**Real-Time Social Media Sentiment Analysis for Brand Reputation Management**

Organization: Verve Bridge

Internship Type: Real-time, Stipend-based

Tools Used: Python, Power BI

**Project Overview**: This project focuses on developing a system for real-time sentiment analysis of social media content, specifically targeting brand reputation management. The goal is to monitor social media discussions, analyze sentiments towards a brand, identify key topics, and provide actionable insights to help brands manage their reputation effectively.

**Objectives**:

* Perform sentiment analysis of social media posts (positive, negative, neutral).
* Detect trending topics and discussions related to the brand.
* Set up real-time monitoring for sentiment shifts.
* Provide insights for brand reputation management through visualizations and alerts.

**Key Tasks and Implementation** -

**1. Data Collection**:

**Objective**: Collect real-time data from social media platforms to track mentions, comments, and posts related to the brand.

**Tools & Technologies**:

**Kaggle Datasets**: Using datasets from Kaggle to extract and analyze data on brand-related social media discussions.

**2. Data Preprocessing**:

**Objective**: Clean the collected data and prepare it for analysis.

**Tools & Technologies**:

* **Pandas**: For data manipulation and cleaning, including handling missing values, duplicates, and incorrect formats.
* **NumPy**: For numerical operations and handling large datasets efficiently.
* **NLTK / spaCy**: For natural language processing tasks such as tokenization, lemmatization, and part-of-speech tagging.
* **TextBlob**: For performing text analysis, including sentiment extraction and subjectivity analysis.

**3. Sentiment Analysis**:

**Objective**: Classify the sentiment of social media posts (positive, negative, or neutral).

**Tools & Technologies**:

* **TextBlob**: For basic sentiment polarity analysis (positive, negative, neutral).
* **VADER Sentiment Analysis**: For analyzing social media-specific text like slang, emoticons, and short phrases.
* **spaCy**: For advanced sentiment classification using pre-trained models for context-aware analysis.
* **Power BI**: For visualizing sentiment distribution over time and across different social media platforms.

**4. Topic Modeling**:

**Objective**: Identify key topics and themes from the social media content.

**Tools & Technologies**:

* **Gensim (LDA):** For performing Latent Dirichlet Allocation (LDA) to discover underlying topics from the social media text.
* **Matplotlib**: For visualizing the distribution of topics and their occurrence.
* **Seaborn**: For creating attractive statistical plots, such as topic distributions over time.

**5. Real-Time Monitoring**:

**Objective**: Create a dashboard for real-time monitoring of sentiment and topics.

**Tools & Technologies**:

* **Elasticsearch & Kibana**: For storing and querying large volumes of data in real time, and visualizing sentiment and topic trends via interactive dashboards.
* **Power BI**: For creating dynamic, real-time dashboards that display trends in sentiment and key topics.
* **Twilio/Slack API**: For setting up real-time alerts when sentiment drastically shifts or when a topic gains significant traction.

**6. Alerts and Notifications:**

**Objective**: Set up a system that alerts users about significant sentiment changes.

**Tools & Technologies**:

* **Twilio API**: For sending SMS alerts when there is a drastic change in sentiment.
* **Slack API**: For real-time notifications sent directly to a Slack channel.

**Real-Time Sentiment Monitoring Workflow:**

1. **Data Collection:** Use Kaggle datasetsto collect social media data.
2. **Data Preprocessing:** Clean and structure the data using Pandas, NumPy, and preprocess it with NLTK, spaCy, and TextBlob.
3. **Sentiment Analysis:** Classify sentiments using TextBlob, VADER, spaCy, and visualize results in Power BI.
4. **Topic Modeling:** Apply LDA for topic extraction using Gensim, and visualize the results with Matplotlib and Seaborn**.**
5. **Real-Time Monitoring:** Monitor sentiment and topic trends using Elasticsearch and Kibana, and visualize the trends in real-time through Power BI.
6. **Alerts & Notifications:** Set up alerts for sentiment changes using Twilio or Slack API.

**Tools and Technologies Used:**

* Kaggle Datasets: For data collection and analysis.
* Pandas: For data manipulation and cleaning.
* NumPy: For handling large datasets and numerical computations.
* NLTK / spaCy: For natural language processing and sentiment analysis.
* TextBlob: For basic sentiment analysis.
* VADER: For sentiment analysis tailored to social media.
* Gensim (LDA): For topic modeling and uncovering key themes.
* Elasticsearch & Kibana: For real-time data querying and visualization.
* Matplotlib / Seaborn: For creating insightful visualizations.
* Power BI: For real-time dashboards and insights.
* Twilio API: For sending SMS notifications.
* Slack API: For sending notifications to a team on Slack.

**Final Deliverables:**

1. **Data Collection and Preprocessing**: A pipeline that collects and prepares social media data for analysis.
2. **Sentiment Analysis and Topic Modeling System**: A system for classifying sentiment and extracting key topics from the data.
3. **Real-Time Monitoring Dashboard**: A dashboard in Power BI, Kibana, and Elasticsearch for real-time insights into sentiment trends.
4. **Alerts and Notifications System**: A system that sends alerts to users when significant changes in sentiment are detected.
5. **Documentation and Presentation**: A detailed report and presentation summarizing the analysis, methodologies, and insights for stakeholders.