

Paper Proposal 2

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.

(636 Words)

Research Questions:

Video Assistant Referee (VAR) is under scrutiny for a big blunder in decision-making in a recent game played in the English Premier League.^[1] Has the sentiment around VAR always been negative? Or is there support for VAR in general and the negativity is a temporary reaction to a bad decision made by VAR?

Background:

Video Assistant Referee was introduced to improve accuracy and fairness of officiating in the game of football. It is a technology-driven system that enables match officials to review and potentially overturn on-field decisions using video footage.

Proponents argue that VAR has led to more accurate decisions, particularly in cases of offside and determining the ball's location during goal-scoring incidents. It promotes greater accountability among match officials, as they are aided by technology in making decisions. VAR is also seen as a tool to uphold fair play and reduce the chances of wrongful decisions.

There is an opposing school of thought here as well. One of the most significant controversies is the time taken to review decisions. Lengthy delays in matches can disrupt the flow and excitement, leading to frustration among players, fans, and even coaches. VAR decisions can sometimes be subjective^[2], with different referees interpreting the same incident in different ways. This has led to inconsistency in applying the technology. VAR's use in offside decisions, particularly the measurement of body parts and the position of players, has been a source of debate. Some argue that it can be overly precise, leading to goals being disallowed for marginal offside positions. VAR has sparked debates about how it affects the overall fan experience. While some appreciate the pursuit of accuracy, others miss the spontaneous emotions and celebrations associated with goals.

As mentioned in the research question, the use of VAR led to a wrong decision that changed the course of the game, and potentially the entire season for the team having to face the consequences. Studying the general sentiment around VAR can help footballing authorities make necessary adjustments to the system wherever necessary so that both the fair play aspect and the thrill of spontaneity can be balanced.

Data:

Data will be collected from Reddit via scraping through the Reddit API for Python. Since the idea is to analyze general sentiment, and not sentiments of just players or just fans, Reddit is a viable source to gather data. Reddit has been a source of data for football related sentiment analysis in other works as well.^[3] Both a reddit-wide search and search within specific subreddits like r/soccer, r/football etc. will be performed to gather over 500 viable data points (Post titles, post bodies, post comments). VAR was widely accepted by 2018-2019 after it was used in the world cup in 2018. So, the range of gathering data will be from 2018 onwards. Before Reddit is scraped for relevant data, a list of the top 10 most controversial VAR decisions will be made from the web. Special focus will be given to posts around the time of these controversial decisions to see if sentiments become overly polarized around these times.

Method:

VADER (Valence Aware Dictionary and sEntiment Reasoner)^[4] will be the approach taken 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 lexicon includes both polarity scores (positive or negative) and intensity scores. VADER takes into account negations, booster words, and punctuation to make sentiment analysis more context-aware. VADER has been selected as the tool of choice because is easy to use and can provide quick sentiment analysis for short, informal text and it handles nuances like negations and booster words well.

Since VADER works better on shorter texts, longer texts will be split into chunks and their individual compound scores will be combined as a weighted average based on the number of polarizing words found in each chunk.

The results will be explained both statistically (tables, summaries etc.) and visually (charts, plots etc.).

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).