Task 2: Sentiment Analysis on Customer Reviews using Python

Submitted by: Manasi Subhash Khodade
Date: 18/07/2025
Internship Project Report

Objective:

The aim of this task was to perform sentiment analysis on customer reviews to classify them as **Positive**, **Negative**, or **Neutral**, and derive insights using Python.

Tools & Libraries Used:

- Python (Jupyter Notebook or PyCharm)
- pandas, matplotlib, seaborn
- TextBlob (for sentiment analysis)

Dataset Used:

- Amazon Product Reviews Dataset (Clothing)
- Key columns used: `Review Text`, `Rating`, `Recommended IND`

Steps Performed:

1. Data Preprocessing:

- Loaded the CSV dataset using pandas.
- Displayed the first few rows and column names for understanding the structure.
- Focused on the **Review Text** column for sentiment classification.

2. Sentiment Classification using TextBlob:

- Applied TextBlob to each review to calculate sentiment polarity.
- Defined logic to classify:
  - **Polarity > 0** → Positive
  - **Polarity = 0** → Neutral
  - **Polarity < 0** → Negative
- Stored sentiment label in a new column `Sentiment`

3. Visualization & Summary:

- Counted the number of reviews in each sentiment class.
- Created a **bar chart** showing the sentiment distribution.

📊 Figure 1: Bar Chart – Sentiment Distribution

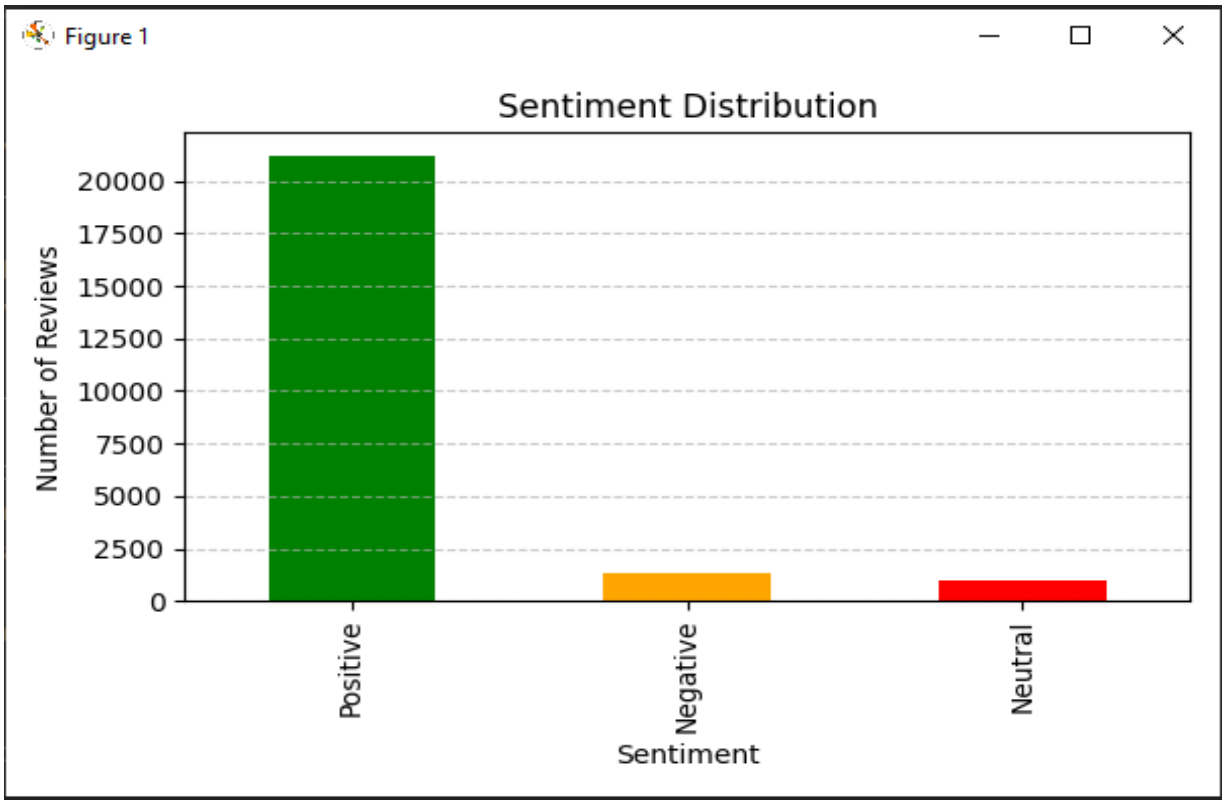

Figure 1: Distribution of Sentiments among Customer Reviews

🥧 Figure 2: Pie Chart – Sentiment Distribution

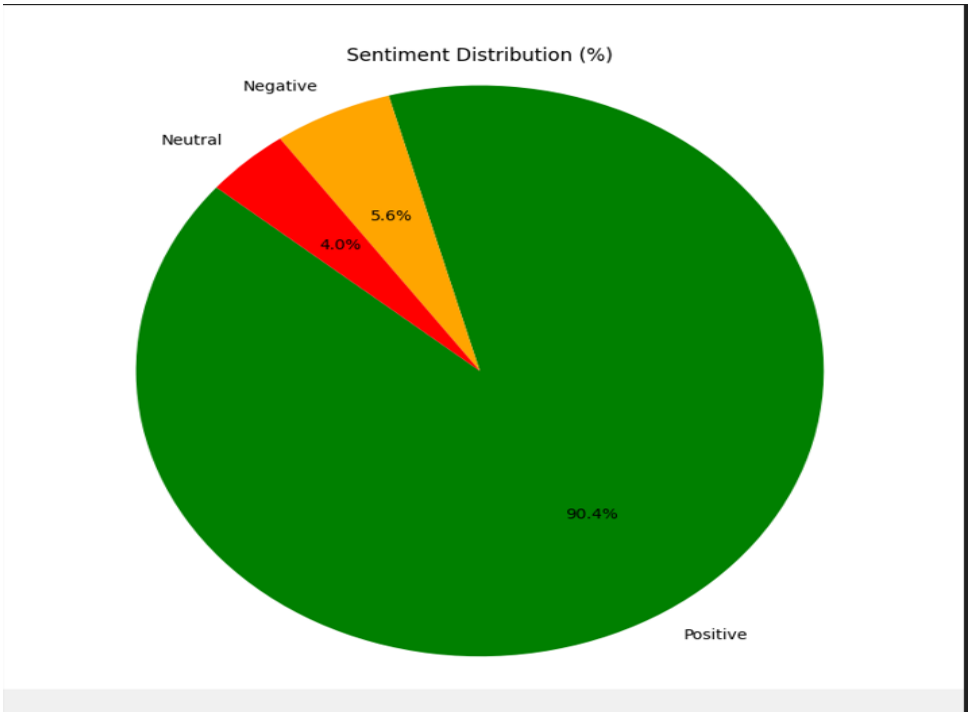

Figure 2: Pie Chart View of Sentiment Classification

- Positive reviews dominated with over **21,000 reviews**, followed by Neutral and Negative.

Sample Output:

```
Sentiment Counts:
Positive    21227
Negative     1315
Neutral       944
```

Sample classified reviews:

Review Text	Sentiment
Absolutely wonderful - silky and sexy...	Positive
I had such high hopes but...	Negative

Key Insights:

- Majority of customer reviews were **positive**, indicating general satisfaction.
- Very few **negative** reviews were observed, suggesting high product quality or customer experience.
- The process shows how **text data** can be structured and analyzed for business use.

◆ Insights:

- Most customer reviews are **Positive (~90%)**.
- Very few reviews are **Negative (~5.6%)** or **Neutral (~4%)**.
- This shows good customer satisfaction with the clothing products.

Conclusion:

This task helped demonstrate basic **Natural Language Processing (NLP)** using Python. It showed how to extract, clean, analyze and visualize customer sentiments — a vital skill for product feedback analysis and marketing teams.

Thank you for reviewing this report. Screenshots and charts from the code output are attached.