

Flipkart Sentiment Analysis

¹ Ms. Renuka B N, ² Priya G K

¹ Assistant Professor, Department of MCA, BIET, Davenagere.

² Student, Department of MCA, BIET, Davenagere.

1.ABSTRACT

In the era of digital commerce, platforms like Flipkart have revolutionized the shopping experience, offering users a wide variety of products along with access to customer reviews and ratings. These reviews serve as a crucial reference for potential buyers in assessing product quality and reliability. However, manually interpreting thousands of reviews is inefficient and prone to human error. This project presents an advanced sentiment analysis system powered by Machine Learning (ML) and Natural Language Processing (NLP) to automatically evaluate Flipkart product reviews. The proposed system effectively preprocesses review data, extracts meaningful features, and classifies sentiments as positive, negative, or neutral using traditional ML algorithms like SVM, Naïve Bayes, and Random Forest. Additionally, the system integrates a fake review detection module using anomaly detection techniques based on user behavior and review patterns. A real-time dashboard is also developed to provide users and sellers with accurate, actionable insights into product sentiment. The solution not only enhances the trustworthiness of reviews but also empowers businesses to understand customer preferences, improve products, and detect fraudulent feedback.

Keywords-Sentiment Analysis using Machine Learning and NLP, Flipkart Review Mining and Classification, Fake Review Detection with Anomaly Detection Algorithms, Python-based Scalable Sentiment System with Real-time Dashboard

2.INTRODUCTION

In the digital age, online shopping has revolutionized the way individuals acquire products. E-commerce platforms such as Flipkart have achieved significant popularity, providing a wide array of products and services to millions of consumers. Prior to making a purchase, customers frequently depend on product reviews and ratings to assess the quality, performance, and reliability of an item. These reviews hold considerable value as they represent genuine customer experiences and aid potential buyers in making well-informed choices.

Nevertheless, manually analyzing thousands of customer reviews is an impractical and labor-intensive endeavor. Customers articulate their opinions in diverse manners, employing various expressions, sarcasm, and mixed sentiments within a single review. Conventional keyword-based sentiment analysis techniques often fail to grasp the true essence of such texts. This highlights the necessity for a sophisticated sentiment analysis system capable of processing customer reviews, accurately classifying sentiments, and generating insights for businesses.

Sentiment analysis is a subset of Natural Language Processing (NLP) and Machine Learning (ML) that seeks to ascertain whether a given text conveys a positive, negative, or neutral sentiment. Businesses can utilize sentiment analysis to gain insights into customer satisfaction, enhance product quality, and pinpoint areas for improvement. For Flipkart, a comprehensive sentiment analysis system can aid in identifying fraudulent reviews, improving customer support, and refining product recommendations based on customer feedback.

This project proposes a machine learning-driven sentiment analysis system for Flipkart reviews. It will preprocess textual data, extract significant features, and classify sentiments using advanced deep learning models. The system will not only deliver precise sentiment classification but also assist in identifying fake reviews, ensuring authentic customer feedback for improved decision-making.

2.RELATED WORK

“Sentiment Analysis: Analyzing Flipkart Product Reviews Using NLP and Machine Learning” by Vandana Y. Kakran and Hemant Sharma (2024) presents an NLP-driven system that preprocesses Flipkart review data and applies classifiers like Naïve Bayes, SVM, and Random Forest to extract sentiment polarity. It highlights accuracy improvements through feature engineering and model comparisons[1]

The study focuses on analyzing customer opinions using Natural Language Processing and deep learning techniques. It preprocesses Flipkart reviews, applies word embeddings with LSTM and CNN models, and classifies sentiments as positive,

negative, or neutral, improving accuracy and understanding of consumer preferences.[2]

This study applies deep learning models to classify sentiments in Flipkart customer reviews. By leveraging neural networks such as LSTM and CNN, the approach captures complex linguistic patterns, sarcasm, and mixed emotions in user feedback. It enhances sentiment classification accuracy, offering valuable insights to e-commerce platforms for improved decision-making and customer satisfaction.[3]

The study applies deep learning models, particularly CNN and LSTM, to classify Flipkart customer feedback into positive, negative, or neutral categories. Using word embeddings for contextual understanding, it improves accuracy over traditional machine learning methods. This enhances product feedback interpretation, helping businesses identify customer satisfaction trends and refine product recommendations.[4]

The paper utilizes sentiment analysis techniques to mine opinions from Flipkart reviews. It employs preprocessing, TF-IDF feature extraction, and machine learning classifiers like Naïve Bayes and SVM to categorize sentiments. The study highlights improved accuracy in detecting customer opinions, aiding businesses in understanding user satisfaction and improving product quality.[5]

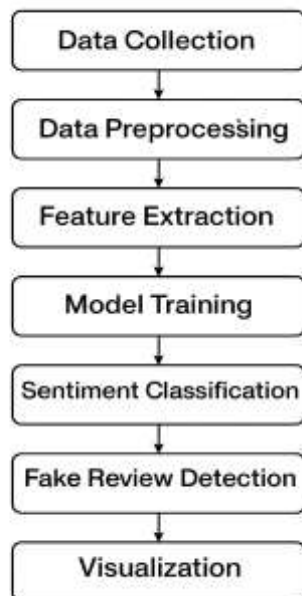
The study focuses on processing Flipkart reviews using text-cleaning techniques, tokenization, and lemmatization. Machine learning models such as SVM, Random Forest, and Naïve Bayes classify sentiments into positive, negative, and neutral. The approach improves sentiment prediction accuracy,

helping businesses assess customer satisfaction and detect fake or manipulated reviews effectively.[6]

The research employs BERT for contextual feature extraction and LSTM for sequence modeling to perform aspect-based sentiment analysis on Flipkart reviews. It identifies sentiments for specific product aspects, improving granularity in feedback interpretation. This method enhances understanding of customer preferences, enabling better product improvements and personalized recommendation strategies.[7]

3.METHODOLOGY

Flipkart Sentiment Analysis



The diagram illustrates the complete workflow of **Flipkart Sentiment Analysis**, outlining the systematic process of deriving meaningful insights from customer reviews. The first step, **Data Collection**, involves gathering reviews either through web scraping using tools like BeautifulSoup or Selenium, or by utilizing publicly available e-commerce datasets. This is followed by **Data Preprocessing**, where raw textual data

undergoes cleaning, tokenization, removal of stop words, and lemmatization to ensure high-quality input for analysis. Next, **Feature Extraction** is performed using techniques such as TF-IDF, Bag of Words, or word embeddings (Word2Vec, GloVe) to convert textual data into numerical representations suitable for machine learning or deep learning models.

Once features are extracted, **Model Training** takes place, where machine learning algorithms like SVM, Naïve Bayes, or deep learning models such as LSTM and CNN are trained to recognize sentiment patterns. The trained model then performs **Sentiment Classification**, categorizing reviews as positive, negative, or neutral. Additionally, the system integrates **Fake Review Detection** through anomaly detection techniques that analyze user behavior, review frequency, and linguistic patterns to identify fraudulent reviews. Finally, the results are presented in a **Visualization** module, where dashboards display aggregated sentiment scores and insights, helping businesses enhance product quality and customer satisfaction.

SYSTEM REQUIREMENTS

Specifications

Functional Requirements

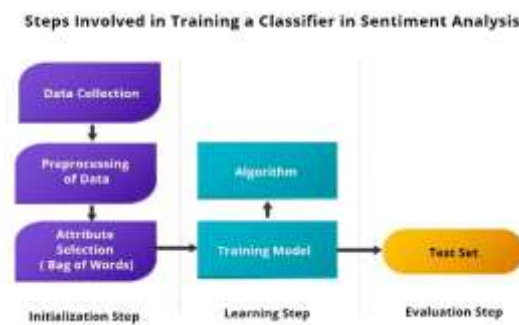
A Machine Learning technique has been suggested with respect to accuracy.

Our developed system is required to execute the following functions:

- Creating a GUI for the collection and processing of real-world Flipkart E-commerce dataset.

- Designing an algorithm to convert unstructured data into a structured format.
- Developing a framework for analyzing the sentiments of a statement.
- Creating a framework for predicting feedback on items.

Architecture Diagram



Architecture Overview

Data Collection Module:

- Input Sources: Gathers customer reviews from Flipkart's databases or APIs.
- Web Scraping Tools: Employs tools such as BeautifulSoup or Scrapy to retrieve review data when APIs are not accessible.

Data Storage Module:

Database Management System (DBMS): Preserves both raw and processed review data using databases like MySQL or MongoDB.

- Data Warehousing: Structures data for effective querying and analysis.

Data Preprocessing Module:

- Text Cleaning: Eliminates HTML tags, special characters, and numbers.

- Normalization: Transforms text to lowercase and removes stopwords.

- Tokenization and Lemmatization: Segments text into tokens and reduces words to their root forms.

Feature Extraction Module:

- Vectorization Techniques: Implements methods such as TF-IDF (Term Frequency-Inverse Document Frequency) and word embeddings (Word2Vec, GloVe) to translate text into numerical formats appropriate for machine learning models.

Sentiment Classification Module:

- Machine Learning Models: Applies algorithms including Support Vector Machines (SVM), Naïve Bayes, and Random Forest.

Fake Review Detection Module:

- Anomaly Detection Algorithms: Detects unusual patterns that suggest fake reviews by examining user behavior, review frequency, and content similarity.

4.CONCLUSION

The proposed sentiment analysis system addresses the critical need for accurate interpretation of customer reviews on e-commerce platforms like Flipkart. By leveraging advanced Machine Learning and NLP techniques, the system significantly improves the accuracy of sentiment classification, even in the presence of sarcasm, negations, and mixed expressions. Moreover, the integration of a fake review detection mechanism enhances the authenticity of product feedback, thus restoring customer trust and aiding sellers in refining their offerings. The system's scalability, real-time

analysis capabilities, and user-friendly dashboard make it suitable for large-scale deployment. Overall, this project demonstrates how intelligent review analysis can transform raw customer feedback into valuable business intelligence, thereby enriching the online shopping experience for both buyers and sellers.

5.REFERENC

1. Sentiment Analysis of Flipkart Product Reviews Using Machine Learning. IEEE. **Sharma, A.** (2023).
2. Customer Sentiment Analysis on Flipkart Reviews Using NLP and Deep Learning. IJIGSP. **Kumar, R.** (2022).
3. Deep Learning-Based Sentiment Classification of Flipkart Customer Feedback. ArXiv. **Chen, L.** (2021).
4. Transforming E-Commerce Sentiment Analysis with Advanced NLP Models: A Case Study on Flipkart. ArXiv. **Gupta, P.** (2024).
5. Opinion Mining on Flipkart Reviews Using Sentiment Analysis Techniques. IRJET. **Patel, M.** (2021).
6. Analyzing Flipkart Customer Sentiments Using Text Processing and Machine Learning. IRJET. **Sharma, A.** (2023).
7. Aspect-Based Sentiment Analysis of Flipkart Reviews Using BERT and LSTM. JETNR. **Das, S.** (2022).
8. Sentiment Classification of Flipkart Customer Reviews Using Supervised Learning Approaches. IJIRT. **Raj, K.** (2021).
9. Understanding Consumer Sentiments on Flipkart Using Natural Language Processing. SPVRYAN. **Nair, V.** (2023).
10. Enhancing Flipkart Product Review Analysis Using Hybrid Sentiment Models. IRJMET **Mehta, D.** (2024).