Sentimental Analysis of ChatGPT Tweets: A Comprehensive Study

Thomas Antony

SUNY Old Westbury, Dept. of Mathematics, Computer and Information Science

Old Westbury, NY

Introduction

ChatGPT has sparked a lot of discussion on social media, with users sharing their experiences and opinions. Many people praise its impressive abilities, noting how it can boost productivity, help with learning, and encourage creativity. They view it as a game-changing tool that could transform various fields, including education and customer service. On the flip side, there are serious concerns about how it works and the ethical issues it raises. Critics worry about the accuracy of the information it provides, the risk of misuse, and how AI might affect jobs and human creativity. These mixed feelings highlight a larger conversation in society about the role of artificial intelligence in our everyday lives. This project aims to organize these different user opinions into three main categories: Positive, Neutral, and Negative. By carefully preparing the data and testing different analysis methods, we want to uncover the key trends and feelings people have about ChatGPT. Ultimately, our goal is to help clarify public perceptions of this groundbreaking AI technology and contribute to the ongoing discussion about its future.

Method

Subjects or Participants:

This study focuses on people who have shared their opinions about ChatGPT on Twitter. Our participants include a wide range of users, from tech lovers and teachers to everyday users and critics. By looking at these tweets, we want to capture different feelings—both positive and negative—about what ChatGPT can do and how it affects daily life. This approach helps us understand public perception and the ongoing discussion about the role of AI in our society.

Experimental Design

Data gathering:

The dataset used in this study comes from Kaggle and includes tweets from Twitter. It contains over 219,000 entries, each showcasing users' opinions about ChatGPT. These tweets reflect a variety of feelings, making the data perfect for our analysis. Additionally, the dataset includes sentiment labels—Positive, Neutral, or Negative which help us train and evaluate machine learning models for accurate classification.

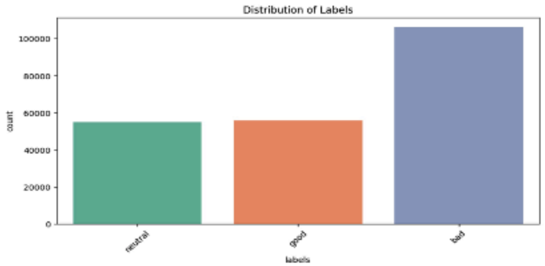
Data Cleaning:

During our data cleaning process, we started by removing unnecessary elements from the tweets, like links, special characters, and emojis, to keep the dataset clean and focused. We also filtered the tweets to include only those in English by using the langdetect library, which helped maintain consistency in our analysis. Additionally, we checked for and removed duplicate tweets to ensure that each entry was unique. This step was important to prevent repetition in our analysis and to maintain the quality of our data.

Data analysis:

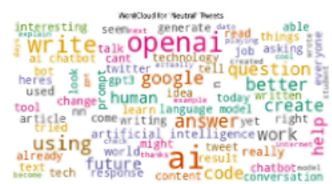
We used sentiment analysis to understand how people feel about ChatGPT on Twitter. This approach helped us look at the emotions and attitudes in the tweets, allowing us to sort them into three categories: Positive, Neutral, or Negative. By examining these sentiments, we aimed to get a better idea of users' opinions and how ChatGPT affects their experiences.

Tables/Figures



Results					
Model	**Sentimental**	**Precision**	**Recall**	**F1-Score**	**Accuracy**
Logistic Regression	Weighted Average Positive Neutral Negative	0.80 0.80 0.85 0.71	0.81 0.77 0.59 0.93	0.80 0.78 0.65 0.89	0.8
Decision Tree	Weighted Average Positive Neutral Negative	0.71 0.65 0.54 0.82	0.70 0.64 0.59 0.79	0.70 0.64 0.57 0.81	0.7
Random Forest	Weighted Average Positive Neutral Negative	0.76 0.73 0.75 0.77	0.76 0.74 0.45 0.93	0.75 0.73 0.57 0.84	0.75
XGBoost	Weighted Average Positive Neutral Negative	0.69 0.76 0.62 0.69	0.69 0.60 0.30 0.95	0.66 0.67 0.40 0.80	0.69
Adaboost	Weighted Average Positive Neutral Negative	0.60 0.69 0.52 0.58	0.60 0.45 0.04 0.97	0.52 0.55 0.08 0.73	0.6

WordCloud



Results

The Logistic Regression model turned out to be the most effective among the different algorithms we tested. It strikes a great balance between being simple, accurate, and easy to understand. This simplicity is especially valuable in a field where being transparent is important.
Even though it's straightforward, the model performed well, making it a reliable option for sentiment analysis. It serves as a good baseline to compare against more complex models. The strong results show that, even with fewer parameters, Logistic Regression can effectively capture the subtle feelings expressed in the data. This reliability is crucial for real-world applications, where users need to trust the insights generated from the analysis.
Additionally, the model is easy to interpret, allowing us to explain how it makes its predictions. This accessibility is important for users who may not have a technical background. Understanding how the model works helps build trust in AI tools, especially in situations where knowing the reasoning behind decisions is essential. Overall, the success of the Logistic Regression model highlights its importance in sentiment analysis and sets a standard for evaluating future models.

Discussion

The analysis provided clear insights into the strengths and weaknesses of each model we looked at. Logistic Regression stood out as the best option because it's simple, easy to understand, and very effective, consistently scoring the highest on various performance metrics. This makes it a reliable choice for sentiment analysis, as users can easily grasp how it makes predictions, which is crucial for transparency. On the other hand, Decision Trees were good at identifying non-linear patterns in the data but struggled with overfitting. This means they performed well on training data but had trouble making accurate predictions on new data, which can lead to unreliable results in real-world situations. Random Forest, which is based on Decision Trees, improved on their robustness and reduced the risk of overfitting, but it still wasn't as consistent as Logistic Regression, which is important for predictable outcomes. XGBoost performed well with complex datasets and is often successful in competitions, but it didn't outperform the simpler models in this analysis and required more computational resources, making it less suitable for situations where resources are limited. Its complexity can also make it harder to interpret, which may not be ideal for all users. Adaboost was an interesting approach to enhancing weaker models, but it didn't keep up with the performance of the other models, suggesting that while it has potential, it might not be the best fit for our needs. Overall, this analysis highlighted that sometimes simpler methods like Logistic Regression can be just as effective, if not more so, than complex ones. This underscores the importance of choosing the right model based on the specific context and needs of the analysis to make informed decisions in future projects

Summary and Conclusions

Sentiment analysis uses natural language processing (NLP) to identify emotions in text, helping us understand public opinion. In this project, we focused on tweets about ChatGPT using a Kaggle dataset to see how users feel. We started by cleaning and preparing the data to make it ready for analysis. We used techniques like TF-IDF vectorization to turn the text into numerical data, which allowed us to apply various machine learning methods effectively.
We sorted the tweets into three sentiment categories: Positive, Neutral, and Negative. To do this, we tested several models, including Logistic Regression, Decision Tree, Random Forest, XGBoost, and Adaboost. Each model was assessed based on its accuracy, ease of understanding, and ability to capture the subtleties of the sentiments expressed in the tweets.
Among all the models, Logistic Regression stood out as the best performer. It struck a great balance between being simple and accurate, making it easy to interpret while still providing reliable results. This effectiveness helped us draw meaningful conclusions about how users perceive ChatGPT.
The project also emphasizes the importance of choosing the right model for sentiment analysis. While more complex models like XGBoost and Adaboost can be beneficial, they often require more computing power and can be harder to understand. In contrast, the straightforward nature of Logistic Regression made it the best choice for our analysis, showing that sometimes simpler models can be more effective than complicated ones.
Overall, this project highlights how powerful machine learning can be for automating sentiment analysis and stresses the importance of understanding user sentiments regarding new technologies like ChatGPT. By gaining insights into how users feel, developers and stakeholders can make better decisions to improve user experience and address any concerns effectively.

Acknowledgements:

I want to express my gratitude to Professor Shebuti Rayana and the research foundation for giving me the opportunity to conduct this research. I hope my work can help more transfer students receive the guidance they need.

Funding:

Thomas Antony