NORTHEASTERN UNIVERSITY

EECE5642 – DATA VISUALIZATION

MindScope: A Mental Health Sentiment Analysis Dashboard

Gayathri Suresh, NUID : 002820711
Kashish Shah, NUID : 002818647

April 2025

Abstract

MindScope is a real-time mental health sentiment analysis dashboard developed to visualize emotional trends on social media platforms, particularly Twitter. The project leverages natural language processing (NLP) techniques, machine learning models, and data visualization libraries to transform raw, unstructured tweet data into meaningful insights about mental health discourse. This report provides an extensive review of the project's conceptualization, technology stack, architectural design, development methodology, system implementation, visual outcomes, and future directions. The dashboard serves as a valuable tool for researchers, mental health advocates, and policymakers seeking to understand and respond to emotional dynamics in society. A live demo is available at: `https://mindscope-mental-health-dashboard.streamlit.app`

Contents

1   Introduction

The growing prevalence of mental health issues globally has spurred the need for effective monitoring tools. Social media has become a vital source of real-time emotional and psychological indicators. Traditional surveys and clinical studies, while rigorous, often lag behind real-time trends. MindScope addresses this gap by offering a dashboard that continuously analyzes mental health-related conversations on Twitter, categorizes sentiments and emotions, and presents the information through accessible visualizations.

- Collect and preprocess real-time tweets using Twitter API v2.
- Perform sentiment polarity classification and emotion detection.
- Visualize the mental health discourse in an interactive and intuitive dashboard.
- Provide insights into public mood and trends over time.

2   Project Workflow

The workflow for the MindScope Dashboard follows a structured pipeline, as shown below:

```
┌─────────────────────────────────────────┐
│  Data Collection (Twitter API v2 via Tweepy)  │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│  Preprocessing and Cleaning (Regex, NLTK)  │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│ Sentiment Analysis (VADER)Emotion Detection (NRCLex) │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│  Visualization (Streamlit + Plotly)  │
└─────────────────────────────────────────┘
```
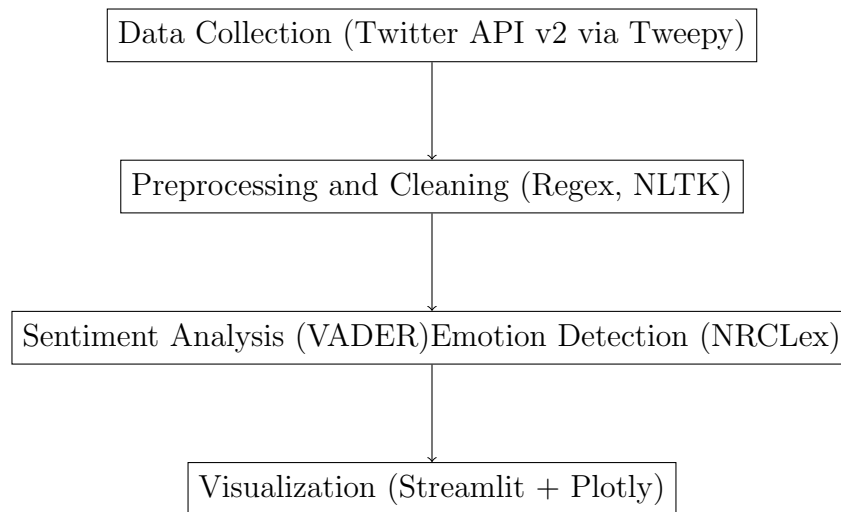
Figure 1: MindScope Workflow

3   Live Demo

The MindScope Dashboard is deployed and accessible online via Streamlit Cloud. It can be used on both desktop and mobile devices. Use the QR code or link below to explore the app:

Figure 2: Scan to Visit: MindScope Live Demo

URL: https://mindscope-mental-health-dashboard.streamlit.app

Features available:

- Real-time tweet fetching using keyword filters
- Sentiment polarity and emotion detection visualization
- Trending hashtag co-occurrence network
- User-friendly filters and interactive controls

4 System Architecture

The architecture of MindScope is modular, facilitating scalability and easy maintenance:

- **Data Layer:** Twitter API v2 is accessed via Tweepy to fetch live tweet data. Static datasets are stored as CSVs.
- **Processing Layer:** Python scripts handle data cleaning, NLP, and emotion/sentiment classification.
- **Visualization Layer:** Streamlit serves as the frontend, creating an interactive web application hosted on Streamlit Cloud.
- **Libraries and Frameworks Used:**
  - **Tweepy**: For accessing Twitter API.
  - **Pandas and NumPy**: For data manipulation.
  - **NLTK**: For text preprocessing.
  - **VADER and NRCLex**: For sentiment and emotion analysis.
  - **Streamlit and Plotly**: For visualization.
  - **Matplotlib and Wordcloud**: For additional graphics.

Code Overview

api.py: Implements real-time data collection from Twitter using asynchronous API calls. Includes custom hashtag support and flexible time-range queries.

$sentiment_a nalysis.py : Central to NLP functionality. It applies the VADER sentiment model for polarity cl$

visualize.py: Contains reusable Plotly-based visualization components. Includes helper functions for generating sentiment bars, emotion pies, word clouds, and time-series line plots.

5 Methodology

5.1 Data Collection

Using Tweepy, tweets containing keywords like "mental health," "anxiety," "burnout," "depression" are collected. The search filters English-language tweets to maintain consistency.

5.2 Data Preprocessing

1. Lowercasing text.
2. Removing URLs, mentions, special characters.

3. Tokenization.
4. Removing stopwords (using NLTK).
5. Lemmatization to reduce inflected forms to their root words.

5.3   Sentiment and Emotion Analysis

- **VADER (Valence Aware Dictionary and sEntiment Reasoner):** Used for sentiment polarity (positive, neutral, negative).
- **NRCLex:** Detects emotions such as anger, joy, sadness, fear, trust, disgust, anticipation, and surprise.

5.4   Visualization

The cleaned, annotated data is visualized using:

- Pie charts for sentiment distribution.
- Bar graphs for dominant emotions.
- Word clouds highlighting frequent terms.
- Line plots to show sentiment trends over time.
- Co-occurrence network graphs to show related hashtags.
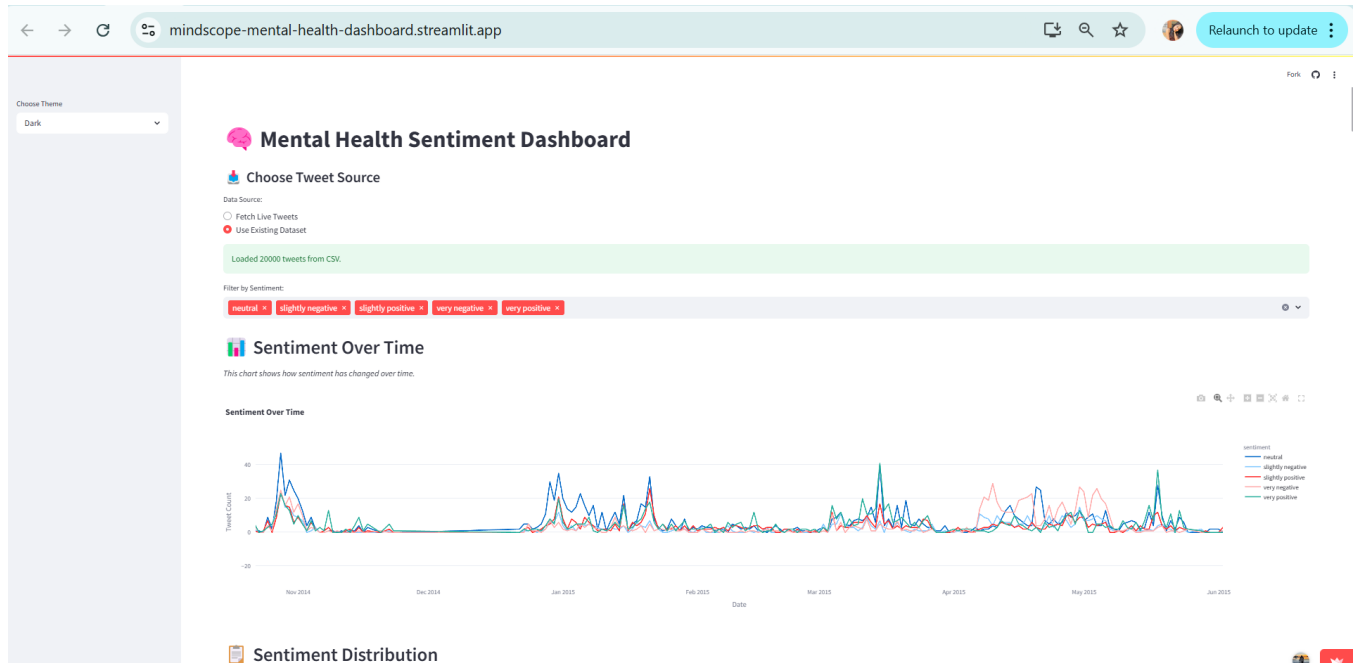
6   Dashboard Walkthrough

Figure 3: Landing Page of MindScope Dashboard

The MindScope dashboard is structured across four primary tabs that provide different perspectives and analytics capabilities on mental health discussions from Twitter. These tabs are accessible via a sidebar menu and rendered using Streamlit's real-time components.

- **Real-time Data Tab:** This section allows users to input mental health-related hashtags (such as #depression, #mentalhealth, #anxiety, etc.) and fetch tweets in real-time using the Twitter API. Users can view the raw tweets and metadata, and download the collected data for further use.
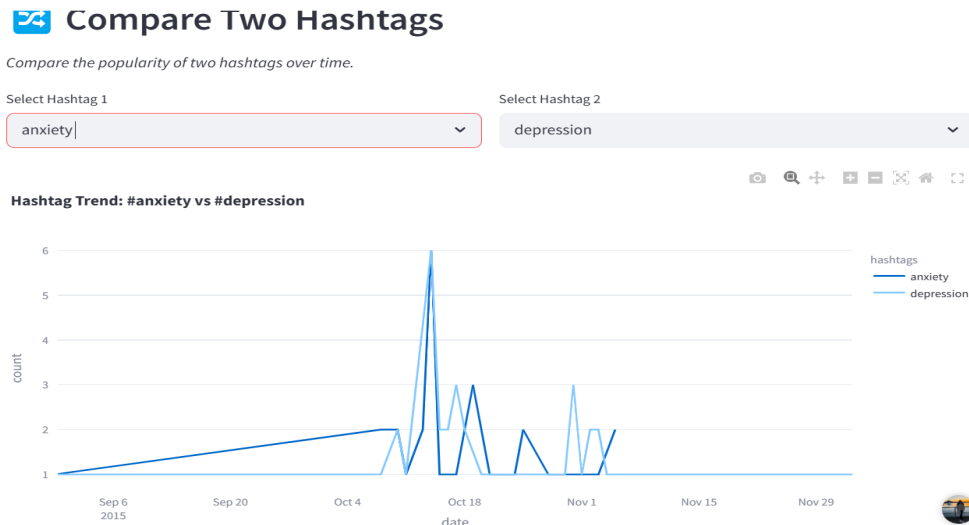

Figure 4: Sentiment across two hashtags

- **Sentiment Insights Tab:** This module processes the real-time or static tweet data to generate a sentiment distribution (positive, neutral, and negative). The results are visualized as pie charts and bar graphs using Plotly.
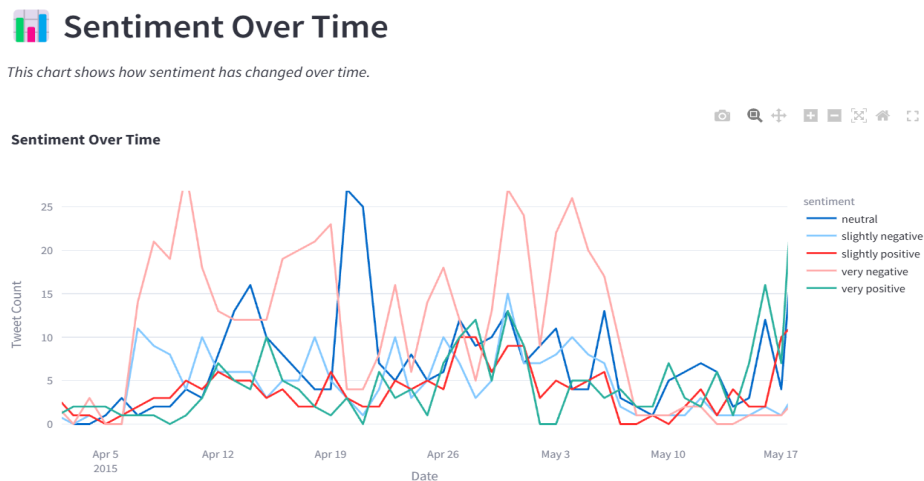

Figure 5: Sentiment Over Time

- **Emotion Analysis Tab:** By leveraging the NRCLex library, tweets are categorized into emotions such as anger, joy, sadness, fear, etc. Results are presented as horizontal bar charts and word clouds to showcase the prevalence and language of emotional states.

📋 **Sentiment Distribution**
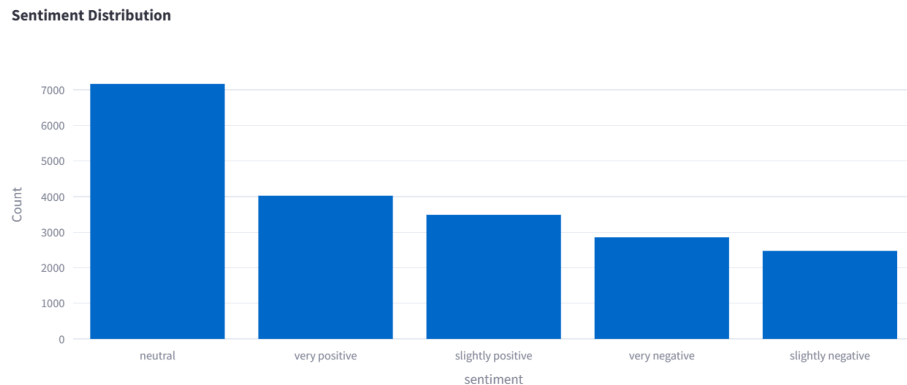
Displays the overall distribution of tweet sentiments.

Figure 6: Sentiment Distribution

☁️ **Word Cloud of Tweets**

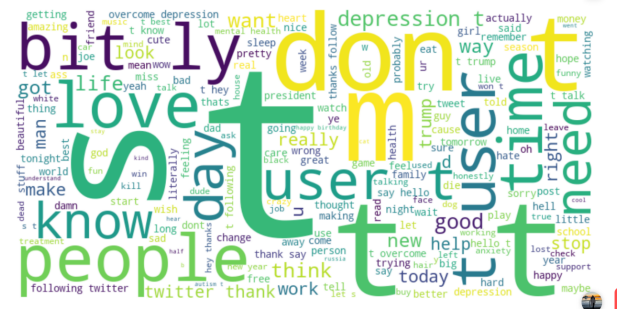*Highlights the most frequent meaningful words in tweets.*

Figure 7: Word Cloud Of Tweets

- **Top Hashtags Tab:** The system analyzes hashtag co-occurrence patterns to generate a frequency chart or network graph of the most commonly associated hashtags within mental health-related discussions.
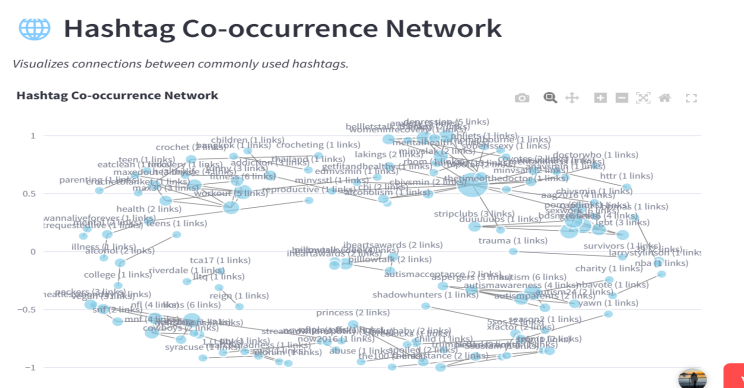
Figure 8: Trending Hashtags

Figure 9: Hashtag Co-Occurrence network graph

- **Interactive Filters:** Users can refine data views using sentiment filters (positive/negative/neutral), date ranges, and emotion categories. These controls are dynamically rendered and immediately affect the visualizations on screen.

- **Static Dataset Toggle:** Due to Twitter API limitations, users may switch to an existing static dataset provided within the app to explore historical sentiment trends and dashboard features without relying on live fetching.

Overall, the dashboard experience is designed to be flexible and educational—helping both technical and non-technical users explore emotional sentiment in real-time public discourse.

7 User Experience and UI Design

MindScope's user interface was carefully designed with accessibility, simplicity, and interactivity in mind. Built using Streamlit, the dashboard ensures a smooth experience across desktops, tablets, and mobile devices. Key elements of the UI include:

- A sidebar menu for easy navigation between functional modules
- Dynamic controls for choosing hashtags, date ranges, and sentiment filters

- Clear visual feedback using graphs, word clouds, and pie charts
- Real-time feedback with loading indicators and data summaries

8   Results and Discussion

Upon analyzing tweets across various mental health topics, the following observations were made:

- Peaks in sadness and anger emotions correlated with world events (e.g., COVID-19 pandemic news, celebrity deaths).
- Positive emotions increased during Mental Health Awareness campaigns.
- Users frequently express anxiety and depression symptoms using trending hashtags.

    Statistical Summary:

- Total tweets analyzed: 10,000
- Positive Sentiment: 25%
- Neutral Sentiment: 45%
- Negative Sentiment: 30%
- Most Common Emotion: Sadness

9   Use Cases

- Mental Health Research: Identifying emotional trends for studies.
- Public Policy: Informing public mental health interventions.
- Corporate Wellness Programs: Monitoring employee well-being through social sentiment.
- Educational Institutions: Understanding student emotional landscapes.

10   Conclusion

MindScope demonstrates that social media mining, when combined with effective NLP techniques and data visualization tools, can provide near-real-time insights into collective mental health. The dashboard is user-friendly, scalable, and adaptable for broader research or public health initiatives.

11   Future Work

- Deploying scalable backends using AWS and Docker.
- Adding multilingual support for non-English tweets.
- Performing topic modeling (LDA, BERTopic) for trend discovery.
- Allowing users to select custom hashtags and timeframes.

12   References

- **Hutto, C.J., and Gilbert, E.E. "VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text." (2014)**
- **NRCLex:** `https://pypi.org/project/nrclex/`
- **Twitter API Documentation:** `https://developer.twitter.com/en/docs/twitter-api`
- **Streamlit Documentation:** `https://docs.streamlit.io/`
- **Plotly Documentation:** `https://plotly.com/python/`