COURSE PROJECT PROPOSAL

Project Title: WhatsApp Chat Analyzer
Course: CS699 Software Lab
Instructor: Bhaskaran Raman
Group Members: Tirthesh Jain (23M0758), Aditya Tomar (23M2103)

Abstract
WhatsApp, boasting over 2 billion active users, is a globally favoured messaging app used for personal, professional, and educational communications. One of the challenges of using WhatsApp for marketing is to understand what kind of content resonates with the users and when is the best time to send messages associated with the same. This is where a 'WhatsApp chat analyser' can be helpful.
This project proposes to develop a WhatsApp Chat Analyzer web application that will allow to upload a WhatsApp chat export file, and the application will provide various insights and visualizations to help users understand their chat patterns, usage statistics, and more.

Features
The following are the key features of our WhatsApp Chat Analyzer:

- User sign-up/login functionality with authentication
- Secure storage of usernames and password hashes in a PostgreSQL database
- The application will provide various plots, charts, and statistics to analyse WhatsApp chats, including:
    - Total messages, words, media, and links shared
    - Monthly, weekly and daily timeline for messages
    - Most active users
    - Most popular words and phrases
    - Word cloud of most used words
    - Sentiment analysis

Targeted Users

Our WhatsApp Chat Analyzer will be beneficial to a variety of users, including:

- Individuals who want to better understand their WhatsApp communication patterns
- Businesses that want to understand what kind of content is suited for their customers on WhatsApp along with identifing the best times to send messages while optimizing their marketing strategies for increased engagement and conversion.
- Researchers who want to study WhatsApp communication patterns

Implementation

The WhatsApp Chat Analyzer will be implemented using the following technologies: **Python, Streamlit, PostgreSQL, nltk, matplotlib, pandas, wordcloud, seaborn, urlextract, emoji, PyCharm etc.**

- Streamlit authenticator library will be used to implement user sign-up/login functionality with authentication.
- Usernames and password hashes will be stored in a PostgreSQL database to ensure security.
- The nltk library will be utilised to perform various natural language processing tasks, such as stop word removal, and sentiment analysis.
- Libraries like matplotlib, pandas, wordcloud, seaborn, and urlextract will aid in the generation of various plots, charts, and statistics.
- Emoji library will be used to parse and display emojis in the WhatsApp chat data.
- Finally, PyCharm as our integrated development environment (IDE).