



FLIPKART REVIEWS SENTIMENT ANALYSIS USING MACHINE LEARNING

Nivethika C¹, Dr. P. Deepika²

¹Student, Department of Artificial Intelligence and Machine Learning, Dr.N.G.P. Arts and Science College, Coimbatore, Tamil Nadu, India

²Associate Professor, Department of Artificial Intelligence and Machine Learning, Dr.N.G.P. Arts and Science College, Coimbatore, Tamil Nadu, India

ABSTRACT

With the rapid growth of e-commerce platforms, customers frequently share their opinions about products through online reviews. These reviews provide valuable insights into product quality, delivery services, and overall customer satisfaction. However, the large volume of user-generated content makes manual analysis difficult. Sentiment analysis, a Natural Language Processing (NLP) technique, helps automatically identify the emotional tone behind textual data. This paper presents a sentiment analysis system designed to classify Flipkart product reviews into positive, negative, and neutral categories using the DistilBERT model. DistilBERT is a lightweight transformer-based model derived from BERT that provides strong contextual language understanding with reduced computational cost. The proposed system includes stages such as data preprocessing, tokenization, model training, evaluation, and visualization. The model is evaluated using performance metrics such as accuracy, precision, recall, and F1-score. Experimental results demonstrate that the DistilBERT model effectively captures contextual relationships between words and produces reliable sentiment predictions. The developed system can assist businesses in analyzing customer feedback efficiently and support data-driven decision-making in e-commerce platforms.

KEYWORDS: Sentiment Analysis, Natural Language Processing, DistilBERT, Transformer Models, Flipkart Reviews, Machine Learning

1. INTRODUCTION

In recent years, e-commerce platforms have transformed the way consumers purchase products and share their experiences online. Websites such as Flipkart allow customers to provide feedback in the form of ratings and textual reviews. These reviews contain valuable information about product performance, delivery experience, customer satisfaction, and service quality. Businesses can use this feedback to improve product design, enhance service quality, and understand customer expectations.

However, the increasing number of online reviews has made manual analysis inefficient and time-consuming. Large e-commerce platforms generate thousands of reviews daily, making it difficult for businesses to analyze them individually. Therefore, automated sentiment analysis systems are required to process and interpret customer feedback efficiently.

Sentiment analysis is a subfield of Natural Language Processing that focuses on identifying the emotional tone expressed in textual content. It classifies text into categories such as positive, negative, or neutral sentiment. Early sentiment analysis methods relied on rule-based systems and traditional machine learning algorithms such as Naïve Bayes, Support Vector Machines, and Logistic Regression. These approaches use techniques such as Bag-of-Words and TF-IDF for feature representation. Although effective for simple tasks, these models often struggle to understand contextual meaning, sarcasm, and complex sentence structures.

Recent advancements in deep learning have significantly improved sentiment analysis performance. Transformer-based

language models such as BERT and DistilBERT have demonstrated strong capabilities in understanding contextual relationships between words in a sentence. DistilBERT is a compressed version of BERT that maintains most of its language understanding ability while reducing model size and computational cost.

This research focuses on developing a sentiment analysis system for Flipkart product reviews using the DistilBERT model. The objective is to classify reviews into positive, negative, and neutral sentiments automatically. The system includes several stages such as data preprocessing, tokenization, model training, evaluation, and visualization. By applying modern NLP techniques, the proposed system provides an efficient solution for analyzing large volumes of customer feedback.

2. LITERATURE SURVEY

Recent studies in sentiment analysis emphasize enhancing deep learning models' accuracy and efficiency for vast online text data from e-commerce, social media, and forums. Essential preprocessing like tokenization, stop-word removal, normalization, and cleaning structures data for better performance, with transformers like DistilBERT using subword tokenization to handle rare words effectively.

A major advance is contextual embeddings in transformer models, which capture word meanings based on surrounding context, outperforming static traditional methods. Fine-tuning pretrained models such as BERT and DistilBERT on domain-specific data (e.g., product reviews) boosts results without starting from scratch.



To address computational demands, techniques like knowledge distillation create lighter models. DistilBERT retains near-BERT performance with reduced size and speed. Visualization tools, including confusion matrices and word clouds, aid in evaluating predictions and spotting sentiment patterns.

Applied in business for review analysis, sentiment tools uncover issues and preferences, though challenges like imbalance, sarcasm, and mixed sentiments persist. This project leverages DistilBERT for scalable classification of Flipkart reviews, building on these transformer advancements.

3. METHODOLOGY

This project uses a straightforward ML pipeline to analyze Flipkart reviews: data collection, preprocessing, tokenization, training, evaluation, and visualization.

Data Collection

We gathered Flipkart product reviews from an online source, stored as CSV files. Each entry includes review text and a sentiment label (positive, negative, or neutral), processed with Python libraries.

Data Preprocessing

Raw reviews often have noise like special characters or inconsistencies, so we clean them up: drop missing values, lowercase everything, strip special chars and extra spaces. This ensures clean, consistent data for training.

Label Encoding

Models need numbers, not words, so we encode sentiments: negative=0, neutral=1, positive=2. This turns it into a multi-class problem.

Train-Test Split

We split data 80/20 (train/test) using stratified sampling to keep sentiment balance in both sets for fair evaluation.

Tokenization

DistilBERT can't handle raw text, so we use its tokenizer to break reviews into subword tokens, convert to IDs, and pad/truncate to uniform length for batch processing.

Model Training

DistilBERT shines here with its context awareness and efficiency. We fine-tune it on our dataset with set hyperparameters (learning rate, batch size, epochs), teaching it to link phrases to sentiments.

Evaluation Metrics

We track accuracy for overall performance, plus precision, recall, and F1-score per class. A confusion matrix visualizes how well it classifies each sentiment.

4. RESULTS AND DISCUSSION

After training the DistilBERT model, the system was evaluated using the testing dataset. The experimental results indicate that the model successfully classifies the majority of reviews into the correct sentiment categories. The performance metrics show balanced accuracy, precision, recall, and F1-score values across the three sentiment classes.

The confusion matrix analysis revealed that positive and negative reviews were classified with high accuracy. However, some neutral reviews were occasionally misclassified because neutral comments often contain mixed or balanced opinions. Despite this challenge, the overall performance remained stable and reliable.

To further analyze the dataset, visualization techniques such as word clouds were generated. Word clouds highlight frequently used words in customer reviews. Positive reviews commonly contained words related to satisfaction and quality, while negative reviews included words associated with dissatisfaction or product issues. These visual insights help better understand customer feedback patterns.

The system also includes an interactive interface that allows users to enter review text and receive instant sentiment predictions.

This real-time prediction capability demonstrates the practical application of the sentiment analysis system in monitoring customer feedback.

This chart provided the risk assessment for the given data. This showed that the pressure level is high for the person.

5. RESULTS AND DISCUSSION

The sentiment analysis system developed using DistilBERT successfully classifies product reviews into positive, negative, and neutral categories. The word cloud visualization highlights frequently occurring words such as *good*, *money*, *okay*, *wonderful*, and *waste*, which represent common customer opinions in the dataset. Positive reviews mainly include terms related to quality, satisfaction, and value for money, indicating a favorable customer experience. In contrast, negative reviews contain words such as *waste*, *poor*, and *disappointed*, reflecting dissatisfaction with product performance or quality. Neutral reviews include balanced expressions like *okay*, *decent*, and *specified*, which indicate average or mixed experiences. The system also provides confidence scores for each sentiment class, allowing users to understand how strongly the model predicts a particular sentiment. Sample results show predictions such as 49% positive, 58.5% negative, and 49% neutral confidence levels for different reviews. The interactive dashboard enables users to input product reviews and instantly obtain sentiment predictions. Overall, the system effectively analyzes customer feedback and can help businesses better understand user opinions and improve their products and services.



Fig No:1 WORD CLOUD



Fig No:2 Positive Sentiment Prediction



Fig No:3 Negative Sentiment Prediction

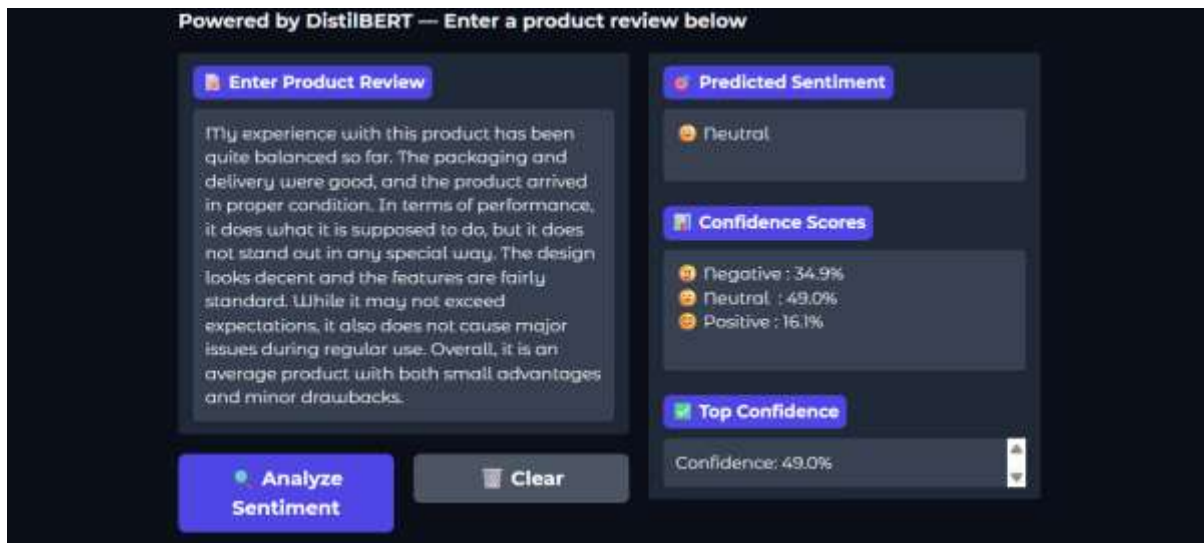


Fig No:4 Neutral Sentiment Prediction

6. CONCLUSION AND FUTURE WORK

This DistilBERT-powered system classifies Flipkart reviews into positive, negative, or neutral sentiments with high accuracy, thanks to smart preprocessing and strong contextual understanding that outshines traditional ML while staying computationally efficient. It equips businesses to automatically sift through customer feedback, uncovering preferences, spotting product flaws, and driving service improvements. Experimental results affirm its reliability for real-world e-commerce analysis. Looking ahead, future enhancements could include multilingual support for global reviews, integration with larger datasets for robustness, and aspect-based sentiment analysis to drill into specifics like "battery life" or "design quality."

REFERENCES

Bibliography

1. L. H. Fung and S. L. M. Belaidan, "Sentiment Analysis in Online Products Reviews Using Machine Learning," *iee*, vol. 18, no. SI05, pp. 914–928, Oct. 2021, doi: <https://doi.org/10.14704/web/v18si05/web18271>.
2. J. Soni, S. Xavier, R. Saxena, and P. Waghela, "Leveraging Web Scraping for Aspect-Based Sentiment Analysis: A Case Study on Flipkart Customer Reviews of Mobile Phones," *International Journal of Research Publication and Reviews*, vol. 5, no. 6, pp. 6747–6752, Jun. 2024, doi: <https://doi.org/10.55248/gengpi.5.0624.1638>.
3. H. Henderi and Q. Siddique, "Comparative Analysis of Sentiment Classification Techniques on Flipkart Product Reviews: A Study Using Logistic Regression, SVC, Random Forest, and Gradient Boosting," *Journal of Digital Market and Digital Currency*, vol. 1, no. 1, pp. 21–42, Jun. 2024, doi: <https://doi.org/10.47738/jdmdc.v1i1.4>.
4. G. Abhinav, Dr. Mahesh, and gc gc, "Sentiment Analysis Study of Human Thoughts using Machine Learning Techniques," *International Journal of Research Publication and Reviews*, vol. 6, no. 8, pp. 5998–6002, Aug. 2025, doi: <https://doi.org/10.55248/gengpi.6.0825.31139>.
5. G. Kaur and R. Dutta, "Sentiment Mining Based on Products Reviews Using Machine Learning," *Journal of Advances and*

Scholarly Researches in Allied Education, vol. 15, no. 5, pp. 185–191, Jul. 2018, doi: <https://doi.org/10.29070/15/57679>.

6. A. H. I. Ahmad Azrir, P. Naveen, and S. C. Haw, "Sentiment Analysis using Machine Learning Models on Shopee Reviews," *Journal of System and Management Sciences*, vol. 14, no. 2, 2024, doi: <https://doi.org/10.33168/JSMS.2024.0213>.
7. L. Ashbaugh and Y. Zhang, "A Comparative Study of Sentiment Analysis on Customer Reviews Using Machine Learning and Deep Learning," *Computers*, vol. 13, no. 12, pp. 340–340, Dec. 2024, doi: <https://doi.org/10.3390/computers13120340>.
8. A. Khan, D. Majumdar, and B. Mondal, "Sentiment analysis of emoji fused reviews using machine learning and Bert," *Scientific Reports*, vol. 15, no. 1, Mar. 2025, doi: <https://doi.org/10.1038/s41598-025-92286-0>.
9. M. M and S. Mehla, "Sentiment Analysis of Movie Reviews using Machine Learning Classifiers," *International Journal of Computer Applications*, vol. 182, no. 50, pp. 25–28, Apr. 2019, doi: <https://doi.org/10.5120/ijca2019918756>.
10. P. Poomka, N. Kerdprasop, and K. Kerdprasop, "Machine Learning Versus Deep Learning Performances on the Sentiment Analysis of Product Reviews," *International Journal of Machine Learning and Computing*, vol. 11, no. 2, pp. 103–109, Mar. 2021, doi: <https://doi.org/10.18178/ijmlc.2021.11.2.1021>.