JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

WHATSAPP CHAT ANALYZER

[1] Selvavani Mathavan, [2]Alishana Thorat, [3]Kanishka Panpatil, [4]Sneha Kushwaha, [5]Prajakta khaire

[1]Student, [2]Student, [3]Student, [4] Student, [5] Assistant Professor
[1]Information technology,
[1]SSJCOE, Mumbai, India

Abstract : This research paper presents a comprehensive analysis of WhatsApp chat data utilizing machine learning techniques. The study focuses on extracting valuable insights such as total message count, media distribution, link sharing frequency, and emoji usage within the context of group conversations. Additionally, it investigates temporal trends through a monthly timeline visualization and identifies the most active participants using a bar graph representation. By employing natural language processing and data visualization methodologies, the study aims to provide a deeper understanding of communication patterns and dynamics within WhatsApp groups, contributing to the field of chat analytics and social network analysis.

***Keywords*- WhatsApp chat text file, NLP, Matplotlib, Seaborn Emoji analysis, Streamlit, etc.**

I. INTRODUCTION

In recent years, the proliferation of messaging applications has revolutionized the way individuals communicate, with WhatsApp emerging as one of the most popular platforms worldwide. The abundance of textual and multimedia content exchanged within WhatsApp groups presents a rich source of data for analysis. This research seeks to delve into the realm of chat analytics within WhatsApp groups, leveraging machine learning techniques to extract meaningful insights. Understanding the dynamics of group conversations, including message distribution, media sharing patterns, and link dissemination, holds significant implications for various fields such as social network analysis, communication studies, and digital sociology. By applying natural language processing algorithms, this study aims to uncover hidden patterns and trends within WhatsApp chat data, shedding light on the intricacies of modern communication. Additionally, visualizations such as line charts and bar graphs will be utilized to provide a clear representation of temporal trends and participant activity levels. Through this research, we endeavor to contribute to the growing body of knowledge surrounding chat analytics and pave the way for further exploration in this field.

This information can only be obtained by some research in the context of the desired device. Relevant data is very much needed because many people are interested machine learning to create models that help solve various problems. These models require well-designed learning a model that better enables these models. Our project provides deep data search in different types of Whats App chats. This project can be useful for various applications such as market research, social studies and psychological studies. The project can provide valuable insights into consumer behaviors, political sentiment and mental health. Research. The paper aims to explore various aspects of this project and its potential applications.[1]

II. LITERATURE REVIEW

A. Examine the utilization and influence of WhatsApp Messenger through a demonstration study.[1]

Extensive research and analysis has delved into the usage and impact of WhatsApp across various demographics. Some studies have focused on its effects on students, while others have focused on specific local populations. For example, a study in South India surveyed individuals between the ages of 18 and 23 and revealed that young people spend about 8 hours a day on WhatsApp and stay online for about 16 hours. The research highlighted WhatsApp's key role in facilitating peer-to-peer communication, with participants routinely exchanging multimedia content such as images, audio and videos. Notably, WhatsApp has emerged as the predominant application used by the youth while using smartphones. Using analytical methods, these investigations aim to explore the features of WhatsApp and determine its positive and negative effects on users.

B. Analysis of the content within WhatsApp chats.[1]

Research studies measure the effectiveness of the WhatsApp application. This study will be a major to explore the possibility of WhatsApp to become the leading mobile application. With the development of technology and the advent of mobile phones, communication has completely changed. The growth of smartphones and social networking apps makes communication faster and easier than ever before. People may no longer have enough money to eat, sleep or dress, but they have a phone in their pocket to interact with family, friends and customers. At the same time, smartphones and instant messaging applications such as WhatsApp, Viber and Skype have taken over the communication world. At the same time, smartphones and have taken over the communication world.

C. Forensic analysis of WhatsApp Messenger.[2]

. User-to-user communication, broadcast messaging, and group conversations are just a few of the several types of communic--ation that WhatsApp offers its users. Users may communicate by exchanging simple text messages, multimedia items (it has voice, video, and picture files), contact cards, and geographical data. Each user has a profile, or collection of data, that include their WhatsApp name, status update, and avatar (a visual file, usually a photo). Each user's profile is kept on a centralized system and may be accessed by other WhatsApp users who have added that person to their contacts. Other services offered by the central systems include message relay, user registration, and authentication.

III. METHODOLOGY

1.Data extraction: Data is extracted from WhatsApp by clicking on export button. After exporting, a text file is created which includes raw data. The raw data is given to WhatsApp chat analyzer to perform preprocessing. [2]

2.Data Collection: Use the "Export Chat" functionality to send the entire conversation in text format to your WhatsApp number. Download the exported chat from your WhatsApp number.[2]

3.Data Preprocessing: In this module data preprocessing is done by removing all the unwanted content from the raw data. Only the data required for analysis is taken into consideration. [2] The chat data is separated like user name, user message, year, month, date, hours, minute, seconds.

4.Data Import: The system uses Python programming language and panda library to import the WhatsApp chat data, which is in a specific format, into a pandas Data Frame.[1]

5.Data Visualization: The system uses data visualization libraries such as matplotlib and seaborn to create visualizations such as bar graphs, and word clouds to help users understand the data.[1]

6.Importing Data into website: Users can choose between overall group analysis and a specific user analysis. The user then clicks the display analysis option to analyze the imported file after choosing the user. Analysis of an imported WhatsApp text file is displayed.[2]

7.Statistical Representation: Various graphical representations are used to display the preprocessed data.[2]

IV. WORKING OF PROJECT

1. Data Collection and Pre-processing: This phase involves collecting WhatsApp chat data from the source and preparing it for analysis. Preprocessing tasks may include it segregates the messages like user name, user message, days, month, year, hours, minutes and second. It organizing data into a structured format suitable for analysis.

2. Feature Extraction: Once the data is cleaned and organized, relevant features are extracted to provide insight into chat activity. This includes calculating metrics such as total words, messages and media shares, links shared. Shared media may include images, videos, audio files, or other attachments exchanged within the chat.

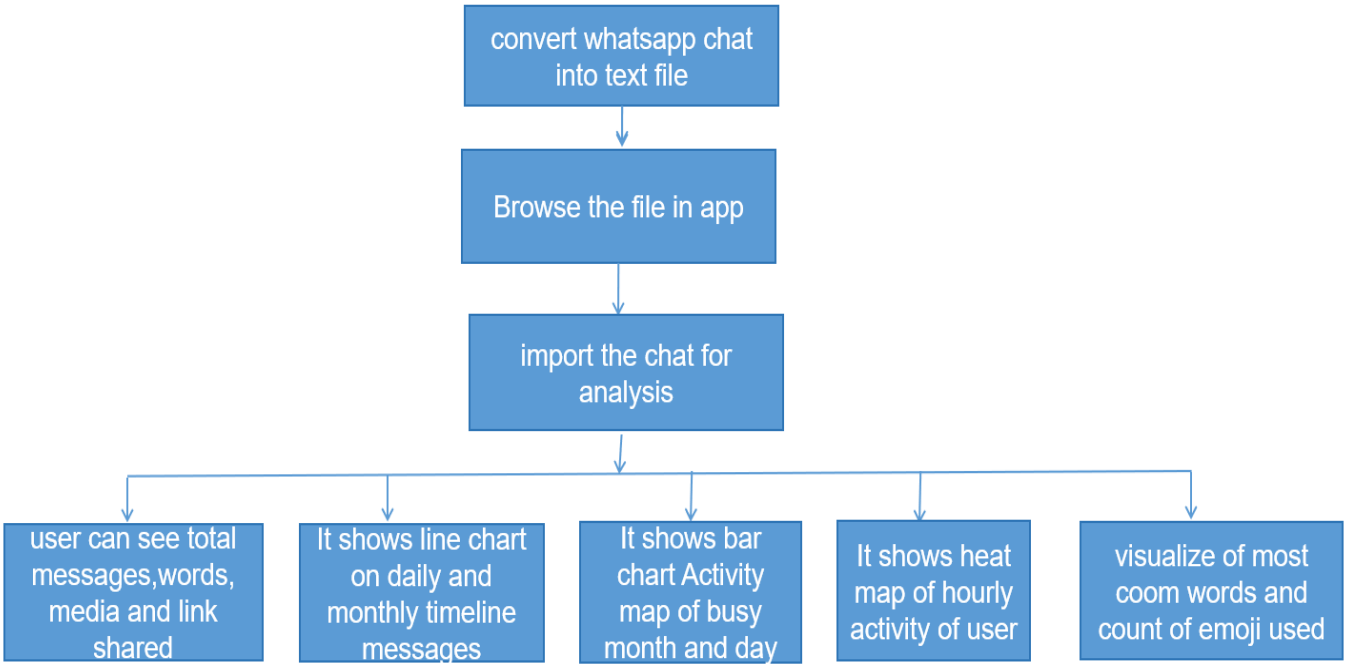
3. Timeline Analysis: Graphical representations such as line graphs or bar graphs are created to visualize chat activity over different time periods. Weekly timelines show trends in messaging behavior throughout the week, while daily timelines provide a more detailed look at activity patterns throughout the day. Analyzing these timelines helps identify peak usage times, quiet periods, and any recurring communication patterns.

4. Identification of the most busiest users: Machine learning algorithms are used to analyze user activity and identify the busiest participants in the chat group. This includes assessing metrics such as message frequency, total posts, and engagement levels to determine who is most actively engaged in the conversation.

5. Analysis of emojis and words: Textual content is analyzed to identify commonly used emojis and count number the emoji. Emoji add context and emotion to messages, so analyzing their frequency provides insight into the overall tone of the conversation. Similarly, word frequency analysis helps identify recurring themes, topics of interest, or commonly used phrases within a chat.

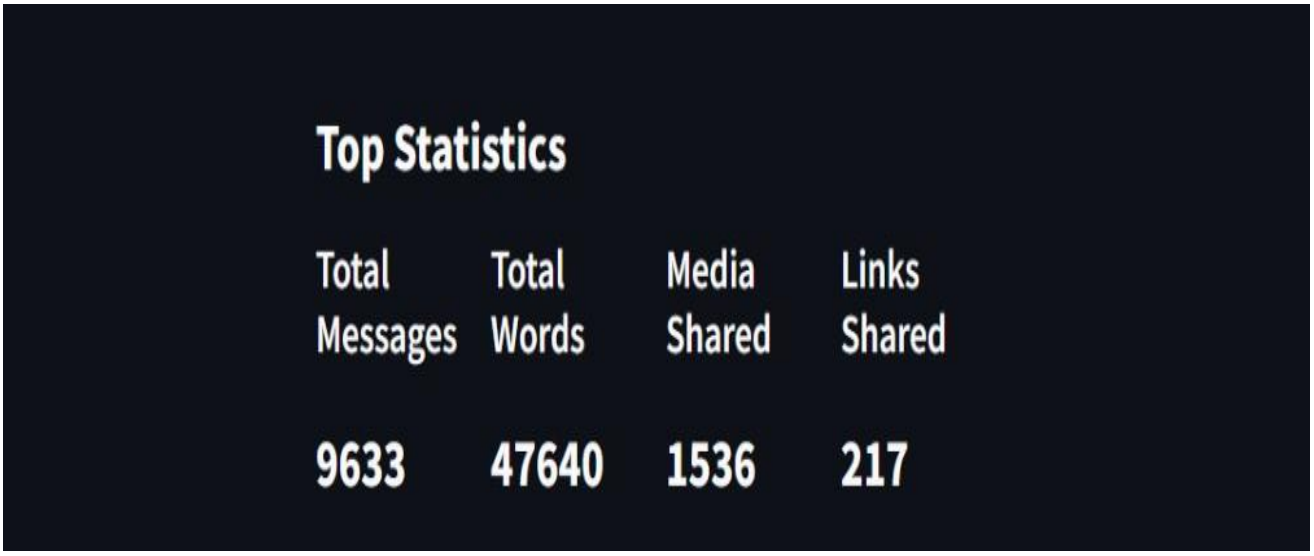
6. Visualization: Insights derived from analysis are presented visually through graphs and word clouds. Charts can include bar charts showing the frequency of messages over time, histogram for showing the busy users, activity map for showing the analyze which time the users are communicating or word clouds showing the most frequently used words in a chat. These visualizations make it easier for users to interpret and understand the data and allow them to gain useful insights from their conversations.

V. FLOWCHART

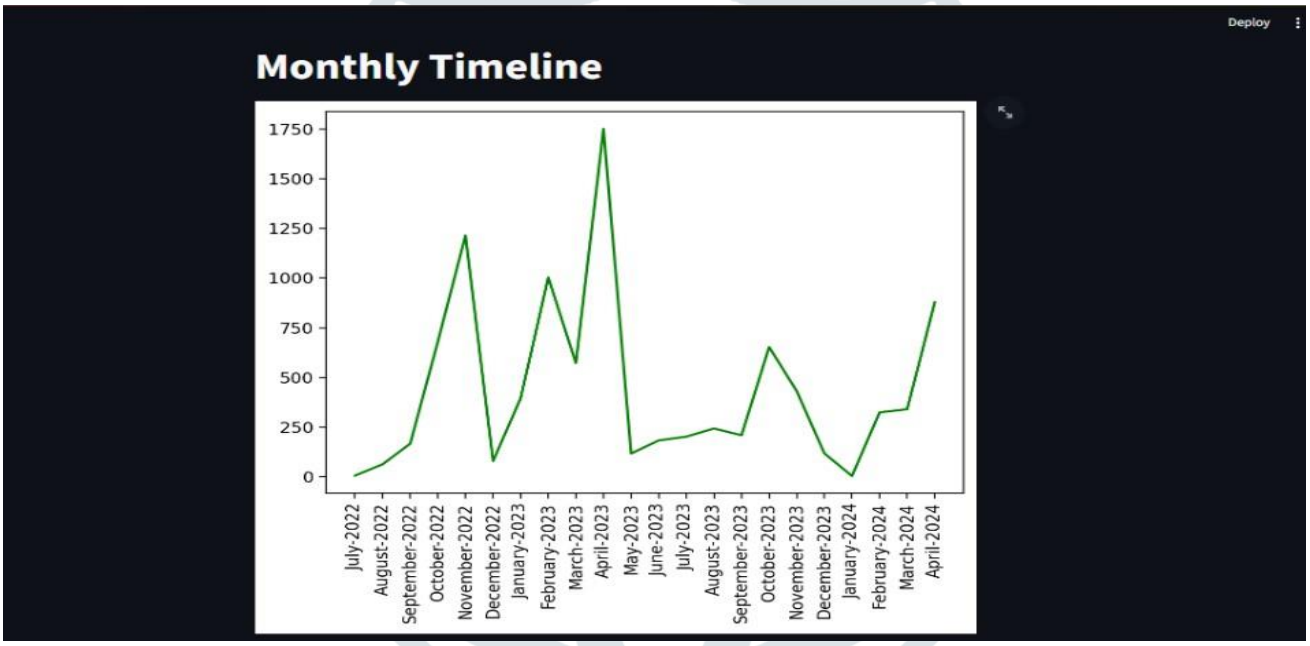


VI. RESULTS

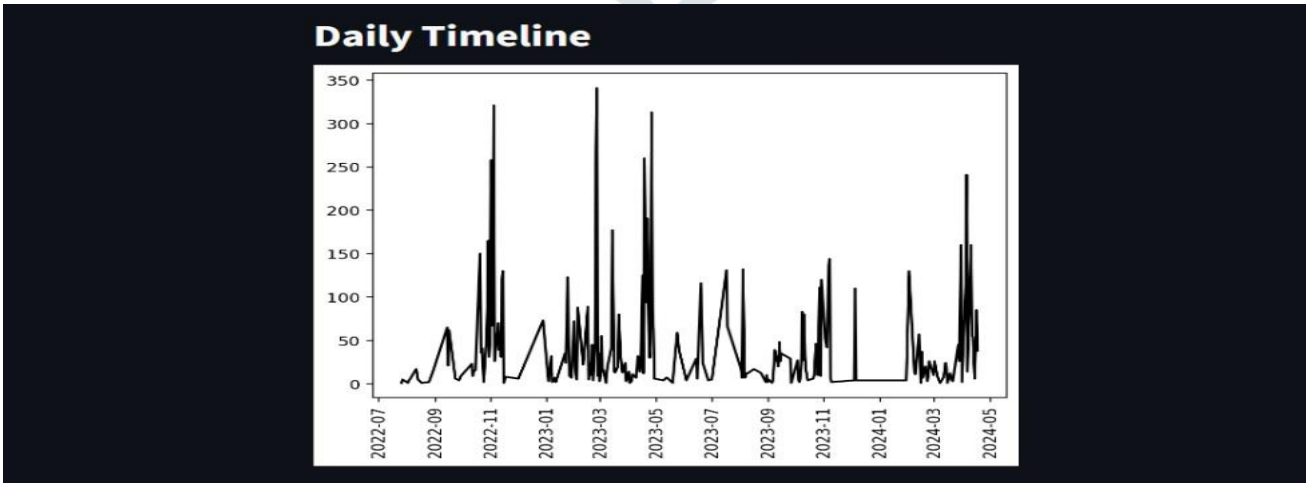
Top statistics:



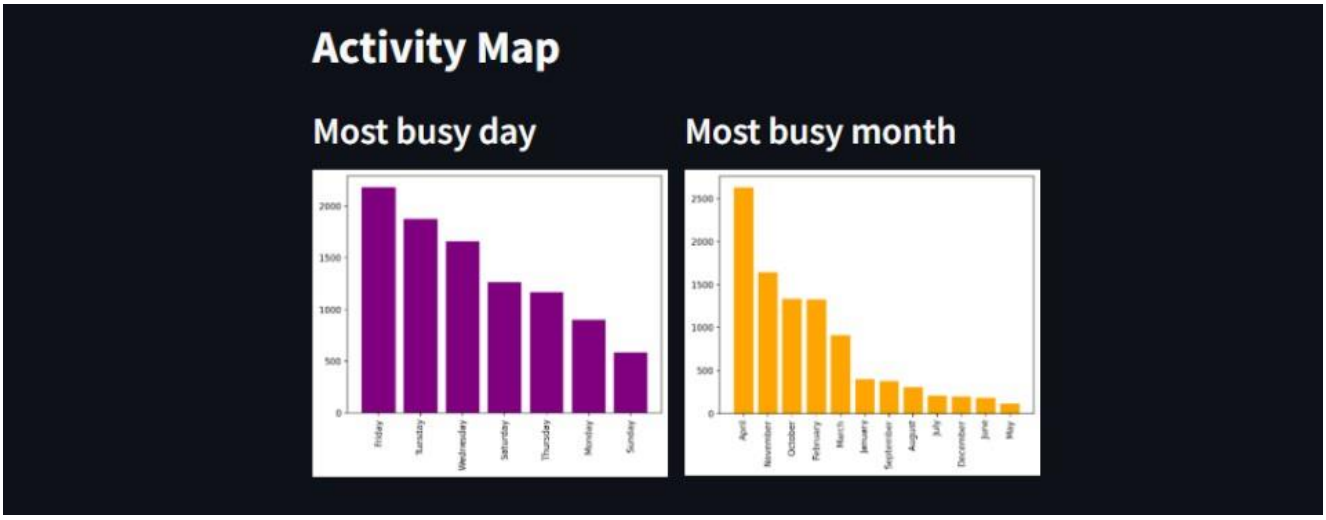
Monthly Timeline:



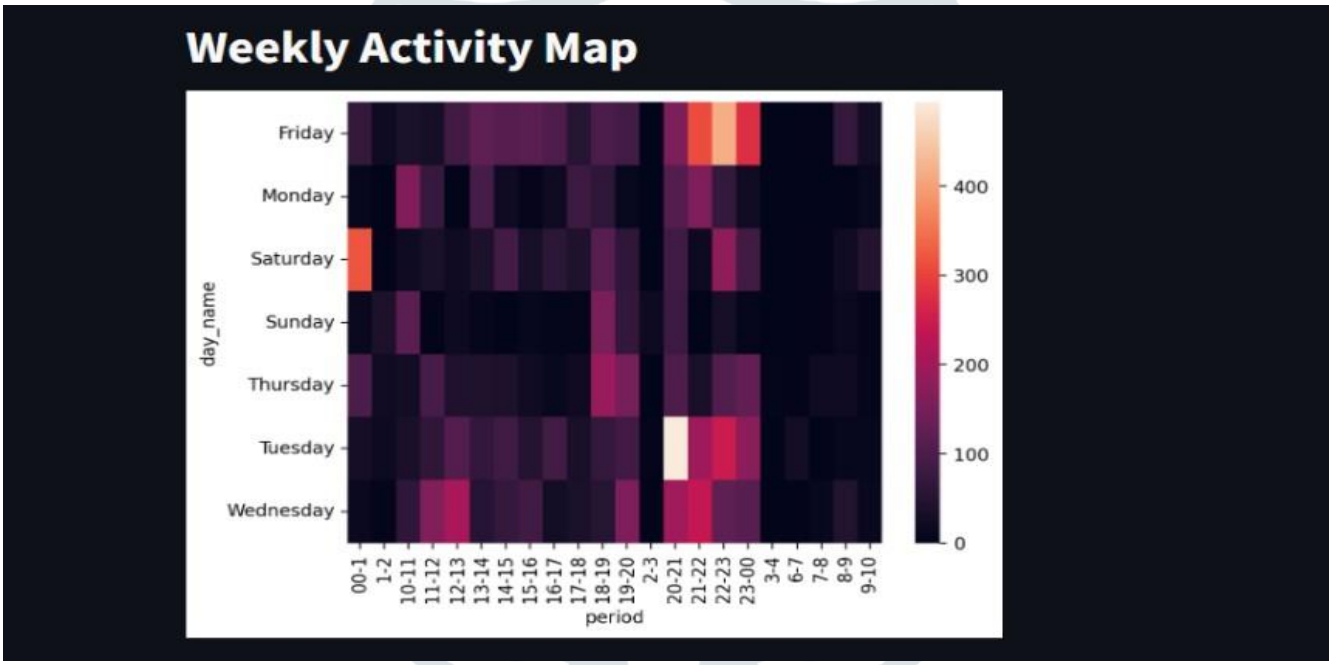
Daily Timeline:



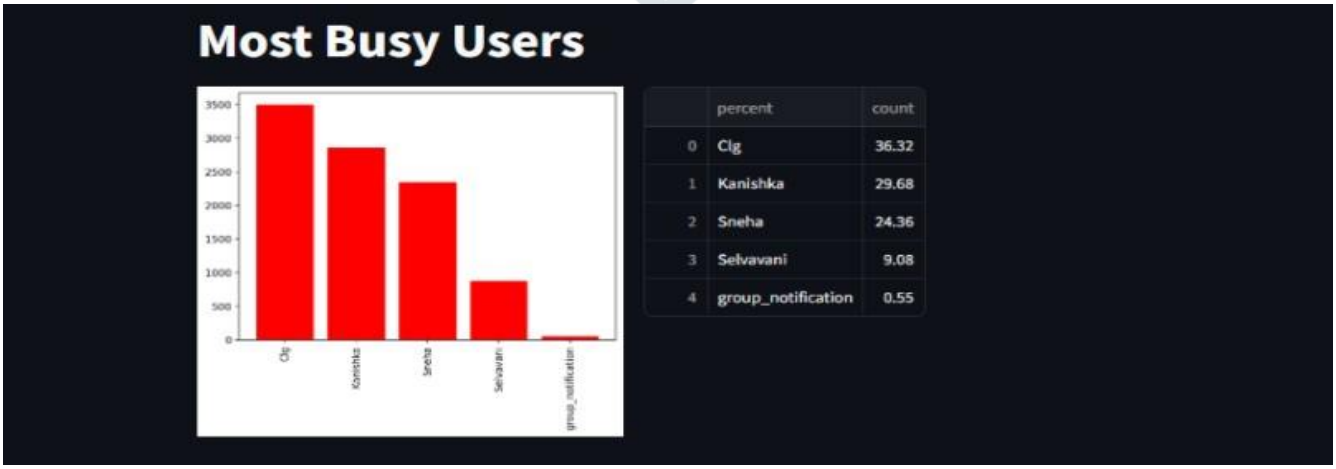
Activity Map:



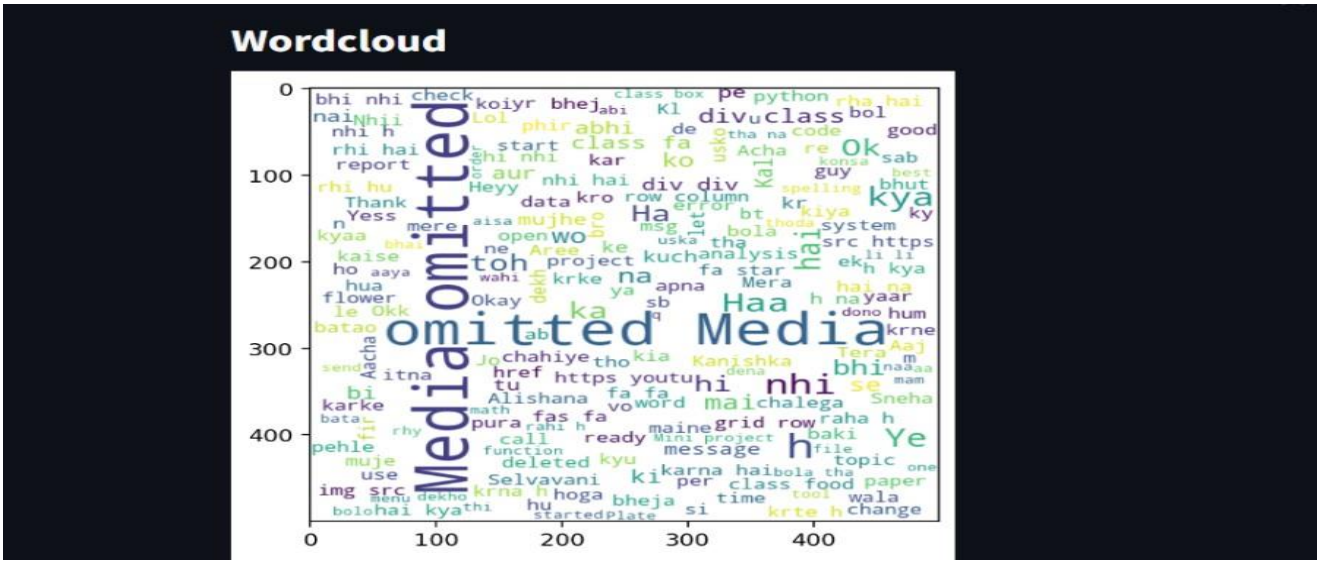
Weekly Activity:



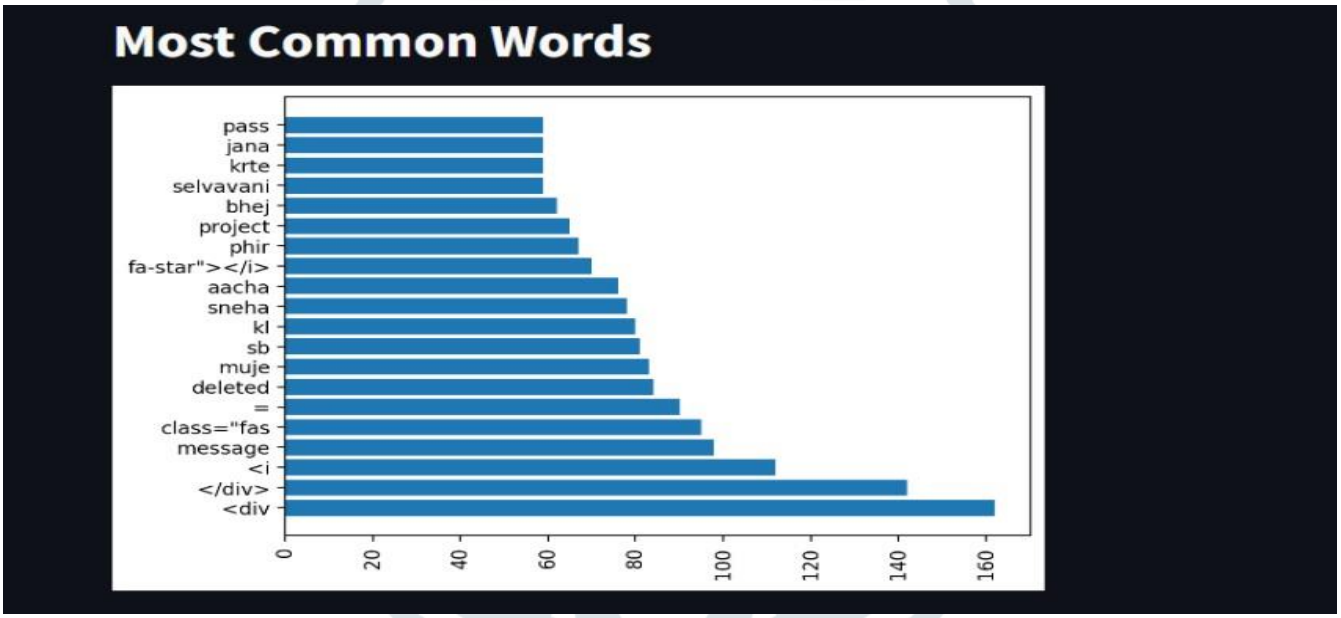
Busy User:



Word Cloud:



Most Common Words:



Emoji Analysis:

VII FUTURE SCOPE

We can make WhatsApp Chat Analyzer even better in the future! One thing we could do is see if people are feeling happy, sad, or just okay in their messages. We could also make it smart enough to understand what topics people are talking about the most.

Additionally, we may be able to provide each person using the Application with their own special insights based on how they chat. Imagine if the app allowed you to play more with the data, such as filtering the news or creating great news. It could also pull data from other places like your calendar or social media to give you even more information. And wouldn't it be nice if the app could predict when you'll chat the most in the future? We could even make it work with other messaging apps, not just WhatsApp. Of course, we make sure everything is private and secure, so your cottages stay safe. Overall, the future of WhatsApp Chat Analyzer is super exciting, with tons of ways to make chatting even more fun and insightful. Chatbots is used fort chat analysis and to feed data into your chatbot for smarter responses based on past conversations. Predictive analyzer of user by using machine learning models to predict future communication patterns or needs based on historical data. Behavioral analyzer of user to Analyze patterns to better understand user behavior, such as predicting when users are likely to start a chat or what topics they prefer to discuss.

VIII CONCLUSION

The WhatsApp Chat Analyzer project is an exciting exploration into the fields of machine learning, natural language processing and data visualization. At its core, it seeks to delve into the intricate details of WhatsApp conversations, uncovering the hidden patterns and insights that lie within. Using advanced algorithms, the project tries to decode the nuances of the language used in these chats, identifying common themes, feelings and communication patterns. Through interactive visualizations and intuitive interfaces, users can gain a deeper understanding of their messaging habits and behavior. Ultimately, the goal is to enrich users' digital communication experiences by providing them with valuable insights that improve the way they interact and engage on the platform.

VII. ACKNOWLEDGMENT

We extend our heartfelt gratitude to Professor. Prajakta Khaire, our supervisor and guide, for her unwavering support and invaluable assistance throughout the completion of our project, the WhatsApp Chat Analyzer. We also thank Dr. Pramod Rodge, our Principal, and Dr. Savita Sangam, our HOD, for providing us with this enriching opportunity and their insightful guidance. Their collective mentorship and encouragement have been indispensable in achieving this milestone, and we are deeply appreciative of their contributions to our success.

REFERENCE

[1] Er. Farzana Khan, Farooq Shaikh, Rohan Shirke, Raj Manjrekar, Aakash Kumar Jha, "WhatsApp Chat Analysis and visualization", International Journal of Advanced Research in Science, Communication and Technology, vol3, issue8, April 2023.

[2] Dr. T. Siva Ratna Sai1, T. Naga Nandini2, M. Harsha Vardhan3, B. Sivaji4, R. Sai Kalyan5, "CHAT ANALYSIS ON WHATSAPP USING MACHINE LEARNING", Journal of Engineering Sciences, Vol 14 Issue 04,2023.

[3] Shaikh Mohd Saqib*1, Prof. Sujata Bhosle*2, "WHATSAPP CHAT ANALYZER", International Research Journal of Modernization in Engineering Technology and Science, Volume:04/Issue:05/May-2022.

[4] Blessing Nwamaka Iduh "WhatsApp Network Group Chat Analysis Using Python Programming", International Journal of Latest Technology in Engineering, Management & Applied Science Volume IX, Issue II, February 2020.

[5] Ravishankara K, Dhanush, Vaisakh, Srajan I S "WhatsApp Chat Analyzer" International Journal of Engineering Research & Technology (IJERT), Vol. 9 Issue 05, May-2020.