iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** XII **Month of publication:** December 2022

DOI: https://doi.org/10.22214/ijraset.2022.47920

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 10 Issue XII Dec 2022- Available at www.ijraset.com

WhatsApp Chat Sentiment Analyzer

Aditya Jadhav[1], Suraj Patil[2], Mujtaba Shaikh[3], Prajwal Pal[4], Prof. Aparna Sawant[5], Prof. Pankaj Kunekar[6]
[1, 2, 3, 4, 5, 6]*Department of IT and Department of MCA, Vishwakarma Institute of Technology Pune, India*

Abstract: While being a widely used method of communication in today's times, WhatsApp is also headed towards being the platform where billions of people share their personal thoughts, emotions and sentiments, most of which is found in the group chats of this application. WhatsApp chat analyzer is a web application which was developed to provide analysis of WhatsApp group chats. Python libraries like matplotlib, re, seaborn, streamlit, pandas and some conceptual knowledge of NLTK are in use here. This is the combination of NLP, NLTK and machine learning. This system makes use of a WhatsApp chat file that is exported from the group chat or an individual user and then analyzes it to give us the analysis.
Keywords: Whatsapp chat, streamlit, matplotlib, machine learning, analysis, NLTK, sentiment, vader.

I. INTRODUCTION

In this paper, we have proposed a WhatsApp Chat Sentiment Analyzer. A WhatsApp conversation contains different types of communications held among the group participants and individual users. The exported chat file can be used to work with various machine learning and NLP models. These technologies provide the right learning experience. This application provides analysis of such data from exported WhatsApp chats. The key advantage of this system is that it is implemented using simple python libraries like seaborn, streamlit, numpy, matplotlib and pandas. These are commonly used for creating data frames and graphs.

Data preprocessing plays an important role in machine learning. We mainly focused on WhatsApp, one of Facebook's big data producers, as it would need a lot of data to make the model more efficient. WhatsApp has claimed that more than 50 billion messages are sent every day. An average user spends about more than 500 minutes a week on the WhatsApp application.

II. PROBLEM STATEMENT

The WhatsApp Chat Sentiment Analyzer is a statistical analysis tool for the WhatsApp chats. Working on exported chat files will help generate different plots of analysis. For example, which other group participants the user interacted with the most. In order to better understand WhatsApp chats on the phone, we propose to use record manipulation techniques.

III. PROPOSED SYSTEM

Much development has been done on the current whatsapp application. Older versions lacked status display, document sharing, and location sharing features. All these features are available in the current version. Older versions could not share images in document format. The system allows users to remotely access their WhatsApp on any web application via the QR code.

In the Whatsapp Chat Sentiment Analyzer, we have created a visualization dashboard which will show us different parameters extracted from the exported chat file.

In the initial stage, the exported chat is cleaned and formatted using numpy before processing on it. Further, with the use of pandas library a data frame is designed which is then used to analyze the data and create meaningful insights. We then use NLTK, specifically the Vader library to analyze the sentiments of the group chat or a specific individual and then visualize that data in statistical representation[3].

IV. FEASIBILITY STUDY

A. Technical Feasibility

1) *Python:* Python is a programming language that is most widely used for its support of various libraries. The following libraries of python are used like numpy, scipy pandas, csv, sklearn, matplotlib, sys, re, emoji, nltk seaborn, etc.
2) *Regex (Regular Expression):* A regular expression is a string that specifies a search pattern within text. Such patterns are typically used by "find" or "find and replace" operations on strings, or by string search algorithms for input validation. Regular expression techniques are developed in theoretical computer science and formal language theory [8].

©IJRASET: All Rights are Reserved

B. Operational Feasibility

1) *Matplotlib:* Matplotlib is an easy-to-use visualization library in Python which consists of multiple charts such as pie, line, bar, graph, scatter plot and histogram. This project uses Matplotlib for various visualizations such as bar charts, line charts, and pie charts are used [2].

2) *Seaborn:* Seaborn is a library primarily used for statistical plotting in Python[1].

3) *Streamlit:* In this project, we will use this library to create beautiful web elements and objects to represent our WhatsApp chat analytics using different kinds of charts and visualizations with Streamlit as a web application[6].

4) *Pandas:* Pandas is a Python library which is primarily used when working on data science along with machine learning. This library provides analytical tools for data manipulation to analyze data for time series analysis and numerical data.[2]

5) *VADER:* VADER stands for Valence Aware Dictionary for Sentiment Reasoning. It is a rule-based sentiment analyzer. It contains a list of lexical features which are generally labeled as per the semantic orientation as positive or negative It is included in the NLTK package and can be applied directly to unlabeled text data. VADER's SentimentIntensityAnalyzer() takes a string and returns a dictionary of three categories:

a) Positive
b) Negative
c) Neutral [5].

V. WORKING

1) Go to the sidebar and click on the 'Browse Files' button.

2) Select an exported WhatsApp chat file to upload for analysis. fig(a) shows the above.

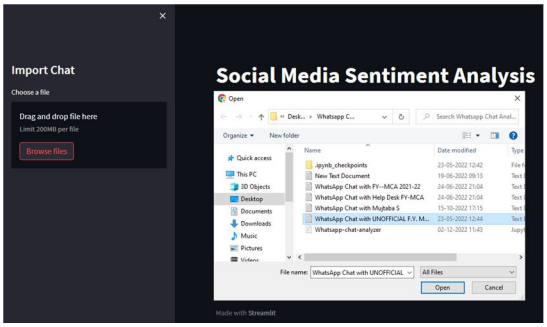


Fig (a). Select chat

3) The user can select for analysis of an entire group chat or for an individual user from the group.

4) After selecting, click on the 'View Analysis' button to analyze the uploaded file.

Fig (b). Imported File has been processed

5) The total number of messages, words, media and links shared within the group is displayed.

6) A daily and monthly timeline of messages using a line chart representation is displayed.

7) An activity map with a bar graph showing the busiest month and day is displayed next.

8) The top 5 busiest users in the group are displayed using representation of charts.

9) WordCloud shows visualization of the most common words used in the group chat.

10) List of frequently used emojis.

11) A pie chart showing the usage of the top five emojis.

VI. OUTPUT AND RESULTS

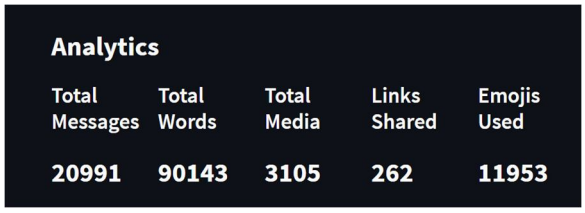


Fig 1. Analytics

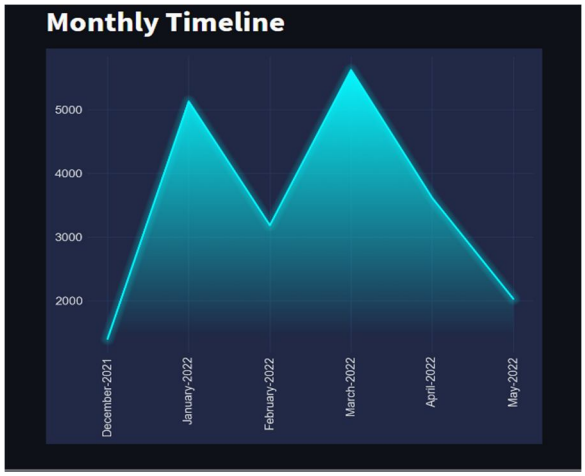


Fig 2. Monthly Timeline

Fig 3. Activity Map

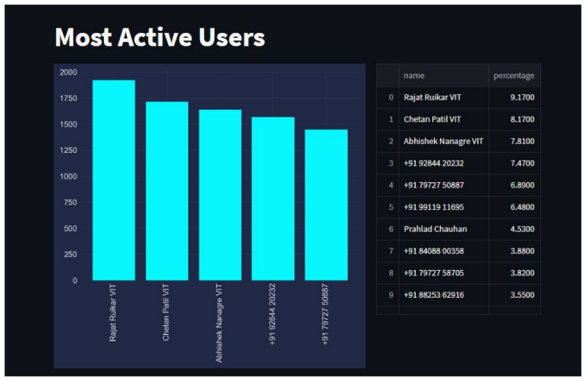


Fig 4. Most Active user

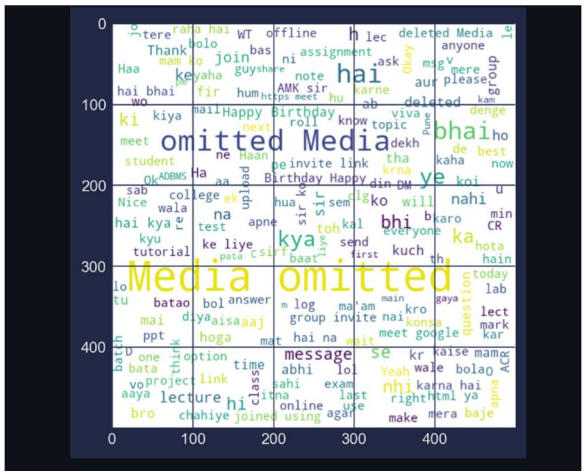


Fig 5. Word Cloud

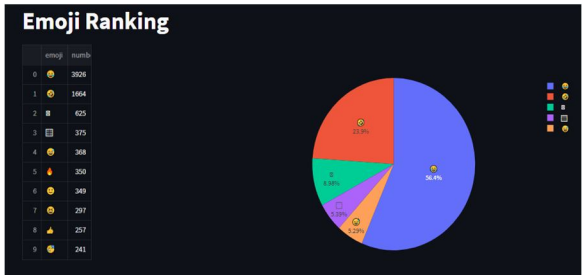


Fig 6. Emoji Ranking

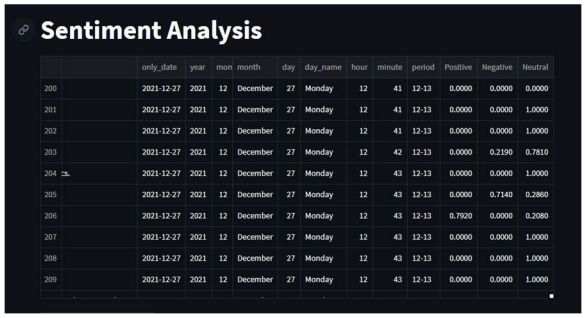


Fig 7. Sentiment scores for each message.

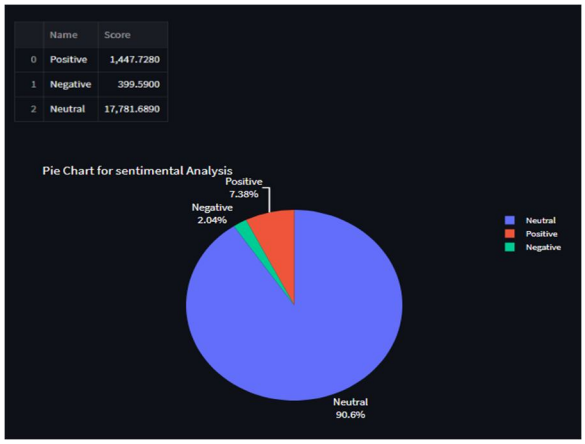


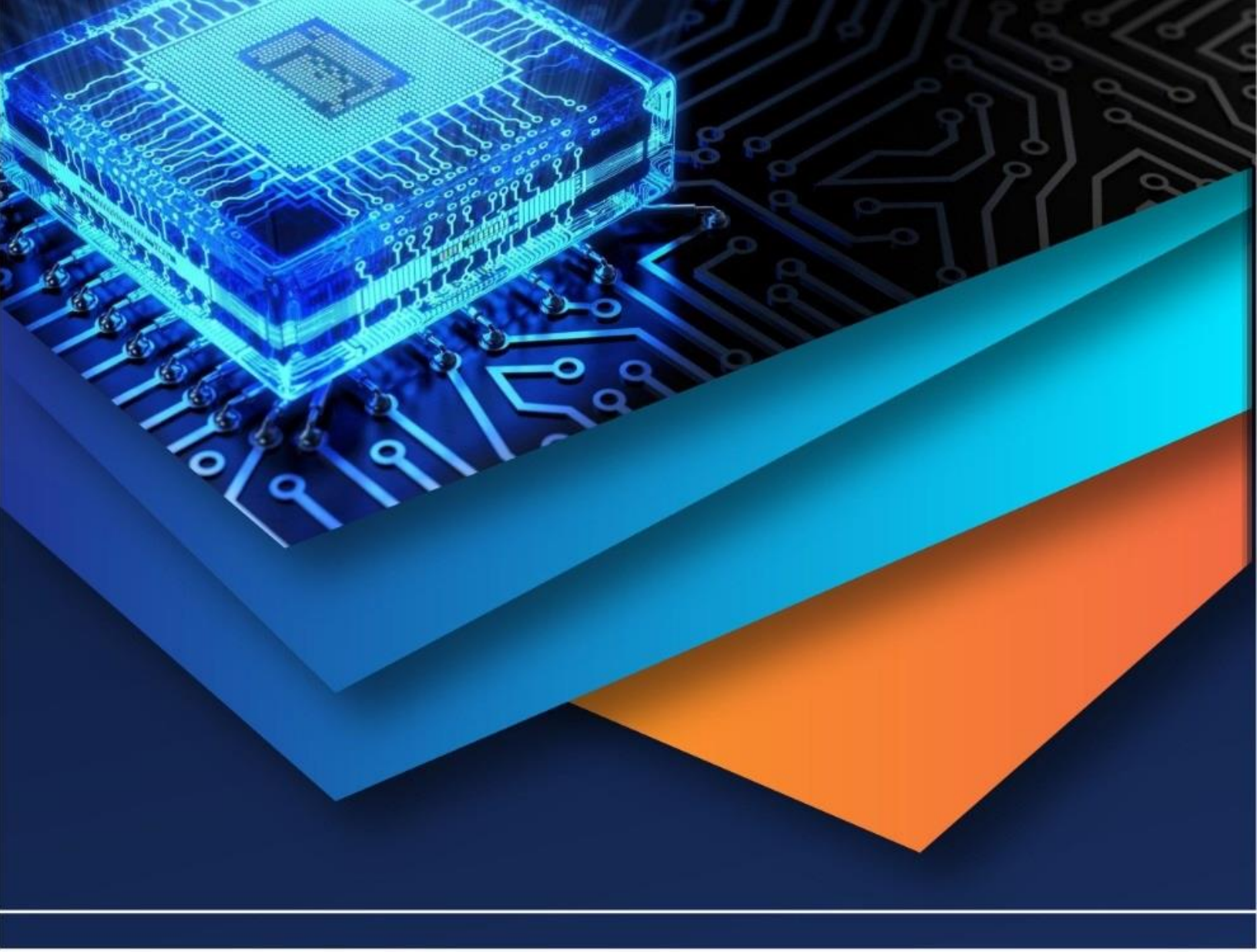
Fig 8. Pie-chart representing sentiment of the group

VII. CONCLUSION

The overarching goals defined in the early stages of requirements analysis are successfully achieved. After implementation, the system provides reliable results. The system is fully user friendly, making it easy for users with limited knowledge of the computer environment to operate the developed system. The system avoids the shortcomings of existing manual systems and completely eliminates the possibility of incorrect data entry due to the system's validation capabilities.

REFERENCES

[1] Marada Pallavi, Meesala Nirmala, Modugaparapu Sravani, Mohammad Shameem. WhatsApp Chat Analysis. International Research Journal of Modernization in Engineering Technology and Science. Volume: 04/Issue:05/May-2022
[2] Shaikh Mohd Saqib. Whatsapp Chat Analyzer. International Research Journal of Modernization in Engineering Technology and Science. Volume: 04/Issue:05/May-2022
[3] K, Ravishankara & Dhanush, & Vaisakh, & S, Srajan. (2020). Whatsapp Chat Analyzer. International Journal of Engineering Research and. V9.10.17577/IJERTV9IS050676.
[4] D.Radha, R. Jayaparvathy, D. Yamini, “Analysis on Social Media Addiction using Data Mining Technique”, International Journal of Computer Applications (0975 – 8887).
[5] https://towardsdatascience.com/sentimental-analysis-using-vader-a3415fef7664
[6] https://www.analyticsvidhya.com/blog/2021/06/build-web-app-instantly-for-machine-learning-using-streamlit/
[7] Meng Cai, “PubMed Central”, PMCID: PMC7944036, PMID: 33732917
[8] E. Larson, "[Research Paper] Automatic Checking of Regular Expressions," 2018 IEEE 18th International Working Conference on Source Code Analysis and Manipulation (SCAM), 2018, pp. 225-234, doi: 10.1109/SCAM.2018.00034.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)