling missing values
 - TF-IDF vectorization for numerical representation
- These steps improve text quality and convert unstructured text into structured numerical form suitable for machine learning models.

6. MODELING AND ANALYSIS

The developed system classifies customer reviews into three major categories:

- Positive Sentiment
- Negative Sentiment
- Neutral Sentiment

After preprocessing and feature extraction, machine learning algorithms are applied for sentiment prediction. The model is trained using labelled review datasets and tested using unseen customer review data for validation.

Performance evaluation includes:

- Accuracy

Precision

Recall

F1-Score

Confusion Matrix

These performance metrics help determine the reliability, consistency, and effectiveness of the proposed sentiment analysis model. Higher values indicate stronger classification performance and better practical usability.

7. RESULTS AND DISCUSSION

The sentiment analysis system successfully classifies customer reviews based on emotional tone and textual opinion patterns. Positive reviews commonly contain words such as excellent, good, satisfied, happy, and amazing. Negative reviews include terms such as poor, bad, disappointed, terrible, and unsatisfied. Neutral reviews usually contain balanced or informational feedback. Visualization techniques such as bar charts, word clouds, confusion matrices, and sentiment distribution graphs improve interpretation of classification results. These graphical outputs help organizations understand customer satisfaction levels quickly and accurately. Compared to manual review analysis, the proposed system improves speed, consistency, scalability, and business intelligence support. The results confirm that automated sentiment analysis significantly improves customer feedback management and supports stronger decision-making processes.

8. ADVANTAGES OF THE PROPOSED SYSTEM

- Reduces manual effort in customer review analysis
- Improves classification speed and consistency
- Supports better product and service improvement
- Helps organizations identify dissatisfied customers quickly
- Improves customer relationship management
- Provides scalable analysis for large review datasets
- Enhances strategic business decision-making
- Offers practical implementation using Python and machine learning
- Supports future integration with dashboards and chatbot systems

9. CONCLUSION

The project “Sentiment Analysis System Using Python” successfully demonstrates how customer feedback can be analysed automatically using Natural Language Processing techniques supported by Python implementation. The system effectively classifies customer reviews into positive, negative, and neutral sentiments and helps organizations understand customer opinions more efficiently. The implementation reduces manual effort, improves response speed, enhances service quality, and supports better business decision-making. It also provides organizations with valuable customer insights for product improvement and customer satisfaction management. Future enhancements may include deep learning models, multilingual sentiment analysis, real-time dashboard integration, chatbot support, and cloud deployment for enterprise-level applications. Overall, the project highlights the practical importance of sentiment analysis in modern business environments and AI-driven decision support systems.

ACKNOWLEDGEMENT

The authors sincerely thank Dr. R. Devi, M.C.A., M.Phil., Ph.D., Professor & Head, for valuable guidance, continuous encouragement, and technical support throughout the successful completion of this project. The authors also thank the Department of Applied Computing & Emerging Technologies and Vels Institute of Science, Technology and Advanced Studies (VISTAS) for providing the academic environment, institutional support, and resources required for this research work.

10. REFERENCES

- [1] Liu, B. (2015). Sentiment Analysis: Mining Opinions, Sentiments, and Emotions.
- [2] Jurafsky, D., & Martin, J. H. (2023). Speech and Language Processing.
- [3] Bird, S., Klein, E., & Loper, E. (2009). Natural Language Processing with Python.
- [4] Raschka, S., & Mirjalili, V. (2020). Python Machine Learning.
- [5] Pedregosa, F. et al. (2011). Scikit-learn: Machine Learning in Python.
- [6] McKinney, W. (2018). Python for Data Analysis.
- [7] Scikit-learn Official Documentation.
- [8] Provost, F., & Fawcett, T. (2013). Data Science for Business.