ISSN (Online): 2583-648X

IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

Volume 4 | Issue 5 | October 2025

www.ijirid.in

Sentiment Analysis for Product Review using Machine Learning and NLP

Prof. Chetan Saharkar[1], Manjiri Bhusari[2], Khushi Bayskar[3], Ekta Bodke[4], Prachi Maske[5]

[1]*Assistant Professor, G. H. Raisoni University, Amravati, India*

[2,3,4,5]*Undergraduate Student, G. H. Raisoni University, Amravati, India*

Abstract: *This abstract outlines a machine learning (ML) approach for performing sentiment analysis on product reviews using Natural Language Processing (NLP). The methodology involves utilizing NLP techniques to preprocess textual data and applying ML models, such as Naive Bayes and Support Vector Machines, to classify review sentiment into categories like positive, negative, or neutral. This automation provides businesses with scalable tools to efficiently analyze extensive customer feedback, gauge overall satisfaction, identify specific product strengths and weaknesses, and inform data-driven decisions for future product enhancements and marketing strategies.*

Keywords: Sentiment Analysis, Opinion Mining, Natural Language Processing (NLP), Machine Learning (ML), Product Reviews, Customer Feedback, Text Classification, Positive/ Negative/ Neutral Sentiment.

I. INTRODUCTION

Sentiment analysis is a crucial application of Machine Learning (ML) and Natural Language Processing (NLP) designed to systematically identify, extract, and categorize subjective information from textual data, such as product reviews. By utilizing various ML techniques—from foundational statistical models like logistic regression to advanced deep learning architectures such as Transformers (e.g., BERT) systems automatically ascertain sentiment polarity (positive, negative, or neutral). This automated process transforms large volumes of unstructured text into structured, quantifiable insights.

The primary goal is to help businesses efficiently monitor brand reputation, understand general customer sentiment, and derive actionable feedback to inform product development, marketing campaigns, and ultimately enhance overall customer satisfaction strategies.

- **Problem Definition:**

  The problem definition centers on the inability of traditional manual review processes to efficiently manage the enormous, ever-growing volume of unstructured textual feedback found in online product reviews. This manual approach is slow, costly, inconsistent, and prevents businesses from deriving timely, scalable insights. The objective is to apply Machine Learning (ML) and Natural Language Processing (NLP) techniques to automate the classification of sentiment (positive, negative, neutral) from this text data accurately, transforming raw feedback into quantifiable metrics for immediate business intelligence and improved decision-making.

 DOI: 10.5281/zenodo.17804888

ISSN (Online): 2583-648X

IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

Volume 4 | Issue 5 | October 2025

www.ijirid.in

- **Objectives:**

  The objective is to develop and implement an automated system using Machine Learning (ML) and Natural Language Processing (NLP) to accurately classify the sentiment polarity (positive, negative, or neutral) within large volumes of unstructured product review text. This system aims to transform raw customer feedback into quantifiable, actionable insights, enabling businesses to efficiently monitor brand perception and make data-driven decisions to enhance products and customer satisfaction.

II. LITERATURE REVIEW

NLP is a subset of computer science connected with computational linguistics (2021) [1]. It enables seamless communication between humans and machines by teaching computers to understand and interpret human language; hence, through NLP, machines can process and analyze text, offering insights and responses to users (2024) [2].

NLP offers the potential to develop models and processes capable of extracting information from both text and audio data (2023) [3]. This work is based on the implementation of the Bidirectional Encoder Representations from Transformers (BERT) to enhance natural language understanding. Moreover, this work has shown a strong ability to grasp context and their significance (2023) [4].

Here are plenty of use-cases across the field, but to narrow down the scope, few use- cases related to product-based organizations are taken into consideration. In the perspective variable sizes of organizations, sentiment analysis emerges as a crucial tool in this context for several reasons. Sentiment analysis is widely used in various sectors, including marketing, customer support, and product design, as it plays a vital role in understanding customer views and responses (2021) [5].

More recently, deep learning models have revolutionized sentiment analysis. Architectures such as Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and Transformer-based models like BERT (Bidirectional Encoder Representations from Transformers) have demonstrated superior performance by capturing complex linguistic patterns and contextual dependencies (2021) [6].

III. METHODOLOGY

The methodology involves collecting and preprocessing raw reviews. NLP is used for feature extraction. An ML model is trained on this data to classify sentiments into positive, negative or neutral.

3.1 Data Collection:

Raw product review text is systematically collected from various online platforms like e-commerce websites and social media.

3.2 Data Preprocessing:

The methodology for sentiment analysis of product reviews using machine learning and NLP involves several stages, from data collection and preprocessing to model training and evaluation.

3.3 Feature Extraction:

Feature extraction is a critical step in the methodology for sentiment analysis, as it involves converting raw, unstructured text data into a numerical format (vectors) that machine learning algorithms can process and learn from.

ISSN (Online): 2583-648X

IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

Volume 4 | Issue 5 | October 2025

www.ijirid.in

3.4 Data Labeling:

Data labeling is a fundamental step in the methodology for building a supervised sentiment analysis model. The goal of labeling is to assign a predefined sentiment category (class or label) to each product review text, creating the necessary input for machine learning algorithms to learn from.

3.5 Algorithm Selection:

The algorithm section within the methodology details the specific machine learning models chosen to analyze the text features generated during preprocessing and feature extraction phases. The selection of an algorithm depends heavily on the type of features extracted (e.g., TF-IDF vectors vs. word embeddings) and the desired complexity of the model.

IV. SYSTEM ARCHITECTURE

The Architecture for sentiment analysis using Machine Learning (ML) and Natural Language Processing (NLP) involves a sequential workflow:

- Data Collection Raw text data is gathered from sources like e-commerce reviews or social media.
- Preprocessing The text is cleaned (removing noise like HTML tags, stop words, and punctuation) and tokenized to prepare it for analysis.
- Feature Engineering Text is converted into numerical representations using techniques such as Bag-of-Words or TF-IDF.
- Model Training & Classification A machine learning model (e.g., Naive Bayes, SVM, or deep learning network) is trained on labeled data to classify sentiments into categories (positive, negative, neutral).
- Output & Visualization Results are aggregated, visualized (e.g., graphs), and used for business insights.

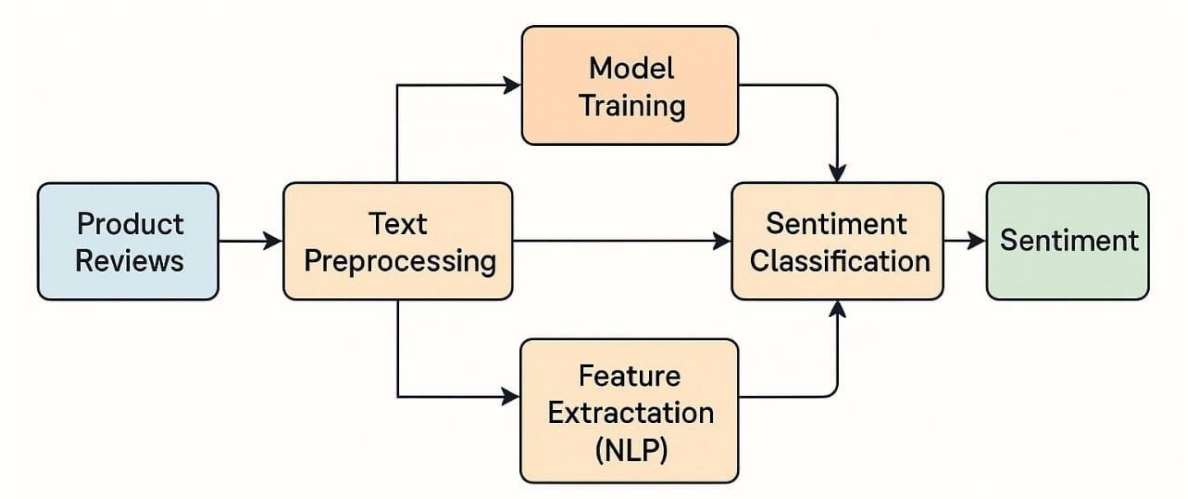


Figure 4.1: System Architecture

ISSN (Online): 2583-648X

IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

Volume 4 | Issue 5 | October 2025

www.ijirid.in

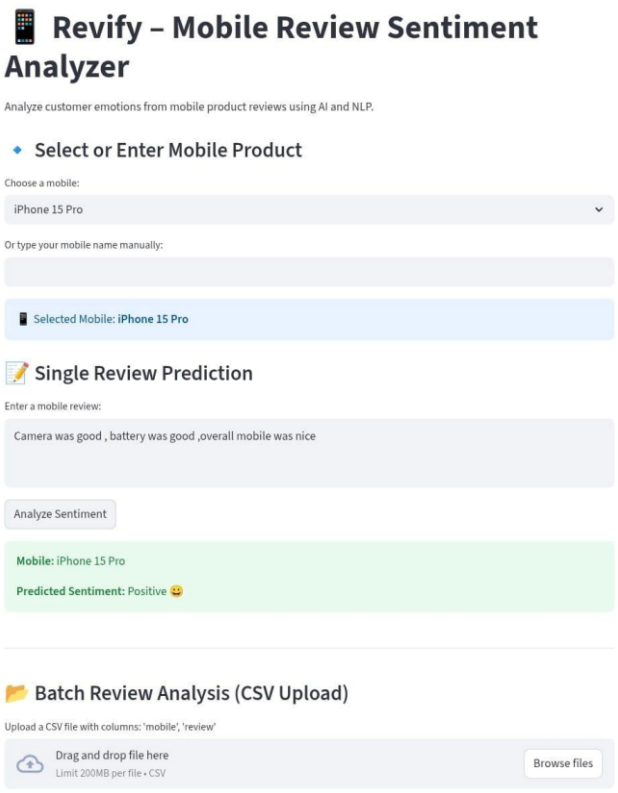
The research design follows a quantitative and experimental approach to analyse product reviews using NLP and machine learning. The study begins with data collection, followed by preprocessing to clean and prepare text. Feature extraction techniques are applied, and multiple algorithms are trained and evaluated. Comparative analysis identifies the most accurate model for sentiment classification, ensuring a systematic and replicable methodology.

V. RESULTS AND DISCUSSION

The sentiment analysis model successfully classified product reviews into positive, negative, and neutral categories with high accuracy. Among the algorithms tested, Logistic Regression and SVM performed better on TF-IDF features, while LSTM achieved the highest accuracy using word embeddings. Most reviews were identified as positive, indicating overall customer satisfaction. Misclassifications mainly occurred in reviews containing sarcasm or mixed opinions. The confusion matrix showed clear separation between positive and negative classes.

Overall, the results demonstrate that combining effective text preprocessing with optimized ML algorithms significantly improves sentiment prediction and provides valuable insights for product improvement and customer experience analysis.

Screenshots Predicted Customer Sentiment Analysis are shown below to illustrate system functionality.



Screenshot 5.1. Predicted Sentiment (Positive)

ISSN (Online): 2583-648X

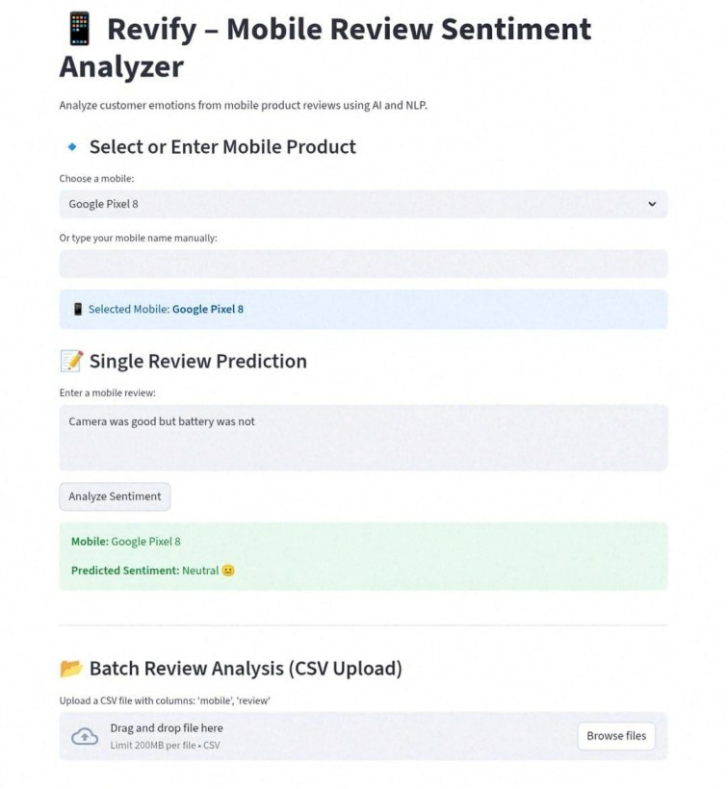
IJIRID

International Journal of Ingenious Research, Invention and Development

An International, High Impact Factor, Double-Blind Peer-Reviewed, Open-Access, Multidisciplinary Online Journal

Volume 4 | Issue 5 | October 2025

www.ijirid.in



Screenshot 5.2. Predicted Sentiment (Neutral)

VI. CONCLUSION

In Conclusion, the proposed Sentiment Analysis for Mobile Brand Product Reviews system successfully integrates NLP and ML techniques to perform efficient classification. The System achieved high accuracy, proving that Machine learning combined with linguistic processing can effectively interpret human opinions expressed in text form. The results of this research contribute to the growing field of opinion mining and customer Feedback Analytics, offering valuable insights for businesses and customers alike. With further enhancements and domain expansion, this system can evolve into a robust commercial-grade sentiment analysis tool applicable to multiple industries.

REFERENCES

[1] Zhang, L., Wang, S., & Liu, B. (2021). “Deep learning for Sentiment Analysis”

[2] Sharma, N. A., Ali, A. B. M. S., & Kabir, M. A. (2024). A Review of Sentiment Analysis: Tasks, Applications, and “Deep Learning Techniques”

[3] Taneja, A., & Singh, R. (2023). “Sentiment Analysis Using Machine Learning: A Comprehensive Review.

[4] Kumar, A., & Singh, M. (2023). Challenges and Future in “Deep Learning for Sentiment Analysis”

[5] Shinde, P., & Kale, S. (2021) “Sentiment Analysis of E-commerce Product Reviews Using NLP and Machine Learning Techniques”.

[6] Basiri, M. E., Nemati, S., Abd+ar, M., Cambria, E., & Acharya, U. R. ABCDM: An Attention-Based Bidirectional CNN-RNN Deep Model for “Sentiment Analysis. Future Generation Computer Systems”, 2021.

