PROJECT REPORT: SOCIAL MEDIA CUSTOMER SENTIMENT ANALYSIS OF PAKISTANI FASHION BRANDS

Eship

KASHIF NASEER

MUHAMMAD BIN IMRAN

1. Introduction

This project aims to analyze customer sentiment toward two popular Pakistani fashion brands—**Sapphire** and **Khaadi**—based on user-generated content from **Instagram**. The data was extracted in **JSON format**, parsed, and processed to evaluate public opinion. This sentiment analysis helps identify what drives customer satisfaction or dissatisfaction and provides insights that can inform better brand strategies.

2. Objectives

- To collect real customer comments from Instagram (JSON format)

- To preprocess and clean the data for analysis

- To perform sentiment classification using **TextBlob**

- To visualize sentiment trends and extract key insights

- To store the results for potential future modeling or decision support

3. Dataset Sources

The datasets used were exported as **JSON files** from Instagram:

1. **Sapphire Instagram Comments**

   o   File: sapphire_instagram.json

2. **Khaadi Instagram Comments**

   o   File: khaadi_instagram.json

Each file included:

- username

- comment_text

- timestamp

- post_id

The data was parsed into structured format using Python.

4. Technologies Used

- **Python** (Pandas, JSON, Matplotlib, Seaborn, TextBlob)

- **NLTK** for natural language processing

- **Instaloader / Custom Scraper** for Instagram data

- **Jupyter Notebook / VS Code** as development platforms

5. Deliverables Summary

1. Data Collection & Preprocessing

- Raw JSON comments collected from brand Instagram pages

- Fields extracted: comment text, timestamps, post info

- Cleaned dataset saved as cleaned_comments.csv

2. Sentiment Analysis

- Sentiment scoring with **TextBlob polarity**

- Labels: **Positive**, **Neutral**, **Negative**

- Output saved to sentiment_comments_final.csv

3. Key Insights

- Positive comments focused on design aesthetics and fast service

- Negative comments highlighted late delivery and high prices

- Insights visualized using plots and word clouds

6. Sentiment Drivers Analysis

Using the labeled sentiment data, we identified key drivers:

Positive Sentiment Drivers

- Common terms: "great quality", "love the design", "reasonable", "amazing collection"

- Customers appreciated fabric, color combinations, and seasonal offerings

Negative Sentiment Drivers

- Repeated issues: "late delivery", "too expensive", "bad customer service", "not received"

- Complaints centered around order fulfillment and pricing

Campaign Relevance

- Positive spikes occurred during sales and new launches

- Negative sentiment increased during overhyped releases and shipping delays

7. Temporal Analysis: Sentiment Over Time

Based on Instagram comment timestamps:

- **Positive peaks**:

- o   Eid and seasonal launches

- o   Giveaway and collaboration campaigns

- **Negative dips**:

  - o   Pre-Eid shipping delays

  - o   Overpriced product launches

Interactive or user-participated campaigns received more engagement and positivity.

8. Visualizations

Generated visual outputs include:

- Pie chart of sentiment distribution

- Bar chart comparing sentiment counts per brand

- Histogram of sentiment polarity scores

- Word clouds of positive and negative comments

- Line chart for temporal sentiment trend

9. Conclusion

This project demonstrates how Instagram comments, when parsed from **JSON files**, can provide actionable insights through sentiment analysis. Both **Sapphire** and **Khaadi** benefit from understanding what customers love and what needs improvement. This method offers a scalable way for fashion brands to listen to their audience and tailor campaigns for better engagement.