VAR in Football: Game-Changer or Controversy Magnet?

Using sentiment analysis to study if football fans find VAR to be a positive element or more of a detriment to the game. (1336 Words)

Ayush Khaneja

ABSTRACT

The introduction of the Video Assistant Referee (VAR) system in football has ignited intense debates, with proponents highlighting its potential to enhance the accuracy of officiating and critics decrying its potential to disrupt the flow of the game. This study uses sentiment analysis on data collected from Reddit to determine if VAR is perceived as an enhancement or a disruption to the sport. The research contributes to the ongoing VAR debate, offering insights into technology's impact on the football tradition.

1. Introduction

The idea behind introducing VAR to football was to improve the accuracy and fairness of officiating. Using various technologies and video footage review, referees are given the power of re-evaluating and potentially overturning their on-field decisions.

While VAR can lead to more accurate decision-making, the delays, the highly mechanical application of the rules, and the subjectivity[2] of each individual referee combine to make an experience that is not always enjoyed by the fans and viewers. The decisions are sometimes still inaccurate and the delays kill the pace of the game, not to mention the loss in the thrill and excitement that comes from spontaneity.

In recent times there has been a lot more talk about VAR, especially since there was a blatant error in a recent game [1]. This added fuel to the already heated debate on VAR. Seeing fan and viewer sentiments on VAR can help decide if VAR is actually even required in the game or football, or are we better off without it.
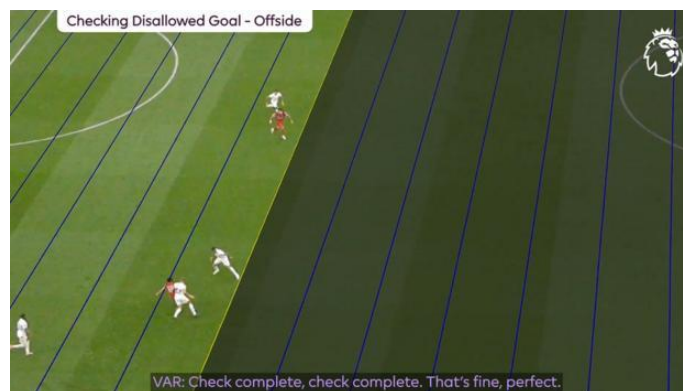

Figure 1: The controversial VAR call [5]

2. Research Questions

The paper aims to answer the following questions:
1. How do the people feel about VAR in general? What is the overall sentiment about the use of VAR?
2. Is the sentiment always consistent throughout over time? Or do bad VAR calls alter the sentiment temporarily?

3. Method

3.1 Data

Data for the research was collected from Reddit through PRAW, the Reddit API Wrapper for Python. Football-related data is a popular choice for sentiment analysis.[7] Using Reddit to gather this data is has also been preferred in past researches.[3], [8] The search queries "VAR" and "VAR Error" were searched for in the following subreddits:
- r/football
- r/soccer
- r/worldcup
- r/PremierLeague
- r/MLS
- r/LaLiga

After going through many subreddits, these were selected because of their popularity and activity.

To get data around specific VAR error incidents, the top 5 were chosen from the internet [6]. The website chosen was selected because it had VAR errors from all across football and not just a particular league. In addition to these 5, the most recent VAR error was also chosen. A search query was written to search for posts related to these specific incidents. This search query was searched for across Reddit.

Match	Competition And Season	Search Query
Arsenal vs Brentford	Premier League 2022-23	"Arsenal vs Brentford VAR error"
West Ham United vs Chelsea	Premier League 2022-23	"West Ham United vs Chelsea VAR error"
Manchester City vs Southampton	Premier League 2020-21	"Manchester City vs Southampton VAR error"
Tottenham Hotspur vs Newcastle United	Premier League 2020-21	"Tottenham Hotspur vs Newcastle United VAR error"
Iran vs Portugal	World Cup 2018	"Iran vs Portugal VAR error"
Tottenham Hotspur vs Liverpool	Premier League 2023-24	"Tottenham Hotspur vs Liverpool VAR error"

Table 1: Specific Search queries to get Reddit data

For all searches, 20 posts and 10 comments were fetched.

Subreddit	Post Count
football	9
LaLiga	1
MLS	9
PremierLeague	19
soccer	20
worldcup	5

Table 2: Subreddit-wise count of posts in the data

3.2 Analysis

Once the post data was obtained, it was manually filtered to remove irrelevant posts and comments, deleted or removed comments, and bot comments. This was done so that only the relevant pieces of data are analyzed. All in all, after the data was filtered, there were over 650 data points (after accounting for missing data).

Post Segment	Count
Title	63
Body	25
Comments	567

Table 3: Distribution of data points

VADER (Valence Aware Dictionary and sEntiment Reasoner)[4] was the approach chosen for sentiment analysis. It is a lexicon and rule-based sentiment analysis tool, and uses a predefined lexicon of words and phrases with associated sentiment scores, which are typically pre-trained. The research in [9] uses VADER to perform sentiment analysis on football data.

The texts in the titles, bodies and comments of the posts were analyzed through VADER. Since the texts were not too long, breaking them into smaller chucks was not required.

An individual dataframe was created to store each type of polarity score. Another dataframe was created to store the aggregate polarity scores for all the comments separately as well.

The VADER compound polarity scores were used to categorize the texts into 5 different categories – Highly Negative, Slightly Negative, Neutral, Slightly Positive, and Highly Positive.

Sentiment Category	Polarity Score Range
Highly Negative	< -0.4
Slightly Negative	>= -0.4 and < -0.1
Neutral	>= -0.1 and >= 0.1
Slightly Positive	> 0.1 and <= 0.4
Highly Positive	> 0.4

Table 4: Breakdown of Sentiment Categories

The next step was to do more analysis through other plots and visualizations. Different density plots and bar plots were made to see trends. Word clouds were also plotted to visualize the most positive and most negative words (most polarizing) in the data
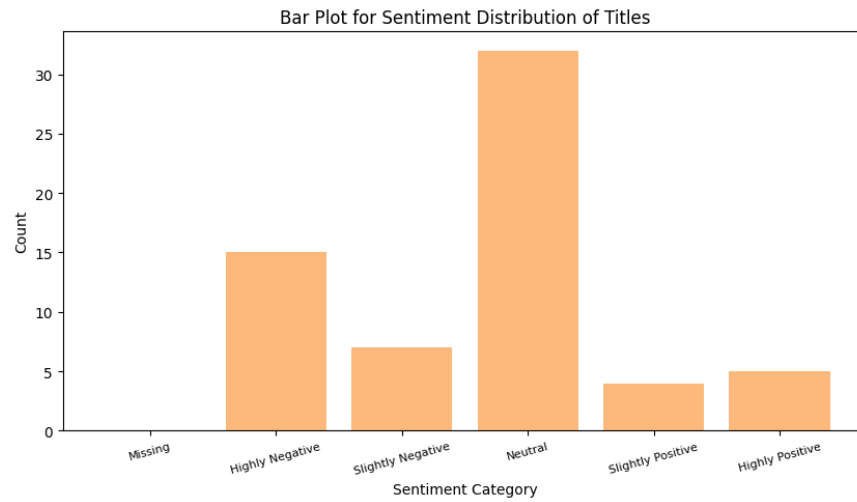
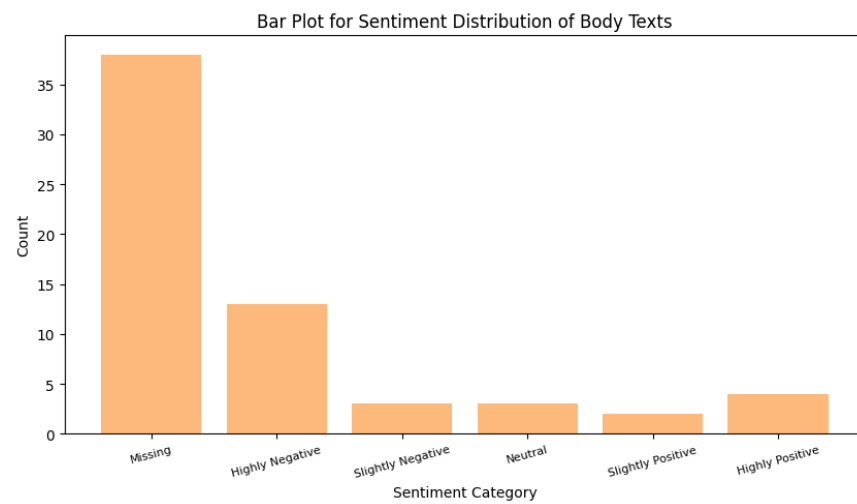Figure 2: Sentiments expressed in post titles

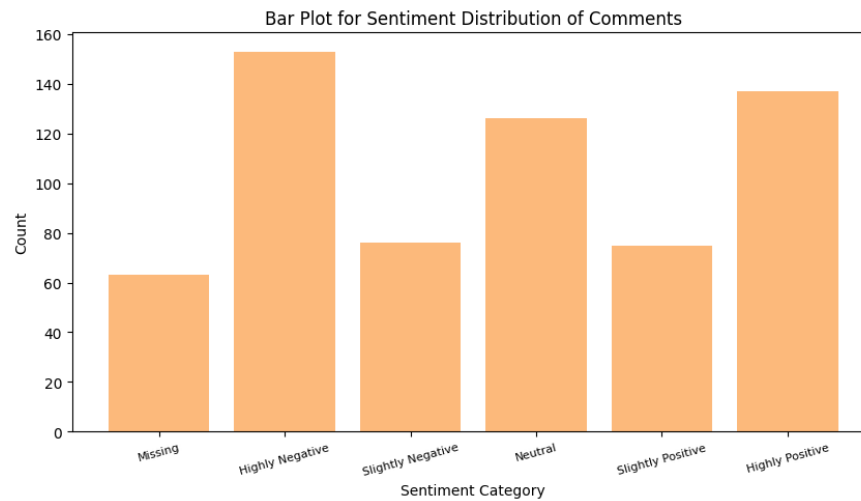Figure 3: Sentiments expressed in post bodies

Figure 4: Sentiments expressed in post comments

To see if the sentiments change around the time a poor VAR decision happens in a game, a timeline chart was also plotted to see the sentiments around VAR over time. The dates of major VAR incidents were also marked on the plot.
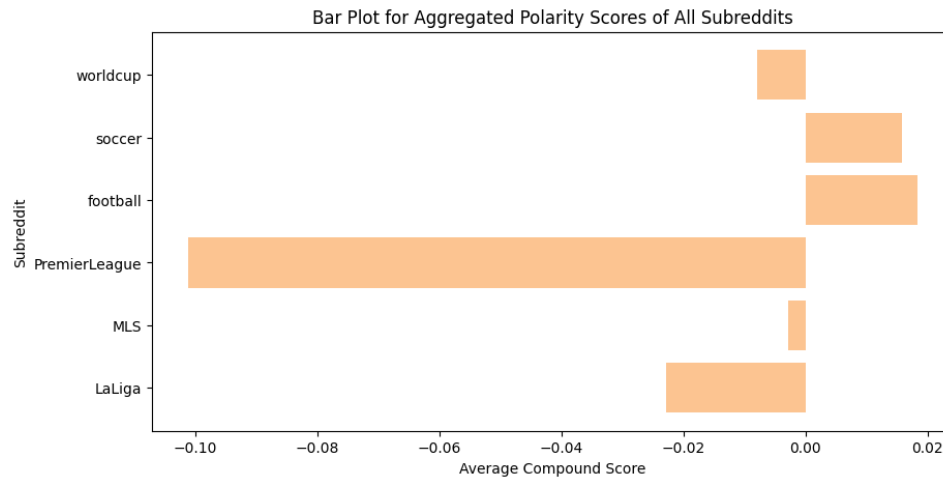
4. Results

Figure 5: Polarity scores across different Subreddits

The r/PremierLeague subreddit seems to have the most negative sentiment towards VAR. It is understandable too, since a lot of incorrect VAR decisions get reported from the Premier League only.
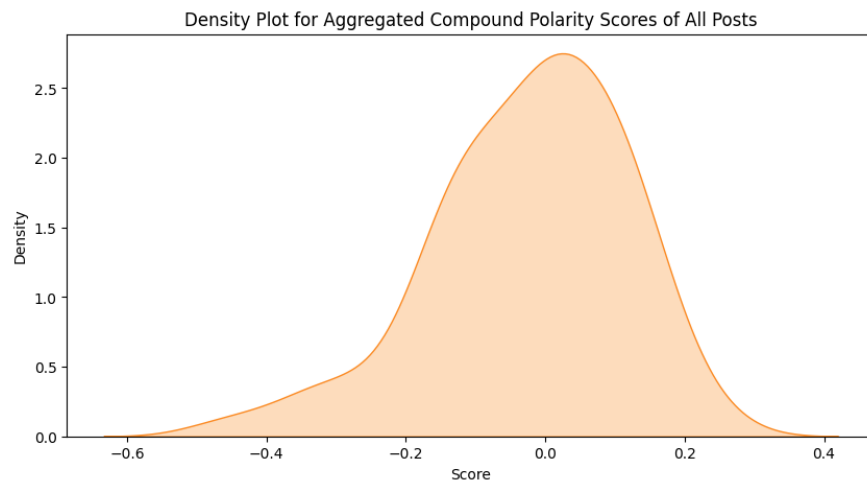

Figure 6: Distribution of aggregated compound polarity scores across posts

Most posts were majorly neutral in their sentiment, but there were more negative posts than positive ones.
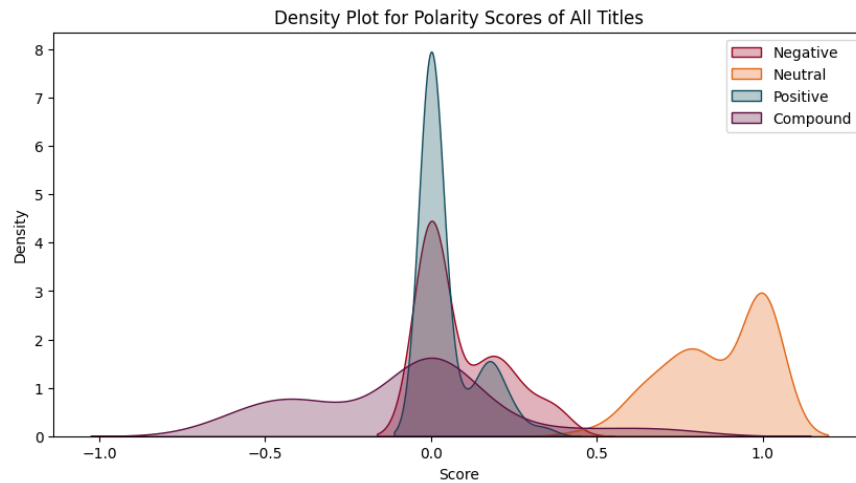
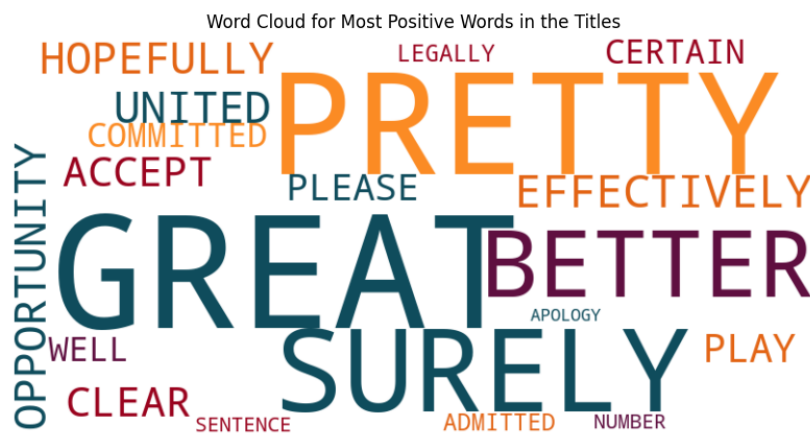Figure 7: Distribution of polarity scores in post titles

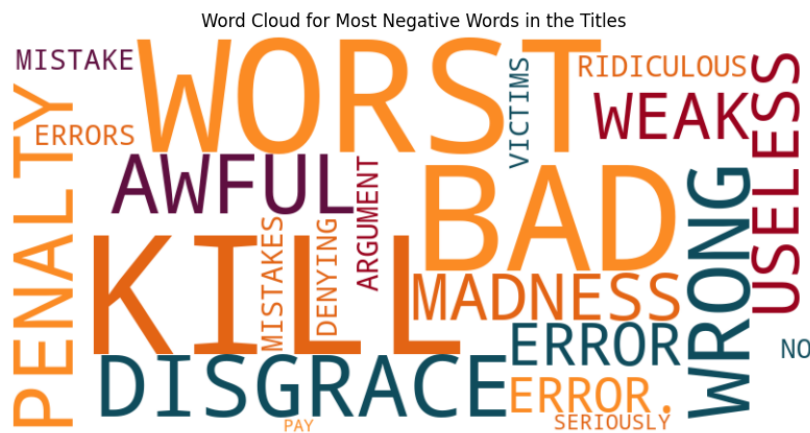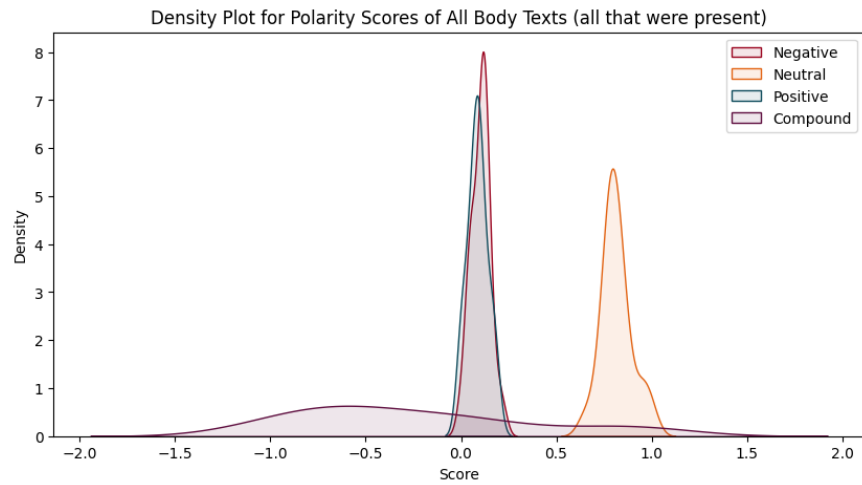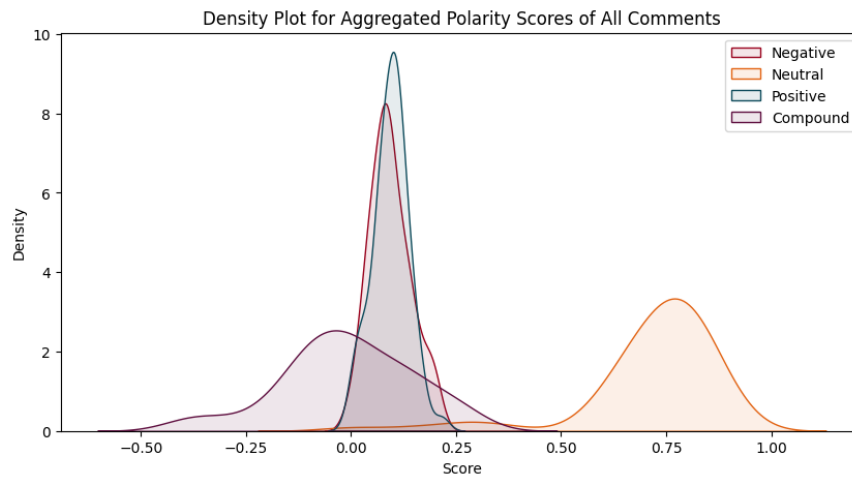Figure 8: Most positively polarized words in the post titles

Figure 9: Most negatively polarized words in post titles

The post titles are mostly of a neutral sentiment as seen in figures 2 and 8. The titles of posts don't seem to be very strongly-worded and probably look to engage people into a more detailed, candid discussion in the comments, but still lean towards a more negative sentiment. The word clouds show the positively polarized and negatively polarized words used in the post titles. Looking at these words it appears that there are more negative words than positive words used in the titles. There is a lot of mention of errors and mistakes in the titles, suggesting that a lot of the post titles are talking about VAR-related oversights.

Figure 10: Distribution of polarity scores in post bodies

Figure 11: Most positively polarized words in the post bodies

Figure 12: Most negatively polarized words in the post bodies

Looking at the post bodies, we see a similar trend on polarity scores as post titles. Mostly neutral sentiments, but more negative opinions than positive ones. A lot of posts did not have bodies, but the ones that did expressed mostly negative sentiments in their content. Looking at the word clouds, the positive words seem to be taken slightly out of context, some of these words might have been used either sarcastically, or were meant to convey a different sentiment. The negative words too, have some words that were meant to convey a different sentiment (eg: attackers might have been taken out of context to mean assailants and not the football player position). The bodies seem to have some expletives as well, showing that stronger sentiments are expressed in the post bodies.

Figure 13: Distribution of polarity scores in post comments

Figure 14: Most positively polarized words in the post comments

Figure 15: Most negatively polarized words in the post comments

The comments show a similar overall trend as the title and body, more negative than positive sentiment. Again, as seen in the post bodies, some words for both positive and negative sentiments might be taken out of context by VADER. The negative sentiments are presented with much more emphasis than the negative words. The comments are the most candid means of expressing views, with use of very strong language and expletives. Despite VADER taking some words or phrases being taken out of context, the sentiment is leaning towards being negatory.
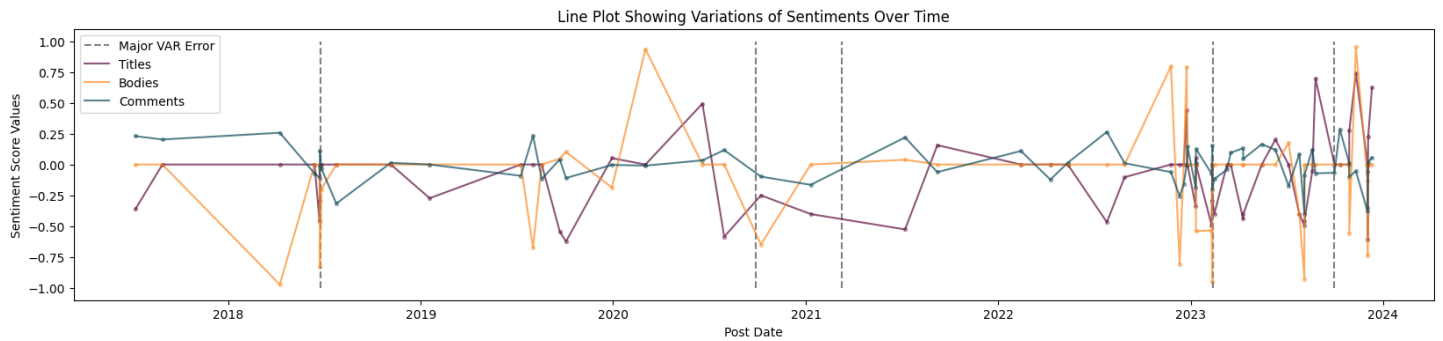
Figure 16: Sentiment changes over time

In the above plot we see that sentiments around VAR are highly fluctuating, but don't really seem to turn more negative after a major VAR blunder takes place. Towards the end we see many more VAR blunders and more posts about VAR as well.

5. Conclusion and Limitations

In conclusion, the sentiment for VAR was more negative than positive. The negative opinions were expressed more strongly through the use of very harsh words, and even expletives. There were posts and comments in support of VAR but they were fewer than the posts and comments expressing dissent for the technology.

Sentiments don't seem to get much more negative after a VAR error, and keep fluctuating over time.

The limitations with the above research are related to the mining of the data and the use of VADER.

Reddit, being a discussion-based forum, has a lot of data that is not all relevant. Even when using very specific search queries, it is a possibility that the data collected is not exactly relevant to the analysis. Another problem is with the inability to search for posts between certain time periods.

VADER has the problem of not understanding specific context for some words, as it just relies on pre-trained information and already known polarities for words. But a lot of times words are used in a different context to express a different sentiment, and VADER does not handle that well.

REFERENCES

[1] Johnson, D. (2023, October 2). The VAR Review: What went wrong for Luis Diaz's offside goal - ESPN. ESPN.com. https://www.espn.com/soccer/story/_/id/38512240/the-var-review-went-wrong-luis-diaz-goal

[2] Smith, R. (2023, October 6). Liverpool, V.A.R. And the problem with process. The New York Times. https://www.nytimes.com/2023/10/06/world/europe/liverpool-var.html

[3] S. Aswath, D. Godavarthi and B. Das, "Analysing Conflicts in Online Football Communities of Reddit," 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), Vellore, India, 2020, pp. 1-6, doi: 10.1109/ic-ETITE47903.2020.386.

[4] Hutto, C., & Gilbert, E. (2014, May). Vader: A parsimonious rule-based model for sentiment analysis of social media text. In Proceedings of the international AAAI conference on web and social media (Vol. 8, No. 1, pp. 216-225).

[5] PGMOL makes audio public of controversial Diaz VAR error. (20233, October 3). BBC Sport. https://www.bbc.com/sport/football/66998252

[6] Tripathi, S. D. (2023, February 14). VAR causes WAR. FootTheBall. https://www.foottheball.com/football-top-10/top-ten-worst-var-decisions-ever-in-football/

[7] Patel, R., & Passi, K. (2020). Sentiment Analysis on Twitter Data of World Cup Soccer Tournament Using Machine Learning. IoT, 1(2), 218–239. MDPI AG. Retrieved from http://dx.doi.org/10.3390/iot1020014

[8] de Sousa Silva, G. H. (2021). Football Players Performance Analysis and Formal/Informal Media: Sentiment Analysis and Semantic Similarity (Doctoral dissertation, ISCTE-Instituto Universitario de Lisboa (Portugal)).

[9] She, J., Swart-Arries, K., Belal, M., & Wong, S. (2023). What Sentiment and Fun Facts We Learnt Before FIFA World Cup Qatar 2022 Using Twitter and AI. arXiv preprint arXiv:2306.16049.