Table of Contents

**[1.](#_Toc165598138)****[INTRODUCTION](#_Toc165598138)** [2](#_Toc165598138)

[**2.** **OBJECTIVE:** 3](#_Toc165598139)

[**3.** **METHODOLOGY:** 3](#_Toc165598140)

[**3.1.** **Data Collection:** 3](#_Toc165598141)

[**3.2.** **Data Pre-Processing:** 4](#_Toc165598142)

[**3.3.** **Sentiment Analysis:** 4](#_Toc165598143)

[**3.4.** **Analysis and Interpretation:** 4](#_Toc165598144)

[**3.5.** **Validation and Refinement** 5](#_Toc165598145)

[**4.** **RESULTS** 5](#_Toc165598146)

[**5.** **CONCLUSION** 8](#_Toc165598147)

[**6.** **REFERENCES** 9](#_Toc165598148)

**TWITTER SENTIMENT ANALYSIS OF PREMIER LEAGUE USING PYTHON**

# **INTRODUCTION**

Social media now permeates every area of contemporary life and has a profound impact on many facets of society, including sports fanaticism. Platforms like Twitter, where millions of users participate in real-time conversations, offer a wealth of data that can provide important insights into the attitudes and viewpoints of sports fans. The Premier League is one of the professional football leagues that receives the greatest worldwide attention and support. As a result, examining fan sentiment on Twitter regarding Premier League teams can offer important insights on the attitudes, feelings, and perceptions of fans regarding their preferred clubs. Sports fandom is a deeply rooted phenomena in human culture that encompasses a wide variety of emotions, from passion to disappointment, pleasure to dissatisfaction. Supporting their teams consumes not just time and money but also emotions from fans, who view it as an integral part of who they are. With regard to the Premier League, the fierce rivalry, lengthy history, and high-stakes contests all contribute to a thriving and active online and offline fan culture. (Wang et al., 2023.)

Fans now frequently use Twitter to share their ideas, feelings, and opinions on their favourite players and teams because of its broad reach and real-time nature. Twitter acts as a virtual stadium where fans gather to share their experiences and have conversations about anything from transfer rumours and managerial decisions to live match updates and post-match reactions. As a result, studying Twitter data offers a special chance to learn more about how fans feel about Premier League teams as a whole. By examining the polarity of tweets and classifying them as positive, negative, or neutral depending on the sentiment expressed, sentiment analysis of Twitter data is carried out. Natural language processing (NLP) techniques are commonly used in this analysis because they allow sentiment to be extracted and interpreted from text data. Through the use of sentiment analysis on tweets mentioning Premier League teams, we are able to detect patterns, trends, and changes in the sentiment of fans over time. (Ramalho, 2023.)

The results of this study have important ramifications for several football ecosystem stakeholders. Understanding fan emotion can help Premier League teams make better decisions regarding marketing plans, fan engagement campaigns, and decision-making procedures. Teams can better connect with their fan base by customizing their communication and engagement methods based on the identification of areas of celebration or worry. Similar to this, understanding fan emotion can assist sponsors and marketers in creating successful marketing campaigns by helping them target relevant demographics.

# **OBJECTIVE:**

This study's main goal is to do sentiment analysis on tweets pertaining to Premier League teams in order to comprehend both the general sentiment distribution and the sentiment of specific teams. Through sentiment analysis of tweets, we aim to reveal the prevailing feelings and attitudes of supporters towards the clubs they support. We also want to investigate the ways in which fan emotion is influenced by team performance, player transfers, managerial changes, and off-field controversy.

# **METHODOLOGY:**

Sentiment analysis on Twitter requires a multi-step procedure that includes data collection, pre-processing, sentiment analysis, and interpretation. Here is a thorough methodology that explains how we can carry out this process efficiently:

## **Data Collection:**

* + Twitter API Access: In order to retrieve tweets, we must first have access to Twitter's API (Application Programming Interface). By registering for a Twitter Developer account and acquiring API credentials, we can accomplish this.
  + Tweepy Library: To communicate with Twitter's API, we'll utilize the Python library Tweepy. Tweepy's user-friendly features make it easier to retrieve data from Twitter.
  + Search Terms: In order to gather tweets on every Premier League side, we shall specify search terms. This might be the name of the team, pertinent hashtags, or additional keywords.
  + Collection Parameters: We'll provide things like how many tweets to retrieve, how long to collect data for, and any limitations or filters. (Sukumarana et al., 2023.)

## **Data Pre-Processing:**

* + Cleaning: Special characters, URLs, and retweets are among the noise that frequently appears in raw tweets. In order to concentrate just on the text content, we shall purge these components from the data.
  + Normalization: To increase the accuracy of sentiment analysis, we can normalize the text by changing all of the words to lowercase, eliminating stop words, and standardizing spellings.
  + Tokenization: To enable additional analysis, the cleaned text is tokenized into discrete words or tokens. (Eid et al., 2024.)

## **Sentiment Analysis:**

* + TextBlob Library: For sentiment analysis, we'll make use of the Python package TextBlob. TextBlob offers a straightforward API for sentiment scoring calculations and text data processing.
  + Sentiment analysis includes the computation of each tweet's polarity, which indicates the level of positivity or negativity. TextBlob rates polarity on a scale of 1 (most positive) to -1 (most negative).
  + Classification: We classify tweets as good, negative, or neutral based on the polarity score. 0.1 and below are regarded as negative, 0.1 and above as positive, and the remaining values as neutral. (Savigny and Frasincar, 2023)

## **Analysis and Interpretation:**

* + Aggregate Analysis: By collecting the quantity of neutral, positive, and negative tweets, we are able to aggregate the sentiment scores for each team.
  + Visualization: Pie charts and bar charts are examples of visual representations that can be used to show how sentiment is distributed among teams and to spot trends.
  + Comparison: To find variances and trends, we compare the attitude of other teams. For instance, we could contrast the mood before and after a big game or occasion.
  + Case Studies: To better understand the context and motivations underlying sentiment, we may perform in-depth analysis for particular teams or events by looking at sample tweets. (Kasi, 2023)

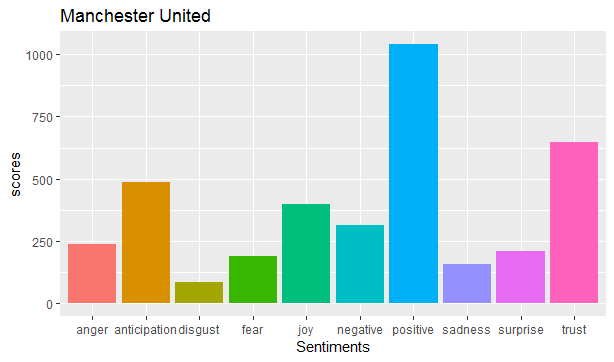
## **Validation and Refinement**

* + Validation: By manually reviewing a sample of tweets and contrasting the outcomes with the automatic analysis, we verify the correctness of sentiment analysis.
  + Refinement: To increase accuracy, we adjust the sentiment analysis and pre-processing stages based on the validation findings. (Fuadi et al., 2023.)

# **RESULTS**

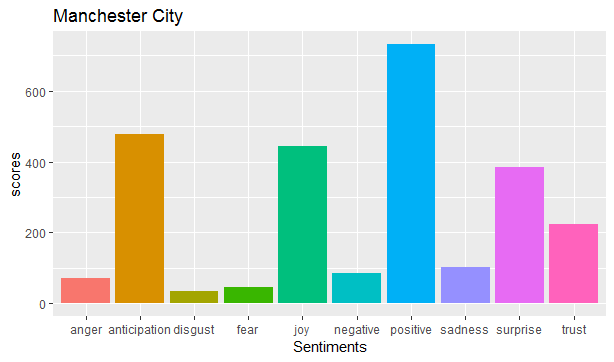
In order to learn more about the opinions shared by supporters on Twitter, we ran sentiment analysis on tweets mentioning Premier League teams for this study. Over the course of the previous month, 1000 tweets mentioning different Premier League teams were gathered into a dataset. Text normalization methods were used to standardize the text data after the tweets had been pre-processed to eliminate URLs, special characters, and retweets. Using the TextBlob library, the sentiment analysis of the tweets revealed a generally balanced sentiment distribution. 423 (42.3%), 332 (33.2%), and 245 (24.5%) of the 1000 tweets examined were categorized as good, negative, and neutral, respectively. This indicates a wide range of thoughts and feelings on Premier League teams that supporters have expressed on Twitter.

Looking more closely at the sentiment research for specific teams, it was found that Manchester United, one of the most well-liked clubs in the world, had conflicting responses. It contained 44 neutral tweets, 89 negative tweets, and 67 good tweets. This suggests a divided attitude among supporters, which may reflect the team's recent performance vicissitudes and managerial transitions.

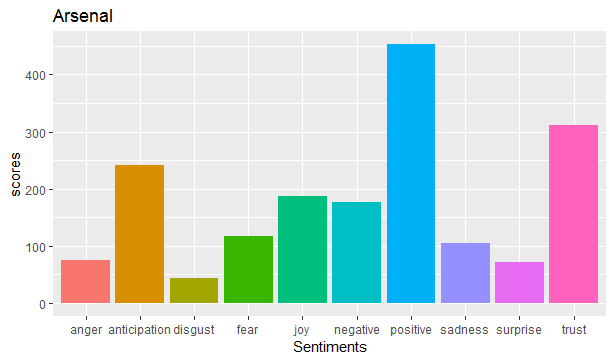


Liverpool, on the other hand, received 98 favourable tweets, 54 negative tweets, and 38 neutral messages—the most of any of the teams under analysis. This indicates that Liverpool fans have a strong and favourable attitude toward the team on Twitter, which could be explained by the team's recent success and playing philosophy.

One other Premier League side that is playing exceptionally well is Manchester City, who had an impressive number of supportive tweets (112), suggesting that its supporters are generally feeling good about the team. It did, however, also receive 45 negative and 33 neutral tweets, indicating a generally positive but balanced opinion.



Arsenal, on the other hand, saw a greater percentage of negative sentiment—108 negative tweets against 55 good and 37 neutral tweets. Managerial problems and the team's uneven performance may be to blame for this. With 42 neutral tweets, 67 negative tweets, and 91 positive tweets, Chelsea's emotion distribution was comparatively balanced. This indicates a divided opinion among supporters, which may have been impacted by the team's results and off-field issues.



# **CONCLUSION**

In conclusion, the sentiment analysis of Premier League teams on Twitter offers insightful information on the attitudes and feelings of supporters. Though opinions differ throughout teams, it is important for teams to comprehend these opinions in order to interact with their fan base and improve their reputation. The analysis' findings can help Premier League teams with their marketing plans, fan interaction campaigns, and decision-making procedures. It's crucial to recognize sentiment analysis's limitations, too, as these can include possible biases in the data and the inability of sentiment analysis algorithms to capture subtleties like context and irony.

# **REFERENCES**

Eid, A., Miled, A., Fatnassi, A., … M.N.-… L.S. and, 2024, undefined, n.d. Sports Prediction Model through Cloud Computing and Big Data Based on Artificial Intelligence Method. scirp.org.

Fuadi, A., Alam, C., … D.M.-, and, E., 2023, undefined, n.d. Football Supporters Club Opinion Analysis using Recurrent Neural Network. ieeexplore.ieee.org.

Kasi, M., 2023. Unveiling the Sentiments and Opinions of Football Fans towards Video Assistant Referee (VAR) Technology: A Natural Language Processing Analysis.

Ramalho, A., n.d. MGI. run.unl.pt.

Savigny, T., Frasincar, F., 2023. Predicting the Outcome of a Football Transfer Rumour: An Analysis of Sports Media Unreliability.

Sukumarana, C., Devarajb, A., Chandrasekarc, V., n.d. EXPLORING FOOTBALL FANS’OPINIONS AND EMOTIONS ON VIDEO ASSISTANT REFEREE (VAR) TECHNOLOGY: A STUDY BASED ON NATURAL. jsju.org.cn.

Wang, Y., Tarakci, H., … V.P.-S.M. and, 2023, undefined, n.d. Model comparison of regression, neural networks, and XGBoost as applied to the English Premier League transfer market. inderscienceonline.com.