

# WhatsApp Chat Analyser

Eshan Bhuse<sup>1\*</sup>, Om Doshi<sup>1</sup>, Sahil Gavande<sup>1</sup>, Anuj Tawari<sup>1</sup>

<sup>1</sup>Department of Computer Engineering, Sardar Patel Institute of Technology, Mumbai, India

**Abstract:** WhatsApp is a well-known social networking platform with more than 2 billion members. WhatsApp is a vital communication tool in people's lives. They use WhatsApp to send text messages expressing their emotions. Since WhatsApp is available in more than 180 countries, code-switching is frequent. Additionally, as emojis have grown more widely used and prevalent, they are now essential for sentiment research. This paper discusses our method for transforming unstructured WhatsApp messages into a structured format that can be used for sentiment analysis using a variety of data mining techniques. We go over how we handle codemixing, the many emojis and the feelings they represent, and how to use this method to conduct basic analysis.

**Keywords:** Machine learning, WhatsApp, sentiment analysis, emoji, library.

## 1. Introduction

This project analyses and processes data. So, the first thing you got to do when using a learning algorithm is to figure out how the model can learn get better. Data pre-processing is super important in the world of machine learning. To make the model work really well, we need a bunch of data. That's why we decided to look at all the data coming from WhatsApp, which is owned by Facebook, a huge data producer.

WhatsApp says there are about 55 billion messages flying around every day. Crazy, right? The average person spends around 195 minutes per week chatting on WhatsApp and is part of lots of groups. With all this data just waiting for us, we need to dig in and find out more about all those messages our phones have to deal with.

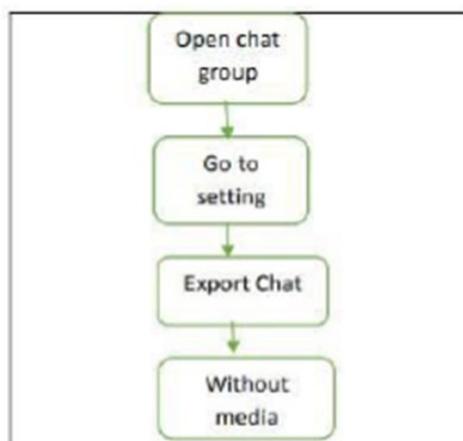


Fig. 1. Block diagram of WhatsApp chat analyzer

This tool collects all that interesting data from your WhatsApp chats and makes cool pie charts and diagrams to show you what's going on. You know what to do - back up your chat and send it to the email address on the site.

## 2. Motivation

Making a WhatsApp chat analyser is a cool project that helps you into how people talk be more efficient, and as a person. You use fancy computer tricks to learn from chats without spying on folks. Looking at chat info can show interesting stuff like what folks about most, how often they text, and when they. This info helps you see who's connected, get more done, and be better at talking to folks. Plus, making this thing helps you get better at something called NLP, which lots of folks want to know about. Sharing this tool with others helps everyone learn about chatting online. It's fun to work on, it's important these days, and it can help you get ahead in your life.

## 3. Literature Survey

Whatsapp Sentiment Analysis Using R - Many messaging apps, such as WhatsApp, are utilized for conversions in the modern internet era. The WhatsApp software is used by all smartphone users for video calling and private and group messaging. Individuals sign up for a range of desired and unwanted WhatsApp groups, which leads to disarray, insufficient storage space, and unreported problems. From mining Twitter to Facebook posts to collecting online reviews, sentiment analysis is increasingly influencing the creation and delivery of more products and services that meet customers' desires. These days, it is a little bit simpler to accomplish thanks to enhanced packages in widely used programming languages. This study uses R to do sentiment analysis on WhatsApp data. This captures the essence of the WhatsApp community, regardless of a successful or unsuccessful conversion. Users can discuss those WhatsApp groups based on the outcome.

WhatsApp Group Chat Analysis by using Machine Learning - In the recent past, sharing information between individuals has been done often via WhatsApp. An important part of internet-based communication is WhatsApp. This study's primary objective is to give a thorough analysis of the WhatsApp group conversations and pinpoint the members of the group who are the most active. The Kaggle repository will be used to get the dataset, which will then be analyzed to determine the most popular message type in group chats, the day of the week with

\*Corresponding author: eshan.bhuse@spit.ac.in

the highest activity, and the age group with the highest activity. The act of cleaning, examining, and modeling the data that will be utilized by this system is part of the data analysis step.

Regression and other machine learning methods will be used for the examination and interpretation of the findings. The implementation will make use of Google Collaboratory.

**WhatsApp Analyzer: A Tool for Collecting and Analyzing WhatsApp Mobile Messaging Communication Data** - WhatsApp Analyzer is a web-based tool that gathers and examines WhatsApp chat records on mobile devices. In order to access the conversation histories—which are encrypted owing to end-to-end encryption during transmission over the Internet and encrypted storage on the mobile device—it thus makes use of WhatsApp's e-mail export capability. Therefore, the primary benefit of the service is the ability to get actual communication data without the bias that comes from participant observation or surveys. Analysis of the gathered communication data yields insightful information on WhatsApp communication and the associated network traffic. Anonymizing all conversation data respects user privacy and encourages people to provide chat logs. Additionally, the sender of the chats can obtain some analyses of each conversation history on a webpage

**Pre-Processing and Emoji Classification of WhatsApp Chats for Sentiment** - With more than 2 billion users registered, WhatsApp is one of the most widely used social networking platforms. WhatsApp is a vital communication tool in people's lives. They text each other about their sentiments via WhatsApp. Since WhatsApp is available in more than 180 countries, code-switching is frequent. Emojis have also been more widely used and necessary during sentiment analysis as a result of this. This article discusses our method for transforming unstructured WhatsApp messages into a structured format that can be used to do sentiment analysis using a variety of data mining techniques. We go over how we handle code-mixing, the many emoji's and the feelings they represent, and how to use this method to do basic analysis.

**A Survey of Various Sentiment Analysis Techniques of Whatsapp** - Sentiment is a type of mental process, feelings, or assessment that results from a certain emotion. Opinion extraction is a method used to examine sentiment. This methodology is utilized to investigate the feelings of various populations based on detailed foundation. Through social media sites, data pertaining to opinions is gathered. Many people follow WhatsApp, which is one of the social networking sites. WhatsApp status updates are a type of message shared by these followers. WhatsApp data is analysed for sentiment using the feature extraction and classification approach. This study reviews and further evaluates a number of SA approaches.

#### 4. Problem Statement

With the use of a web application called WhatsApp Chat Analyser, users with or without technical expertise may examine their chat history. This allows you to obtain information on who is participating in the chat, for example. In general, when is the other person available to talk? Which group member is participating in the chat most actively? Which term does a certain group member use the most? You may obtain

several other insights from this analyser. Obviously, anyone with a WhatsApp account may utilize this web application. He or she can submit their text-based, media-free group or individual conversation to the Web-App after exporting it. It can handle WhatsApp chats in English.

#### 5. Objectives

This project is all about understanding different types of chats. It helps machine learning models by giving them good chat data to work with. To make these models work well, we need good examples. Our project dives deep into WhatsApp chats to analyse them thoroughly.

Exporting a WhatsApp chat as a text file doesn't take up much space or internet. The metrics we look at include how messages each user sends on average, the total number of messages, how many users there are, how users are active, who chats the most, word clouds, the busiest day, chat analysis by date, and more. The WhatsApp chat analyser focuses more on emoji's and words than other metrics. But it's just as fun as any other WhatsApp chat analyser.

#### 6. Methodology

##### *Module 1: Data collection*

Gaining information from messages exchanged on WhatsApp is the first step. You may accomplish this by using the WhatsApp app to export the chat history as a text file.

##### *Module 2: Data pre-processing*

Before being analysed for analysis, data must first be pre-processed after it has been gathered. This include doing things like deleting stop words, transforming the text to lowercase, and discarding extra letters.

##### *Module 3: Data Transformation*

This data frame will serve as the database, resulting in the raw string has to be poured into the fields for date, time, user messages, and more (using regular expressions) and organized tabularly (using Pandas).

##### *Module 4: Data Exploration*

In this instance, a classification model is trained by interpreting and analysing the gathered data. Following that, the model is assessed and serialized.

1. Top Statistics: Shows all messages, word count, and shared media links.
2. Monthly timeline: This graph displays the frequency of chats for a specific month.
3. The daily timeline: this graph shows the total amount of messages sent during the day.
4. Activity Map: shows you your busiest and least busy days of the month.
5. Weekly Activity Map: Displays the peak conversation times. Term cloud: the ubiquitous and often used term.
6. The Busiest Users
7. Emoji use analysis, including a list of the most popular ones.

##### *Module 5: Data Visualization*

Following analysis, the data will be shown on the screen as a graph, such as a line, bar, and heat graph and many other data



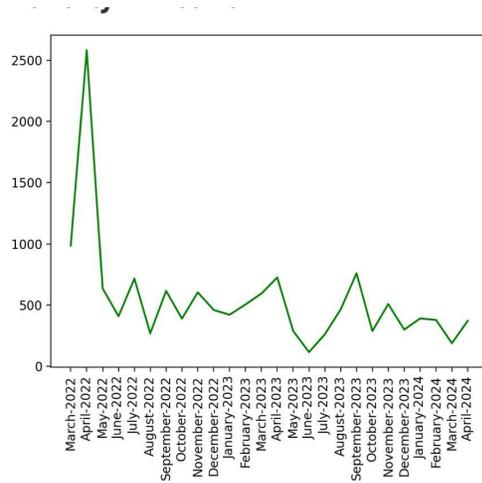


Fig. 4.

### 9. Future Work

Voice and Video Chat analysis: enlargement to include evaluation of voice and video chats, permitting insights into non-textual communication. Integration with other structures: Integration with different messaging systems past WhatsApp, which include fb Messenger, Telegram, or Slack, to provide a complete messaging evaluation solution.

### 10. Conclusion

In conclusion, it can be stated that the competencies of the WhatsApp application and the energy of the python programming language in imposing something records analysis intended, cannot be overemphasized. This work became to speak about the WhatsApp software and python libraries, to create an evaluation of a WhatsApp chat. We propose to hire dataset manipulation strategies to have a higher know-how of WhatsApp chat present in our telephones. It suggests maximum used emoji and word which again and again most instances. It tracks our communicate and analyzes how a good deal time we're spending. The system changed into executed with python, and the python libraries that have been carried out consists of, NumPy, Pandas, Matplotlib and Seaborn. In the end of the paintings anticipated outcomes were obtained and the analysis turned into in a position to reveal the level of participation of the various people on the given WhatsApp institution chat. On serious notice this device has the potential to investigate any WhatsApp chat enter into it.

### References

- [1] J. N. Singh, Y. Kumar, A. Srivastava, D. Baghel and K. Al-Attabi, "WhatsApp Chat Analyzer," *2024 1st International Conference on Advances in Computing, Communication and Networking (ICAC2N)*, Greater Noida, India, 2024, pp. 78-83.