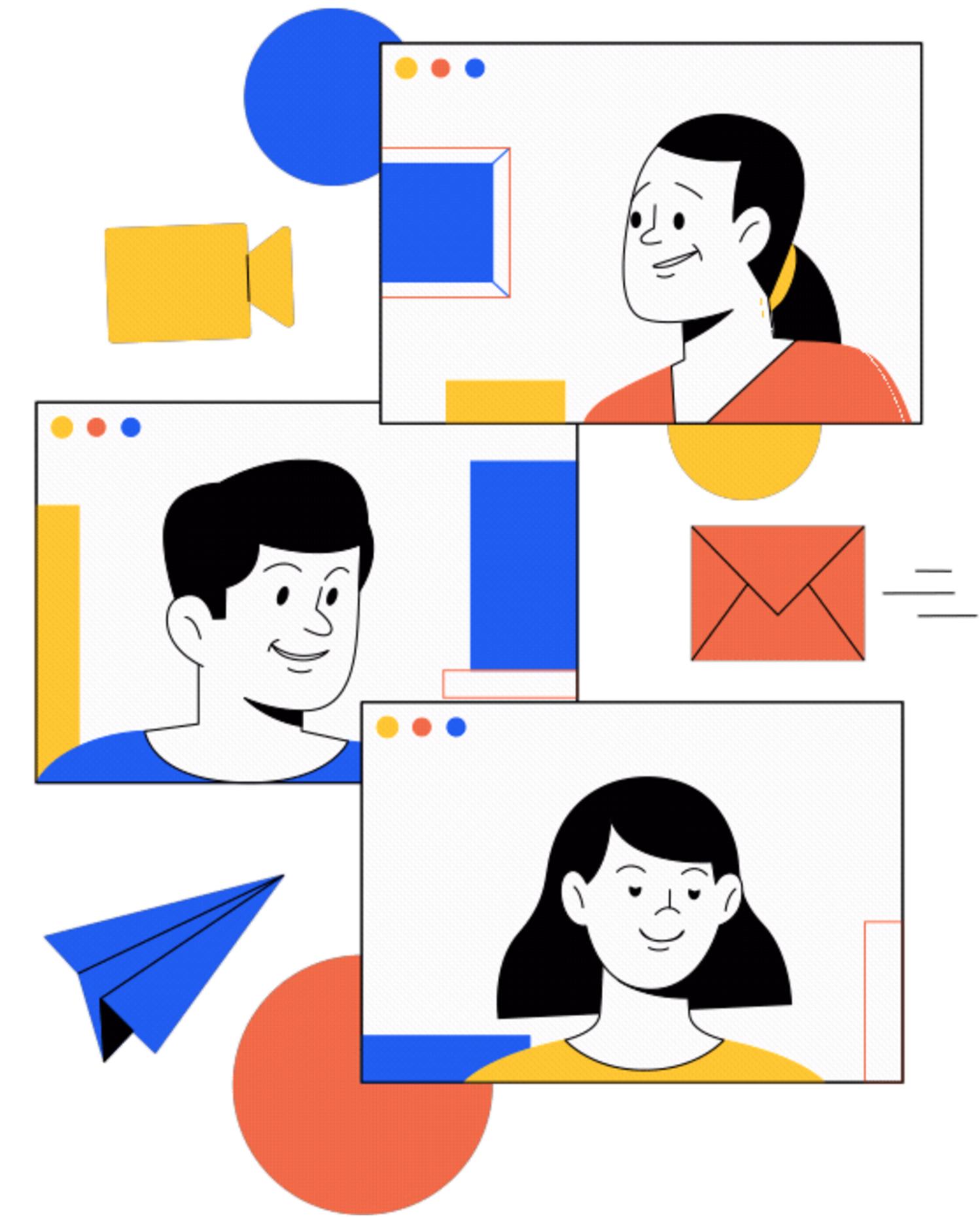


SOCIAL MEDIA SENTIMENT ANALYSIS (NLP)



Problem Statement

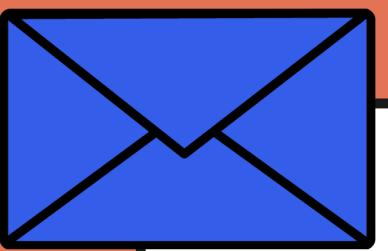
Objective:

Analyze social media data (tweets/reviews) using NLP to classify user sentiments as Positive, Negative, or Neutral.

Goal:

Understand public perception of a brand and identify sentiment trends to inform marketing and customer strategies.

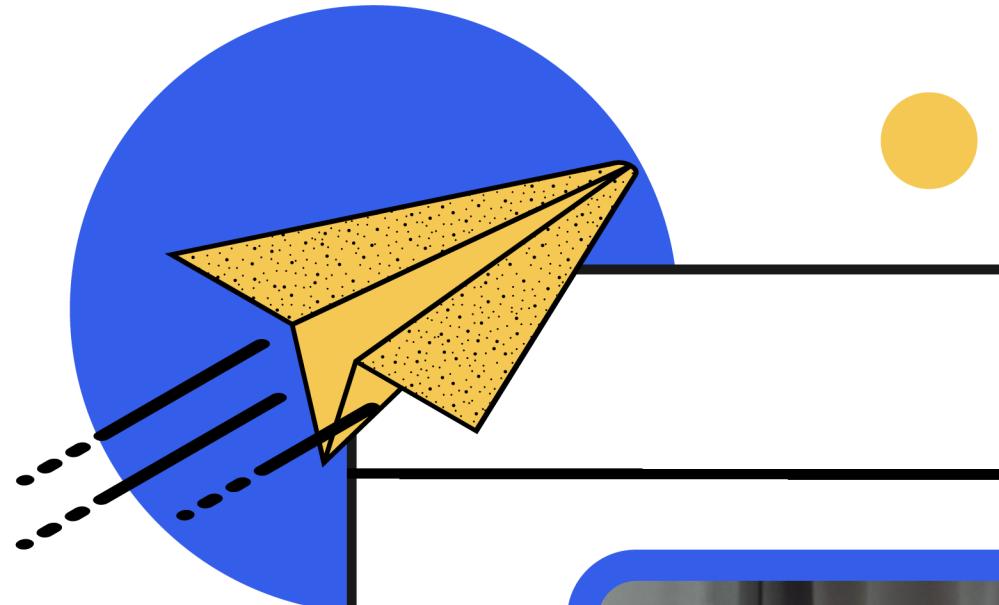




DATASET OVERVIEW

- 📁 **Source:** Kaggle - Social Media Sentiments Dataset
- 📊 **Size:** ~30,000+ tweets (can vary by chosen dataset)
- 📅 **Key Columns:**
 - **text:** Original tweet or review
 - **sentiment:** (original label)
 - **clean_text:** Preprocessed text (created)
 - **sentiment_vader:** Classified sentiment using VADER





TOOLS & LIBRARIES USED



 **Google Colab - Python Environment**
 **NLP & Sentiment:**
NLTK
TextBlob
VADER SentimentIntensityAnalyzer

 **Visualization:**
WordCloud
Matplotlib
Seaborn

TASK1 – DATA COLLECTION

- **✓ Loaded dataset from Kaggle into Colab**
- **✓ Verified structure and size**
- **✓ Checked for nulls, duplicates, and basic patterns**



TASK 2 - PREPROCESSING

Lowercased Text

Converted all text to lowercase for uniformity (e.g., "Happy" → "happy").

Removed Noise

Eliminated URLs, mentions (@user), hashtags, punctuation, and numbers using regex.

Tokenization & Stopword Removal

Split sentences into words and removed common stopwords like the, is, and.

Lemmatization

Reduced words to their root form (e.g., "running" → "run").

Final cleaned text saved in clean_text column

TASK 3 – SENTIMENT CLASSIFICATION

Used two models:

- TextBlob → based on polarity (-1 to +1)
- VADER → social media-optimized sentiment scoring

Classified each tweet as:

Positive

Neutral

Negative

Results stored in 'sentiment_vader' & 'sentiment_textblob'

TASK 4 - VISUALIZATIONS

AB
CD Word Cloud

Created using the WordCloud library.

Displays the most frequent and impactful words in the cleaned tweets.

Larger words indicate higher frequency.

Helps identify key themes and topics in user sentiment.

↗ Sentiment Over Time

Plotted sentiment values (Positive = 1, Neutral = 0, Negative = -1) over the index of tweets.

Used a line plot via Seaborn to visualize sentiment trends.

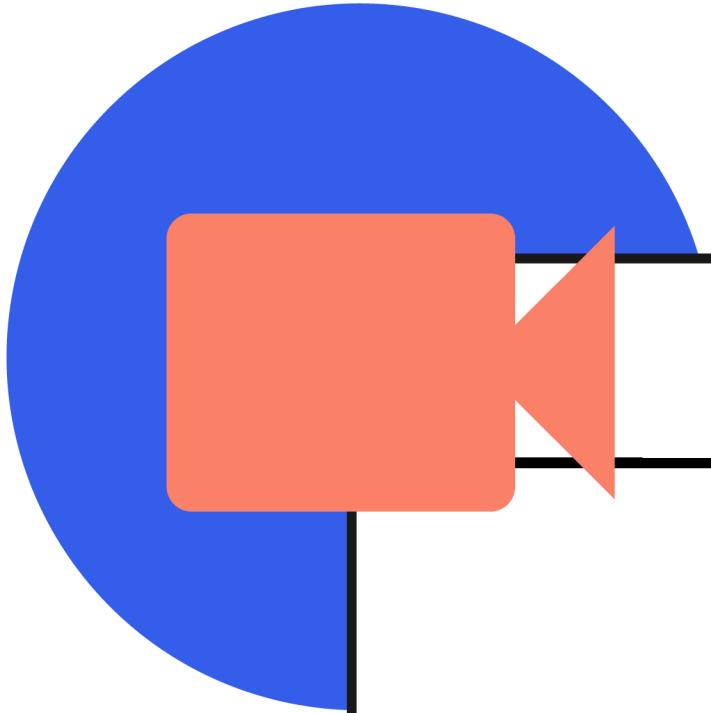
Although tweets had no timestamps, index values approximate posting order.

✓ Tools Used:

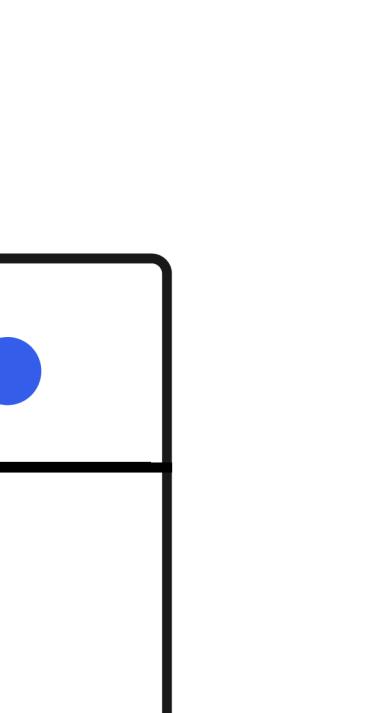
WordCloud

Seaborn for Line Plot

Matplotlib for plot rendering



TASK 5 – TREND ANALYSIS & SPIKES



 **Converted sentiments to scores:**

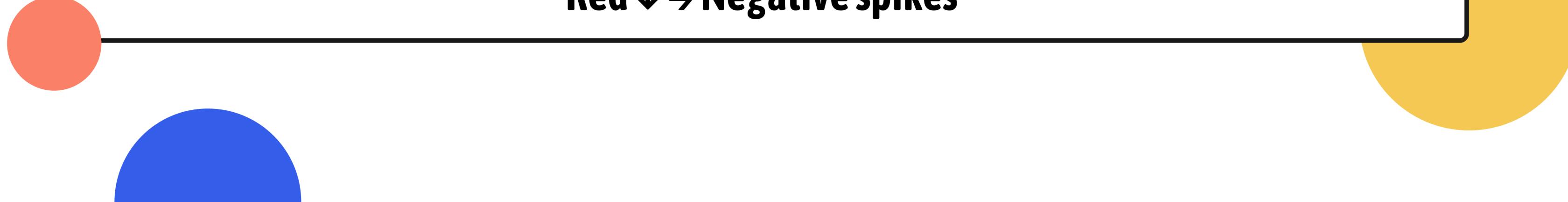
Positive = 1, Neutral = 0, Negative = -1

 **Used rolling average to smooth the trend**

 **Marked sentiment spikes:**

Green ↑ → Positive spikes

Red ↓ → Negative spikes





KEY INSIGHTS

**AB
CD MOST FREQUENT WORDS:**
FROM THE WORD CLOUD, COMMON WORDS LIKE “SERVICE,” “DELAY,” “AWESOME,”
“PRICE” STOOD OUT, REVEALING KEY THEMES IN USER DISCUSSIONS.

● DOMINANT SENTIMENT:
MAJORITY OF TWEETS WERE POSITIVE, INDICATING OVERALL FAVORABLE PUBLIC
OPINION.

● SPIKE POINTS:
SHARP SENTIMENT SHIFTS (POSITIVE OR NEGATIVE) HIGHLIGHT MOMENTS WHEN
USERS REACTED STRONGLY — OFTEN DUE TO EVENTS, ISSUES, OR
ANNOUNCEMENTS.

● TAKEAWAY:
TRACKING TOP WORDS AND SPIKES HELPS BRANDS UNDERSTAND USER CONCERNS,
IMPROVE COMMUNICATION, AND RESPOND EFFECTIVELY.





RECOMMENDATIONS

📢 RESPOND QUICKLY TO NEGATIVE TWEETS

Address customer complaints or negative feedback publicly and promptly. Fast, transparent responses help rebuild trust and show that the brand cares.

🙌 Leverage Positive User-Generated Content (UGC)

Share and highlight tweets or posts from happy customers. UGC builds social proof and encourages others to share positive experiences.

📈 Monitor Key Sentiment-Triggering Keywords

Keep a close eye on words like “delay,” “awesome,” “expensive”. These terms often drive emotional reactions and can guide product or service improvements.

👥 Use Influencer Collaborations Strategically

Partner with influencers to improve public perception, especially during or after negative sentiment spikes. Their reach can help flip the overall narrative.

🎯 Create Educational & Value-Driven Content

Sharing tips, guides, and behind-the-scenes content can build trust and brand authority. It also keeps engagement high during neutral sentiment phases.

CONCLUSIONS

- Successfully carried out Sentiment Analysis on social media data using Natural Language Processing (NLP) techniques.
- Effectively visualized sentiment trends across tweets and identified emotional spikes, revealing key moments of public reaction.
- Based on insights, provided actionable recommendations to help brands improve customer engagement, manage reputation, and enhance user satisfaction.





FUTURE SCOPE

⌚ Real-Time Tweet Streaming using Twitter API

Implement Twitter API to fetch live tweets, enabling real-time sentiment tracking and immediate reaction to public opinion.

🧠 Topic Modeling with LDA

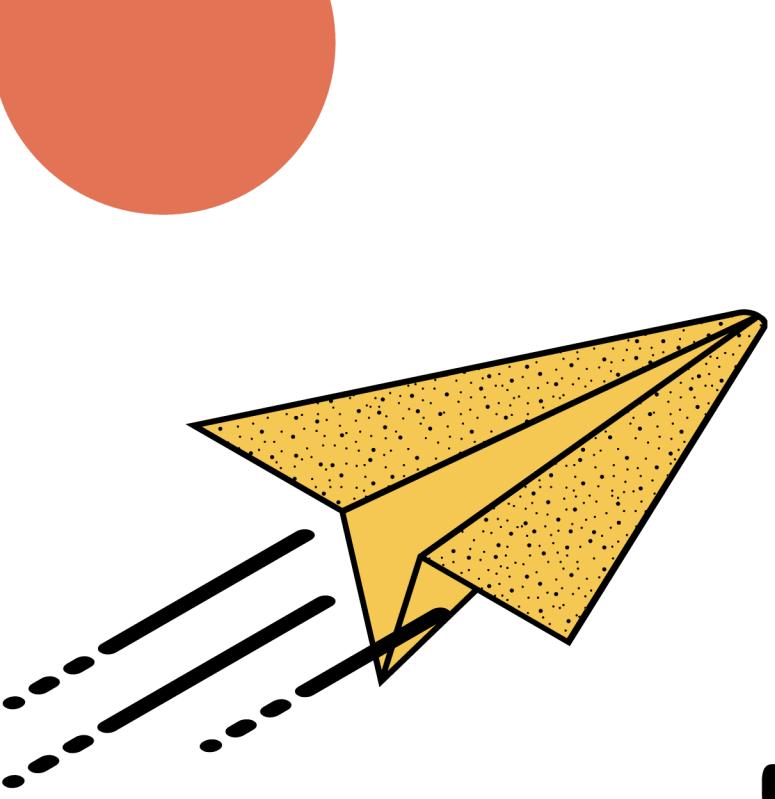
Apply Latent Dirichlet Allocation (LDA) to uncover hidden topics behind sentiment shifts and understand the reasons behind user emotions.

📊 Live Sentiment Dashboards

Build interactive dashboards using tools like Plotly, Streamlit, or Power BI to monitor sentiment in real-time and generate instant reports.

🌐 Multilingual Sentiment Analysis

Expand sentiment detection to support tweets in regional or international languages using multilingual NLP models like BERT or mBERT.



THANK YOU!

