International Journal of Research Publication and Reviews

NLP-Based Mental Health Analysis Using WhatsApp Chat

Dr. Manoj Kumar[1], Shreshth Aggarwal[2]

[1]Assistant Professor, Maharaja Surajmal Institute, New Delhi, India
[2]Student, Maharaja Surajmal Institute, New Delhi, India
 [1] manoj.rke77@gmail.com, [2] shreshth76@gmail.com

ABSTRACT

WhatsApp is a popular social media platform which is widely used all over the world to communicate and connect as it allows people to discuss in groups. The discussion logs of WhatsApp can reveal information about user behavior which also tells about the Mental Health of a person. Analysis of the sentiments can help the users in decision making and in the diagnosis the problem of mental health all over the world. This research study is a data analysis initiative designed to uncover patterns in communication, user interactions, and engagement within WhatsApp conversations. By leveraging natural language processing (NLP) and data visualization techniques such as charts, graphs, and word clouds present the data insights in a clear and accessible format. The study aims to evaluate the content, sentiment, and present the data insights in a clear and accessible format. Future sentiment analysis systems aim to detect a wider range of emotions and moods, incorporating contextual understanding and cultural nuances for more nuanced interpretations. These advancements can enhance behavioral analysis, and mental health assessment. In healthcare, sentiment analysis could aid early detection of mental health issues by identifying patterns in expressions and language usage. The analysis of the WhatsApp conversations using NLP and data visualization provides insights into user sentiment, engagement trends, and communication patterns. It supports mental health by identifying early concerns and promoting supportive digital interactions.

Keywords – Analyzer, Chats , Mental Health, Sentiments, Decision Making, Natural Language Processing (NLP)

1. INTRODUCTION

Technology is deeply ingrained across various aspects of life, including social, political, and educational domains. [1]. WhatsApp is the most popular messaging apps globally having more than 2 billion users worldwide [2]. People use WhatsApp for daily personal communication, including text messaging, voice, and video calls. It is favored for its ease of use, reliability, and features like end-to-end encryption, which ensures privacy [3]. In the education sector, WhatsApp is used to facilitate group discussions, share educational materials, and support remote learning. Its accessibility and user-friendly interface make it a valuable tool for teachers and students alike [4]. The usage of WhatsApp has surpassed the other social media apps like WeChat, Facebook Messenger, QQ, SnapChat, and Telegram [5].

The advent of the internet and its associated technologies has revolutionized information exchange. Platforms like WhatsApp, Twitter, Facebook, and blogs have become vital spaces for users to share valuable opinions on various topics. This creates a need to analyze these views and sentiments to determine the relevance of the shared information or to predict the likelihood of future events. Sentiment analysis, often referred to as opinion mining, is the process of detecting and extracting emotions or opinions from textual data. This computational approach is widely used to interpret sentiments expressed in written content and has become valuable in areas such as marketing, customer service, and social media analysis. As a branch of natural language processing, opinion mining aims to gauge public opinion on specific topics or events [6]. Techniques employed in sentiment analysis include machine learning algorithms, deep learning frameworks, and lexicon-based methods [7].

The status of mental health in India has been a growing concern in recent years. While mental health issues have always existed, there's been a significant increase in awareness and recognition of these issues. The status of mental health in India is a complex issue influenced by various socio-economic, cultural, and healthcare factors. While there have been strides in awareness and initiatives to address mental health challenges, significant gaps and challenges persist. The impact of WhatsApp on mental health in India is a multifaceted issue influenced by various factors such as social dynamics, communication patterns, and technological usage. While WhatsApp has become an integral part of daily communication for millions of Indians, its influence on mental health has raised several concerns including following.

1) Information Overload and Stress: The constant flow of messages, notifications, and information on WhatsApp can contribute to information overload and digital stress. The pressure to respond promptly to messages and stay updated on group conversations can increase anxiety and overwhelm individuals, negatively impacting their mental health [8].

2) Misinformation and Rumor Propagation: WhatsApp has been implicated in the spread of misinformation, fake news, and rumors, which can have detrimental effects on mental health. False information about health, politics, and social issues circulated on WhatsApp can lead to confusion, fear, and anxiety among users, undermining their trust in information sources and exacerbating mental distress [9].

3) Cyberbullying and Harassment: WhatsApp groups and chats can become platforms for cyberbullying, harassment, and online abuse. Cases of cyberbullying through derogatory messages, hate speech, and offensive content shared on WhatsApp have been reported, leading to psychological distress and emotional trauma for the victims [10].

4) Privacy Concerns and Digital Addiction: The privacy implications of WhatsApp, including data privacy breaches and concerns about surveillance, can contribute to feelings of insecurity and stress among users. Additionally, excessive use of WhatsApp and other social media platforms can lead to digital addiction, with individuals spending excessive amounts of time online, neglecting real-life interactions, and experiencing withdrawal symptoms when offline [11].

Through this study, it is aimed to empower individuals by providing insights into their chat patterns, enabling them to understand their behavior's impact on mental health. By analyzing communication dynamics on platforms like WhatsApp, users can gain valuable insights that help them make informed decisions to improve their well-being. This initiative not only fosters self-awareness but also promotes proactive steps towards healthier digital interactions. With the data-driven understanding of their communication habits, individuals can identify areas for improvement or modification, ultimately enhancing their mental health and overall quality of life. By leveraging natural language processing (NLP) and data visualization techniques such as charts, graphs, and word clouds present the data insights in a clear and accessible format, the project aims to evaluate the content, sentiment, and present the data insights in a clear and accessible format. Future sentiment analysis systems aim to detect a wider range of emotions and moods, incorporating contextual understanding and cultural nuances for more nuanced interpretations.

Remaining sections of the paper is arranged as follows. Section 2 surveys the related works on this topic. Section 3 details the methodology used, Section 4 contains the results and discussion. Section 5 concludes the paper with some recommendations for future work.

2 RELATED WORKS

[12] explores the usage patterns and benefits of WhatsApp in small and medium enterprises, providing insights into chat analysis within a business context. Another study [13] compares the behaviors and usage patterns between WhatsApp and traditional SMS, offering valuable data for understanding group dynamics in chats. The comprehensive text [14] covers various techniques and applications of sentiment analysis, providing foundational knowledge that can be applied to chat analysis. Although focused on Twitter, the techniques discussed in [15] for sentiment analysis in social media can be adapted to WhatsApp chat data. Another study [16] provides insights into feature-based sentiment analysis which can be applied to WhatsApp chat analysis to extract sentiments from customer reviews or feedback.

pSenti, a concept level sentiment analysis system combine lexicon-based and machine learning-based techniques in opinion mining. System demonstrates the improved accuracy classifying sentiment polarity and detecting sentiment intensity compared to purely lexicon-based systems. Testing on two real-world datasets—CNET software reviews and IMDB movie reviews—showed that pSenti outperformed other hybrid approaches, surpassing advanced systems like Senti Strength [17].

3 METHODOLOGY

The development of the WhatsApp chat analyzer incorporates a meticulously crafted methodology that merges Agile and Lean practices, ensuring iterative progress and responsiveness to evolving user needs. Agile Development principles drive the iterative development and feedback integration process. By structuring the project into sprints, the focus remains on specific analysis components like user engagement metrics, sentiment analysis, and language trend identification. This approach facilitates adaptability and a user-centric focus, as flexibility is prioritized to adjust functionalities based on continuous user feedback.

Collaboration and communication play pivotal roles in Agile Development. Cross-disciplinary teams comprising data analysts, user experience designers, and developers collaborate to enrich analysis and solution ideation. Regular review meetings ensure alignment with project objectives and immediate issue resolution. Lean Methodology complements Agile practices by emphasizing value stream mapping and Minimal Viable Product (MVP) development. Streamlined data processing ensures a smooth flow from data extraction to visualization, enabling prompt generation of insightful analytics. The MVP approach prioritizes essential features for initial functional product release, allowing for subsequent incremental enhancements based on user input.

The system implementation leverages Python as the primary programming language, predominantly used for backend development, data processing, and core functionalities of WhatsInsight. Frameworks and libraries such as NLTK (Natural Language Toolkit) are employed for sentiment analysis, word frequency trends, emoji analysis, and tokenization. Streamlit serves as the primary web application framework for creating and deploying the user interface, ensuring a seamless user experience. Pandas is utilized for efficient data manipulation and analysis tasks, while Matplotlib and Seaborn are employed for data visualization, including activity maps, timelines, and graphical representations. Additional libraries and frameworks, such as UrlExtract for extracting URLs from WhatsApp chats and WordCloud for generating visual representations of word frequency patterns, enhance the analytical capabilities of the system. Specialized data structures are implemented using Collections.

Development tools such as PyCharm, an integrated development environment (IDE), facilitate coding, debugging, and project management for WhatsInsight. Git, a version control system, enables collaborative coding and tracking changes throughout the development process, ensuring code integrity and version management.

System maintenance involves regular updates, version control, and monitoring for performance optimization. Version control systems like Git manage changes efficiently, while regular updates ensure system stability and security. Monitoring mechanisms track system performance, resource utilization, and potential bottlenecks, allowing for algorithm and process optimization.

Evaluation and improvement are integral to system maintenance. User feedback integration mechanisms gather insights for continuous improvement. Performance metrics aligned with project objectives assess system effectiveness, enabling data-driven decisions for enhancements. Testing and quality assurance ensure system reliability through rigorous testing, bug tracking, and resolution.

Deployment and rollout strategies involve phased deployment and user onboarding to ensure smooth integration of new features and user familiarity with the system. Comprehensive training materials and user onboarding sessions facilitate seamless adoption of the system.

In conclusion, the WhatsApp chat analyzer combines Agile and Lean methodologies with a robust set of programming languages, frameworks, libraries, and development tools to deliver a responsive and efficient solution for sentiment analysis. Through iterative development, cross-disciplinary collaboration, and continuous improvement, the analyzer meets evolving user needs and fosters a data-driven approach to decision-making.

4 RESULTS AND DISCUSSION

A application has been developed which is analyzed in this section. Figure 1 shows the interface for the users to enter the .txt file for the chat analysis.
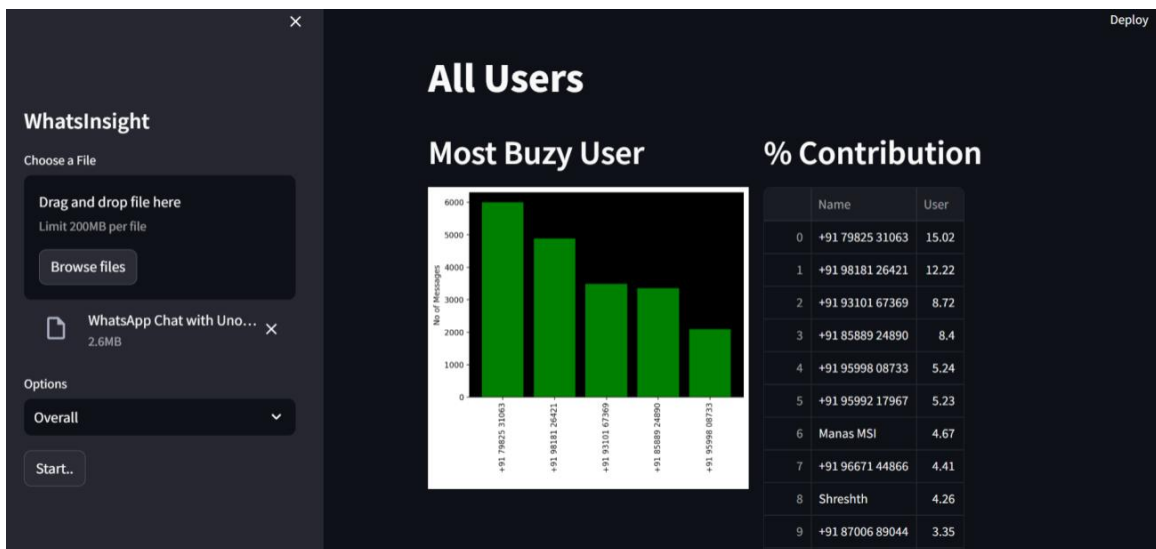

Fig. 1: Input Interface of WhatsApp Analyzer

Figure 2 shows the dropdown list from the group chats uploaded by the user which also allows the user to select an individual from the group.
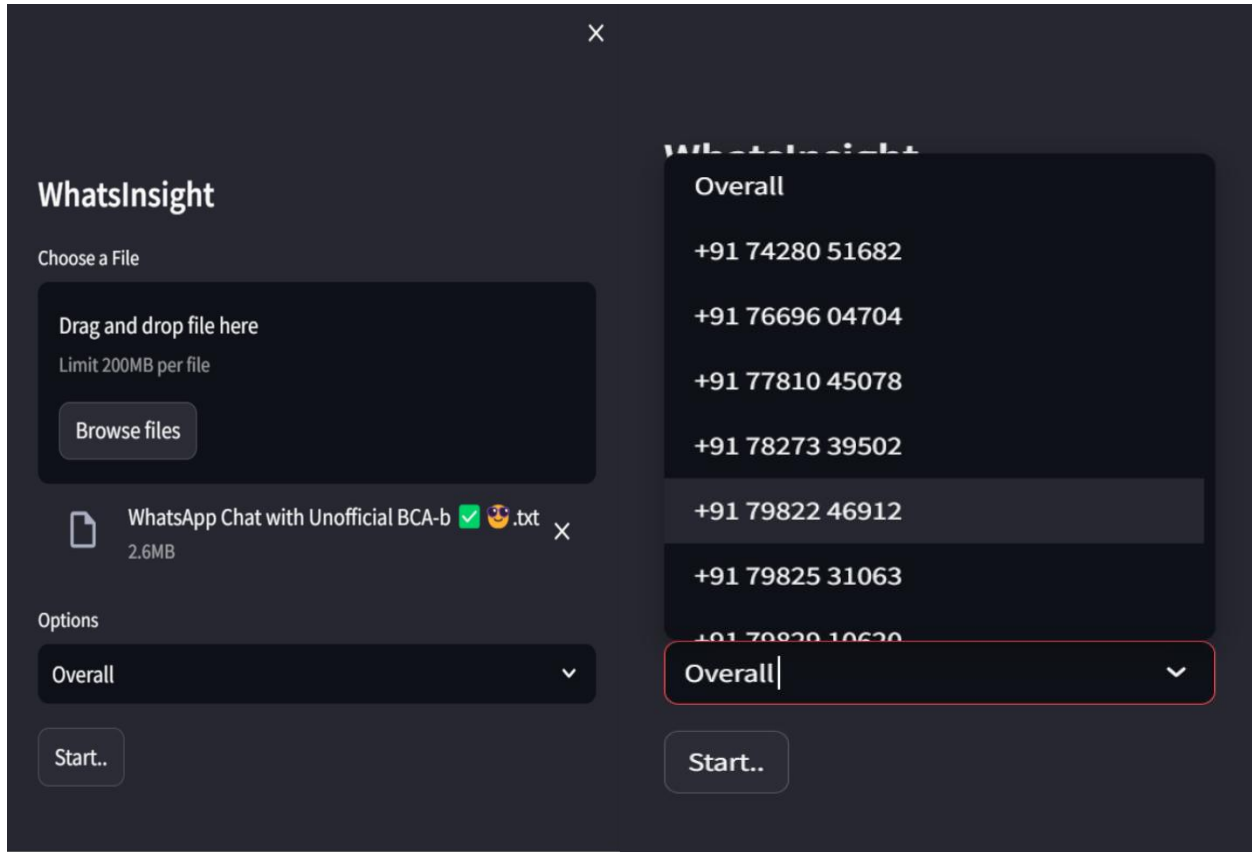
Fig. 2: Drop down list to select an individual for analysis

Figure 3 shows the data frame of the chats uploaded by the user with various

attributes.   **Fig.**
3:Data frame of the chats uploaded by the user

Figures 4-11 show the visual analysis of the chats uploaded in the form of charts , bar graph , pie charts and Heat map (confusion matrix) to show the Most busy user , %Contribution , Timeline , weekly activity , words visualization & emoji analysis .
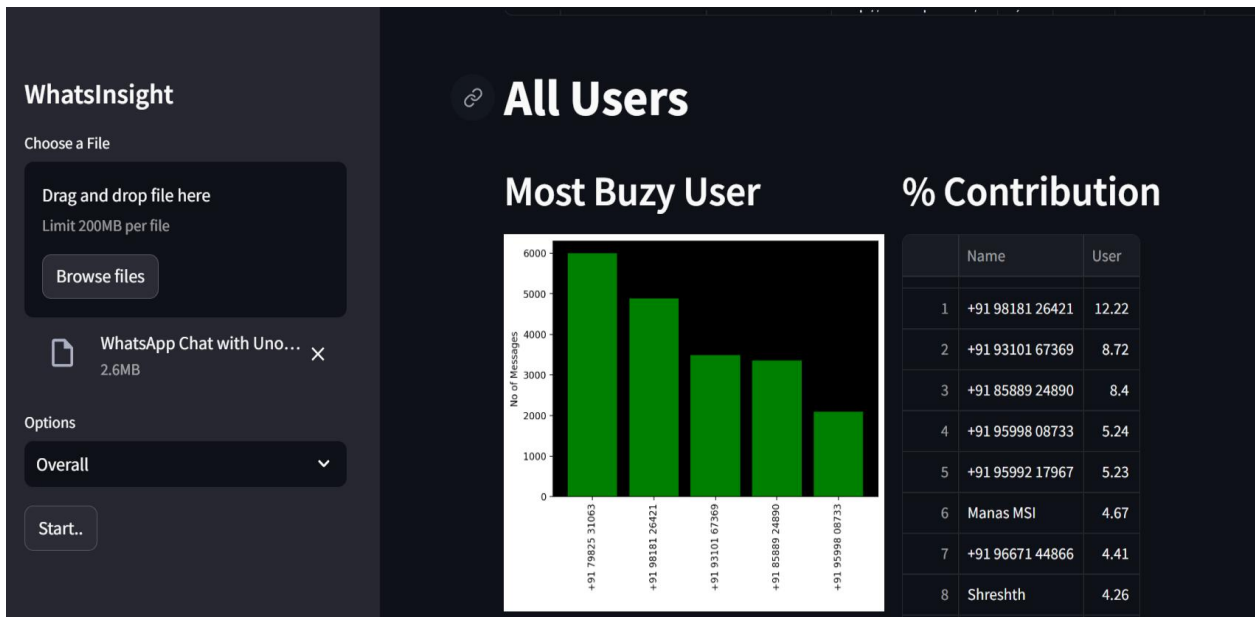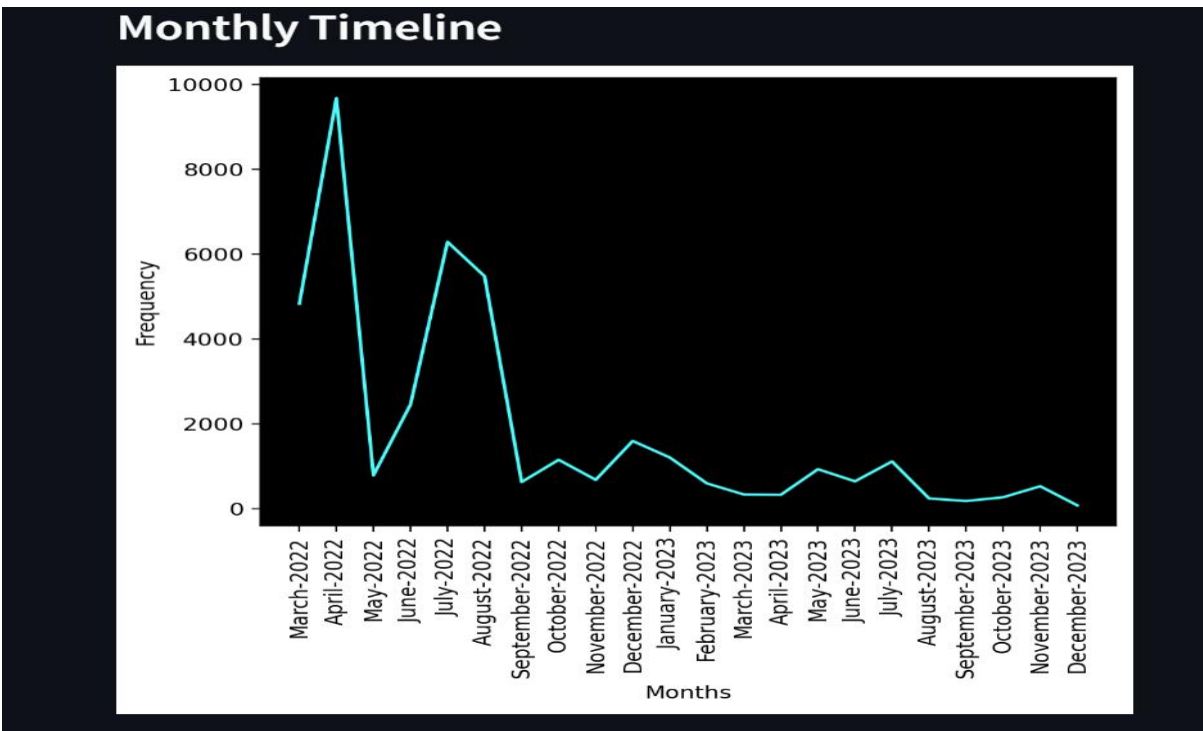
Fig. 4: Analysis of the chats using bar graph

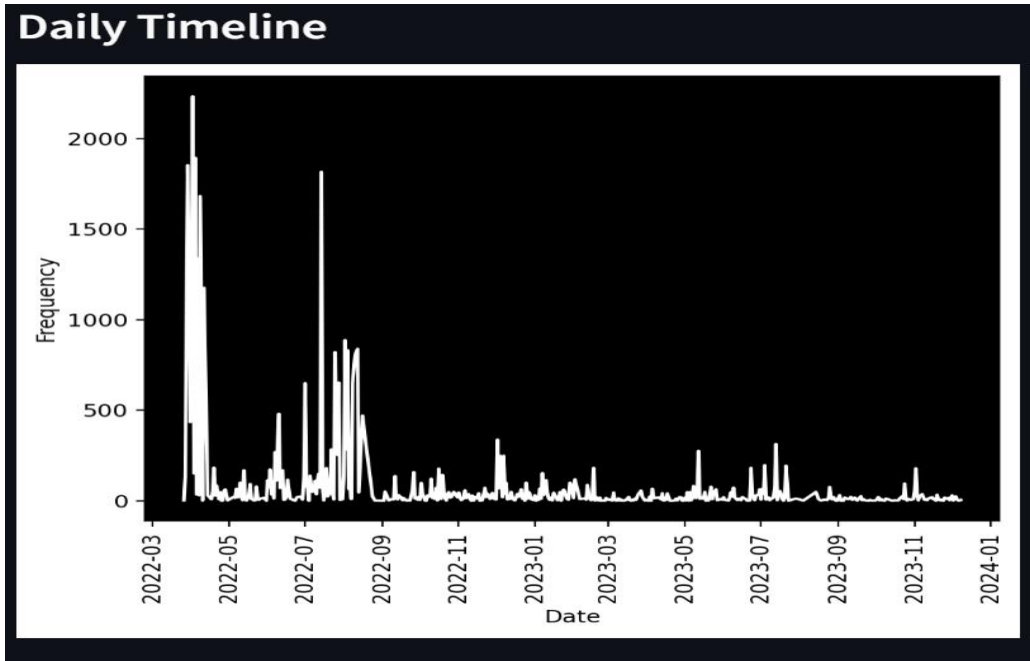**Fig. 5**

Month wise frequency timeline

Fig. 6: Day wise frequency timeline
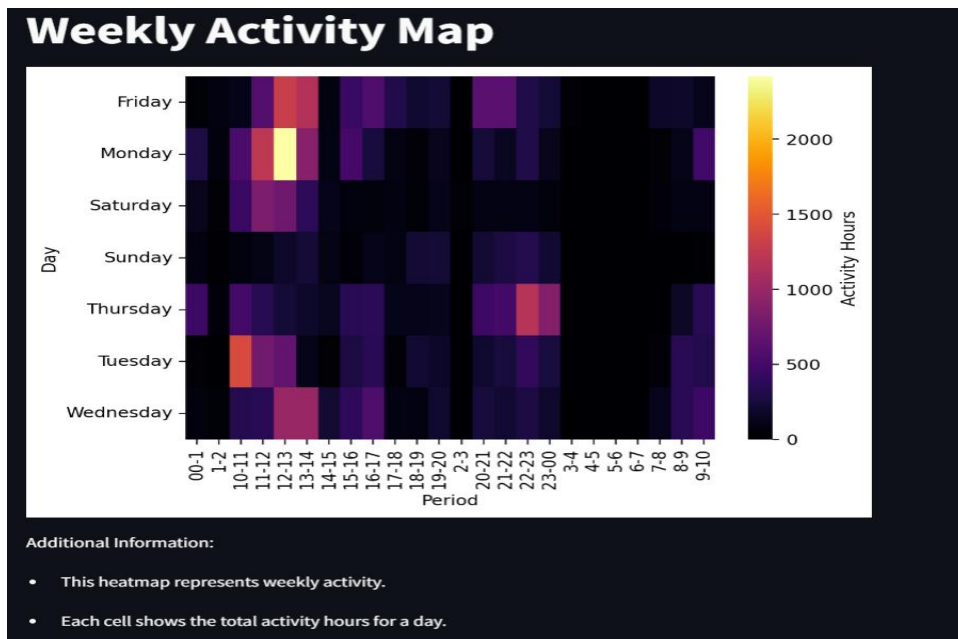
Fig.7: Activity visualization using heat map

Fig.8: Visualization using Word Cloud

Fig. 9: Most frequent emojis

Fig. 10: Sentiment Analysis

Fig. 11: Positive users vs Negative users using bar graph

The figures show the most Crucial and Important part of the study. Sentiment Analysis shows the visual representation of the polarity count through graph representing users and words helping people in decision making.
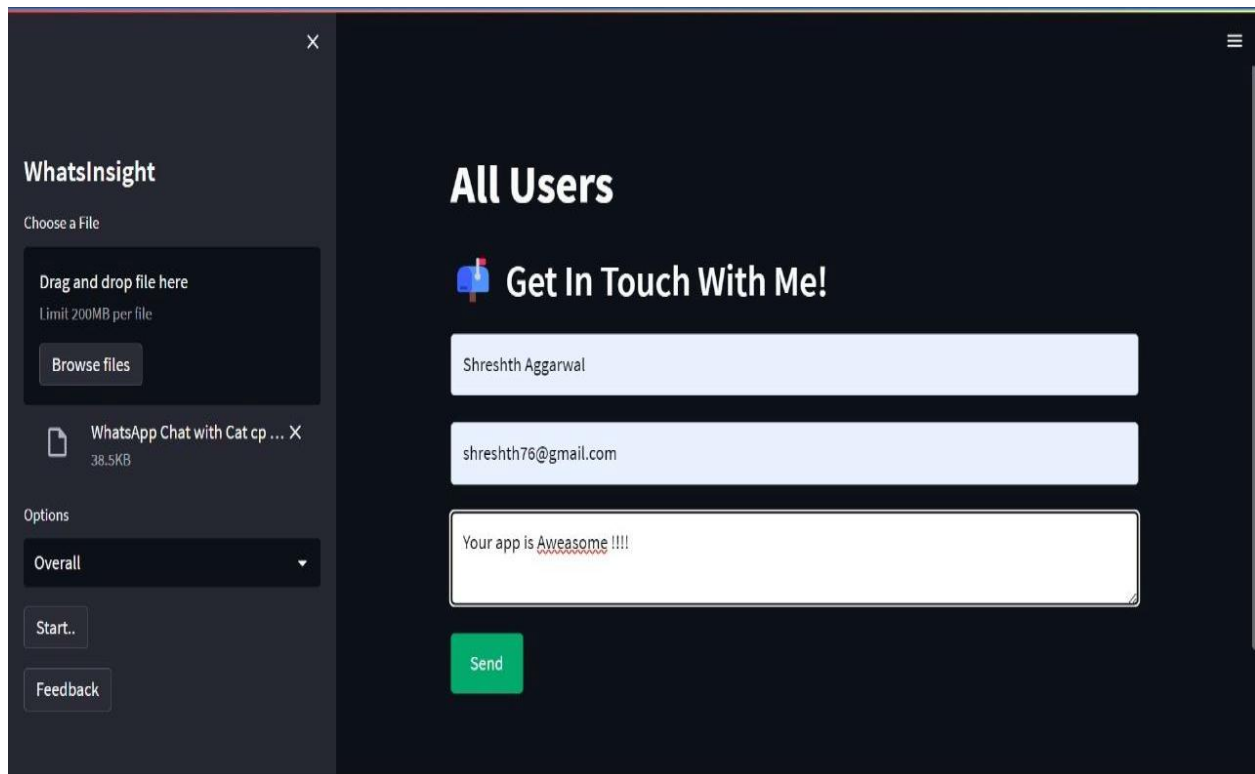
Fig. 12: Feedback panel

The application also has a panel to get the users' feedback to make some improvements and changes in the application. The continuous evolution will help in enhancing the user experience gradually.

5 CONCLUSION & FUTURE SCOPE

The developed application supports proactive mental health measures by identifying potential concerns early and facilitating timely interventions. By focusing on sentiment analysis, keyword and topic identification, and conversation patterns, it recognizes shifts in user mood or signs of distress, promoting healthier digital interactions. In conclusion, this application is a powerful tool for addressing mental health challenges through advanced WhatsApp data analysis, playing a crucial role in fostering safer, more supportive online communities.

The application can evolve significantly with future enhancements, such as advanced emotion and mood recognition, improved contextual understanding, and integration with behavioral psychology for nuanced sentiment analysis. Enhanced healthcare applications could enable early detection of mental health issues. Emphasis on ethical and privacy considerations, AI and machine learning integration, advanced NLP techniques, and real-time visualizations will improve accuracy and adaptability. Future developments include multi-platform support, collaborative analysis, strengthened security, and scalability optimizations. These advancements will make this application a more sophisticated and powerful tool for in-depth chat data analysis, catering to diverse user needs.

6 References

[1] Wahyuni, S., & Febianti, K. (2019). The use of WhatsApp group discussion to improve students' writing achievement. *Indonesian Educational Administration and Leadership Journal*, *1*(1), 45-51.

[2] Statista. (2021). Number of WhatsApp users worldwide from 2013 to 2021. Retrieved from Statista

[3] Pew Research Center. (2019). Mobile Messaging and Social Media 2019. Retrieved from Pew Research Center

[4] Ajani, O. A., & Khoalenyane, N. B. (2023). Using WhatsApp as a tool of learning: A systemic literature review of prospects and challenges. *International Journal of Innovative Technologies in Social Science*, (3 (39)).

[5] Batra, B. (2016). News communication through WhatsApp. *International Journal of Informative & Futuristic Research*, *3*(10), 3725-3733.

[6] Kapase, H., Galande, K., Sonna, T., Pawar, D., & Salunke, D. (2018). A review on: Sentiment polarity analysis on Twitter data from different Events. *International Research Journal of Engineering and Technology (IRJET)*, *5*(3), 1479-1482.

[7] Turney, P. D. (2002). Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*.

[8] Liang, L., Li, C., Meng, C., Guo, X., Lv, J., Fei, J., & Mei, S. (2022). Psychological distress and internet addiction following the COVID-19 outbreak: Fear of missing out and boredom proneness as mediators. *Archives of Psychiatric Nursing*, *40*, 8-14.

[9] Mahlous, A. R. (2024). The Impact of Fake News on Social Media Users During the COVID-19 Pandemic, Health, Political and Religious Conflicts: A Deep Look. *International Journal of Religion*, *5*(2), 481-492.

[10] Özgür, H. (2020). A Systematic Review on Cyberbullying Interventions and Preventions. *Shanlax International Journal of Education*, *9*(1), 11-26.

[11] Andreassen, C. S., Torsheim, T., Brunborg, G. S., & Pallesen, S. (2012). Development of a Facebook addiction scale. *Psychological reports*, *110*(2), 501-517.

[12] Richard, M., Osunsan, O. K., Byamukama, V., & Timbirimu, M. Integration of WhatsApp and its Influence on Small and Medium Enterprise Performance in Buloba Town, Wakiso, Uganda.

[13] Church, K., & De Oliveira, R. (2013, August). What's up with WhatsApp? Comparing mobile instant messaging behaviors with traditional SMS. In *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services* (pp. 352-361).

[14] Liu, B. (2012). *Sentiment analysis and opinion mining. Synthesis lectures on human language technologies, 5 (1), 1-167.*

[15] Pak A, P. (2010). Twitter as a Corpus for sentiment analysis and opinion mining, Proceedings of the International Conference on Language Resources and Evaluation (LREC'10). *Valletta, Malta*, *1320*, 1326.

[16] Bafna, K., & Toshniwal, D. (2013). Feature based summarization of customers' reviews of online products. *Procedia Computer Science*, *22*, 142-151.

[17] Mudinas, A., Zhang, D., & Levene, M. (2012, August). Combining lexicon and learning based approaches for concept-level sentiment analysis. In *Proceedings of the first international workshop on issues of sentiment discovery and opinion mining* (pp. 1-8).