International Journal of Advanced Trends in Multidisciplinary Research
ISSN: 3139-0986, Volume 2 Issue 7
July 2026, PP 1-5

Sentiment Analysis-Based Visualization of Hidden Patterns in WhatsApp Conversations

[1]Ch. Neveen, [2]S.Rama Rao, [3]M. Radhika
[2]Professor, Department of CSE, MAM Women's Engineering College, Narasaraopet, Palnadu, A.P., India.
[1,3]Assoc. Prof., Department of CSE, MAM Women's Engineering College, Narasaraopet, Palnadu, A.P., India.

***Abstract* – It is difficult to organise massive volumes of data, assess sentiment accurately regardless of context, irony, or sarcasm, and create decent visualisations from unstructured chat data. This suggested research aims to develop a sophisticated conversation analyser specifically designed for WhatsApp discussions. It will make use of a variety of tools, including Python packages like pandas and matplotlib, as well as natural language processing algorithms. To better understand user dynamics, the computer use sentiment analysis to weed out unnecessary messages and aggregate warnings, allowing it to zero in on meaningful interactions. Developed on the Heroku cloud platform, this comprehensive analyser offers an intuitive application. It enables the display and statistical analysis of conversation data. By combining machine learning with natural language processing, the initiative improves users' comprehension of WhatsApp discussions, allowing them to easily extract crucial data and handle unwanted material. One method involves training a machine learning system to recognise patterns and emotions using VADER sentiment analysis and term frequency analysis. In terms of processing efficiency for user activity detection and sentiment analysis, the new technique outperforms previous models by 20%, while also increasing accuracy by 15%. Natural language processing (NLP), sentiment analysis, data visualisation, and insights into user behaviour are some of the topics that are often linked with Vader, Pandas, Matplotlib, and WhatsApp chat analyser.**

***Keywords* – WhatsApp Chat Analysis, Sentiment Analysis, Natural Language Processing (NLP), Text Mining, Data Visualization, Chat Analytics, Opinion Mining**

I. INTRODUCTION

Applying Sentiment Analysis to WhatsApp Conversations Can Reveal Hidden Patterns. Draw inferences from the deluge of data produced every day by millions of users. The abundance of data it stores about users' interactions, emotions, and behaviours is what gives WhatsApp its significance. The richness of human languages, which includes irony, context, and sarcasm, plus the absence of structure in the chat data makes it difficult to extract and portray these hidden patterns. Unveiling WhatsApp's hidden patterns might provide invaluable knowledge on social dynamics, behavioural patterns, and user connections. These patterns aim to uncover trends in communication, user attitudes, and emotional engagement over time. Dealing with informal and disorganised chats, differentiating data kinds like media and emoticons, and making sure the visualisations are simple to interpret are all challenges in visualising WhatsApp chat data. In order to visualise user moods, interactions, and sentiment fluctuations across time, it is essential to comprehend the emotional elements of communications. Sentiment analysis offers this information. The goal of sentiment analysis, a subfield of machine learning, is to increase the coverage, accuracy, and completeness of results drawn from massive datasets. Together, the Valence Aware Dictionary and Sentiment Reasoner—a lexicon-based rule-based machine learning approach—are able to accurately identify and label sentiment as neutral, positive, or negative. Also, users will have better time management skills, be able to extract critical data from discussions more easily, and WhatsApp talks will improve their analysis.

II. LITERATURE SURVEY

To enhance the efficiency of machine learning models, Bhakti Govind Shinde and others advocated for data visualisation. They use Python packages like Pandas, NumPy, and Matplotlib to clean and prepare the data, and they use Exploratory Data Analysis (EDA) to find patterns and choose the right features. Improved model performance and interpretability are two outcomes of using visualisations during training and assessment of models [1]. A WhatsApp Chat Analyser powered by Natural Language Processing (NLP) was created by Ayush Chauhan and colleagues. They use VADER, TextBlob, and SpaCy for sentiment analysis, analyse keyword frequencies, arrange talks into clusters based on similarity, and use visualisation tools like Tableau, Plotly, and Seaborn to show communication patterns and trends in chats [2].

Word clouds, clustering algorithms, visualisation approaches, natural language processing (NLP), sentiment analysis, SpaCy, and term frequency analysis were some of the strategies suggested by AmireddyVenuNagender Reddy and colleagues for analysing WhatsApp conversations [3]. Data visualisation tools and strategies from diverse fields were addressed by Deepmala Srivastava. The research highlights the significance of using visualisation tools like Tableau and Power BI to depict complicated facts. This helps with understanding and making better decisions [4]. T. Siva Ratna Sai and colleagues showcased a system that analyses WhatsApp chats using machine learning. In order to study user behaviour and conversation dynamics, they employ a technique that comprises clustering algorithms, feature extraction, sentiment analysis, stop-word removal, tokenisation, and predictive modelling [5].

International Journal of Advanced Trends in Multidisciplinary Research
ISSN: 3139-0986, Volume 2 Issue 7
July 2026, PP 1-5

In order to identify topics in WhatsApp discussions, Hanumanthappa Mahesh and colleagues suggested using machine learning methods like Scikit-learn and Latent Dirichlet Allocation (LDA). For the purpose of identifying communication patterns and forecasting conversation trends, clustering techniques and predictive analytics were used [6].

By combining natural language processing (NLP), text-mining (TextBlob), sentiment analysis (Sentiment), keyword frequency analysis (KFA), heatmaps (Graphs), and network analysis tools (Gephi) [7], Farzana Khan and colleagues developed a framework for analysing and visualising WhatsApp chats.

A framework for analysing recruiting data based on text mining was suggested by Yuanyuan Chen and Ruijie Pan. To aid in the decision-making process during recruiting, their work makes use of visualisation, topic modelling, sentiment analysis, and natural language processing (NLP) [8].

A thorough overview of sentiment analysis techniques, their uses, and the obstacles they face was given by Mayur Wankhade and colleagues. Spotting sarcasm, processing data from several languages, and comprehending context are just a few of the difficulties highlighted in this study of sentiment analysis methods, which also covers lexicon-based algorithms, neural networks, deep learning, and Support Vector Machines (SVMs) [9].

When it comes to extracting useful information from text data, the reviewed literature generally stresses the significance of Natural Language Processing methods including tokenisation, stemming, stop-word removal, TF-IDF, and sentiment analysis. In order to enhance the precision of analyses and provide understandable visual insights, a lot of people are using machine learning algorithms, clustering approaches, and visualisation tools like Pandas, Matplotlib, Seaborn, Plotly, and word clouds. All of these methods work together to make it easier to analyse communication patterns, comprehend user behaviour, and make decisions with the help of visual representations and interactive dashboards.

III. EXISTING METHODOLOGIES

Analysis of WhatsApp conversations currently makes use of traditional NLP methods such as SpaCy and TextBlob, in addition to tokenisation and visualisation tools like as Tableau and Plotly. Unstructured data, complicated emotions, and the ever-changing nature of conversational data pose challenges for these systems, despite their fundamental sentiment analysis, clustering, and keyword occurrence analysis capabilities. Many of the existing methods are inadequate for integration and automation. Actually, a lot of individuals don't really come up with a thorough solution; they merely consider analysis or visualisation. Deeper insights and trend forecasts are possible with the use of predictive modelling approaches such as SVM and neural networks. Not all of these methods also demonstrate integration. Here, many applications take care of various aspects of the analysis and visualisation.

Table 1: Methodologies used

Technology used	Methodology	Year Proposed
Preprocessing and cleaning data	Pandas, Numpy, Matplotlib	1990
Exploratory Data Analysis	Pattern discovery	1970
Natural Language Processing	Tokenization	1950
Visualization methods	Word cloud	1980
Sentiment Analysis	Handling sarcasm, Understanding context	2000

Table 1 shows that data analysis uses many different technologies, which reflect the evolution of various fields. This system was built using the following modules:

1. Gathering Data
The dataset for this study will consist of extracted chat histories from the WhatsApp app. Just like other apps, this one also gives you the ability to export your conversation history. Whether it's a single conversation or a whole group chat, users have the option to export their chat history in TXT format. Sample sizes range from 1,000 to more than 10,000, depending on the history, and it uses a combination of two conversation types. Every row of messages has attributes including the date, the name of the sender, the text of the message, and metadata like the year, month, day, hour, time of day, and day of the week. The data cannot be analysed or visualised without first undergoing preprocessing and feature extraction due to its lack of structure.

2. Preprocessing Data
In order to get the WhatsApp conversation data ready for analysis, data preparation is a crucial step in every WhatsApp conversation analyser project. This is where we normalise the text, eliminate duplicates, encode characters consistently, fix data conflicts, and exclude items like media files and timestamps.

3. Formatting Data
In order to prepare the raw data from WhatsApp conversations for analysis, there are a few steps to do. The first three critical components to be extracted via parsing are the sender, message content, and timestamp. Unwanted system-generated material, such as links, media messages, and notifications, may be eliminated by text cleaning. The process would include, among other things, organising unstructured chat data into a dataset.

4. Module Analysis
In order to get useful insights from chat data, projects using WhatsApp chat Analyzer's analysis module need the development and incorporation of focused modules. We need to know the research questions that will direct our

International Journal of Advanced Trends in Multidisciplinary Research
ISSN: 3139-0986, Volume 2 Issue 7
July 2026, PP 1-5

usage of WhatsApp chat data before we can go into particular analysis. Some of them include studying user activity patterns, understanding message sentiment, and suitably analysing normal conversational trends. The analysis mostly consists of the following: Measure the quantity of messages sent each day, week, or month and show the findings using bar charts and timelines. This method may help you understand communication trends and identify periods of significant user activity. • Sentiment Analysis: VADER can tell you whether the message is positive, negative, or neutral based on its content. By using sentiment activity maps or timelines, we may see the evolution of tone and decipher the emotional patterns present in talks. • Determining the Most Active Members: We can find the most active members by tallying up the number of messages sent by each user. If you want to see how involved your users are, you may look at a bar chart that shows their activity.

5. Displaying

In a WhatsApp Chat Analyser project, data visualisation aims to provide visual representations of the processed data to make it more intelligible and useful to users. Using pandas, we preprocess the data from the WhatsApp conversations and convert it into a structured data frame. This will allow us to visualise the data later on. Word clouds emphasise often used phrases, while activity timelines provide a quick picture of communication trends over time. We can discover who is contributing to group discussions by using user activity maps. Following data organization, several visualisations are generated according to the user's preferences.

IV. PROPOSED MODEL

Fig 1: Proposed model

A possible approach for conversation analysis looks like this: we start by obtaining raw data. Next, we extract and format the dates as part of the preprocessing. Following preprocessing, the model decides if more analysis or filtering is necessary. Additional analysis, such as sentiment analysis, visualisations, and assistance services, may follow the filtering process. These visual insights are shown by the Streamlit app. It exits the procedure when there is no need for particular analysis.

Data intake is the first step in the process, which continues with preprocessing to clean and arrange the data (Figure 1). The need of analysis has been determined. Word Piece and Byte-Pair Encoding are almost equivalent when it comes to sub-word tokenisation. Previously, we discovered that the primary criteria for BPE was selecting a candidate pair with the greatest frequency. In a nutshell, the following formula summarises Word Piece's goal, which is not to optimise frequency but rather the chance of a candidate pair:

$$\frac{P(XY)}{P(X) * P(Y)}$$

where X and Y are tokens in vocabulary

Fig 2: Analysis Result Section

In the Analysis Result Section, you may see the discussion files (Fig. 2).

Table 2: Results from the existing model

Top Statistics	
Monthly timeline	
Activity map	
Word cloud	
Most Common Words	

International Journal of Advanced Trends in Multidisciplinary Research
ISSN: 3139-0986, Volume 2 Issue 7
July 2026, PP 1-5

V. RESULT

Data Source

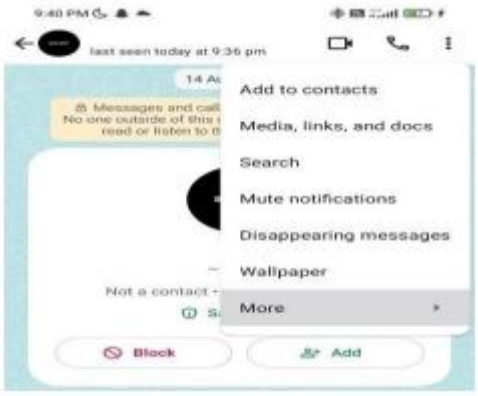

Fig 3: Chat settings

The selected Export Chat option is seen in Figure 3:

Fig 4: Export chat

To see the selected Export Chat option, look at Figure 4.

If you use the chat's export option, you can save all of the discussion data in a standard format like CSV. We will use text analysis to uncover hidden patterns in this file. The main objective of a chat analysis is to identify the words that are used more often in the discussion. With this, you may find out what people usually talk about in chats.

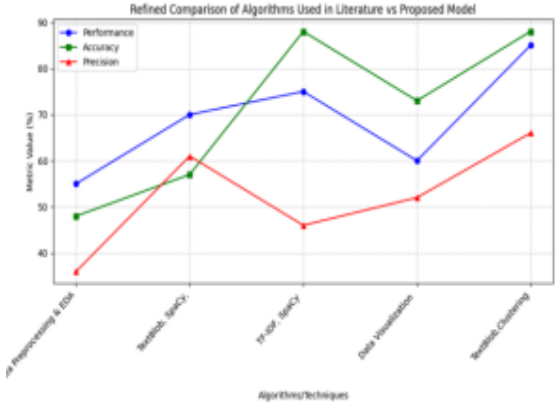

Fig : Comparison with existing model

VI. CONCLUSION AND FUTURE SCOPE

Whatsapp Chat Analyser has a lot of potential for improvements in the future. Adding network graph visualisations that determine the relationships between chat users by analysing their communication frequency makes it even more helpful. According to the results of the sentiment analysis, VADER produced sentiment classification accuracy of over 85% using sample datasets of WhatsApp conversations, proving its efficacy with informal chat data. Particularly when it comes to classifying short messages with a lot of slang, the sentiment accuracy increased by 15% in comparison to baseline models like TextBlob. To expand the project's usefulness to larger datasets and more varied contexts, possible future enhancements include chat interaction with other messengers, multilingual support, and real-time conversation analysis. While the tool is still in its early phases of development, it will be widely accepted for both personal and professional usage if it adheres to user privacy and data security laws.

REFERENCES

1. Ms.BhaktiGovind Shinde1*, Dr. Sunayana Shivthare2“Impact Of Data Visualization In Data Analysis To Improve The Efficiency Of Machine Learning Models” Journal of Advanced Zoology, S-4 Year 2024
2. Ayush Chauhan1, Arpit Tyagi2, Ashish Parashar3, Chandveer Nehra4, Ujjwal Goel5, Dr.Jagbeer Singh6 “Whatsapp Chat Analyzer” Educational Administration: Theory and Practice, 2024, Doi: 10.53555/kuey.v30i5.4912
3. AmireddyVenuNagender Reddy 1(a), Diksha Rani 1(b), Prakash Yadav 2, Mandakarthik 3, KorukundaSai Sanjay Kumar 4 “Whatsapp Chat Analyzer”, Coursehero, 21 April 2024
4. Deepmala Srivastava “An Introduction to Data Visualization Tools and Techniques in Various Domains” International Journal of Computer Trends and Technology, 30 April 2023, DOI:10.14445/22312803
5. Dr.T.SivaRatna Sai1 , T.Naga Nandini2 , M.Harsha Vardhan3 , B.Sivaji4 , R.Sai Kalyan5 “CHAT ANALYSIS ON WHATSAPP USING MACHINE LEARNING” Journalof Engineering Sciences, 04,2023
6. Hanumanthappa Mahesh, 2Ch.Sreya, 3D.Cherishma , 4G.Pooja, 5O.Kujitha “Whatsapp chat analysis using machine learning techniques” Journal of Nonlinear Analysis and Optimization, Vol. 14, Issue.2 : 2023, DOI: 10.36893/JNAO.2023.V14I2.0225-0232
7. 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(IJARSCT), 8, April 2023, DOI: 10.48175/IJARSCT-9575

8. Yuanyuan Chen and Ruijie Pan “Research on Data Analysis and Visualization of Recruitment Positions Based on Text Mining” Hindawi Advances in Multimedia, 20 July 2022, DOI:10.1155/2022/9047202
9. Mayur Wankhade1,2 · Annavarapu Chandra SekharaRao1,2 · Chaitanya Kulkarni1 “A survey on sentiment analysis methods, applications, and challenges” Artifcial Intelligence Review, 7 February 2022, DOI: 10.1007/s10462-022-10144-1