"Are we ready for AI?"
Sentiment Analysis of Tweets on ChatGPT

Geetansh Saxena
Matriculation Number - 6322801
geetansh.saxena@student.uni-tuebingen.de

Lalitha Sivakumar
Matriculation Number - 6300674
lalitha.sivakumar@student.uni-tuebingen.de

Abstract

The aim of this paper is to find an answer to the question- "Is the World ready for AI?" on the basis of tweets that are related to ChatGPT. We perform sentiment analysis for the tweets that contain information related to ChatGPT, in order to capture sentiments of people from diverse professions. Three different models are used to analyse and compare the results to avoid any bias in drawing the overall conclusion.
GitHub Repo: https://github.com/SGeetansh/sentiment-ChatGPT

1 Introduction

ChatGPT by OpenAI has revolutionized how people think about AI. We analyze the sentiments of tweets containing the hashtag 'chatgpt' to gain insight on public perception and readiness for AI integration in society. By using the twitter bios of people, we filter out the people from technical fields of work (5) and report their views about the uprising of ChatGPT.

1.1 Dataset Description

The data is obtained from Konrad B, Kaggle Grandmaster (1). The dataset is a collection of 84,552 tweets containing the hashtag 'chatgpt', posted in the period between 05 December 2022 to 04 January 2023. Following are the keys present in the dataset: username', 'text', 'user location', 'user description', 'user created', 'user followers', 'user friends', 'user favourites', 'user verified', 'date', 'hashtags', 'source'.

1.2 Sentiment Analysis and Models Used

Sentiment analysis is the process of determining the emotional tone in a given text (5). In this study, tweets are classified into positive, negative, and neutral sentiment towards ChatGPT. Three different sentiment analysis models were used to avoid any biases: TextBlob (4) (lexicon-based and machine learning-based).VADER (3) (lexicon-based for social media text). H5 file - Sentiment140 (2)(1.6M annotated tweets with pre-labeled sentiment; in this study, it is referred as 1.6M).

2 Sentiment Distribution

The results of sentiment prediction performed by the above described models are summarised in the 3 pie-charts in figure 1. We observe that VADER is more biased towards positive predictions, and the statistics predicted by 1.6M and TextBlob are almost similar. More importantly, the majority sentiments towards ChatGPT and the acceptance of it, has been either positive or neutral. A small section of the population (about 14%) have negative views on ChatGPT.

3 Observing Everyday Sentiments Over Time

The visualization in 2 depicts the daily distribution of sentiments expressed in the tweets about ChatGPT. The bar graph shows the fluctuation in sentiments for each day, giving us insights into the dynamic nature of opinions about the technology. Observing the graph, we can see that positive sentiments have dominated throughout the study period, with a consistent proportion of the three

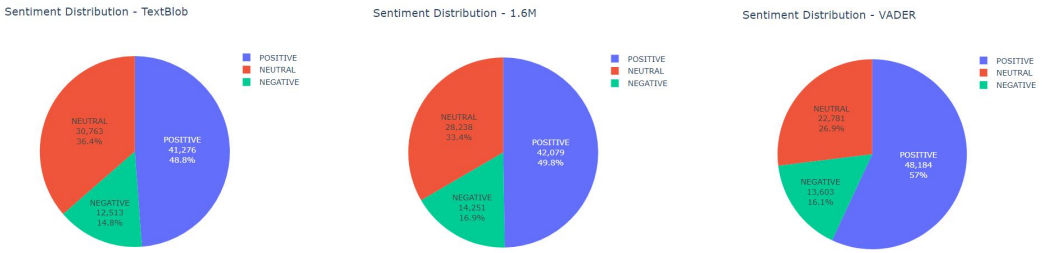

Figure 1: Sentiment Distribution

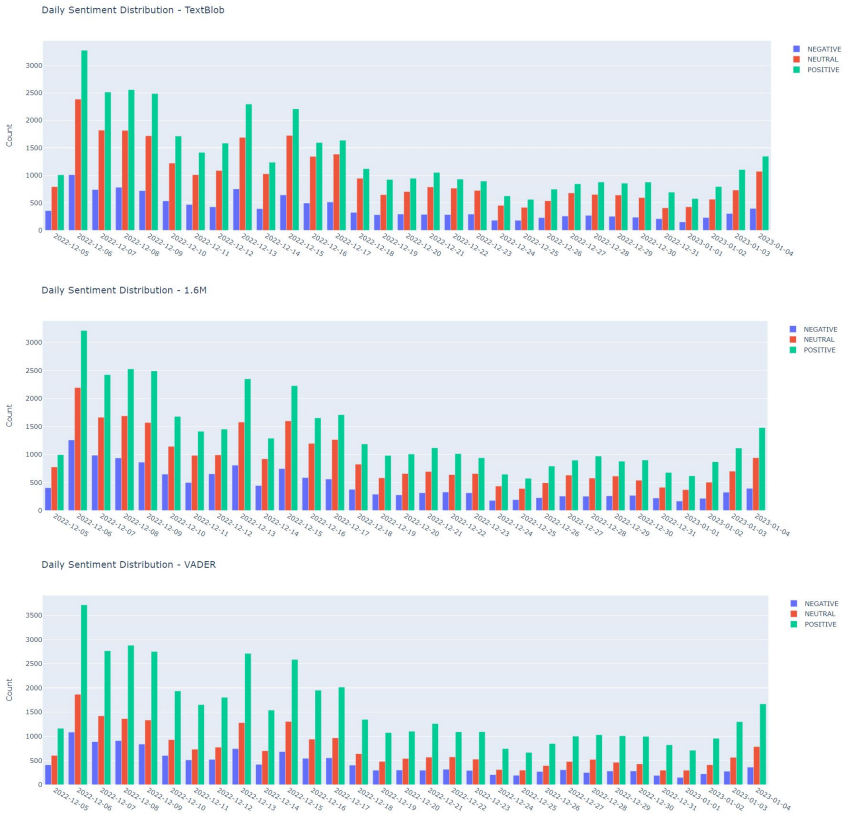

Figure 2: Trend of changes in the sentiments on a daily basis

sentiments (positive, negative, and neutral) across all days. This graph also provides an overview of the trend of relevance and popularity of ChatGPT on Twitter over time, indicating the level of adoption and normalization of its use in people's daily lives.

4 Word Frequency Analysis

Since the further statistical analysis of all the 3 models were giving simalr results, we decided to showcase only the results from textblob model.

Based on the word frequency graph for tweets with positive sentiments (figure 4), it appears that the public is interested in the topic, potentially seeking information or discussing its pros and cons. Words such as "new," "like," "good," "great," and "better" suggest a positive perception, while words such as "asked," "using," and "code" suggest a desire for understanding and knowledge. The presence of the word "Google" could indicate that people might be looking at ChatGPT as an alternative.

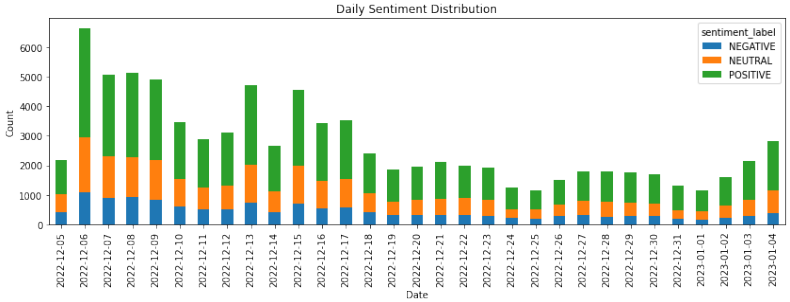

Figure 3: Daily tweet count with proportions

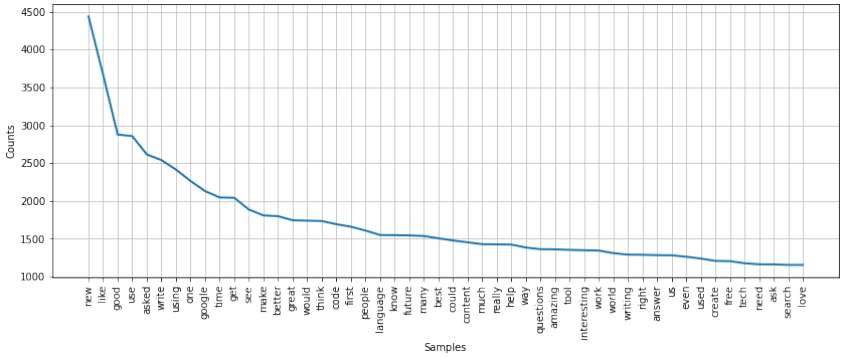

Figure 4: Word frequency diagram for tweets with positive sentiments

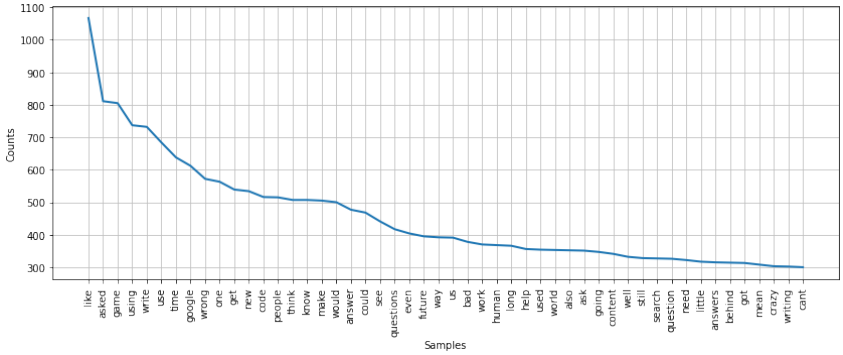

Figure 5: Word frequency diagram for tweets with negative sentiments

For the negative sentiments, the word frequency graph 5 suggests that the public may have concerns with ChatGPT. Words such as "wrong," "people," and "questions" indicate that there may be criticisms or doubts about the topic. The presence of words such as "asked," "using," "write," and "code" suggest that people are actively seeking information or trying to understand the topic. The words "time," "future", "human" and "bad" suggest that people may be concerns about the long-term implications of the topic. Though some words indicate doubts, none of them represent fear.

5 People in Science and Technology

The refinement of the population was carried out by selecting individuals who identified as being in the tech field, as indicated by keywords such as 'coder', 'programmer', 'engineer', 'data scientist', 'developer', etc. in their Twitter bios. A sentiment analysis was performed on this filtered population in comparison to the entire population. Despite the widespread use of ChatGPT by individuals from both technical and non-technical backgrounds, no significant difference was observed in the sentiment distribution among those in the tech field. However, a notable shift in sentiment was observed around

a specific date, where neutral sentiments appeared to convert to positive sentiments, as seen in figure 7. This could be attributed to multiple updates aimed at reducing server load, UX and preserving previous chats, hence accommodating more users.

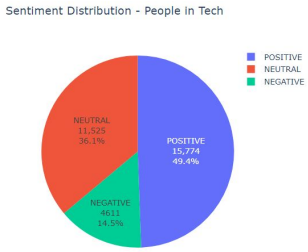

Figure 6: Distribution of sentiments among the tech population

Figure 7: Percentage change trend of sentiments among the tech population

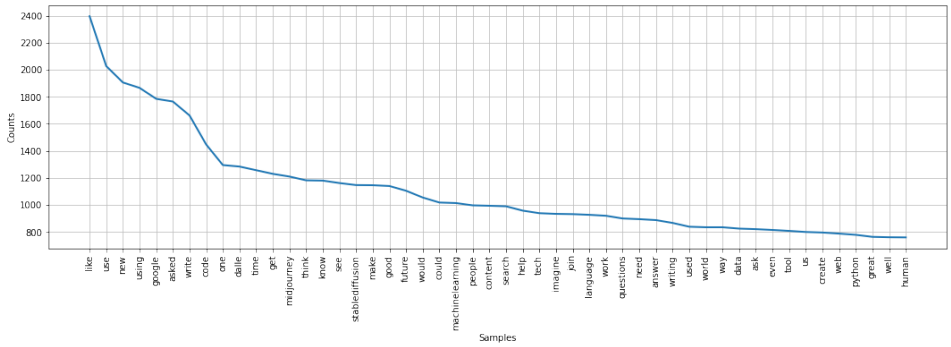

Figure 8: Word frequency diagram for people in science and technology

6 Conclusion and Scope

The project, "Are we ready for AI?", aimed to assess public sentiment towards advanced AI technology by analyzing tweets about ChatGPT from 5th December 2022 to 4th January 2023. We employed 3 sentiment analysis models and synthesized our findings in this report. Our findings reveal a small segment of the population expressing concern about AI implications. This is a natural reaction and should be acknowledged. Nevertheless, almost 50% of the population holds a positive view of AI, and there has been a shift from neutral to positive sentiments among the tech community. Based on these results, we conclude that the general public is receptive to integrating AI tools into their lives and would welcome further advancement in AI technology. Our analysis suggests that, with ongoing improvements and new opportunities, the population is ready for a gradual adoption of AI.

7 References

References

[1] Konrad B. ChatGPT: the Tweets. 2022. Kaggle. https://www.kaggle.com/datasets/konradb/chatgpt-the-tweets.

[2] Paolo Ripamonti. Twitter Sentiment Analysis: Data. 2023. Kaggle. https://www.kaggle.com/code/paoloripamonti/twitter-sentiment-analysis/data.

[3] NLTK Developers. VaderSentiment. 2023. PyPI. https://pypi.org/project/vaderSentiment/.

[4] Steven Loria. TextBlob. 2023. PyPI. https://pypi.org/project/textblob/.

[5] F. S. Rana and A. N. Singh. Fundamentals of Sentiment Analysis and Its Applications. 2016. ResearchGate. https://www.researchgate.net/publication/300965436_Fundamentals_of_Sentiment_Analysis_and_Its_Applications.