Assessment 3: WebCrawler and NLP System

MA5851: Data Science Master Class 1

Name: Rijo Kuruvilla

Student ID: 14171011

I. Abstract

Soccer, deeply ingrained in society and beyond, thrives in the English Premier League (EPL), captivating fans of all ages. This report investigates the profound impact of Video Assistant Referee (VAR) implementation on referees, teams, and the Premier League Football Association (FA). VAR, introduced to assist on-field referees in critical decision-making, such as reviewing penalties, disallowed goals, and other pivotal moments, is scrutinized within the context of football officiating. This study delves into the realm of football officiating with VAR's assistance, aiming to classify VAR decision sentiments as FOR, AGAINST, or NEUTRAL, discerning whether these decisions favor the Home team or Away Team. Data from EPL matches spanning 2019 to 2023 is meticulously collected, employing natural language processing (NLP) techniques for information preprocessing. Tasks include text pre-processing, incorporating stop word removal, lemmatization, and TF-IDF analysis. Subsequently, machine learning algorithms, such as Logistic Regression, Random Forest, Support Vector Machines and Gradient Boosting, are deployed to categorize and assess sentiments related to VAR descriptions and outcomes. The study rigorously evaluates the performance of each machine learning algorithm, utilizing metrics such as accuracy, precision, recall, and F1 score. The anticipated findings aim to provide valuable insights for EPL teams, referees, and the FA, pinpointing areas for improvement in VAR application and enhancing the overall football experience for players, fans, and stakeholders alike.

II. Overview of the Issue

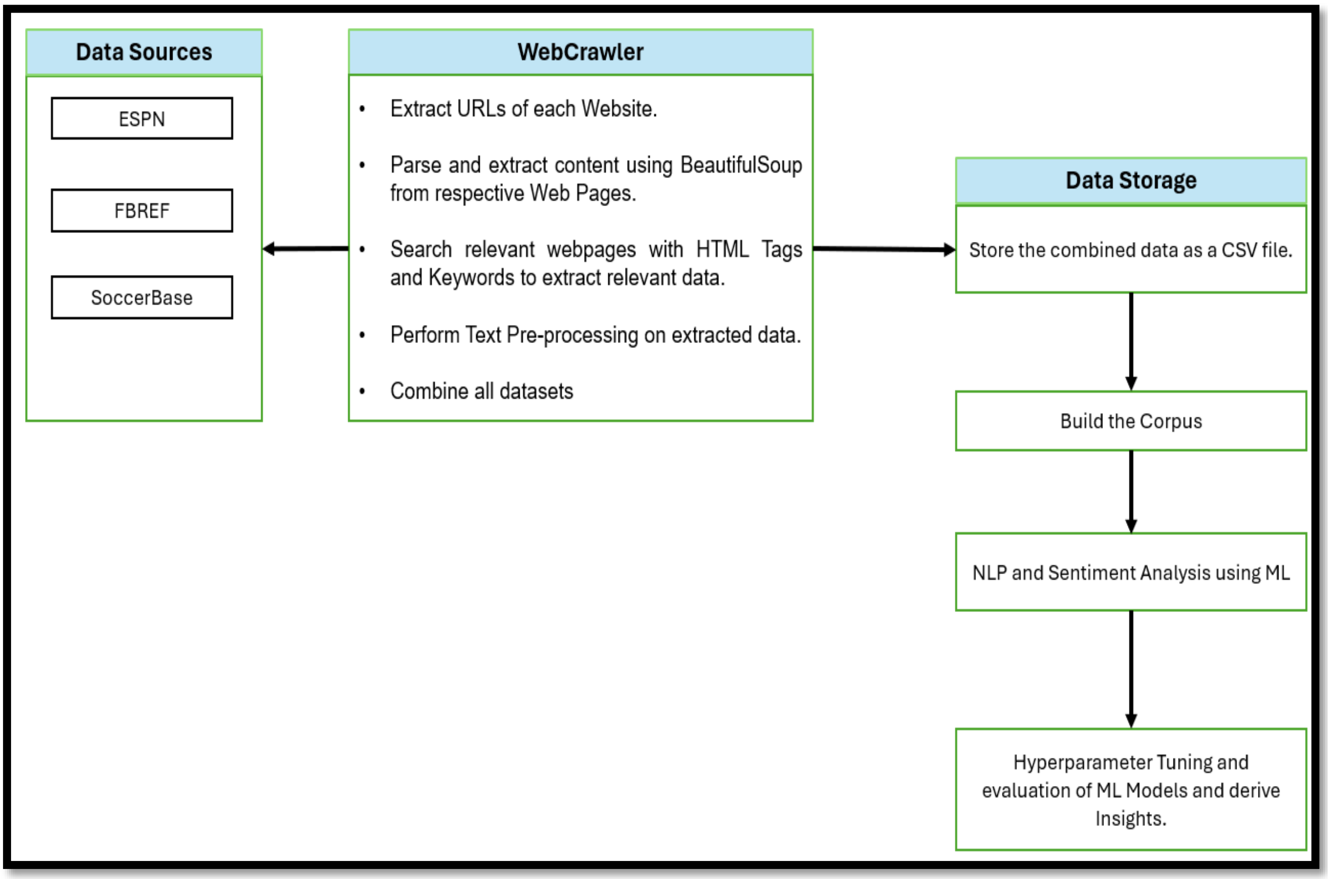
Created by the Royal Netherlands Football Association, VAR made its debut in The 2018 FIFA World Cup (Rowe-Willcocks et al., 2023). It was introduced to assess goals and offences leading up to a goal, penalty decisions and red card decisions and cases of mistaken identity . This report delves into the impact of VAR on match outcomes, scrutinizing incidents that VAR was called upon to review. Figure 1 displays the workflow of the WebCrawler and the NLP workflow.

Sentiment Analysis, also known as Opinion mining, is a branch of research that examines how people behave, process, and feel about particular issues, events, organizations, goods, services, or their corresponding characteristics (Liu, 2012, as cited in Liaqat et al., 2022). Applied here to the context of EPL matches, Sentiment Analysis aims to categorize VAR decisions as either FOR, AGAINST, or NEUTRAL for the respective teams. The objective is to provide practical insights for referees, team managers, coaches, and the Football Association (FA) to make informed decisions.

Navigating through the vast amount of match details and VAR incident reports requires efficient handling. To efficiently process the extensive data, various NLP techniques are applied, including text cleaning, lemmatization, and Term Frequency Inverse Term Document Frequency (TF-IDF). These techniques not only enhance the quality of the data by reducing noise but also contribute to the accuracy of the results, ensuring a robust foundation for subsequent analyses.

Figure 1

WebCrawler, NLP and ML Workflow


Data Sources
ESPN
FBREF
SoccerBase
WebCrawler
• Extract URLs of each Website.
• Parse and extract content using BeautifulSoup from respective Web Pages.
• Search relevant webpages with HTML Tags and Keywords to extract relevant data.
• Perform Text Pre-processing on extracted data.
• Combine all datasets
Data Storage
Store the combined data as a CSV file.
Build the Corpus
NLP and Sentiment Analysis using ML
Hyperparameter Tuning and evaluation of ML Models and derive Insights.

Machine Learning models, such as Logistic Regression, Random Forest Classifier, and Gradient Boost Classifier, are then employed to classify sentiments from VAR incident descriptions. Evaluation metrics, including Accuracy, Precision, Recall, and F1 score, play a crucial role in assessing the performance of these models. This comprehensive evaluation aids in determining the most effective classifier, contributing to a thorough sentiment analysis.

The implementation of VAR in the EPL since 2019 has garnered both criticism and acclaim. Its correct application has the potential to assist referees, rectify errors, and offer valuable insights for team managers assessing player performance. End-of-season reviews conducted by FA representatives help refine VAR's role in the league, addressing challenges and ensuring optimal use.

For instance, a team facing relegation may experience an incorrectly disallowed goal. VAR becomes instrumental in rectifying the decision, potentially influencing the team's final standings, and saving them from relegation. This study contributes valuable insights, facilitating ongoing improvements in VAR application for the benefit of players, fans, and stakeholders in the English Premier League, fostering a more informed and efficient football experience.

III. WebCrawler:

- ***Websites Consumed and Website/Data Copyright considerations***

In conducting web crawling analysis, we selected three key websites to gather football-related statistics comprehensively. ESPN (Johnson, 2019-2023), FBREF (FBREF, 2019-2023), and SoccerBase (SoccerBase, 2019-2023) contributed distinct datasets, enriching the analytical framework.

FBREF, a valuable resource, supplied detailed team positions, scores, and fixture information from 2019 to 2023. The decision to use FBREF was based on its wealth of data and consideration of copyright issues. FBREF allows web scraping but enforces rate limiting for responsible usage, emphasizing ethical data extraction (FBREF, n.d.).SoccerBase played a pivotal role in extracting referee statistics, covering information on referees, games officiated, and card frequencies. Its selection was driven by the relevance of referee data to the broader analysis. ESPN contributed by providing VAR incident reports, offering insights into contentious decisions during matches. Considering copyright implications, a careful approach was taken aligning with each website's policies. While FBREF explicitly allows scraping with restrictions, ESPN and SoccerBase lacked clear guidelines, necessitating a cautious interpretation of fair use principles. This transparent approach ensures ethical considerations guide the data collection process.

All three websites provide rich natural language data through match statistics and metadata. Each has a glossary of terms simplifying the retrieval of pertinent information. For example, both FBREF and ESPN offer team names, player names, and VAR incident descriptions alongside numerical insights. The contextual information from this natural language data enhanced the analysis, enabling a deeper understanding of the football domain. Player-specific data and detailed match descriptions facilitated a more complex analysis, revealing broader relationships within the dataset.

The choice of these websites aligns with the intricacies of the data science scenario and its ethical implications. The interaction of these datasets underscores the need for a comprehensive approach in navigating web crawling and data analysis, addressing the investigation's complexities.

- ***Methodology and Limitations of applying the web crawler/scraper***

Python libraries such as Requests and Beautiful Soup were used to extract data from FBREF and Soccer Base. Requests is a powerful tool for initiating HTTP requests and handling responses whereas, Beautiful Soup coupled with Requests provides an easy-to-use interface for parsing HTML and XML documents (ZenRows, 2023). Beautiful Soup is easy to use, flexible and robust, which makes it a go-to tool for web scrapping. Selenium WebDriver on the other hand is an open-source browser automation tool often used for web scraping. Like Beautiful Soup, it is not only easy to use but also aids in scraping web pages dynamically, mimicking human action (ZenRows, 2023). Selenium interacts with the web page's JavaScript code and waits for elements to load before scraping the data.

The framework of each website and the content to be retrieved can significantly affect a web crawlers' efficiency and effectiveness. When pulling data from FBREF and SoccerBase, customized functions were created to navigate to the specific URL's, parse the HTML content present in those URL's, locate each season's standings, scores and fixtures, referee statistics and extract those specific column headers and data rows. For

ESPN, Selenium WebDriver was utilized to scrape VAR incident details for multiple Premier League seasons. Given the absence of explicit links on the current page to navigate to the next season, individual URLs were created for each season. Selenium efficiently located web elements using XPATH, enabling the extraction of data for each season separately.

In extracting data from FBREF and SoccerBase, various measures have been implemented to ensure the quality and optimization of the respective web crawlers. Robust exception handling, incorporating try and except blocks, addresses potential errors in both cases, enhancing the resilience of the web crawlers in handling unforeseen situations such as HTTP errors, connection issues, timeouts, and general request exceptions. Introducing a sleep timer (time delay) of 15 seconds after each data extraction strategically mitigates the risk of IP blocking, aligning with the Website Copyright permissions discussed earlier, and optimizing the efficiency of the web crawler. Additionally, in the SoccerBase code, the script validates the existence of the target table on the webpage before proceeding with data extraction, further contributing to the reliability and quality assurance of the web crawler when scraping referee statistics. These collective measures reflect best practices, ensuring reliable, efficient, and optimized data extraction from FBREF and SoccerBase.

For Selenium, a wait time (waittime) of 30 seconds and a sleep timer of 0.5 seconds is induced contributing to the pacing of the web scraping process, allowing for appropriate responses from both the browser and server. Secondly, the WebDriver options are configured, including arguments such as 'no-sandbox' and 'ignore-certificate-errors,' enhancing the reliability and compatibility of the web scraping process. These measures collectively ensure a robust and optimized data extraction experience from the specified URLs.

- ***Methodology of storing harvested data***

Storing the harvested data is an essential characteristic of any NLP analysis post web-scrapping. The way a file is stored can affect the scalability and the accessibility of the data. For analyzing this issue, the harvested data is stored in a Comma Separated Values (CSV) file format. CSV files are widely used by many organizations, have a straightforward file structure, and can be compressed precisely and easily (Gepard, 2023), making it a go-to option for storing data from web pages.

IV. Data Wrangling

- ***Cleaning, normalization, feature extraction of the sourced data*.**

Data wrangling is a critical phase in data preparation, encompassing cleaning, normalization, and feature extraction to ensure its suitability for analysis. Extracting data from various websites introduced challenges, notably inconsistencies in team abbreviations, such as 'Manchester United' being spelled as 'Manchester Utd' on one site and 'Manchester United' on another. To maintain transparency and consistency, each team underwent meticulous mapping to its original name.

To refine data quality, a systematic approach was adopted. Standardizing all words to lowercase promoted uniformity, and regular expressions (regex) played a crucial role in removing HTML tags and replacing non-alphabetic characters with spaces. This regex application effectively cleansed the text of extraneous characters like brackets, commas, parentheses, square brackets, and slashes, resulting in a cleaner and more standardized dataset. Tokenization segmented the text into individual words or tokens, a critical step for subsequent analysis. Additionally, lemmatization was employed to obtain the base form of words, considering their meanings, resulting in a more intricate and accurate representation (Gillis, 2023).

The target variable, labeled 'VAR Results,' includes three sentiment categories: 'FOR,' 'AGAINST,' and 'NEUTRAL.' During additional exploratory data analysis (EDA), it was discovered that only two instances of 'NEUTRAL' VAR incidents were recorded. Upon further examination, these neutral results were reclassified as 'FOR' sentiments to ensure a more accurate representation of sentiment aligned with the overall goal of sentiment analysis. This adjustment aimed to enhance the accuracy of sentiment representation, aligning with the overarching goal of refining categorizations for improved precision and reliability in sentiment analysis.

Feature Extraction identifies relevant features in the text data for predicting sentiments. The TF-IDF vectorizer was employed to convert the pre-processed text corpus into numerical vectors. TF-IDF, widely used in information retrieval and text mining, evaluates the relationship for each word in the collection of documents (Kim & Gil, 2019). Hyperparameters for TF-IDF, such as maximum features, minimum, and maximum document frequency, were tuned to enhance model performance. Finally, the dataset was randomly divided, with 75% for training and 25% for testing, using stratified sampling to maintain consistent category proportions (Kim & Gil, 2019).

This diligent approach in data wrangling and sentiment analysis aims for precision and reliability, with each step contributing to the overall quality of the analysis. The adoption of lemmatization and strategic adjustments in sentiment categorization aligns with the overarching goal of refining data for improved analysis outcomes. The application of the TF-IDF vectorizer reflects a comprehensive effort to extract meaningful features from the text, optimizing the dataset for sentiment analysis. The random division of the dataset, coupled with stratified sampling, ensures robust model training and testing, enhancing the overall quality of sentiment predictions.

- ***Summary and visualization of the harvested data. Preliminary EDA is acceptable in this section as well***.

Descriptive Statistics of the Sample

Table 1

Descriptive statistics of the Original Datasets

Dataset Information	**Values / Statistics**
Column names	'Year', 'Home_Team', 'Away_Team', 'VAR_Description', 'Time_of_Incident', 'VAR_Result', 'Home_Team_Stadium', 'Referee', 'Score', 'Top_Scorers', 'Goalkeepers', 'Notes_combined', 'Referee_Origin', 'Total_Yellow_Cards_by_Referee', 'Total_Red_Cards_by_Referee', 'Total_Games_Officiated_by_Referee'
Target Variable (Sentiments)	VAR_Results(FOR , AGAINST, NEUTRAL)
Total Number of Home Teams	25
Total Number of Away Teams	26
Total Number of Unique Referees	31

Descriptive Statistics of the Corpus: Corpus here indicates the combined data merged and pre-processed from all three websites. Table 2 displays the descriptive statistics of the corpus.

Table 2

Descriptive statistics of the Corpus

Corpus Information	**Values / Statistics**
Dimensions	545 rows x 65 columns
Top 5 words	('league', 399), ('goal', 283), ('stadium', 273)
Document length	Max: 438, Min: 248, Mean: 321.411, Std Dev: 40.364
Average Word Length/ VAR Description	59.89

Sample Visualizations

Figure 2 illustrates the frequency distribution of VAR Incidents and their outcomes from the 2019 to 2023 seasons. Notably, the 2021/2022 season recorded the highest number of incidents, with 86 results in favor (FOR) and 75 results against the teams under consideration. In Figure 3, a bar plot depicts the frequency distribution of sentiment categories. The sentiments 'FOR' and 'AGAINST' are closely balanced, with 280 instances favoring and 263 instances against. Notably, only two instances of Neutral Incidents were recorded. In football, Neutral Incidents do not occur, as decisions are made either for or against the concerned team. As detailed in Section IV, these 'neutral' instances have been appropriately reclassified as 'FOR' to ensure a more accurate representation of sentiment. Moving to Figure 4, referee statistics are presented across all seasons. Referee Anthony Tailor stands out with an average of 150 games officiated, and a combined total of 536 yellow and red cards given. In contrast, Thomas Bramall and Tony Harrington officiated the fewest games and displayed the lowest number of cards, suggesting their status as newcomers to the Premier League officiating scene.

Figure 2

VAR Results over the years

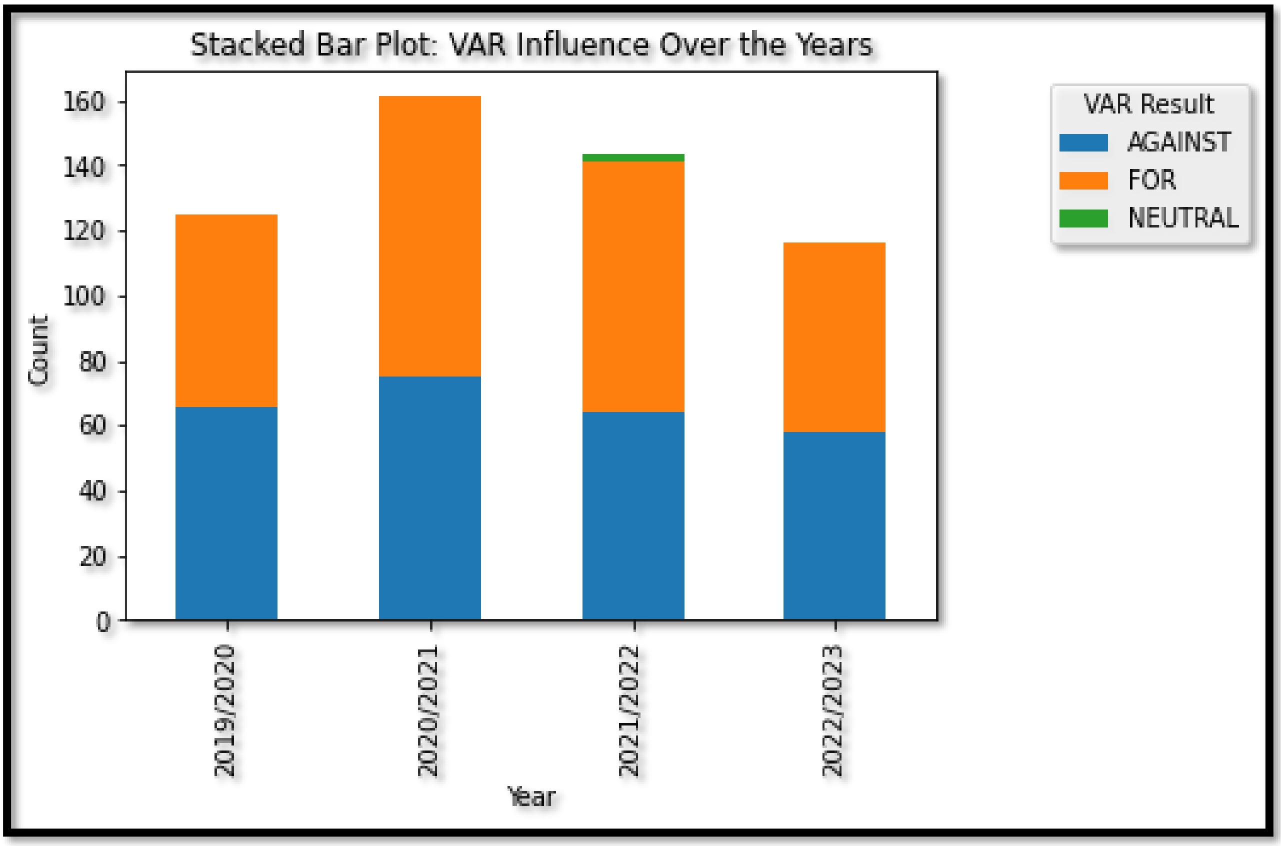


Figure 3

Frequency results of the VAR Sentiment Categories

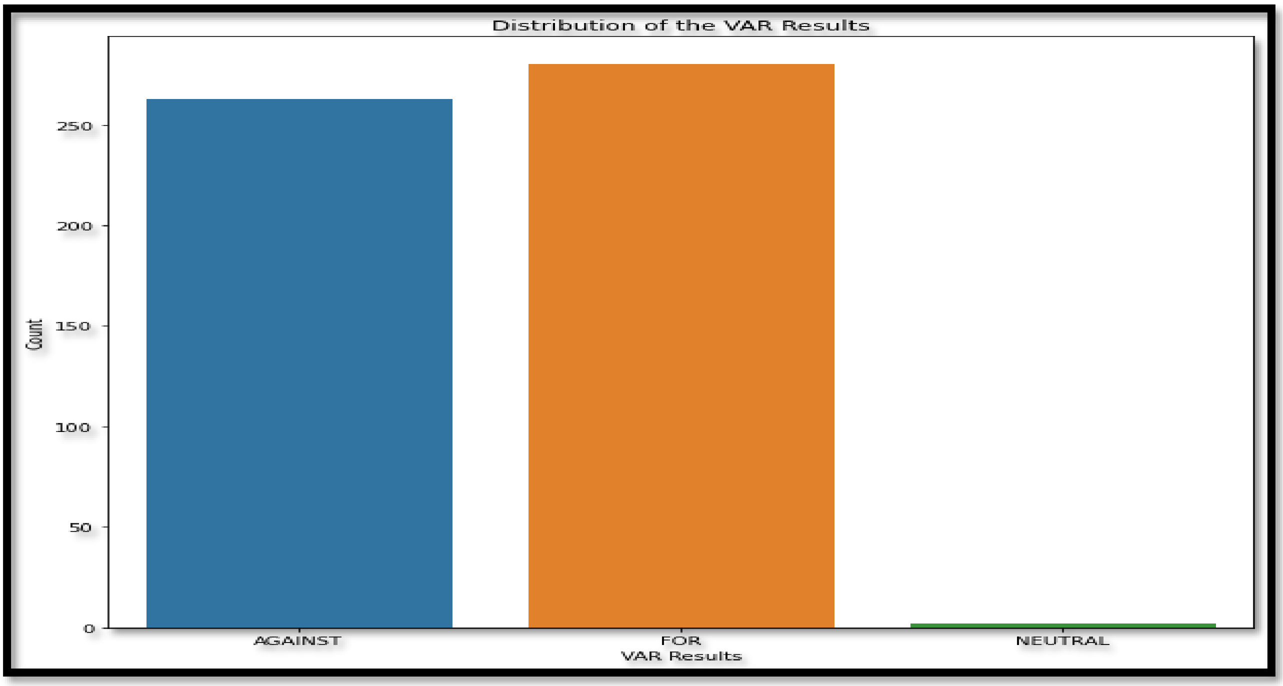
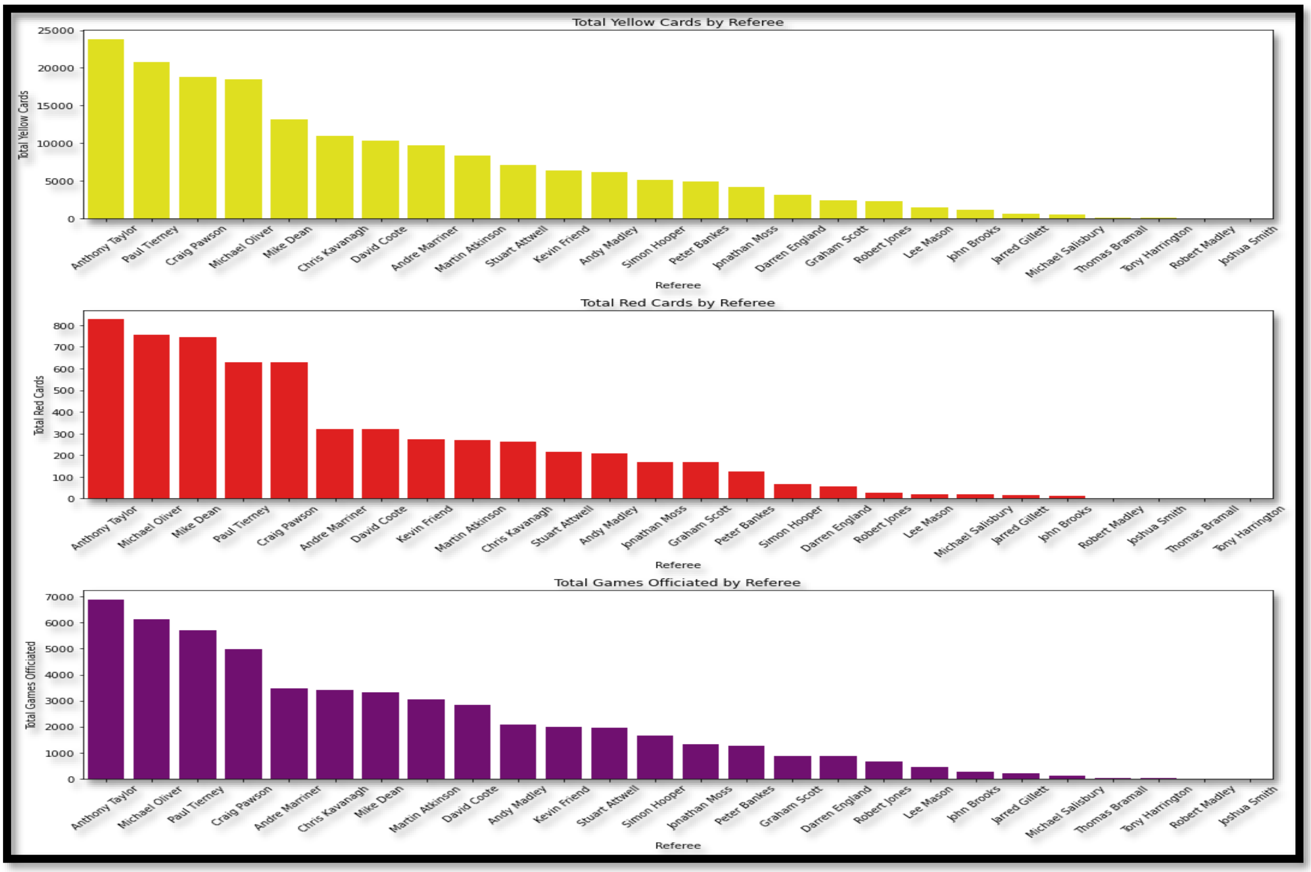


Figure 4

Referee statistics from Seasons 2019 - 2023



Corpus Visualizations

The histogram shown in Figure 5 provides a visual representation of the distribution of VAR Incident descriptions. The histogram, demonstrating a right-skewed distribution, highlights the significant variability in the dataset even though the average length of each VAR description is roughly 321 words. This skewness suggests that there aren't many descriptions that are noticeably longer than average; these could be examples of descriptions that are more intricate or detailed. These examples could be further analyzed to improve the performance of VAR in the upcoming seasons.

Figure 6 presents the Top 30 words in the corpus, highlighting 'league' (399), 'goal' (280), and 'stadium' (250) as the most frequent. Concurrently, Figure 7 showcases a comprehensive word cloud representing sentiments 'FOR' and 'AGAINST' in the context of VAR Description details. Given the balanced nature (as displayed in Figure 3) of these sentiments, a singular word cloud suffices, as the vocabulary associated with VAR is expected to overlap between both perspectives.

Figure 5

Frequency Distribution of the VAR Incident Descriptions

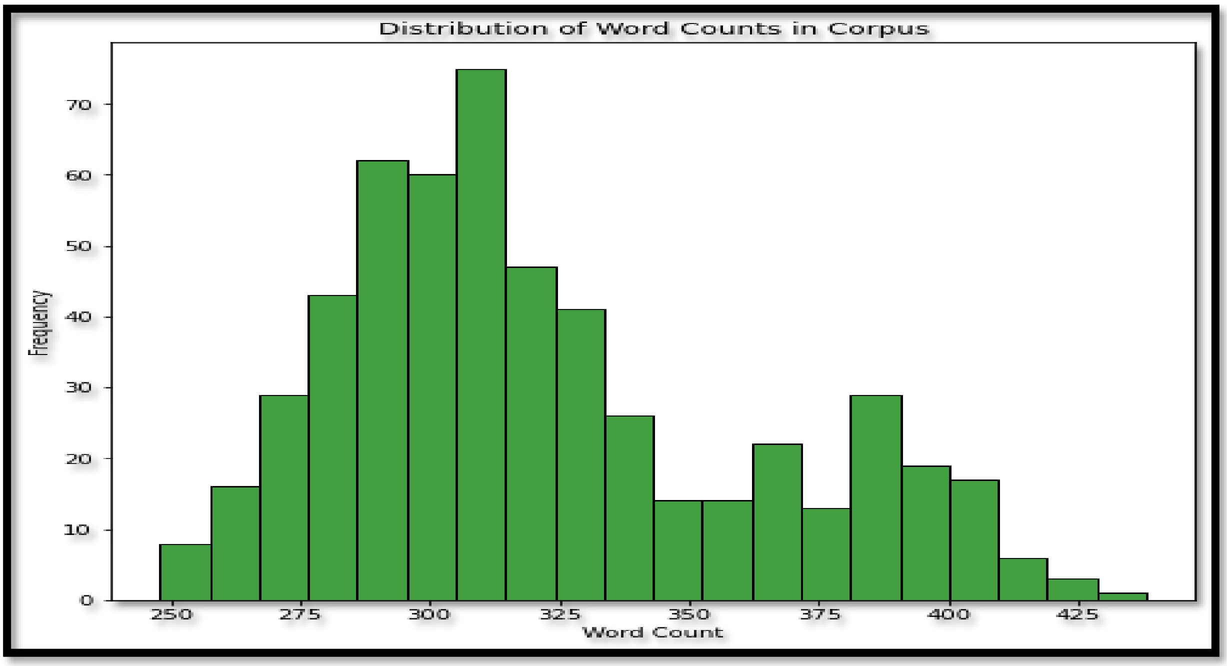


Figure 6

Top 30 words in the Corpus

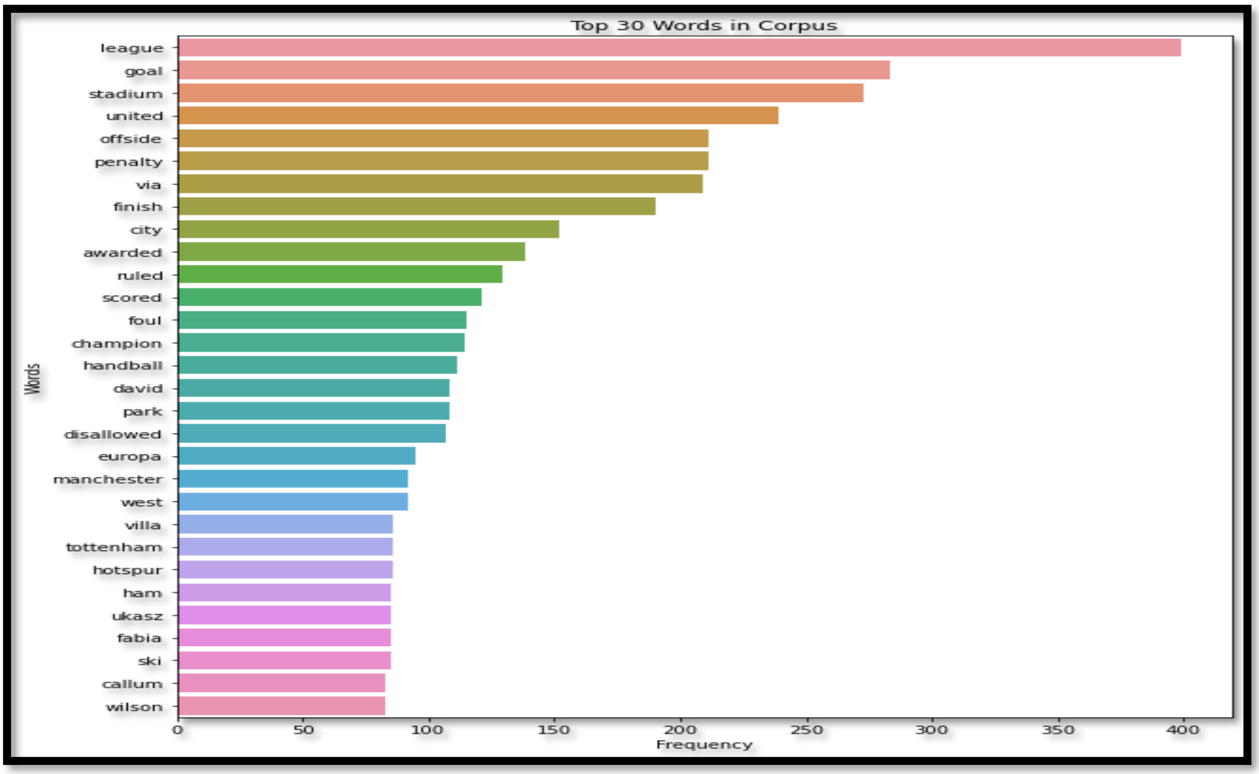
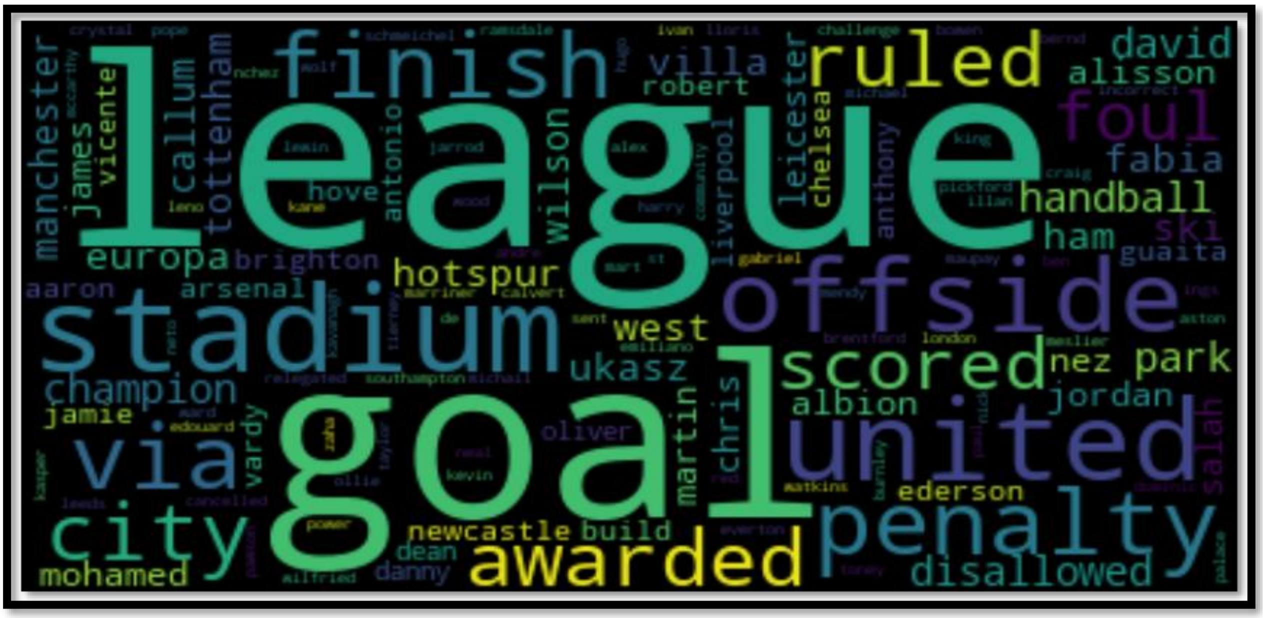


Figure 7

Word Cloud



- ***Corpus Limitations and Sampling Biases***

    - Corpus Limitations: The VAR-related corpus faces challenges due to variations in team abbreviations across websites and a relatively short timeframe since its introduction in 2019, limiting the depth and diversity of textual data. The presence of a limited number of instances for the 'NEUTRAL' class further restricts the corpus's representativeness, impacting the comprehensiveness of sentiment analysis. In addition, incomplete or biased datasets, unreviewed incidents, and the selective impact of VAR on certain teams introduce limitations to the VAR analysis.
    - Sampling Biases: The stratified sampling approach aims to maintain category proportions, yet it may not fully account for nuanced data distributions, potentially introducing biases. Reclassifying 'NEUTRAL' incidents into 'FOR' may impact the balance and accuracy of sentiment classification, influencing the overall performance of the machine learning model in the English Premier League's VAR context. Further, biases can emerge from incomplete data, uneven representation of teams, and variations in VAR application and referee interpretations across matches, complicating the generalizability of VAR assessments.

V. Machine Learning

- ***Specification and justification of the implementation of the ML model***.

Sentiment analysis in VAR incidents can be effectively addressed by machine learning, which uses algorithms to extract subtle sentiments from textual descriptions. ML improves the precision of sentiment classification by training models on pre-processed data, allowing for a more insightful comprehension of the emotional tone in football.-relevant material. Table 3 outlines key hyperparameters crucial for optimizing model performance, enabling systematic exploration of various configurations. As discussed in section IV, hyperparameters such as '*max_features*,' '*min_df*,' and '*max_df*' were tuned for TF-IDF, impacting model complexity and term filtering based on document frequency. The Random Forest Classifier, Gradient Boost Classifier, Logistic Regression, and Support Vector Classifier were employed for sentiment classification.

An Ensemble model using a Voting Classifier was implemented with the aforementioned models using a pipeline. The pipeline integrates these classifiers and hyperparameter tuning is performed for each model which includes '*n_estimators*, '*max_depth'* for Random Forest Classifier and Gradient Boost Classifier, and regularization *parameter (C)* for Logistic Regression and SVM, respectively. The ensemble approach combines the strengths of different classifiers, enhancing the robustness and generalization of the model. Moreover, stratified K-fold cross-validation was used to ensure that the model is validated on diverse subsets of the training data, preventing overfitting. The hyperparameters tuned are used to maximize accuracy, and are also assessed for precision, recall and F1 score. The best hyperparameters are then extracted and using these best hyperparameters the final model is built to classify the sentiments.

Table 3

Hyperparameters for the ML models

Sr. No	Hyper-parameters	Range	Mean Accuracy CV Score
1.	TF-IDF Vectorizer max_featurers	[50,100,150]	-
2.	TF-IDF Vectorizer min_df	[0.2, 0.4, 0.6]	-
3.	TF-IDF Vectorizer max_df	[0.5, 0.7, 0.9]	-
4.	Random Forest Classifier (max_depth)	[100,150,200]	54.43
5.	Random Forest Classifier (n_estimators)	[10,20,30]	
6.	Gradient Boost Classifier (max_depth)	[100,150,200]	56
7.	Gradient Boost Classifier (n_estimators)	[10,20,30]	
8.	Logistic Regression Classifier	[0.1, 1, 5]	46.323
9.	Support Vector Classifier	[0.1, 1, 2]	51.72

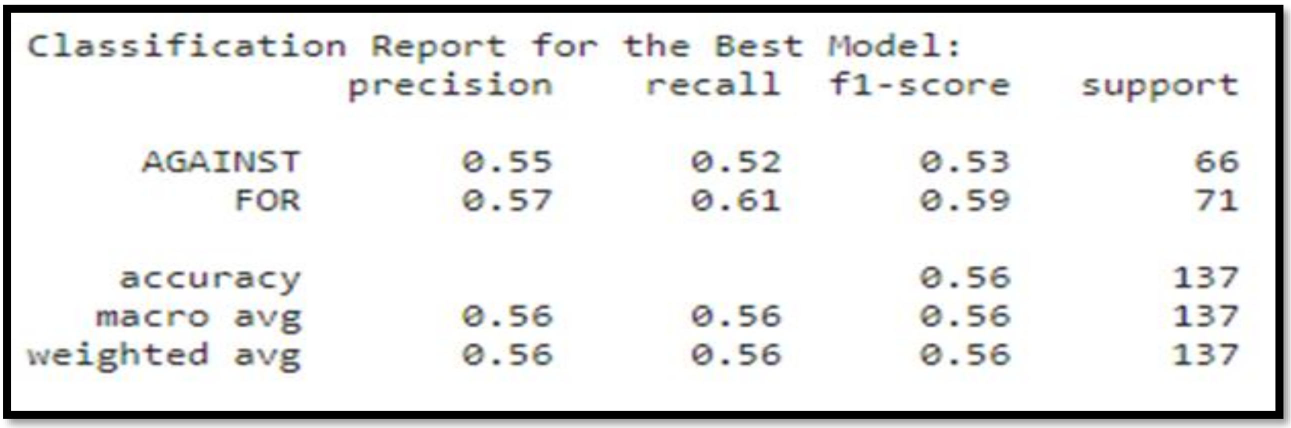
To conduct sentiment analysis, the computation environment comprises Python 3.8.5 as the software platform, utilizing essential libraries such as pandas, numpy, and scikit-learn. The operating system is Windows 10, and the development environment is Jupyter Notebook accessed via Anaconda Navigator. Despite producing commendable results, the analysis proved computationally expensive, requiring approximately 114.7 minutes to obtain optimal hyperparameters. This underscores the significance of computational efficiency in real-world applications. This machine learning approach draws upon established theories in sentiment analysis and classification algorithms. For further depth, references to previous works such as (Sukumarana et al., 2023 and Keshav et al., 2020) could be explored. The ensemble method and hyperparameter tuning strategies align with best practices, contributing to the model's robustness and generalization.

- ***Evaluation and visualization of the machine learning model performance*.**

  In the context of sentiment analysis for VAR incidents, hyperparameter tuning identified optimal settings, with the Gradient Boost Classifier emerging as the top-performing model, achieving a mean CV accuracy of 56%. Figure 8 presents a comprehensive classification report, detailing key scores from the model's predictions on the test data. Recall assesses the classifier's capacity to predict positive cases, precision measures the accuracy of positive predictions (true positives), and the F1-score integrates the two to provide a thorough evaluation of performance (Kanstrén, 2023).

Figure 8

Classification report for the Gradient Boost Classifier

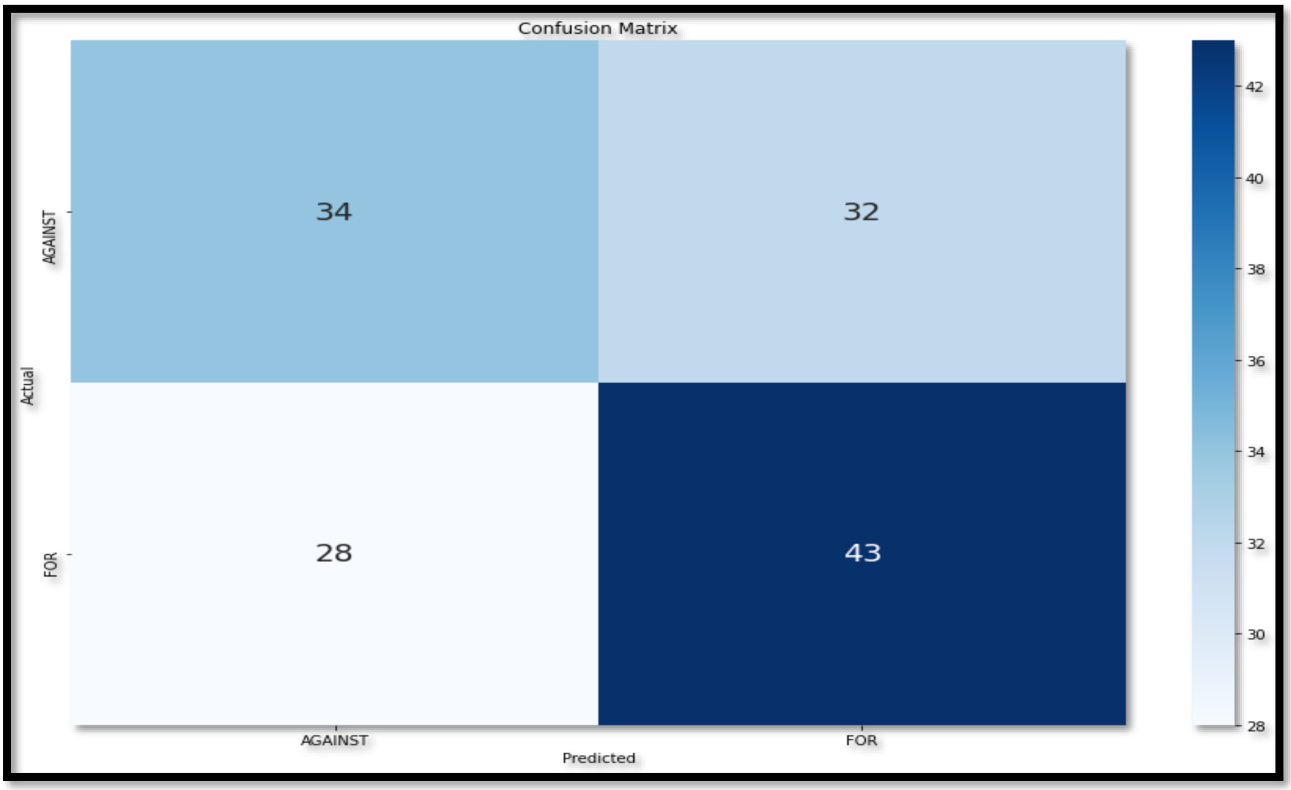


```
Classification Report for the Best Model:
              precision    recall  f1-score   support

     AGAINST       0.55      0.52      0.53        66
         FOR       0.57      0.61      0.59        71

    accuracy                           0.56       137
   macro avg       0.56      0.56      0.56       137
weighted avg       0.56      0.56      0.56       137
```

For the "AGAINST" sentiment class, the model exhibited a balanced precision of 55% and recall of 52%, indicating correct predictions 55% of the time while capturing 52% of actual "AGAINST" instances out of a total of 66. In the case of the "FOR" sentiment class, the precision and recall were 57% and 61%, respectively, out of a total of 71. The F1 score for 'FOR' is 53% whereas for 'AGAINST' is 59%.

Figure 9

Confusion matrix for the Gradient Boost Classifier



The confusion matrix, visualized in Figure 9 via a heatmap, offers insights into the machine learning model's overall performance. The model correctly classified 34 out of 62 instances of "decisions that went against the Home Team" sentiments, while misclassifying

32 out of 75 instances of "decisions that went in favor of the Home Team." While the model exhibits a reasonable ability to discern sentiments in VAR incidents, there exists room for improvement, particularly in minimizing misclassifications.

This extensive evaluation underscores the model's capacity to distinguish between sentiment classes, contributing valuable insights for refining sentiment analysis in the context of VAR-related textual data.

- ***Effect of the data limitations and sampling biases on the machine learning performance***.

The machine learning performance in sentiment analysis of VAR incident descriptions is influenced by notable data limitations and sampling biases. Challenges arise from variations in team abbreviations across websites and the limited instances available for the 'NEUTRAL' class. The decision to reclassify 'NEUTRAL' incidents as 'FOR' introduces potential bias. Moreover, the stratified sampling approach, while aiming to maintain category proportions, may not fully address nuanced data distributions. Moreover, VAR was only introduced in the Premier League in 2019, making it relatively new, and the limited available data or information poses challenges in extending the analysis, introducing data limitations. Instances of incorrect VAR decisions underscore the evolving nature of this technology, adding complexity to the analysis. These complexities emphasize the imperative for cautious consideration and the exploration of expanded datasets to fortify the model's robustness and generalizability.

References

[1] AWS. (n.d.). *What is sentiment analysis?* Amazon Web Services, Inc. https://aws.amazon.com/what-is/sentiment-analysis/#:~:text=Sentiment%20analysis%20is%20the%20process,social%20media%20comments%2C%20and%20reviews

[2] DataQuest. (2022, May 2). *Web Scraping Football Matches From The EPL With Python*. YouTube. https://www.youtube.com/watch?v=Nt7WJa2iu0s

[3] FBREF. (n.d.). *Bot/Scraping/Crawler traffic on sports-reference.com sites*. Sports-Reference.com. https://www.sports-reference.com/bot-traffic.html

[4] Gepard. (2023, February 9). *Enhance your data exchange with the help of CSV files*. Gepard PIM. https://gepard.io/connections/csv

[5] Gillis, A. S. (2023, March). *What is Lemmatization? | Definition from TechTarget*. Enterprise AI. https://www.techtarget.com/searchenterpriseai/definition/lemmatization#:~:text=Lemmatization%20is%20the%20process%20of,processing%20(NLP)%20and%20chatbots

[6] Johnson, D. (2019). *How VAR decisions affected every Premier League club*. ESPN. Retrieved December 2023, from https://www.espn.in/football/story/_/id/37575919/how-var-decisions-affected-every-premier-league-club

[7] Kanstrén, T. (2023, August 4). A look at precision, recall, and F1-score. Medium. https://towardsdatascience.com/a-look-at-precision-recall-and-f1-score-36b5fd0dd3ec

[8] Keshav, R., V, K. A., Priyam, S., & S, M. K. (2020). Applications of Artificial Intelligence in the Game of Football: The Global Perspective. *International Refereed Social Sciences Journal*, *11*(2). http://dx.doi.org/10.18843/rwjasc/v11i2/03

[9] Kim, S., & Gil, J. (2019, August 26). *Research paper classification systems based on TF-IDF and LDA schemes*. SpringerLink. https://doi.org/10.1186/s13673-019-0192-7

[10] Leung, K. (2021, September 23). *Analyzing English Premier League VAR football decisions*. Medium. https://towardsdatascience.com/analyzing-english-premier-league-var-football-decisions-c6d280061ebf

[11] Liaqat,, M. I., Hassan,, M. A., Shoaib, M., Khurshid, S. K., & Shamseldin, M. A. (2022, August 31). *Sentiment analysis techniques, challenges, and opportunities: Urdu language-based analytical study*. PeerJ Computer Science. https://doi.org/10.7717/peerj-cs.1032

[12] Rowe-Willcocks, H., Micallef, C., & Stratton, C. (2023, October 17). *All about the virtual assistant referee - VAR*. The Sun. Retrieved December 14, 2023, from https://www.thesun.co.uk/sport/5278073/var-world-cup-video-assistant-referee-decisions-nations-league/

[13] SoccerBase. (n.d.). *Latest football betting odds | Ref stats | Place your bet today*. Football Betting | Place Your Football Bet Today | Soccer Base. Retrieved December 2023, from https://www.soccerbase.com/referees/home.sd

[14] Sukumarana, C., Bento Devaraj, A. B., Chandrasekar, V., & Marimuthu, R. K. (2023). EXPLORING FOOTBALL FANS' OPINIONS AND EMOTIONS ON VIDEO ASSISTANT REFEREE (VAR) TECHNOLOGY: A STUDY BASED ON NATURAL LANGUAGE PROCESSING. *JOURNAL OF SOUTHWEST JIAOTONG UNIVERSITY*, *58*(1). https://jsju.org.cn/pdf/1009.pdf

[15] ZenRows. (2023, October 14). *Selenium vs. BeautifulSoup in 2023: Which Is Better?* https://www.zenrows.com/blog/selenium-vs-beautifulsoup#selenium