Customer Sentiment Analysis Report

Github

Introduction

Customer Sentiment Analysis is a technique used to understand customer opinions and emotions from textual data such as reviews. In this project, sentiment analysis is performed on customer reviews of Apple iPhone 15 (128 GB) collected from the Flipkart website. The main purpose of this analysis is to evaluate customer satisfaction and understand customer behavior based on their feedback.

Objective

The objectives of this project are:

- To collect real customer reviews from Flipkart.
- To clean and preprocess the review data.
- To perform sentiment analysis using Natural Language Processing (NLP).
- To classify customer reviews into **Positive** and **Negative** sentiments.
- To analyze customer behavior using ratings and review length.
- To extract simple and meaningful insights.

Data Collection (Web Scraping)

The data was collected using **Python-based web scraping** techniques.

Tools Used:

- Requests
- BeautifulSoup

Process:

- Reviews were scraped from multiple pages of the Flipkart product review section.
- Pagination was handled to collect reviews from several pages.
- A total of **~289 customer reviews** were successfully collected.

Extracted Fields:

- Reviewer Name
- Rating (1–5)
- Review Text

The collected data was saved in CSV format for further analysis.

Data Cleaning and Preprocessing

To ensure quality analysis, the dataset was cleaned and prepared using **Pandas**.

Steps Performed:

- Removed duplicate reviews
- Handled missing values
- Converted review text to lowercase
- Removed punctuation, special characters, and extra spaces
- Created a cleaned review column for analysis

Additional features were engineered:

- **Polarity Score** (from sentiment analysis)
- **Sentiment Label** (Positive / Negative)
- **Review Length** (number of characters in each review)

Sentiment Analysis

Sentiment analysis was performed using **TextBlob**.

Methodology:

- Polarity score ranges from **-1 (negative)** to **+1 (positive)**
- Reviews were classified as:
    - **Positive:** Polarity ≥ 0.1
    - **Negative:** Polarity < 0.1

Each review was assigned a sentiment label and polarity score.

Data Analysis and Insights

Based on the analysis, the following insights were observed:

- The majority of customer reviews are **positive**, indicating high satisfaction with the product.
- Higher ratings (4 and 5 stars) strongly align with positive sentiment polarity.
- Positive reviews are generally **short and appreciative**, focusing on features like camera quality, performance, and premium design.
- Negative reviews are **fewer in number** but are usually more detailed and descriptive.
- Customers tend to write **longer reviews when they are dissatisfied**, highlighting specific issues.

- Review length acts as an indicator of dissatisfaction, as negative reviews are often longer than positive ones.
- Some reviews show neutral or low polarity despite high ratings, showing that ratings alone may not fully capture customer sentiment.

These findings confirm that **text-based sentiment analysis provides deeper insights than ratings alone**.

Visualization Summary

The following key visual analyses were performed:

- **Sentiment Distribution:** Count of positive vs negative reviews
- **Average Rating vs Sentiment:** Positive sentiment corresponds to higher average ratings
- **Review Length Analysis:** Negative reviews have higher average length compared to positive reviews

These visualizations help clearly communicate customer behavior and sentiment patterns.

Insights

- Customers are highly satisfied with the iPhone 15 overall.
- Camera quality and performance are major positive drivers.
- Dissatisfaction mainly arises from expectations, pricing, or specific performance concerns.
- Long negative reviews provide valuable feedback for product improvement.

Recommendations

Based on the analysis:

- Focus on addressing issues mentioned in negative reviews to improve customer experience.
- Use long negative reviews as a feedback source for quality improvement.
- Combine ratings with sentiment analysis for a more accurate understanding of customer opinions.
- Businesses should monitor sentiment trends regularly for proactive decision-making.

Conclusion

This project successfully demonstrates how web scraping and sentiment analysis can be used to extract meaningful insights from customer reviews. The combination of ratings, textual sentiment, and review length provides a comprehensive understanding of customer feedback. Such analysis can be effectively used by businesses to enhance products, improve customer satisfaction, and make data-driven decisions.

"The complete project code and data are available on GitHub."