Decoding Ronaldo's Twitter Chronicle: A Comparative Analysis of Tweet Spikes and Sentiment Scores

Ariel Lutati   Anay Contractor   Gauri Barar

December 9, 2023

Project Description and Data Collection

Our project aims to qualitatively discern patterns in Twitter discussions about Cristiano Ronaldo. We will delve into a dataset comprising over 400 thousand tweets sourced from Kaggle, covering pivotal moments in Ronaldo's career during the Twitter era. These tweets encapsulate events such as transfers, victories, Balon D'Or wins, and Champions League triumphs.

Objectives and Interests

Our primary objective is to compare spikes in tweet volume and their overall sentiment scores to significant moments in Ronaldo's career. Ariel will focus on refining the dataset by cleaning and utilizing Regex to bring nuance to sentiment analysis, especially in cases where words like "nasty" may convey positive sentiments. Gauri will run the sentiment analysis, creating distributions of sentiment scores, while Anay is responsible for resampling using bootstrapping to generate reliable point estimates of sentiment scores over time.

Class Material for Analysis

We will leverage the power of `R` for data cleaning, sentiment analysis, and bootstrapping. Ariel's expertise will be pivotal in preprocessing the data and introducing context to sentiment analysis. Gauri's role involves running the sentiment analysis and creating distributions of sentiment scores. Anay's expertise lies in resampling with bootstrapping to ensure robustness in our sentiment score estimates.

Responsibilities

- **Ariel Lutati:** In charge of data cleaning, Ariel will use Regex to assess sentiment, considering combinations of words and adding context to the sentiment analysis.

- **Anay Contractor:** Responsible for resampling with bootstrapping, Anay will generate point estimates of sentiment scores from the bootstrap sample, ensuring statistical reliability.

- **Gauri Barar:** Gauri will run the sentiment analysis, creating distributions of sentiment scores. She will contribute to the interpretation of sentiment trends over time.

Conclusion

The conclusion of our project will be presented on the day of the project presentation. We anticipate uncovering valuable insights into the correlation between tweet spikes and sentiment scores on Twitter regarding Cristiano Ronaldo, enriching our understanding of the dynamics surrounding key moments in his remarkable career.