

ARTIFICIAL INTELLIGENCE

PROJECT REPORT

**Title:** WhatsApp Chat Analyzer

Members:

Sahir Mansoor 21K-3449

Wajih Hyder 21I-2957

**Introduction**

The WhatsApp Chat Analyzer is a data visualization tool designed to analyze WhatsApp chat data. It provides insights into various aspects of the chat, such as message frequency, most active users, popular words, and emoji usage. The tool enables users to upload their WhatsApp chat data and obtain visual representations of key metrics and trends within the chat.

**Features**

**Upload WhatsApp Chat Data:** Users can upload their WhatsApp chat data directly to the tool.

**Message Statistics:** Provides statistics such as total messages, words, media shared, and links shared.

**Timeline Analysis:** Offers insights into chat activity over time, including monthly and daily timelines.

**Activity Map**: Visualizes chat activity on a weekly and monthly basis.

**Most Active Users:** Identifies the most active users in the chat.

**Word Cloud:** Generates a word cloud representing the most common words used in the chat.

**Most Common Words:** Presents a bar chart showing the most commonly used words in the chat.

**Emoji Analysis:** Analyzes emoji usage and presents the most frequently used emojis.

**Technologies Used**

**Python:** The entire project is built using Python programming language.

**Streamlit:** Streamlit is used for creating the web-based user interface.

**Pandas:** Pandas is utilized for data manipulation and analysis.

**Matplotlib:** Matplotlib is used for creating various plots and visualizations.

**Seaborn:** Seaborn is used for enhancing the aesthetics of the visualizations.

**WordCloud:** WordCloud library is used to generate word clouds.

**Emoji:** Emoji library is used to handle emoji-related operations.

**Regular Expressions (Regex):** Regex is used for text preprocessing and pattern matching.

**Project Structure**

**app.py:** This file contains the main Streamlit application code, including UI layout, data processing, and visualization rendering.

**helper.py:** This file contains helper functions responsible for processing data and generating visualizations.

**preprocessor.py:** This file contains functions for preprocessing the raw WhatsApp chat data.

**stop\_hinglish.txt:** This file contains a list of stop words in the Hinglish language, used for filtering out common words in the chat.

**wordcloud.png:** This is the output file where the generated word cloud image is saved.

**Workflow**

**Upload Data:** Users upload their WhatsApp chat data using the file uploader.

**Data Preprocessing:** The uploaded data is preprocessed to extract relevant information such as message content, user names, and timestamps.

**Statistical Analysis:** Various statistics such as message frequency, word count, and media sharing are computed.

**Visualization:** The processed data is visualized using plots and charts to provide insights into chat activity, word usage, and emoji usage.

**Interactive UI:** The Streamlit UI allows users to interact with the visualizations and customize the analysis based on their preferences.