

|  |
| --- |
| Sentiments Analysis On Twitter |
|  |
| May 17  AI Project  Authored by: Aizaz Ullah Khan Niazi |



# Introduction

|  |
| --- |
| Social media platforms have become a significant source of data for understanding public sentiment and opinion. Twitter, being one of the most popular platforms, provides an abundance of user-generated content that can be leveraged for sentiment analysis. This report explores the application of sentiment analysis techniques to analyze the emotions expressed on Twitter. |
|  |
| **Methodology:**  *To conduct sentiment analysis on Twitter, the following steps were taken:*   1. *Data Collection: A large sample of tweets was collected using Twitter's API. The data collection focused on a specific timeframe and specific keywords or hashtags related to the subject of interest. I manually can’t collect that data, so I imported this from a file called,*          1. *Preprocessing: The collected tweets underwent preprocessing steps to remove noise and irrelevant information. This involved removing URLs, mentions, special characters, and stop words. Additionally, tokenization, stemming, and lemmatization were applied to standardize the text.*   *#Checking null values*    *#droping some features*    *#removing substrings using re libraries#telling the sentimental types*  *#making a countplot*  *#understanding better using pieplot*  *#findings*   1. *Sentiment Analysis Techniques: Several sentiment analysis techniques were considered, including rule-based approaches, machine learning-based approaches, and hybrid approaches. In this analysis, a machine learning-based approach using a pre-trained model was employed. The model was trained on a large annotated corpus to classify tweets into positive, negative, or neutral sentiments.*     *Now frequent used positive sentimental words*    *Now frequent used negative sentimental words*    *Now frequently used neutral words*    *d. Evaluation: To assess the performance of the sentiment analysis model, a manually annotated subset of tweets was used as a gold standard. Metrics such as accuracy, precision, recall, and F1 score were calculated to evaluate the model's effectiveness.* |

Findings:

a. Sentiment Distribution: The sentiment analysis revealed the distribution of sentiments expressed on Twitter. The analysis indicated the percentage of tweets classified as positive, negative, or neutral, providing an overall understanding of the sentiment landscape.

b. Sentiment Trends: By analyzing the sentiment over time, it was possible to identify trends and fluctuations in public sentiment towards a specific topic. This information can be valuable for businesses, brands, and policymakers to gauge public opinion and adapt their strategies accordingly.

c. Influential Users and Hashtags: The sentiment analysis also identified influential users and hashtags associated with different sentiments. By tracking the sentiment of influential users or monitoring the impact of specific hashtags, organizations can gain insights into the factors shaping public sentiment.

Applications and Benefits:

Sentiment analysis of Twitter can offer several benefits and applications:

a. Brand Monitoring: Companies can monitor the sentiment towards their brand in real-time, identify issues or negative sentiment, and take proactive measures to address them.

b. Product Feedback and Market Research: Sentiment analysis enables businesses to extract feedback and opinions about their products or services, helping them understand customer needs and preferences.

c. Crisis Management: During crises or emergencies, sentiment analysis can help gauge public sentiment, identify potential issues, and respond promptly to mitigate negative sentiments.

d. Reputation Management: Public figures, politicians, and organizations can analyze Twitter sentiment to manage their reputation, understand public perception, and address concerns or criticisms.

Limitations and Challenges:

a. Contextual Understanding: Sentiment analysis algorithms may struggle to comprehend the nuanced meaning behind sarcasm, irony, or cultural references, leading to misclassification.

b. Data Quality: Twitter data often contains noise, spam, or irrelevant content, which can impact the accuracy of sentiment analysis results.

c. Sample Bias: Twitter users may not represent the entire population, and there may be biases in terms of demographics, geographic location, or user behavior, leading to skewed sentiment analysis results.

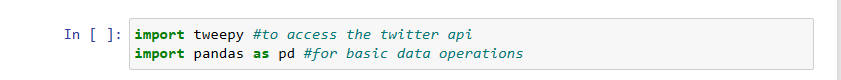
d. Privacy and Ethical Considerations: Analyzing tweets for sentiment raises privacy concerns, as user data is being processed without explicit consent. It is crucial to handle user data responsibly and adhere to ethical guidelines to protect user privacy.

e. Misinterpretation of Emotions: Sentiment analysis models might struggle to accurately interpret complex emotions or mixed sentiments expressed in tweets, leading to oversimplification or misinterpretation of the true sentiment.

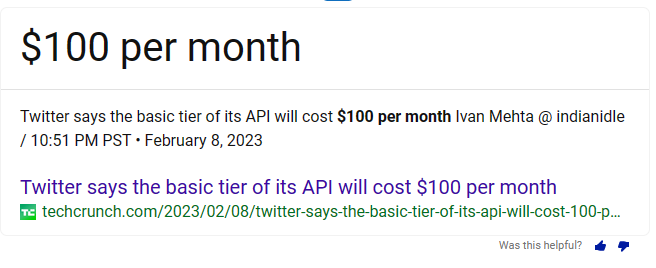
f. Real-Time Nature of Twitter: Twitter is a platform known for its fast-paced and real-time nature. Sentiment analysis models may face challenges in keeping up with the rapid influx of new tweets and capturing the evolving sentiment accurately.

*The future Expansion:*

Now for further, we can integrate with the live API and make the analysis of the tweets at any time as we want.



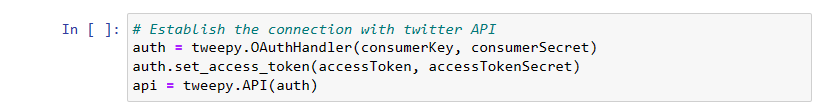
Now, as we tried to access the real api key and number but it costs almost 100$ per month. Here is the proof,



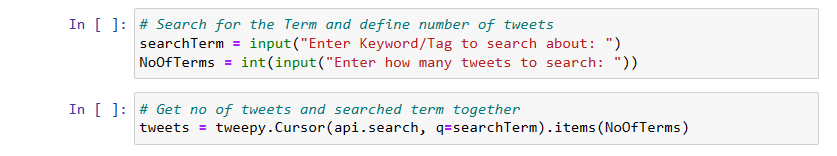
But anyway, we tried to move on by the assumed key,



Now using the tweepy,



And further for daily tweet numbers,



It can be further extended but as for limitations it didn’t, unfortunately.

*Conclusion*

Sentiment analysis of Twitter provides valuable insights into public sentiment and opinion. By leveraging the abundance of user-generated content, organizations can gain a deeper understanding of how people perceive brands, products, events, and public figures. Despite the limitations and challenges associated with sentiment analysis on Twitter, it remains a powerful tool for businesses, researchers, and policymakers to make data-driven decisions, tailor their strategies, and engage with the public more effectively.