BA723 Business Analytics Capstone

ChatGPT Sentiment Analysis

Submitted by:

Justine Tinio (301241837)

Submitted to:

Savita Seharawat

David Parent

Due date: July 17, 2023

Analytics Startup Plan

Synopsis: ***This document provides a high-level walkthrough of the activities required to guide completion of the analysis.***

Project	*ChatGPT Sentiment Analysis*
Requestor	*Savita Seharawat, David Parent*
Date of Request	*July 4, 2023*
Target Quarter for Delivery	*August 18, 2023*
Epic Link(s)	*N/A*
Business Impact	*To further understand customer opinions and feedback on Chat GPT usage.*

Table of Contents

1.0 Business Opportunity Brief

On November 30, 2022, OpenAI unveiled the ChatGPT (Chat Generative Pre-Trained Transformer) artificial intelligence chatbot. Since then, a variety of sectors and applications have made extensive use of it, including virtual assistants, content creation, language learning, information sharing, research, entertainment, and gaming. It also allows normal language interactions and provides responses that are comparable to those of humans. Additionally, it has recently garnered a lot of discussion and interest on social media platforms.

Sentiment analysis on ChatGPT presents a valuable business opportunity. By analyzing user sentiments, businesses such as OpenAI can gain insights into customer feedback, customize responses, ensure quality control, evaluate model performance, and track user sentiments over time. This enhances the user experience, enables personalized interactions, filters inappropriate content, and provides valuable insights for strategic decision-making. The target market for sentiment analysis on Chat GPT includes businesses in customer service and support, e-commerce, and retail, social media and digital marketing, market research and consumer insights, language learning, and education, as well as entertainment and gaming industries.

The goal of this project is to analyze the sentiment of tweets to determine the public's overall sentiment toward ChatGPT's performance. Additionally, we aim to analyze tweet patterns to identify underlying trends and patterns. Furthermore, based on their sentiment, we intend to determine potential general subjects of conversation in tweets.

1.1 Supporting Insights

Due to its massive following, Twitter has garnered a lot of attention as one of the most important venues for people to publicly voice their opinions on current events and their personal lives (Li et al. 2013). (Pang & Lee 2008; Poria et al. 2014; Taboada et al. 2011; Liu 2012) Several research have examined sentiment analysis and opinion mining on Twitter. Pure sentiment analysis has given way to sentiment and subjective analysis (Pang & Lee 2004), while machine learning or lexicon-based approaches have given way to more sophisticated hybrid methods (Serrano-Guerrero et al. 2015). Similarly, simple text mining techniques have given way to more complex symbol and feature recognition (Liu 2012), pure sentiment analysis has given way to sentiment and subjective analysis, and more. This study tested research questions related to sentiment towards ChatGPT compared to other AI language models and if it varies across different user groups or demographics.

Many different businesses, including customer service, content creation, and virtual assistants, have used AI language models like ChatGPT. Advances in AI, particularly in natural language processing, have substantially aided these models' capacity to grasp and produce writing that is human-like. However, ethical issues have been raised due to its ongoing development. Discussions on bias removal, ethical AI deployment, and assuring the models' beneficial and safe use have all gained significant importance. To improve their efficiency and flexibility, AI models are frequently customized and fine-tuned for tasks or domains. As a result, businesses constantly need to modify the models to suit their requirements. To improve user experiences and streamline workflows, AI language models are also being easily included in apps like chatbots, virtual assistants, and content creation tools.

As of September 2021, OpenAI stands out as a frontrunner among its principal rivals. The company that created ChatGPT, OpenAI, has made a name for itself as a leader in the field of AI research and development. In addition to emphasizing the significance of creating safe and useful

AI, they provide AI models like GPT-3. The Dialogflow platform from Google, which enables programmers to create conversational agents and chatbots with natural language comprehension skills, is one of the industry's other top rivals. For the creation of chatbots and virtual assistants, Microsoft's Azure Cognitive Services, especially Language Understanding (LUIS), are frequently employed. With IBM's Watson Assistant, programmers can use natural language processing to build AI-powered virtual assistants. Using speech and text interactions, developers can create conversational user interfaces for apps using Amazon's Amazon Lex service.

1.2 Project Gains

Businesses can gain a lot from sentiment analysis of ChatGPT-related tweets. Companies can acquire important insights to boost revenue by examining consumer opinions, preferences, and satisfaction levels. Businesses can capitalize on its benefits by determining which features are popular with customers by identifying positive sentiment. Additionally, sentiment analysis identifies areas for development, assisting programmers in resolving client issues and raising the standard of the AI model. Better user experiences, higher levels of customer satisfaction, and positive word-of-mouth ultimately result from this iterative process of refining, which also drives revenue growth.

Sentiment research also allows for time and money savings. By eliminating the need for manual inspection and automating the analysis of tweets referencing ChatGPT, organizations may save precious resources and respond quickly to client feedback. Organizations can effectively discover patterns, trends, and pain spots by utilizing AI-powered sentiment analysis, which enables quicker decision-making and prompt action to meet customer needs.

A proactive strategy is essential to setting priorities in relation to organizational objectives and key performance indicators (KPIs). A dedication to continually enhancing the user experience is shown through acting on sentiment analysis data and customer feedback. The development team can improve the performance of the model, raise customer satisfaction, and foster loyalty by addressing identified pain areas. Customers should be concerned because their opinions are appreciated and their issues are actively handled, which leads to an AI model that more closely matches their requirements and expectations.

There are consequences to think about if sentiment analysis is not used or customer sentiment is neglected. Ignoring feedback can lead to customer dissatisfaction, negative sentiment spread, and diminished trust and loyalty. Missed revenue opportunities, a decline in competitiveness, and a detrimental effect on company objectives and KPIs can occur from failing to respond to customer complaints and implement essential modifications. Prioritizing sentiment analysis supports customer-centric goals and promotes brand reputation, long-term growth, and customer pleasure.

It's vital to remember that depending on the industry, environment, and implementation, the precise results, advantages, and implications of sentiment analysis for ChatGPT-related tweets may change. It is advised to carry out in-depth research, collect current data, and carry out analyses within the context of the organization to obtain exact insights on revenue increases, quality improvements, cost and time savings, and alignment with company goals and KPIs.

2.0 Analytics Objective

The project's main objective is to answer the question "How can sentiment analysis algorithms effectively process and interpret the nuanced emotions expressed in short-form tweets to improve ChatGPT's response accuracy and enhance the user experience?".

The deliverable of this project will not only contain the results of the analysis but also a suggested operational model or roadmap for a pilot project targeted at strengthening its response precision and user experience. To accomplish these goals, this thorough strategy will include a variety of recommendations. First, an iterative feedback loop can be used to train and improve ChatGPT over time. By doing this, it will be able to better grasp and interpret the nuanced emotions represented in tweets. Additionally, user-specific sentiment profiles that account for unique sentiment preferences and customize interactions can be made to further enhance the chatbot's answers. An automated system can be created to categorize and properly handle such comments to address specific user problems or critiques discovered through sentiment analysis, leading to continual improvements in ChatGPT's performance. Additionally, a frequent sentiment analysis monitoring method can be built to keep tabs on user sentiment. This allows for the tracking of changes in user sentiment over time and makes it easier to proactively identify any potential problems or emerging patterns.

Key questions:

1. What is the overall sentiment expressed in tweets discussing ChatGPT?
2. How does sentiment towards ChatGPT vary across different user groups or demographics?
3. How does sentiment towards ChatGPT compare to other AI language models?
4. Are there any specific user concerns or criticisms expressed in tweets about ChatGPT?

2.1 Other related questions and Assumptions:

Assumptions:

The recent tremendous growth in ChatGPT's popularity has led to several conversations and references on social media sites. This assumption is supported by the finding that it has gained popularity and traction in online communities, with an increase in the number of people conversing, exchanging stories, and voicing opinions about the language model. It is assumed that this increased degree of participation and conversation demonstrates ChatGPT's importance and influence in the public domain.

1. Sentiment analysis through short-form tweets is an accurate reflection of how users feel about ChatGPT in general. According to this supposition, as tweets are a common way for people to express their thoughts and feelings, they offer a trustworthy snapshot of how users feel about ChatGPT as a whole. The underlying premise is that sentiments conveyed in tweets are real and accurately reflect users' views, understandings, and interactions with the language model.
2. Algorithms for sentiment analysis can accurately decipher subtle emotions expressed in tweets. This premise holds that sophisticated natural language processing tools coupled with sentiment analysis algorithms may successfully capture and comprehend the nuanced emotional nuances conveyed within the constrained character space of tweets. The underlying presumption is that sentiment analysis models have been trained on a variety of datasets, giving them the ability to recognize and understand a wide range of emotional expressions.

3. The opinions shared in tweets are indicative of how the public feels about ChatGPT as a whole. This presumption argues that the opinions expressed in tweets when analyzed and aggregated collectively, offer a trustworthy portrayal of the overall attitude towards ChatGPT that is now prevalent in society. It suggests that valid conclusions regarding the general public's opinion on the language model can be made by examining a large enough sample of tweets.

2.2 Success measures/metrics

Based on the project's primary goal and the expected results, the success metrics can be determined. Here are a few possible success indicators:

1. Sentiment analysis accuracy: Evaluate how well the algorithms for sentiment analysis can discern the emotions represented in tweets about ChatGPT. Standard evaluation criteria like precision, recall, and F1 score can be used to assess this.
2. Sentiment Distribution: Examine how the sentiment of tweets about ChatGPT is distributed. This includes calculating the ratio of the dataset's expressed favorable, negative, and neutral sentiments. A balanced distribution that matches the actual emotion of consumers can be used to gauge success.
3. User Engagement and Satisfaction: After adopting sentiment analysis insights, evaluate user engagement and satisfaction with ChatGPT. Metrics like user retention, session length, and user feedback can be used to track this. An improved user experience is indicated by positive improvements in these measures.
4. Determine underlying trends and patterns in the sentiments expressed in tweets relating to ChatGPT. The capacity to unearth important insights and spot important changes or new user sentiment over time is a key indicator of success.
5. Alignment with User Feedback: Consider how closely the results of the sentiment analysis match up with what users have to say about ChatGPT. This can be established by contrasting the findings of sentiment analysis with user surveys, reviews, or other direct sources of feedback.
6. Assess the performance of the operational model or blueprint that is suggested and put out in the project. This can be done by looking at how well the suggested tactics, like incorporating sentiment feedback into ChatGPT's training process or creating user-specific sentiment profiles, are implemented and what effect they have.

To evaluate the project's success and impact, it is crucial to develop precise metrics that are in line with its objectives.

2.3 Methodology and Approach

Type of Analysis: Sentiment Analysis

Sentiment analysis is a method of examining textual information to ascertain the emotional undertone of the discourse. This method is frequently applied in social media monitoring to determine how the general population feels about a range of subjects, including companies and goods. We may use sentiment analysis to examine tweets about ChatGPT and ascertain how people feel generally about the model.

Methodology:

1. **Data collection:** A sizable collection of tweets containing mentions of the model would first need to be gathered. To gather this data, we can make use of a variety of tools and methods, including web scraping and API connections.

2. **Data Preparation:** After datasets have been gathered, we will move on to data preparation and put the data into a format that can be used. By classifying the independent/feature variables and the dependent/target variables in this stage, we will learn more about our datasets. For reference, a glossary explaining the terms will be created.

3. **Data Wrangling:** In this step, we will tidy up the data and look for any duplicate or missing values. If any of these exist, we shall mark them as "null" or "NaN." This also includes removing useless data, such as Tweet ID and Permalink.

4. **Data analysis:** Since the majority of the data are numerical, we must use the model to look for linear relationships between them.

5. **Train the model:** At this stage, we determine whether the model can use all the parameters and produce reliable results. The accuracy of predictions and outcomes generated will increase as more variables and their relationships become available.

6. **Put the model to the test:** We'll fit all the inputs to the model and evaluate its accuracy. It will go on to the deployment step once it has provided all the data required to remedy the company's problem.

Output: The output will be a set of insights, rules, and strategic recommendations that will help us to evaluate ChatGPT's effectiveness.

3.0 Population, Variable Selection, considerations

Audience/population selection: A selection of tweets with the hashtag #chatgpt are included in this dataset. This offers an insight into the online discussion surrounding the ChatGPT language model and can be utilized for a number of machine learning and natural language processing tasks, including sentiment analysis, topic modeling, and others. It enables comprehension of the neighborhood, the intensity of curiosity, and ChatGPT usage.

Observation window: April 3-16, 2023

Inclusions: The tweets, which connect in various ways to the ChatGPT language model, were harvested from Twitter. For each tweet in the dataset, the following details are included: User data (username, user ID, location, etc.) for tweets, timestamp of a tweet, Favorite and retweet counts, and hashtags incorporated into tweet URLs.

Exclusions: Observations and tweets that don't mention ChatGPT.

Data Sources: We would first need to compile a sizable dataset of tweets including mentions of the model to do a sentiment analysis of tweets regarding ChatGPT (41,003 tweets, 20 features/columns). Once we have the information, we can classify the tweets into those with good, negative, or neutral sentiments using machine learning methods. The findings of this analysis can help us enhance the model in a number of ways and offer insightful information about how the general public perceives ChatGPT.

Audience Level: The target audience for this project would likely be intermediate to advanced, presuming a basic familiarity with sentiment analysis principles and methodologies, as well as expertise with Python programming for developing sentiment analysis algorithms and assessing their effectiveness. Additionally, a thorough understanding of model evaluation and data pretreatment would be necessary. The project entails tackling sentiment analysis-related risks and difficulties, which can call for a certain amount of NLP and machine learning proficiency.

Variable Selection:

Here are the variables that will be gathered from the dataset:

a. tweet_id: Unique identifier for a tweet.
b. tweet_created: Date and time when the tweet was created.
c. tweet_extracted: Date and time when the tweet was extracted or retrieved.
d. text: The actual text content of the tweet.
e. lang: Language in which the tweet is written.
f. user_id: Unique identifier for the user who posted the tweet.
g. user_name: Name of the user who posted the tweet.
h. user_username: Username or handle of the user who posted the tweet
i. user_location: Location mentioned in the user's profile.
j. user_description: Description or bio provided by the user in their profile.
k. user_created: Date and time when the user's account was created.
l. user_followers_count: Number of followers the user has.
m. user_following_count: Number of accounts the user is following.
n. user_tweet_count: Total number of tweets posted by the user.
o. user_verified: Indicates whether the user's account is verified (True/False).
p. source: The source or platform from which the tweet was posted.
q. retweet_count: Number of times the tweet has been retweeted.
r. like_count: Number of times the tweet has been liked.
s. reply_count: Number of replies received by the tweet.
t. impression_count: Number of times the tweet has been seen or displayed.

Derived Variables: N/A

Assumptions and data limitations:
Data assumptions can be found in Section 2.1 of the paper.

Data Limitations:

When extracting tweets from Twitter, there are several data limitations to consider. These include API rate limits that restrict the number of requests, limited access to historical tweets beyond a certain timeframe, availability of sampled data rather than the complete set of tweets, potential data filtering and privacy restrictions, variable data quality and integrity, and limited user information provided by the API. Awareness of these limitations is crucial for ensuring accurate analysis and interpretation of the extracted tweet data.

4.0 Dependencies and Risks

Risk	Likelihood (based on historical data)	Delay (based on historical data)	Impact
Insufficient labeled training data	Medium	Low	High *Lack of labelled training data can have a significant negative effect on the project. To create correct models, training data is essential to sentiment analysis algorithms. Lack of labelled data can result in limited coverage of different sentiment expressions, which lowers the precision of the sentiment analysis findings. This might influence ChatGPT's overall effectiveness by misclassifying sentiments, leading to inaccurate interpretations. Due to the project's high impact, collecting enough and a variety of training data is essential for successfully developing the sentiment analysis models.*
Technical limitations of sentiment analysis algorithms	Low	Medium	High *Technical constraints on sentiment analysis algorithms may have a significant effect. Sentiment analysis algorithms might not always be able to correctly decipher complex emotions, sarcasm, or context-specific attitudes in tweets. As a result, sentiment analysis may be less accurate, which could influence the overall dependability of the results used to train ChatGPT. High impact indicates that overcoming technical obstacles and extending the algorithms' functionality are essential if accurate sentiment analysis and improved ChatGPT response accuracy are to be achieved.*
Inaccurate representation of	High	Low	Medium *The effects of inaccurately representing sentiment due to*

sentiment due to linguistic nuances			*language quirks can be moderate. It might be difficult to effectively detect and understand emotions when tweets are written in several languages, dialects, and cultures. The efficiency of sentiment analysis for ChatGPT may be impacted by inaccurate representations of sentiment that lead to misunderstandings and wrong analysis. For sentiment analysis to be more accurate and for ChatGPT to better comprehend user sentiment, linguistic nuance must be understood and addressed.*
Ethical considerations and potential biases in sentiment analysis	Medium	Medium	Medium *Potential biases and ethical considerations can have a medium-sized impact on sentiment analysis. Due to skewed training data or built-in biases, sentiment analysis algorithms may unintentionally induce biases. This could lead to unfair or biased analysis, which would harm ChatGPT's reputation and user experience. To maintain the integrity and trustworthiness of sentiment analysis results, it is critical to resolve ethical issues, ensure fairness, and minimize biases.*
Lack of user engagement and feedback for evaluation	Low	Low	Low *Lack of user participation and comments for evaluation may have minimal effects. User interaction and feedback are important for gauging the success of sentiment analysis and pinpointing areas where ChatGPT's response accuracy needs to be improved. However, a lack of user input or participation may make it more difficult to evaluate and improve the sentiment analysis procedure. Low impact means that, even if user participation and input are ideal, the project can still advance and meet its*

			goals using different evaluation techniques, if necessary.

5.0 Deliverable Timelines

Item	Major Events / Milestones	Description	Scope	Days	Date
1.	Kick-off / Formal Request	Class orientation and alignment	Gather project requirements, define roles and responsibilities, establish project objectives and timeline.	2	July 4, 2023
2.	Analysis Plan	Project set-up Definition of parameters and deliverables	Define the project scope, goals, deliverables, and milestones. Identify data sources and analysis techniques.	6	July 11 – July 17, 2023
3.	Exploratory Data Analysis	Review of dataset characteristic s and data preprocessing	Analyze and understand the dataset, perform data cleaning, handle missing values, and explore statistical properties.	6	July 18-July 24, 2023
4.	Model Development	Model specification, development and tuning	Design and develop sentiment analysis models using suitable algorithms and techniques. Optimize model parameters and evaluate performance.	13	July 25-August 7, 2023
5.	Governance	Model risk management and mitigation	Identify potential risks and challenges associated with the models. Implement risk management strategies and develop mitigation plans.	12	August 1-August 12,2023
6.	Documentation	Preparation of final draft and presentation materials	Document the project process, methodologies, findings, and insights. Prepare a final report and presentation materials for stakeholders.	14	August 1-August 14,2023

7.	Peer Feedback	Presentation to peers	Present the project progress, findings, and challenges to peers for feedback and constructive discussions.	1	August 13,2023
8.	Presentation to Releasing Authority	Presentation to faculty panel.	Present the project outcomes, methodology, and results to a faculty panel for evaluation and feedback.	3	August 14-16, 2023
9.	Revisions	Application of revisions and recommendations from the panel	Incorporate feedback and recommendations from the faculty panel into the project. Revise models, analysis, and documentation as necessary.	3	August 14-16,2023
10.	Portfolio	Incorporate capstone project on our respective portfolios	Include the completed project as a showcase of skills and achievements in individual portfolios.	1	August 18, 2023

6.0 References

Artificial intelligence in urban planning and design. (2022). In *Elsevier eBooks*. https://doi.org/10.1016/c2019-0-05206-5

Hauthal, E., Burghardt, D., Fish, C., & Griffin, A. L. (2020). Sentiment analysis. In *Elsevier eBooks* (pp. 169–177). https://doi.org/10.1016/b978-0-08-102295-5.10593-1

Li, L., Goodchild, M. F., & Xu, B. 2013. Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. cartography and geographic information science, 40: 61–77.

Liu, B. 2012. Sentiment Analysis and Opinion Mining. San Rafael: Morgan & Claypool Publishers.

Mobility patterns, big data and transport analytics. (2019). In *Elsevier eBooks*. https://doi.org/10.1016/c2016-0-03572-6

Pang, B., & Lee, L. 2004. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In: Proceedings of the 42nd annual meeting on Association for Computational Linguistics, 2004. Association for Computational Linguistics, p. 271.

Poria, S., Gelbukh, A., Cambria, E., Hussain, A., & Huang, G.-B. 2014. EmoSenticSpace: A novel framework for affective common-sense reasoning. Knowledge-Based Systems, 69: 108–123.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. 2011. Lexicon-based methods for sentiment analysis. Computational linguistics, 37: 267–307.

Serrano-Guerrero, J., Olivas, J. A., Romero, F. P., & Herrera-Viedma, E. 2015. Sentiment analysis: A review and comparative analysis of web services. Information Sciences, 311: 18–38.

Song, Z., & Xia, J. (Cecilia). (2016). Spatial and Temporal Sentiment Analysis of Twitter data. In C. Capineri, M. Haklay, H. Huang, V. Antoniou, J. Kettunen, F. Ostermann, & R. Purves

(Eds.), *European Handbook of Crowdsourced Geographic Information* (pp. 205–222). Ubiquity Press. http://www.jstor.org/stable/j.ctv3t5r09.20