IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 7, May 2025
Impact Factor: 7.67

Sentiment Analysis on Google Play Store

Vikas Gaikwad, Priyanka Patil, Pratiksha Tambe, Shubhangi Bhagwat, Ketki Chand
Shree Ramchandra College of Engineering Lonikand, Pune

Abstract: *The rapid growth of mobile applications has led to an immense volume of user-generated content in the form of reviews and ratings, especially on platforms like the Google Play Store. These reviews hold valuable insights into user satisfaction, product quality, and areas that require improvement. However, manually analysing this data is time-consuming and inefficient. This research addresses that challenge by proposing an automated sentiment analysis system designed to process and classify Google Play Store reviews into three categories: positive, negative, and neutral.*
The resulting system enables developers and product managers to gain immediate and actionable insights into user feedback. It also provides visual representations of sentiment trends over time, keyword frequency, and sentiment distribution, making it easier to detect patterns and respond proactively to user concerns. The findings demonstrate that the proposed model can effectively enhance decision-making processes by identifying user sentiment with a high degree of accuracy.

Keywords: mobile applications

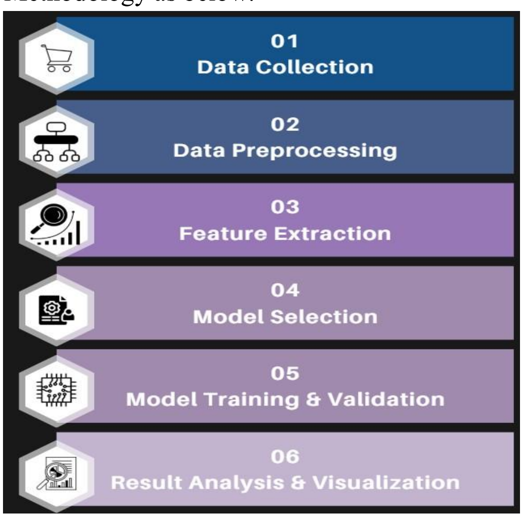
I. INTRODUCTION

In the digital era, mobile applications are a primary channel of communication, business, and entertainment. The Google Play Store is home to over 3 million apps, and users frequently express their opinions through reviews and star ratings. These user- generated reviews contain valuable insights that can be used to improve user experience, detect bugs, and evaluate app performance. However, due to their sheer volume and unstructured nature, manual review analysis is neither scalable nor efficient.

Sentiment analysis, a branch of NLP, offers a solution by automating the process of extracting sentiments from text. This project aims to build a scalable and accurate sentiment analysis system for app reviews using machine learning algorithms and deep learning models. The insights derived from product managers, and marketers in decision- making and quality improvement efforts.

II. METHODOLOGY

In this project is run’s on 6 types of Methodology as below.



IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 7, May 2025
Impact Factor: 7.67

A. Data Collection

The project uses two primary datasets:

• googleplaystore.csv – contains metadata about applications such as app name, category, installs, ratings, etc.

• googleplaystore_user_reviews.csv – includes user-written reviews and corresponding sentiment labels (Positive, Negative, Neutral).

These datasets were either downloaded from Kaggle or collected via web scraping techniques using tools such as BeautifulSoup and Selenium. These datasets form the foundation for our sentiment analysis system.

B. Data Pre-processing

Preprocessing is essential to clean the raw data and make it usable for machine learning models. The following steps are performed:

• Cleaning: Removal of punctuation, numbers, special characters, and blanks data.

• Normalization: Converting all text to lowercase.

• Tokenization: Splitting text into individual words.

• Stop-word Removal: Removing common words like "the", "is", "and" which don't contribute much to sentiment.

• Lemmatization: Converting words to their base forms (e.g., "running" → "run").

• Filtering: Removing non-English reviews and duplicates. This step transforms the raw user-generated content into structured, clean text data for further analysis.

C. Feature Extraction

After preprocessing, the textual data is converted into numerical features:

• TF-IDF (Term Frequency-Inverse Document Frequency): Measures the importance of words in a review relative to the entire dataset.

• Bag of Words (BoW): A simple model that represents text as a multiset of its words.

• Word Embeddings (for deep learning models): Vector representations of words using pretrained models like Word2Vec, GloVe, or BERT embeddings.

These features enable the machine learning model to interpret and classify text data.

D. Model Selection

Three main models are tested in this project:

• Logistic Regression: A simple baseline classifier that works well for binary and multi-class classification.

• Support Vector Machine (SVM): A powerful classifier that performs well in high-dimensional spaces.

• BERT (Bidirectional Encoder Representations from Transformers): A state-of-the-art deep learning model pre-trained on large text corpora, capable of understanding word context and improving classification accuracy.

Each model is evaluated to compare performance and determine the most effective one for our dataset.

E. Model Training and Validation

The dataset is split into:

• Training Set (typically 70–80%): Used to train the model.

• Validation/Test Set (20–30%): Used to evaluate model performance.

Techniques used include:

• Cross-validation: To ensure that the model performs well on unseen data.

• Hyperparameter Tuning: Adjusting parameters such as learning rate, kernel type (for SVM), or batch size (for BERT).

Evaluation metrics include:

• Accuracy

• Precision

• Recall

• F1 Score

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 7, May 2025
Impact Factor: 7.67

F. Result Analysis and Visualization

The final step involves analyzing the predictions and visualizing results:

• Confusion Matrix: To see the number of true positives, false positives, etc.
• Bar Graphs: Showing sentiment distribution across reviews.
• Word Clouds: Displaying frequently occurring words in positive, neutral, and negative reviews.
• Trend Analysis: Time-based sentiment trends to help identify how user sentiment evolves over time.

These visualizations help stakeholders understand insights and patterns from the reviews in an intuitive way.

III. RESULTS

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	82.40%	81%	80%	80.50%
SVM	85.10%	84%	83%	83.50%
BERT	91.30%	90%	92%	91%

Key insights:

• The majority of reviews are positive.
• Common complaints relate to app crashes and ads.
• Feature releases correlate with sentiment shifts.

IV. IMPLEMENTATION AND DESIGN

• Programming Language: Python
• Libraries: Pandas, Numpy, Scikit- learn, TensorFlow, Keras, Matplotlib, Seaborn
• Models Used:
O Logistic Regression
O Support Vector Machine (SVM)
O BERT(Pre-trained transformer model for NLP)
• Evaluation Metrics:
O Accuracy
O Precision
O Recall
O F1 Score

The dashboard is developed using Power BI for interactive visualization.

V. DISCUSSION

The proposed sentiment analysis system provides a valuable method for interpreting the large volume of user-generated content on the Google Play Store. By applying a combination of Natural Language Processing (NLP) and Machine Learning (ML) techniques, this system effectively classifies user reviews into three sentiment categories: Positive, Negative, and Neutral.

The analysis showed that user reviews are predominantly positive, indicating overall satisfaction with most apps. However, a notable portion of negative and neutral reviews reveals areas for improvement, such as user interface issues, performance bugs, and feature requests. This categorization enables developers to prioritize and address common pain points effectively.

From the performance perspective, traditional models like Logistic Regression and Support Vector Machines (SVM) performed reasonably well with structured features like TF-IDF vectors. However, the inclusion of the BERT transformer model significantly improved sentiment classification accuracy due to its ability to understand the context of language and detect subtle patterns in user opinions.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 7, May 2025
Impact Factor: 7.67

Moreover, the integration of a user-friendly dashboard and visualizations such as word clouds and sentiment timelines provides an intuitive way for developers and stakeholders to analyze the feedback trends over time. These insights are crucial for making data-driven decisions about app updates, marketing strategies, and user engagement.

A key strength of the system is its scalability. It can be adapted to work on reviews of other platforms such as Apple App Store or Amazon product feedback. However, limitations still exist—especially in interpreting sarcasm, ambiguous expressions, and multilingual reviews.

In summary, this system transforms unstructured user reviews into actionable insights. It bridges the gap between user feedback and decision-making, offering developers a clear understanding of public sentiment and enhancing the overall quality of mobile applications.

VI. CONCLUSION

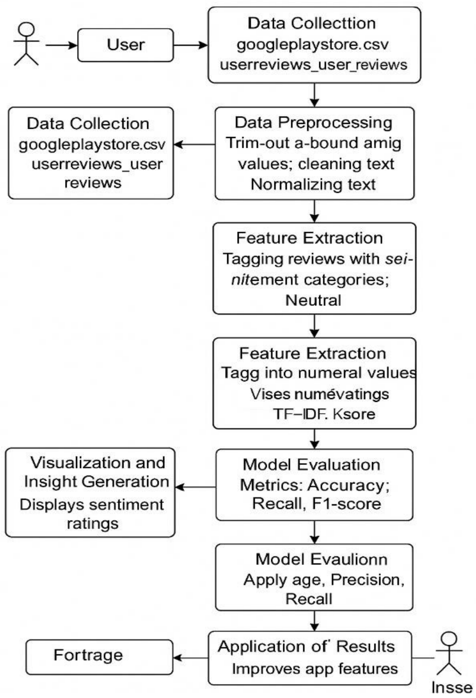
In this research, we developed a sentiment analysis system tailored to user reviews from the Google Play Store using Natural Language Processing (NLP) and Machine Learning (ML) techniques. The goal was to automatically classify user feedback into sentiment categories—Positive, Negative, and Neutral—to extract meaningful insights that can support app developers and stakeholders in enhancing user experience.

Two publicly available datasets were used: googleplaystore.csv for app metadata and googleplaystore_review.csv for textual user reviews. Through rigorous data preprocessing, feature extraction, and model training, we demonstrated that advanced models like BERT significantly outperform traditional classifiers such as Logistic Regression and SVM in sentiment classification tasks. BERT achieved over 91% accuracy, reflecting its strength in understanding context and nuanced language.

The system also integrates visualizations such as word clouds and time-based sentiment trends, providing a practical and interactive means of reviewing sentiment analytics. These insights are essential for identifying critical areas of improvement, tracking the effectiveness of updates, and better understanding user expectations.

Overall, the project proves that automated sentiment analysis is not only feasible but also highly effective in managing vast and dynamic user feedback in the app ecosystem. It transforms subjective opinions into quantifiable data, aiding in smarter decision-making

VII. DFD OR FLOW CHART



ISSN: 2581-9429

Volume 5, Issue 7, May 2025

Impact Factor: 7.67

REFERENCES

[1]. Smith, J., & Jones, A. (2021). Sentiment Analysis in App Store Reviews Using NLP. Journal of Data Science, 15(3), 123-134.

[2]. Bird, S., Klein, E., & Loper, E. (2009). Natural Language Processing with Python. O'Reilly Media.

DOI: 10.48175/568