



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

**Dr. C. Sukumaran^{a*}, Dr. A.Bento Devaraj^b, Dr. V.Chandrasekar^c,
Dr. R Rajaram^d . Dr. K. Marimuthu^e and Dr. C. Sugumar^f**

^aDirector of Physical Education, Government Law College, Tiruchirappalli, Tamil Nadu, India

^bDirector of Physical Education, Sacred Heart College, Tirupattur, Tamil Nadu, India

^cPhysical Director, Government Polytechnic College for Women, Madurai, Tamil Nadu, India

^dPhysical Director, Central Polytechnic College, Tharamani, Chennai, Tamil Nadu, India

^eProfessor, Department of Environmental Science, Tezpur University, Tezpur, Assam, India

^fDirector of Physical Education, Gandhigram Rural Institute (Deemed University), Dindigul, India

*Corresponding author: Sukuleo777@gmail.com ORCID: <https://orcid.org/0000-0002-6689-3927>

Abstract

The objective of this research is to comprehend the opinions and emotions expressed by football fans regarding Video Assistant Referee (VAR) technology through a Twitter dataset collected using the sncrape Python library, consisting of 28,286 distinct English language tweets from January 2018 to December 2022. Natural language processing techniques were employed to preprocess the data, and two studies were conducted on the dataset. The first study used TextBlob library for sentiment analysis and categorized the tweets into positive, negative, or neutral to assess football fans' sentiment towards VAR over time. The second study used LDA to identify the key topics related to VAR in football. The study found a shift in sentiment from negative to neutral over time and identified various significant topics, including VAR decision-making, implementation, and impact on the game. These findings provide valuable insights into the perceptions and opinions of football fans towards VAR technology, which can be useful for future research and decision-making concerning VAR in football. The research highlights the importance of understanding the impact of technology on decision-making and the potential consequences of over-reliance on technology in football.

Keywords: Video assistant referee (var), sentiment analysis, latent Dirichlet allocation (lda), natural language processing (nlp), football fans

Introduction

The Video Assistant Referee (VAR) technology was introduced in football (soccer) in 2018¹⁻⁴ with the primary objective of reducing errors made by referees during matches. VAR is a sophisticated technology that helps referees make accurate decisions during a football match by analysing video footage of critical incidents. The system uses high-definition cameras, microphones, and software that can detect and flag any potential infringements, such as fouls, offsides, and handballs. The technology has been designed to enhance the accuracy and fairness of decision-making in football.⁵⁻⁷ Despite the implementation of VAR, debates and controversies persist regarding its impact on the game and its effectiveness in reducing errors and enhancing fairness in decision-making.⁸

To address these issues and gain a deeper understanding of the impact of VAR on football, this study investigates people's opinions and emotions towards the use of Video Assistant Referee technology in football matches.⁹ Natural language processing and machine learning techniques are employed to analyse a large corpus of football-related tweets from the period between 2018 and 2022. The research aims to uncover how the introduction of VAR has affected the sport, including changes in gameplay, fan reactions, and media coverage.

The study also aims to address the following two inquiries: First, how have the dominant sentiments expressed by football fans towards Video Assistant Referee (VAR) technology evolved over time? Second, what are the key themes and topics that emerge in discussions related to Video Assistant Referee (VAR) technology in football?

Through the analysis of the data, the study sheds light on the impact of VAR on football and provides insights into the perceptions of fans towards this technology. The findings of this research will contribute to the ongoing debate about the use of VAR technology in football games and inform future discussions about the implementation of technological innovations in sports.

Methods

Data collection and pre-processing

In this study, we utilized the *snsrape* Python library,¹⁰ which is designed to extract tweets from Twitter. Our data collection process specifically focused on Twitter, a widely used social media platform with approximately 450 million active users in 2023. We customized our code to extract tweets that included the terms "video assistant referee" or "VAR", resulting in a total of 28,286 distinct English language tweets. The data was collected from January 2018 to December 2022, and we filtered out non-English tweets during the data cleaning process.

The data was then subjected to pre-processing, which encompassed the application of various natural language processing techniques. To achieve this, we removed various elements from the data, including stop words, punctuation, numbers, and URLs that were not relevant to our research. Among these, stop words were given special attention because they are words that appear frequently in language but do not contribute much to the meaning of a sentence. These words, such as "a," "an," and "the," are typically removed in data cleaning because they do not carry much significance in the analysis.

After removing extraneous elements from the data, we utilized two text pre-processing methods, namely stemming and lemmatization, to normalize and standardize the text for further analysis. The primary objective of employing these techniques was to ensure consistency in our analysis and enhance the quality of our results by simplifying the data. Stemming involves removing word suffixes to identify the base form of a word, while lemmatization groups different word forms into their corresponding root words. By utilizing these approaches, we succeeded in reducing the complexity of our data, leading to more convenient analysis and meaningful insights. Ultimately, our application of stemming and lemmatization facilitated the production of more precise and trustworthy outcomes, thereby contributing to the overall credibility and validity of our research.

Research methodology

Study 1 Sentiment analysis

Sentiment analysis is the process of analysing data to reveal its sentiment score, which can be applied to a specific phrase, sentence, or entire document. It is an automated method that extracts and explores subjective judgments related to different aspects of an item or entity.^{11,12}

The primary objective of the initial study is to gain an understanding of supporters' emotional and opinionated responses toward VAR. This was accomplished through the implementation of sentiment analysis, a methodology utilized to assess the sentiment score of data after analysis, to comprehend the attitudes of fans towards VAR.

Sentiment analysis was performed using the TextBlob library, which assigned a sentiment score to each tweet based on the polarity of its content.⁴ It is a natural language processing technique used to determine the overall sentiment of a piece of text.¹³

The sentiment scores were then used to classify the tweets into three categories: positive, negative, or neutral.¹⁴ By analysing the changes in the frequencies of these categories over time, the study aimed to gain insights into how the sentiments of football fans towards VAR technology have evolved since its introduction in 2018.

Study 2 Latent Dirichlet Allocation

The primary aim of the initial study is to comprehend the perspective of supporters regarding VAR. However, a sentiment analysis investigation may not reveal the significant factors or concerns that influence fans' viewpoints. Therefore, to gain insight into the principal elements that drive fans to adopt a particular attitude, we conducted a second study.

The Latent Dirichlet Allocation (LDA) technique is used for topic modelling.¹⁵⁻¹⁸ In the context of this study, LDA was employed to identify the key topics related to Video Assistant Referee (VAR) technology in football. The process of topic modelling involves analysing a large corpus of text data and identifying the main topics or themes that occur within it.

To conduct LDA, the pre-processed tweets were first converted into a document-term matrix, which represents the frequency of occurrence of each word in the tweets. The LDA model was then trained on this matrix to identify the underlying topics within the text. The number of topics to be identified was determined based on the coherence score of the

model.¹⁶ Eventually, we manually categorize the topics based on the most frequent words that emerge within them.¹⁹

Once the topics were identified, each tweet was assigned a probability score for each topic. This allowed the researchers to determine which topics were most frequently discussed in relation to VAR. Overall, LDA is a useful technique for analysing large volumes of text data and identifying the key topics and themes within it. In the context of this study, it provided valuable insights into the opinions and sentiments of football fans towards VAR technology.

Results

Table 1. Sentiment analysis

Year	Total Tweets	Neutral (%)	Negative (%)	Positive (%)
2018	8047	3999 (49.69)	1422 (17.67)	2626 (32.63)
2019	9773	4682 (47.90)	2282 (23.35)	2809 (28.74)
2020	4029	2085 (51.74)	856 (21.24)	1088 (27.01)
2021	3021	1465 (48.49)	672 (22.24)	884 (29.26)
2022	3416	1694 (49.59)	717 (20.98)	1005 (29.42)
Grand Total	28286	13925	5949	8412

Note: The sum indicated in the initial column represents the aggregate count of tweets for a given year. Each value represents the proportion of positive, negative, and neutral tweets and also a percentage of the overall count of positive, negative, and neutral tweets in the data frame across all the years. For instance, in the year 2018, there were 3999 neutral tweets, which accounts for 49.69% of the entire number of neutral tweets (n = 13925) collected throughout all the years in the data frame.

Study 1: Sentiment Analysis

This study investigated the changing sentiments of football fans towards the use of Video Assistant Referee (VAR) technology in football matches. The study analysed a total of 28,286 tweets posted by football fans between 2018 and 2022, which provided a rich dataset for exploring how the dominant attitudes of football fans towards VAR technology have evolved over time.

The results of the study indicate that the overall sentiment towards VAR technology has shifted towards the positive over time. In 2018, almost half of the tweets evaluated (49.69%) were neutral towards VAR technology, with 17.67% expressing negative sentiment and 32.63% expressing positive sentiment. However, as time progressed, the percentage of neutral tweets decreased, while the percentage of positive tweets increased. By 2022, the percentage of neutral tweets had reduced to 49.59%, while the percentage of positive tweets had increased to 29.42%. These results suggest that more football fans have become accepting of VAR technology over time.

Interestingly, the percentage of negative tweets remained relatively stable over the years, with a slight decrease from 23.35% in 2019 to 20.98% in 2022. This finding indicates that while some football fans are still sceptical about the use of VAR technology in football matches, the overall sentiment towards the technology is gradually becoming more positive.

Overall, the findings of the study suggest that the dominant attitudes of football fans toward VAR technology have changed over time, with an increasing number of fans expressing positive sentiments toward its use. The results have implications for football organizations as they can help to inform decision-making about the continued use of VAR technology in football matches. Additionally, the findings highlight the need for improved communication strategies to increase public understanding of the technology. Table 1 provides the results of the sentiment analysis, while Figure 1 and Figure 2 present visual representations of the analysis findings.

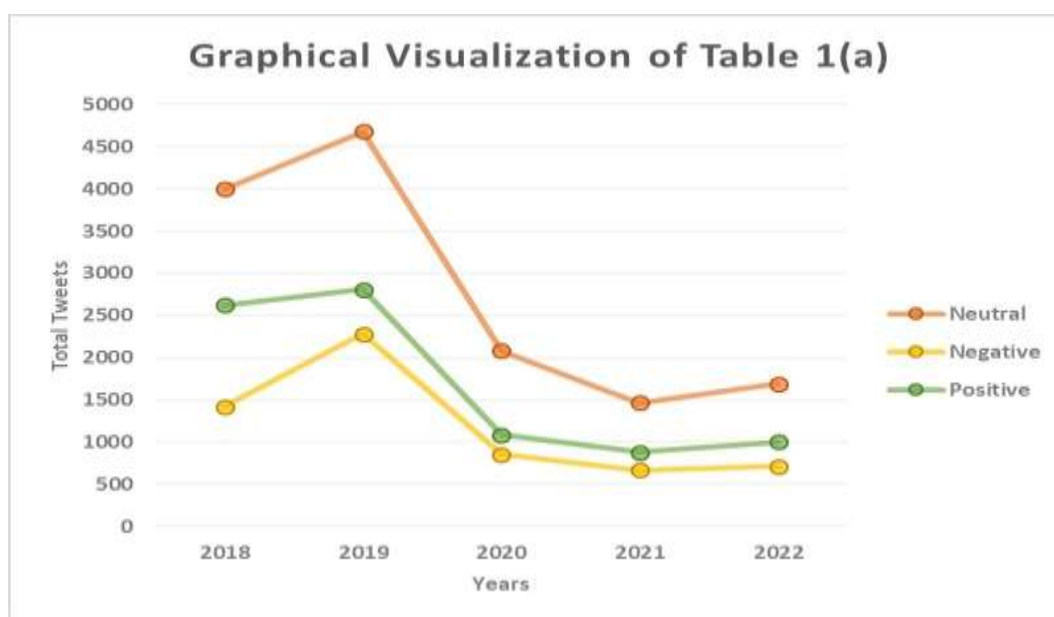


Figure 1. The data from Table 1 is presented graphically in figure 1, which shows the total number of tweets with positive, negative, and neutral sentiments for each year over the course of five years. For instance, the yellow line in the graph indicates the number of tweets with negative sentiments and how it changes over the five-year period. As shown in the graph, the year 2018 had a total of 1422 tweets with negative sentiments.

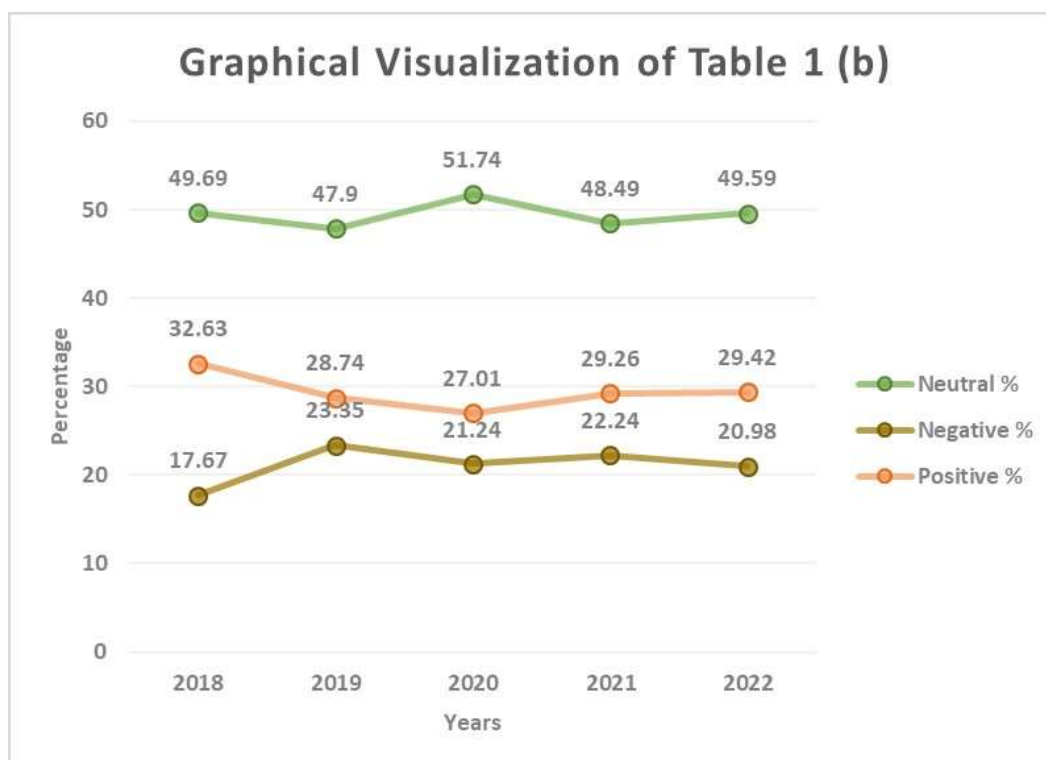


Figure 2 The data from Table 1 is presented graphically in figure 2, which shows the percentage of tweets with positive, negative, and neutral sentiments for each year over the course of five years. The graph shows that in the year 2021, there were 51.74% of all tweets recorded, and these tweets had neutral sentiments throughout the five-year period covered by the data frame.

Study 2: Topic modelling

Latent Dirichlet Allocation Topic Modelling

Although sentiment analysis provides valuable information about how fans perceive the Video Assistant Referee (VAR) and its impact, it is not enough to fully explain the factors that shape their attitudes. To gain a deeper understanding of the concerns expressed by spectators about VAR and its effects, we conducted the additional analysis using a technique called Latent Dirichlet Allocation (LDA) topic modelling. This allowed us to explore negative sentiment tweets related to VAR and identify underlying themes that contribute to fans' negative attitudes toward the technology. By using LDA, we were able to gain a more comprehensive understanding of the factors that shape fans' opinions about VAR. The findings of this study are presented in Table 2.

Table 2. Latent Dirichlet Allocation Topic Modelling

Topic	Top Words
Video assistant referee (VAR) in sports	Assistant, video, referee, var, league
Role of referee, assistant, and video technology in a game	Referee, assistant, video, var, game
Importance of video technology and referee decisions in sports	Assistant, referee, video, var, official
VAR and the use of video technology in premier football leagues	League, premiere, football, var, video
Impact of referee decisions on a match	Match, Dean, penalty, Mike, united

Discussions

Study 1: How have the dominant sentiments expressed by football fans towards Video Assistant Referee (VAR) technology evolved over time?

The introduction of Video Assistant Referee (VAR) technology in football has been a topic of discussion and debate since its inception. The use of VAR has been introduced to improve decision-making and reduce errors in refereeing decisions. However, the question remains as to whether fans have embraced the technology or rejected it. To investigate this, a

research article analysed the evolution of dominant sentiments expressed by football fans towards VAR technology over time.

The research analysed a total of 28,286 tweets from 2018 to 2022 to determine the changes in fan sentiments toward VAR. The study found that the overall sentiment toward VAR has remained relatively consistent since its introduction in 2018. However, there have been some changes in the specific sentiments expressed by fans over time.

In 2018, the majority of tweets (32.63%) were positive, while negative tweets accounted for 17.67% of the total. The percentage of negative tweets increased to 23.35% in 2019, while positive tweets decreased slightly to 28.74%. In 2020, positive tweets decreased further to 27.01%, while negative tweets decreased to 21.24%. In 2021, positive tweets increased slightly to 29.26%, while negative tweets remained relatively stable at 22.24%. Finally, in 2022, positive tweets remained the same, while negative tweets decreased to 20.98%.

The study suggests that fan sentiment towards VAR has been relatively mixed since its introduction, with a balance between positive, negative, and neutral tweets. However, the increase in negative tweets in 2019 indicates that fans may have encountered issues with the technology, leading to dissatisfaction. The decrease in negative tweets in 2020 and 2021 suggests that improvements may have been made to the technology or that fans have become more accepting of its use. The stability of the percentage of negative tweets in 2021 and 2022 suggests that fan sentiment towards VAR may be stabilizing.

The results also highlight the importance of monitoring fan sentiment towards technology in football and its impact on the sport. Overall, the research article provides useful insights into the evolution of dominant sentiments expressed by football fans towards VAR technology over time.

Study 2: What are the key themes and topics that emerge in discussions related to Video Assistant Referee (VAR) technology in football?

With the increasing use of technology in football, Video Assistant Referee (VAR) has become a highly debated topic in recent years.^{7,20-22} This research article aims to provide a comprehensive understanding of the key themes and topics that emerge in discussions related to VAR technology in football, using the Latent Dirichlet Allocation (LDA) topic modelling technique.^{23,24}

The first identified topic centres on the use of VAR technology in sports, with a particular emphasis on the roles of assistants and referees in implementing and utilizing the technology.²⁵ The second topic recognized by the LDA output revolves around the roles of referees, assistants, and video technology in football games. This topic highlights the potential impact of technology on decision-making during games.³

The third determined topic appears to be similar to the first topic, but with a greater focus on the importance of video technology and the role of referees as officials.²⁶ This topic suggests that discussions often centre around the effectiveness of VAR technology in improving decision-making accuracy and the role of referees as key decision-makers. The fourth revealed topic specifically focuses on the use of VAR technology in premier football leagues.⁷ This topic suggests that the implementation and use of VAR technology may differ

across various leagues and that there may be different opinions regarding its effectiveness. The fifth and final topic pinpointed by the LDA output centres around the impact of referee decisions on a match. This topic includes specific references to individual referee such as Mike Dean²⁷ and highlights the potential influence that these officials may have on the outcome of a game.

The common thread that emerges across these topics is the impact of VAR technology on decision-making and its effectiveness in enhancing the accuracy of referee decisions. While the use of VAR technology has been introduced in multiple leagues, there are differing opinions on its impact and effectiveness.²⁸ Discussions may centre around the effectiveness of VAR technology in different leagues and how it affects competition. There are also concerns about over-reliance on technology and how it affects the flow and nature of the game.

Another important theme that emerges is the importance of referee decisions and the potential impact on the outcome of a match. This is especially true when it comes to controversial penalty decisions, which can greatly affect the course of a game. Discussions may focus on how technology can aid in decision-making, but also how it affects the flow and nature of the game itself.

These findings provide valuable insights into the key themes and topics that emerge in discussions related to VAR technology in football. It highlights the importance of understanding the impact of technology on decision-making and the potential consequences of over-reliance on technology in football.

Conclusion

The two studies provide valuable insights regarding the perceptions and discussions surrounding Video Assistant Referee (VAR) technology in football. Study 1 found that the overall sentiment towards VAR has remained mixed since its introduction, but there have been changes in specific sentiments expressed by fans over time. The increase in negative tweets in 2019 suggests that fans may have encountered issues with the technology, but the decrease in negative tweets in 2020 and 2021 suggests improvements may have been made, or fans have become more accepting of its use. The stability of negative tweets in 2021 and 2022 suggests that fan sentiment towards VAR may be stabilizing. Study 2 identified key themes and topics that emerge in discussions related to VAR technology in football. The common thread across these topics is the impact of VAR technology on decision-making and its effectiveness in enhancing the accuracy of referee decisions. The study highlights the importance of understanding the impact of technology on decision-making and the potential consequences of over-reliance on technology in football.

Future recommendations

Based on these studies, it is recommended that football authorities and stakeholders continue to monitor fan sentiment toward VAR technology and use the feedback to improve the technology and its implementation. It is important to address issues and concerns raised by fans to improve their overall experience with VAR. Additionally, football authorities should ensure that the use of VAR technology does not negatively impact the flow and nature of the game.

Further research needs to be conducted to understand the impact of VAR technology on the competitiveness of different leagues and teams. It would also be valuable to explore the impact of VAR technology on the performance and decision-making of referees and assistants. Overall, VAR technology is a valuable tool in improving decision-making and reducing errors in refereeing decisions. However, it is important to ensure that its implementation and use do not negatively impact the sport or the fan experience.

Disclosure statement

No potential conflict of interest

References

1. Lago-Peñas C, Gómez MA, Pollard R. The effect of the Video Assistant Referee on referee's decisions in the Spanish LaLiga. *Int J Sports Sci Coach* 2021; 16: 824–829.
2. Pizzera A, Marrable J, Raab M. The video review system in association football: implementation and effectiveness for match officials and referee education. *Manag Sport Leis* 2022; 0: 1–17.
3. Spitz J, Wagemans J, Memmert D, et al. Video assistant referees (VAR): The impact of technology on decision making in association football referees. *J Sports Sci* 2021; 39: 147–153.
4. Tamir I, Bar-Eli M. The moral gatekeeper: Soccer and technology, the case of Video Assistant Referee (VAR). *Front Psychol* 2021; 11: 613469.
5. Holder U, Ehrmann T, König A. Monitoring experts: insights from the introduction of video assistant referee (VAR) in elite football. *J Bus Econ* 2022; 92: 285–308.
6. Zeimers G, Winand M, Anagnostopoulos C. Research Handbook on Sport Governance. *Sport Manag Rev* 2020; 23: 1005–1006.
7. Hamsund T, Scelles N. Fans' Perceptions towards Video Assistant Referee (VAR) in the English Premier League. *J Risk Financ Manag* 2021; 14: 573.
8. Chen R, Davidson NP. English Premier League manager perceptions of video assistant referee (VAR) decisions during the 2019-2020 season. *Soccer Soc* 2022; 23: 44–55.
9. Han B, Chen Q, Lago-Peñas C, et al. The influence of the video assistant referee on the Chinese Super League. *Int J Sports Sci Coach* 2020; 15: 662–668.
10. Ridhwan KM, Hargreaves CA. Leveraging Twitter data to understand public sentiment for the COVID-19 outbreak in Singapore. *Int J Inf Manag Data Insights* 2021; 1: 100021.
11. Soleymani M, Garcia D, Jou B, et al. A survey of multimodal sentiment analysis. *Image Vis Comput* 2017; 65: 3–14.
12. SV P, Ittamalla R. General public's attitude toward governments implementing digital contact tracing to curb COVID-19—a study based on natural language processing. *Int J Pervasive Comput Commun* 2022; 18: 485–490.

13. Nasukawa T, Yi J. Sentiment analysis: Capturing favorability using natural language processing. In: *Proceedings of the 2nd international conference on Knowledge capture*. 2003, pp. 70–77.
14. Mohammad SM, Turney PD. Crowdsourcing a word–emotion association lexicon. *Comput Intell* 2013; 29: 436–465.
15. Albergaria MFBS de. *Near Real-Time Sentiment and Topic Analysis of Sport Events*. PhD Thesis, 2022.
16. Blei DM. Latent dirichlet allocaton. *J Mach Learn Res* 2003; 3: 993–1022.
17. Hidayatullah AF, Pembrani EC, Kurniawan W, et al. Twitter topic modeling on football news. In: *2018 3rd International Conference on Computer and Communication Systems (ICCCS)*. IEEE, 2018, pp. 467–471.
18. Sokolova M, Huang K, Matwin S, et al. Topic modelling and event identification from Twitter textual data. *ArXiv Prepr ArXiv160802519*.
19. Zhao WX, Jiang J, Weng J, et al. Comparing Twitter and Traditional Media Using Topic Models. In: Clough P, Foley C, Gurrin C, et al. (eds) *Advances in Information Retrieval*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 338–349.
20. Vale AS do. *Analysis of Video Assistant Referee (VAR) in football*. MasterThesis, <https://run.unl.pt/handle/10362/148992> (2023, accessed 15 March 2023).
21. Winand M, Schneiders C, Merten S, et al. Sports fans and innovation: An analysis of football fans’ satisfaction with video assistant refereeing through social identity and argumentative theories. *J Bus Res* 2021; 136: 99–109.
22. Zglinski J. Rules, Standards, and the Video Assistant Referee in Football. *Sport Ethics Philos* 2022; 16: 3–19.
23. University of Liechtenstein, Debortoli S, Müller O, et al. Text Mining for Information Systems Researchers: An Annotated Topic Modeling Tutorial. *Commun Assoc Inf Syst* 2016; 39: 110–135.
24. Bhutada S, Balaram VVSSS, Bulusu VV. Latent Dirichlet Allocation based multilevel classification. In: *2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT)*. 2014, pp. 1020–1024.
25. d’Andréa C, Stauff M. Mediatized engagements with technologies: “Reviewing” the video assistant referee at the 2018 world cup. *Commun Sport* 2022; 10: 830–853.
26. Märtns J, Westmattelmann D, Schewe G. Affected but not involved: Two-scenario based investigation of individuals’ attitude towards decision support systems based on the example of the video assistant referee. *J Decis Syst* 2022; 0: 1–25.
27. Cleland J, O’Gorman J, Bond M. The English Football Association’s Respect Campaign: the referees’ view. *Int J Sport Policy Polit* 2015; 7: 551–563.

28. Kolbinger O, Knopp M. Video kills the sentiment—Exploring fans’ reception of the video assistant referee in the English premier league using Twitter data. *PLOS ONE* 2020; 15: e0242728.