Natural Language Processing
Emotion-Based Conversational Filter Project

Presented by

B. Swarna 2203A51540

D. Sucharitha 2203A51279

1.Introduction:

The goal of this project is to analyze and filter conversations based on their emotional content. By processing a dataset of conversations with emotional labels, the project helps identify, visualize, and export conversations filtered by specific emotions. This allows users to gain insights from emotional expressions in text data.

2. Objectives:

The main objectives of the project are to:

- Analyze conversations based on their emotional content.
- Visualize the distribution of emotions in the dataset.
- Provide interactive filtering to display conversations of a selected emotion.
- Allow users to export filtered conversation data for further analysis or reporting.

3. Methodology:

The methodology of the project follows a step-by-step process:

- **Loading Dataset:** The dataset containing conversations and their corresponding emotions is loaded using Pandas.

- **Data Cleaning:** The dataset is checked for missing or inconsistent values. Any missing data is removed to ensure the dataset's integrity.
- **Emotion Count:** The frequency of each emotion (e.g., Joy, Sadness, Anger) is counted using the `Counter` function.
- **Visualization:** The project generates visual representations (bar charts and pie charts) to show the distribution of emotions within the dataset using the Matplotlib library.
- **Filtering Conversations:** Conversations are filtered by the user's specified emotion. Only conversations with the selected emotion are displayed.
- **Text Preprocessing:** Optionally, text data is cleaned (e.g., lowercase conversion, extra spaces removal) to maintain consistency and improve accuracy.
- **Export Data:** Filtered data is saved into Excel files for future use, analysis, or sharing.
- **Interactive Mode:** Users can input specific emotions to dynamically filter and view relevant conversations from the dataset.

4. Algorithms and Techniques:

The following techniques and algorithms were used in the project:

- **Data Filtering:** Conversations are filtered based on emotions using simple conditional selection in Pandas.

```
filtered_data = df[df['Emotion'] == 'Joy']
```

- **Text Preprocessing:** Basic text cleaning techniques such as converting the text to lowercase and removing unnecessary spaces are used to standardize the input.

```
df['Text'] = df['Text'].str.lower().str.strip()
```

- **Emotion Counting:** The `Counter` class from the Python `collections` module is used to count the occurrences of each emotion in the dataset.

```
emotion_counts = Counter(df['Emotion'])
```

- **Data Visualization:** Emotion distribution is visualized using Matplotlib to produce bar charts and pie charts for easier understanding of the dataset's emotional breakdown.

```
df['Emotion'].value_counts().plot(kind='bar',
title='Emotion Distribution', color='skyblue')
```

5. Workflow:

The workflow of the project is as follows:

- **Load the Dataset:** Import the dataset into the program.
- **Perform Exploratory Data Analysis (EDA):** Analyze the dataset to get an overview of emotions and their frequencies.
- **Filter Conversations by Emotion:** Based on user input, filter conversations that correspond to a particular emotion.
- **Visualize the Emotion Distribution:** Create visual representations of how the emotions are distributed within the dataset.
- **Provide the Option to Save Filtered Conversations:** Allow users to save the filtered dataset to Excel files.
- **Allow Interactive Filtering by User Input:** Enable users to choose emotions dynamically and interact with the dataset.

6. Conclusion:

This project effectively filters and analyzes conversations based on emotions. It provides valuable insights into the emotional distribution of conversations and allows users to filter, view, and export the data based on specific emotions. The interactive filtering and data export functionality make the project suitable for further analysis and reporting.