

WHATSAPP CHAT ANALYZER USING MACHINE LEARNING

Chavan Ganesh Baban*, Abhishek Khodave, Shubham Kharade**, Saurabh Khartode**, Pradip Khude****

*(Assistant Professor, Information Technology, SVPM's College of Engineering Malegaon(bk), Baramati

Email: ganesh99222@gmail.com)

** (UG Students, Information Technology, SVPM's College of Engineering Malegaon(bk), Baramati)

Email: abhishekhattatraykhodave@gmail.com shubhamkharade625@gmail.com

saurabhkharode73@gmail.com

pradipkhude1912@gmail.com)

Abstract:

Nowadays, messages are the primary form of communication. Through WhatsApp, a lot of information has been shared. With over 650 million active users, WhatsApp is the most widely used communication application. Everyone has utilised it extensively, but mainly business people and young people. Users are able to examine the WhatsApp group chat or individual chat using a variety of analysis tools. Users genuinely want to examine their discussion for a variety of reasons. The goal of this study project is to do a flirt analysis and a time analysis.

Keywords — Sentimental Analysis, Data mining, Emotional, Natural Language Processing

1. INTRODUCTION

We live in a world where instant messaging has become a dominant mode of communication, used for personal, professional, and social purposes. Analysing WhatsApp chats can provide invaluable insights into the content, context, and sentiment of conversations, shedding light on communication dynamics, language usage, and even user behaviour.

The objectives of this paper are to provide a comprehensive overview of WhatsApp chat analysis and its potential applications, discuss the methodologies employed in analysing WhatsApp chats, and highlight the significance of the findings for various domains. We will delve into the data collection process, pre-processing techniques, NLP methods, and statistical analysis utilized to uncover insights from WhatsApp chats.

Furthermore, this paper will present the findings from WhatsApp chat analysis, including patterns and trends identified in the chat data, sentiment analysis results, key topics of discussion, and other meaningful insights. Data visualizations, such as word clouds, charts, and graphs, will be used to illustrate the findings and make them accessible to readers.

We will also discuss the implications of WhatsApp chat analysis findings, including how the insights can be applied in different domains such as marketing, customer relationship management, social sciences, and beyond. Understanding the dynamics of WhatsApp chats can offer businesses and researchers valuable information for improving communication strategies, enhancing customer engagement, and gaining deeper insights into human behaviour in the digital realm.

2. MOTIVATION

WhatsApp has become a staple of modern life for communicating with others and exchanging ideas. The current industrial aim is to create a better Chatbot that enhances the human experience. Understanding client responses so becomes a reasonable expectation. To accomplish this, the business chat must comprehend the customer's computational linguistics, context, and tone. A document's or a sentence's intention or attitude can be binary (positive, negative). Enabling sentiment analysis in WhatsApp chat assists with both advance study of services and malfunctions as well as fine-tuning responses in accordance with user mood. Every day, WhatsApp group conversation generates a lot of data. Each user transmits a sizable amount of text via the network. This information is used by the proposed chat analysis system to discover and forecast the user's emotion or mood at any given moment. By reading the user's chat messages, a person can accurately forecast the user's mood. We must translate this textual data into features that a computer can comprehend in order to allow a system to carry out a similar activity. After the features have been chosen, a training set is used to train the system's classifiers.

3. OBJECTIVE

The objective of WhatsApp chat analysis is to gain insights and understanding from the chat conversations that occur on the platform. This can involve various goals, such as understanding communication patterns among participants, analysing the sentiment or emotional tone of the

conversations, identifying key topics or themes, analysing user behaviour, identifying trends and patterns, and deriving meaningful insights and recommendations based on the findings.

4. APPLICATIONS

Applications of this work is to develop net based mostly WhatsApp chat analyser backed by powerful tools python and Django. Proposed system will helpful for:

- 1.Educational Counsellors
- 2.Psychologist
- 3.Business CRM(Customer Relationship Management)

5. LITERATURE SURVEY

Sentiment Analysis of WhatsApp Chats: A Comparative Study (2018): This research article presents a comparative study of sentiment analysis techniques applied to WhatsApp chats. It discusses various methods for sentiment analysis, such as lexicon-based, machine learning, and deep learning approaches, and evaluates their performance on WhatsApp chat data. This study can provide insights into the effectiveness of different sentiment analysis techniques for WhatsApp chat analysis.[1]

WhatsApp Chats Analysis for Understanding Users' Information Needs (2017): This paper focuses on analysing WhatsApp chat conversations to understand users' information needs. It presents a methodology for collecting and analysing WhatsApp chats to extract information needs, user queries, and other relevant insights. The findings of this study can shed light on the types of information exchanged in WhatsApp chats and can be useful for applications such as information retrieval and recommendation systems.[2]

Group Chat Analysis on WhatsApp: A Case Study of Information Sharing and Social Interaction (2019): This research article presents a case study of group chat analysis on WhatsApp, focusing on information sharing and social interaction. It discusses the patterns and dynamics of group conversations, including the frequency and type of messages exchanged, roles of group members, and communication patterns. This study can provide insights into group dynamics and communication patterns in WhatsApp chats.[3]

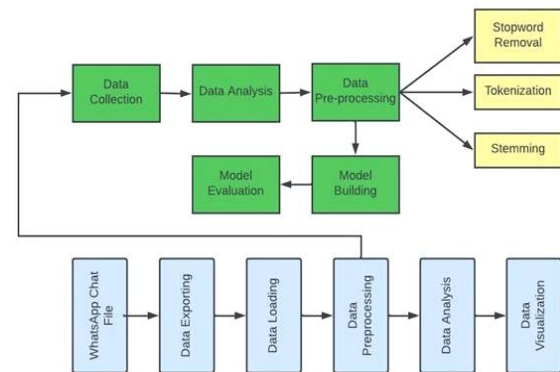
Visualizing WhatsApp Chats: An Exploratory Study (2018): This paper presents an exploratory study on visualizing WhatsApp chats to gain insights into the content and structure of conversations. It discusses various visualization techniques, such as word clouds, bar charts, and network graphs, and demonstrates how these visualizations can provide meaningful insights into WhatsApp chat conversations. This study can provide ideas for visualizing WhatsApp chat data in an informative and visually appealing manner.[4]

Analysing WhatsApp Chats: A Systematic Literature Review (2020): This systematic literature review provides an overview of existing research on WhatsApp chat analysis. It summarizes the methodologies, findings, and applications of WhatsApp chat analysis in various domains, including social

sciences, business intelligence, and customer relationship management. This review can provide a comprehensive understanding of the current state of knowledge and identify research gaps in the field of WhatsApp chat analysis. In conclusion, a literature survey for WhatsApp chat analysis can provide insights into the current state of knowledge, methodologies, findings, and applications in the field. It helps to identify existing research and publications, and can guide further research and applications in analysing WhatsApp chats for various purposes. [5]

6. ARCHITECTURE

The architecture for WhatsApp chat analysis typically involves several key steps, including data collection, data pre-processing, feature extraction, analysis, and visualization.



Data Collection: The first step in WhatsApp chat analysis is collecting the chat data. This can be done by obtaining the chat logs or chat export from WhatsApp, which typically come in the form of text files. The data may include chat messages, timestamps, sender and recipient information, and other relevant metadata.

Data Pre-processing: After collecting the chat data, the next step is to pre-process the data to clean and transform it into a suitable format for analysis. This may involve removing irrelevant information, handling missing data, normalizing timestamps, and converting the data into a structured format such as a data frame or a database.

Feature Extraction: Once the data is pre-processed, the next step is to extract relevant features from the chat data. This can include extracting text features such as words, sentences, or emojis, as well as metadata features such as message frequency, sender/receiver information, and timestamps. Feature extraction is an important step that determines the information that will be used for analysis.

Analysis: After feature extraction, various analysis techniques can be applied to gain insights from the WhatsApp chat data. This can include techniques such as natural language processing (NLP), sentiment analysis, topic modelling, network analysis, and statistical analysis, depending on the objectives of the analysis. These techniques can help uncover patterns, trends, and meaningful insights

from the chat data, such as sentiment trends, frequently discussed topics, communication patterns, and user behaviour.

Visualization: Visualization is an essential part of WhatsApp chat analysis, as it allows for easy interpretation and communication of the findings. Data visualizations, such as word clouds, bar charts, line charts, and network graphs, can be used to visually represent the results of the analysis. Visualization techniques can help to effectively communicate insights and make the findings accessible to a wider audience.

Finally, the results of the analysis are interpreted and conclusions are drawn based on the findings. The insights obtained from the analysis can be used to draw conclusions, make recommendations, and identify areas for further research or applications. The conclusions and recommendations can be summarized in the conclusion section of the analysis report.

Overall, the architecture for WhatsApp chat analysis involves data collection, data pre-processing, feature extraction, analysis, visualization, interpretation, and conclusion. It is a multi-step process that involves applying various techniques and methods to extract meaningful insights from the WhatsApp chat data and draw conclusions based on the findings.

7. ALGORITHM

Naive Bayes: The Bayes Theorem is the foundation of the Naive Bayes algorithm, a contingency machine learning technique utilized in a wide range of classification applications. Nave Bayes is a fantastic illustration of how the most straightforward solutions are frequently the most crucial. Despite recent breakthroughs in machine learning, it has proven to be not just quick, accurate, and dependable but also easy. It has been successfully employed for a variety of purposes, but language process (NLP) difficulties are one area where it shines. The mathematician's theorem may enable the probabilistic machine learning algorithm Nave Bayes, which is often used for classification problems. This essay will help us understand the Naive Bayes algorithm and other crucial ideas, eliminating any room for uncertainty. For a hypothesis with two occurrences A and B, the MAP is

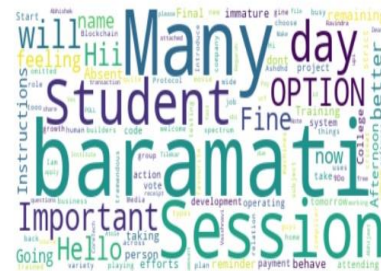
$$\begin{aligned} \text{MAP}(A) &= \max(P(A - B)) \\ &= \max(P(B - A) * P(A) / P(B)) \\ &= \max(P(B - A) * P(A)) \end{aligned}$$

8. RESULT

The results of WhatsApp chat analysis can vary depending on the specific techniques and methods used, as well as the objectives and data available for analysis. Here are some potential results that can be obtained from WhatsApp chat analysis:

Communication patterns: Analysis of the chat conversations may reveal communication patterns, such as the frequency and type of interactions among participants, the most active participants, and the flow of information within the conversations. This can provide insights into how participants engage with each other and how information is shared.

Sentiment analysis: Sentiment analysis can reveal the overall sentiment or emotional tone of the chat conversations. It may identify the proportion of positive, negative, and neutral messages, as well as sentiment trends over time or among different participants. This can provide insights into the emotions expressed in the conversations and the overall sentiment of the discussions.



1. Group Chatting Stats

1. Total Number of Messages
2. Total Number of Media Messages
3. Total Number of Links
4. Total Number of Emojis

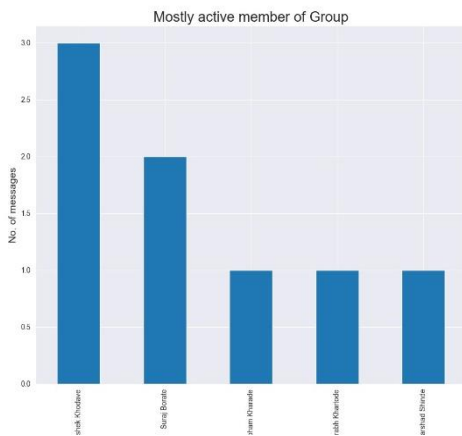
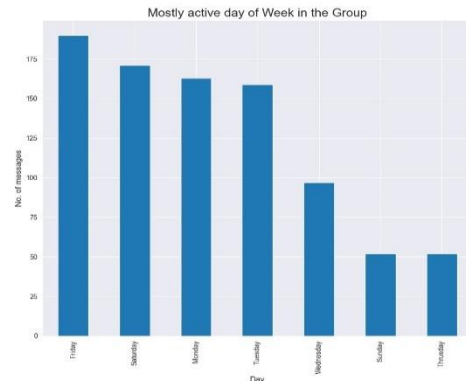
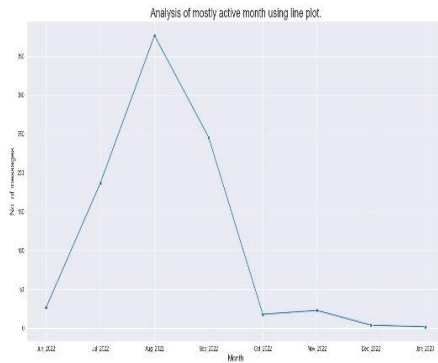
2. Word Cloud

3. Graph

- Top 5 Active Members
- Active Day of Week
- Active Months Line Plot
- Active Times of Day

4. Sentiment Analysis

Topic identification: Topic modelling techniques can help identify and extract key topics or themes discussed in the chat conversations. This can provide insights into the main subjects of the conversations, the frequency of different topics, and the distribution of topics among different participants. Topic identification can help identify common interests, concerns, or trends in the chat conversations.



Trends and patterns: Analysis of trends and patterns in the chat conversations may uncover frequently used words or phrases, emerging trends, and shifting dynamics within the conversations. This can provide insights into changing conversation topics, trends in communication patterns, and evolving dynamics among participants.

Insights and recommendations: The results of WhatsApp chat analysis can generate meaningful insights and recommendations based on the findings. These insights may guide improvements in communication strategies, identify areas of concern or improvement, and suggest strategies for better engagement or user experience.

It's important to note that the results of WhatsApp chat analysis are highly dependent on the data available for analysis, the methods and techniques used, and the specific objectives of the analysis. Careful interpretation and analysis of the results are necessary to draw meaningful conclusions and make relevant recommendations

9. CONCLUSION

Analysing WhatsApp chat conversations, organizations and individuals can better understand how participants engage with each other, identify common interests or concerns, uncover sentiment trends, understand user behaviour, and identify emerging topics or trends. These insights can inform communication strategies, improve user experience, and guide decision-making in various contexts, such as business, customer service, research, and social interactions.

However, it's important to note that WhatsApp chat analysis should be conducted with careful consideration of privacy and ethical considerations, ensuring compliance with relevant laws and regulations. Interpretation of the results should be done cautiously, taking into account the limitations of the analysis methods, the context of the data, and the specific objectives of the analysis.

In summary, WhatsApp chat analysis can provide valuable insights and understanding from the chat conversations on the platform, and when conducted appropriately, can yield meaningful results that can

10. ACKNOWLEDGMENT

We'd like to express our gratefulness to Prof. Chavan G.B., our BE dissertation supervisor, whose valuable suggestions and support greatly contributed to the completion of this paper. We would also like to thank Prof. Gawade J.S., Head of Department, and Honourable Principal Prof. Dr Mukane S.M. for providing us with the opportunity and resources to undertake this project.

11. REFERENCES

- [1] V.Selina Annie Retna, P. Brundha; G RajKumar; "People's Behaviour Analysis of Chat Message using Natural Language Processing". IEEE ICICV 2021
- [2] MeylanWongkar, ApriandyAngdresey; "Sentiment Analysis Using Naive Bayes Algorithm of the Data Crawler: Twitter", 2019 IEEE.
- [3] Sonika Dahiya, Astha Mohta, Atishay Jain; "Text Classification based Behavioural Analysis of WhatsApp Chats", ICCES 2020.
- [4] Aditi Saluja, Sonika Dahiya; "Pre-Processing and Emoji Classification of WhatsApp Chats for Sentiment Analysis". I-SMAC 2020.
- [5] Jishnu Sri Ojas Wy Akella, LN Yashaswy Akella; "Sentiment Analysis Using Naïve Bayes Algorithm: With Case Study", ICICT 2018.
- [6] MADHURA JAYARATNE, BUDDHI JAYATILLEKE; "Predicting Personality Using Answers to Open-Ended Interview Questions", IEEE June 22, 2020.
- [7] Rick Cents, Nhien-A Le-Khac; "Towards A new Approach to Identify WhatsApp Messages", IEEE 2020.
- [8] Shikha Tiwari, Anshika Verma; "Social Media Sentiment Analysis on Twitter Datasets". (ICACCS) 2020
- [9] R Raja Subramanian, Nukala Akshith, Srikar Amara; "A Survey on Sentiment Analysis". IEEE 2021