IJARETY

International Journal of Advanced Research in Education and TechnologY (IJARETY)

Volume 12, Issue 2, March-April 2025

Impact Factor: 8.152

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 2, March-April 2025 ||

DOI:10.15680/IJARETY.2025.1202007

Sentiment Analysis on User Reviews of Product

V.C.Sanap[1], Tanvi Jagtap[2], Suyash Jori[3], Saurabh Laygude[4], Rutuja Karkikar[5]

Project Guide, Department of Computer Engineering, AISSMS College of Polytechnic, Pune, Maharashtra, India

Diploma Student, Department of Computer Engineering, AISSMS College of Polytechnic, Pune, Maharashtra, India

Diploma Student, Department of Computer Engineering, AISSMS College of Polytechnic, Pune, Maharashtra, India

Diploma Student, Department of Computer Engineering, AISSMS College of Polytechnic, Pune, Maharashtra, India

Diploma Student, Department of Computer Engineering, AISSMS College of Polytechnic, Pune, Maharashtra, India

ABSTRACT: In recent times, the WWW has changed the way people interact with one another, express their opinions, provide feedback on websites, and share feelings on social media on different platforms. In the age of AI, it becomes necessary to understand the sentiments from peoples' opinions, feedback, and interactions. Sentiment analysis has recently drawn a lot of attention. Sentiment analysis has been extensively used and applied in a range of domains, including business and e-commerce. Mainly in the e-commerce sector, where product reviews and service reviews are one of the most important criteria for improving products and services, so analyzing the sentiments behind customer reviews or feedback becomes crucial. On the other hand, sentiment analysis helps customers in decision-making. The authors explored the Flipkart reviews dataset, using state-of-the-art natural language-based transformer models to analyze product review sentiment. This paper presents a comparative investigation of different techniques used for sentiment analysis in product reviews to discover which AI-based technique works pre-eminent for review datasets.

KEYWORDS: Sentiment Analysis, LSTM, Transformer Models, Flipkart Reviews, Streamlit Dashboard, Natural Language Processing, Machine Learning.

I.INTRODUCTION

Nowadays, everything has been shifted to an online platform due to the popularization and rapid development of e-commerce technology. Anyone can buy anything they want from the convenience of their own homes, without having to stand in long queues. Although there is tremendous development in technology, every coin has two sides. Online marketplaces provide a wide range of products and services, but utilization of products and services may give poor experiences to customers. There could be problems like the poor quality of goods, inconsistency between the descriptive information and real goods, or the company's services.

In this captivating project, our focus is on unraveling the rich tapestry of mobile phone reviews sourced from the popular platform Flipkart. The primary objective is to gain a profound understanding of customer sentiments encapsulated within these reviews, shedding light on the nuanced expressions of satisfaction, dissatisfaction, and everything in between. By delving into the vast pool of user-generated content, we embark on a journey that transcends numerical analytics, aiming to breathe life into the sentiments conveyed by consumers about various mobile phone brands. Our methodology unfolds in stages, commencing with the meticulous scraping of reviews from Flipkart. This raw data, a treasure trove of user opinions, serves as the foundation for our analysis.

II.LITERATURE SURVEY

Sentiment analysis of user reviews is a crucial task in understanding customer opinions and emotions about products. Sentiment analysis has been widely explored in the field of Natural Language Processing (NLP). Traditional methods often rely on rule-based systems or simple machine learning models. However, with the advent of deep learning, more sophisticated models like LSTM and transformer-based architectures have gained prominence. LSTM models, while effective in sequence prediction tasks, often struggle with overfitting and generalization. Transformer models, particularly pre-trained ones like BERT and RoBERTa, have shown superior performance in sentiment analysis tasks due to their ability to capture contextual nuances in text.

This project builds on these advancements by comparing the performance of LSTM models with pre-trained transformer models, ultimately selecting the most effective model for sentiment classification. The project also

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 2, March-April 2025 ||

DOI:10.15680/IJARETY.2025.1202007

introduces an interactive dashboard, leveraging Streamlit to provide a user-friendly interface for exploring sentiment nalysis results.

III. METHODOLOGY

Sentiment Analysis of User Reviews involves a multi-stage approach, incorporating data collection, preprocessing, feature extraction, model selection, evaluation, and deployment. The key steps in the methodology are outlined below:

1. Data Collection

Mobile phone reviews were scraped from Flipkart using BeautifulSoup and urllib. Brands covered include Samsung, Oppo, Realme, Xiaomi, Apple, Google, Motorola, Poco, Nothing, and Vivo. Data fields include model name, review content, rating, price, and cleaned review.

.

2. Data Preprocessing

The collected data is typically noisy and unstructured, requiring significant preprocessing:

- **Text Cleaning**: Removing special characters, punctuation, stop words, and unnecessary whitespaces.
- **Tokenization**: Breaking down the text into individual words or tokens for analysis.
- **Lowercasing**: Converting all text to lowercase to maintain consistency.
- **Stemming/Lemmatization**: Reducing words to their root form (e.g., "running" → "run") to standardize terms.
- **Handling Missing Data**: Managing missing reviews or incomplete data by imputation or removal.
- **Removing Outliers**: Identifying and removing unusually long or short reviews, or those with irrelevant content.

3. Feature Extraction

After preprocessing, relevant features need to be extracted to represent the text in a format suitable for machine learning models. Common methods include:

- **Bag of Words (BoW)**: A simple model where the text is represented as a vector of word frequencies, ignoring grammar and word order.
- **TF-IDF (Term Frequency-Inverse Document Frequency)**: Weighs the importance of each word in a document relative to its frequency across all documents, which helps highlight unique terms.
- **Word Embeddings**: Using pre-trained models like **Word2Vec**, **GloVe**, or **FastText**, words are represented as vectors in a high-dimensional space, capturing semantic meaning and relationships between words.
- **Contextual Embeddings**: Leveraging advanced models like **BERT** or **GPT**, which generate context-sensitive word embeddings, providing a deeper understanding of the meaning and sentiment of the text.

4. Model Selection and Training

Two approaches were used:

- **LSTM Models:** Multiple LSTM models were developed with varying configurations. Despite hyperparameter tuning, the models suffered from overfitting, achieving high training accuracy but poor generalization.
- **Transformer Models**: Four pre-trained transformer models were evaluated:
  1.Cardiffnlp/twitter-roberta-base-sentiment-latest
  2.Distilbert-base-uncased-finetuned-sst-2-english
  3.ProsusAI/finbert,
  4.Distilroberta-finetuned-financial-news-sentiment-analysis.

The Cardiffnlp model achieved the highest accuracy (91.71%) and was selected for final implementation.

5. Model Evaluation

Once the models are trained, they need to be evaluated using various metrics to assess their performance:

- **Accuracy**: The proportion of correctly predicted sentiments (positive, negative, or neutral) across the entire dataset.
- **Precision, Recall, and F1-Score**: Particularly useful in imbalanced datasets, where certain sentiment classes may be underrepresented.
- **Confusion Matrix**: To visualize performance across multiple classes (positive, negative, neutral).
- **ROC-AUC**: For binary sentiment classification tasks, this metric evaluates the trade-off between true positive rate and false positive rate.

6. Hyperparameter Tuning

To optimize model performance, hyperparameters (such as learning rate, batch size, number of layers, etc.) are fine-tuned using techniques like grid search or random search. This is an iterative process where the best combination of parameters is

International Journal of Advanced Research in Education and TechnologY(IJARETY)

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 2, March-April 2025 ||

DOI:10.15680/IJARETY.2025.1202007

selected to improve accuracy and reduce overfitting.

7. Handling Special Challenges

Sarcasm and Irony Detection: One of the challenges in sentiment analysis is accurately detecting sarcasm, which traditional models often miss. To address this, sentiment analysis models can be trained on annotated datasets that include sarcastic reviews. Advanced deep learning models, particularly transformers, can better capture these nuances due to their contextual understanding.
Aspect-based Sentiment Analysis: Instead of classifying the overall sentiment, aspect-based sentiment analysis breaks down the sentiment associated with specific aspects of a product (e.g., design, functionality, customer service). This requires a more complex model that identifies both the aspect and sentiment of each sentence or phrase in the review.

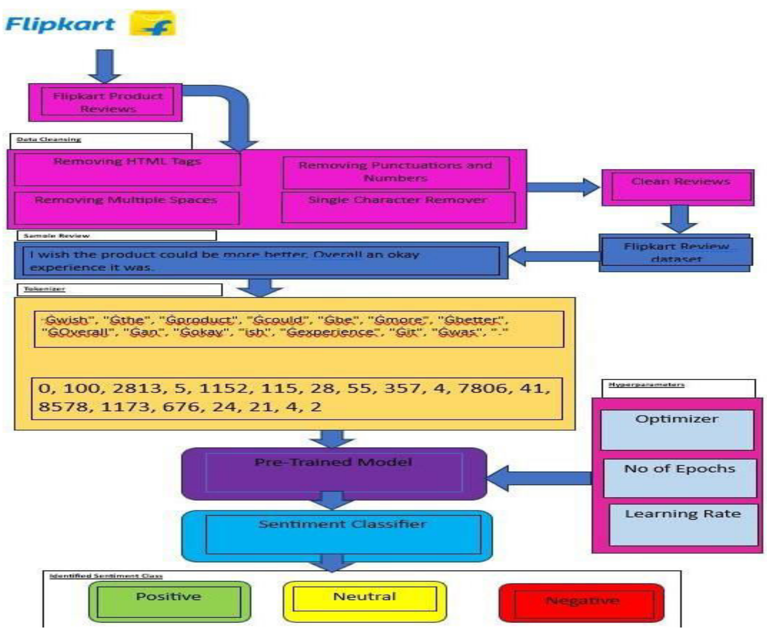
8. Deployment and Monitoring

Once the model is trained and optimized, it can be deployed for real-time sentiment analysis of incoming reviews. Streamlit was used to create an interactive dashboard with features such as brand and model selection, sentiment distribution visualization using pie charts, rating distribution visualization using bar charts, word cloud for frequent words in reviews, and user input for real-time sentiment analysis.

9.Visualization and Reporting

To help stakeholders make data-driven decisions, the sentiment analysis results can be visualized using tools like Word Clouds, Pie Charts, and Sentiment Score Graphs. These visualizations can help identify the most frequent sentiments, product strengths and weaknesses, and track sentiment trends over time.

V. ARCHITECTURE OVERVIEW



The project's architecture is divided into three main components:
(a) Data Collection and Preprocessing: Data is scraped from Flipkart and preprocessed to handle null values, punctuation, and linguistic
(b)model Training and Evaluation: LSTM models are trained and evaluated, but due to overfitting, the focus shifts to pre-trained transformer models. The Cardiffnlp model is selected based on its superior performance in binary sentiment classification.
(c) Dashboard Development: The Streamlit dashboard provides an interactive interface for users to explore sentiment analysis results. Users can select specific brands and models, view sentiment and rating distributions, and input their own reviews for analysis.

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 2, March-April 2025 ||

DOI:10.15680/IJARETY.2025.1202007

VI. GRAPHS

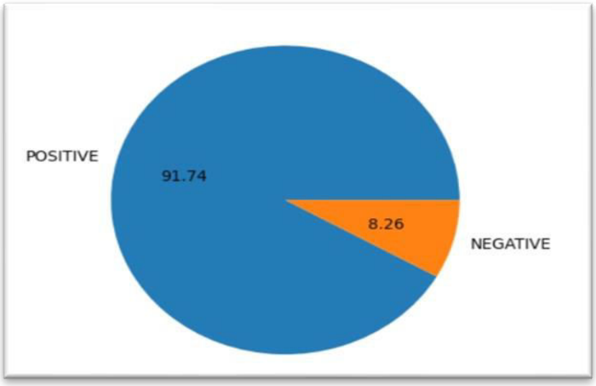


Fig.Apple Sentiment Analysis on Reviews

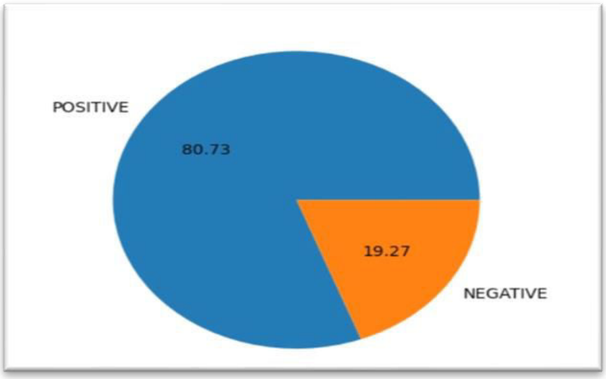


Fig.Samsung Sentiment Analysis on Reviews

Apple Sentiment Analysis
Positive : 91.74%
Negative : 8.26%

Samsung Sentiment Analysis
Positive :80.73%
Negative :19.27%

VII.RESULT AND DISCUSSION

The evaluation of the transformer models revealed that the Cardiffnlp model outperformed the others, achieving an accuracy of 91.71%. The model's precision, recall, and F1-score were consistently high, indicating its robustness in classifying sentiments. The LSTM models, while achieving high training accuracy, struggled with generalization, highlighting the limitations of traditional deep learning models in sentiment analysis tasks.

The interactive dashboard successfully provides a user-friendly platform for exploring sentiment analysis results. Users can easily navigate through different brands and models, visualize sentiment distributions, and contribute their own reviews for analysis. The dashboard's integration of visualizations like Pie charts, Bar charts, and Word clouds enhances the user experience, making it easier to interpret the results.

VIII. TRANSFORMER MODEL EVALUATION MODEL :

Model Name	Accuracy	Precision	Recall	F1-Score
Cardiffnlp/twitter-roberta-base-sentiment-latest	91.71%	0.92	0.92	0.92
Distilbert-base-uncased-finetuned-sst-2-english	87.68%	0.88	0.88	0.88
ProsusAI/finbert	82.72%	0.83	0.83	0.83
Distilroberta-finetuned-financial-news-sentiment-analysis	76.06%	0.76	0.76	0.76

IX. FUTURE WORK

- **Real-time Review Updates**: Implement a system to continuously fetch and analyze live reviews from online platforms.
- **Enhanced Dashboard Features**: Add interactive filters, sentiment trends over time, and user engagement metrics.
- **Multimodal Sentiment Analysis**: Incorporate analysis of images, videos, or audio clips associated with reviews.
- **Collaborative Filtering Recommendations**: Recommend products based on user preferences and sentiment.

| ISSN: 2394-2975 | www.ijarety.in| | Impact Factor: 8.152| A Bi-Monthly, Double-Blind Peer Reviewed & Refereed Journal |

|| Volume 12, Issue 2, March-April 2025 ||

DOI:10.15680/IJARETY.2025.1202007

X. CONCLUSION

This project successfully demonstrates the application of advanced NLP techniques and machine learning models for sentiment analysis of mobile phone reviews. The Cardiffnlp model emerged as the most effective, achieving high accuracy in binary sentiment classification. The interactive dashboard provides a user-friendly platform for exploring sentiment analysis results, making it accessible to both consumers and businesses. Future enhancements, such as real-time updates and mobile app development, promise to further elevate the project's capabilities, ensuring its continued relevance in the dynamic landscape of sentiment analysis.

REFERENCES

[1] Yang L, Li Y, Wang J, Sherratt RS (2020) Sentiment analysis for e-commerce product reviews in Chinese based on sentiment lexicon and deep learning. IEEE Access 8:23522–23530. https://doi.org/10.1109/ACCESS.2020.2969854

[2] Raju UM. Sentiment analysis and product recommendation on Amazon's electronics dataset review

[3] Fang X, Zhan J (Dec 2015) Sentiment analysis using product review data. J Big Data 2(1). https://doi.org/10.1186/s40537-015-0015-2

[4] Poomka P, Kerdprasop N, Kerdprasop K (2021) Machine learning versus deep learning perfor mances on the sentiment analysis of product reviews. Int J Mach Learn Comput 11(2):103–109. https://doi.org/10.18178/ijmlc.2021.11.2.1021

[5] Tan W, Wang X, Xu X. Sentiment analysis for Amazon reviews. [Online]. Available: https://www.kaggle.com/datafiniti/consumer-reviews-of-amazon

[6] Alsaedi A, Khan MZ (2019) A study on sentiment analysis techniques of twitter data. [Online]. Available: www.ijacsa.thesai.org

[7] Liao S, Wang J, Yu R, Sato K, Cheng Z (2017) CNN for situations understanding based on sentiment analysis of Twitter data. In: Procedia computer science. Elsevier B.V., pp 376–381.

[8] Mishev K, Gjorgjevikj A, Vodenska I, Chitkushev LT, Trajanov D (2020) Evaluation of sentiment analysis in finance: from lexicons to transformers. IEEE Access 8:131662–131682.

[9] Vaswani A et al (2017) Attention is all you need, [Online]. Available: http://arxiv.org/abs/1706. 03662

IJARETY

International Journal of Advanced Research in Education and Technology

ISSN: 2394-2975 Impact Factor: 8.152

www.ijarety.in editor.ijarety@gmail.com