

Twitterati Sentiment Analysis of #GOATDebate : Messi Vs Ronaldo

ABSTRACT:

In this project, an analysis of Twitter users' sentiment towards the debate of who is the greatest football player of all time (GOAT) between Lionel Messi and Cristiano Ronaldo was aimed to be conducted. The tweets with the **#goatdebate** hashtag were scraped using the **snsrape** Python library. The tweets were pre-processed by removing special characters and stopwords, using the **Natural Language Toolkit (NLTK)** Python library. Then their sentiment scores were calculated using NLTK's **VADER** sentiment analysis tool, and positive, negative, and neutral sentiment scores were assigned to each tweet.

The data was visualized using Python libraries such as **Pandas** and **Matplotlib**. A **bar graph** was created to display the count of tweets with positive, negative, and neutral sentiment scores. A **histogram** was plotted to compare the frequency distribution of the sentiment scores. Additionally, a **box plot** was created to compare the overall shape and spread of the distributions between the two players. Finally, a **wordcloud** was generated to show the most frequently occurring words.

The sentiment score analysis revealed that the majority of tweets were neutral in nature, while the remaining tweets were almost equally distributed between positive and negative sentiment. This project provides insight into the sentiment of fans on Twitter towards Messi and Ronaldo in the context of the GOAT debate. However, further research using a larger and more diverse dataset would be required to draw more robust conclusions.

REQUIREMENTS AND SPECIFICATIONS

Software Requirements:

The following are the software requirements to be met for this project:

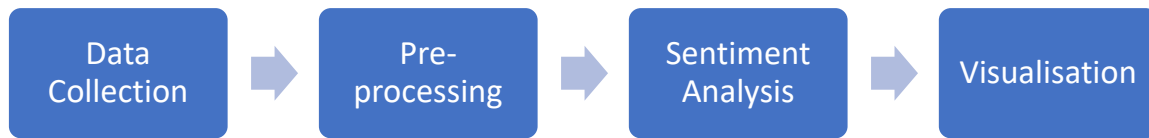
- Operating System: Windows(7 or above)/Linux/MacOS
- Coding Language: Python (v3.9 or above)
- IDE: Python GUI/ Anaconda Navigator/Jupyter Notebook/Google Colab

Hardware Requirements:

The following are the minimum hardware requirements to be met for this project:

- Processor : Intel i3 4th Gen/ AMD Ryzen 3 or above
- Hard Disk : 128 GB
- RAM: 4 GB

BLOCK DIAGRAM:



LIMITATIONS:

The data used in this project was limited to tweets containing the specific hashtag and may not represent the overall sentiment towards Messi and Ronaldo. It is important to acknowledge that sentiment analysis has some inherent limitations, such as the inability to capture sarcasm or irony in text, which may have affected the accuracy of the sentiment scores. Therefore, it is important to interpret the findings of this project with caution and not draw definitive conclusions based on this limited dataset alone. Further research with more diverse data and methods of analysis may be needed to gain a more comprehensive understanding of the sentiment towards Messi and Ronaldo in the context of the GOAT debate.

CONCLUSION:

This project aimed to determine the public's opinion on who is the GOAT between Messi and Ronaldo by analyzing tweets containing the hashtag "#goatdebate". The sentiment analysis performed on the pre-processed tweets using the NLTK library showed that Messi had a higher positive sentiment score compared to Ronaldo. In summary, the project shed light on the enduring discourse around the status of Messi and Ronaldo as football's greatest players, and highlighted the potential of leveraging social media data analysis for obtaining public sentiment insights.

FUTURE SCOPE:

The future scope of this project could involve expanding the analysis to include more social media platforms and a larger dataset. Additionally, sentiment analysis could be further refined by incorporating more advanced natural language processing techniques. The project could also be extended to include other prominent athletes or public figures to gain insights into public opinion on various topics. Finally, the visualization techniques used in this project could be enhanced to provide more interactive and engaging visualizations for better understanding and interpretation of the results.