Technische Universität München

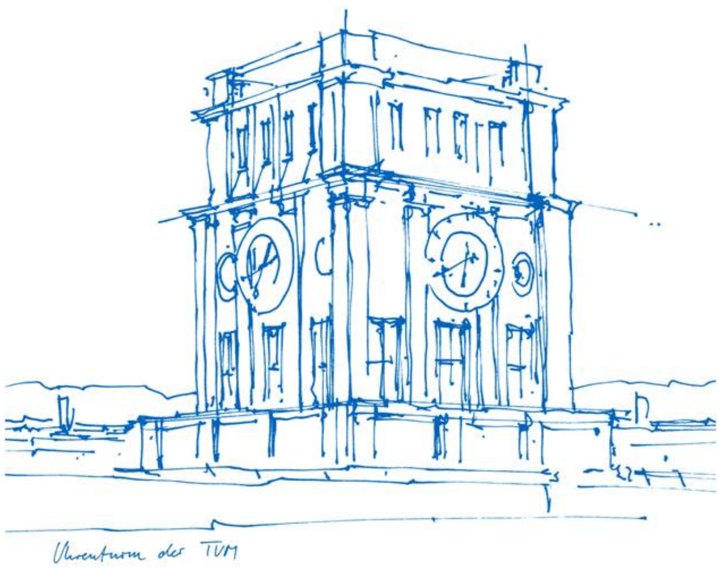
Machine Learning and Lexicon-Based Sentiment Analysis of Twitter Responses to Video Assistant Referees in the Premier League during the 2019-2020 Season

Master's Thesis For Completion of Degree:
M. Sc. Sport and Exercise Science

TUM Department of Sport and Health Sciences
Technical University of Munich

Supervisor:
Dr. Otto Kolbinger
Chair of Performance Analysis
and Sports Informatics
Department of Sport and
Health Sciences

Submitted by:
Melanie Knopp
Melanie.Knopp@tum.de
Arcisstraße 21,
80333 München



Submitted:
Munich, July 6, 2020

Acknowledgment

First of all, I would like to thank my thesis advisor Dr. Otto Kolbinger for giving me the opportunity to write this thesis in a manner most fitting to me and for his patience and support. Whenever I ran into trouble with my writing or needed coding or other assistance, he pointed me in the right direction. Additionally, his understanding of my situation was very much appreciated.

An additional thanks goes to my instructors both at Pepperdine University and at the Technical University of Munich for providing me with the tools and knowledge necessary not only to be successful in my career but also to develop a passion for this field.

Special acknowledgment also goes to Marcel Knopp (TUM Biomedical Computing) for assisting with the processing of the data files as he has a much more powerful computational environment than my laptop.

I would also like to acknowledge Dr. Michael Knopp for his valuable comments on this thesis.

Furthermore, I would like to express my gratitude to Dr. Daniel Ruiz for believing in me every step of the way and for being so understanding.

Finally, I would like to thank both the Menter's, particularly Chris, and my family for their continuous support and encouragement and for giving me a balance to the situation.

This accomplishment would not have been possible without all of them. Thank you!

Munich, July 6, 2020 MELANIE KNOPP

I. Abstract

Introduction: In soccer, the video assistant referee (VAR), which refers to a match official who reviews video footage and communicates with the head referee, has relatively recently been introduced to the game. While previous research has examined the descriptive statistics of the impact of VAR on soccer, so far, no research has examined fans' reaction to its implementation. One way to study how fans feel about its introduction to soccer is to use Twitter analysis collecting input from many different fans all around the world. Therefore, this report examines the use of Twitter as a technology to obtain data from fans in real time coordinated through various time zones in response to live events as its happening during a broadcasted sports event. In more detail, the aim of this study is to analyze how sports fans on Twitter react during soccer games in response to VAR used in the English Premier League via sentiment analysis to examine the emotions demonstrated in the published tweets.

Methods: For this study, the Twitter data, made up of 643,251 tweets posted during the examined games of the 2019-2020 English Premier League season, was collected and analyzed. Two different text mining techniques were applied for both event detection and sentiment analysis. The first method consisted of a frequency analysis of the content of tweets to determine if and when the VAR was involved and a lexicon-based system to calculate the corresponding sentiment. The second method attempted to do both event detection of VAR incidents and sentiment analysis with the help of a supervised machine learning algorithm. Since the machine learning method performed better, this was used to categorize or label each tweet accordingly.

Results: The results showed VAR tweets have a significantly lower sentiment value than the collected non-VAR tweets. Furthermore, when games were split into quantiles and examined, it was found that VAR tweets were overrepresented in quantiles with low average sentiment values. When examining triggers to VAR peaks, the sentiment of the period *before* the trigger was found to be significantly higher than the period directly *after* the trigger as well as higher still than the *later* period.

Conclusions: Distinct results emerged suggesting fans on Twitter express a significant negative opinion towards VAR usage during the examined games. Indicating that, in this case, the main objective of VAR to correct clear and obvious errors without significantly interrupting the game is not being achieved. While the findings in this research are based specifically on the Premier League's usage of VAR, implications can be highlighted for other leagues and tournaments about the potential reactions from fans if VAR is administered in the same way.

Keywords: Video Assistant Referee, Soccer, Football, Twitter, Social Media, Text Mining, Event Detection, Sentiment Analysis, Sentiment Lexicon, Machine Learning

II. Table of Contents

III. Table of Figures

IV. List of Tables

1. Introduction

1.1 Video Assistant Referee

During a soccer match, referees are faced with the difficult task to make the correct judgment, according to the rule book, in situations that involve fast movements, multiple players, different cues coming from different sources, and sometimes limited visibility (Lex, Pizzera, Kurtes, & Schack, 2015). With all this going on, referees' jobs are susceptible to human error due to the limits of the referees' perception which can easily but unintentionally result in judgment errors and biases (Lago-Peñas, Rey, & Kalén, 2019). In the past years, there has been an increased usage of technologies during sports games, giving spectators more real-time information about an event, however, these are not being used to assist referees, whose decisions are highly constrained at that moment by human limitations. For example, commentators, coaches, and spectators all find themselves in the position of scrutinizing the referees' decision, while having access to modern technologies allowing them to watch actions again in slow-motion replay as well as seeing an incident from various different angles (Ugondo & Tsokwa, 2019). Therefore, it is no surprise that multiple technological aids are continually being introduced to assist officiates in several different sports in an attempt to reduce the incidence of controversial decisions and lead to fairer competition; in soccer, this includes the previously introduced goal line technology, vanishing spray, and the video assistant referee (Kolbinger & Lames, 2017; Kolbinger & Link, 2016; Lago-Peñas et al., 2019; Oudejans et al., 2000; Ugondo & Tsokwa, 2019).

As hard as referees try to have a flawless performance, a referee's job is always going to be somewhat subjective as they are only human. Previous studies have demonstrated that a referees' decision can be influenced by things such as social pressure, the strength of the teams, the match status, and crowd noise (Boyko, Boyko, & Boyko, 2007; Dohmen, 2008; Garicano, Palacios-Huerta, & Prendergast, 2005; Riedl, Strauss, Heuer, & Rubner, 2015; Sutter & Kocher, 2004; Unkelbach & Memmert, 2010). For example, referees were found to show a type of bias for the home team with officiates adding a significantly greater amount of stoppage time after the second half when the home team was behind by one goal than when it was ahead by one goal (Garicano et al., 2005). To determine how crowd noise affects a referees' decision, one study found that referees who watched videotaped footage of potential fouls with noise were significantly affected by the crowd with fewer fouls being called for the away team (Nevill, Balmer, & Williams, 2002). Additional suggested biases also include referees favoring superior teams when they are playing in close games

(Lago-Penas & Gomez-Lopez, 2016). One way to try to avoid such biases is to introduce a form of video assisted replays.

While various sports including American football, rugby, golf, and ice hockey have consistently been using video assisted replays already for many years, referees in soccer have only recently begun using videos to assist them in their decisions (Lago-Peñas et al., 2019). One reason for this delay is that the governing bodies of soccer have long been opposed to the introduction of any technology that removes the responsibility from the referees including both goal line technology and instant replay by arguing that such technology undermines the game as soccer "must be played in the same way no matter where you are in the world" (Fédération Internationale de Football Association, 2010). Additional concerns were also raised about how such technologies would detract from the atmosphere of debating controversial goals and other crucial decisions that most fans' of the sport treasure (Winand & Fergusson, 2018). Even with this reluctance, controversial instances continued to be present in the game forcing the governing bodies to find a solution and not appear to be out of touch. This at first brought about the introduction of goal line technology into the game of soccer in 2012, whose success led the way to introduce the use of video refereeing (Sport, 2016; Winand & Fergusson, 2018). In soccer, the video assistant referee (VAR), which refers to a match official who reviews video footage and communicates with the head referee about decisions made on the field, was first introduced into the game in 2016 by the International Football Association Board (IFAB) as a system that would bring greater fairness to the game by correcting "clear and obvious errors" or "serious missed incidents" while not causing a significant interruption under the philosophy of "minimum interference, maximum benefit" (International Football Association Board, 2016, 2017, 2019; Medeiros, 2018). The main criticism that came about in its introduction is that using video-replay in soccer will disrupt the flow and pace of the game due to the additional stopping and starting that would be required (Dyer, 2015; Svantesson, 2014). Since then this system has been trialed in various soccer associations and tournaments with the Fédération Internationale de Football Association (FIFA) incorporating it into the 2018 World Cup and has quickly become a part of the professional game (International Football Association Board, 2018; Medeiros, 2018). Finally, for the 2019-2020 season, the Premier League also introduced VAR into its competitive matches (Premier League, 2018b).

VAR is used as a technology that supports referees in decision making and can be used in four different situations: goals, penalties, red card incidents, and mistaken identities (Lago-Peñas et al., 2019). In a VAR review, the video footage obtained during the specific incident in question is reviewed by another referee with the help of additional video analysis tools, who then communicates with the head referee on the field via a headset until the appropriate action is taken (Fédération Internationale de Football Association, 2018; Lago-Peñas et al., 2019). During a soccer match, the VAR team automatically checks every on-field referee decision falling into one of the four reviewable categories (International Football Association Board, 2018). Based on the information obtained from the VAR, the referee has three options: they can immediately overturn the original call based on the recommendations from the VAR, they can review the incident themselves on a pitch-side monitor, or they can stick with their initial decision (Platt, 2019). During this process, the VAR gives recommendations and assists with reviews but the head referee on the field is always the one who has the final say (Standard Sport, 2018).

On average, it was found that during an international match, a referee needs to make more than 130 observable decisions, all of which could be influenced by a variety of different factors (Helsen, Gilis, & Weston, 2006). In soccer, inaccurate decisions made from the referee are relatively frequent due to the nature of the job and can have a direct impact on the final result of an important match. Therefore, erroneous decisions can result in significant financial implications for many different stakeholders including clubs, managers, and players (Helsen et al., 2006; Kolbinger & Lames, 2017). For example, when it comes to calling offsides during a soccer game, previous studies have examined the human perceptional limitations in determining a correct offsides position (Helsen et al., 2006; Kolbinger & Lames, 2017). Here, perceptional limitations can be due to poor positioning where the assistant referee is not in step with the offside line and therefore cannot correctly determine a call (Oudejans et al., 2000). Therefore, implementing new technological aids like VAR are important as they contribute to enhance the quality of refereeing (Lago-Peñas et al., 2019).

Previously, the VAR system was introduced in other top European soccer leagues including the Bundesliga (German first division) and the Serie A (Italian first division) during the 2017-18 season (Lago-Peñas et al., 2019). Research conducted with these two leagues during the aforementioned season had examined how the introduction of the VAR system influenced the elite soccer game and found that overall the game is not significantly modified (Lago-Peñas et al., 2019). Although this study found that after the implementation of VAR there was a significant decrease in the number of offsides, fouls and yellow cards, there was an increase in the number of minutes added to the playing time in the first half (15 seconds) and the full game (20 seconds), but not in the second half, and additional individual differences were found when comparing seasons with and without VAR in the Italian Serie A and the German Bundesliga (Lago-Peñas et al., 2019). These differences were explained away by other reasons. This study declared that the extra time that the VAR system adds to the game of soccer was not substantial, and the decrease in the number of fouls and yellow cards could be explained by the fact that players were playing less aggressively and had to be much more careful with their behavior during the game with the new officiating aids (Lago-Peñas et al., 2019). A similar trend in reducing the extent of rule violations was found when vanishing spray, which assists referees in enforcing specific rules by providing a temporary visual marker, was introduced to soccer (Kolbinger & Link, 2016).

However, certain drawbacks of VAR must also be considered. One study looking at how people perceive videos that are played in slow motion discovered that slow motion video review can distort reality and cause the viewer to perceive an action, such as violent contact in professional football or even murder footage, as more intentional (Caruso, Burns, & Converse, 2016). This slow motion bias occurs because it causes viewers to believe that the actor had more time to respond, even when they are aware of how slowed down the actual clock is (Caruso et al., 2016). A similar study focusing particularly on the effect video speed has on soccer referees found that in situations like offsides, viewing replays in slow motion may have an added value on the decisional accuracy of the referee as they require specific evaluations of spatial and temporal landmarks (Spitz, Moors, Wagemans, & Helsen, 2018). However, in foul play situations, this study also found that referees penalized players more severely when watched in slow motion compared to regular speeds as it increases the perceived intent of violent action (Spitz et al., 2018).

1.2 Social Media and Twitter in Sports

Social media has previously been defined as "the sharing of information, experiences, and perspectives through community-oriented websites" and pertains to a wide variety of devices, virtual worlds, and social networking sites that as a realm is consistently growing (Weinburg, 2009). Social media is said to represent a portion of the overall population, however, services such as Twitter comprises only a small part of that (Billings, 2014). With the emergence of social media, communication methods in multiple different industries have drastically changed partly due to the speed at which information can be shared and the range of people who can be reached even in real time beyond the local venue (Witkemper, Blaszka, & Chung, 2016). One societal institution, particularly impacted by the relatively recent arrival of social media is the world of sports as more and more media organizations, teams and athletes use this platform to connect with their audiences (Hutchins, 2011; Pedersen, 2014). For example in the realm of mediated sports, social media presents its users with the potential to connect on a digital platform in real time either before, simultaneously, or after consuming sports content either in person or through other mediated channels (Hutchins, 2011). Users here are defined as an individual, groups of individuals, or organizations that produce and/or consume content that is presented through social media (Witkemper et al., 2016).

Since its introduction in 2006, one microblogging platform in particular, Twitter, has emerged as a persistent channel in sport communication, as it allows its members to connect with millions of other users worldwide who share a similar enthusiasm for a player, team, or sporting event (Clavio & Kian, 2010). Microblogging can be described as a form of blogging that entails the conveyance of personal opinion, news, and/or ideas in an online setting via short bursts of content (Clavio & Kian, 2010). The platform itself boasts about having an average of 500 million tweets posted per day, coming from their 330 million monthly active users of which 145 million are active daily (Lin, 2019). Twitter demographics reveal that 38% of all Twitter users worldwide are between the ages of 18 and 29, while 26% are 30-49 years old, this means 64% of all users are between 18 and 49 years old (Aslam, 2020; Lin, 2019). These users have a ratio of male to female being roughly two to one with 66% of users being male and 34% are female (Lin, 2019). On Twitter, users are equipped with a unique stage to express themselves: tweets are short, limited to 280 characters, and are therefore an ideal way for people to directly react to live events for instant communication among individuals who may or may not have necessarily been connected before (Zhao, Zhong, Wickramasuriya, & Vasudevan, 2011b). This set up has allowed Twitter to become particularly prominent in a variety of uses in addition to sports including interpersonal (Marwick & Boyd, 2010), crisis communication (Twitter Inc, 2020; Woodford, 2013), and political debate, among others (Ott, 2017; Shamma, Kennedy, & Churchill, 2009).

Additionally, with the hashtag, a unique attribute of Twitter, users are able to connect and interact over a particular phenomenon while creating a digital archive for this topic that can then be used by researchers (Rodriguez, 2016). On the platform, a distinct term or phrase is preceded by "#," and is incorporated into the tweet. Upon publication, these hashtags identify tweets related to the same topic (Frederick, Burch, & Blaszka, 2015; Rodriguez, 2016). Research has also shown that hashtags are a tool preferred predominately by the daily user as a way to convey a specific opinion or comment to a greater audience (Frederick et al., 2015). Ordinarily, a user's tweet is limited to the people who 'follow' (subscribe to) that user and then only at that particular moment, meaning generally a user's communicative reach is limited to the size of their own social network (Marwick & Boyd, 2010). However, when published containing hashtags, tweets, in turn, address the entire community of users following a specific topic or, in this case, a sporting event (Highfield, Harrington, & Bruns, 2013). It is important to note, however, that the usage of hashtags is not guaranteed, users can change their settings to only allow those in their follower network to see their comments or discuss current events without the usage of a hashtag (Highfield et al., 2013). Additionally, it is also possible that for some major events multiple hashtags exist meaning not all comments may be documented together (Highfield et al., 2013). Within the sports industry, the utilization of hashtags has increased with athletes, teams, sports organizations, and fans creating unique, event-specific hashtags to direct conversations on the platform and to centralize communication (Frederick et al., 2015). This occurs often for largescale sporting events such as the Olympics, European Championships, or the FIFA World Cup (Frederick et al., 2015; Panzarino, 2012; Stranges, 2019). In fact, during the 2018 World Cup, FIFA specifically teamed up with Twitter to generate more excitement about the tournament by creating a unique mediated sport event that engaged their fans with everything from Twitter polls and trivia contests to exclusive emojis and FIFA-curated moments (Stranges, 2019).

With its unique features, it is no surprise that Twitter has become a significant permanent fixture in the sports industry (Pegoraro, 2010; Wertheim, 2011). Sports is one of the most read topics on social media, with Twitter usage topping the list in this context when compared to other similar platforms (Mitchell, Kiley, Gottfried, & Guskin, 2013; Shearer, Barthel, Gottfried, & Mitchell, 2015). This is partly because, as a platform, Twitter itself has also been pushing its usage in the realm of sports with major sporting events constantly breaking records for the most tweeted about events (Rodgers, 2014). Overall, Twitter is one of the most popular and preferential social media platforms to be used by organizations, media outlets, athletes and sports fans alike all around the world (Boehmer, 2015; Tiago, Tiago, Faria, & Couto, 2016).

Professional athletes utilize Twitter for broadcasting and connecting with their fans in all sorts of different ways whether its interactivity, diversion, information sharing, content sharing, promotional, or fanship (Frederick, Lim, Clavio, Pedersen, & Burch, 2014; M. Hambrick, Simmons, Greenhalgh, & Greenwell, 2010; Williams, Chinn, & Suleiman, 2014). For example, one time Lance Armstrong posted on Twitter asking his fans to come along for a bike ride at a clarified time and location. Several hours later, more than 1,000 cyclists showed up to participate in an afternoon bike ride with the famous cyclist. Armstrong embraced this opportunity with his followers and spent the whole time shaking hands, speaking to fans, and having his picture taken giving fans a once in a lifetime opportunity for this interactivity with their sports role model (Cromwell, 2009). Another study examining Lance Armstrong and Serena William's Twitter usages found that these celebrity athletes used Twitter for promotional messages for corporate sponsors and products, charitable organizations, and their own personal activities 12% of the time (M. E. Hambrick & Mahoney, 2011).

In addition to athletes, most well-known professional team homepage's features an eminent Twitter link to gather thousands of their supporters (Gibbs, O'Reilly, & Brunette, 2014). Sports organizations, use this platform to promote their team, build relationships with their supporters, increase fan engagement, and drive revenue by using it as a marketing platform for merchandise and displaying game attendance, for example (Armstrong, Delia, & Giardina, 2016; O'Shea & Alonso, 2011; Williams et al., 2014). This is particularly prominent in soccer as soccer teams completely dominate the list of top 10 most followed sporting institutions on social media (Martín, 2019). Real Madrid sits at the top place of Twitter with 33 million followers (Martín, 2019). Real Madrid has achieved this digital success by exclusively revealing the club's starting line up on Twitter and keeping its followers up to date with live text, key club news items, and leading photo and video content (Madrid, 2018).

For sports fans, Twitter is an important channel for a more personal connection to their favorite sports organizations and athletes than what occurs through mass media broadcast (Williams et al., 2014). When following their beloved sports organizations on Twitter, fans get easy access to instantaneous information about the team from both official sources as well as insider information or behind-the-scenes activity (Williams et al., 2014). All of this has been shown to increase fan involvement and strengthen individuals' bonds to an organization allowing them to feel closely connected to the team (Williams et al., 2014).

Researchers have argued that studying Twitter is of vital importance in order to better understand both sport's relation to society and vice versa and has therefore recently become a major academic research area (Sanderson, 2014). The social media platform has been used by scholars to examine, the interactions among spectators, fans, athletes, and sports teams at major sporting events such as the 2012 London Summer Olympics (Frederick et al., 2015), the Sochi 2014 Winter Olympics (Girginova, 2016), as well as both men's and women's FIFA World Cups (Hayes Sauder & Blaszka, 2018; Yu & Wang, 2015). Additionally, sport-specific research has examined Twitter usage patterns among professional athletes (Frederick et al., 2014; M. Hambrick et al., 2010; Hull & Lewis, 2014; Kassing & Sanderson, 2010; Lebel & Danylchuk, 2012; Pegoraro, 2010), college athletes (Browning & Sanderson, 2012), fans (Clavio & Kian, 2010; Frederick et al., 2014), organizations (Clavio, Burch, & Frederick, 2012; Sanderson, 2011), and sporting events (Blaszka, Burch, Frederick, Clavio, & Walsh, 2012; L. R. Smith & Smith, 2012). Furthermore, mediated platforms provide a space of overlap to analyze different types of language social theories as one study used Twitter to determine the cultural specific connotation of words debated to be allowed by fans for sports events (Rodriguez, 2016). Here, researchers showed the potential of using Twitter to determine the connotation of a word used in a chant that some considered to be a homophobic slur while others merely saw it as a distraction technique by fans (Rodriguez, 2016). All of these studies combined show the potential this platform has for all kinds of research in the realm of sports.

1.3 Twitter as a Backchannel for Television and Current Events

When compared to other social media platforms, Twitter is the most relevant source of information for event-related use, as the primary goal of the individual on Twitter is to follow and contribute to the mass-oriented discussion of current events rather than acquiring sports-related information for which Facebook was found to be more appropriate (Boehmer, 2015). The briefness, rapidness, and potential global impact of tweets allow for broad discussions about current events happening on television, with Twitter claiming that up to 95% of public social media conversations about such topics occur on their platform (Graver, 2012; Zhao et al., 2011b). In such uses, opposed to initial concerns, Twitter does not replace the traditional media channels including conventional broadcasting and online mainstream media but rather acts to complement the experience and is therefore considered a type of backchannel to live programming (Harrington, Highfield, & Bruns, 2012). This allows for the incorporation of an ‘active’ audience, which was first introduced in media studies several decades ago, and has encouraged programming stations to recognize this medium as an important part of the audience experience (Harrington et al., 2012).

Researchers have used this data generated by the active Twitter audience to gain an advanced, vast, and instantaneous measure that represents “empirical evidence … of how other people make sense of the world” (McKee, 2003, p. 15). While it is clear that Twitter does not provide a true representation of any one population, previous research has declared that it can be used to provide insight into what users are talking about at a particular point in time (Ovadia, 2009). By studying the audience’s reactions to key moments of a live broadcast, researchers have been able to explore audiences opinions of candidates in a political debate, for example (Shamma et al., 2009).

This concept of social television through applications has also become increasingly popular during live sporting broadcasts (Fan, Billings, Zhu, & Yu, 2019). For example, during the 2018 World Cup, Twitter recorded a total of 115 billion impressions of tweets, referring to the number of times tweets related to this World Cup appeared to users (Bavishi & Filadelfo, 2018). These individuals often use a second screen, which allows them to simultaneously engage in both types of media; posting their own commentary to express their emotional response or looking for additional information about an event as it unfolds live on television for example (Boehmer, 2015; Highfield et al., 2013). Most of the time this second screen refers to a phone, for example, as 80% of Twitter users are registered on their mobile device (Aslam, 2020). One study found that 79% of participants used second screens for social media interaction such as posting or commenting on Twitter while watching a sports event or other live broadcasts on a primary screen (Cunningham & Eastin, 2017).

Various motivations have been researched as to why second screens are used and include: the opportunity for increasing social interaction with others and also to reduce the fear of missing out on additional information about the phenomenon (Larkin & Fink, 2016). Such practices have been reported to enhance users' experience by providing them with additional options to engage thereby increasing their overall enjoyment (Highfield et al., 2013; Winter, Krämer, Benninghoff, & Gallus, 2018).

The activity of Twitter users during specific televised events such as a World Cup Final or UEFA Champions League match makes up a unique category of Twitter use. This clearly differs from other major events that also generate a high rate of tweets per second such as major news breaking events including a devastating natural disaster, for example. Activity on Twitter during sports events is dedicated to fans wanting to offer their own real-time, relatively unmediated commentary in reaction to live events for a communal discussion (Highfield et al., 2013). For such events, Twitter becomes a sort of "unofficial extension" of a particular broadcast (Highfield et al., 2013; Panzarino, 2012).

1.4 Who is on Twitter, Why are they using it, and to Whom are they talking?

When it comes to fans posting on the microblogging platform, these individuals generate what is called user-generated content (Witkemper et al., 2016). User-generated content refers to content that is produced by the general public rather than by paid professionals and is made publicly available primarily through the internet (Daugherty, Eastin, & Bright, 2008). It has been researched that creators publish this content because it helps them comprehend their environment, the topic at hand, and/or ultimately themselves as it offers them a sense of understanding and a way to process (Daugherty et al., 2008). Authors participate on such platforms as it can minimize their own self-doubts and present users with a sense of belonging (Daugherty et al., 2008).

One reason why people use specific social media platforms has been explained by the uses and gratification approach. In the study of communication methods, this approach is an audience-centered method often applied to assess how individuals use new methods of media to fulfill certain personal needs (Fisher, 1978). This is based on the idea that individuals participate in certain media behavior for a specific purpose: in this realm, the audience members are able to select from a variety of different options to determine which one is best suited for their needs (Rubin, 2009; Tan, 1985). This approach has already been applied to various internet platforms and social-media technologies including Twitter (Blaszka et al., 2012; G. Chen, 2011; Clavio & Kian, 2010; M. Hambrick et al., 2010). Since people are picking Twitter and sticking with it, it means Twitter is meeting the needs of its users in some way (G. Chen, 2011).

To determine what motivates individuals to use Twitter, one study, in particular, applied this model and focused on the gratification of Twitter to meet individual's needs to connect with others (G. Chen, 2011). This study shows, similar to other previously published research, that people who specifically seek out Twitter as a preferred social media platform are doing so to fulfill the basic human need to connect with others by using this computer medium to create a sense of virtual community (G. Chen, 2011; Clavio & Kian, 2010; M. Hambrick et al., 2010; Johnson & Yang, 2009). These findings support the idea "that Twitter is not just virtual noise of people talking at each other," as some critics have previously claimed, but instead, people distinctly seek out this platform to fulfill their need to connect with others (G. Chen, 2011). Another study elaborated on these findings by researching various different motives and found that users were most satisfied with Twitter's ability to help them pass time, meet new people, participate in discussions, communicate with many different people at the same time, express themselves freely, and get insight into what others are up to (Johnson & Yang, 2009). While the platform was originally created for its social aspects, data suggests that many users see it also as an information source and a way to share information (Blaszka et al., 2012; Johnson & Yang, 2009). The high motivation for information sharing via Twitter has made clear the satisfaction individuals gain from providing information to others (I. Liu, Cheung, & Lee, 2010).

While uses and gratifications research has been plentiful in the general realm of internet technologies, limited studies have focused particularly on the sport-specific context. One study that utilized this approach when examining Twitter usage trends among professional athletes, found they use the platform mainly for interaction and expressions of diversion in the form of entertainment or distraction (M. Hambrick et al., 2010). This study also found that information sharing and content-related tweets were used moderately, while promotional and fanship tweets were rather infrequent (M. Hambrick et al., 2010). When using this theoretical approach from the fan perspective, one study found that the primary motivation for following the Twitter feed of a retired female athlete was her expertise in her given sport (Clavio & Kian, 2010). This was followed by a fondness for this athlete's writing style. An additional factor analysis conducted as to why individuals follow specifically this athlete revealed three main factors including organic fandom, functional fandom, and interactivity (Clavio & Kian, 2010). In this study, organic fandom describes personally oriented fandom explained by a liking for the athlete herself and includes things like viewing the athlete as a role model and being interested in her career and expertise in her sport (Clavio & Kian, 2010). Functional fandom is more related to the impersonal elements of fandom like being interested in purchasing the athlete's products and is more business-related (Clavio & Kian, 2010). Finally, interactivity refers to using Twitter to interact with the athlete as well as with other fans (Clavio & Kian, 2010).

Another model used to determine what sports fans are looking for on Twitter, found that the platform fulfills four primary satisfactions: interaction, promotion, live game updates, and news (Gibbs et al., 2014). Interaction tweets were related to fan discussion that involved the teams or fans of a game or encouraged discussion about what was happening (Gibbs et al., 2014). Promotional messages and marketing tweets were included in the promotion category (Gibbs et al., 2014). Live game updates had its own category and includes score, substitution, and time updates, more along the lines of a traditional media platform for sports games (Gibbs et al., 2014). News tweets included things like team-roster updates and links to videos about the game such as pre-game thoughts (Gibbs et al., 2014). Another study categorizing tweets differently found the four main motivational factors that affected fans' usage of Twitter in a positive manner were information, entertainment, to pass the time, and fanship (Witkemper et al., 2016). All these studies suggest that Twitter users consist of various individuals all using the platform for different reasons.

When it comes to users posting about sporting events, one study found that individuals who feel that they themselves are considered sports experts are more likely to turn to and post on Twitter while watching a sports event on television (Boehmer, 2015).

Previous research has also been conducted to determine who social media users believe their audience is. On Twitter, while each user has their specific set of followers, the vast majority of Twitter accounts and communication are public meaning anyone can be reading them (Marwick & Boyd, 2010). This means, as with most other forms of computer-mediated communication, a tweet's actual readers may be different from the audience it's producers had imagined when constructing the tweet (Marwick & Boyd, 2010). One study researching how Twitter users imagine their audiences found that, when generalized, individuals address their friends, fans, and often themselves either in a form of a diary or record of their lives or they use the platform as a way to express opinions for themselves (Marwick & Boyd, 2010).

One study examined specifically who was using the sports hashtag, #WorldSeries, during the 2011 World Series, and how it was being used on Twitter (Blaszka et al., 2012). This study found that most of the individuals using this hashtag during the event were laypersons (Blaszka et al., 2012). When examining the content of these tweets, researchers found that most were centered on fanship—with tweets rooting for a specific team for example—or interactivity—with fans engaging mostly with each other, or league officials and asking questions (Blaszka et al., 2012). This usage of Twitter to satisfy fanship and interactivity aligns with what was found in the previously mentioned research (Clavio & Kian, 2010; M. Hambrick et al., 2010). When comparing the results of this analysis to other studies, by examining solely a hashtag, users were found to be tweeting primarily as an act of personal expression (Blaszka et al., 2012).

1.5 Event Detection with Twitter

Previous studies have shown that Twitter not only gives useful information about major social and physical events including things like earthquakes, stocks, celebrity deaths, and presidential elections but also events that occur during sports games such as touchdowns in American football or goals in a soccer game (Bollen, Mao, & Zeng, 2011; P. Earle, Bowden, & Guy, 2012; Lucas et al., 2017; Sankaranarayanan, Samet, Teitler, Lieberman, & Sperling, 2009; Shamma et al., 2009; Zhao et al., 2011b). This event detection information can be successfully obtained from Twitter through application interfaces, language processing, and text mining (P. Earle et al., 2012).

Twitter has been shown to be particularly useful for detecting physical events such as an earthquake. Immediately after an earthquake occurs sometimes thousands of tweets are posted describing the shaking effects (P. Earle et al., 2012). Such notifications are in fact sometimes available even before the seismically derived estimates distributed by the U.S. Geological Survey and have been demonstrated as a possible method for detection with a low rate of false triggers (P. Earle et al., 2012; P. Earle, M. Guy, R. Buckmaster, C. Ostrum, S. Horvath and A. Vaughan 2010). While Twitter is in no way a replacement for other well-established systems, Twitter's international reach, easy and quick access, and short communication methods show the potential in supplementing other established systems with the network (P. Earle et al., 2012).

In sports, major game events can also be detected by analyzing tweets collected during the game, such as detecting events during an American National Football League (NFL) game (Zhao et al., 2011b). This can be done by examining the post rate, the frequency at which new tweets are being published on Twitter, during a sports match. Event detection here occurs by measuring the post rate of a specific time window as a ratio of the post rate in the second half of the window compared to the first (Zhao, Zhong, Wickramasuriya, & Vasudevan, 2011a). This was proved successful in detecting pre-determined events such as touchdowns, interceptions, fumbles, and field goals during nearly a hundred NFL games in real-time when these events produced various peaks by comparing post rate to game minute (Zhao et al., 2011b).

Similar methods have also been applied to soccer for the World Cup. As a sport, soccer has some unique advantages over others in this context, as soccer, for example, does not allow for commercial breaks during gameplay, which is particularly useful when trying to relate social media posts to game events (Lucas et al., 2017). For example, one study examined the potential of using Twitter data to automatize the process of creating highlight videos from a sporting event like the World Cup (Hannon, McCarthy, Lynch, & Smyth, 2011). This study implemented two different techniques including frequency-based summaries, where specific events were identified with a higher than normal tweet volume being posted during specific game minutes, and content-based summaries, where specific terms are set as a trigger to select the desired sequences of the game (Hannon et al., 2011). In this case study, the users who watched the automatically created highlight videos were satisfied with the summary of the game showing the potential of implementing Twitter data in various different ways (Hannon et al., 2011). A similar methodology was once again used to create summarization videos during the 2014 World Cup and one study analyzed specifically the precision of using social media streams to detect these highlights and found bursts of tweets can accurately identify them during a soccer match (Jai-Andaloussi, Mohamed, Madrane, & Sekkaki, 2014). Similarly, one exploratory study has demonstrated how Twitter analysis can be used in soccer describing how, by dynamically analyzing tweets per minute in a game, specific events in the match such as halftime or a goal can easily be detected (Lucas et al., 2017). This study collected tweets that contained at least one of the official hashtags of the World Cup, as was released by FIFA, for the two teams playing in a match from which the tweets per minute were recorded (Lucas et al., 2017). As explained by the researchers, this data showed peaks for particular events of the game including things like game start, halftime, goals, and end of the game (Lucas et al., 2017).

1.6 Sentiment Analysis

Sporting matches are known to evoke strong emotions from fans that change continuously throughout a game and therefore sports have been considered an emotional laboratory for investigating various social and psychological theories (Lucas et al., 2017; Yu & Wang, 2015). Research has claimed that fans view mediated sports specifically to be stimulated allowing these individuals to experience emotional arousal and release both their positive and negative emotions (Raney, 2006). Here sports games act as an entertainment source that can bring its spectators joy and happiness when one's team is winning or anger and sadness when the opposing team scores, for example (Yu & Wang, 2015). Such differing events means sports fans' emotions vary as a function of what is happening during the game (Yu & Wang, 2015).

With the growing emergence of social media and platforms such as Twitter allowing people to react to live events of a sports match, future research can be expanded to include greater fan input. Previous research in social media and electronic communication has reported that one of the main motivations for sports fans to use social media during a game is for emotional release, and a significant relationship was found with sports fans' game enjoyment and their objective to use social media (X. Wang, 2013, 2015). With this in mind, previous studies have declared the potential in using social media to measure viewers' reactions and sentiment changes to mediated sports programs (Yu & Wang, 2015).

Sentiment analysis is one of the fields of natural language processing (Ljajić, Ljajić, Spalević, Arsic, & Vučković, 2015). Various sentiment analysis techniques have previously been applied to a variety of different fields including business, education, and politics (Ceron, Curini, Iacus, & Porro, 2014; Pang & Lee, 2008; Tumasjan, Sprenger, Sandner, & Welpe, 2011). A growing number of publications have already used Twitter to examine users' responses through social media sentiments (B. Liu, 2012; Pang & Lee, 2008).

Previous studies have demonstrated the possibility of using sentiment analysis for tweets to quantify the general emotion of the audience. Sentiment in sport has been described to be connected to an individual's involvement in an event, their attachment to the specific sport type, and their identification as a fan (Funk & James, 2001; Holmes et al., 2007; Sutton, McDonald, Milne, & Cimperman, 1997). Sentiment analysis attempts to evaluate individuals' emotions, attitudes, or opinions in response to different issues (B. Liu, 2012). Most research involving sentiment analysis is based on either existing linguistic resource packages or machine learning. Such linguistic resources include pre-made libraries and packages that are made up of lists of positive and negative words that can then be counted up to determine the expected overall sentiment of a text: either favorable, unfavorable, or neutral (Taboada, Brooke, Tofiloski, Voll, & Stede, 2011; Yu & Wang, 2015). While this methodology is an update from the traditional, time-consuming method of using human coders to analyze each text, using such methods can be problematic as some contextual information, sarcastic phrases, and words with double meanings could be incorrectly interpreted (B. Liu, 2012; Yu & Wang, 2015).

Another study claimed that game sports such as soccer are one of the best situations to analyze sentiment due to the routine nature of the sport because what is good for one team is bad for the other (Jai-Andaloussi, Mourabit, Madrane, Chaouni, & Sekkaki, 2015). In a study about the 2014 World Cup, Yu and Wang (2015) examined how U.S. soccer fans' emotions change on Twitter in response to real-time events that happened when the U.S. Men's National Soccer Team was playing. This study found that emotional patterns reflected the status of their team during the game with negative emotions increasing when the opponent scored and decreasing when the U.S. team scored (Yu & Wang, 2015). Another study used a similar methodology to see how theories of human emotion relate to sentiment analysis during various games in the 2014 World Cup (Lucas et al., 2017). This study found that games with higher tweets per minute were also found to have a higher percentage of negative tweets (Lucas et al., 2017). Additionally, games that ended up having a bigger difference between the competing teams than expected were found to also have a higher percentage of negative tweets (Lucas et al., 2017). Therefore, these researchers suggest that excitement, as communicated via Twitter, relates to expressions of negative emotions as opposed to positive or neutral ones (Lucas et al., 2017). Collectively, these studies indicate the potential such a method has to research particularly sports fans' sentiment.

Sentiment analysis is, however, not as straight forward as it might seem. Previous studies have distinguished tweets into three different categories: either objective tweets or subjective tweets made up of either positive or negative sentiment (Barbosa & Feng, 2010). This categorization is completed with two different tasks including identifying sentiment expressions in the tweet, and then determining the polarity of the expressed sentiment whether positive or negative (Davidov, Tsur, & Rappoport, 2010). But even this is not always as simple as it sounds, because a single sentence may contain both subjective and objective clauses (T. Wilson, Wiebe, & Hwa, 2004). The subjectivity of tweets is also dependent on the inclusion of Twitter specific clues such as emoticons, exclamations, and upper case words (Barbosa & Feng, 2010).

When it comes to sentiment analysis for tweets, researchers should distinguish between sentiment on a sentence level and sentiment on an aspect level as both are possible within the limitations of tweets. Sentence level sentiment classification can be regarded as an intermediate step when it comes to the task of assigning an overall sentiment and many limitations must still be noted (B. Liu, 2012). When looking at text on a sentence level additional details might still be missing. For example, it is possible to have complex sentences where different targets have different sentiments (B. Liu, 2012). An example of this is the sentence: *I like football, but I don't like handball,* here the author is expressing different opinions about different sports mentioned in the two parts of the sentence, where one is clearly positive and the other negative. Similarly, while a sentence may have an overall positive or negative tone, specific components can be expressing opposite opinions (B. Liu, 2012). For example, researchers consider this sentence as positive: *Despite the captain not performing his* *best, the team is doing well* (B. Liu, 2012). Here it is true that the overall tone of the author might be positive, however, it contains a negative sentiment about the captain (B. Liu, 2012). And finally, when it comes to comparative sentences, for example: *Spain plays better soccer than Italy,* sentence level sentiment classification is unable to deal with opinions presented in this way (B. Liu, 2012).

For this reason, to ensure a complete sentiment analysis, researchers often focus on specific aspects, or individual parts, and classify those (B. Liu, 2012). For this, the steps are to extract the specific aspect that is to be evaluated and then assign it with the correct sentiment (B. Liu, 2012). On this level, two main approaches are often used in sentiment analysis: the lexicon-based approach and supervised or machine learning approach (B. Liu, 2012). Due to the nature of the language, there are endless different ways people can express positive or negative opinions which greatly limits the accuracy of possible learning algorithms that try to account for various styles of speech (B. Liu, 2012). This is particularly difficult in the world of politics or sports as this includes a complex mixture of subjective opinions as well as many sarcastic sentences (B. Liu, 2012). In relation to sporting events, on Twitter, many users utilize sarcasm which is described as a sophisticated form of speech where writers write the opposite of what they actually mean (B. Liu, 2012). This is very difficult to deal with in written text but can make a huge difference in sentiment analysis because when used it would appear that something may be positive, but is actually negative or vice versa (B. Liu, 2012). With human coding of sentiment analysis, researchers can incorporate contextual information to help determine if sarcasm is being used or not, and even then it is far from perfect. This common speech style is present on Twitter and was found to greatly limit the accuracy of machine learning sentiment analysis systems and is therefore often simply ignored by some researchers (Fan et al., 2019; B. Liu, 2012).

An additional problem in analyzing social media is all the data noise present (B. Liu, 2012). Tweets are full of all kinds of spelling, grammatical, and punctuation errors, all of which make automatic sentiment analysis more difficult as most natural language processing tools require clean data for accurate results (B. Liu, 2012).

1.7 Lexicon-Based Approach For Sentiment Analysis

As mentioned, the lexicon-based approach is one of two commonly used approaches for sentiment analysis. This method is founded on the fact that most indicators of emotion or sentiment come from opinion words that are commonly used to express positive or negative sentiments and are therefore essential tools for sentiment analysis for obvious reasons (B. Liu, 2012). For example, words such as *incredible*, *great*, and *amazing* clearly express a positive sentiment, while on the other hand words like *horrible*, *bad*, and *hopeless* are rather associated with a negative sentiment. A collection of such words is called a sentiment lexicon and can be created in various different ways and be made up of multiple different words (B. Liu, 2012). The lexicon-based method then uses this sentiment lexicon made up of words with their associated sentimental quantitative assignment measured either in a range or simply with positive or negative ratings and incorporates this to compute a sentiment score for a specified document, whether it's a phrase, a sentence, or multiple sentences (Taboada et al., 2011). Here, any words that are not included in the respective lexicon are simply ignored.

While sentiment lexicons are important for sentimental analysis, they are oftentimes considered to be insufficient due to various different issues. First of all, different words have different sentimental meanings in different domains (B. Liu, 2012). For example, the word *soft* used on an online sports community to describe a hockey player would be an insult with a negative sentiment, however, when referring to toy animals it is said to have a positive connotation (Hamilton, Clark, Leskovec, & Jurafsky, 2016). Secondly, it is fairly common that a sentence may contain sentiment words that might not express any sentimental meaning (B. Liu, 2012). For example the sentences: *Can you tell me which soccer cleats are good?* and *If I find a good pair of soccer cleats, I will buy them,* both contain the sentiment word *good*, but in neither case is this word being used to express any sort of opinion on a specific pair of soccer cleats. Thirdly, sarcastic sentences, for example: *What a great pair of cleats! They ripped after just two days*, are particularly difficult to deal with both with or without sentiment words and are, unfortunately for this project, quite common in sports (B. Liu, 2012). Fourthly, many sentences that lack sentimental words are still used to imply opinions and these can be written in many different ways (B. Liu, 2012). One example, *This shoe gets untied every two minutes,* implies a negative sentiment about the shoe due to a problem with the shoelaces without explicitly stating sentimental words. These four are just some of the major challenges present to researchers that use this method.

Nonetheless, various studies have used such lexicons for sentiment analysis. One reports states that when it comes to using sentiment lexicons to classify the sentiment of a document, while sentiment words are able to classify an average of 60% of cases, with more in some domains and less in others, the remaining multitude of cases are very diverse, and isolated, which makes it hard to determine any possibility of a pattern (B. Liu, 2012). Essentially, there seems to be an indefinite number of ways that individuals can use to express either positive or negative opinions (B. Liu, 2012).

One study looking at the NFL introduced the feasibility of this method to extract sentiment responses in sports and how this can be used to describe a match (Zhao et al., 2011a). This study found a simple lexicon-based method effective in identifying sentiment reactions of fan's tweets to major events happening during an NFL game. Here, after detecting an event, this lexicon was used to confirm and recognize what type of event it was and therefore also helped to suppress false alarms for event detection. Then, to continue to grow its lexicon, these researchers, classified synonyms and antonyms of already analyzed sentiment words based on the extracted tweets. This study claims that the limited vocabulary involved in tweets of sports games makes the lexicon-based analysis an effective method for event and sentiment detection (Zhao et al., 2011a). With this method in place, researchers then compared sentiments between fans of different teams by also extracting the team names that appear in the game-related tweets. During the 2011 Super Bowl, at the beginning of the game, both teams' tweets had about the same positive sentiment, then as the game progressed and specific events took place, the trends started to fluctuate apart as they impacted the fans' sentiment oppositely depending which team they support (Zhao et al., 2011a). The feasibility of this method was also found for different types of sports: in soccer for a UEFA (the Union of European Football Associations) Champions League game where relatively few goals occur as well as in an NCAA (National Collegiate Athletic Association) men's basketball tournament where scoring events are common (Zhao et al., 2011a).

Another study even attempted to create their own soccer-specific sentiment lexicon from soccer associated tweets posted in relation to the 2014 FIFA World Cup and the UEFA Champions League 2016/2017 (Aloufi & Saddik, 2018). In the end, this lexicon consisted of 3,479 words and was created using various general sentiment lexicons along with known soccer-related tweets (Aloufi & Saddik, 2018). When tested for its performance, this lexicon performed better than other lexicons used in the experiment (Aloufi & Saddik, 2018).

1.8 Machine Learning Approach For Sentiment Analysis

Machine learning is an implementation of artificial intelligence that uses algorithms to provide systems with the ability to automatically learn through experience without being explicitly programmed (Expert System Team, 2017). Machine learning algorithms work by building a mathematical model that identifies characteristics of interest and patterns in the data and learns from a specific pre-rated document, known as “training data”, from which it builds an automatic text classifier that can then be used to make predictions or decisions (Bishop, 2006; Sebastiani, 2001). One category of such algorithms is supervised learning. Here, starting with the evaluation of the previously labeled data, the learning algorithm builds a theorized function that it uses to predict the output values (Expert System Team, 2017). In this manner, the system is able to analyze and produce outputs for any new input it is given after sufficient learning has taken place (Expert System Team, 2017). The algorithm can also circle back and compare its own output with the correct, pre-labeled output and can modify the model accordingly based on any errors that were found (Expert System Team, 2017). While some limitations still exist, this process enables researchers with the unique ability to analyze massive quantities of data to an extent that was not feasible before (Expert System Team, 2017).

To make use of machine learning, an algorithm is necessary to calculate the desired output from an input (Fan et al., 2019). For example, if a robot is put in charge of sorting through bananas and apples in a grocery store, the first step would be to classify specific features that distinguish the two fruits such as color, hardness, size, and weight. With these features in mind, fruit samples are labeled with their respective values and this data is fed into the machine as training data. The machine then uses this data to build a mathematical model from which it will try to determine the correct output of the training data. By comparing its results to the pre-established labels, the correct importance is assigned to the different variables, strengthening the algorithm until it is able to correctly classify apples from bananas and can then be used to sort through unclassified inputs (Fan et al., 2019).

Machine learning has previously been applied in a variety of different fields including customer evaluation, medicine, general game playing, natural language processing and many more (Daelemans & Hoste, 2002; Deo, 2015; Hastie, Tibshirani, & Friedman, 2009; Prasasti & Ohwada, 2014; Samuel, 1959). One of the big fields in natural language processing is text classification. This is the process of organizing and categorizing free-text according to its content into a set of pre-defined assignments such as for example, sentiment (Hastie et al., 2009). Previous studies have shown the potential machine learning has specifically for sentiment analysis (Zafarani, Abbasi, & Liu, 2014). This approach has also entered the world of sports research. For example, one study, previously mentioned, analyzed the sentiment of tweets with machine learning algorithms to create an automatic textual summarization of a soccer match (Jai-Andaloussi et al., 2014).

Scholars have also adopted such methods in analyzing soccer tweets posted during the FIFA 2018 World Cup (Fan et al., 2019). This study looked to determine how English fans behaved in response to how their team was performing and if they tended to associate themselves with a successful team, while also dissociating themselves with an unsuccessful team (Fan et al., 2019). This study showed the successful implementation of machine learning in determining the sentiment of English fan’s tweets and discovered that English fans tended to bask in the reflected glory of their team when they were successful by scoring goals, saving goals, or taking free-kicks (Fan et al., 2019). Additionally, this study found a lower team identification, lower national identification, and lower sentiment when the opposing team scored (Fan et al., 2019).

1.9 Aim

While previous research has examined the descriptive statistics of the impact of VAR on soccer, so far, no research has been published validating the system since its introduction or examining fans' reaction to its implementation (Lago-Peñas et al., 2019). Therefore, this project aims to research how sports fans on Twitter react during soccer games in response to VAR used in the English Premier League. Soccer is one of the many sports that experiences tensions between the fans of the game and the governing authorities of the sport that are trying to impose new structures and guidelines (Rodriguez, 2016). While FIFA and the International Football Association Board have approved the introduction of VAR technologies into soccer, there has been limited evaluation of this system on the game itself (Premier League, 2018b).

One way to study how fans feel about the introduction of this system is to use Twitter analysis allowing for input from many different fans from all around the world. This work focuses on the English Premier League. This league is one of the most prestigious leagues in the world and draws the highest global television audience of any soccer league with broadcasts in 212 territories reaching 643 million homes with a potential TV audience of 4.7 billion people (Dubber & Worne, 2015; Ebner, 2013; Premier League, 2018a). Additionally, an advantage of this league is that all teams have a presence on Twitter with an official account that attracts millions of subscribers. For example, the official Manchester United account has 21.7 million followers, Arsenal has 15.7 million followers and the official Premier League account has 22 million followers (Arsenal, 2020; Manchester United, 2020; Premier League, 2020). Since Twitter is not a Premier League specific forum, it can be expected that the collected data set features contributions from a variety of different sources including fans of the teams involved in each game as well as more casual viewers and general sports fans. Generally, due to the nature of sports fans on Twitter, any type of text analysis is not without imperfections. In this research, two different methods of event detection and sentiment analysis are introduced: both a frequency and lexicon-based analysis of tweets as well as a machine learning approach. This experiment attempts to analyze how Twitter users react during soccer games when the VAR is involved in the Premier League through sentiment analysis to examine the emotions demonstrated in the published tweets.

2. Methods

2.1 Sample

The data set for this study consisted of most matches in the 2019/2020 Premier League season in the period from November 23, 2019 (Matchweek 13) to March 9, 2020 (Matchweek 29). The start date marks the moment when the code was developed allowing for quick processing of the subsequent games. The end date was determined by the governmental authorities because, after this, further games were delayed due to the COVID-19 outbreak (Brown, 2020). Unfortunately, not all matches in this timespan could be considered as some matches were lost due to technical difficulties.

Table 1. Descriptive Statistics of Sample

Total Sample					
	N	**Games with Official VAR Incident**		**Games without Official VAR Incident**	
		Count	%	Count	%
Number of Games	129	85	66	44	34
Information per Game					
	Total	(per Game)	(per Game)	**Range** (per Game)	
	N	**Mean**	**SD**	**Min**	**Max**
Number of Tweets	643,251	4,986.4	6,186.4	284	40,127

This study collected Twitter data from 130 matches. One game had to be excluded due to the poor performance of the algorithm for this game, further details will be discussed later due to readability. Therefore, 129 matches were analyzed for this report. Not all matches had a particular incidence that required the assistance of the VAR, although 66% of them had at least one occurrence (n=85). From all collected games, a total of 643,251 tweets were analyzed for an average of 4,986 tweets per game. However, the range of tweets per game is very broad, with the smallest number of tweets collected for a single game being for the Watford vs Burnley match with 284, while the largest number of tweets collected was for the Chelsea vs Manchester United match with 40,127. The descriptive data of the collected sample can be found in Table 1. The tweets in question were entirely or mostly in English.

2.2 Data Access: Twitter API

To obtain this desired data from each game individually, multiple steps had to be taken. A Twitter application programming interface, or API, is a platform that allows users to search the web for specified tweets. This is another important feature of Twitter as it means third-party developers as well as academia and business-led market researchers can access valuable data through the platform (Makice, 2009). For this project, the `RTweet` Package in R was used as the API (Cohen & Kühne, 2019; Kearney, 2019). This package allows for a streamlined approach to interacting with Twitter's API and assists in converting the returned information into tabular data structures (Kearney, 2019). Through this interface, tweets can be extracted either in real-time or, as for this project, after the game (Appendix A).

To be used properly the API required the correct input so the desired tweets could be extracted. This requires descriptive information from each match including the official time of the kickoff, the amount of stoppage time for each half, and the official hashtag. The approach used in this study focuses solely on tweets containing the single official hashtag as published on the official Premier League game page. Therefore, it must be noted that not all messages related to the games in question were collected as some users who are tweeting about and commenting on the game do not use the specific hashtag in question. The official hashtags are made up of the team's three-letter codes with the home team code first and then the away team. For example, when Chelsea (CHE) plays against Arsenal (ARS) the hashtag would be #CHEARS. All game data regarding timing was collected from the Premier League website itself (Premier League, 2019-2020). Additionally, to determine which matches had VAR instances, the term 'VAR' was used to search the live text comments of each game on the Premier League Website. (Appendix L).

When given the necessary information, the API extracted all these tweets in the specified time interval that were tagged with the official game hashtag and combined them into a database. The time interval was set to the official starting time of the match as posted on the Premier League website and then a time span of two hours. Additional information was collected to match the tweets to the game minute. This was done by using the 45 minutes of playing time of each half plus the additional stoppage time before both halftime and fulltime in addition to accounting for the 15 minute halftime break. By calculating this time in comparison to the official starting time, the game minute for each tweet was recorded. For this, tweets posted in the second when the game started were considered to be posted in minute 0, otherwise, the minute groups were counted up. For example, a tweet posted in the first 30 seconds is considered to be posted within the first minute and is therefore in the one-minute group. By doing it in this way, the lag between different TV broadcasters and streaming speeds are also partially accounted for. Additional manipulation of the API data also resulted in each tweet having its own unique ID allowing this tweet to be tracked through various data frames as the tweet is analyzed. Additional data pre-processing steps done directly on the tweets extracted by the Twitter API included removing URLs, mentions, the official game hashtag that every tweet has, and converting all the text to lowercase.

Table 2. Sample exported tweets using the Twitter API from the Arsenal vs Brighton and Hove Albion match with the official hashtag #ARSBHA

New ID	Text	Created at	Location	Minute Since Kick Off
6427	coyg!!! we winning this. 💪🏽 (Szeth-son-son-Vallano, 2019)	2019-12-05 20:18:19	Akoka, Lagos	4
11300	why aren't var picking up these fouls in the box against lacazette? (Andrade, 2019)	2019-12-05 20:26:50	Bailiwick of Jersey	12
10918	are arsenal really that bad now that a team like brighton can come to the emirates and dominate them all over the park… (Sam, 2019)	2019-12-05 20:45:34		31
1210	var is killing the passion in football (Dan, 2019)	2019-12-05 21:38:11	Clock End, Emirates Stadium	84
6697	i have said it countless times, arsenal problem is not the coach, it's the players, they lack the zeal to win games. (Adeleke, 2019)	2019-12-05 21:53:15		99
6682	this club is a joke (Papoose, 2019)	2019-12-05 21:53:20	Nigeria	99

Table 2 illustrates some examples of tweets containing the hashtag #ARSBHA that are obtained from the Twitter API from the match Arsenal vs Brighton and Hove Albion. This table shows example data that can be extracted using the above-explained method. This includes an ID to identify corresponding tweets, the text of the tweet that can be further analyzed, the time of the tweet which is used to determine in what minute since kick off the tweet was posted, and finally the location of the tweet was also collected but not analyzed within this project. This series of tweets, shown in Table 2, will be used in the following pages to demonstrate how the text is manipulated in steps to obtain the desired information. Other information also extracted from the used API, but was declared irrelevant for this study, includes, favorite count, retweet count, and the screen name of the user. It should be noted that when extracting the tweets, specific variables can be included or not, for example, retweets—referring to messages that are reposted from another user—can be included or not, as well as users mentioning other users. Both of these variables were excluded via the API while extracting the tweets.

To further analyze the VAR count and mean sentiment per minute of the match various text mining methods were used to interpret some of the big data collected in a way that would allow for an understanding of the results (Silge & Robinson, 2017).

Text mining refers to the practice of converting text into a structured data set which can then be used for analysis (Silge & Robinson, 2017). In this project, text mining was used on the tweets to determine those related to VAR and to be able to assign a sentiment to each tweet. While previously this was done with human coders individually correctly coding each tweet, in the current realm of social media such methods are greatly limited due to the large amount of tweets that can be produced during just a single game (Yu & Wang, 2015). Therefore, for this project, two commonly used methods for big data were considered including the lexicon-based approach and machine learning.

2.3 Lexicon-Based Approach

Typical lexicon-based analysis works by splitting up different words in a text and examining each one individually (B. Liu, 2012). The overall sentiment of a sentence, or in this case a tweet, is then determined by averaging the scores of all the labeled sentiment words in that tweet. (Appendix B).

2.3.1 General Data Pre-Processing for Lexicon-Based Approach

The first step after extracting the tweets was to prepare the data. Given that Twitter messages are known to contain useless or non-usable information, various pre-processing attempts exist to clean up some of these messages. As mentioned, some pre-processing occurred as the tweets were extracted with the API, so those applied changes still remain. Next, each tweet was split up into individual words. In this step, emojis were removed. The stop words, referring to common words that most software has been programmed to ignore, are then also filtered out (Cohen & Kühne, 2019). For this project, additional stop words were removed such as general words related to the premier league such as *premier*, *league*, *pl*, *englishpremierleague*, *epl*, *football*, *club*, etc. as well as those words connected to the various teams such as team names, nicknames, mascots, and common phrases found in tweets. For example, the stop words for the Liverpool team included *liverpool*, *liv*, *lfc*, *reds*, *liverpoolfc*, *weareliverpool*, etc. Another data preparation step is referred to as stemming. Here, algorithms strip common suffixes of the words presented in the text (Jurka). This means that, for example, words like *winning*, *killing*, and *picking*, as seen in the tweets in Table 2 become *win*, *kill*, and *pick*. The R `dplyr` package was also throughout this research to prepare data (Wickham, François, Henry, & Müller, 2020). Returning to the previously mentioned example, Table 3 shows four of the tweets with stop words taken out and broken down into individual words. Once in this form, further analysis can be done.

Table 3. Sample text-mined data using the lexicon-based approach extracted from some of the tweets shown in Table 2

New ID	Word	Minute Since Kick Off
6427	win	3.32
11300	aren't	11.83
11300	var	11.83
11300	pick	11.83
11300	foul	11.83
11300	box	11.83
11300	lacazett	11.83
1210	var	83.18
1210	kill	83.18
1210	passion	83.18
6682	joke	98.33

2.3.2 Event Detection of VAR Incidents

Part of the text mining also includes event detection. Here, a combination of frequency and content help identify VAR incidents. For this study, to get the VAR count, the number of times *var* appears in the mentioned data frame during that minute is summed up. This means that when there is no tweet mentioning the VAR the value is 0. Depending on the game and the number of tweets collected, the peaks in VAR count help identify VAR instances or situations in which fans wanted VAR to be involved.

2.3.3 Sentiment Analysis

Sentiment analysis is considered a type of text mining that attempts to determine the subjectivity and the opinion of a text (Liske, 2018). Unlike the more factual information of if a tweet contains references to the VAR or not, opinions and sentiments are more difficult to characterize as they are highly subjective (B. Liu, 2012). Initially in this project, emotions were measured at a word level; that is, words in a pre-defined library were considered and quantified. Here, to extract sentiment from the recorded tweets, R's `tidytext` library package was used with the `AFINN` lexicon (DeQueiroz et al., 2020; Nielsen, 2017). This lexicon provides a sentiment score by assigning a value from a scale of -5 to -1 and +1 to +5 (0 is not included) to the 2,477 different words in this lexicon (Nielsen, 2011). Here, negative numbers refer to words with negative sentiment and positive numbers refer to words with positive sentiment. The package only works by assigning a sentiment rating to words that already exist in the library, additional words in the tweet that are not in the library are ignored. To get a feel for this library used here, from the words obtained, *superb* has a value of 5, *win* and *brilliant* are ranked as 4, *perfect* and *impress* are 3, *top* and *solid* are 2, *yes* and *wish* are 1, *stop* and *anti* are -1, *ruin* and *hopeless* are -2, *worst* and *lost* are -3, *fraud* and *piss* are -4, and *b**ch* is -5. Table 4 shows the sentiment assignment of the sample extracted tweets in Table 2.

Table 4. Sample sentiment assignment data extracted from the tweets shown in Table 2.

New ID	Word	Minute Since Kick Off	Value
6427	win	3.32	+4
10918	bad	30.57	-3
1210	kill	83.18	-3
6697	lack	98.25	-2
6697	win	98.25	+4
6682	joke	98.33	+2

These individual word sentiments can then be used to show the final ratings of each tweet (Table 5).

Table 5. Sample tweets with their respective ratings for both sentiment and VAR using the lexicon-based method.

New ID	Text	Sentiment	VAR
6427	coyg!!! we winning this. 💪 (Szeth-son-son-Vallano, 2019)	+4	0
11300	why aren't var picking up these fouls in the box against lacazette? (Andrade, 2019)	n.a. (0)	1
10918	are arsenal really that bad now that a team like brighton can come to the emirates and dominate them all over the park... (Sam, 2019)	-3	0
1210	var is killing the passion in football (Dan, 2019)	-3	1
6697	i have said it countless times, arsenal problem is not the coach, it's the players, they lack the zeal to win games. (Adeleke, 2019)	+1	0
6682	this club is a joke (Papoose, 2019)	+2	0

2.3.4 Finalized Data for Export

Once all the words have their assigned sentiment, the mean sentiment per minute of the game is calculated and saved along with the VAR count per that minute of the game. If no tweets were posted in a specific minute since kick off no sentiment could be counted and therefore the sentiment for these minutes was filled in as a 0. Table 6 shows a sample of the final data that is obtained for each game and that can be exported.

Table 6. VAR count and respective mean sentiment shown for various game minutes of the full data set of the Arsenal vs Brighton match as calculated via the lexicon-based method.

Minute Since Kick Off	VAR Count	Minute Since Kick Off.1	Mean Sent	Game Minute
4	0	4	1.06	4
12	17	12	-0.18	12
31	13	31	-0.57	31
84	27	84	-1.29	69
99	0	99	-0.81	84

In this table, note the differences between the time columns (Table 6). The *Game Minute* column accounts for stoppage and halftime, while the *Minute Since Kickoff* column does not. The different sentiments as well as the VAR count from the examples above are incorporated here.

Finally, a graph made up of the mean sentiment and the VAR count in the tweets per minute for every minute of the game is created (Figure 1).

2.3.5 Limitations of the Lexicon-Based Method

While this method shows potential in obtaining fan opinion about what happens during a soccer match, a critical look at the steps used to obtain the desired data reveals the constraints of this approach. When comparing the original tweets, shown in Table 2, to the sentiment analysis data shown in Table 4 obvious inadequacies become evident. In the end, only limited words are being considered in the sentiment analysis, due to the library that was being used, instead of the whole tweet, which leads to some inaccurate results. For example, for the tweet saying “why aren’t var picking up these fouls in the box against Lacazette?”, the sentiment analysis claims this to be a neutral tweet even though in whole there is an exclaimed negative sentiment towards VAR (Andrade, 2019). Not only might the current method claim more tweets as being neutral than there are, sometimes the sentiment is quantified as the opposite of what it actually is. Take the two final tweet examples, “I have said it countless times, arsenal problem['s] is not the coach, it’s the players, they lack the zeal to win games” and “this club is a joke” (Table 2) (Adeleke, 2019; Papoose, 2019). In context, both could be described to be portraying negative sentiment. However, looking at the sentiment rating of these tweets both have positive values, the first one a value of +1 taken as the average value of the word *lack* with -2 and of the word *win* with +4, and the second tweet with the value of +2 corresponding to the word *joke*. Additionally, when examining the words individually in the library a word like *shoot* has been assigned a value of -1, which is often inaccurate when used in the context of soccer. For example, with the current library, the tweet referring to a player who is a “baller as he certainly can cross, pass, and shoot” would be considered to be expressing a negative sentiment with a value of -1 (l. smith, 2020). Previous studies have mentioned the same limitation by stating that sentiment analysis is dependent on the domain to which it is being applied because a word can convey different sentiments and meanings depending on its context (Zhaoxia Wang, Tong, Ruan, & Li, 2016). Another rather large drawback of the sentiment analysis used here is that emojis were not considered.

Unfortunately, these limitations are not restricted to the sentiment analysis alone. Some tweets are clearly related to VAR in a game without explicitly using the word *VAR*. These are ignored, as here, only the number of times the phrase *VAR* appears is counted. Consider this tweet during the Arsenal vs Brighton game: "[why is this taking so long.] David Luiz was yards offsides. assistant referee was asleep" (Okelana, 2019). Here there is no direct mention of VAR, however, in the context of this game is it clear that this fan is complaining that the VAR review for the goal is taking too long. Also, when a tweet uses VAR multiple times it is counted as multiple occurrences of VAR tweets. For example this tweet from the same game, "var comes into life to disallow a goal!! where was var for our penalties in the first half???" is counted twice in reference to VAR (Stokesie, 2019). Additionally, this method does not account for variations of VAR that could be used in tweets. For example, the tweet: "I'm failing to comprehend the usefulness of v.a.r" is, with this method, not counted (HK, 2020).

Both of these rather large drawbacks suggest that further considerations should be made about the sentiment ratings and VAR count methods used here.

2.4 Machine Learning Method

To overcome some of the mentioned limitations and improve the results of this study, machine learning was also implemented to create an algorithm that will more accurately identify the sentiment and also pick out tweets related to VAR even when it is not explicitly stated. Previous studies have demonstrated the potential this method has in examining soccer games (Fan et al., 2019).

Here, the same export from the Twitter API was used as described previously. Therefore, continuing with the previous example, here the data started in the same form as seen in Table 2.

2.4.1 Preparing Training Data for Supervised Learning

The first step for the machine learning method is to create a set of training data. Here, two researchers characterized a specified set of tweets based on sentiment and relevance to VAR in response to different game situations. For this, 4,583 tweets were collected and labeled from various different game situations. These situations were determined based on the various ways a VAR could be involved in a soccer match. This included the four general categories: goal, penalty, red card, and mistaken identity. For a goal, the VAR could help in confirming it or overturning it based on a possible handball, foul, or offsides call. Additionally, a penalty kick can be confirmed, awarded, or canceled with the input from the VAR. Finally, a red card could also be awarded, overturned, and can be related to a mistaken identity case. None of the examined games, however, had an incident of mistaken identity, so that case was excluded. All the other possibilities resulted in project researchers labeling

23 different games to be used for the training data. About 200 tweets, were labeled from each of these 23 games. For these subsets, a random sample of 100 tweets was collected from the described VAR incident and another random sample of 100 tweets was taken from the rest of the game. The boundaries of the VAR incident was set depending on the information obtained from the live-ticker of the respective game and confirmed by the VAR count tweet peaks identified with the previous lexicon-based method (Premier League, 2019-2020). For this, depending on the number of tweets collected for the game in question, either just the peak minute of the incident or the peak range of the incident was used for the VAR training data. For games with limited tweets related to specific incidents, less than 100 tweets were labeled.

For these tweets, two different categories were labeled: VAR and sentiment. For the first category, the tweets related to the VAR were labeled as a 1, and those not related were labeled as a 0. The sentiment was coded by using 1 for positive, 0 for neutral, and -1 for negative. In determining the labeling, researchers used a variety of context cues to help decipher the tweets including the time stamp, the other tweets written by the same author, and other contextual information that could help. Such contextual clues are particularly helpful as this study also considered sarcasm when coding the tweets. It should be noted that emojis were also examined to help determine the sentiment.

Along with basic intuition of if a tweet is positive or negative, some rules were established to help researchers label the tweets accurately. First of all, it was determined that if the researchers could not make sense of the tweet due to spelling errors or general gibberish, the tweet would be labeled as neutral. Any sort of "go boys," "go team," or cheering was considered positive. Hints at conspiracy theories, for example, VAR favoring one team over another, were considered negative. When a tweet referred to a team member or team doing something bad, this was considered negative and vice versa. For example, the tweet: "keeper should have saved that anyway" was labeled as negative (Shepherd, 2019). Additionally, since the tweet was considered as a whole, if there was a positive part and a negative part, this would result in a general neutral rating. Finally, if the text of a tweet could go both ways but no other contextual information gives a hint as to what the intended sentiment is, the tweet will be labeled as neutral. For example, the tweet "are. you. kidding. me?" could be interpreted as positive or negative and would require additional context such as the team this fan supports, or what was happening in the game when this was posted, to determine which one it is meant to be (Walker, 2020).

Table 7. Sample tweets from the labeled training data with the appropriate categorization for VAR and sentiment

Tweet	VAR	Sentiment	Explanation
Never a foul...goalkeepers get protected way too much (P. Adams, 2020)	1	-1	Author referring to a VAR overturned goal due to a foul and disagrees
Why do they give big games to such clowns to referee? (Richard, 2020)	1	-1	Contextual information revealed this was posted during a VAR incident
var and man city a love story that one (Ebénn, 2019)	1	-1	Sarcasm, here Man City goal is overturned via VAR due to offside call
See in that instance var has worked, no chance you can argue that's an intentional handball (Guest, 2020)	1	1	Author is agreeing with the VAR decision made here
That's purely onside (Roksie, 2020)	1	1	During a VAR decision looking at a possible offside, here VAR declares it onside and author agrees
Goal or offside? (uhuru jr, 2020)	1	0	Posted during a VAR incident while waiting for the decision to be made
var completes the check and no hand ball (Indy Football, 2020)	1	0	Objective tweet about a VAR decision
I like Salah, so I won't say anything (Singhal, 2020)	0	-1	User implies something negative happened but is a fan of Salah so won't explicitly say anything bad about him
Feels like it should be a double digit goal scoring half (Nick, 2020)	0	-1	Suggesting score should be much higher if team would actually score their shots
Finally the team won..... which I don't support (nafees, 2019)	0	-1	Sarcasm here, author does not support winning team so not actually happy
Can [we] please win this. Lets go anything is possible. (Brahmbhatt, 2020)	0	1	Author is cheering for their team
Chelsea tried though (Èkó, 2019)	0	1	Even though Chelsea lost, positive response
What an incredible match, easily match of the season so far (sully!, 2020)	0	1	All around positive sentiment
Today had 53% possession. Proud of this young team. Unlucky to of lost today (Khalid, 2019)	0	1	Author is proud therefore positive sentiment
Hold on tight (Ike, 2020)	0	0	Could be interpreted either way so considered neutral following the rules
Man City and Chelsea will have a great match. And Martin Atkinson will make it frustrating, so I'm excited to see how this plays out. (RecWand, 2019)	0	0	Two different parts have two different ratings and no one part outweighs the other making it neutral

Table 7 shows samples of some of the labeled tweets that were included in the training data and their respective categorization. Here, it should be mentioned that in some cases, a tweet appears to refer to a posted image or a graphic interchange format (GIF) as are commonly added in tweets, however, the tweet extraction method used here did not include the image URL. Although this was not that common, in cases where it did occur it sometimes made sentiment assignment rather difficult. For example, the tweet "me every time Pablo Mari made a pass to Pépé" appears to refer to an accompanying image. In this case, the sentiment cannot be correctly labeled without knowing the contents of the image (E.G, 2020).

For the training data, the inter-rater-reliability for labeling the tweets showed quality scores. Calculated values for Cohen's Kappa reached an excellent value of 0.99 for determining if the tweet was related to a VAR incident and a good value of 0.88 for determining the correct sentiment of the tweet.

As done in other studies, another part of the training data included creating a customized stop words dictionary (Aloufi & Saddik, 2018). The stop words dictionary used here included all names specifically found in the training data tweets with the exact spellings and nicknames found in these tweets. Names here included the team names as well as nicknames and mascots, player names, league officials or commentators, stadium names, and others commonly referred to during soccer events. (Appendix H).

It should also be noted that emojis in the tweets were converted into a text version so that they could also be included as variables in the machine learning algorithms. For this they were converted from the emoji symbol into their 8-bit Unicode Transformation Format (UTF8) which turns each emoji into an 8-bit code, for example, the broken heart emoji becomes `'<f0><9f><92><94>'`. From this a dictionary was used to translate this set of numbers and letters into a text version that could be used by the machine learning algorithm, for this example, the final text included in the training data is then `'*broken_heart*'`. It should be noted that this method also accounts for the different skin tones of the emojis, for example, different versions of the `'*thumbs_up*'` emoji also include variations from `'*thumbs_up_light_skin_tone*'` to `'*thumbs_up_dark_skin_tone*'`. (Appendix G and I).

2.4.2 Three Different Text Classification Learning Algorithms

For machine learning to function, it involves creating a model that uses the pre-labeled training data, from which it can then sort new data and make predictions accordingly. Various different types of such models have already been used and researched for machine learning systems. For this text classification problem, the three different existing supervised learning approaches commonly used include the naïve Bayes classification, support vector machines, and random forests (Aloufi & Saddik, 2018; Hastie et al., 2009; B. Liu, 2012).

All scripts are generally structured the same way and firstly include a data preparation part. Here, the customized stop words dictionary is removed, and the data is cleaned. The data input then occurs in the form of a document-term matrix, a common approach in text classification (Feinerer, 2019). In this matrix, the rows are made up of the tweets which are referred to as the documents and the columns are made up of the terms or words in the document (Feinerer, 2019). A normal document-term matrix is made up of 1s and 0s. When the word in the column header occurs in the tweet the matrix cell is filled in with a 1 and if not, then a 0. The completed training data with all the labeled tweets was converted into such a document-term matrix. This matrix is then the input for the learning algorithm. Once the data is inputted, the next step is to run the learning algorithm. Finally, once the model has learned, the last part is to apply this to the test set. From this, the system is able to get a measure for how good the model is in the form of a confusion matrix where it compares the predicted outcomes to the actual outcomes (Table 8). With these values, depending on the goals of the learning algorithms, the best performing one for the intended purpose can be selected.

Table 8. Confusion Matrix

	Positive Label	**Negative Label**
Positive Prediction	True Positives	False Positive
Negative Prediction	False Negatives	True Negative

2.4.3 Naïve Bayes Classifier

The naïve Bayes classifier is a probability-based classifier, that calculates the probability that a document belongs to a specific class (Aloufi & Saddik, 2018). This classifier can be exemplified with the question of whether to play golf today or not. This question can take particular variables into account such as temperature and wind (Gandhi, 2018). From the training data, depending on the probabilities that golf was previously played when mild temperatures occurred and wind was in the forecast, the algorithm proportionally assigns either a *yes* or *no* to the *play* variable (Gandhi, 2018). This classifier is referred to as naïve because it holds the assumption that each feature variable in a document is conditionally independent from the other features meaning that the presence of one feature does not affect the other (Aloufi & Saddik, 2018). This classifier also assumes that all the predictor variables have an equal effect on the outcome (Gandhi, 2018).

For the coding of this classifier, the data preparation occurs with the help of the `quanteda` package and data cleaning happens with the document-term matrix settings (Benoit et al., 2020). The model here is trained with the help of the `quanteda.textmodel` package (Benoit et al., 2020). Further details can be found in Appendix C.

2.4.4 Support Vector Machine

The support vector machine (SVM) is a classifier method that attempts to classify the extremes of a database and selects an optimized hyperplane that "maximizes the margin between the closest instances of two classes" (Aloufi & Saddik, 2018). To visualize this consider trying to distinguish between dogs and cats. This algorithm works by looking for a cat that most closely resembles a dog and a dog that most closely resembles a cat. Based on these extremes, the algorithm creates a linear separation, called a hyperplane, from which it then predicts whether a picture is of a cat or a dog (Markowetz, 2004). This separating hyperplane uses the maximal margin, meaning the edges of the hyperplane pushes up against the extremities of the data points (Markowetz, 2004). This classifier is often used for text classification and has previously shown strong performances in this field (Z. Wang, Sun, Zhang, & Li, 2006).

To code this classifier, the data preparation was done with the help of the `tm` text mining package (Feinerer, 2019). The model here is trained using the `caret` package (Kuhn, 2020). Further details can be found in Appendix D.

2.4.5 Random Forest

The random forest (RF) classifier is an ensemble classifier that is made up of multiple full-grown decision trees (Aloufi & Saddik, 2018). A decision tree is a commonly used tree-like structure that visually represents decisions with each internal node representing a particular test of an attribute possibly by asking a question, each branch then represents possible outcomes of this test, and each end node or leaf represents some sort of decision that is made after computing through all the attributes (Brid, 2018). Each tree in this forest is built from a bootstrapped version of the training data, meaning a random subset of variables is considered at each step (Yiu, 2019). From this data, for each tree, two random options are selected as candidates for each node, it is determined which one of these two does the best job separating the samples, then that variable is used for that node and the process continues until a full decision tree is made. However, the number of variables considered in each step can change based on the learning process. This cycle is then repeated creating a variety of trees in a random forest ultimately resulting in lower correlation and more diversification (Yiu, 2019). At the end, when new data is inputted, this algorithm runs the data down these trees and keeps track of the predictions made. In the end, the algorithm predicts based on the majority of votes consisting of all the individual trees (Yiu, 2019). For the data preparation of this classifier, once again the `tm` text mining package is used (Feinerer, 2019). The rest of the algorithm is conducting with the help of the `randomForest` package (Breiman, Cutler, Liaw, & Wiener, 2018) (Appendix E).

The most common methods then applied specifically to improve the accuracy of the learning algorithms for random forests is called boosting. This method is primarily used to reduce bias and variance that occurs in supervised learning of random forests (Gupta, 2017). Here, each decision tree is reduced to a stump, meaning a single level decision tree tries to classify the data points (Gupta, 2017). All these stumps are considered weak learners as they can only take one feature into account and are considered to be only slightly better than chance (Sharma, 2019). These are all then combined into a strong learner, which is able to then achieve excellent performance. Boosting specifically pays attention to misclassified data points in order to ensure that higher importance is given to these points until they too are correctly interpreted (Gupta, 2017). In the end, all the weak learner predictions are combined and their assigned weights are accounted for to achieve a single final weighted prediction (Sharma, 2019).

One of the computing packages for boosting used in this project is `xgboost`, whose main aim is to increase the speed, efficiency, and achieve a higher accuracy of the model (T. Chen & Guestrin, 2016). With boosting packages additional tuning specifically, meant for random forests can be done. Tuning refers to a way to optimize parameters for a machine learning algorithm and is explained in detail below in the tuning section (Malik, 2018). In the `xgboost` package, this includes a variety of different variables that can all affect the performance of the learning algorithm. The `eta` variable is the measure of the step size of each boosting step and acts to control the learning rate to prevent overfitting by making the boosting process more conservative (T. Chen et al., 2020). In short, a decision tree is overfitted if it is able to produce a highly accurate output on the included training data, but low accurate output on additional new test data (Sharma, 2019). The `max_depth` variable is simply the maximum depth of a tree. The `min_child_weight` variable refers to the minimum sum of weights of all observations required in a child. Higher values here will prevent the model from learning a series of relations that could be highly specified to the particular sample that has been selected for that tree. If the tree-building step results in a leaf node for which the sum of the weights is less than `min_child_weight`, the building process will prevent further partitioning (T. Chen et al., 2020). The `subsample` variable is a subsample ratio of the training data (T. Chen et al., 2020). When this value is 0.5, for example, it means that `xgboost` randomly collects half of the data to grow trees, which therefore prevents overfitting (T. Chen et al., 2020). Finally, the variable `colsample_bytree`, represents the subsample ratio of features randomly selected that are used when constructing each tree in the forest (T. Chen et al., 2020). This determines which features or columns of the training set will be used to build each tree.

The second boosting package used is the `gbm` package, which stands for a generalized boosted regression model and also involves various variables that can be tuned (Greenwell, Boehmke, Cunningham, & Developers, 2019). In this package, the variable `interaction.depth` specifies the maximum depth of each tree which refers to the number of edges from the node to the tree's root node (Greenwell et al., 2019). The `shrinkage` parameter in this package is a learning rate measure or step-size reduction applied to each new tree in an expansion that determines how much better a new tree must be when compared to an old tree (Greenwell et al., 2019). The `m.minobsinnode` variable is an integer that specifies the minimum number of observations in the terminal nodes of a tree (Greenwell et al., 2019). Here, a higher value means smaller trees and a value of one means the algorithm continues until there is only one observation in each terminal node. The variable `bag.fraction` refers to the fraction of observations that are to be randomly selected to form the next tree, thereby including randomness into the tree-building process (Greenwell et al., 2019). `Train.fraction` refers to the fraction of the training set observations that will be used to fit the `gbm` model (Greenwell et al., 2019). Finally, the `cv.folds` variable refers to the number of cross-validation folds the model will perform (Greenwell et al., 2019). Cross-validation is a technique used to estimate the performance of a specific model and works by randomly dividing the original sample into the specified n-fold equal subsets (Fan et al., 2019). Then, in a repetitive process, one of these subsets are extracted for testing while the remaining are used as the training data until all subsets have been used for testing. In the end, the average score of all the cross-validation results is used to estimate the performance of the specific algorithm (Fan et al., 2019).

2.4.6 Generalized Tuning

Tuning is the crucial process that one goes through in an attempt to optimize the parameters that can impact a machine learning model to improve its accuracy and result in a better performance (Malik, 2018). The learning algorithms depend on the parameters set before, and there are many different ways these can be changed that can tune the results of the algorithm for the better making this an exhaustive process.

One way often used in text classification is to use what is called Term Frequency-Inverse Document Frequency (TF-IDF). This feature works by reducing the weight of more frequent words that appear in a given text, as such extremely common words are considered to have a non-relevant meaning, and is therefore often used to improve a learning algorithm (Aloufi & Saddik, 2018; Fan et al., 2019). When implemented in this project, however, no improvements were noted.

Another attempt to improve the learning algorithm included conducting the experiment with a binary classification as was also attempted in other studies (Aloufi & Saddik, 2018). For this, the neutral and positive classes were combined due to these two classifications causing the most trouble throughout the algorithms. This would still show a change of negative sentiment. However, once again, in this project, no significant improvements in the algorithm were noted with a binary classification of sentiment.

Other studies have also tried to improve their algorithm by converting the document-term matrix into a probability model, where instead of 0 and 1s indicating if the term is present in the tweet document or not, the matrix is filled with probabilities of that term being in a given document (B. Liu, 2012). This unfortunately also did not improve the results.

Additionally, this project attempted to remove less frequent words found in the training data text however, here once again the results just continued to get worse. Another attempt included splitting the training data and only using that of one of the two researchers involved but once again no improvements were found. Other simple methods often used involve including or removing punctuation and digits and also stemming words, as described above in the lexicon-based method, however only minor changes occurred here. The only thing that improved the results was excluding the customized stop words dictionary created for the training data.

In an attempt to tune the algorithms for this project, the computer spent days running through different models to find the best method.

2.4.7 Accuracy Results of Machine Learning Method

To determine the success of a machine learning algorithm, mainly four distinguishable measures are used: *accuracy*, *precision*, *recall*, and *F1 score* (Joshi, 2016). These four measurements come from the so-called confusion matrix, shown in Table 8, indicating the four different outputs that are possible from an algorithm including also type 1 (false negatives) and type 2 (false positive) errors. The *accuracy* measure describes how good the predictions of the algorithm are and is simply a ratio of the correctly predicted observations to the total number of observations (Joshi, 2016). Here, a *balanced accuracy* measurement accounts for an unequal number of trials in the different categories by calculating the average of the proportion of correct predictions in each class individually (Brodersen, Ong, Stephan, & Buhmann, 2010; Etzel, 2015). *Precision* measures how good the model is at predicting positive outcomes and is, therefore, the ratio of correctly predicted positive observations, so true positives, to the total number of predicted positive observations (Joshi, 2016). The *recall* variable measures how sensible the model is at predicting positive outcomes and is measured as a ratio of the correctly predicted positive observations to all the observations in the positive prediction class (Joshi, 2016). Finally, the *F1 score* is a measure of the weighted average of both the *precision* and *recall* measurement (Joshi, 2016). This score takes both the false positive and false negatives into account (Joshi, 2016).

Table 9. Validation Results of Different Machine Learning Algorithms

Methods	Variables	VAR	Sentiment		
			Negative	Neutral	Positive
Naïve Bayes	Overall Accuracy	0.81		0.67	
	Balanced Accuracy	0.81	0.77	0.81	0.70
	Precision	0.74	0.87	0.58	0.48
	Recall	0.97	0.65	0.81	0.58
	F1 Score	0.84	0.75	0.68	0.52
Support Vector Machine	Overall Accuracy	0.93		0.66	
	Balanced Accuracy	0.93	0.75	0.76	0.70
	Precision	0.97	0.79	0.65	0.46
	Recall	0.93	0.67	0.69	0.57
	F1 Score	0.95	0.72	0.67	0.51
Random Forest – xgboost	Overall Accuracy	0.95		0.74	
	Balanced Accuracy	0.95	0.81	0.83	0.77
	Precision	0.98	0.84	0.69	0.63
	Recall	0.94	0.74	0.80	0.66
	F1 Score	0.96	0.79	0.74	0.65

In calculating the validation results of the different algorithms for the different categories, an interesting trend was revealed (Table 9). It was found that all validation results for the positive sentiment category were less than both the neutral and the negative sentiment categories.

After comparing the different algorithms, the boosted random forest approach gave the highest accuracy and best performance for classifying both the VAR relevance and the sentiment with machine learning. The final code for sentiment analysis that leads to the best overall accuracy score of 0.74 used the boosted random forest method optimized with `xgboost` (Appendix F). To get such a high score, along with the various optimized tuning parameters (Table 10), it was found that instead of building each tree from a large share of variables defined by the `colsample_bytree` parameter for which the value is by default usually close to 100%, the code for this project builds each tree based on a small fraction of variables, here 5%. For text classification algorithms, as was found here, this modification may be useful as document-term matrices for text usually have a huge amount of variables. Additional optimization included removing punctuation, numbers, and the customized stop words dictionary, as well as stemming words. The final code for VAR analysis that leads to the best overall accuracy score of 0.95 also used the `xgboosted` random forest method but had different optimized tuning parameters as shown in Table 10.

Table 10. Optimized tuning parameters for xgboosted machine learning algorithm

	VAR	**Sentiment**
`eta`	0.2	0.2
`max_depth`	15	30
`min_child_weight`	1	1
`subsample`	1	0.9
`colsample_bytree`	0.65	0.05
`nfold`	--	5
`objective`	'bonary:logistic'	'multi:softprob'
`num_class`	--	3
`silent`	1	1

2.5 Method Selection: Lexicon-Based vs Machine Learning

To examine how the two described methods compare in terms of finalized results, both methods were applied to the Arsenal vs Brighton game used as an example throughout this report.

2.5.1 Lexicon-Based Results of Arsenal vs Brighton Match

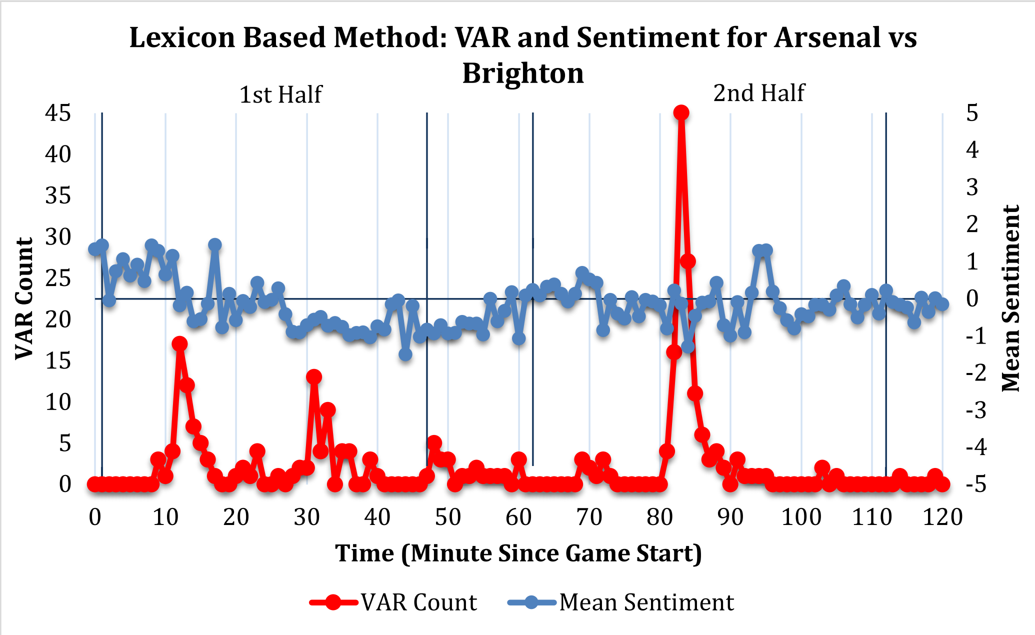


Figure 1. Exported data from the Arsenal vs Brighton and Hove Albion match using the lexicon-based method showing the dynamic VAR Twitter count and mean sentiment of the tweets over the duration of the game.

Starting with the lexicon-based method, shown in Figure 1 is a visual representation of the final data exported from the Arsenal vs Brighton game using this methodology. This figure shows the dynamic VAR Count as well as the mean sentiment through the match as split up by different game phases, 1st half, halftime, and 2nd half. Looking at the figure, the ongoing fluctuation of the mean sentiment throughout such a game becomes clear, however, further analysis is necessary to reveal trends. Comparing this figure data to the descriptive data of what is happening during a game, shows the potential of this methodology.

Generally, at the beginning of each match, the sentiment tends to be high with optimism and support being expressed for the teams, as also portrayed in this game here (Figure 1). The first VAR tweet peak in this game comes in the 12th game minute. Looking at the corresponding tweets, also shown in Table 2 as well as what happened during the game at this point, it becomes clear that during an offensive attack by Arsenal, fans believe Brighton players are fouling the Arsenal players and the situation should be reviewed by VAR. When comparing the sentiment of the tweets directly before this VAR peak to during the peak, the sentiment drops from 1.43 to -

0.18. The next increase in tweets directly citing VAR occurs in the 31st minute. Here once again Arsenal fans are complaining about fouls not being picked up in the box for potential penalty kicks. Here, during the VAR tweets that are a bit more spread out, the trend shows a more consistent negative sentiment when compared to the fluctuations occurring in the rest of the game. Finally, during the second half, the VAR peak here corresponds to an Arsenal goal overturned by VAR after it was determined that the Arsenal player was offsides. This single spike in VAR tweets is accompanied by a decrease in sentiment from 0.21 to -1.29.

2.5.2 Machine Learning Results of Arsenal vs Brighton Match

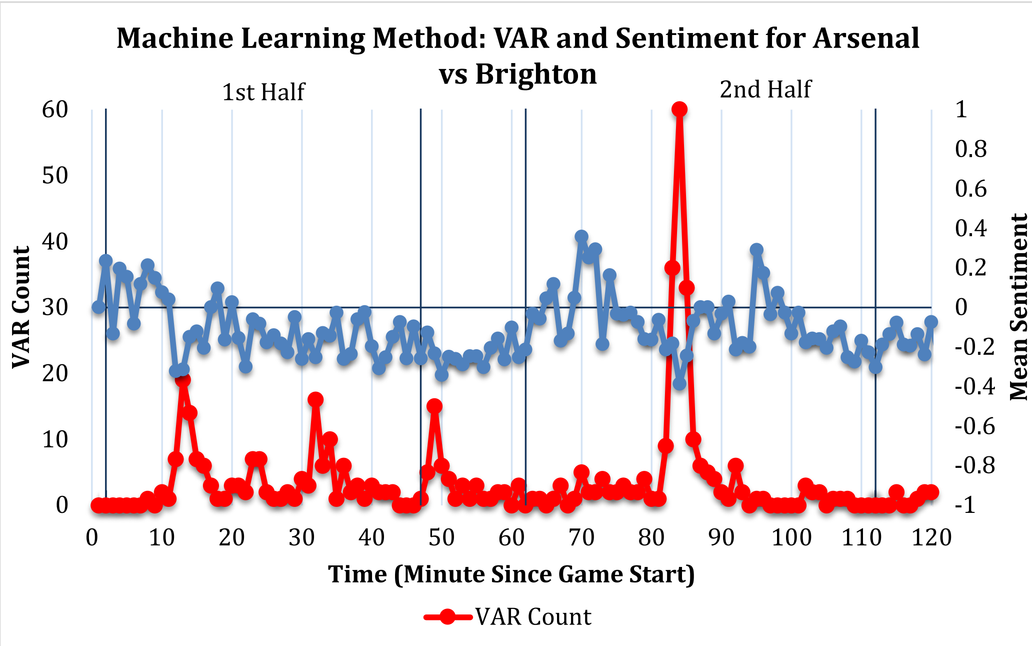


Figure 2. Exported data from the Arsenal vs Brighton and Hove Albion match using the machine learning approach showing the dynamic VAR Twitter count and mean sentiment of the tweets over the duration of the game.

When comparing the lexicon-based data to the machine learning results of the same Arsenal vs Brighton game, as shown in Figure 2, some interesting things can be seen. Starting with the VAR count throughout the game, as expected the general count of VAR tweets is higher. This is because the machine learning method is no longer just counting occurrences of the term *VAR*, and instead uses an algorithm to identify instances of VAR even when it is not explicitly stated. Additionally, the general trend of the VAR count line is similar in both methods. The same two instances previously mentioned around the 12th and 31st minutes are also visible here with another incident appearing shortly after halftime that did not appear as clearly in the lexicon-based method. When looking at the tweets it appears that after the halftime whistle is blown individuals are reflecting on instances they believed should have been reviewed by the VAR during the game but were not. Additionally, some of these tweets are referring to a yellow card warning given right before the halftime whistle was blown.

When examining the sentiment of this game with the machine learning method, the general trend appears different with more fluctuations which may be due to the different range of sentiment labels. Inspecting the specific sentiment changes during the VAR spikes shows that during the first VAR spike in the 12th minute of the game the sentiment goes from 0.15 to -0.32 from before the peak to the height of this spike, a similar trend was observed in the lexicon-based method. Moving on to the next VAR spike found in both methods during the 31st minute, here for the machine learning, looking solely at the peak value, the sentiment decreases from -0.16 to -0.25. Generally, at this point in the game, the sentiment stays mostly negative and, as mentioned, there are a lot of fluctuations whereas the sentiment calculated by the lexicon-based method was also generally negative but did not experience as many fluctuations. Finally, the major VAR peak that also corresponds to a VAR incident of a goal from Arsenal getting turned over in the second half during the 83rd minute since kickoff also is accompanied by a sentiment decrease from -0.06 before any VAR tweets were present to -0.38 during the height of the peak. A similar decrease in sentiment during this incident is also present in the lexicon-based method.

When examining other peaks present in the graph from the machine learning method, additional details become apparent that were not as prevalent in the lexicon-based graph. For example, here there are two peaks in sentiment in the second half one in the 70th minute since kickoff, shortly after halftime, and one in the 95th minute since kickoff. These two peaks correspond to goals scored in the second half. The first peak occurs at the same time as Arsenal scored an equalizing goal to make the game 1-1. The second peak in the second half occurred at the same time Brighton scored their second goal thereby restoring their leading against Arsenal 1-2 and was also present in the lexicon-based method. Interestingly, neither method resulted in any indication of the first goal scored by Brighton in the 35th minute.

2.5.3 Accuracy Results of Both Methods

While comparing the outcomes of the two methods above, some interesting things are revealed and to ensure the best of the two methods were chosen, quantitative analysis was also considered.

As previously discussed there are obvious limitations of the lexicon-based method which greatly negatively affected the accuracy of the obtained data and suggested using an additional method to obtain better results. In an attempt to quantify the accuracy of the lexicon-based method, two different lexicons were used to analyze the sentiment of a 4,583 tweet pre-labeled training set described above. These two include the `AFINN` and the `bing` lexicons (Hu & Liu, 2004; Nielsen, 2011). For this to work, only the tweets that include at least one word in the respective lexicon can be included in the analysis, this can be problematic considering how short tweets are. As a result, the training sets were reduced by a little bit more than half. For the `AFINN` dictionary, the training set was reduced to 2,388 tweets and for the `bing` dictionary to 2,491 tweets. This is a major argument to not use the lexicon approach. Additionally, for this lexicon-based approach, the accuracy levels were found to be at best 0.48 for sentiment analysis and 0.88 for VAR count, which were declared to be less than desirable and much worse than other methods involved (Table 11).

Comparing these results to the data obtained from the machine learning method, it becomes clear that for this project, the machine learning algorithm should be used. First of all, the machine learning approach addresses some of the previously mentioned limitations of the lexicon-based method therefore already improving the results in that way. Additionally, the accuracy results obtained from the optimized random forest machine learning algorithm were found to be 0.74 for sentiment analysis and 0.95 for VAR count (Table 11). While this is clearly better than the obtained results from the lexicon-based method, further research revealed that the machine learning algorithms used here were sufficient for the scope of this project.

When comparing the accuracy values of sentiment results of the algorithm used in this project to similar studies found in literature, our values appear to be in line with other research. One study also looking at soccer tweets for sentiment analysis in a method similar to the one presented in this project, found the best accuracy values for one game to be 0.75, and another game to 0.73 (Fan et al., 2019). Another study, comparing a variety of different algorithms for sentiment analysis in soccer tweets, looked at lexicons within a machine learning approach and found the average accuracy in this approach to be 0.43 for the `AFINN` lexicon and 0.54 for the `bing` lexicon (Aloufi & Saddik, 2018). These researchers found the average accuracy of the different machine learning algorithms, also used here, to be 0.59, but were able to improve it by changing to a binary classification with a random forest algorithm to an accuracy value of 0.74 (Aloufi & Saddik, 2018).

Table 11. Validation results of the different methods used in this project. Included here are the two different lexicons, as well as the best performing machine learning algorithm. Note for both the lexicon-based methods the VAR is purely a count of the number of tweets for which the word *var* appears, and therefore the values are the same.

Methods	Variables	VAR	Sentiment		
			Negative	Neutral	Positive
Lexicon-Based - `AFINN`	Overall Accuracy	0.88		0.45	
	Balanced Accuracy	0.92	0.67	0.57	0.57
	Precision	0.99	0.36	0.68	0.30
	Recall	0.85	0.67	0.40	0.37
	F1 Score	0.92	0.47	0.50	0.33
Lexicon-Based - `bing`	Overall Accuracy	0.88		0.48	
	Balanced Accuracy	0.92	0.71	0.57	0.60
	Precision	0.99	0.43	0.68	0.30
	Recall	0.85	0.73	0.40	0.42
	F1 Score	0.92	0.54	0.50	0.35
Machine Learning	Overall Accuracy	0.95		0.74	
	Balanced Accuracy	0.95	0.81	0.83	0.77
	Precision	0.98	0.84	0.69	0.63
	Recall	0.94	0.74	0.80	0.66
	F1 Score	0.96	0.79	0.74	0.65

2.6 Data Processing

Based on the superior method identified with the optimized algorithm—the random forest machine learning method with xgboosting—each game first had to be analyzed individually. This started with the originally exported data files from the Twitter API, meaning this is the same starting point as the example of the tweets extracted from the Arsenal vs Brighton match presented throughout this report (Table 2). Detailed information about this code can be found in Appendix G and I, and the entire data processing procedure is summarized in Figure 3.

2.6.1 General Pre-Processing of Game Tweets

With the correct game loaded into code important game variables such as halftime and fulltime were accounted for so the correct game minute for each tweet could be identified. Various functions then transformed the data, so it could run through the algorithm. Firstly, the emojis had to be translated to text, then various different methods worked to clean the text in these tweets including to remove punctuation, numbers, the customized stop words dictionary made for this project, typical English stop words, and to stem each word (Table 12). It is important to note that the general English stop words were removed from the test data but not from the training data.

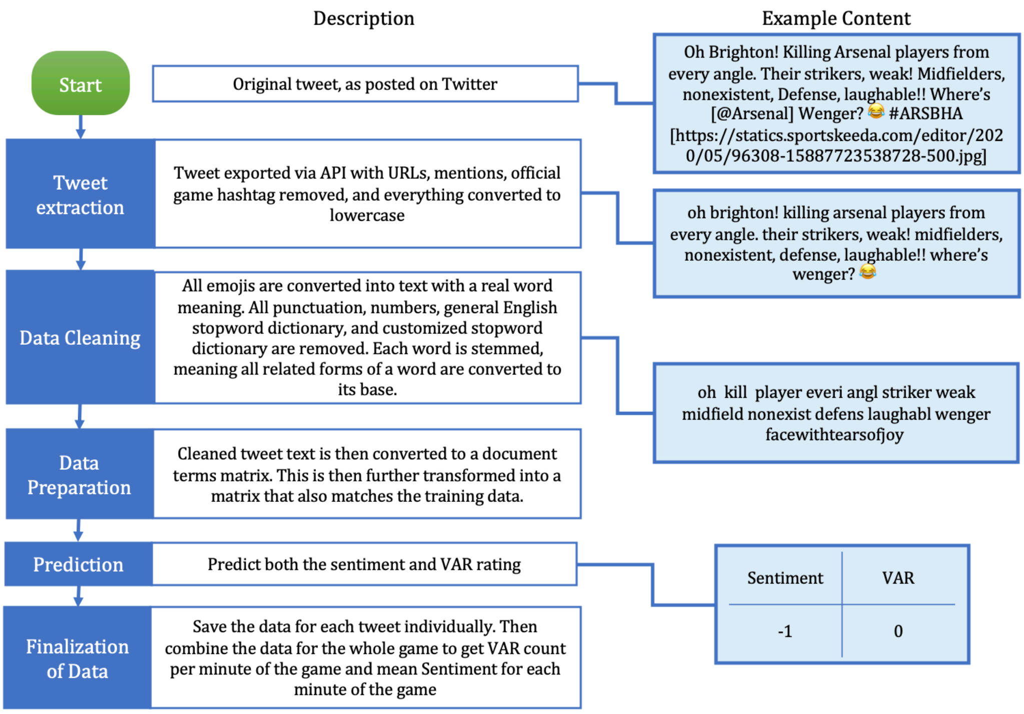


Figure 3. Data processing steps with the example of a sample tweet from the Arsenal vs Brighton match. (••||, 2019)

These cleaned up tweets then were transformed into a document-term matrix for the algorithm to work properly. For this, each document, in this case, each tweet, represents the rows of the matrix and each term in the document represents the columns. The matrix is simply filled in by counting the occurrences of the terms in each document. Table 13 shows how this is set up for two of the tweets from the example in Table 12.

Table 12. Sample cleaned tweets for the machine learning method from the Arsenal vs Brighton examples shown in Table 2. Note here the emojis are converted to a text form, punctuation, numbers, the customized stop words, and general English stop words are all removed, and each word is converted to its base form through stemming.

Original Exported Text	Cleaned Text
coyg!!! we winning this. 💪🏽 (Szeth-son-son-Vallano, 2019)	coyg win flexedbicepsmediumskinton
why aren't var picking up these fouls in the box against lacazette? (Andrade, 2019)	whi arent var pick foul box
are arsenal really that bad now that a team like brighton can come to the emirates and dominate them all over the park... (Sam, 2019)	realli bad now team like can come domin
var is killing the passion in football (Dan, 2019)	var kill passion
i have said it countless times, arsenal problem is not the coach, it's the players, they lack the zeal to win games. (Adeleke, 2019)	said countless time problem coach player lack zeal win game
this club is a joke (Papoose, 2019)	club joke

Table 13. Two samples of cleaned tweets and the respective document-term matrix. Note here the rows refer to the documents, in this case, the tweets, and the columns refer to the terms in the document with the number representing the count of that term in that document.

Cleaned Text	Document-Term Matrix							
	arent	box	foul	pick	var	whi	kill	passion
whi arent var pick foul box	1	1	1	1	1	1	0	0
var kill passion	0	0	0	0	1	0	1	1

This matrix is then converted into a matrix that matches the columns of the training data. For this, a new matrix was made with the same columns referring to the terms in the training data and the same rows to the testing data of the game referring to the documents that were to be analyzed, with all cells starting as 0. For each document of the test data, the code looped over the columns of the document-term matrix of the game data until it found a column, or term in this case, that was also present in the training document-term matrix which is made up of 4,233 columns. When this occurred the code copies the value from the game document-term matrix, referring to a count of how many times that term appeared in the tweet currently being examined, into the result matrix.

2.6.2 Prediction Algorithms For Game Tweets

Table 14. Probability results from sample tweets of Arsenal vs Brighton. These probabilities were calculated using a supervised machine learning algorithm. The rating column is determined by the category with the maximum probability.

Original Exported Tweet	**Sentiment Probabilities**				**VAR Probabilities**		
	Negative -1	**Neutral 0**	**Positive +1**	**Rating**	**Other 0**	**VAR 1**	**Rating**
coyg!!! we winning this. 💪🏽 (Szeth-son-son-Vallano, 2019)	0.19	0.25	0.56	1	0.01	0.99	0
why aren't var picking up these fouls in the box against lacazette? (Andrade, 2019)	0.86	0.09	0.05	-1	1.00	0.00	1
are arsenal really that bad now that a team like brighton can come to the emirates and dominate them all over the park... (Sam, 2019)	0.44	0.19	0.37	-1	0.01	0.99	0
var is killing the passion in football (Dan, 2019)	0.75	0.06	0.19	-1	1.00	0.00	1
i have said it countless times, arsenal problem is not the coach, it's the players, they lack the zeal to win games. (Adeleke, 2019)	0.48	0.18	0.34	-1	0.00	1.00	0
this club is a joke (Papoose, 2019)	0.61	0.22	0.17	-1	0.17	0.83	0

Once completed, the new result matrix is then inserted into a function along with the optimized sentiment model that will predict sentiment values for each tweet by calculating the probabilities that the tweet falls into each sentiment category, shown in Table 14. The highest probability is then used to assign each tweet a value. These categories include a positive, negative, and neutral rating.

Once the sentiment is determined, the same matrix is run through another function with the optimized VAR model to categorize each tweet into a VAR category the same way (Table 14). The categories here include VAR and no VAR (other).

Here, the selected algorithm performed particularly poorly on one special case game. This was the Manchester United vs Manchester City (MUNMCI) game. Here, after Manchester United won, many fans were tweeting the phrase "Manchester is red" referring to their team winning the Manchester derby. This phrase was labeled with this algorithm as a VAR reference. This special case resulted in this particular game to be excluded from all analysis, as mentioned previously.

2.6.3 Finalized Data for Export for Each Game

Finally, once each tweet has a sentiment and VAR rating all this data needs to be combined to get a final export table made up of the total VAR count and the average sentiment of all the tweets for each minute since kickoff, which with the help of the known stoppage times in each half is converted to game minutes (Table 15). Another finalized exported table, that was saved for each game, was made up of the raw file containing all the information about each tweet as exemplified in Table 2, but also included the newly calculated sentiment rating and VAR count of each tweet individually, which is useful for a deeper analysis of the game. This process than needed to be repeated for each tweet of the 129 games.

Table 15. Final export data for the machine learning method with VAR count and respective mean sentiment shown for various game minutes of the full data set of the Arsenal vs Brighton match.

Minute Since Kick Off	VAR Count	Mean Sent	Game Minute
4	0	0.15	4
12	19	-0.31	12
31	16	-0.25	31
84	33	-0.24	69
99	0	-0.13	84

The best way to run all these games with the complex algorithm for processing each individually was found to run games in parallelization, so multiple games could be running at once on different threads with different instances of the programming software on a computer that has sufficient RAM and CPU cores. While the programming software here was convenient for this project, it is not the most efficient option to run analysis at this level and further research could look into optimizing the coding for situations where even more tweets and games are analyzed to ensure the time it takes to run the code does not exponentially increase. (Appendix O).

2.7 Data Analysis

Once each tweet is processed from the Twitter platform and VAR count and sentiment rating information for each game is combined, further analysis was done to see how sentiment changes in response to VAR calls. For this, three different approaches were used to analyze what sentiment trends are present in relation to VAR calls. Also used were different functions from the R packages `stats`, `irr`, and `cocor` (Aslam, 2020; Diedenhofen, 2016; Gamer, Lemon, Fellows, & Singh, 2019). (Appendix J and K).

The first approach looked through all the tweets of each game and combined those related to VAR and those that were not and calculated the mean sentiment value for these two different categories while also keeping track of the number of tweets in each (Table 16). Once all this data is recorded for each game, the average sentiment of all the tweets related to VAR and the average sentiment of all tweets not related to VAR is calculated also accounting for the high variance in the number of tweets for the different games. Statistical significance was determined using a paired t-test to compare sentiment values of the two different categories of tweets.

Table 16. Example results of the first approach to analyze the data for the Arsenal vs Brighton game, including the number of VAR tweets and other (non-VAR) tweets as well as the mean sentiment of each group accordingly with the calculation accounting for the different tweet numbers for the different categories

A	B	C	D	E	F	G
Game	**Num VAR Tweets**	**Mean Sent VAR**	**Num Other Tweets**	**Mean Sent Other**	**Num * Sent VAR (B * C)**	**Num * Sent Other (D * E)**
ARSBHA	431	-0.72	11,091	-0.12	-310	-1300

Table 17. Example outcome of the second approach to analyze the data showing the results for the quantile calculations for the Arsenal vs Brighton game. Each period is made up of five minutes and the periods including halftime (HT) and the end of the game (End) is also noted.

Game	Period	Num Tweets	Mean Sentiment	VAR Count	Percent of VAR	Over Threshold	Percent Rank of Mean Sent
ARSBHA	1	214	0.10	0	0.0	TRUE	65.6
ARSBHA	2	153	0.12	4	2.6	TRUE	70.2
ARSBHA	3	190	-0.22	53	27.9	TRUE	9.6
ARSBHA	4	174	-0.05	11	6.3	TRUE	33.7
ARSBHA	5	234	-0.15	19	8.1	TRUE	18.2
ARSBHA	6	294	-0.18	11	3.7	TRUE	13.2
ARSBHA	7	423	-0.17	39	9.2	TRUE	13.8
ARSBHA	8	1151	-0.15	11	1.0	TRUE	17.0
ARSBHA	9	721	-0.17	4	0.6	TRUE	13.5
ARSBHA	10 HT	777	-0.25	31	4.0	TRUE	6.7
ARSBHA	11 HT	644	-0.27	9	1.4	TRUE	5.6
ARSBHA	12 HT	466	-0.20	8	1.7	TRUE	11.6
ARSBHA	13 HT	413	-0.02	3	0.7	TRUE	40.0
ARSBHA	14	372	0.13	11	3.0	TRUE	71.9
ARSBHA	15	422	0.04	13	3.1	TRUE	53.1
ARSBHA	16	241	-0.09	10	4.1	TRUE	27.4
ARSBHA	17	357	-0.24	148	41.5	TRUE	7.5
ARSBHA	18	183	-0.02	18	9.8	TRUE	40.5
ARSBHA	19	164	-0.04	10	6.1	TRUE	37.4
ARSBHA	20	498	-0.06	0	0.0	TRUE	32.0
ARSBHA	21	669	-0.17	8	1.2	TRUE	15.0
ARSBHA	22	503	-0.21	2	0.4	TRUE	10.7
ARSBHA	23 End	1252	-0.16	2	0.2	TRUE	16.6
ARSBHA	24 End	1007	-0.18	6	0.6	TRUE	13.3

The second approach split each game into five-minute periods and calculated the mean sentiment as well as the number of tweets and the VAR count for each of these periods (Table 17). To ensure the periods of interest had a sufficient number of tweets for a proper analysis, certain exclusion criteria were first established. This decision was made by looking at the figures for all the games, from which it was noticed that even though games that comparatively had a low number of tweets, during certain interesting periods there are still enough tweets to analyze what is happening. Thus, by setting a threshold of the number of tweets for the different quantiles of the game, having to exclude complete games upfront is avoided. This threshold was set to 25 tweets per five minutes, meaning therefore that no one tweet's rating can theoretically change the average sentiment by more than 8%. Once this information is recorded, and the periods under the threshold are excluded. Then the periods are ranked as a percentage by that periods' mean sentiment considering the whole data set. This percentage rank is then categorized into different 5% quantile intervals and was used to get the average VAR count of periods in these quantiles.

Once the calculations were done all VAR counts were converted to percentage share of VAR within that period. This method was used to determine how VAR tweets are represented in sentiment based quantiles. Statistical analysis to determine if this the data followed an even distribution was done with a Kolmogorov-Smirnov test.

The third approach focuses specifically on the VAR peaks that occur during each game and what is happening with sentiment trends around these peaks (Table 18). For this, the first step is to define a VAR peak. After testing multiple criteria that included most true peaks while leaving out excess noise, the best identification definition for a trigger of a peak included: the share of tweets referring to the VAR for this minute had to be above 10%; the number of tweets referring to the VAR in this minute needed to increase by more than the natural logarithm of the total number of tweets for this game when compared to the previous minute; there has been no VAR trigger in the previous six minutes; and if there was a VAR peak at a previous time in this game, the share of tweets referring to VAR had to drop below 10% before a new trigger could be identified. It should also be noted that all peaks that happened during halftime or after the fulltime game whistle were excluded because it was found that during these periods of the game individuals often tweeted a summary or commented about the game as a whole instead of focusing on particular events and could therefore skew the results. Additionally, all peaks that occurred less than 15 minutes before the end of data collection were excluded to ensure enough data was present to analyze these peaks as desired. To ensure this method is sufficient to identify peaks, two researchers checked through randomly selected games to ensure all identified triggers of a peak were there for a reason. Once each trigger is identified, this acts as a representation of when a VAR incident starts, and therefore allows the sentiment rating right *before* the trigger to be compared to the sentiment ratings *after* the trigger. For this, however, the minute before the trigger was also excluded to eliminate the early tweets referring to a VAR incident that could skew the results. With this in mind, the mean sentiment of the five-minute period *before* the trigger was calculated and compared to the mean sentiment of the five-minute period *after* and including the trigger as well as a *later* measurement starting six minutes after the trigger for another five-minute period. Here also the number of tweets in each period was calculated because, similar to the quantile analysis, if any of the periods for a trigger had a total number of tweets below 25, the trigger and respective peak were excluded from the analysis. For statistical analysis, the mean sentiment of the period *before* the trigger was compared to the mean sentiment *after* the trigger, as well as the *later* period with a one-way repeated-measures analysis of variance (ANOVA) test.

Table 18. Example outcomes of the third approach to analyze the data showing the results for the peak trigger calculations for the Arsenal vs Brighton game. The time value depends on the classification, for a *before* trigger row the time is the start minute of the five-minute period, in the trigger row just the trigger minute is included, in the *after* and *later* trigger row the time corresponds to the end of the five-minute period.

Game	Minute Since Kickoff	Num Tweets	Mean Sentiment	VAR Count	Classification
ARSBHA	6	153	0.12	4	Before Trigger
ARSBHA	12	48	-0.31	19	Trigger
ARSBHA	16	185	-0.18	49	After Trigger
ARSBHA	21	191	-0.11	10	Later Trigger
ARSBHA	25	294	-0.17	9	Before Trigger
ARSBHA	31	83	-0.25	16	Trigger
ARSBHA	35	423	-0.17	39	After Trigger
ARSBHA	40	1151	-0.15	11	Later Trigger
ARSBHA	76	241	-0.09	10	Before Trigger
ARSBHA	82	88	-0.18	36	Trigger
ARSBHA	86	351	-0.22	145	After Trigger
ARSBHA	91	184	-0.07	18	Later Trigger

2.8 Author Contributions

For the methods, O.K. wrote the Twitter API code used to export the tweets, while M.K. executed the exportation of the games. O.K. formed the basic text analysis code with M.K. making adjustments as needed during execution. Both distributed the work of creating the training data. O.K. prepared the basis of the machine learning code for the training data and created the machine learning model. M.K. adjusted and executed the code to allow it to run for all the games. O.K. devised the presented idea for data analysis, while M.K. carried out the analysis writing the appropriate code to do so.

3. Results

With all the tweets labeled in terms of sentiment and VAR, interesting descriptive statistics about the sample were revealed. Here, the sentiment was ranked with a -1 referring to a negative sentiment, a 0 referring to a neutral sentiment, and a +1 for a positive sentiment. Additionally, VAR rating was a 1 for tweets related to VAR and a 0 for those that were not.

The total number of tweets analyzed related to VAR ended up being 58,211. However, the total VAR tweets per game ranged widely with the game Chelsea vs Manchester United having the most with 8,053 tweets relating to VAR in one game, compared to the minimum tweets related to VAR being 2 during the Wolverhampton vs Burnley match. This leads to an average of 451 tweets per game that are related to the VAR. Additionally, while each game had at least 1 tweet related to VAR, the number of tweets during a given VAR peak in a game also varied greatly. Here, the minimum number of tweets in a single VAR peak was 1 tweet per minute, while the maximum was 707 per minute.

When considering the analyzed VAR tweets, 75.4% of them were labeled as having a negative sentiment, while only 12.8% were labeled as neutral, and 11.9% as positive (Table 19). This is a much different distribution when compared to the other non-VAR tweets where 31.4% of tweets were labeled in the negative category, while 29.2% were neutral, and 39.4% were positive.

Table 19. Descriptive information of all analyzed tweets including count and percentages of the total number of collected tweets as well as additional percentages describing data

Total Sample							
	Sentiment						
	Negative		Neutral		Positive		Total
	Count	% of total	Count	% of total	Count	% of total	
VAR Tweets	43,881	6.8	7,424	1.2	6,906	1.1	58,211
Other Tweets	183,699	28.6	170,774	26.5	230,567	35.8	585,040
Total	227,580	35.4	178,198	27.7	237,473	36.9	643,251

Additional Percentages			
	Sentiment		
	Negative	Neutral	Positive
Percentage of VAR tweets that fall into sentiment category	75.4 %	12.8 %	11.9 %
Percentage of other tweets that fall into sentiment category	31.4 %	29.2 %	39.4 %
Percentage of sentiment category tweets related to VAR	19.3 %	4.17 %	2.91 %

The first data analysis approach examined the mean sentiment score of tweets related to the VAR and the score of tweets not related to the VAR. This method found the mean sentiment of the 58,211 tweets related to the VAR to be -0.64 and the mean sentiment of the 585,040 tweets not related to the VAR to be 0.08 (Figure 4). The mean sentiment of all 643,251 tweets examined in this study was found to be 0.02. These values take into account the varying numbers of tweets in each category. (Appendix M).

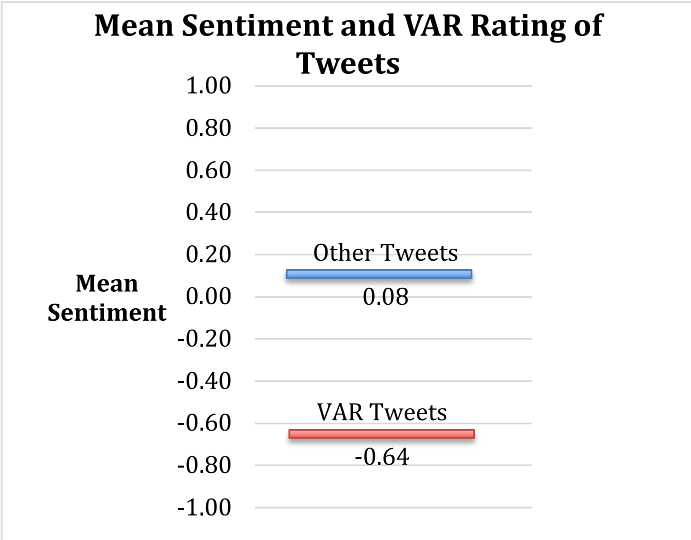


Figure 4. Mean sentiment rating of tweets categorized as being related to VAR incidents and those not.

Table 20. Paired t-test results to examine the relation between sentiment and the relatedness to VAR. Values here are paired by game and do not account for tweet number differences between the two groups.

Descriptive Summary

	n	Mean	Standard Deviation
Mean Sent VAR	129	-0.593	0.183
Mean Sent Other	129	0.098	0.098

Paired Samples T-Test

	t	df	p
Mean Sentiment VAR - Mean Sentiment Other	-42.651	128	< .001

A paired t-test was performed to examine the relation between sentiment and the relatedness to VAR (Table 20). A significant difference was found between the mean sentiment of VAR tweets (M = -0.593, SD = 0.18) compared to the mean sentiment for the other non-VAR tweets (M = 0.098, SD = 0.098; $t_{(128)} = -42.651$, $p < .001$), with the sentiment for VAR tweets being lower. Note the values here are paired by game and do not account for the tweet number differences between the two groups, which is the case in Figure 4.

The second approach, which split each game into five-minute periods with their respective mean sentiment as well as VAR count and ranked these periods by the percentage of mean sentiment, found that VAR tweets were overrepresented in quantiles with low average sentiment values. As seen in Figure 5, the lowest quantile, made up of the lowest 5% ranked sentiment periods, has the highest average VAR share with 43.4%. The second lowest quantile, made up of the 5-10% sentiment periods had the second highest VAR share with 15.8%. This trend shows that among the lowest ranked sentiment periods of the combined data set, the VAR share is the highest. Additionally, a Kolmogorov-Smirnov test revealed that the distribution of the average VAR count in the quantiles is significantly different from an even one (D(20) = 0.82, p < .001).

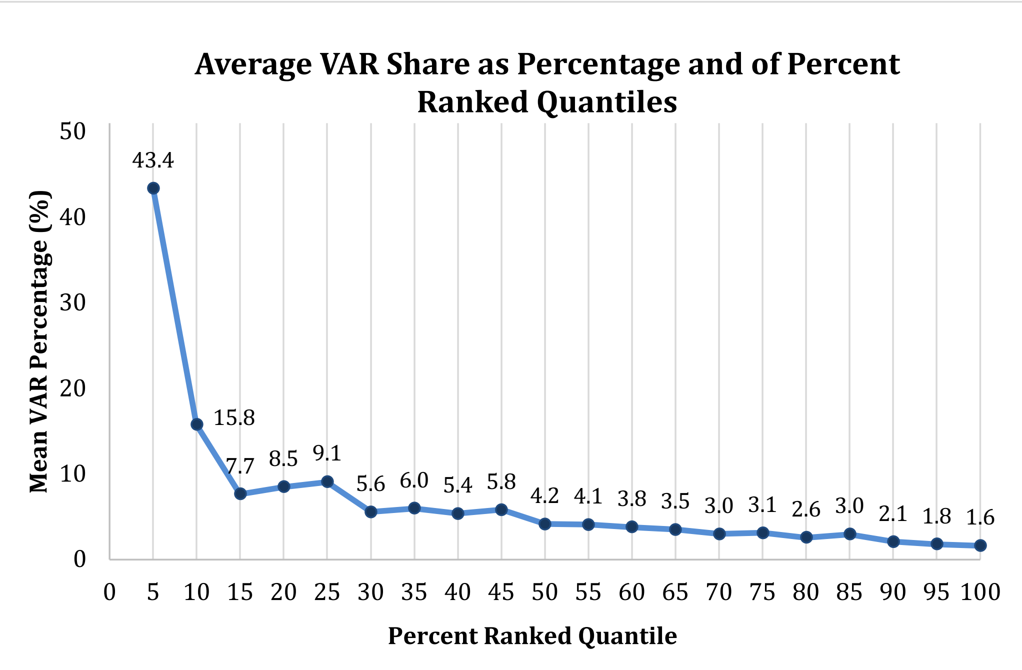


Figure 5. Average VAR share of different percent ranked quantiles including all periods over the threshold of the combined data set of all games. Each quantile is made up of 5% intervals.

The third approach, focused specifically at VAR peaks and how sentiment changed in the five minutes *before* the trigger of the VAR peak compared to the five minutes *after* and including the trigger of the VAR peak, and also compared to the *later* period starting at the sixth minute after the trigger (Figure 6). (Appendix N).

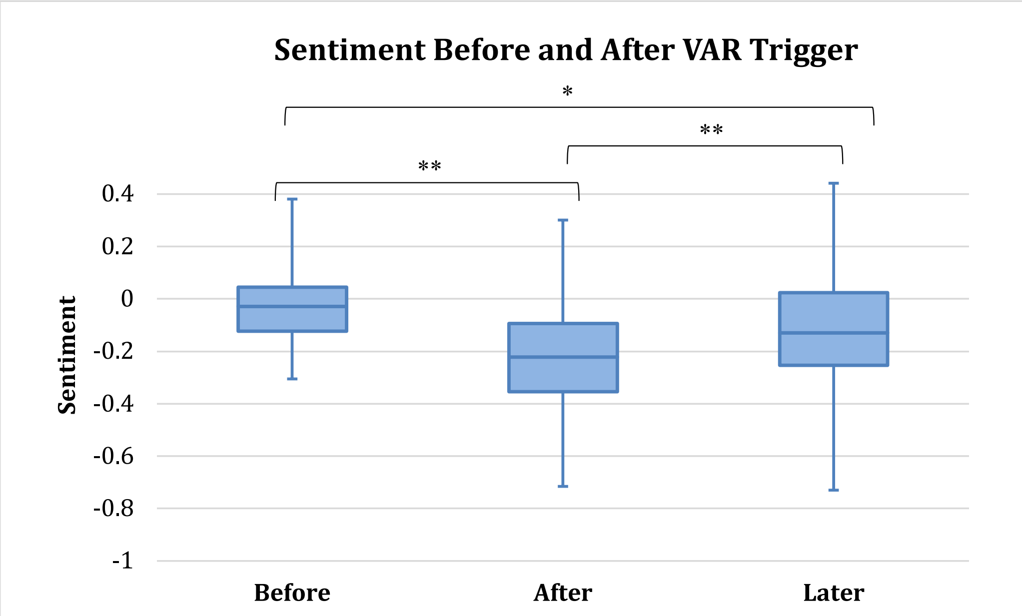


Figure 6. Box plot representing sentiment comparisons *before*, *after*, and six minutes *later* than the found VAR triggers. ** Significant difference p < .001 * Significant difference p < .01

A repeated measures analysis of variance (ANOVA) showed a significant difference between the mean sentiment of the three different time periods in relation to the trigger: *before*, *after*, and *later* [$F(2,180) = 29.831$, $p < .001$]. Post hoc tests using Bonferroni correction revealed that the mean sentiment of the period *before* the trigger (-0.029 ± 0.141) was significantly different ($p < .001$) to the period *after* the trigger (-0.223 ± 0.230). Additionally, the period *after* the trigger was significantly different ($p < .001$) to the *later* period (-0.131 ± 0.236) staring six minutes after the trigger. Finally, the mean sentiment of the period *before* the trigger was found to be significantly different than the *later* period ($p = .001$). Therefore, it can be concluded that mean sentiment changes in response to a VAR trigger with periods *before* and *after* such a trigger have statistically significant differences, as well as the *later* period starting six minutes after the trigger (Table 21).

Table 21. Statistical analysis results from an ANOVA comparing the sentiment in five-minute periods *before*, *after* a VAR trigger, and the *later* period starting six minutes after the trigger.

Descriptive Summary

	N	Mean	SD
Mean Sentiment Before	91	-0.029	0.141
Mean Sentiment After	91	-0.223	0.230
Mean Sentiment Later	91	-0.131	0.236

Within Subject Effect

	Sum of Squares	df	Mean Square	F	p
Sentiment	1.699	2	0.850	29.831	< 0.001
Residual	5.127	180	0.028		

Post Hoc Comparisons - Sentiment

Mean Sentiment		Mean Difference	95% CI of Mean Difference		SE	t	p_{bonf}
			Lower	Upper			
After	**Before**	-0.193	-0.262	-0.125	0.028	-6.865	< 0.001
	Later	-0.092	-0.137	-0.047	0.018	-5.010	< 0.001
Before	**Later**	0.101	0.034	0.168	0.027	3.695	0.001

4. Discussion

The purpose of this study was to analyze how soccer fans' react on Twitter specifically in response to VAR in the English Premier League through a textual and sentimental analysis of extracted tweets during these soccer matches. In order to also analyze the best way to obtain the data of interest, various exploratory methods were examined along the way.

First of all, from a methodological perspective, this study offers further insight into the application and potential of the described methodology to extract fan opinion in response to sports games. Twitter APIs are a useful and relatively simple way to extract data from social media that can be used to address important questions such as sentiment towards different events that take place during a match. While this report joins other research in recognizing the potential of using Twitter to obtain data about sports fans' sentiment, this project goes on to show, what such data, particularly when exported in large quantities can reveal about sports' event (Aloufi & Saddik, 2018; Fan et al., 2019; Jai-Andaloussi et al., 2015; Lucas et al., 2017; Yu & Wang, 2015; Zhao et al., 2011a). When it comes to analyzing this amount of big data, part of this project was to compare different methodologies found in literature and determine which approach is most appropriate in this context. While the lexicon-based method may be less tedious, the validation values found here deemed the method inappropriate to obtain quality information from the extracted Twitter data. In contrast, using machine learning methods leads to a much better performance. Therefore, this study reveals the potential of using Twitter APIs and machine learning techniques to extract and analyze big data in the form of great quantities of tweets.

The descriptive information about this analyzed big data set of tweets reveals some interesting findings. When looking at the sentiment classification label of all of the tweets, the highest percentage was made up of positive tweets, then negative, then neutral. Even with these differences, the make-up of the group of tweets as a whole is fairly well distributed when compared to the make-up of the group of tweets related to VAR. Here, an overwhelming majority (75%) of VAR tweets were placed in the negative category. One study that examined tweets from the 2014 World Cup, found that excitement on Twitter during a soccer game relates to expressions of negative emotions rather than positive ones (Lucas et al., 2017). While, as mentioned, the combined data found here is comparatively well distributed among the different sentiment categories, tweets related to VAR normally happened during a more concentrated time period showing a similar pattern to what was found in the study that games with higher tweets per minute were found to also have a higher percentage of negative tweets (Lucas et al., 2017).

To address the main aim of this study, various different analyses were done on the collected Twitter data. First of all, a significant difference was found for the sentiment rating of all tweets related to VAR (-0.64) and those that are not (0.08). Specifically, an overwhelming majority of VAR related tweets were labeled with a negative sentiment. This is an interesting finding because, it clearly illustrates that Twitter users generally feel negatively in response to topics related to VAR, for example, either in the form that VAR was involved and they disagree with the call, have complaints about part of the process, or that VAR was not involved in a play and they believe it should have been. Comparatively this finding appears to be rather extreme considering one study looking at Twitter fan responses to the England team during the 2018 World Cup, where researchers found there to be a sentiment difference of 0.26, using the same scale as used here, among fans when their team was winning when compared to losing (Fan et al., 2019).

The second analysis method found that quantiles with the lowest percent ranking in terms of sentiment tended to have a higher average share of VAR related tweets. This is an interesting finding because this method takes into account all negative sentiment tweets when finding the quantiles with the most negative sentiment. Even here with all the negative comments that occur during a game when, for example, the other team scores or an individual's own team is not doing well, VAR is still overrepresented in the lowest ranked sentiment quantiles in terms of count of VAR related tweets.

Finally, the third analysis found a significant difference between the sentiment value *before* an identified VAR trigger peak, that of the five-minute period *after* the trigger, as well as the 6-10 minute *later* period after the identified trigger of the peak. This finding shows not only that these identified triggers clearly mark a change in sentiment values of collected tweets in this timeframe but also that even in the *later* period 6-10 minutes after the trigger, the effect of VAR involvement or lack thereof continues to influence the conversations happening around the event on Twitter.

Overall this study displays clear significant negative sentiment on Twitter regarding the VAR in the English Premier League in the 2019-2020 season. Further examination reveals why this could be the case and also shows the general controversial nature of VAR in the examined season of the Premier League also found in the analyzed tweets.

From the start, one of the main concerns commonly expressed about how the spirit of the soccer game will change with VAR was focused around goal celebrations (Arastey, 2019). Fans as well as leaders in the world of soccer, have all addressed the issue that VAR reviews for goals would take the joy out of the game as no one is able to celebrate properly until the review is completed and by then the spontaneous joy is already lost thereby ruining the atmosphere in stadiums (Reuters Staff, 2019; Silva, 2020). This similar concern was also found in the analysis of the extracted tweets, for example: "This var actually takes the fun away from the game. Imagine scoring a goal and you are afraid of celebrating because you fear it will be ruled out by VAR" (chux, 2020); "Celebrating a goal a few minutes after scoring isn't even sweet. #var" (Fridays254, 2020); and "tired of not celebrating goals properly. VAR should not exist" (Zαcκ, 2020). These are all examples where the author may agree with the VAR decision to award the goal in the end but disagrees with the process because it ruins the goal scoring moment.

Additionally, interestingly, while some supporters of VAR, like Arsene Wenger, believe that "video will help the referee, not question their authority [as] it will give them more credit, more authority and fewer mistakes", one study found a different result (ESPN, 2012). This study surveyed subjects from an online panel and evaluated how VAR influences the spectators' perception of the quality, flow, outcome, and enjoyment of soccer matches and also how observers perceive referees' performance, credibility, and authority (Anik, 2019). These participants mentioned that they believe that VAR involvement in soccer leads referees to make more mistakes and take more risks. It was found that a referee who made the correct call in a match without using VAR was considered to be more competent. Alternatively, a referees' reputation was negatively affected when they used VAR, even when they made the correct decision to overturn the original call (Anik, 2019). This was also reflected in some of the tweets considered in this study, for example in the Arsenal vs West Ham United (ARSWHU) match, there was a VAR incident where an offside call was overruled into a goal for Arsenal. For one of the tweets of the game: “Why did that VAR check take so long? Seemed pretty clear cut to everyone except for Martin Atkinson” (ConsistentlyCompetentBlogs, 2020), the author is saying they believe in the usage of the VAR because the initial call by the referee, Martin Atkinson, was wrong, however, they do not understand why it takes so long. Shown here is an example of how, while individuals may support the usage of VAR, they still have complaints about the process and believe it also shows the incompetence of the referee.

The complaints against VAR are broad in spectrum, and while it may have improved the accuracy of refereeing, critics argue it has overstepped its bounds of correcting “clear and obvious errors” making the sport worse off as a result (Bushnell, 2019). Similar criticism was already being expressed for the implementation of VAR in the 2018 World Cup after which reporters complained that VAR fails to correct human error and instead only adds to the controversies because human judgment is still part of the process (Stinson, 2018). This was a concern from the beginning for implementing more technology into the game of soccer due to the general strict and subjective nature of soccer’s laws (Arastey, 2019). One such acknowledged incident occurred in the Chelsea vs Tottenham game (CHETOT). Here in the second half, Tottenham Spurs attacker Lo Celso stepped on the outstretched leg of Chelsea player Azpilicueta as both players competed for a 50-50 ball (Casey, 2020). While the referee called it a foul, VAR intervened to check for a possible red card for a serious foul play. At the time, VAR decided no card should be given. This was met with outrage from fans on Twitter who believed the incident should be awarded a red card: “That is a crystal clear red for Lo [Celso]… a career ending challenge. A ridiculous call. Shocking. Atrocious. I'm running out of words.” (Sen, 2020); “If var won't send Lo Celso off for that... Then why is var even in the game?” (Matt, 2020); and “VAR is trash. That was a straight red card” (Tkay, 2020). In fact, when looking at the collected data during this incident the sentiment of tweets from before the trigger to the height of the incident dropped by a rating of 0.66, showing clear negative opinion in response to this call. Then in the game, a half an hour later, the sports commentator informed those watching that VAR officials had admitted their decision not to send off Lo Celso was the wrong one (Casey, 2020). During this point in the game data, a VAR peak emerges again as fans on Twitter are once again outraged at the VAR system: “VAR is so poor it's actually embarrassing… how can you openly admit 20 minutes after there should be a red card? so bad” (Harris, 2020); “VAR admitting that they made a mistake by not sending off Lo Celso tells you everything you need to know about VAR.

Absolute joke" (Pavey, 2020); and "Yet another VAR failure, this time admitted by Stockley Park themselves" (Ben, 2020). Here, in response to the admitted wrong VAR call, with many people tweeting about it, the sentiment once again notably decreases by 0.45.

The complaints against VAR, however, do not stop there. In the examined season of the English Premier League, VAR has reversed some marginal, previously undetectable errors like an offsides call by a measure of centimeters only made possible by virtually adding lines to the video (Bushnell, 2019). Critics argue that while offsides is supposed to be factual, it is currently anything but as determining where exactly a player's shoulder ends and an arm begins as well as when exactly a ball left a player's foot are all human decisions and per the current rules, any part of the body that can play the ball can be declared to be offsides by mere millimeters (Bushnell, 2019; White, 2020).

In fact, this controversy is a hot topic in the examined season of the Premier League with players, managers, fans, and even the sports' governing body going so far as to call for an immediate change for the way the league is using VAR (De Menezes, 2019). In response to goals being ruled out by armpits or toes considered to be offside, the International Football Association Board suggested that the way the Premier League is currently using VAR specifically for offsides calls is not what the system was meant to be used for: "if something is not clear on the first sight, then it's not obvious and it shouldn't be considered. Looking at one camera angle is one thing but looking at 15, trying to find something that was potentially not ever there, this was not the idea of the VAR principle. It should be clear and obvious" (De Menezes, 2019). Some suggestions have even come about as now being a good time to look at the laws of the game, particularly for example the offsides rule to try to avoid goals being disallowed over a matter of centimeters, as such microscopic video analysis of offsides goes against the original spirit of the game (Arastey, 2019; Panja, 2019). This controversy is also just as present for the extracted Twitter data: "Enough is enough with these scientifically suspect offside lines. Change the law - if lines are needed it should be deemed level and onside. Common sense." (Darke, 2019); "What a joke. Half his toe is offside. Explain to me how [a player] gains an advantage for a headed goal by his toe being offside by an inch. Ridiculous. I've had enough with VAR now. It takes away too much, and gives nothing back" (L. J. Wilson, 2020); and "and VAR calls it off, what a joke, another armpit offside" (Nelson, 2019).

These opinions shown here by Twitter users about the Premier League is not limited to fans of the online platform. Similar opinions were also expressed openly during the examined season by members of the Premier League itself. For example, Chelsea manager Frank Lampard has described VAR as a "passion killer," West Ham midfielder Declan Rice has made a statement claiming that almost every professional wants the league to scrap the system, additionally Jose Mourinho, the manager of the Tottenham Spurs team has stated: "this is England, this is the Premier League, this is the best competition in the world, [and when we try to change things like adding VAR], we are killing the best league in the world" (Reuters Staff, 2019; Silva, 2020). Some players also showed outrage specifically regarding the usage of VAR for offsides calls, for example, Wolves captain Conor Coady spoke out against VAR saying, "a lot of people are going to tell me that they have come to the right decision and they might have. But what is it, an armpit, that's offside, or a toe, or something like that? … It's horrible for me, it's tough to take" (De Menezes, 2019). Fans are also involved showing their dissatisfactions during games with anti-VAR chants and banners being found at soccer grounds across the country (Silva, 2020). With all this considered, it is no surprise that the analyzed data from Twitter shows a similar trend with the sentiment towards VAR being overwhelmingly negative.

Another interesting finding from this study demonstrated how some of the initial concerns of VAR might not actually be as bad as stakeholders believed. Before its implementation, it had been suggested that VAR in soccer will affect the nature of the sport's fanbase as it will remove the much enjoyed debate amongst fans about the accuracy and appropriateness of an officiating decision that has become a key part of the soccer supporter experience and satisfaction (Arastey, 2019; Fédération Internationale de Football Association, 2010). One of the famous examples that helped initiate the talks for implementing goal line technology into the game of soccer was Frank Lampard's disallowed goal at the 2010 World Cup (Singh, 2013; Winand & Fergusson, 2018). Here, although the match official made the wrong decisions, this researcher argues that the controversy surrounding this decision ended up enhancing many supporters' enjoyment of the match (Nlandu, 2012). Critics claim that fans would rather watch a free flowing imperfect game than a perfectly refereed match where they sit in silence, thereby addressing the game of soccer as entertainment rather than purely a sporting machine (Arastey, 2019). While this may have been considered the case before the implementation of VAR, the examined Premier League games ignited a large amount of debate around the usage of VAR as it is by no means controversy free and fans continue to question its appropriateness as well as the outcome of decisions.

For example, in the match of Leicester City vs Everton (LEIEVE), after an assumed foul, the referee awarded Leicester City with a penalty kick. After this incident was reviewed via VAR, the original decision from the referee was overturned and no penalty was given. In response, there was a debate about the incident on Twitter. On one side there are plenty of people arguing the foul should be awarded a penalty, for example: "If he gets tackled there it could break his leg.… It's a foul" (m i k e, 2019); and "don't know how that was overturned. Penalty all day long" (Niall C, 2019). On the other hand, plenty of people are also agreeing with the VAR decision that no penalty should be called: “Never in a million years was that a pen! No contact and he got the ball” (R a c h, 2019); and “definitely not a pen. No contact whatsoever” (MCDMcMC, 2019). Another side of the debate was also mentioned: “Strange one. Was difficult to tell if contact or not. Would have thought that means referee original decision stands” (Briers, 2019). For this incident many fans were also bringing up the point that if the original penalty call did not stand, that means the player dove, or simulated getting fouled, which is punishable in the premier league: “If that’s not a pen, it’s surely a yellow for trying to con the ref?” (Kayfabe, 2019); “correct VAR decision, no booking for the dive though?” (G G, 2019; Premier League, 2017). This is just one of many examples of how the debate spirit among fans still continues to be present even with the implementation of VAR.

All in all, while VAR was introduced into the game of soccer to make it fairer, this study brings to light the viewed imperfections by fans who believe that the way it is being used in the Premier League during the examined season does not fulfill that goal. As one Twitter user put it: “this is totally unfair, VAR has totally spoilt [soccer]” (Stranger, 2020).

4.1 Limitations

While some interesting trends and the potential of such a system are described here, due to the nature of this study, it is not without limitations. As previously mentioned, the lexicon-based approach has specific limitations itself which resulted in using a machine learning approach to analyze the data. Additional steps that have its own limitations include general limitations of the Twitter API, as well as other limitations that appeared throughout the study.

4.1.1 Limitations of Twitter API

There are some general limitations that exist when using a Twitter API that must be noted. First of all, this method here was limited to the English language for multiple reasons. As was tested, English tweets are a lot more plentiful than those of any other language. Additionally, since the methodology here involved rating specific words, in the English language a tweet will be rated based on the general sentiment due to the structure of the English language. In a language like German, it is common to use a double negative to express a positive sentiment rendering this system false. This is why the project was originally conducted for the English Premier League, however even if most of the fans tweeting about these games are English, the Premier League is followed worldwide and therefore foreign language tweets are still present during some of the analyzed matches and are often unable to be categorized. This limitation is commonly noted in other research pertaining to Twitter content and future research efforts could try to incorporate multiple languages to get more internationally relevant responses (Frederick et al., 2015).

Secondly, due to the used `RTweet` API, there is a 6-9 day limit to extract the tweets from the games (Kearney, 2019). This limitation of the program resulted in some games being lost.

Thirdly, at the beginning of this project, the `RTweet` API had a rate limit that prevented usable data from being obtained from big games, the solution for this was discovered only towards the end of the project.

Another limitation to consider is that there is a fan-based bias on Twitter, this means that for example the more well-known teams in the Premier League such as Manchester United also have a larger Twitter following than other teams. In fact, research has suggested that Manchester United has a following of 659 million supporters meaning that as much as 10% of the world's population supports this team (Dubber & Worne, 2015; Prior, 2013). Of these supporters, 21.7 million follow the official Manchester United Twitter account (Manchester United, 2020). For this project, this means that when this team is playing, the sentiment might more largely reflect that of this team than general fan opinion. Previous studies have found that fanship intensifies one's involvement with events and results of a game and therefore may affect the respective sentiment analysis (Yu & Wang, 2015). One way future research can account for this problem is to also analyze the location stamps of the tweets, as done by other studies (Fan et al., 2019; Yu & Wang, 2015). Although for these studies, location stamps were used to distinguish among national teams which are different from clubs, as here many fans are not necessarily tweeting from the home location of their club and this could also be problematic depending on where the game is being played to account for fans in the stadium. Another way to distinguish fans is to analyze the words in the tweet related to one team or another. One study did this to distinguish between teams and found that team names appear in over 60% of game-related tweets (Zhao et al., 2011a). For this to be done correctly, analyzers much also account for the action verbs associated with teams, for example, *go team A* and *beat team A*, suggest two opposite sentiments towards team A. This would, however, allow for researchers to identify how the sentiment of different teams is changing in response to events happening during the game (Zhao et al., 2011a).

In addition, in this project, only the tweets that used the official hashtag were considered, however, many more people are commenting on the game without using the analyzed hashtag. One study accounted for this limitation by using a predetermined lexicon made up of game terminology and team names to extract tweets related to specific games and found 10 keywords to be effective in extracting the game related tweets (Zhao et al., 2011a). A similar method was used in another study with multiple keywords or hashtags being used to extract more tweets than just those containing the official hashtag of the event (Fan et al., 2019).

Additionally, due to the nature of Twitter, it is difficult to determine the configuration of the users for specific events and how they compare from one event to another, for example. This is also a commonly expressed limitations in social media research (Frederick et al., 2015). There are a total of 13.7 million Twitter users in the UK, which represents only 24% of the UK population with the proportion using the platform for sports content expected to be even lower (Aslam, 2020). Compare this to the about 70% of the UK population, so around 47 million people, that watched the previous 2018-2019 Premier League season (Premier League, 2019). This limitation means that even if the majority of data collected came from England, input was only collected from a proportion of the population that watched the Premier League and is active on Twitter, meaning results here are not generalizable.

Additionally, the Twitter activity in this project does not differentiate the sentiment reaction of fans vs neutral audience members and therefore may result in a bias of opinion that may not necessarily portray the reactions of the wider audience. Other studies differing between the two have found divergent tweeting patterns affirming the above mentioned limitation (Highfield et al., 2013).

Finally, some games have a very low general tweet count for the game and therefore also a low VAR tweet count so nothing can be concluded from these cases. As mentioned before, one way to increase the tweets per game is to increase the keywords and hashtags being used to extract the tweets from the Twitter API, which would increase the tweet count. However, considering the wide range in numbers of supporters of the teams in the English Premier League it is possible that some games will still need to be excluded due to a low amount of Twitter activity.

4.1.2 Limitations of Machine Learning

Other limitations are due specifically to the machine learning approach taken here. For this, potential bias in the machine learning algorithm may be present as the researchers labeling the data were not blinded and had additional contextual information to ensure the tweets were properly labeled. Since this project also analyzed for sarcasm in tweets, additional contextual information was necessary and allowed researchers to be more confident of their sentiment label and VAR label of the training data.

When taking a closer look at the tweets labeled with the machine learning method, one thing in particular becomes noticeable. Of the 4,583 tweets that were used for the training data, 1,425 of them were labeled as relating to a VAR incident.

Of these, 203 tweets were labeled to have a neutral sentiment, 230 tweets were labeled as a positive sentiment and 992 were labeled as having a negative sentiment. While the rules for these labels were agreed upon by the researchers and therefore deemed correct, the huge proportion, 70%, of labeled VAR tweets being negative could have skewed the machine learning algorithm to label specifically negative tweets as being related to VAR. Therefore, it is possible the results presented here are skewed in a negative direction. It must be noted labeled training data was determined based on games that fulfilled certain requirements with random tweets taken from these games, so it can also be said generally there are more negative tweets relating to the VAR than positive and neutral. Unfortunately, for this issue, similar to the chicken and egg problem, it is not clear exactly what logic is the issue: are more tweets being labeled as VAR because they are considered to have a negative sentiment, or are, in general, more negative tweets related to VAR incidents. To further test this theory and clarify this issue, future studies could make sure to have a relatively even distribution in labeled training data for the three different sentiment categories for VAR related tweets to see if this changes the outcome.

Another potential limitation related to the training data of the machine learning algorithm is that the researchers here are not trained psychologists and therefore have no real training in declaring sentiment. However, for this project, a general understanding of if a tweet is positive or negative seemed sufficient. Further, neither researcher has extensive experience understanding English culture and therefore, some tweets containing many slang phrases and cultural references were difficult to properly analyze.

4.1.3 Additional Study Limitations

Unfortunately, the limitations of this study are not only constrained in the methodology used.

Moreover, sportscasters have been described as professional gatekeepers and embellishers who are responsible for creating an added effect in sports media by inserting commentary and opinions into a broadcast and thereby further dramatizing the event (Comisky, Bryant, & Zillmann, 1977). Previous research has shown that sports commentary influences the viewer's perception of an event that takes place during a sports match mostly by intensifying the drama (Comisky et al., 1977). Therefore, sports commentary can be said to operate as a framework offering viewers with not only additional understanding and knowledge of the presented gameplay but also potentially altering their perceptions of it (Frederick, Lim, Chung, & Clavio, 2013). When it comes to VAR calls, for example, viewers may be influenced one way or another due to the comments they hear from the commentators. Additionally, it should be noted that commentators' names were often found in the tweets, suggesting that viewers do pay attention to what the commentators are saying.

4.2 Outlook

Along with the aforementioned changes that can be implemented to improve the results from this study due to the limitations, further outlook based on what was found here could reveal additional findings.

Firstly, it could also be interesting to compare these results to games where VAR is not implemented to see how sentiment changes just to referee calls, or other events during a soccer game and how that compares to sentiment changes when the VAR system is part of the game.

Further analysis can also be done to determine how the sentiment changes in response to different types of VAR calls, for example, the trend might be found to be more negative if a goal is disallowed due to involvement of VAR than if it is something like a confirmed offsides call. As previously mentioned, an additional analysis could also show how the sentiment of the team changes in response to various incidents. This could be interesting to see how different teams with referee calls affecting them differently feel about the intervention of the VAR in these situations.

With these distinctions clarified, further analysis can also be done with this data in social science research specifically in the field of behavioral economics. One study has previously shown that Twitter data and sentiment analysis techniques can be used to get new insights into existing theories of what it is that makes sport exciting (Lucas et al., 2017). This study found that, on Twitter, excitement during a sports match is related to expressions of negative emotions (Lucas et al., 2017). Another study also used Twitter data to research behavior theories in sport by examining how social identity and team identification changes depending on the success of a fan's team during a soccer match (Fan et al., 2019). These studies have already exemplified the potential Twitter data has for a social scientist to analyze sports fans' behavior. Therefore, with the data acquired here, further analysis into theories of emotions could also be done. Specifically, examining the change in fans' response to a sequence of events of the game, in this case involving the VAR in a soccer match, as well as how this reaction may differ over multiple games or an entire season, for example. An additional point of interest could also be how reactions to events change based on other events that happened during the same game, for example, are fans of one team going to react differently after consecutive VAR calls or after the opposing team had a similar VAR situation. Sports economists could also use the data generated in this study to research what events happening in a soccer match, in particular, makes it more exciting and addictive to fans and how the implementation of VAR might change that (Szymanski, 2003).

Finally, in order to make the results of this study more generalizable, other national or international competitions should be further explored in the future. While the English Premier League is one of the most watched soccer leagues in the world, the English subculture may also be present on Twitter, leading to specific reactions to sports events, which could result in sentimental feedback that may not be internationally generalizable as different cultures may react differently. Additionally, how the VAR system was implemented in the Premier League is different to other leagues and international competitions.

5. Conclusion

This work describes the usage of fan generated data extracted from Twitter to examine the effect the introduction of VAR has on the Premier League. Shown here is the methodology that can be used to produce large databases on which text analysis can be done. Therefore this report joins other previously published articles in demonstrating the potential of Twitter as a unique information source regarding sporting events. Once this database was established various attempts were used for event detection of VAR and to determine the sentimental response of fans to these incidents. One way to do such analysis is to use the text mining strategy of a lexicon-based approach however apparent shortcomings limit the practical application of this method and suggest that further development needs to be done to improve parts of the analysis. Therefore, an additional text mining strategy involving an optimized machine learning approach was deemed more appropriate. With this methodology, clear results emerged suggesting fans on Twitter express a significant negative opinion towards VAR usage during the examined games of the Premier League. This indicates that the main objective of VAR to correct clear and obvious errors without significantly interrupting the game was not achieved during its introduction into the Premier League. Comparing these findings to other articles describing the complaints of fans during the same games shows similar negative trends, confirming these findings. While the findings presented here are limited specifically to the Premier League's usage of VAR, implications can be highlighted for other leagues and tournaments about the potential reactions from fans if VAR is administered the same way.

References

••|| [@DynaEkwueme]. (2019, December 5). Oh Brighton! Killing Arsenal players from every angle. Their strikers, weak! Midfielders, nonexistent, Defense, laughable!! Where's Arsenal Wenger? #ARSBHA [Tweet]. Retrieved from https://twitter.com/DynaEkwueme/status/1202691944603930630

Adeleke, C. [@AweelowBaba]. (2019, December 5). I have said it countless times, Arsenal problem is not the coach, it's the players, they lack the zeal to win games. #ARSBHA [Tweet]. Retrieved from https://twitter.com/AweelowBaba/status/1202707533770702848

Aloufi, S., & Saddik, A. E. (2018). Sentiment Identification in Football-Specific Tweets. *IEEE Access, 6*, 78609-78621. doi:10.1109/ACCESS.2018.2885117.

Andrade, D. [@d_andradesport]. (2019, December 5). Why aren't VAR picking up these fouls in the box against Lacazette? #ARSBHA [Tweet]. Retrieved from https://twitter.com/d_andradesport/status/1202685783360102400

Anik, L. (2019, July 31). *How Soccer's Video Assistant Referee (VAR) Influences Belief in Human Referees' Competence*. Behavioral Scientist. https://behavioralscientist.org/how-soccers-video-assistant-referee-var-influences-belief-in-human-referees-competence/

Arastey, G. M. (2019, November 6). *Application of Video Technology in Football Refereeing - VAR*. Sport Performance Analysis. https://www.sportperformanceanalysis.com/article/application-of-video-technology-in-football-refereeing-var

Armstrong, C. G., Delia, E. B., & Giardina, M. D. (2016). Embracing the Social in Social Media: An Analysis of the Social Media Marketing Strategies of the Los Angeles Kings. *Communication & Sport, 4*(2), 145-165. doi:10.1177/2167479514532914.

Arsenal [@Arsenal]. (2020). [Twitter Profile]. Retrieved from https://twitter.com/Arsenal

Aslam, S. (2020, February 10). *Twitter by the Numbers: Stats, Demographics & Fun Facts*. Omnicore. https://www.omnicoreagency.com/twitter-statistics/

Barbosa, L., & Feng, J. (2010). Robust Sentiment Setection on Twitter from Biased and Noisy Data. In *Proceedings of the 23rd International Conference on Computational Linguistics: Posters* (pp. 36–44). Beijing, China: Association for Computational Linguistics.

Bavishi, J., & Filadelfo, E. (2018, July 17). *Insights into the 2018 #WorldCup conversation on Twitter*. blog.twitter.com. https://blog.twitter.com/en_us/topics/events/2018/2018-World-Cup-Insights.html

Ben [@BenMichael1992]. (2020, February 22). Yet another VAR failure, this time admitted by Stockley Park themselves #CHETOT [Tweet]. Retrieved from https://twitter.com/BenMichael1992/status/1231220514703761410

Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., MuLler, S., . . . Fellows, I. (2020). Package ‘quanteda’: Quantitative Analysis of Textual Data. [R Documentation]. Retrieved from https://cran.r-project.org/web/packages/quanteda/quanteda.pdf

Billings, A. (2014). Power in the Reverberation: Why Twitter Matters, But Not the Way Most Believe. *Communication & Sport, 2*(2), 107-112. doi:10.1177/2167479514527427.

Bishop, C. (2006). *Pattern Recognition and Machine Learning* (1 ed.). New York: Springer-Verlag.

Blaszka, M., Burch, L. M., Frederick, E. L., Clavio, G., & Walsh, P. (2012). #WorldSeries: An Empirical Examination of a Twitter Hashtag During a Major Sporting Event. *International Journal of Sport Communication, 5*(4), 435. doi:10.1123/ijsc.5.4.435.

Boehmer, J. (2015). Does the Game Really Change? How Students Consume Mediated Sports in the Age of Social Media. *Communication & Sport, 4*(4), 460-483. doi:10.1177/2167479515595500.

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science, 2*(1), 1-8. doi:10.1016/j.jocs.2010.12.007.

Boyko, R. H., Boyko, A. R., & Boyko, M. G. (2007). Referee bias contributes to home advantage in English Premiership football. *Journal of Sports Sciences, 25*(11), 1185-1194. doi:10.1080/02640410601038576.

Brahmbhatt, Y. [@yashbrahmbhatt]. (2020, January 19). Can @ManUtd please win this. Lets go anything is possible. #LIVMUN [Tweet]. Retrieved from https://twitter.com/yashbrahmbhatt/status/1218951522341486592

Breiman, L., Cutler, A., Liaw, A., & Wiener, M. (2018). Package 'randomForest'. [RDocumentation]. Retrieved from https://cran.r-project.org/web/packages/randomForest/randomForest.pdf

Brid, R. S. (2018, October 26). *Introduction to Decision Trees*. Medium. https://medium.com/greyatom/decision-trees-a-simple-way-to-visualize-a-decision-dc506a403aeb

Briers, G. [@BriersGary]. (2019, December 1). Strange one. Was difficult to tell if contact or not. Would have thought that means Referee original decision stands. #LEIEVE [Tweet]. Retrieved from https://twitter.com/BriersGary/status/1201186074837565448

Brodersen, K. H., Ong, C. S., Stephan, K., & Buhmann, J. (2010, August 23). The Balanced Accuracy and Its Posterior Distribution. *Paper presented at the 2010 International Conference on Pattern Recognition*. (pp. 3121-3124). doi:10.1109/ICPR.2010.764.

Brown, L. (2020, March 13). *Premier League suspended over coronavirus*. Independent. https://www.independent.co.uk/sport/football/premier-league/premier-league-cancelled-coronavirus-suspended-liverpool-what-happens-epl-games-a9399666.html

Browning, B., & Sanderson, J. (2012). The Positives and Negatives of Twitter: Exploring How Student-Athletes Use Twitter and Respond to Critical Tweets. *International Journal of Sport Communication, 5*(4), 503. doi:10.1123/ijsc.5.4.503.

Bushnell, H. (2019, December 29). *VAR needs a rethink. Soccer should look to the NFL to fix it*. Yahoo Sports. https://sports.yahoo.com/var-premier-league-soccer-replay-review-system-offside-210847300.html?guccounter=1&guce_referrer=aHR0cHM6Ly93d3cuZ29vZ2xlLmNvbS8&guce_referrer_sig=AQAAANPo3oUwRaNl1LHPOP75ytCZNXQwAmBGNbDcYR0uAx-SHcbvtZkXdWtWDjvWR_eU2omcsAmy7cAwRPm0cr8R9RTY9B9YPgHuGMvLdKrQ3MxnNCx3HO7BMw6xcEPxlkh3Bi7g7hWf2DXY2HX7Xapd7AX5DctezZow7NycfknMqksW

Caruso, E. M., Burns, Z. C., & Converse, B. A. (2016). Slow motion increases perceived intent. *Proceedings of the National Academy of Sciences of the United States of America, 113*(33), 9250-9255. doi:10.1073/pnas.1603865113.

Casey, G. (2020, February 22). *VAR officials admit mid-game that they made 'human error' in not red-carding Lo Celso*. The 42. https://www.the42.ie/var-chelsea-v-spurs-5017894-Feb2020/

Ceron, A., Curini, L., Iacus, S. M., & Porro, G. (2014). Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France. *New Media & Society, 16*(2), 340-358. doi:10.1177/1461444813480466.

Chen, G. (2011). Tweet this: A uses and gratifications perspective on how active Twitter use gratifies a need to connect with others. *Computers in Human Behavior, 27*(2), 756-772. doi:10.1016/j.chb.2010.10.023.

Chen, T., & Guestrin, C. (2016, August). XGBoost: A Scalable Tree Boosting System. *Paper presented at the Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, California, USA. (pp. 785-794). doi:10.1145/2939672.2939785.

Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., . . . Li, Y. (2020). Package 'xgboost'. [R Documentation]. Retrieved from https://cran.r-project.org/web/packages/xgboost/xgboost.pdf

Chux [@chuxmac47]. (2020, February 22). This VAR actually takes the fun away from the game. Imagine scoring a goal and you are afraid of celebrating because you fear it will be ruled out by VAR. #CHETOT [Tweet]. Retrieved from https://twitter.com/chuxmac47/status/1231199491082084352

Clavio, G., Burch, L. M., & Frederick, E. L. (2012). Networked Fandom: Applying Systems Theory to Sport Twitter Analysis. *International Journal of Sport Communication, 5*(4), 522-538. doi:10.1123/ijsc.5.4.522.

Clavio, G., & Kian, T. M. (2010). Uses and Gratifications of a Retired Female Athlete's Twitter Followers. *International Journal of Sport Communication, 3*(4), 485-500. doi:10.1123/ijsc.3.4.485.

Cohen, D., & Kühne, S. (2019, July 19). *Collecting and Analyzing Twitter Data Using R*. Method Bites: Blog of the MZES Social Science Data Lab. https://www.mzes.uni-mannheim.de/socialsciencedatalab/article/collecting-and-analyzing-twitter-using-r/

Comisky, P., Bryant, J., & Zillmann, D. (1977). Commentary as a Substitute for Action. *Journal of Communication, 27*(3), 150-153. doi:10.1111/j.1460-2466.1977.tb02141.x.

Consistentlycompetentblogs [@BlogsCompetent]. (2020, March 7). Why did that VAR check take so long? Seemed pretty clear cut to everyone except for Martin Atkinson #ARSWHU [Tweet]. Retrieved from https://twitter.com/BlogsCompetent/status/1236330469500825604

Cromwell, G. (2009, August 9). *Armstrong's Tweet turns out more than 1,000 riders for a jaunt around Dublin*. VeloNews. https://www.velonews.com/news/armstrongs-tweet-turns-out-more-than-1000-riders-for-a-jaunt-around-dublin/

Cunningham, N., & Eastin, M. (2017). Second Screen and Sports: A Structural Investigation Into Team Identification and Efficacy. *Communication & Sport, 5*(3), 288-301. doi:10.1177/2167479515610152.

Daelemans, W., & Hoste, V. (2002, May 29-31). Evaluation of machine learning methods for natural language processing tasks. *Paper presented at the LREC*

2002: Third International Conference on Language Resources and Evaluation, Las Palmas, Spain.

Dan [@DanAFC_]. (2019, December 5). VAR is killing the passion in Football. #ARSBHA [Tweet]. Retrieved from https://twitter.com/DanAFC_/status/1202703741767602185

Darke, I. [@IanDarke]. (2019, December 28). Enough is enough with these scientifically suspect offside lines. Change the law - if lines are needed it should be deemed level and onside. Common sense. [Tweet]. Retrieved from https://twitter.com/IanDarke/status/1210988810747682818

Daugherty, T., Eastin, M. S., & Bright, L. (2008). Exploring Consumer Motivations for Creating User-Generated Content. *Journal of Interactive Advertising, 8*(2), 16-25. doi:10.1080/15252019.2008.10722139.

Davidov, D., Tsur, O., & Rappoport, A. (2010, August). Enhanced sentiment learning using Twitter hashtags and smileys. *Paper presented at the Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, Beijing, China. (pp. 241–249).

De Menezes, J. (2019, December 31). *Premier League is using VAR wrong for offside reviews, suggests IFAB general secretary*. Independent. https://www.independent.co.uk/sport/football/premier-league/premier-league-var-wrong-offside-review-ifab-goals-replays-a9264306.html

Deo, R. C. (2015). Machine Learning in Medicine. *Circulation, 132*(20), 1920-1930. doi:10.1161/circulationaha.115.001593.

Dequeiroz, G., Fay, C., Hvitfeldt, E., Keyes, O., Misra, K., Mastny, T., . . . Silge, J. (2020). Package ‘tidytext’. [RDocumentation]. Retrieved from https://cran.r-project.org/web/packages/tidytext/tidytext.pdf

Diedenhofen, B. (2016). Package ‘cocor’. Retrieved from https://cran.r-project.org/web/packages/cocor/cocor.pdf

Dohmen, T. J. (2008). The Influence of Social Foces: Evidence from the Behavior of Football Referees. *Economic Inquiry, 46*(3), 411-424. doi:10.1111/j.1465-7295.2007.00112.x.

Dubber, J., & Worne, J. (2015, October). *Playing the game: The soft power of sport*. British Council. https://www.britishcouncil.org/research-policy-insight/insight-articles/playing-game-soft-power-sport

Dyer, B. (2015). The controversy of sports technology: a systematic review. *SpringerPlus, 4*, 524. doi:10.1186/s40064-015-1331-x.

E.G [@EGCOYG]. (2020, March 7). Me everytime pablo mari made a pass to pepe #ARSWHU #coyg [Tweet]. Retrieved from https://twitter.com/EGCOYG/status/1236311606708985856

Earle, P., Bowden, D., & Guy, M. (2012). Twitter earthquake detection: Earthquake monitoring in a social world. *Annals of geophysics, 54*(6), 8. doi:10.4401/ag-5364.

Earle, P., M. Guy, R. Buckmaster, C. Ostrum, S. Horvath and A. Vaughan (2010). OMG Earthquake! Can Twitter improve earthquake response? *Seismological Research Letters, 81*(2), 246-251. doi:10.1785/gssrl.81.2.246.

Ebénn [@lema_yian]. (2019, November 23). #MCICHE VAR and Man City a love story that one [Tweet]. Retrieved from https://twitter.com/lema_yian/status/1198321816084926465

Ebner, S. (2013, July 2). History and time are key to power of football, says Premier League chief. *The Times*. https://www.thetimes.co.uk/article/history-and-time-are-key-to-power-of-football-says-premier-league-chief-3d3zf5kb35m

Èkó, M. [@mogajicares]. (2019, November 23). Chelsea tried though #MCICHE [Tweet]. Retrieved from https://twitter.com/mogajicares/status/1198321576472719360

Espn. (2012, April 13). *Wenger Wants Video Refs*. ESPN. https://www.espn.com/soccer/news/story/_/id/1048508/arsenal-boss-arsene-wenger-calls-for-video-assistance-to-referees

Etzel, J. (2015, December 18). *Balanced Accuracy: What and Why?* MVPA Meanderings: Musings on fMRI Multivariate Pattern Analysis. http://mvpa.blogspot.com/2015/12/balanced-accuracy-what-and-why.html

Expert System Team. (2017, May 6). What is Machine Learning? A definition. *Expert System*. Retrieved from https://expertsystem.com/machine-learning-definition/

Fan, M., Billings, A., Zhu, X., & Yu, P. (2019). Twitter-Based BIRGing: Big Data Analysis of English National Team Fans During the 2018 FIFA World Cup. *Communication & Sport, 8*(3), 317-345. doi:10.1177/2167479519834348.

Fédération Internationale De Football Association. (2010, March 11). *FIFA's position on technology in football*. FIFA.com. https://www.fifa.com/who-we-are/news/fifa-position-technology-football-1179851

Fédération Internationale De Football Association. (2018). *Video Assistant Referees (VAR)*. FIFA.com. https://football-technology.fifa.com/en/media-tiles/video-assistant-referee-var/

Feinerer, I. (2019). Introduction to the tm Package Text Mining in R. [R Documentation]. Retrieved from https://cran.r-project.org/web/packages/tm/vignettes/tm.pdf

Fisher, B. A. (1978). *Perspectives on human communication*. New York: Macmillan.

Frederick, E., Burch, L., & Blaszka, M. (2015). A Shift in Set: Examining the Presence of Agenda Setting on Twitter During the 2012 London Olympics. *Communication and Sport, 3*(3). doi:10.1177/2167479513508393.

Frederick, E., Lim, C., Chung, J., & Clavio, G. (2013). Determining the Effects of Sport Commentary on Viewer Perceptions, Attitudes, Beliefs, and Enjoyment through Violence Justification. *Journal of Sports Media, 8*(1), 65-86. doi:10.1353/jsm.2013.0009.

Frederick, E., Lim, C. H., Clavio, G., Pedersen, P. M., & Burch, L. M. (2014). Choosing Between the One-Way or Two-Way Street: An Exploration of Relationship Promotion by Professional Athletes on Twitter. *Communication & Sport, 2*(1), 80-99. doi:10.1177/2167479512466387.

Fridays254 [@Fridays254]. (2020, March 7). Celebrating a goal a few minutes after scoring isn't even sweet. #VAR #ARSWHU [Tweet]. Retrieved from https://twitter.com/Fridays254/status/1236331200999981056

Funk, D., & James, J. (2001). The Psychological Continuum Model: A Conceptual Framework for Understanding an Individual's Psychological Connection to Sport. *Sport Management Review, 4*(2), 119-150. doi:10.1016/S1441-3523(01)70072-1.

G G [@galvingreengg]. (2019, December 1). Correct VAR decision, no booking for the dive though? #LEIEVE [Tweet]. Retrieved from https://twitter.com/galvingreengg/status/1201185719236145153

Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2019). Package 'irr'. [RDocumentation]. Retrieved from https://cran.r-project.org/web/packages/irr/irr.pdf

Gandhi, R. (2018, May 5). *Naive Bayes Classifier*. Medium. https://towardsdatascience.com/naive-bayes-classifier-81d512f50a7c

Garicano, L., Palacios-Huerta, I., & Prendergast, C. (2005). Favoritism Under Social Pressure. *Review of Economics and Statistics, 87*(2), 208-216.

Gibbs, C., O'reilly, N., & Brunette, M. (2014). Professional Team Sport and Twitter: Gratifications Sought and Obtained by Followers. *International Journal of Sport Communication, 7*(2), 188-213. doi:10.1123/ijsc.2014-0005.

Girginova, K. (2016). New Media, Creativity, and the Olympics:A Case Study into the Use of #NBCFail During the Sochi Winter Games. *Communication & Sport, 4*(3), 243-260. doi:10.1177/2167479515576101.

Graver, F. (2012). *90 days at Twitter: What I learned from the inside*. Paper presented at the Ad Age Social Engagement/Social TV Conference, Los Angeles, CA.

Greenwell, B., Boehmke, B., Cunningham, J., & Developers, G. (2019). Package 'gbm': Generalized Boosted Regression Models. [RDocumentation]. Retrieved from https://cran.r-project.org/web/packages/gbm/gbm.pdf

Guest, J. [@theJEG]. (2020, January 18). See in that instance var has worked, no chance you can argue that's an intentional hand ball. #MCICRY [Tweet]. Retrieved from https://twitter.com/theJEG

Gupta, P. (2017, May 30). *Boosting the accuracy of your Machine Learning models*. Medium. https://towardsdatascience.com/boosting-the-accuracy-of-your-machine-learning-models-f878d6a2d185

Hambrick, M., Simmons, J., Greenhalgh, G., & Greenwell, C. (2010). Understanding Professional Athletes’ Use of Twitter: A Content Analysis of Athlete Tweets. *International Journal of Sport Communication, 3*(4), 454–471. doi:10.1123/ijsc.3.4.454.

Hambrick, M. E., & Mahoney, T. Q. (2011). 'It's incredible – trust me': exploring the role of celebrity athletes as marketers in online social networks. *International Journal of Sport Management and Marketing, 10*(3/4), 161-179. doi:10.1504/ijsmm.2011.044794.

Hamilton, W. L., Clark, K., Leskovec, J., & Jurafsky, D. (2016). Inducing Domain-Specific Sentiment Lexicons from Unlabeled Corpora. *Proceedings of the Conference on Empirical Methods in Natural Language Processing. Conference on Empirical Methods in Natural Language Processing, 2016*, 595-605. doi:10.18653/v1/D16-1057.

Hannon, J., Mccarthy, K., Lynch, J., & Smyth, B. (2011). *Personalized and automatic social summarization of events in video*. Paper presented at the Proceedings of the 16th international conference on Intelligent user interfaces, Palo Alto, CA, USA.

Harrington, S., Highfield, T., & Bruns, A. (2012). More than a backchannel: Twitter and television. *Participations, 10*(1), 405-409. Retrieved from http://snurb.info/files/2014/More%20than%20a%20Backchannel.pdf.

Harris, D. [@DMRHarris44]. (2020, February 22). VAR is so poor it's actually embarrassing Face with tears of joy how can you openly admit 20 minutes after there should be a red card? So bad. Rolling on the floor laughing #CHETOT [Tweet]. Retrieved from https://twitter.com/DMRHarris44/status/1231219868084686848

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Second ed.): Springer New York.

Hayes Sauder, M., & Blaszka, M. (2018). 23 Players, 23 Voices: An Examination of the U.S. Women's National Soccer Team on Twitter During the 2015 World Cup. *Communication & Sport, 6*(2), 175-202. doi:10.1177/2167479516685579.

Helsen, W., Gilis, B., & Weston, M. (2006). Errors in judging "offside" in association football: Test of the optical error versus the perceptual flash-lag hypothesis. *Journal of sports sciences, 24*(5), 521-528. doi:10.1080/02640410500298065.

Highfield, T., Harrington, S., & Bruns, A. (2013). Twitter as a Technology for Audiencing and Fandom. *Information, Communication & Society, 16*(3), 315-339. doi:10.1080/1369118X.2012.756053.

Hk [@HKM316]. (2020, January 18). I'm failing to comprehend the usefulness of V.A.R #ManCity #EPL #MCICRY [Tweet]. Retrieved from https://twitter.com/HKM316/status/1218572208945213443

Holmes, D., Alpers, G. W., Ismailji, T., Classen, C., Wales, T., Cheasty, V., . . . Koopman, C. (2007). Cognitive and emotional processing in narratives of women abused by intimate partners. *Violence Against Women, 13*(11), 1192-1205. doi:10.1177/1077801207307801.

Hu, M., & Liu, B. (2004, August 22-25). Mining and summarizing customer reviews. *Paper presented at the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Seattle, Washington, USA.

Hull, K., & Lewis, N. P. (2014). Why Twitter Displaces Broadcast Sports Media: A Model. *International Journal of Sport Communication, 7*(1), 16-33. doi:10.1123/ijsc.2013-0093.

Hutchins, B. (2011). The Acceleration of Media Sport Culture. *Information, Communication & Society, 14*(2), 237-257. doi:10.1080/1369118X.2010.508534.

Ike, K. [@Kofi_Kwarteng]. (2020, January 18). Hold on tight @CPFC #MCICRY [Tweet]. Retrieved from https://twitter.com/Kofi_Kwarteng/status/1218571097387544576

Indy Football [@IndyFootball]. (2020, February 16). VAR completes the check and no hand ball from Engels #AVFC #COYS #AVLTOT [Tweet]. Retrieved from https://twitter.com/IndyFootball/status/1229053806484303874

International Football Association Board. (2016, March 5). *130th Annual General Meeting of the IFAB*. The IFAB. http://static-3eb8.kxcdn.com/documents/644/124828_140218_2016AGM_Minutes_v1.0.pdf

International Football Association Board. (2017, April). *Video Assistant Referees (VARs): Implementation handbook for competitions*. https://www.docdroid.net/W2cATO8/var-handbook-v8-final-pdf#page=2

International Football Association Board. (2018). Laws of the game 2018/19. 134-142. Retrieved from https://img.fifa.com/image/upload/khhloe2xoigyna8juxw3.pdf

International Football Association Board. (2019, March 2). *133rd Annual General Meeting of The IFAB: sensible changes to make the game better*. The IFAB. https://www.theifab.com/news/133rd_annual_general_meeting

Jai-Andaloussi, S., Mohamed, A., Madrane, N., & Sekkaki, A. (2014). Soccer Video Summarization Using Video Content Analysis and Social Media Streams. *Paper presented at the Proceedings of the 2014 IEEE/ACM International Symposium on Big Data Computing*. (pp. 1–7). doi:10.1109/bdc.2014.20.

Jai-Andaloussi, S., Mourabit, I., Madrane, N., Chaouni, S., & Sekkaki, A. (2015, December). Soccer Events Summarization by Using Sentiment Analysis. *Paper presented at the International Conference on Computational Science and Computational Intelligence (CSCI)*, Las Vegas, NV, USA. (pp. 398-403). doi:10.1109/CSCI.2015.59.

Johnson, P., & Yang, S.-U. (2009). *Uses and gratifications of Twitter: An examination of user motives and satisfaction of Twitter use*. Paper presented at the Communication Technology Division of the annual convention of the Association for Education in Journalism and Mass Communication, Boston, Massachusetts, USA.

Joshi, R. (2016). Accuracy, Precision, Recall & F1 Score: Interpretation of Performance Measures. *Exsilio Solutions* Retrieved from https://blog.exsilio.com/all/accuracy-precision-recall-f1-score-interpretation-of-performance-measures/

Jurka, T. P. *wordStem*. RDocumentation. https://www.rdocumentation.org/packages/RTextTools/versions/1.4.2/topics/wordStem

Kassing, J. W., & Sanderson, J. (2010). Fan-athlete interaction and Twitter tweeting through the Giro: A case study. *International Journal of Sport Communication, 3*(1), 113-128. doi:10.1123/ijsc.3.1.113.

Kayfabe, D. [@daveteachingsci]. (2019, December 1). If that's not a pen, it's surely a yellow for trying to con the ref? #LEIEVE [Tweet]. Retrieved from https://twitter.com/daveteachingsci/status/1201185779902504961

Kearney, M. W. (2019). rtweet: Collecting and analyzing Twitter data. *Journal of Open Source Software, 4*(42), 1829. doi:10.21105/joss.01829.

Khalid, F. [@FrankKhalidUK]. (2019, November 23). Today @ChelseaFC had 53% possession & have @ManCity a game. Proud of this young team. Unlucky to of lost today. #MCICHE [Tweet]. Retrieved from https://twitter.com/FrankKhalidUK/status/1198321470805618689

Kolbinger, O., & Lames, M. (2017). Scientific approaches to technological officiating aids in game sports. *Current Issues in Sport Science (CISS)*. doi:10.15203/CISS_2017.001.

Kolbinger, O., & Link, D. (2016). The use of vanishing spray reduces the extent of rule violations in soccer. *SpringerPlus, 5*(1), 1572-1572. doi:10.1186/s40064-016-3274-2.

Kuhn, M. (2020). Package ‘caret’. [R Documentation]. Retrieved from https://cran.r-project.org/web/packages/caret/caret.pdf

Lago-Penas, C., & Gomez-Lopez, M. (2016). The Influence of Referee Bias on Extra Time in Elite Soccer Matches. *Perceptual and Motor Skills, 122*(2), 666-677. doi:10.1177/0031512516633342.

Lago-Peñas, C., Rey, E., & Kalén, A. (2019). How does Video Assistant Referee (VAR) modify the game in elite soccer? *International Journal of Performance Analysis in Sport, 19*(4), 646-653. doi:10.1080/24748668.2019.1646521.

Larkin, B. A., & Fink, J. S. (2016). Fantasy Sport, FoMO, and Traditional Fandom: How Second-Screen Use of Social Media Allows Fans to Accommodate Multiple Identities. *Journal of Sport Management, 30*(6), 643-655. doi:10.1123/jsm.2015-0344.

Lebel, K., & Danylchuk, K. (2012). How Tweet It Is: A Gendered Analysis of Professional Tennis Players' Self-Presentation on Twitter. *International Journal of Sport Communication, 5*(4), 461-480. doi:10.1123/ijsc.5.4.461.

Lex, H., Pizzera, A., Kurtes, M., & Schack, T. (2015). Influence of players' vocalisations on soccer referees' decisions. *European Journal of Sport Science, 15*(5), 424-428. doi:10.1080/17461391.2014.962620.

Lin, Y. (2019, November 30). *10 Twitter Statistics Every Marketer Should Know in 2020*. Oberlo. https://www.oberlo.com/blog/twitter-statistics

Liske, D. (2018, March 29). *Tidy Sentiment Analysis in R*. DataCamp. https://www.datacamp.com/community/tutorials/sentiment-analysis-R

Liu, B. (2012). Sentiment Analysis and Opinion Mining. *Synthesis Lectures on Human Language Technologies, 5*(1), 1-167. doi:10.2200/S00416ED1V01Y201204HLT016.

Liu, I., Cheung, C., & Lee, M. (2010). Understanding Twitter Usage: What Drive People Continue to Tweet. *Paper presented at the Pacific Asia Conference on Information Systems 2010*, Taipei, Taiwan.

Ljajić, A., Ljajić, E., Spalević, P., Arsic, B., & Vučković, D. (2015). *Sentiment analysis of textual comments in field of sport*. Paper presented at the 24th International Electrotechnical and Computer Science Conference ERK 2015, Portoroz, Slovenia.

Lucas, G. M., Gratch, J., Malandrakis, N., Szablowski, E., Fessler, E., & Nichols, J. (2017). GOAALLL!: Using sentiment in the world cup to explore theories of emotion. *Image and Vision Computing, 65*, 58-65.

M I K E [@mikebetsonstuff]. (2019, December 1). If he gets tackled there it could break his leg… it's a foul #LEIEVE [Tweet]. Retrieved from https://twitter.com/mikebetsonstuff/status/1201185513543221248

Madrid, R. (2018, April 5). *Real Madrid have reached 30 million followers on Twitter*. Real Madrid. https://www.realmadrid.com/en/news/2018/04/real-madrid-have-reached-30-million-followers-on-twitter

Makice, K. (2009). *Twitter API: Up and Running* (1 ed.). Sebastopol, CA: O'Reilly Media, Inc.

Malik, F. (2018, December 24). *How To Fine Tune Your Machine Learning Models To Improve Forecasting Accuracy?* Medium. https://medium.com/fintechexplained/how-to-fine-tune-your-machine-learning-models-to-improve-forecasting-accuracy-e18e67e58898

Manchester United [@ManUtd]. (2020). [Twitter Profile]. Retrieved from https://twitter.com/ManUtd

Markowetz, F. (2004). Classification by Support Vector Machines. *Practical DNA Microarray Analysis 2004*: Max-Planck-Institute for Molecular Genetics - Computational Molecular Biology.

Martín, D. (2019, December 29). *Real Madrid lead the world in social media followers*. as. https://en.as.com/en/2019/12/29/football/1577606426_305803.html

Marwick, A. E., & Boyd, D. (2010). I Tweet Honestly, I Tweet Passionately: Twitter Users, Context Collapse, and the Imagined Audience. *New Media & Society, 13*(1), 114-133. doi:10.1177/1461444810365313.

Matt [@MatthewSkilton1]. (2020, February 22). If VAR won't send Lo Celso off for that... Then why is VAR even in the game? #CHETOT [Tweet]. Retrieved from https://twitter.com/MatthewSkilton1/status/1231212281817636866

Mcdmcmc [@Dannymc5]. (2019, December 1). Definitely not a pen. No contact whatsoever. #LEIEVE [Tweet]. Retrieved from https://twitter.com/Dannymc5/status/1201185941966217222

Mckee, A. (2003). *Textual analysis: A beginner's guide* (1 ed.): SAGE Publications Ltd.

Medeiros, J. (2018). *The inside story of how FIFA's controversial VAR system was born*. Wired. https://www.wired.co.uk/article/var-football-world-cup

Mitchell, A., Kiley, J., Gottfried, J., & Guskin, E. (2013). *The role of news on Facebook: Common yet incidental*. Pew Research Center. http://www.journalism.org/2013/10/24/the-role-of-news-on-facebook

Nafees, S. [@nafeesitees]. (2019, November 23). finally the team won..... which I don't support #MCICHE [Tweet]. Retrieved from https://twitter.com/nafeesitees/status/1198321312130719744

Nelson [@AzumaNelson]. (2019, December 29). and VAR calls it off, what a joke, another armpit offside #MCISHU [Tweet]. Retrieved from https://twitter.com/AzumaNelson/status/1211353914177536000

Nevill, A., Balmer, N., & Williams, A. (2002). The Influence of Crowd Noise and Experience upon Refereeing Decisions in Football. *Psychology of Sport and Exercise, 3*(4), 261-272. doi:10.1016/S1469-0292(01)00033-4.

Niall C [@Nialler67]. (2019, December 1). Don't know how that was overturned. Penalty all day long #LEIEVE [Tweet]. Retrieved from https://twitter.com/Nialler67

Nick [@DoubleBluffNick]. (2020, February 16). #AVLTOT feels like it should be a double digit goal scoring half [Tweet]. Retrieved from https://twitter.com/DoubleBluffNick/status/1229054001578205184

Nielsen, F. Å. (2011, May). A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. *Paper presented at the ESWC2011 Workshop on 'Making Sense of Microposts': Big things come in small packages*. (pp. 93-98). Retrieved from http://arxiv.org/abs/1103.2903.

Nielsen, F. Å. (2017). *afinn project*. Preseneaeion. DTU Compute. Technical University of Denmark. Retrieved from http://www2.imm.dtu.dk/pubdb/views/edoc_download.php/6975/pdf/imm6975.pdf

Nlandu, T. (2012). The Fallacies of the Assumptions Behind the Arguments for Goal-Line Technology in Soccer. *Sport, Ethics and Philosophy, 6*(4), 451-466. doi:10.1080/17511321.2012.745896.

O'shea, M., & Alonso, A. D. (2011). Opportunity or obstacle? A preliminary study of professional sport organisations in the age of social media. *International Journal of Sport Management and Marketing, 10*(3-4), 196-212. doi:10.1504/ijsmm.2011.044790.

Okelana, O. [@Yincar]. (2019, December 5). See what they've been checking since. David Luiz was yards offside. Assistant Referee was asleep #ARSBHA [Tweet]. Retrieved from https://twitter.com/Yincar/status/1202703487248846871

Ott, B. L. (2017). The age of Twitter: Donald J. Trump and the politics of debasement. *Critical Studies in Media Communication, 34*(1), 59-68. doi:10.1080/15295036.2016.1266686.

Oudejans, R. R., Verheijen, R., Bakker, F. C., Gerrits, J. C., Steinbruckner, M., & Beek, P. J. (2000). Errors in judging 'offside' in football. *Nature, 404*(6773), 33. doi:10.1038/35003639.

Ovadia, S. (2009). Exploring the Potential of Twitter as a Research Tool. *Behavioral & Social Sciences Librarian, 28*(4), 202-205.

P. Adams [@Est1879]. (2020, January 19). Never a foul…Goalkeepers get protected way too much #LIVMUN [Tweet]. Retrieved from https://twitter.com/Est1879

Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval, 2*(1–2), 1-135. doi:10.1561/1500000011.

Panja, T. (2019, November 23). V.A.R. Check: In Premier League, Almost No One Is Happy. *The New York Times*. https://www.nytimes.com/2019/11/23/sports/premier-league-var-mike-riley.html

Panzarino, M. (2012, June 22). *Twitter launches special sports section for Euro 2012 Championship after NASCAR partnership*. The Next Web. https://thenextweb.com/twitter/2012/06/22/twitter-launches-second-special-sports-section-for-euro-2012-championship-after-nascar-partnership/%7B%7B%20linkUrl/%7B%7B%20linkUrl/

Papoose [@papoose502]. (2019, December 5). This club is a joke #ARSBHA [Tweet]. Retrieved from https://twitter.com/papoose502/status/1202707553630732288

Pavey, R. [@Rhi_Pavey]. (2020, February 22). VAR admitting that they made a mistake by not sending off Lo Celso tells you everything you need to know about VAR. Absolute joke #CHETOT [Tweet]. Retrieved from https://twitter.com/Rhi_Pavey/status/1231220065439244288

Pedersen, P. M. (2014). A Commentary on Social Media Research From the Perspective of a Sport Communication Journal Editor. *Communication & Sport, 2*(2), 138-142. doi:10.1177/2167479514527428.

Pegoraro, A. (2010). Look Who's Talking—Athletes on Twitter: A Case Study. *International Journal of Sport Communication, 3*(4), 501-514. doi:10.1123/ijsc.3.4.501.

Platt, O. (2019, November 2). *What is VAR? The video assistant referee system's Champions League & Premier League future*. Goal. https://www.goal.com/en/news/what-is-var-the-video-assistant-referee-systems-world-cup/19m696jq7onm618n3v9oqs02ab

Prasasti, N., & Ohwada, H. (2014, May 27-29). Applicability of machine-learning techniques in predicting customer defection. *Paper presented at the 2014 International Symposium on Technology Management and Emerging Technologies*. (pp. 157-162). doi:10.1109/ISTMET.2014.6936498.

Premier League. (2017, Aug 3). *What's new this season: Bans for diving*. Premier League. https://www.premierleague.com/news/444152

Premier League. (2018a, May 11). *Entertaining audiences*. Premier League. https://www.premierleague.com/news/686489

Premier League. (2018b). *Video Assistant Referees to be introduced in the next campaign after agreement at shareholders' meeting*. Premier League. https://www.premierleague.com/news/912537

Premier League. (2019, May 23). *Fans watch Premier League in record numbers*. Premier League. https://www.premierleague.com/news/1225124

Premier League. (2019-2020). *Premier League - Results*. Premier League. https://www.premierleague.com/results

Premier League [@premierleague]. (2020). [Twitter Profile]. Retrieved from https://twitter.com/premierleague

Prior, E. (2013, February 18). *Do Man Utd really have 659m supporters?* BBC News. https://www.bbc.com/news/magazine-21478857

R a C H [@Ra_hill2]. (2019, December 1). Never in a million years was that a pen! No contact and he got the ball #LEIEVE [Tweet]. Retrieved from https://twitter.com/Ra_hill2/status/1201185975537418240

Raney, A. A. (2006). Why we watch and enjoy mediated sports. . In J. B. Authur A. Raney (Ed.), *Handbook of sports and media* (pp. 313-329). New York: Lawrence Erlbaum Associates.

Recwand [RecWand]. (2019, November 23). Man city and Chelsea will have a great match. and Martin Atkinson will make it frustrating, so I'm excited to see how all this plays out. [Tweet]. Retrieved from https://twitter.com/RecWand

Reuters Staff. (2019, December 23). *Soccer-Mourinho bemoans Son ban and VAR use in Premier League*. Reuters. https://de.reuters.com/article/soccer-england-tot-brh-preview/soccer-mourinho-bemoans-son-ban-and-var-use-in-premier-league-idUKL4N28X352

Richard, T. A. [@raphael223]. (2020, January 19). Why do they give big games to such clowns to referee? Clown face #LIVMUN [Tweet]. Retrieved from https://twitter.com/raphael223/status/1218940215848587264

Riedl, D., Strauss, B., Heuer, A., & Rubner, O. (2015). Finale furioso: referee-biased injury times and their effects on home advantage in football. *Journal of Sports Sciences, 33*(4), 327-336. doi:10.1080/02640414.2014.944558.

Rodgers, S. (2014, July 14). Insights into the #WorldCup conversation on Twitter. Retrieved from https://blog.twitter.com/2014/insights-into-the-worldcup-conversation-on-twitter

Rodriguez, N. S. (2016). #FIFAputos: A Twitter Textual Analysis Over “Puto” at the 2014 World Cup. *Communication & Sport, 5*(6), 712-731. doi:10.1177/2167479516655429.

Roksie, U. [@Olakhalifa_]. (2020, March 7). That's purely onside #ARSWHU [Tweet]. Retrieved from https://twitter.com/Olakhalifa_/status/1236329838497206272

Rubin, A. M. (2009). *Uses-and-gratifications perspectives on media effects* (3rd ed.). New York, NY: Lawrence Erlbaum.

Sam [@Profdog123]. (2019, December 5). Are Arsenal really that bad now that a team like Brighton can come to the Emirates and dominate them all over the park… #ARSBHA. Retrieved from https://twitter.com/Profdog123/status/1202690498340827142

Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development, 3*(3), 210-229. doi:10.1147/rd.33.0210.

Sanderson, J. (2011). To Tweet or Not to Tweet: Exploring Division I Athletic Departments’ Social-Media Policies. *International Journal of Sport Communication, 4*(4), 492-513. doi:10.1123/ijsc.4.4.492.

Sanderson, J. (2014). What do we do with Twitter? *Communication & Sport, 2*(2), 127-131. doi:10.1177/2167479514527429.

Sankaranarayanan, J., Samet, H., Teitler, B. E., Lieberman, M. D., & Sperling, J. (2009). TwitterStand: news in tweets. *Paper presented at the Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, Seattle, Washington. (pp. 42–51). doi:10.1145/1653771.1653781.

Sebastiani, F. (2001). Machine Learning in Automated Text Categorization. *ACM Computing Surveys, 34*(1), 1-47. doi:10.1145/505282.505283.

Sen, K. [@KaranSen]. (2020, February 22). That is a crystal clear red for Lo Celdo… a career ending challenge. A ridiculous call. Shocking. Atrocious. I'm running out of words. #CHETOT [Tweet]. Retrieved from https://twitter.com/KaranSen

Shamma, D. A., Kennedy, L., & Churchill, E. F. (2009). Tweet the debates: understanding community annotation of uncollected sources. *Paper presented at the Proceedings of the first SIGMM workshop on Social media*, Beijing, China. (pp. 3–10). doi:10.1145/1631144.1631148.

Sharma, M. (2019, January 19). *Gentle Introduction of XGBoost Library*. Medium. https://medium.com/@imoisharma/gentle-introduction-of-xgboost-library-2b1ac2669680

Shearer, E., Barthel, M., Gottfried, J., & Mitchell, A. (2015, July 14). *The Evolving Role of News on Twitter and Facebook*. Pew Research Center. https://www.journalism.org/2015/07/14/the-evolving-role-of-news-on-twitter-and-facebook

Shepherd, P. [@paulshep1980]. (2019, November 23). Keeper should have saved that anyway #MCICHE [Tweet]. Retrieved from https://twitter.com/paulshep1980/status/1198320902783545346

Silge, J., & Robinson, D. (2017). *Text Mining with R: A Tidy Approach* (1 ed.). Sebastopol, CA: O'Reilly Media.

Silva, R. (2020, January 23). *VAR is a step too far —it's taking the joy out of watching football*. EveningStandard. https://www.standard.co.uk/comment/comment/var-watching-football-fans-joy-a4342536.html

Singh, H. (2013). In Search of a Level Playing Field: Rules, Morals and the Jurisprudence of Sport. *International Journal of Sport and Society, 3*(1), 81-90. doi:10.18848/2152-7857/CGP/v03i01/53896.

Singhal, V. [@vkerdos]. (2020, January 19). I like #Salah. So, I won't say anything. #LIVMUN #PL [Tweet]. Retrieved from https://twitter.com/vkerdos/status/1218961615782965249

Smith, L. [@sessionar08]. (2020, February 17). What a win. Could be a massive 3 points. Got a little bit of good luck along the way. #brunofernades is class. Can see how good he is already. Certainly can cross, pass and shoot #baller

#CHEMUN [Tweet]. Retrieved from https://twitter.com/sessionar08/status/1229525803559510021

Smith, L. R., & Smith, K. D. (2012). Identity in Twitter's Hashtag Culture: A Sport-Media-Consumption Case Study. *International Journal of Sport Communication, 5*(4), 539-557. doi:10.1123/ijsc.5.4.539.

Spitz, J., Moors, P., Wagemans, J., & Helsen, W. F. (2018). The impact of video speed on the decision-making process of sports officials. *Cognitive Research: Principles and Implications, 3*(1), 16. doi:10.1186/s41235-018-0105-8.

Sport, B. (2016, April 19). *Gianni Infantino: Video referees 'to be used at 2018 World Cup'*. BBC Sport. https://www.bbc.com/sport/football/36085289

Standard Sport. (2018, November 15). *VAR explained: How the Video Assistant Referee system works*. EveningStandard. https://www.standard.co.uk/sport/football/var-explained-how-the-video-assistant-referee-system-works-a3991351.html

Stinson, S. (2018, June 26). *'VAR is bulls—t': Video review's honeymoon phase comes to sudden, spectacular end at World Cup*. National Post. https://nationalpost.com/sports/soccer/var-is-bulls-t-video-reviews-honeymoon-phase-comes-to-sudden-spectacular-end-at-world-cup

Stokesie [@rfstokes44]. (2019, December 5). VAR comes into life to disallow an @Arsenal goal!! Where was VAR for our penalties in the first half??? #ARSBHA [Tweet]. Retrieved from https://twitter.com/rfstokes44/status/1202703673979260932

Stranger [@AkhY_Herpheez]. (2020, February 17). This is totally unfair, VAR has totally spoilt football tbh #CHEMUN [Tweet]. Retrieved from https://twitter.com/AkhY_Herpheez/status/1229520361454211077

Stranges, E. (2019). *10 Ways FIFA brought the World Cup to life on Twitter*. Twitter: Media. https://media.twitter.com/en_us/articles/case-study/2018/10-ways-fifa-brought-the-world-cup-to-life-on-twitter.html

Sully! [@irishdummy01]. (2020, February 16). What an incredible match, easily match of the season so far #AVLTOT @PLinUSA @premierleague

@NBCSportsSoccer@EPLBible [Tweet]. Retrieved from https://twitter.com/irishdummy01/status/1229072454007640067

Sutter, M., & Kocher, M. (2004). Favoritism of agents - The case of referees' home bias. *Journal of Economic Psychology, 25*(4), 461-469. doi:10.1016/S0167-4870(03)00013-8.

Sutton, W. A., Mcdonald, M. A., Milne, G. R., & Cimperman, J. (1997). Creating and Fostering Fan Identification in Professional Sports. *Sport Marketing Quarterly, 6*(1), 15-22.

Svantesson, D. J. B. (2014). Could technology resurrect the dignity of the FIFA World Cup refereeing? *Computer Law & Security Review, 30*(5), 569-573. doi:10.1016/j.clsr.2014.07.004.

Szeth-Son-Son-Vallano [@Stoajoe20]. (2019, December 5). COYG!!! We winning this. #ARSBHA [Tweet]. Retrieved from https://twitter.com/Stoajoe20/status/1202683640230109185

Szymanski, S. (2003). The Economic Design of Sporting Contests. *Journal of Economic Literature, 41*(4), 1137-1187. doi:10.1257/002205103771800004.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-Based Methods for Sentiment Analysis. *Computational Linguistics, 37*(2), 267-307. doi:10.1162/COLI_a_00049.

Tan, A. S. (1985). *Mass Communication Theories and Research*. New York, NY: Macmillan.

Tiago, T., Tiago, F., Faria, S. D., & Couto, J. P. (2016). Who is the better player? Off-field battle on Facebook and Twitter. *Business Horizons, 59*(2), 175-183. doi:10.1016/j.bushor.2015.11.003.

Tkay [@tkaydrifts]. (2020, February 22). VAR is trash. That was a straight red card. #CHETOT [Tweet]. Retrieved from https://twitter.com/tkaydrifts/status/1231212173684355072

Tumasjan, A., Sprenger, T. O., Sandner, P. G., & Welpe, I. M. (2011). Election Forecasts With Twitter: How 140 Characters Reflect the Political Landscape.

Social Science Computer Review, 29(4), 402–418. doi:10.1177/0894439310386557.

Twitter Inc. (2020, April 3). Coronavirus: Staying safe and informed on Twitter. Retrieved from https://blog.twitter.com/en_us/topics/company/2020/covid-19.html

Ugondo, P. I., & Tsokwa, M. (2019). Interpreting Video Assistant Referee and Goal-Line Technology Communication: The Pitch-Based Referees Perspectives. *International Journal of Trend in Scientific Research and Development, 3*(4), 1058-1062. doi:10.31142/ijtsrd23914.

Uhuru Jr [@EmanuelUhuru]. (2020, March 7). Goal or offside?? #ARSWHU [Tweet]. Retrieved from https://twitter.com/EmanuelUhuru/status/1236329969887916044

Unkelbach, C., & Memmert, D. (2010). Crowd Noise as a Cue in Referee Decisions Contributes to the Home Advantage. *Journal of Sport & Exercise Psychology, 32*(4), 483-498. doi:10.1123/jsep.32.4.483.

Walker, C. [@ClintHWalker]. (2020, January 18). You. Are. Kidding. Me? #MCICRY [Tweet]. Retrieved from https://twitter.com/ClintHWalker/status/1218558789563056128

Wang, X. (2013). Applying the integrative model of behavioral prediction and attitude functions in the context of social media use while viewing mediated sports. *Computers in Human Behavior, 29*(4), 1538-1545. doi:10.1016/j.chb.2013.01.031.

Wang, X. (2015). Using attitude functions, self-efficacy, and norms to predict attitudes and intentions to use mobile devices to access social media during sporting event attendance. *Mobile Media & Communication, 3*(1), 75-90. doi:10.1177/2050157914548932.

Wang, Z., Sun, X., Zhang, D., & Li, X. (2006, August 13-16). An Optimal SVM-Based Text Classification Algorithm. *Paper presented at the 5th IEEE International Conference on Machine Learning and Cybernetics*, Dalian, China. (pp. 1378-1381). doi:10.1109/ICMLC.2006.258708.

Wang, Z., Tong, V., Ruan, P., & Li, F. (2016, December). Lexicon Knowledge Extraction with Sentiment Polarity Computation. *Paper presented at the IEEE 16th International Conference on Data Mining Workshops*, Barcelona, Spain. (pp. 978-983). doi:10.1109/ICDMW.2016.0142.

Weinburg, T. (2009). *The New Community Rules: Marketing on the Social Web* (1 ed.). Sebastopol, CA: O'Reilly Media.

Wertheim, J. (2011, July 14). *Twitter is now a permanent part of the sports firmament*. Sports Illustrated. https://www.si.com/more-sports/2011/07/14/twitter-thinkpiece

White, M. (2020, February 19). *Why Arsene Wenger's new plan for a 'daylight' offside rule to fix VAR wouldn't work*. FourFourTwo. https://www.fourfourtwo.com/features/offside-rule-law-var-new-reform-arsene-wenger-daylight-proposal-problems

Wickham, H., François, R., Henry, L., & Müller, K. (2020). Package 'dplyr'. [RDocumentation]. Retrieved from https://cran.r-project.org/web/packages/dplyr/index.html

Williams, J., Chinn, S. J., & Suleiman, J. (2014). The value of Twitter for sports fans. *Journal of Direct, Data and Digital Marketing Practice, 16*(1), 36-50. doi:10.1057/dddmp.2014.36.

Wilson, L. J. [@LukeWilson511]. (2020, February 17). What a joke. Half his toe is offside. Explain to me how Giroud gains an advantage for a headed goal by his toe being offside by an inch. Ridiculous. I've had enough with VAR now. It takes away too much, and gives nothing back. #CFC #CHEMUN [Tweet]. Retrieved from https://twitter.com/LukeWilson511/status/1229521028562464768

Wilson, T., Wiebe, J., & Hwa, R. (2004). Just how mad are you? finding strong and weak opinion clauses. *Paper presented at the Proceedings of the 19th national conference on Artifical intelligence*, San Jose, California. (pp. 761–767).

Winand, M., & Fergusson, C. (2018). More decision-aid technology in sport? An analysis of football supporters' perceptions on goal-line technology. *Soccer & Society, 19*(7), 966-985. doi:10.1080/14660970.2016.1267629.

Winter, S., Krämer, N. C., Benninghoff, B., & Gallus, C. (2018). Shared Entertainment, Shared Opinions: The Influence of Social TV Comments on the Evaluation of Talent Shows. *Journal of Broadcasting & Electronic Media, 62*(1), 21-37. doi:10.1080/08838151.2017.1402903.

Witkemper, C., Blaszka, M., & Chung, J. (2016). Establishing a Typology of Social Media Uses in the Sport Industry:A Multidimensional Scaling Study. *Communication & Sport, 4*(2), 166-186. doi:10.1177/2167479514544951.

Woodford, D. (2013, November 27). *New Users in Times of Crisis (and Revolution)*. Mapping Online Publics. https://mappingonlinepublics.net/2013/11/27/new-users-in-times-of-crisis-and-revolution/

Yiu, T. (2019, June 12). *Understanding Random Forest: How the Algorithm Works and Why it Is So Effective*. Medium. https://towardsdatascience.com/understanding-random-forest-58381e0602d2

Yu, Y., & Wang, X. (2015). World Cup 2014 in the Twitter World: A big data analysis of sentiments in U.S. sports fans' tweets. *Computers in Human Behavior, 48*, 392-400. doi:10.1016/j.chb.2015.01.075.

Zafarani, R., Abbasi, M. A., & Liu, H. (2014). *Social Media Mining: An introduction*: Cambridge University Press.

Zhao, S., Zhong, L., Wickramasuriya, J., & Vasudevan, V. (2011a, June 29). Analyzing Twitter for Social TV : Sentiment Extraction for Sports. *Paper presented at the 2nd International Workshop on Future of Television*, Lisbon, Portugal. (pp. 8). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.414.6401&rep=rep1&type=pdf.

Zhao, S., Zhong, L., Wickramasuriya, J., & Vasudevan, V. (2011b). *Human as Real-Time Sensors of Social and Physical Events: A Case Study of Twitter and Sports Games*. Retrieved from https://arxiv.org/pdf/1106.4300.pdf

Zαcκ [@Zack5cott]. (2020, February 22). Tired of not celebrating goals properly. VAR should not exist. #CHETOT [Tweet]. Retrieved from https://twitter.com/Zack5cott/status/1231198847290019841

Appendix

Appendix A – Twitter API Data Extraction Code

```
library(tidytext)
library(rjson)
library(dplyr)
library(ggplot2)
library(purrr)
library(stopwords)
library(lsa)
library(reshape2)
library(magrittr)
install.packages("rtweet")
library(rtweet)

general_stopwords<-c("premier", "league","premierleague","pl","training","english", "epl",
"englishpremierleague", "fc", "footballclub",
"season","game","club","football","match","it's","it’s","men's","men","men'","vs","v","boys",
"weekend","lads", "soccer","english",0:10, 2019)
arsenal_stopwords<-c("arsenal", "ars", "afc", "gunners", "arsenalfc")
astonvilla_stopwords<-c("aston", "villa", "avfc", "astonvilla", "astonvillafc", "villpark",
"astonvillafootballclub", "avl", "bella", "lion")
bournemouth_stopwords<-c("bournemouth", "afc", "afcb", "bou", "cherries", "boscombe",
"athletic")
brighton_stopwords<-c("brighton", "hove", "albion", "bhafc", "bha", "gully", "seagull",
"seagulls")
burnley_stopwords<-c("burnley", "bur", "burnleyfc", "bfc", "bee", "clarets")
chelsea_stopwords<-c("chelsea", "chelseafc", "che", "cfc", "stamford", "lion", "blues",
"pensioners")
crystalpalace_stopwords<-c("crystal", "palace", "crystalpalace", "cry", "eagles", "glaziers",
"cpfc")
everton_stopwords<-c("everton", "eve", "blues", "toffees", "evertonfootballclub", "efc",
"evertonfc")
leicestercity_stopwords<-c("leicster", "city", "lei", "foxes", "leicsterfootballclub",
"leicstercityfc", "lcfc")
liverpool_stopwords<-c("liverpool", "liv", "lfc", "reds", "weareliverpool", "liverpoolfc",
"upthereds")
mancity_stopwords<-c("man", "city", "manchester", "mci", "manchestercity", "citizens", "sky",
"blues", "skyblues", "mancity", "mcfc")
manutd_stopwords<-c("man", "utd", "manchester", "united", "manutd", "mufc", "mun",
"reddevils", "red", "devils")
newcastle_stopwords<-c("new", "castle", "magpies", "united", "utd", "newcastle", "nufc",
"newcastleunited")
norwich_stopwords<-c("norwich", "nor", "canaries", "yellows", "ncfc")
sheffieldutd_stopwords<-c("sheffield", "utd", "united", "blades", "sufc", "shu",
"sheffieldunited")
southampton_stopwords<-c("south", "ampton", "sou", "saints", "saintsfc", "southampton",
"southamptonfc")
tottenham_stopwords<-c("tot", "spurs", "tottenham", "hotspur", "tottenhamhotspurs",
"tottenhamspurs", "thfc", "coys")
watford_stopwords<-c("watford", "wat", "hornets", "watfordfc", "wtc")
westham_stopwords<-c("westham", "whu", "west", "ham", "united", "irons", "hammers",
"academy", "coyi")
wolves_stopwords<-c("wolves", "wolverhampton", "wanderers", "wol", "wolvesacademy", "wwfc")

##Information you need to update for each game
start_date <- as.POSIXct('2019-12-05 21:15:00')
end_date <- as.POSIXct('2019-12-05 23:15:00')
gameHashTag <- "#ARSBHA"
startHalfTime <- 47
endHalfTime <- 62
endGame <- 112
ThisGame_stopwords<-bind_rows(data.frame(word = c(general_stopwords, arsenal_stopwords,
brighton_stopwords)))
##At end change VARSentimentDataFrame to Game Specific

Twitter.df<-search_tweets(q=gameHashTag,include_rts = FALSE, lang = "en", n=18000)%>%
 select(text,favorite_count,retweet_count,screen_name,created_at,location)%>%
  mutate(created_at = as.POSIXct(created_at, format = "%a %b %d %H:%M:%S +0000 %Y"))%>%
```

```
  filter(created_at >= start_date & created_at <= end_date)
Twitter.df$text<-gsub("http[^[:space:]]*", "", Twitter.df$text)
Twitter.df$text<-gsub("@\\S+","", Twitter.df$text)
Twitter.df$text<-gsub(gameHashTag,"", Twitter.df$text)
Twitter.df$text<-tolower(Twitter.df$text)
Twitter.df[,7]<-1:nrow(Twitter.df)
colnames(Twitter.df)[7]<-"NewID"
Twitter.df[,8]<-as.numeric(round(difftime(Twitter.df$created_at,start_date)/60,2))
Twitter.df[,8]<-ceiling(Twitter.df[,8])
colnames(Twitter.df)[8]<-"MinuteSinceKickOff"
#sort by time so that later can get the correct row number for time
Twitter.df<-Twitter.df[order(Twitter.df$MinuteSinceKickOff),]
wordsTwitter<-Twitter.df %>%
  dplyr::select(NewID,text,created_at) %>%
  unnest_tokens(word, text)
#Remove Stopwords
stop_english <- data_frame(word = stopwords_en)
wordsTwitter<-wordsTwitter %>%
  anti_join(stop_english)
#remove other hashtags and team stopwords
wordsTwitter<-wordsTwitter %>%
  anti_join(ThisGame_stopwords)
#Stemming
wordsTwitter <- wordsTwitter %>%
  mutate(word=wordStem(.$word,language = "en"))
wordsPLGame_i <- wordsTwitter
#Convert Time ##Problem here with start.time<-wordsPLGame_i$created_at[1], due to lag in
program so just use start time recorded from internet
start.time <-start_date
wordsPLGame_i[,4]<-round(difftime(wordsPLGame_i$created_at,start.time)/60,2)
colnames(wordsPLGame_i)[4]<-"MinuteSinceKickOff"
```

Appendix B – Lexicon-Based Sentiment Analysis and VAR Count Frequency Code

```
#Words per Minute Graph for VAR, need to add 1 for the 0th minute
VARproMinute<-data.frame(nrow =
ceiling(wordsPLGame_i$MinuteSinceKickOff[nrow(wordsPLGame_i)])+1, ncol=2)
colnames(VARproMinute)[1]<-'Minute'
colnames(VARproMinute)[2]<-'VARCount'
for (i in 0:ceiling(wordsPLGame_i$MinuteSinceKickOff[nrow(wordsPLGame_i)])+1){
  VARproMinute[i,1]<-i-1
  VARproMinute[i,2]<-sum(wordsPLGame_i$word=="var" &
ceiling(wordsPLGame_i$MinuteSinceKickOff)==i-1)
}
VARproMinute$Minute<-as.numeric(VARproMinute$Minute)
ggplot(data=VARproMinute, aes(x=Minute,y=VARCount))+ geom_line()
#Sentiment Calculation
PLGame_i_sentiment <- wordsPLGame_i %>%
  inner_join(get_sentiments("afinn"))
PLGame_i_sentiment[,6]<-ceiling(PLGame_i_sentiment$MinuteSinceKickOff)
colnames(PLGame_i_sentiment)[6]<-"MinGroup"
senproMin_everyMin<-data.frame(nrow =
ceiling(wordsPLGame_i$MinuteSinceKickOff[nrow(wordsPLGame_i)])+1, ncol=2)
senproMin_PLGame_i<-PLGame_i_sentiment %>%
  group_by(MinGroup) %>%
  summarize(meanSent = mean(value))
#fill in the meanSent right away, using j to control the row number of the shorter data.frame
j=1
for (i in 0:(as.integer(ceiling(wordsPLGame_i$MinuteSinceKickOff[nrow(wordsPLGame_i)])))+1){
  senproMin_everyMin[i,1]<-i-1
  senproMin_everyMin[i,2]<-ifelse(as.integer(senproMin_PLGame_i$MinGroup[j])==(i-
1),senproMin_PLGame_i$meanSent[j],NA)
  j=ifelse(is.na(senproMin_everyMin[i,2]),j,j+1)
}
colnames(senproMin_everyMin)[1]<-'Minute'
colnames(senproMin_everyMin)[2]<-'meanSent'

#Draw Graph
VARandSentiment<-cbind(VARproMinute,senproMin_everyMin)
VARandSentiment[,5]<-"Break"
VARandSentiment[1,5]<-0
VARandSentiment[2:46,5]<-c(1:45)
VARandSentiment[62:113,5]<-c(46:97)
VARandSentiment[114:122,5]<-"End"
colnames(VARandSentiment)[5]<-"GameMinute"
#Make Plot
ggplot(VARandSentiment, aes(x = Minute))+
  geom_line(aes(y = VARCount, color = "VARCount"))+
  geom_line(aes(y = meanSent, color = "meanSent"))+
  xlab("Time")+
  scale_y_continuous(sec.axis = sec_axis(~.*.2, name = "SentimentScore"))+
  theme(legend.position = c(0.2, 0.1))+
  geom_vline(xintercept = 1)+
  geom_vline(xintercept = startHalfTime, linetype="dashed")+
  geom_vline(xintercept = endHalfTime, linetype="dashed")+
  geom_vline(xintercept = endGame)+
  geom_text(x=11, y=2.5, label="1st half")+
  geom_text(x=72, y=2.5, label="2nd half")
```

Appendix C – Naïve Bayes Classifier Code

```
nb_optimalization<-function(TrainingData){
  #Returns Confusion Matrix, Method Naive Bayes Classifier with stemming, and punctuation,
lower case, and numbers removed
  #Data preparation with quanteda
  #TrainingData and TestData as Random Samples of coded data with ID
  #Cleaning occurs here with DTM-Settings
  TrainingDataCorpus<-corpus(TrainingData)
  set.seed(300)
  id_train <- sample(1:nrow(TrainingData), 1000, replace = FALSE)
  docvars(TrainingDataCorpus, "id_numeric") <- 1:ndoc(TrainingDataCorpus)
  dtmTraining<-corpus_subset(TrainingDataCorpus,id_numeric %in% id_train) %>%
    dfm(stem=TRUE,remove_punct=TRUE, tolower=TRUE, remove_numbers=TRUE)
  dtmTest<-corpus_subset(TrainingDataCorpus,!id_numeric %in% id_train) %>%
    dfm(stem=TRUE,remove_punct=TRUE,tolower=TRUE, remove_numbers=TRUE)

  #Model creation (quanteda) und checking. dfm_match (quanteda) necessary to match variables
for both DTMs
  testmodel<-textmodel_nb(dtmTraining,docvars(dtmTraining,"Coding"))
  dfmat_matched <- dfm_match(dtmTest, features = featnames(dtmTraining))
  actual_class <- docvars(dfmat_matched, "Coding")
  predicted_class <- predict(testmodel, newdata = dfmat_matched)
  tab_class <- table(actual_class, predicted_class)
  return(confusionMatrix(tab_class,mode = "everything"))
}
```

Appendix D – Support Vector Machine Code

```
svm_optimalization <- function(TrainingData) {
  #Returns Confusion Matrix, Method is Support Vector Machine with punctuation removed, and
stemming
  #Text preparation with tm.
  #After Preprocessing need a data frame in DTM format and Raw data,
  #where Coding as.factor must be used.
  TrainingData$Coding <- as.factor(TrainingData$Coding)
  TrainingDataCorpus <- Corpus(VectorSource(TrainingData$text))
  TrainingDataCorpus <-
    tm_map(TrainingDataCorpus, content_transformer(tolower))
  TrainingDataCorpus <-
    tm_map(TrainingDataCorpus, removePunctuation)
  TrainingDataCorpus <- tm_map(TrainingDataCorpus, stemDocument)
  dtmTraining <- DocumentTermMatrix(TrainingDataCorpus)
  dtmTraining.df <- as.data.frame(as.matrix(dtmTraining))
  colnames(dtmTraining.df) <-
    make.names(colnames(dtmTraining.df), unique = TRUE)
  TrainingData <-
    cbind(Coding = TrainingData$Coding, dtmTraining.df)

  #Extract the Test- and Trainings-Data in this order
  set.seed(300)
  id_train <- sample(1:nrow(TrainingData), 1000, replace = FALSE)
  TestData <- TrainingData[-id_train,]
  TrainingData <- TrainingData[id_train,]

  #Model Creation und Checking, confusion matrix as result
  testmodel <-
    train(Coding ~ ., data = TrainingData, method = 'svmLinear3')
  svm_predict <- predict(testmodel, na.omit(TestData))
  actual_class <- TestData$Coding
  tab_class <- table(actual_class, svm_predict)
  return(confusionMatrix(tab_class, mode = "everything"))
}
```

Appendix E – Random Forest Code

```
RandomForest_optimalization <- function(TrainingData) {
  #Returns Confusion Matrix, Method Random Forest with punctuation, and numbers removed, and
words stemmed
  #Text preparation with tm
  #After Preprocessing need a data frame in DTM format and Raw data
  #where Coding as.factor must be used. This is the last step in this block of code, lines
before that all text preparation
  TrainingData$Coding <- as.factor(TrainingData$Coding)
  TrainingDataCorpus <- Corpus(VectorSource(TrainingData$text))
  TrainingDataCorpus <-
    tm_map(TrainingDataCorpus, content_transformer(tolower))
  TrainingDataCorpus <- tm_map(TrainingDataCorpus, removePunctuation)
  TrainingDataCorpus <- tm_map(TrainingDataCorpus, stemDocument)
  TrainingDataCorpus <- tm_map(TrainingDataCorpus, removeNumbers)
  dtmTraining <- DocumentTermMatrix(TrainingDataCorpus)
  dtmTraining.df <- as.data.frame(as.matrix(dtmTraining))
  colnames(dtmTraining.df) <-
    make.names(colnames(dtmTraining.df), unique = TRUE)
  TrainingData <- cbind(Coding = TrainingData$Coding, dtmTraining.df)

  #Extract the Test- and Trainings-Data in this order
  set.seed(300)
  id_train <- sample(1:nrow(TrainingData), 1000, replace = FALSE)
  TestData <- TrainingData[-id_train, ]
  TrainingData <- TrainingData[id_train, ]

  #Model Creation und Checking, confusion matrix as result. n tree is controlled
  testmodel <- randomForest(Coding~.,data = TrainingData, ntree = 20)
  predicted_class<-predict(testmodel, newdata=TestData)
  actual_class<-TestData$Coding
  tab_class <- table(actual_class, predicted_class)
  return(confusionMatrix(tab_class,mode = "everything"))
}
```

Appendix F – Optimized Random Forest with xgboost Code

```
xgBoost_optimalization <- function(TrainingData) {
  # Returns Confusion Matrix, Method xgboost based on Random Forest with customized stopwords
removed but not set english stopwords
  # Text preparation with tm
  # After Preprocessing need a data frame in DTM as Matrix
  TrainingDataCorpus <- Corpus(VectorSource(TrainingData$text))
  TrainingDataCorpus <-
    tm_map(TrainingDataCorpus, content_transformer(tolower))
  TrainingDataCorpus <- tm_map(TrainingDataCorpus, removePunctuation)
  TrainingDataCorpus<-tm_map(TrainingDataCorpus,removeWords, stopwords_complete)
  TrainingDataCorpus <- tm_map(TrainingDataCorpus, stemDocument)
  TrainingDataCorpus<-tm_map(TrainingDataCorpus,removeWords, stopwords_complete)
  TrainingDataCorpus <- tm_map(TrainingDataCorpus, removeNumbers)
  dtmTraining <- DocumentTermMatrix(TrainingDataCorpus)
  dtmTraining.df <- as.matrix(dtmTraining)


  # Extract the Test- and Trainings-Data in this order
  set.seed(123)
  id_train <- sample(1:nrow(TrainingData), round(0.75*nrow(TrainingData),0), replace = FALSE)
  TrainingData_xgb <- dtmTraining.df[id_train, ]
  TestData_xgb <- dtmTraining.df[-id_train, ]
  TrainingValues <- TrainingData$Sentiment[id_train]+1

  # Model Creation und Checking
  # Controlled Parameter partially under tuning, partially in the Training-Formula.
  # For Multiclass_Classification should be objective = "multi:softprob" or "multi:softmax",
  # as well num_class = Count of possible results
  # (For binary classification objective = "binary:logistic" and num_class not set)
  # Tuning variables for both models found in Table 10
  tuning <- list(
    eta = .20,
    max_depth = 30,
    min_child_weight = 1,
    subsample = 0.90,
    colsample_bytree = 0.05
  )
  testmodel <-
    xgboost(
      data = TrainingData_xgb,
      params = tuning,
      nrounds = 145,
      label = TrainingValues,
      #early_stopping_rounds = 20,
      verbose = 0,
      nfold = 5,
      objective = "multi:softprob",
      num_class = 3
    )

  pred <- predict(testmodel, TestData_xgb, type = "response")
  # Validation of this class chosen based on highest probability
  # Happens here with a Matrix
  predicted_class <- matrix(pred,
                             nrow = 3,
                             ncol = length(pred) / 3) %>%
    t() %>%
    data.frame() %>%
    mutate(label = TrainingData$Sentiment[-id_train] + 2,
           max_prob = max.col(., "last"))
  tab_class <- table(predicted_class$label,predicted_class$max_prob)
  return(confusionMatrix(tab_class, mode = "everything"))
}
```

Appendix G – Functions Used for the Machine Learning Method

```
#All functions involved in the machine learning process

# install.packages("xgboost")
# install.packages("tm")
# install.packages("gbm")
# install.packages("caret")
# install.packages("tidytext")
# install.packages("randomForest")

library(tm)
library(dplyr)
library(xgboost)
library(gbm)
library(caret)
library(randomForest)
library(tidytext)
library(stringr)
library(ggplot2)
options(stringsAsFactors = FALSE)

read_emojis<-function(Tweets.df, ascii_TRUE = TRUE){
  #Converts emojis and unicode into "*emoji_chr* " style
  #needs Tweets.df and information (TRUE, FALSE) if text has been transformed to ascii
already
  #needs packages stringr, dplyr
  #Step 1: transformation to ascii and manual transformation for non emojis
  if (ascii_TRUE==FALSE){
    Tweets.df$text<-iconv(Tweets.df$text, from = "latin1", to = "ascii", sub = "byte")
  }
  Tweets.df$text<-gsub("<e2><80><99>","'",Tweets.df$text)
  #would be Pound sign
  Tweets.df$text<-gsub("<c2><a3>","dollar ",Tweets.df$text)
  #would be Euro sign
  Tweets.df$text<-gsub("<e2><82><ac>","euro ",Tweets.df$text)
  Tweets.df$text<-gsub("<e2><80><bd>","*interrobang* ",Tweets.df$text)
  Tweets.df$text<-gsub("<c2><a0>","",Tweets.df$text)
  Tweets.df$text<-gsub("<e2><80><98>","'",Tweets.df$text)
  #Random han sign...
  Tweets.df$text<-gsub("<e4><b9><81>","'",Tweets.df$text)
  Tweets.df$text<-gsub("<c2><b4>","'",Tweets.df$text)
  Tweets.df$text<-gsub("<e2><80><a6>","... ",Tweets.df$text)
  Tweets.df$text<-gsub("<cb><b3>",". ",Tweets.df$text)
  #Would be accent aigu (acute accent) over e
  Tweets.df$text<-gsub("<c3><a8>","e",Tweets.df$text)
  #Would be accent grave over e
  Tweets.df$text<-gsub("<c3><a9>","e",Tweets.df$text)
  #Would be accent aigu (acute accent) over u
  Tweets.df$text<-gsub("<c3><ba>","u",Tweets.df$text)
  #Would be accent circumflex over e
  Tweets.df$text<-gsub("<c3><aa>","e",Tweets.df$text)
  Tweets.df$text<-gsub("<e2><80><93>","-",Tweets.df$text)
  #Next two Would be single left and right-pointing quotation marks
  Tweets.df$text<-gsub("<e2><80><b9>","",Tweets.df$text)
  Tweets.df$text<-gsub("<e2><80><ba>","",Tweets.df$text)
  #Next two would be left and right-pointing double angle quotation marks
  Tweets.df$text<-gsub("<c2><ab>","",Tweets.df$text)
  Tweets.df$text<-gsub("<c2><bb>","",Tweets.df$text)
  #Next two Would be normal quotation marks
  Tweets.df$text<-gsub("<e2><80><9c>","",Tweets.df$text)
  Tweets.df$text<-gsub("<e2><80><9d>","",Tweets.df$text)
  Tweets.df$text<-gsub("<e2><80><94>","*em_dash* ",Tweets.df$text)
  #Would be cedilla under previous letter
  Tweets.df$text<-gsub("<d9><8d>","",Tweets.df$text)
  #Would be n with tilde
  Tweets.df$text<-gsub("<c3><b1>","n",Tweets.df$text)
  #would be ü
  Tweets.df$text<-gsub("<c3><bc>","u",Tweets.df$text)
  #would be ö
```

```
  Tweets.df$text<-gsub("<c3><b6>","u",Tweets.df$text)
  #would be g with breve
  Tweets.df$text<-gsub("<c4><9f>","g",Tweets.df$text)
  #Would be combining double over line over previous letter
  Tweets.df$text<-gsub("<cc><bf>","",Tweets.df$text)
  #Would be funny i point
  Tweets.df$text<-gsub("<d9><8f>","",Tweets.df$text)
  #Would be line over and under space
  Tweets.df$text<-gsub("<d9><8e>","",Tweets.df$text)
  #Looks like a casual q
  Tweets.df$text<-gsub("<c5><9f>","q",Tweets.df$text)
  #Would be crossed out l
  Tweets.df$text<-gsub("<c5><82>","l",Tweets.df$text)
  #Would be i with two points
  Tweets.df$text<-gsub("<c3><af>","i",Tweets.df$text)
  #Would be fish-like looking thing over c
  Tweets.df$text<-gsub("<d9><8c>","",Tweets.df$text)
  #Would be macron under next letter
  Tweets.df$text<-gsub("<d9><90>","",Tweets.df$text)
  #Would be double accent aigu over previous letter
  Tweets.df$text<-gsub("<d9><8b>","",Tweets.df$text)
  #Would make something completely weird out of the following letter
  Tweets.df$text<-gsub("<d9><8b><d9><8c><d9><8d><d9><8e><d9><8f><d9><90>","",Tweets.df$text)
  #Would make something completely weird out of the following letter
  Tweets.df$text<-gsub("<d9><8c><d9><8e><d9><8f><d9><90>","",Tweets.df$text)
  Tweets.df$text<-gsub("<e2><80><bb>","*reference_mark* ",Tweets.df$text)
  #Would be a inverted exclamation mark
  Tweets.df$text<-gsub("<c2><a1>","*inverted_!* ",Tweets.df$text)
  #Would be a with circle over it
  Tweets.df$text<-gsub("<c3><a5>","a",Tweets.df$text)
  #Fancy smiles that are not emojis.
  #Requires flipping backslashes first. Needs fixed to override "conflicting arguments" and
double bs for single bs
  Tweets.df$text<-gsub("\\","/",Tweets.df$text,fixed=TRUE)
  Tweets.df$text<-gsub("<c2><af>/_(<e3><83><84>)_/<c2><af>","*Smiley* ",Tweets.df$text,
fixed=TRUE)
  emDict_raw<-dictionary
  for (i in 1:nrow(Tweets.df)){
    if (str_count(Tweets.df$text[i],"<")>0){
      for (j in 1:nrow(emDict_raw)){
        Tweets.df$text[i]<-
gsub(emDict_raw$r_encoding[j],paste("*",emDict_raw$description[j],"* ",sep =
""),Tweets.df$text[i])
      }
    }
  }
  #remove emoji_style (varaiation selection-16) code. remove text_style (vs-15) code
  Tweets.df$text<-gsub("<ef><b8><8f>","",Tweets.df$text)
  Tweets.df$text<-gsub("<ef><b8><8e>","",Tweets.df$text)
  return(Tweets.df)
}


#Create and clean corpus
#clean corpus function
clean_text <- function(PredData) {

  TestDataCorpus<-Corpus(VectorSource(PredData$text)) # might need = sign here instead
  TestDataCorpus<-tm_map(TestDataCorpus, content_transformer(tolower))
  TestDataCorpus<-tm_map(TestDataCorpus, removePunctuation)
  TestDataCorpus<-tm_map(TestDataCorpus, removeWords, stopwords_complete)
  TestDataCorpus<-tm_map(TestDataCorpus, stemDocument)
  TestDataCorpus<-tm_map(TestDataCorpus, removeWords, stopwords_complete)
  TestDataCorpus<-tm_map(TestDataCorpus, removeNumbers)
  TestDataCorpus<-tm_map(TestDataCorpus, removeWords, stopwords("english"))
  return(TestDataCorpus)
}


#create dtm as dataframe
dtm <- function(TestDataCorpus){
  dtmPrediction<-DocumentTermMatrix(TestDataCorpus)
  dtmPrediction.df<-as.data.frame(as.matrix(dtmPrediction)) #turn into dataframe
  return (dtmPrediction.df)
```

```
}

#new dtm needs to match dtmTraining.df
#same columns and same order for both data

matrix_columns_same <- function(dtmTraining.df,dtmPrediction.df){
  #dtmPredictionOrdered.df<-matrix(0, nrow = nrow(dtmPrediction.df), ncol =
ncol(dtmTraining.df))
  #rownames(dtmPredictionOrdered.df) <- rownames(dtmPrediction.df)
  #colnames(dtmPredictionOrdered.df) <- colnames(dtmTraining.df)
  result_matrix<-matrix(0, nrow = nrow(dtmPrediction.df), ncol = ncol(dtmTraining.df))
  rownames(result_matrix) <- rownames(dtmPrediction.df)
  colnames(result_matrix) <- colnames(dtmTraining.df)
  for (row in 1:nrow(dtmPrediction.df)) {
    for(col in 1:ncol(dtmPrediction.df)){
      skip_to_next <- FALSE
      tryCatch(result_matrix[,colnames(dtmPrediction.df[col])], error = function(e)
{skip_to_next <<- TRUE
      })
      if(skip_to_next) {next()}
      else {
        if(dtmPrediction.df[row,col] > 0)
        {
          result_matrix[row,colnames(dtmPrediction.df)[col]]<-dtmPrediction.df[row,col]
        }
      }
    }
  }
  return(result_matrix)
}

predictSentimentData <- function(dtmPredictionOrdered)
{
  predGame_sent <- predict(Model_Sentiment, newdata=dtmPredictionOrdered, type = "response")

  predicted_class_sentiment <- matrix(predGame_sent,
                                      nrow = 3,
                                      ncol = length(predGame_sent) / 3) %>%
    t() %>%
    data.frame() %>%
    mutate(max_prob = max.col(., "last")-2) #subtract 2 here

  return(predicted_class_sentiment)
}

predictVARData <- function(dtmPredictionOrdered)
{
  predGame_var <- predict(Model_VAR, newdata=dtmPredictionOrdered, type = "response")

  predicted_class_var <- matrix(predGame_var,
                                nrow = 1,
                                ncol = length(predGame_var) / 1) %>%
    t() %>%
    data.frame()
  predicted_class_var$noVar <- 1 - predicted_class_var$.
  predicted_class_var$max_prob <- abs(max.col(predicted_class_var, "last")-2)

  return(predicted_class_var)
}

combineGameData<-function(PredData, startHalfTime, endGame)
{
  #VAR per Minute
  VARperMinute<-data.frame(nrow = ceiling(PredData$MinuteSinceKickOff[nrow(PredData)])+1,
ncol=2)
  colnames(VARperMinute)[1]<-'MinuteSinceKickOff'
  colnames(VARperMinute)[2]<-'VARCount'
  for (i in 0:ceiling(PredData$MinuteSinceKickOff[nrow(PredData)])+1){
    VARperMinute[i,1]<-i-1
    VARperMinute[i,2]<-sum(PredData$VAR & ceiling(PredData$MinuteSinceKickOff)==i-1)
  }
  VARperMinute$MinuteSinceKickOff<-as.numeric(VARperMinute$MinuteSinceKickOff)
  #Sentiment per Minute
```

```
  senPerMin_everyMin<-data.frame(nrow =
ceiling(PredData$MinuteSinceKickOff[nrow(PredData)])+1, ncol=2)
  senPerMin<-PredData %>%
    group_by(MinuteSinceKickOff) %>%
    summarize(meanSent = mean(Sentiment))

  j=1
  for (i in 0:(as.integer(ceiling(PredData$MinuteSinceKickOff[nrow(PredData)])))+1){
    senPerMin_everyMin[i,1]<-i-1
    senPerMin_everyMin[i,2]<-ifelse(as.integer(senPerMin$MinuteSinceKickOff[j])==(i-
1),senPerMin$meanSent[j],NA)
    j=ifelse(is.na(senPerMin_everyMin[i,2]),j,j+1)
  }
  colnames(senPerMin_everyMin)[1]<-'MinuteSinceKickOff'
  colnames(senPerMin_everyMin)[2]<-'meanSent'

  VARandSentiment<-cbind(VARperMinute,senPerMin_everyMin$meanSent)
  colnames(VARandSentiment)[3]<-'meanSent'

  #Add game minute column with correct stoppage time
  endHalfTime <- startHalfTime + 15
  stoppageTime1stHalf <- startHalfTime-45
  stoppageTime2ndHalf <- endGame-90 -(startHalfTime-45)
  VARandSentiment[,4]<-"Break"
  VARandSentiment[1,4]<-0
  VARandSentiment[2:(startHalfTime+1),4]<-c(1:startHalfTime)
  VARandSentiment[(endHalfTime+1):(endGame+1),4]<-c(45:(endGame-15-stoppageTime1stHalf))
  VARandSentiment[(endGame+2):122,4]<-"End"
  colnames(VARandSentiment)[4]<-"GameMinute"

  return(VARandSentiment)
}
#completed predicted data

# draw graph
drawGraph<-function(VARandSentiment, startHalfTime, endGame)
{
  endHalfTime <- startHalfTime + 15
  #maxVAR<-max(VARandSentiment$VARCount, na.rm = TRUE)
  #yScaling <- -1 + maxVAR
  ggplot(VARandSentiment, aes(x = MinuteSinceKickOff))+
    geom_line(aes(y = VARCount, color = "VARCount"))+
    geom_line(aes(y = meanSent, color = "meanSent"))+
    xlab("Time")+
    scale_y_continuous(sec.axis = sec_axis(~.*.06, name = "SentimentScore"))+
    theme(legend.position = c(0.2, 0.1))+
    geom_vline(xintercept = 1)+
    geom_vline(xintercept = startHalfTime, linetype="dashed")+
    geom_vline(xintercept = endHalfTime, linetype="dashed")+
    geom_vline(xintercept = endGame)+
    geom_text(x=11, y=2.5, label="1st half")+
    geom_text(x=72, y=2.5, label="2nd half")

}
```

Appendix H – Customized Stopwords

```
[1] "mahrez"           "city"             "mancity"          "mcfc"
  [5] "man"              "pep"              "bernardo"         "sterling"
  [9] "manchester"       "agueri"           "benjamin"         "mendy"
 [13] "sergio"           "aguero"           "guardiola"        "kev"
 [17] "fernandinho"      "kdb"              "mci"              "stones"
 [21] "riad"             "ayew"             "cpfc"             "crystal"
 [25] "palace"           "cry"              "reidawald"        "wilf"
 [29] "cenk"             "tosun"            "jairo"            "riedewald"
 [33] "james"            "mcarthur"         "gary"             "cahill"
 [37] "zaha"             "townsend"         "cenktosun"        "hodgson"
 [41] "crystalpalace"    "jordan"           "ayew's"           "tomkins"
 [45] "wilfred"          "milivojevic"      "andros"           "christian"
 [49] "benteke"          "joel"             "ward"             "glaziers"
 [53] "c"                "guaita"           "roy"              "max"
 [57] "van"              "aanholt"          "mccarthy"         "tonks"
 [61] "vicente"          "guaita's"         "selhurst"         "park"
 [65] "baldock"          "connor"           "wickham"          "brandon"
 [69] "pierrick"         "wilfried"         "martin"           "kelly"
 [73] "hogson"           "sakho"            "arsenal"          "sheffield"
 [77] "southampton"      "wolves"           "astonvilla"       "liverpool"
 [81] "utd"              "united"           "west"             "ham"
 [85] "everton"          "brighton"         "aston"            "villa"
 [89] "newcastle"        "sounew"           "premier"          "league"
 [93] "norwich"          "wolverhampton"    "shunor"           "wolbha"
 [97] "watford"          "keown"            "cryar"            "jorginho"
[101] "saido"            "berahino"         "brom"             "anfield"
[105] "tottenham"        "bournemouth"      "livs"             "livsou"
[109] "bouavl"           "newnor"           "wateve"           "whubri"
[113] "spurs"            "newlei"           "bhache"           "soutot"
[117] "andrea"           "pirlo"            "arsmun"           "livshu"
[121] "norcry"           "buravl"           "whubou"           "mcieve"
[125] "briche"           "livshe"           "marcus"           "rashford"
[129] "lingard"          "cryars"           "evebri"           "leicester"
[133] "chelsea"          "burnley"          "evebha"           "munnor"
[137] "chebur"           "wolnew"           "graham"           "scott"
[141] "premierleague"    "atkinson"         "pl"               "football"
[145] "epl"              "giroud"           "frankfurt"        "craig"
[149] "pawson"           "peter"            "walton"           "cresswells"
[153] "graig"            "dury"             "john"             "barnes"
[157] "attwell"          "bobby's"          "stormdennis"      "desmond"
[161] "andy"             "madley"           "maddison"         "dele"
[165] "alli"             "darren"           "england"          "chicharito"
[169] "lee"              "mason"            "southgate"        "stockley"
[173] "michael"          "oliver"           "paul"             "tierney"
[177] "ihenacho's"       "mount"            "emirates"         "tom"
[181] "daley"            "keane"            "moyes"            "irons"
[185] "haller"           "timo"             "werner"           "sebastian"
[189] "jarred"           "bowen"            "antonio"          "felipe"
[193] "anderson"         "fornals"          "fabianski"        "soucek"
[197] "westham"          "diop"             "hammers"          "mark"
[201] "noble"            "nobles"           "lanzini"          "aubameyang"
[205] "leno"             "ozil"             "alexandre"        "lacazette"
[209] "gooners"          "martinelli"       "laca"             "lacaaaaaaa"
[213] "lacaas"           "afc"              "lacazetteeeeee"   "auba"
[217] "arteta"           "dani"             "ceballos"         "pablo"
[221] "mari"             "pepe"             "saka"             "özil"
[225] "willock"          "nketiah"          "laccazete"        "mikel"
[229] "mabeba"           "aubameyang's"     "dhaka's"          "abangyermam"
[233] "aubameyamg"       "pierre"           "emerick"          "xhaka"
[237] "david"            "luiz"             "bernd"            "torreira"
[241] "guendouzi"        "nicolas"          "emery"            "artertaball"
[245] "mustafi"          "zimmerman"        "tettey"           "canaries"
[249] "vrancic"          "buendia"          "hernandez"        "drmic"
[253] "rupp"             "farke"            "otbc"             "cantwell"
[257] "ncfc"             "sam"              "byram"            "jamal"
[261] "lewis"            "tim"              "krul"             "duda"
[265] "pukki"            "allison"          "becker"           "mclean"
[269] "buendia's"        "tettey's"         "norwichcity"      "zimmermann's"
[273] "arrons"           "no"               "carrow"           "road"
[277] "grant"            "hanley"           "nor"              "stiepermann"
[281] "farke's"          "todd"             "mane"             "lfc"
```

```
[285] "alisson"        "sadio"          "joor"          "salah's"
[289] "liverpoolfc"    "ynwa"           "maneeeeee"     "sadiomane"
[293] "reds"           "salah"          "sade"          "jurgen"
[297] "klopp's"        "keita's"        "firmino"       "sadiooooooo"
[301] "ohmanemane"     "henderson"      "mo"            "salad"
[305] "naby"           "keita"          "sadie"         "dijk"
[309] "robbo"          "trent"          "milner"        "robertson"
[313] "divock"         "origi"          "scousers"      "london"
[317] "stadium"        "king"           "egypt"         "egyptian"
[321] "divooockk"      "origiiii"       "ox"            "curtis"
[325] "jones"          "alexander"      "arnold"        "chamberlain"
[329] "alexandera"     "gini"           "chargeado"     "surly"
[333] "shaq"           "oxlade"         "klopp"         "karius"
[337] "bobby"          "tyrone"         "mings"         "grealish"
[341] "douglas"        "ming's"         "nyland"        "guilbert"
[345] "albrighton"     "pereira"        "terry's"       "villans"
[349] "wesley"         "avfc"           "grealish's"    "wesley's"
[353] "jack"           "wes"            "marcou"        "kodjia"
[357] "elmo"           "jt"             "hourihane"     "stevens"
[361] "sheff"          "sheffieldunited" "sander"       "berge"
[365] "lunny"          "sufc"           "blayards"      "lys"
[369] "mousset"        "billy"          "sharp"         "norwood"
[373] "dean"           "wilder"         "ollie"         "ben"
[377] "nigel"          "pearson"        "billing"       "ake"
[381] "fraser"         "prowse"         "alex"          "sou"
[385] "ings"           "moussa"         "djenepo"       "stuart"
[389] "armstrong"      "redmond"        "jan"           "bednarek"
[393] "shane"          "long"           "st"            "mary's"
[397] "ralph"          "hasenhuttl"     "danny"         "oriol"
[401] "romeu"          "nathan"         "saints"        "saintsfc"
[405] "armstrong's"    "fernandez"      "dubravka"      "saint"
[409] "maximin"        "mina"           "holgate"       "kasper"
[413] "schmeichel"     "marco"          "silva"         "moise"
[417] "kean"           "ev"             "pereyra"       "masina"
[421] "kabasele"       "kabs"           "deulofeu"      "foster"
[425] "vicarage"       "deeney"         "cathcart"      "doucoure"
[429] "chalobah"       "hornets"        "sarr"          "doucouré"
[433] "wat"            "gerard"         "ismaila"       "troy"
[437] "snodgrass"      "issa"           "hammmers"      "jimenez"
[441] "raul"           "wwfc"           "wol"           "jonny's"
[445] "neto"           "raúl"           "jiménez"       "nuno"
[449] "doherty"        "jonny"          "dendoncker"    "adama"
[453] "traore"         "pedro"          "patricio"      "boly"
[457] "mgw"            "saïss"          "saiss"         "moutinho"
[461] "molineux"       "jota"           "neves"         "diogo"
[465] "patricio's"     "conor"          "coady"         "vinagre"
[469] "johnny"         "otto's"         "mou's"         "lcfc"
[473] "leicestercity"  "jamie"          "vardy"         "madison"
[477] "foxes"          "brendon"        "rodgers"       "evans"
[481] "kings"          "power"          "ndidi"         "filbert"
[485] "way"            "praet"          "hamza"         "choudhury"
[489] "leister"        "reina"          "harvey"        "iheanacho"
[493] "lestah"         "chilwell"       "chiwell"       "sidibe"
[497] "richarlison"    "kelechi"        "perez"         "turf"
[501] "moor"           "mee"            "sean"          "dyche"
[505] "chris"          "wood"           "190five"       "afcvcfc"
[509] "arscfc"         "arsche"         "arsenalfc"     "arsvsche"
[513] "betway"         "cfc"            "chelseafc"     "davidluiz"
[517] "ffscout"        "ktbffh"         "redarmy"       "supersunday"
[521] "abraham"        "azpilicueta"    "chambers"      "christensen"
[525] "dl"             "emerson"        "jody"          "joe"
[529] "kante"          "kepa"           "lamptey"       "maitland"
[533] "niles"          "moun"           "n'golo"        "nelson"
[537] "neville"        "ngolo"          "pulisic"       "roman"
[541] "tammy"          "tariq"          "tomori"        "zouma"
[545] "bbcfootball"    "btsport"        "miranda"       "mrwestham"
[549] "ndombele"       "odoi"           "skypl"         "whtid"
[553] "aaron"          "boruc"          "bridge"        "carney"
[557] "cockney"        "cresswell"      "cressy"        "eddie"
[561] "fredericks"     "hazard"         "howe"          "ibe"
[565] "karren"         "levy"           "londres"       "nobes"
[569] "pellegrini"     "solanke"        "stamford"      "sullivan"
[573] "bhafc"          "coyi"           "cryshu"        "munwol"
[577] "pldk"           "plonnbc"        "wapl"          "whu"
```

```
[581] "whubha"          "whufc"          "albion"          "angelo"
[585] "brings"          "glen"           "glenn"           "gross"
[589] "hove"            "jarrod"         "laporte"         "masuaku"
[593] "mcmessi"         "montoya"        "mooy"            "murray"
[597] "ogbonna"         "pascal"         "potter"          "propper"
[601] "rudiger"         "snoddy"         "trossard"        "webster"
[605] "yerry"           "bergwijn"       "josemourinho"    "totmci"
[609] "bbc"             "bruno"          "ederson"         "gondogan"
[613] "gundogan"        "leroy"          "lloris"          "manuel"
[617] "mike"            "mourinho"       "neuer"           "nicky"
[621] "raheem"          "sane"           "sonny"           "tyler"
[625] "walker"          "weaver"         "winksy"          "fridaynightfootball"
[629] "gw22"            "nbcsports"      "optussport"      "ppl20"
[633] "pplcup"          "rice"           "sheffieldutd"    "shuwhu"
[637] "superleague"     "suwh"           "westhamfamily"   "westhamutd"
[641] "adrian"          "balbeuna"       "besic"           "carlos"
[645] "carragher"       "declan"         "drury"           "hendo"
[649] "lucas"           "lundstram"      "mcbunni"         "mcburnie"
[653] "moyesy"          "oli"            "pickford"        "prem"
[657] "robert"          "tevez"          "vvd"             "godfrey"
[661] "nopukkinoparty"  "pitchsidetalk"  "pukkiparty"      "adam"
[665] "callum"          "cook"           "dom"             "dominic"
[669] "elmohamady"      "erling"         "etse"            "hart"
[673] "hull"            "jefferson"      "lambert"         "lerma"
[677] "mazy"            "norwhich"       "ondrej"          "pike"
[681] "rambo"           "ramsdale"       "rico"            "ryan"
[685] "smith"           "steve"          "teemu"           "wilson"
[689] "bha"             "coyb"           "efc"             "evertonfc"
[693] "ffelivecards"    "ancelotti"      "bernard"         "bissaka"
[697] "blues"           "calvert"        "lewin"           "lowin"
[701] "carlo"           "coleman"        "coote"           "dcl"
[705] "digne"           "duffy"          "dunk"            "goodison"
[709] "jahanbakhsh"     "maty"           "maupay"          "richy"
[713] "seamus"          "stephens"       "theo"            "toffees"
[717] "walcott"         "wan"            "wee"             "afcb"
[721] "afcbournemouth"  "ffelivegoals"   "fpl"             "sanmiguel"
[725] "utv"             "anthony"        "biling"          "bjorn"
[729] "cherries"        "dan"            "davis"           "ezri"
[733] "francis"         "gosling"        "grelish"         "jeff"
[737] "konsa"           "lermas"         "nakamba"         "philip"
[741] "samagoal"        "samatta"        "targett"         "taylor"
[745] "trezequet"       "vitality"       "wanyama"         "5livepremsunday"
[749] "avl"             "coys"           "fantasyfootball" "footballfrenzy"
[753] "fplcommunity"    "fplfyi"         "gw26"            "thfc"
[757] "alderweireld"    "alderwiereld"   "ali"             "aurier"
[761] "celso"           "rosie"          "davies"          "davinson"
[765] "dd"              "defo"           "defoe"           "dier"
[769] "drinkwater"      "el"             "engels"          "eriksen"
[773] "ffs"             "ghazi"          "gilbert"         "greslish"
[777] "harry"           "hause"          "hotspur"         "hotspurlive"
[781] "hugo"            "joelinton"      "jose"            "kane"
[785] "kayode"          "kortney"        "lemme"           "lingaard"
[789] "lo"              "moura"          "mufc"            "namkamba"
[793] "raina"           "sancho"         "serge"           "simon"
[797] "son"             "steven"         "tanganga"        "toby"
[801] "tomboita"        "vila"           "villians"        "winks"
[805] "zva"             "jorghinho"      "frank"           "uzil"
[809] "arsen"           "mesut"          "stoke"           "arswhu"
[813] "kenni"           "murphi"         "wolbri"          "davidrosi"
[817] "matt"            "tissier"        "bergjwin"        "eric"
[821] "engel"           "avltot"         "dane"            "lazi"
[825] "trezeguet"       "trump"          "calvertlewin"    "munmci"
[829] "paulo"           "richarlson"     "gea"             "shaw"
[833] "degea"           "bobbi"          "morinho"         "manu"
[837] "lindelof"        "virgil"         "gunnar"          "ole"
[841] "solskjaer"       "william"        "yazini"          "bbcsport"
[845] "manc"            "manutd"         "martial"         "fred"
[849] "lfcmun"          "livmun"         "perreira"        "mata"
[853] "bouwol"          "evenor"         "watbur"          "whutot"
[857] "citychelsea"     "azpi"           "lpool"           "willian"
[861] "pamoja"          "tushikili"      "elijah"          "gideon"
[865] "mose"            "bruyn"          "tammyn"          "mcicri"
[869] "arsshu"          "sissoko"        "norbou"          "souwol"
[873] "dsilva"
```

Appendix I – Code for Predicting Game Data from Machine Learning

```
source("LearningFunctions.R")
source("ImportCode.R")

#Game Data

## Sample Data for each game block incase needed

 # Insert official hashtag instead of XXXXXX
 game<-"XXXXXX"
 load("XXXXXX.RData")
 startHalfTime<-getStartHalfTime(game)
 endGame<-getEndGame(game)
 PredData <- XXXXXX.df
 PredData <- read_emojis(PredData, ascii_TRUE = FALSE)
 TestDataCorpus<-clean_text(PredData)
 dtmPrediction.df<-dtm(TestDataCorpus)
 dtmPredictionOrdered.df<-matrix_columns_same(dtmTraining.df, dtmPrediction.df)
 dtmPredictionOrdered <- as.matrix(dtmPredictionOrdered.df)
 predicted_class_sent<-predictSentimentData(dtmPredictionOrdered)
 predicted_class_sent_XXXXXX <- predicted_class_sent
 save(predicted_class_sent_XXXXXX, file = "XXXXXX_predicted_class_sent.RData")
 PredData$Sentiment = predicted_class_sent$max_prob
 predicted_class_var<-predictVARData(dtmPredictionOrdered)
 predicted_class_var_XXXXXX <- predicted_class_var
 save(predicted_class_var_XXXXXX, file = "XXXXXX_predicted_class_var.RData")
 PredData$VAR = predicted_class_var$max_prob
 PredData_XXXXXX <- PredData
 VARandSentiment<-combineGameData(PredData, startHalfTime, endGame)
 VARandSentiment_XXXXXX<-VARandSentiment

 save(VARandSentiment_XXXXXX, file = "XXXXXX_VARandSentiment.RData")
 save(PredData_XXXXXX, file = "XXXXXX_PredData.RData")
```

Appendix J – Functions Used for Data Analysis

```
library(dplyr)
library(purrr)
library(readr)
#library(plyr)

SentimentAndCount <- function(game)
{
  filename <- paste("PredData_", game,  sep ="")
  PredData <- eval(as.name(filename))

  analysis_return <- data.frame(nrow = 1, ncol = 5)
  VARandSent_Combined <- PredData[9]
  VARandSent_Combined[2] <- PredData[10]
  VARSent <-data.frame(ncol =1)
  OtherSent <-data.frame(ncol=1)
  numVARTweets <- 0
  numOtherTweets <- 0
  meanSentVAR <- 0
  meanSentOther <- 0
  for (i in 1:nrow(VARandSent_Combined)) {
    if(VARandSent_Combined[i,2] == 1)
    {
      numVARTweets<-numVARTweets+1
      VARSent[numVARTweets,1] <-VARandSent_Combined[i,1]
    }
    else
    {
      numOtherTweets<-numOtherTweets+1
      OtherSent[numOtherTweets,1]<-VARandSent_Combined[i,1]
    }
  }
  meanSentVAR<- as.numeric(colMeans(VARSent))
  meanSentOther <- as.numeric(colMeans(OtherSent))
  analysis_return[1] <- game
  analysis_return[2] <- numVARTweets
  analysis_return[3] <- meanSentVAR
  analysis_return[4] <- numOtherTweets
  analysis_return[5] <- meanSentOther
  colnames(analysis_return)[1]<- "game"
  colnames(analysis_return)[2]<- "numVARTweets"
  colnames(analysis_return)[3]<- "meanSentVAR"
  colnames(analysis_return)[4]<- "numOtherTweets"
  colnames(analysis_return)[5]<- "meanSentOther"

  return(analysis_return)
}

combineAllTweets <- function(game)
{
  filename <- paste("PredData_", game,  sep ="")
  PredData <- eval(as.name(filename))
  VARandSent_Combined <- PredData[9]
  VARandSent_Combined[2] <- PredData[10]
  return(VARandSent_Combined)
}

getSentiment<-function(VARandSent_Combined)
{
  analysis_return <- data.frame(nrow = 1, ncol = 4)
  VARSent <-data.frame(ncol=1)
  OtherSent <-data.frame(ncol=1)
  numVARTweets <- 0
  numOtherTweets <- 0
  meanSentVAR <- 0
  meanSentOther <- 0
  for (i in 1:nrow(VARandSent_Combined)) {
    if(VARandSent_Combined[i,2] == 1)
    {
      numVARTweets<-numVARTweets+1
      VARSent[numVARTweets,1] <-VARandSent_Combined[i,1]
```

```
    }
    else
    {
      numOtherTweets<-numOtherTweets+1
      OtherSent[numOtherTweets,1]<-VARandSent_Combined[i,1]
    }
  }
  meanSentVAR<- as.numeric(colMeans(VARSent))
  meanSentOther <- as.numeric(colMeans(OtherSent))
  analysis_return[1] <- numVARTweets
  analysis_return[2] <- meanSentVAR
  analysis_return[3] <- numOtherTweets
  analysis_return[4] <- meanSentOther
  colnames(analysis_return)[1]<- "numVARTweets"
  colnames(analysis_return)[2]<- "meanSentVAR"
  colnames(analysis_return)[3]<- "numOtherTweets"
  colnames(analysis_return)[4]<- "meanSentOther"

  return(analysis_return)
}

quantileAnalysis<-function(gameData)
{ #gameData = row of GameData related to this game
  library(dplyr)
  library(plyr)
  game <- toString(gameData[1,1])
  startHalfTime <- as.numeric(gameData[1,2])
  startHalfTime_period <- as.integer(startHalfTime/5)+1
  endHalfTime_period <- as.integer((startHalfTime+15)/5)+1
  endGame <- as.numeric(gameData[1,3])
  endGame_period <- as.integer(endGame/5)+1
  filename <- paste("PredData_", game,  sep ="")
  PredData <- eval(as.name(filename))

  analysis_return <- data.frame(ncol = 6)
  periodMin <- 5
  threshold <- 25
  period <- 1

  VARandSent_Combined <- PredData[9] #Sentiment
  VARandSent_Combined[2] <- PredData[10] #VAR
  VARandSent_Combined[3] <- PredData[8] #Minute Since Kickoff
  for(i in 1:nrow(VARandSent_Combined)){ #To get period value
    VARandSent_Combined[i,4] <- as.integer((VARandSent_Combined[i,3]-1)/periodMin)+1
  }
  colnames(VARandSent_Combined)[4] <- "Period"
  freqTable <-count(VARandSent_Combined$Period)
  detach(package:plyr) #need to remove so group_by function works properly
  senPerPeriod <- VARandSent_Combined %>%
    group_by(Period) %>%
    summarise(meanSent = mean(Sentiment))
  i<-1
  for (i in 1:ceiling(VARandSent_Combined$Period[nrow(VARandSent_Combined)])) {
    analysis_return[i,1] <- game
    if(period >= startHalfTime_period && period <= endHalfTime_period) {
      periodName <- paste(period, " HT", sep = "")
    } else if(period >= endGame_period) {
      periodName <- paste(period, " End", sep ="")
    } else {
      periodName <- period
    }
    analysis_return[i,2] <- periodName
    analysis_return[i,3] <- freqTable[i,2]
    analysis_return[i,4] <- senPerPeriod[i,2]
    analysis_return[i,5] <- sum(VARandSent_Combined$VAR & ceiling(VARandSent_Combined$Period)
== i)
    if(analysis_return[i,3] < threshold){
      analysis_return[i,6] <- FALSE
    } else { analysis_return[i,6] <- TRUE}
    period <- period+1
  }
  colnames(analysis_return)[1]<- "game"
  colnames(analysis_return)[2]<- "period"
  colnames(analysis_return)[3]<- "numTweets"
```

```
  colnames(analysis_return)[4]<- "meanSent"
  colnames(analysis_return)[5]<- "VARCount"
  colnames(analysis_return)[6]<- "OverThreshold"

  return(analysis_return)
}

peakIdentification <- function(gameData)
{
  library(plyr)
  library(dplyr)
  game <- toString(gameData[1,1])
  startHalfTime <- as.numeric(gameData[1,2])
  endHalfTime <- startHalfTime+15
  endGame <- as.numeric(gameData[1,3])
  filename <- paste("PredData_", game,  sep ="")
  PredData <- eval(as.name(filename))
  filename_VarSent <- paste("VARandSentiment_", game, sep = "")
  VARandSent_Data <- eval(as.name(filename_VarSent))

  VARandSent_Combined <- PredData[9] #Sentiment
  VARandSent_Combined[2] <- PredData[10] #VAR
  VARandSent_Combined[3] <- PredData[8] #Minute Since Kickoff

  freqTable <-count(VARandSent_Combined$MinuteSinceKickOff)
  #ensure frequency value for each minute
  if(nrow(freqTable) != 121) {freqTable <- freqTable_everyMin(freqTable)}

  totalTweetsGame <- sum(freqTable$freq)
  bool_PeakLast6 <- FALSE
  sixMinCount<-0
  bool_PrevPeak <- FALSE
  bool_PeakFound <- FALSE
  analysis_return <- data.frame()
  bool_shareDropBelow10 <- TRUE
  lastMinVARCount <- 0
  bool_moreThanlogPattern <- FALSE
  bool_duringStoppedTime <-FALSE
  peakTriggerData <- data.frame(nrow=1, ncol=6)
  for (i in 0:115) {
    if((i >= startHalfTime && i < endHalfTime) || i >= endGame){
      bool_duringStoppedTime <- TRUE
    } else { bool_duringStoppedTime <- FALSE }
    VARCount <- sum(VARandSent_Combined$VAR & ceiling(VARandSent_Combined$MinuteSinceKickOff)
== i)

    #share of tweets referring to VAR for this min above 10%
    if(freqTable[i+1,2] == 0) {bool_shareAbove10 <- FALSE
      } else { bool_shareAbove10 <- (VARCount/freqTable[i+1,2] > 0.10)}
    #num of VAR tweets inc more than natural log of total number of tweets for game compared
to prevMin
    bool_moreThanlogPattern <- (VARCount > log(totalTweetsGame)+lastMinVARCount)
    #no peak in last 6 min
    if(bool_PeakLast6){
      sixMinCount <- sixMinCount +1
      if(sixMinCount>6)
        bool_PeakLast6 <- FALSE }
    #if peak before in game, share needs to drop below 10% before new peak
    if(bool_PrevPeak) {
      if(!bool_shareAbove10){
        bool_shareDropBelow10 <- TRUE
      } }
    #search for peak
    if(!bool_PeakLast6 && bool_shareDropBelow10 && bool_shareAbove10 &&
!bool_duringStoppedTime && bool_moreThanlogPattern)
    {
      #bool_inPeak <- TRUE
      bool_PeakFound <- TRUE
      bool_PeakLast6 <- TRUE
      sixMinCount <- 1
      bool_PrevPeak <- TRUE
      bool_shareDropBelow10 <- FALSE
      peakTriggerData[1] <- game
      peakTriggerData[2] <- i
```

```
      peakTriggerData[3] <- freqTable[i+1,2]
      peakTriggerData[4] <- VARandSent_Data[i+1,3]
      peakTriggerData[5] <- VARandSent_Data[i+1,2]
      peakTriggerData[6] <- TRUE
      analysis_return <- rbind(analysis_return, peakTriggerData)
    }
     lastMinVARCount <- VARCount
  }

  if(bool_PeakFound)
  {
    colnames(analysis_return)[1]<- "game"
    colnames(analysis_return)[2]<- "MinuteSinceKickoff"
    colnames(analysis_return)[3]<- "numTweets"
    colnames(analysis_return)[4]<- "meanSent"
    colnames(analysis_return)[5]<- "VARCount"
    colnames(analysis_return)[6]<- "PeakFound"
  }
  return(analysis_return)
}


freqTable_everyMin <- function(freqTable)
{
  freqTable_everyMin<-data.frame()
  j=1
  for (i in 1:121){
    freqTable_everyMin[i,1]<-i-1
    freqTable_everyMin[i,2]<-ifelse(as.integer(freqTable$x[j])==(i-1), freqTable$freq[j],NA)
    if(is.na(freqTable_everyMin[i,2])) {
      freqTable_everyMin[i,2] <- 0
    } else {j <- j+1}
  }
  colnames(freqTable_everyMin)[1] <- "MinuteSinceKickoff"
  colnames(freqTable_everyMin)[2] <- "freq"
  return(freqTable_everyMin)
}

peakAnalysis <- function(triggerData)
{
  game <- toString(triggerData[1,1])
  triggerMinuteSinceKickOff <- as.numeric(triggerData[1,2])
  numTweets <- as.numeric(triggerData[1,3])
  meanSentTrigger <- as.numeric(triggerData[1,4])
  VARCount <- as.numeric(triggerData[1,5])

  filename <- paste("PredData_", game,  sep ="")
  PredData <- eval(as.name(filename))

  freqTable <-count(PredData$MinuteSinceKickOff)
  if(nrow(freqTable) != 121) {freqTable <- freqTable_everyMin(freqTable)}
  #change for different durations
  periodDuration <- 5
  peakPeriodData <- data.frame()

  #Before Trigger
  periodBeforeTime <- triggerMinuteSinceKickOff - periodDuration-1
  numTweetsBefore<-0
  VARCountBefore <- sum(PredData$VAR & ceiling(PredData$MinuteSinceKickOff) >=
periodBeforeTime & ceiling(PredData$MinuteSinceKickOff) < periodBeforeTime+5)
  VARandSent_Combined_Before <- subset(PredData, MinuteSinceKickOff >= periodBeforeTime &
MinuteSinceKickOff < periodBeforeTime+5, Sentiment)
  meanSentBefore<-as.numeric(colMeans(VARandSent_Combined_Before))
  for (i in 1:5) {
    numTweetsBefore <- numTweetsBefore + freqTable[periodBeforeTime+i,2]
    }
  peakPeriodData[1,1] <- game
  peakPeriodData[1,2] <- periodBeforeTime
  peakPeriodData[1,3] <- numTweetsBefore
  peakPeriodData[1,4] <- meanSentBefore
  peakPeriodData[1,5] <- VARCountBefore
  peakPeriodData[1,6] <- "Before Trigger"

  #Trigger
```

```
  peakPeriodData[2,1] <- game
  peakPeriodData[2,2] <- triggerMinuteSinceKickOff
  peakPeriodData[2,3] <- numTweets
  peakPeriodData[2,4] <- meanSentTrigger
  peakPeriodData[2,5] <- VARCount
  peakPeriodData[2,6] <- "Trigger"

  #After Trigger
  periodAfterTime <- triggerMinuteSinceKickOff + periodDuration-1
  numTweetsAfter<-0
  VARCountAfter<- sum(PredData$VAR & ceiling(PredData$MinuteSinceKickOff) >=
triggerMinuteSinceKickOff & ceiling(PredData$MinuteSinceKickOff) <= periodAfterTime)
  VARandSent_Combined_After <- subset(PredData, MinuteSinceKickOff >=
triggerMinuteSinceKickOff&  MinuteSinceKickOff <= periodAfterTime, Sentiment)
  meanSentAfter<-as.numeric(colMeans(VARandSent_Combined_After))
  for (i in 1:5) {
    numTweetsAfter <- numTweetsAfter + freqTable[triggerMinuteSinceKickOff+i,2]
  }

  peakPeriodData[3,1] <- game
  peakPeriodData[3,2] <- periodAfterTime
  peakPeriodData[3,3] <- numTweetsAfter
  peakPeriodData[3,4] <- meanSentAfter
  peakPeriodData[3,5] <- VARCountAfter
  peakPeriodData[3,6] <- "After Trigger"

  #6 min After Trigger
  periodAfterTime_later <- periodAfterTime+5
  numTweetsAfter_later<-0
  VARCountAfter_later<- sum(PredData$VAR & ceiling(PredData$MinuteSinceKickOff) >
periodAfterTime & ceiling(PredData$MinuteSinceKickOff) <= periodAfterTime_later)
  VARandSent_Combined_After_later <- subset(PredData, MinuteSinceKickOff > periodAfterTime &
MinuteSinceKickOff <= periodAfterTime_later, Sentiment)
  meanSentAfter_later<-as.numeric(colMeans(VARandSent_Combined_After_later))
  for (i in 1:5) {
    numTweetsAfter_later <- numTweetsAfter_later + freqTable[periodAfterTime+1+i,2]
  }

  peakPeriodData[4,1] <- game
  peakPeriodData[4,2] <- periodAfterTime_later
  peakPeriodData[4,3] <- numTweetsAfter_later
  peakPeriodData[4,4] <- meanSentAfter_later
  peakPeriodData[4,5] <- VARCountAfter_later
  peakPeriodData[4,6] <- "Later Trigger"


  colnames(peakPeriodData)[1] <- "game"
  colnames(peakPeriodData)[2] <- "MinuteSinceKickoff"
  colnames(peakPeriodData)[3] <- "numTweets"
  colnames(peakPeriodData)[4] <- "meanSent"
  colnames(peakPeriodData)[5] <- "VARCount"
  colnames(peakPeriodData)[6] <- "Classification"

  return(peakPeriodData)
  }
```

Appendix K – Code for Data Analysis

```
source("DataAnalysis_Functions.R")
library(dplyr)
library(purrr)
library(readr)
library(data.table)

#library(plyr)
#exclude MUNMCI
GameData <- GameData[-c(127),]

#Analysis 1: Grouped VAR and NonVAR Sentiment Values
dataAnalysis_VARvsOther_Test_2 <- data.frame(nrow = 130, ncol = 5)
dataAnalysis_VARvsOther_Test_2 <-SentimentAndCount(toString(GameData[1,1]))
for (i in 2:nrow(GameData)){
  dataAnalysis_VARvsOther_Test_2<-rbind(dataAnalysis_VARvsOther_Test_2,
SentimentAndCount(GameData[i,1]))
}
save(dataAnalysis_VARvsOther_Test_2, file = "dataAnalysis_VARvsOther_Test_2.RData")

#Analysis 2: Quantile analysis
dataQuantileAnalysis <- data.frame()
dataQuantileAnalysis <- quantileAnalysis(GameData[1,])
for (i in 2:nrow(GameData)){
  dataQuantileAnalysis<-rbind(dataQuantileAnalysis, quantileAnalysis(GameData[i,]))
}
dataQuantileAnalysis<-dataQuantileAnalysis_2
save(dataQuantileAnalysis, file = "dataQuantileAnalysis.RData")
write.csv(dataQuantileAnalysis, "dataQuantileAnalysis.csv")

#Analysis 3: Peak analysis (with MUNMCI excluded)
    #identification
dataPeakAnalysis_identification <- data.frame()
dataPeakAnalysis_identification <- peakIdentification(GameData[1,])
for (i in 2:nrow(GameData)) {
  dataPeakAnalysis_identification <- rbind(dataPeakAnalysis_identification,
peakIdentification(GameData[i,]))
}
save(dataPeakAnalysis_identification, file = "dataPeakAnalysis_identification.RData")
write.csv(dataPeakAnalysis_identification, "dataPeakAnalysis_identification.csv")

#analysis  5 min with later period too
dataPeakAnalysis_5minLater <- data.frame()
dataPeakAnalysis_5minLater <- peakAnalysis(dataPeakAnalysis_identification[1,])
for (i in 2:nrow(dataPeakAnalysis_identification)) {
  dataPeakAnalysis_5minLater <- rbind(dataPeakAnalysis_5minLater,
peakAnalysis(dataPeakAnalysis_identification[i,]))
}
save(dataPeakAnalysis_5minLater, file = "dataPeakAnalysis_5minLater.RData")
write.csv(dataPeakAnalysis_5minLater, "dataPeakAnalysis_5minLater.csv")


#Combine all predData tweets
PredData_All_2 <- data_frame()
for(i in 1:nrow(GameData)) {
  game <- toString(GameData[i,1])
  filename <- paste("PredData_", game,  sep ="")
  PredData <- eval(as.name(filename))
  PredData[11] <- game
  PredData_All_2 <- rbind(PredData_All_2, PredData)
}
colnames(PredData_All_2)[11] <- "game"
save(PredData_All_2, file = "PredData_All_2.RData")
write.csv(PredData_All_2, "PredData_All_2.csv")


 #All VAR and sentiment together
VARandSentiment_All <- data_frame()
for(i in 1:nrow(GameData)) {
  game <- toString(GameData[i,1])
  filename <- paste("VARandSentiment_", game,  sep ="")
  VARandSentiment <- eval(as.name(filename))
```

```
  VARandSentiment[5] <- game
  VARandSentiment_All <- rbind(VARandSentiment_All, VARandSentiment)
}
colnames(VARandSentiment_All)[5] <- "game"
save(VARandSentiment_All, file = "VARandSentiment_All.RData")
write.csv(VARandSentiment_All, "VARandSentiment_All.csv")


#neg tweets related to VAR
uncountedTweets <- data_frame()
countVAR_neg <- 0
countTotal_neg <- 0
countVAR_neu <- 0
countTotal_neu <- 0
countVAR_pos <- 0
countTotal_pos <- 0
uncounted <- 0
for (i in 1:nrow(PredData_All_2)) {
  if(PredData_All[i,10] == 1) #then VAR
  {
    if(PredData_All_woMUNMCI[i,9] == 0) #sentiment rating neu
    { countVAR_neu <- countVAR_neu + 1
    countTotal_neu <- countTotal_neu + 1
    } else if(PredData_All_woMUNMCI[i,9] == 1) { # pos
      countVAR_pos <- countVAR_pos + 1
      countTotal_pos <- countTotal_pos + 1
    } else if(PredData_All_woMUNMCI[i,9] == -1) { # neg
      countVAR_neg <- countVAR_neg + 1
      countTotal_neg <- countTotal_neg + 1
    } else {
      uncounted <- uncounted +1
      uncountedTweets <- rbind(uncountedTweets, PredData_All[i,])
    }
  } else if(PredData_All_woMUNMCI[i,10] == 0){
    if(PredData_All_woMUNMCI[i,9] == 0) { #sentiment rating neu
      countTotal_neu <- countTotal_neu + 1
    } else if(PredData_All_woMUNMCI[i,9] == 1) { # pos
      countTotal_pos <- countTotal_pos + 1
    } else if(PredData_All_woMUNMCI[i,9] == -1) { # neg
      countTotal_neg <- countTotal_neg + 1
    } else {
      uncounted <- uncounted +1
      uncountedTweets <- rbind(uncountedTweets, PredData_All[i,])
    }
  }
  else
  {
    uncounted <- uncounted +1
    uncountedTweets <- rbind(uncountedTweets, PredData_All[i,])
  }}
```

Appendix L – Game Data

Table 22. Game Data

Day	Time kick-off (local time)	Game Hashtag	VAR Used?	Start Half Time	End Game	Total Number of Tweets
5-Dec-19	21:15	BOUWOL	Yes	48	112	11522
29-Dec-19	15:00	WHUTOT	Yes	48	115	13048
23-Feb-20	17:30	ARSSOU	Yes	49	117	10248
16-Feb-20	17:30	BHALEI	No	47	111	10699
18-Jan-20	16:00	CRYLIV	Yes	48	114	5029
23-Nov-19	16:00	EVENOR	Yes	47	113	8870
7-Mar-20	16:00	WATBUR	Yes	51	116	5841
12-Jan-20	17:30	MCICHE	Yes	48	114	6740
25-Nov-19	21:00	AVLNEW	No	48	112	1555
26-Dec-19	16:00	NEWMCI	Yes	48	113	470
16-Feb-20	15:00	LIVBHA	Yes	47	112	5390
21-Jan-20	20:30	BURCRY	No	48	115	553
18-Jan-20	16:00	CHEWHU	Yes	48	115	287
28-Dec-19	13:30	TOTBOU	Yes	47	115	1439
1-Jan-20	13:30	SOUWAT	No	48	115	4954
23-Nov-19	16:00	MUNAVL	Yes	49	114	339
8-Feb-20	18:30	NORARS	No	49	114	584
26-Dec-19	16:00	WOLSHU	Yes	48	112	6247
1-Feb-20	16:00	LEIEVE	Yes	48	116	679
21-Jan-20	20:30	CRYBOU	Yes	48	113	423
7-Dec-19	16:00	BURMCI	No	48	112	4561
12-Jan-20	15:00	CHEAVL	No	48	113	1713
23-Nov-19	16:00	SOUNOR	No	47	112	516
2-Feb-20	15:00	WOLWHU	No	48	112	7514
1-Jan-20	13:30	LEIWAT	Yes	48	115	1762
22-Feb-20	16:00	LIVEVE	Yes	49	113	559
30-Nov-19	16:00	SHUNEW	Yes	48	115	312
19-Jan-20	15:00	ARSBHA	No	47	112	2151
3-Dec-19	21:15	EVECHE	No	48	115	3292
28-Dec-19	20:45	BOULIV	No	49	112	5437
7-Mar-20	18:30	TOTBUR	No	49	112	2477
21-Jan-20	21:15	WATCRY	No	47	113	23796
4-Dec-19	20:30	WHUARS	No	51	119	3433
11-Jan-20	16:00	CRYBHA	Yes	47	111	3822
8-Mar-20	15:00	TOTCHE	No	50	119	11234
17-Feb-20	21:00	CHESOU	Yes	48	113	40127
26-Dec-19	16:00	CRYWHU	No	46	112	5586
22-Feb-20	13:30	EVEBUR	Yes	47	112	15699
30-Nov-19	16:00	LEILIV	Yes	47	111	4040

11-Jan-20	13:30	MUNNEW	Yes	48	112	9122
16-Dec-19	20:45	AVLNOR	Yes	49	115	285
3-Dec-19	20:30	BOUARS	No	47	112	1435
23-Nov-19	16:00	TOTBHA	Yes	48	113	6692
22-Feb-20	16:00	BURMUN	No	49	115	1035
1-Feb-20	16:00	NEWEVE	Yes	47	112	618
21-Jan-20	20:30	SOUCRY	No	47	112	385
7-Mar-20	16:00	BHABOU	Yes	47	111	485
26-Dec-19	16:00	NORTOT	No	48	116	819
11-Jan-20	16:00	WATAVL	Yes	49	116	433
26-Dec-19	16:00	WHULEI	No	49	114	529
7-Dec-19	13:30	ARSCHE	No	51	119	10116
8-Feb-20	13:30	MCISHU	No	47	112	2262
21-Jan-20	20:30	BHACHE	Yes	49	113	2003
23-Nov-19	16:00	BURAVL	Yes	48	118	729
9-Mar-20	21:00	MCIEVE	Yes	48	113	3755
1-Feb-20	13:30	NEWLEI	Yes	49	113	8771
1-Dec-19	17:30	NORCRY	Yes	47	113	4362
26-Dec-19	21:00	SOUTOT	No	49	115	12060
22-Feb-20	18:30	WATWOL	Yes	50	117	4726
11-Jan-20	16:00	WHUBOU	Yes	49	114	1634
4-Dec-19	20:30	LIVSHU	Yes	47	110	814
22-Jan-20	20:30	SHUWHU	Yes	48	114	2395
30-Nov-19	16:00	CHEBUR	No	50	113	4602
7-Mar-20	13:30	CRYARS	No	48	115	9935
4-Dec-19	21:15	EVEBHA	Yes	48	114	13834
19-Jan-20	17:30	LEISOU	Yes	48	115	36127
2-Jan-20	21:00	MUNNOR	No	47	111	6123
1-Feb-20	16:00	TOTLIV	Yes	49	114	6942
24-Feb-20	21:00	WOLNEW	Yes	51	116	14180
23-Nov-19	18:30	BOUWAT	Yes	49	114	13008
18-Jan-20	16:00	AVLMCI	Yes	46	108	4189
1-Jan-20	18:30	NEWCHE	Yes	51	117	3317
29-Dec-19	19:00	ARSSHU	Yes	48	112	2014
19-Feb-20	20:30	BHAAVL	No	48	114	2745
1-Dec-19	17:30	MCICRY	No	48	115	9550
22-Jan-20	21:15	NORBOU	Yes	48	114	14759
26-Dec-19	18:30	SOUWOL	No	47	114	8573
11-Jan-20	16:00	WATTOT	Yes	48	113	6837
23-Feb-20	15:00	WHUEVE	Yes	47	112	11504
1-Feb-20	18:30	BURLEI	No	47	112	8336
18-Jan-20	18:30	LIVMUN	No	48	112	4991
28-Dec-19	16:00	AVLWAT	No	48	114	1388
1-Jan-20	16:00	SHUMCI	Yes	47	113	1538
30-Nov-19	13:30	CHEARS	No	48	113	6551
1-Feb-20	16:00	CRYSOU	No	49	114	454
1-Dec-19	15:00	BOUBHA	Yes	47	113	11385

18-Jan-20	16:00	EVENEW	Yes	47	113	541
1-Jan-20	18:30	LEIWHU	Yes	51	116	745
15-Feb-20	18:30	MUNBUR	Yes	48	113	6190
28-Dec-19	18:30	WOLLIV	Yes	48	113	7952
22-Feb-20	16:00	WHULIV	Yes	48	112	399
9-Feb-20	15:00	MUNWOL	No	48	113	1205
21-Jan-20	20:30	NEWNOR	No	48	113	2916
5-Dec-19	20:30	BOUAVL	Yes	48	113	1557
7-Mar-20	16:00	CRYSHU	No	47	114	457
10-Jan-20	21:00	LEICHE	Yes	48	114	2890
22-Feb-20	16:00	LIVSOU	No	47	112	807
15-Feb-20	13:30	WATEVE	Yes	50	116	1363
28-Dec-19	16:00	WHUBHA	No	47	114	464
7-Mar-20	16:00	BURARS	Yes	47	111	1054
4-Dec-19	20:30	TOTMCI	No	50	116	381
1-Jan-20	16:00	BHAWAT	Yes	47	112	3762
30-Nov-19	18:30	EVECRY	Yes	49	113	794
18-Jan-20	16:00	SHUBOU	Yes	49	114	705
26-Dec-19	13:30	WOLLEI	Yes	49	115	3952
30-Nov-19	16:00	NORLIV	Yes	48	112	1980
7-Dec-19	16:00	SOUBUR	Yes	51	116	3026
22-Dec-19	17:30	ARSNEW	Yes	47	114	7543
11-Jan-20	18:30	AVLTOT	Yes	49	115	12733
2-Feb-20	17:30	CHEMUN	Yes	49	115	17283
28-Dec-19	16:00	MCIWHU	Yes	47	112	699
23-Nov-19	16:00	CRYNEW	Yes	47	114	284
7-Dec-19	16:00	SOUAVL	Yes	50	116	339
1-Feb-20	16:00	CHETOT	Yes	47	112	1052
18-Jan-20	13:30	SHUBHA	Yes	48	113	3573
1-Jan-20	16:00	BURBOU	Yes	47	112	769
9-Dec-19	21:00	LEIMCI	Yes	48	112	10282
1-Feb-20	16:00	ARSEVE	Yes	51	117	543
1-Jan-20	18:30	MUNWAT	Yes	47	111	1558
18-Jan-20	16:00	LIVWHU	Yes	47	113	710
28-Dec-19	18:30	ARSWHU	Yes	47	113	1163
29-Jan-20	20:45	BURTOT	Yes	48	113	5065
23-Nov-19	13:30	CHEEVE	Yes	46	109	7805
7-Mar-20	16:00	CRYWAT	No	49	115	292
14-Feb-20	21:00	LEIAVL	Yes	47	110	2679
23-Jan-20	21:00	LIVBOU	Yes	50	114	12490
11-Jan-20	16:00	SHUNOR	Yes	47	111	699
1-Dec-19	15:00	SOUNEW	Yes	49	112	706
4-Dec-19	20:30	WOLBHA	No	48	111	549

Appendix M – Sentiment Rating of VAR Tweets vs Other

Table 23. Sentiment Rating of VAR Tweets vs Other

	Game	Num VAR Tweets	Mean Sent VAR	Num Other Tweets	Mean Sent Other
1	ARSBHA	431	-0.719	11 091	-0.117
2	ARSCHE	566	-0.754	12482	0.081
3	ARSEVE	196	-0.485	10052	0.199
4	ARSNEW	163	-0.718	10536	0.213
5	ARSSHU	323	-0.647	4706	0.081
6	ARSSOU	367	-0.741	8503	-0.086
7	ARSWHU	748	-0.341	5093	0.062
8	AVLMCI	174	-0.466	6566	0.057
9	AVLNEW	27	-0.667	1528	0.034
10	AVLNOR	10	-0.600	460	0.122
11	AVLTOT	390	-0.518	5000	0.097
12	AVLWAT	6	-0.833	547	0.077
13	BHAAVL	8	-0.500	279	0.143
14	BHABOU	453	-0.784	986	0.175
15	BHACHE	127	-0.669	4827	0.165
16	BHALEI	32	-0.563	307	0.244
17	BHAWAT	16	-0.938	568	0.144
18	BOUARS	63	-0.635	6184	0.028
19	BOUAVL	33	-0.333	646	0.135
20	BOUBHA	4	0.250	419	0.215
21	BOULIV	20	-0.400	4541	0.282
22	BOUWAT	46	-0.630	1667	0.061
23	BOUWOL	56	-0.464	460	0.165
24	BURARS	121	-0.785	7393	-0.123
25	BURAVL	718	-0.777	1044	0.086
26	BURBOU	191	-0.497	368	0.234
27	BURCRY	13	-0.462	299	0.187
28	BURLEI	124	-0.597	2027	0.122
29	BURMCI	59	-0.390	3233	0.170
30	BURMUN	231	-0.736	5206	0.159
31	BURTOT	191	-0.681	2286	-0.045
32	CHEARS	986	-0.465	22810	0.048
33	CHEAVL	45	-0.711	3388	0.156
34	CHEBUR	287	-0.362	3535	0.257
35	CHEEVE	62	-0.645	11172	0.304
36	CHEMUN	8053	-0.587	32074	0.014
37	CHESOU	28	-0.786	5558	-0.119
38	CHETOT	3941	-0.740	11758	0.127
39	CHEWHU	123	-0.374	3917	0.000

40	CRYARS	1799	-0.599	7323	0.054
41	CRYBHA	32	-0.688	253	0.004
42	CRYBOU	180	-0.506	1255	0.037
43	CRYLIV	1516	-0.555	5176	0.170
44	CRYNEW	19	-0.526	1016	0.018
45	CRYSHU	55	-0.345	563	0.020
46	CRYSOU	18	-0.944	367	0.136
47	CRYWAT	20	-0.600	465	0.095
48	CRYWHU	5	-0.600	814	0.119
49	EVEBHA	101	-0.594	332	0.105
50	EVEBUR	8	-1.000	521	0.198
51	EVECHE	87	-0.678	10029	-0.051
52	EVECRY	25	-0.680	2237	0.165
53	EVENEW	11	-0.364	1992	0.093
54	EVENOR	19	-0.579	710	-0.001
55	LEIAVL	694	-0.816	3061	0.017
56	LEICHE	230	-0.552	8541	0.029
57	LEIEVE	904	-0.295	3458	0.064
58	LEILIV	400	-0.448	11660	0.247
59	LEIMCI	1075	-0.646	3651	0.005
60	LEISOU	219	-0.584	1415	0.194
61	LEIWAT	182	-0.797	632	0.180
62	LEIWHU	221	-0.652	2174	0.028
63	LIVBHA	154	-0.409	4448	0.196
64	LIVBOU	945	-0.772	8990	0.104
65	LIVEVE	433	-0.538	13401	0.205
66	LIVMUN	4700	-0.726	31427	-0.027
67	LIVSHU	197	-0.523	5926	0.259
68	LIVSOU	529	-0.681	6413	0.207
69	LIVWHU	447	-0.577	13733	0.147
70	MCICHE	900	-0.720	12108	0.111
71	MCICRY	492	-0.563	3697	0.098
72	MCIEVE	709	-0.626	2608	0.057
73	MCISHU	770	-0.740	1244	0.057
74	MCIWHU	137	-0.796	2608	-0.057
75	MUNAVL	220	-0.636	9330	-0.059
76	MUNBUR	140	-0.493	14619	-0.151
78	MUNNEW	111	-0.477	8462	0.193
79	MUNNOR	142	-0.225	6695	0.245
80	MUNWAT	945	-0.304	10559	0.256
81	MUNWOL	156	-0.718	8180	-0.107
82	NEWCHE	99	-0.798	4892	-0.024
83	NEWEVE	70	-0.729	1318	0.218
84	NEWLEI	16	-0.563	1522	0.005
85	NEWMCI	147	-0.388	6404	0.194

86	NEWNOR	5	-0.600	449	0.067
87	NORARS	1039	-0.581	10346	-0.037
88	NORBOU	92	-0.239	449	0.158
89	NORCRY	76	-0.224	669	0.112
90	NORLIV	326	-0.641	5864	0.236
91	NORTOT	3877	-0.792	4075	-0.040
92	SHUBHA	34	-0.735	365	0.219
93	SHUBOU	15	-0.867	1190	0.164
94	SHUMCI	354	-0.718	2562	0.054
95	SHUNEW	333	-0.294	1224	0.104
96	SHUNOR	8	-0.625	449	0.107
97	SHUWHU	1215	-0.749	1675	-0.039
98	SOUAVL	8	-0.750	799	0.006
99	SOUBUR	185	-0.778	1178	0.160
100	SOUCRY	104	-0.635	360	0.133
101	SOUNEW	162	-0.444	892	0.056
102	SOUNOR	3	-0.667	378	0.196
103	SOUTOT	372	-0.672	3390	-0.050
104	SOUWAT	36	-0.861	758	0.156
105	SOUWOL	53	-0.377	652	0.206
106	TOTBHA	681	-0.708	3271	0.069
107	TOTBOU	136	-0.662	1844	0.220
108	TOTBUR	20	-0.600	3006	0.251
109	TOTCHE	986	-0.419	6557	0.059
110	TOTLIV	962	-0.703	11771	0.122
111	TOTMCI	4098	-0.654	13185	0.094
112	WATAVL	40	-0.300	659	0.083
113	WATBUR	25	-0.160	259	0.135
114	WATCRY	21	-0.571	318	0.019
115	WATEVE	32	-0.688	1020	0.112
116	WATTOT	508	-0.636	3065	-0.011
117	WATWOL	132	-0.439	637	0.127
118	WHUARS	213	-0.709	10069	0.003
119	WHUBHA	41	-0.488	502	0.229
120	WHUBOU	140	-0.379	1418	0.159
121	WHUEVE	34	-0.500	676	0.022
122	WHULEI	57	-0.439	1106	0.047
123	WHULIV	238	-0.408	4827	0.201
124	WHUTOT	364	-0.668	7441	0.090
125	WOLBHA	2	-0.500	290	0.028
126	WOLLEI	1204	-0.764	1475	-0.012
127	WOLLIV	232	-0.560	12258	0.188
128	WOLNEW	10	-0.800	689	0.016
129	WOLSHU	27	-0.889	773	0.107
130	WOLWHU	6	-0.833	543	0.085

Appendix N – Sentiment Rating of VAR Tweets in Relation to Trigger

Table 24. Sentiment Rating of VAR Tweets in Relation to Trigger

	Game	Minute Since Kickoff	Num Tweets	Mean Sent	VAR Count	Classification
1	WHUTOT	7	182	0.027	14	Before Trigger
2	WHUTOT	13	110	-0.527	40	Trigger
3	WHUTOT	17	334	-0.434	94	After Trigger
4	WHUTOT	22	199	-0.161	9	Later Trigger
5	WHUTOT	102	202	-0.198	9	Before Trigger
6	WHUTOT	108	55	-0.145	16	Trigger
7	WHUTOT	112	239	-0.071	71	After Trigger
8	WHUTOT	117	424	0.097	19	Later Trigger
9	ARSSOU	24	190	-0.047	12	Before Trigger
10	ARSSOU	30	68	-0.279	21	Trigger
11	ARSSOU	34	264	-0.231	53	After Trigger
12	ARSSOU	39	205	-0.127	15	Later Trigger
13	ARSSOU	84	225	-0.244	6	Before Trigger
14	ARSSOU	90	73	-0.110	12	Trigger
15	ARSSOU	94	684	-0.161	110	After Trigger
16	ARSSOU	99	834	-0.168	40	Later Trigger
21	CRYLIV	39	165	-0.006	0	Before Trigger
22	CRYLIV	45	109	-0.174	34	Trigger
23	CRYLIV	49	953	-0.374	600	After Trigger
24	CRYLIV	54	579	-0.366	324	Later Trigger
25	CRYLIV	95	132	-0.061	18	Before Trigger
26	CRYLIV	101	137	0.219	17	Trigger
27	CRYLIV	105	713	0.224	50	After Trigger
28	CRYLIV	110	393	0.181	21	Later Trigger
33	CHEWHU	81	155	0.039	3	Before Trigger
34	CHEWHU	87	65	-0.231	15	Trigger
35	CHEWHU	91	316	-0.082	69	After Trigger
36	CHEWHU	96	161	-0.118	9	Later Trigger
37	TOTBOU	36	59	0.000	4	Before Trigger
38	TOTBOU	42	21	-0.476	11	Trigger
39	TOTBOU	46	82	-0.378	41	After Trigger
40	TOTBOU	51	66	-0.030	19	Later Trigger
41	MUNAVL	25	395	-0.122	8	Before Trigger
42	MUNAVL	31	119	-0.235	24	Trigger

43	MUNAVL	35	549	-0.231	56	After Trigger
44	MUNAVL	40	407	-0.192	16	Later Trigger
45	NORARS	19	385	-0.018	1	Before Trigger
46	NORARS	25	173	-0.145	42	Trigger
47	NORARS	29	941	-0.128	265	After Trigger
48	NORARS	34	1028	-0.208	392	Later Trigger
49	NORARS	39	272	-0.202	36	Before Trigger
50	NORARS	45	55	-0.164	19	Trigger
51	NORARS	49	814	-0.077	76	After Trigger
52	NORARS	54	1021	-0.231	25	Later Trigger
53	LEIEVE	28	116	-0.026	6	Before Trigger
54	LEIEVE	34	57	-0.439	21	Trigger
55	LEIEVE	38	513	-0.343	331	After Trigger
56	LEIEVE	43	180	-0.383	90	Later Trigger
61	CRYBOU	16	35	0.143	0	Before Trigger
62	CRYBOU	22	51	-0.255	24	Trigger
63	CRYBOU	26	257	-0.358	116	After Trigger
64	CRYBOU	31	96	-0.260	15	Later Trigger
65	LEIWAT	35	30	0.100	3	Before Trigger
66	LEIWAT	41	27	-0.667	21	Trigger
67	LEIWAT	45	99	-0.657	74	After Trigger
68	LEIWAT	50	52	-0.173	21	Later Trigger
69	LIVEVE	21	1037	0.293	11	Before Trigger
70	LIVEVE	27	84	-0.131	21	Trigger
71	LIVEVE	31	503	-0.229	133	After Trigger
72	LIVEVE	36	1292	0.268	38	Later Trigger
77	ARSBHA	6	153	0.124	4	Before Trigger
78	ARSBHA	12	48	-0.313	19	Trigger
79	ARSBHA	16	185	-0.178	49	After Trigger
80	ARSBHA	21	191	-0.110	10	Later Trigger
81	ARSBHA	25	294	-0.173	9	Before Trigger
82	ARSBHA	31	83	-0.253	16	Trigger
83	ARSBHA	35	423	-0.173	39	After Trigger
84	ARSBHA	40	1151	-0.152	11	Later Trigger
85	ARSBHA	76	241	-0.087	10	Before Trigger
86	ARSBHA	82	88	-0.182	36	Trigger
87	ARSBHA	86	351	-0.219	145	After Trigger
88	ARSBHA	91	184	-0.071	18	Later Trigger
89	WHUARS	35	374	-0.166	35	Before Trigger
90	WHUARS	41	172	-0.169	22	Trigger
91	WHUARS	45	682	-0.260	67	After Trigger
92	WHUARS	50	467	-0.225	6	Later Trigger

93	TOTCHE	41	165	-0.176	16	Before Trigger
94	TOTCHE	47	60	-0.150	28	Trigger
95	TOTCHE	51	821	-0.048	267	After Trigger
96	TOTCHE	56	692	0.027	78	Later Trigger
97	TOTCHE	76	181	-0.061	10	Before Trigger
98	TOTCHE	82	96	-0.135	29	Trigger
99	TOTCHE	86	900	-0.148	305	After Trigger
100	TOTCHE	91	532	-0.226	104	Later Trigger
101	TOTCHE	102	244	-0.078	11	Before Trigger
102	TOTCHE	108	49	-0.449	16	Trigger
103	TOTCHE	112	260	-0.112	33	After Trigger
104	TOTCHE	117	404	0.149	17	Later Trigger
105	LEILIV	82	258	0.089	9	Before Trigger
106	LEILIV	88	102	0.216	26	Trigger
107	LEILIV	92	1321	0.225	224	After Trigger
108	LEILIV	97	2280	0.321	49	Later Trigger
109	TOTBHA	20	48	0.188	1	Before Trigger
110	TOTBHA	26	98	-0.520	70	Trigger
111	TOTBHA	30	408	-0.581	303	After Trigger
112	TOTBHA	35	128	-0.250	62	Later Trigger
113	BURMUN	76	149	-0.007	7	Before Trigger
114	BURMUN	82	51	-0.157	17	Trigger
115	BURMUN	86	183	-0.251	49	After Trigger
116	BURMUN	91	249	-0.032	37	Later Trigger
117	BHABOU	71	42	0.024	4	Before Trigger
118	BHABOU	77	53	-0.604	38	Trigger
119	BHABOU	81	403	-0.687	317	After Trigger
120	BHABOU	86	88	-0.636	61	Later Trigger
121	NORTOT	29	116	-0.103	3	Before Trigger
122	NORTOT	35	188	-0.351	75	Trigger
123	NORTOT	39	2170	-0.698	1749	After Trigger
124	NORTOT	44	881	-0.730	695	Later Trigger
125	NORTOT	93	107	-0.150	10	Before Trigger
126	NORTOT	99	47	-0.106	20	Trigger
127	NORTOT	103	324	-0.117	80	After Trigger
128	NORTOT	108	156	-0.038	32	Later Trigger
129	ARSCHE	29	497	-0.032	15	Before Trigger
130	ARSCHE	35	168	-0.137	17	Trigger
131	ARSCHE	39	574	-0.047	47	After Trigger
132	ARSCHE	44	359	-0.067	46	Later Trigger
133	ARSCHE	92	375	-0.128	13	Before Trigger
134	ARSCHE	98	93	-0.333	20	Trigger

135	ARSCHE	102	482	-0.320	94	After Trigger
136	ARSCHE	107	919	-0.037	60	Later Trigger
145	MCISHU	65	28	-0.107	2	Before Trigger
146	MCISHU	71	77	-0.221	21	Trigger
147	MCISHU	75	355	-0.513	188	After Trigger
148	MCISHU	80	132	-0.424	62	Later Trigger
149	BURAVL	9	41	-0.073	1	Before Trigger
150	BURAVL	15	24	-0.542	15	Trigger
151	BURAVL	19	341	-0.716	282	After Trigger
152	BURAVL	24	190	-0.595	127	Later Trigger
153	MCIEVE	8	127	0.079	0	Before Trigger
154	MCIEVE	14	64	0.141	12	Trigger
155	MCIEVE	18	368	-0.188	207	After Trigger
156	MCIEVE	23	359	-0.579	246	Later Trigger
157	MCIEVE	99	108	-0.028	17	Before Trigger
158	MCIEVE	105	34	-0.235	19	Trigger
159	MCIEVE	109	101	-0.287	38	After Trigger
160	MCIEVE	114	84	0.119	4	Later Trigger
165	SOUTOT	62	55	-0.255	4	Before Trigger
166	SOUTOT	68	33	0.030	13	Trigger
167	SOUTOT	72	288	-0.503	192	After Trigger
168	SOUTOT	77	122	-0.262	29	Later Trigger
173	SHUWHU	105	53	0.094	1	Before Trigger
174	SHUWHU	111	58	0.017	9	Trigger
175	SHUWHU	115	1094	-0.588	799	After Trigger
176	SHUWHU	120	520	-0.604	399	Later Trigger
177	CHEBUR	21	151	-0.219	43	Before Trigger
178	CHEBUR	27	57	-0.281	24	Trigger
179	CHEBUR	31	366	0.019	90	After Trigger
180	CHEBUR	36	148	-0.061	34	Later Trigger
181	CRYARS	78	256	-0.105	22	Before Trigger
182	CRYARS	84	74	-0.216	32	Trigger
183	CRYARS	88	1042	-0.349	654	After Trigger
184	CRYARS	93	737	-0.351	375	Later Trigger
189	LEISOU	77	39	-0.077	6	Before Trigger
190	LEISOU	83	26	-0.346	17	Trigger
191	LEISOU	87	108	-0.537	79	After Trigger
192	LEISOU	92	41	-0.439	17	Later Trigger
193	LEISOU	102	127	0.173	1	Before Trigger
194	LEISOU	108	31	-0.226	13	Trigger
195	LEISOU	112	153	-0.026	70	After Trigger
196	LEISOU	117	204	0.319	18	Later Trigger

197	MUNNOR	62	161	0.161	0	Before Trigger
198	MUNNOR	68	87	0.046	19	Trigger
199	MUNNOR	72	708	0.301	58	After Trigger
200	MUNNOR	77	787	0.079	8	Later Trigger
201	TOTLIV	33	388	0.008	15	Before Trigger
202	TOTLIV	39	306	0.229	52	Trigger
203	TOTLIV	43	1204	0.010	300	After Trigger
204	TOTLIV	48	510	-0.141	71	Later Trigger
205	TOTLIV	79	362	0.039	8	Before Trigger
206	TOTLIV	85	79	-0.152	26	Trigger
207	TOTLIV	89	525	-0.229	163	After Trigger
208	TOTLIV	94	432	-0.139	36	Later Trigger
209	AVLMCI	101	333	-0.036	4	Before Trigger
210	AVLMCI	107	54	0.111	11	Trigger
211	AVLMCI	111	361	0.280	30	After Trigger
212	AVLMCI	116	198	0.187	8	Later Trigger
213	ARSSHU	82	151	0.020	0	Before Trigger
214	ARSSHU	88	81	-0.136	39	Trigger
215	ARSSHU	92	309	-0.204	130	After Trigger
216	ARSSHU	97	167	0.120	18	Later Trigger
217	MCICRY	12	104	0.029	3	Before Trigger
218	MCICRY	18	43	-0.372	24	Trigger
219	MCICRY	22	180	-0.317	87	After Trigger
220	MCICRY	27	82	-0.049	8	Later Trigger
221	MCICRY	83	87	-0.080	2	Before Trigger
222	MCICRY	89	25	-0.080	11	Trigger
223	MCICRY	93	360	-0.333	234	After Trigger
224	MCICRY	98	187	-0.203	64	Later Trigger
225	WATTOT	8	96	0.010	12	Before Trigger
226	WATTOT	14	27	-0.481	11	Trigger
227	WATTOT	18	117	-0.214	26	After Trigger
228	WATTOT	23	94	-0.053	5	Later Trigger
229	WATTOT	23	85	-0.176	1	Before Trigger
230	WATTOT	29	32	-0.500	12	Trigger
231	WATTOT	33	215	-0.419	86	After Trigger
232	WATTOT	38	138	-0.232	33	Later Trigger
233	WATTOT	81	135	-0.304	30	Before Trigger
234	WATTOT	87	39	0.179	14	Trigger
235	WATTOT	91	370	-0.003	86	After Trigger
236	WATTOT	96	171	-0.041	17	Later Trigger
237	BURLEI	80	85	-0.129	3	Before Trigger
238	BURLEI	86	34	0.029	10	Trigger

239	BURLEI	90	294	0.143	30	After Trigger
240	BURLEI	95	90	0.033	9	Later Trigger
241	LIVMUN	19	1116	-0.005	30	Before Trigger
242	LIVMUN	25	269	-0.138	39	Trigger
243	LIVMUN	29	3345	-0.420	1483	After Trigger
244	LIVMUN	34	1964	-0.377	649	Later Trigger
245	SHUMCI	30	79	-0.127	1	Before Trigger
246	SHUMCI	36	45	0.044	15	Trigger
247	SHUMCI	40	432	-0.169	142	After Trigger
248	SHUMCI	45	195	-0.303	64	Later Trigger
249	CHEARS	21	325	-0.098	1	Before Trigger
250	CHEARS	27	313	-0.003	68	Trigger
251	CHEARS	31	2863	-0.101	333	After Trigger
252	CHEARS	36	1530	-0.164	98	Later Trigger
253	CHEARS	64	413	-0.133	23	Before Trigger
254	CHEARS	70	66	-0.333	17	Trigger
255	CHEARS	74	384	-0.232	52	After Trigger
256	CHEARS	79	391	-0.130	30	Later Trigger
257	LEIWHU	65	63	-0.016	2	Before Trigger
258	LEIWHU	71	44	-0.159	17	Trigger
259	LEIWHU	75	197	-0.218	80	After Trigger
260	LEIWHU	80	74	-0.095	18	Later Trigger
261	WHULIV	28	153	-0.020	0	Before Trigger
262	WHULIV	34	60	0.050	30	Trigger
263	WHULIV	38	399	0.113	110	After Trigger
264	WHULIV	43	214	0.075	33	Later Trigger
265	LEICHE	16	329	-0.061	1	Before Trigger
266	LEICHE	22	76	-0.263	28	Trigger
267	LEICHE	26	314	-0.111	53	After Trigger
268	LEICHE	31	288	0.063	2	Later Trigger
269	LIVSOU	16	191	-0.042	16	Before Trigger
270	LIVSOU	22	51	-0.275	13	Trigger
271	LIVSOU	26	212	-0.085	31	After Trigger
272	LIVSOU	31	181	0.072	9	Later Trigger
273	LIVSOU	26	173	0.052	12	Before Trigger
274	LIVSOU	32	66	-0.045	21	Trigger
275	LIVSOU	36	365	-0.255	150	After Trigger
276	LIVSOU	41	157	-0.064	24	Later Trigger
277	LIVSOU	67	349	0.080	57	Before Trigger
278	LIVSOU	73	60	0.000	13	Trigger
279	LIVSOU	77	221	-0.081	47	After Trigger
280	LIVSOU	82	564	0.441	26	Later Trigger

285	BURARS	27	220	-0.218	4	Before Trigger
286	BURARS	33	50	-0.200	16	Trigger
287	BURARS	37	257	-0.206	30	After Trigger
288	BURARS	42	309	-0.139	2	Later Trigger
289	TOTMCI	8	210	-0.019	11	Before Trigger
290	TOTMCI	14	132	-0.402	84	Trigger
291	TOTMCI	18	791	-0.503	509	After Trigger
292	TOTMCI	23	307	-0.241	110	Later Trigger
293	TOTMCI	32	173	-0.052	17	Before Trigger
294	TOTMCI	38	75	-0.267	15	Trigger
295	TOTMCI	42	1824	-0.375	894	After Trigger
296	TOTMCI	47	2514	-0.337	1088	Later Trigger
297	TOTMCI	75	284	-0.106	33	Before Trigger
298	TOTMCI	81	224	-0.098	57	Trigger
299	TOTMCI	85	1818	0.191	200	After Trigger
300	TOTMCI	90	772	0.194	44	Later Trigger
301	WOLLEI	40	69	-0.087	2	Before Trigger
302	WOLLEI	46	50	-0.220	17	Trigger
303	WOLLEI	50	669	-0.689	507	After Trigger
304	WOLLEI	55	353	-0.654	266	Later Trigger
305	NORLIV	80	179	0.112	7	Before Trigger
306	NORLIV	86	42	-0.262	13	Trigger
307	NORLIV	90	237	-0.165	53	After Trigger
308	NORLIV	95	329	0.170	7	Later Trigger
313	ARSNEW	82	322	0.283	0	Before Trigger
314	ARSNEW	88	65	-0.046	19	Trigger
315	ARSNEW	92	289	0.100	41	After Trigger
316	ARSNEW	97	253	0.138	5	Later Trigger
317	AVLTOT	38	156	0.071	5	Before Trigger
318	AVLTOT	44	48	-0.125	12	Trigger
319	AVLTOT	48	344	-0.137	134	After Trigger
320	AVLTOT	53	466	-0.114	136	Later Trigger
321	CHEMUN	16	458	-0.039	4	Before Trigger
322	CHEMUN	22	123	-0.081	16	Trigger
323	CHEMUN	26	1221	-0.342	468	After Trigger
324	CHEMUN	31	1186	-0.265	127	Later Trigger
325	CHEMUN	30	1228	-0.305	58	Before Trigger
326	CHEMUN	36	160	-0.094	21	Trigger
327	CHEMUN	40	1563	-0.228	201	After Trigger
328	CHEMUN	45	1272	-0.192	92	Later Trigger
329	CHEMUN	70	531	-0.066	29	Before Trigger
330	CHEMUN	76	226	-0.062	68	Trigger

331	CHEMUN	80	2719	-0.454	1449	After Trigger
332	CHEMUN	85	1568	-0.373	618	Later Trigger
333	CHEMUN	91	2160	0.064	153	Before Trigger
334	CHEMUN	97	497	-0.068	135	Trigger
335	CHEMUN	101	4269	-0.265	2253	After Trigger
336	CHEMUN	106	1993	-0.282	799	Later Trigger
337	MCIWHU	18	166	-0.157	9	Before Trigger
338	MCIWHU	24	55	-0.564	16	Trigger
339	MCIWHU	28	221	-0.330	58	After Trigger
340	MCIWHU	33	181	-0.011	11	Later Trigger
341	CHETOT	12	630	0.381	12	Before Trigger
342	CHETOT	18	317	0.167	52	Trigger
343	CHETOT	22	925	0.128	113	After Trigger
344	CHETOT	27	443	-0.081	16	Later Trigger
345	CHETOT	64	1103	0.291	5	Before Trigger
346	CHETOT	70	254	0.000	87	Trigger
347	CHETOT	74	2565	-0.543	1815	After Trigger
348	CHETOT	79	1068	-0.453	574	Later Trigger
349	CHETOT	93	427	-0.089	54	Before Trigger
350	CHETOT	99	174	-0.080	22	Trigger
351	CHETOT	103	843	-0.332	298	After Trigger
352	CHETOT	108	725	-0.257	250	Later Trigger
353	BURBOU	72	31	0.258	2	Before Trigger
354	BURBOU	78	26	-0.308	19	Trigger
355	BURBOU	82	111	-0.216	70	After Trigger
356	BURBOU	87	46	-0.239	28	Later Trigger
357	LEIMCI	24	87	-0.149	1	Before Trigger
358	LEIMCI	30	33	-0.364	18	Trigger
359	LEIMCI	34	191	-0.393	94	After Trigger
360	LEIMCI	39	103	-0.019	13	Later Trigger
361	LEIMCI	35	103	-0.019	13	Before Trigger
362	LEIMCI	41	110	-0.109	53	Trigger
363	LEIMCI	45	429	-0.340	190	After Trigger
364	LEIMCI	50	211	-0.223	68	Later Trigger
365	LEIMCI	73	94	0.096	5	Before Trigger
366	LEIMCI	79	44	-0.136	18	Trigger
367	LEIMCI	83	924	-0.186	357	After Trigger
368	LEIMCI	88	308	-0.237	109	Later Trigger
369	LEIMCI	94	126	-0.222	11	Before Trigger
370	LEIMCI	100	99	-0.101	22	Trigger
371	LEIMCI	104	295	-0.108	46	After Trigger
372	LEIMCI	109	117	0.060	4	Later Trigger

373	MUNWAT	36	291	-0.261	2	Before Trigger
374	MUNWAT	42	109	-0.037	35	Trigger
375	MUNWAT	46	1000	0.220	166	After Trigger
376	MUNWAT	51	539	0.163	52	Later Trigger
377	MUNWAT	66	144	0.153	6	Before Trigger
378	MUNWAT	72	205	0.005	90	Trigger
379	MUNWAT	76	782	-0.143	359	After Trigger
380	MUNWAT	81	1790	0.397	109	Later Trigger
381	LIVWHU	98	1412	0.206	25	Before Trigger
382	LIVWHU	104	240	0.138	36	Trigger
383	LIVWHU	108	969	0.102	153	After Trigger
384	LIVWHU	113	449	0.109	20	Later Trigger
385	ARSWHU	79	186	-0.081	0	Before Trigger
386	ARSWHU	85	58	-0.259	26	Trigger
387	ARSWHU	89	297	-0.290	80	After Trigger
388	ARSWHU	94	296	-0.311	36	Later Trigger
389	ARSWHU	91	292	-0.281	38	Before Trigger
390	ARSWHU	97	100	-0.220	49	Trigger
391	ARSWHU	101	880	0.009	424	After Trigger
392	ARSWHU	106	292	-0.007	73	Later Trigger
393	BURTOT	68	136	0.007	19	Before Trigger
394	BURTOT	74	32	-0.531	11	Trigger
395	BURTOT	78	121	-0.488	36	After Trigger
396	BURTOT	83	46	0.022	5	Later Trigger
397	LEIAVL	73	54	-0.130	0	Before Trigger
398	LEIAVL	79	35	-0.257	11	Trigger
399	LEIAVL	83	610	-0.500	396	After Trigger
400	LEIAVL	88	288	-0.514	177	Later Trigger
401	LIVBOU	4	221	0.136	2	Before Trigger
402	LIVBOU	10	243	0.012	40	Trigger
403	LIVBOU	14	1458	-0.236	375	After Trigger
404	LIVBOU	19	793	-0.202	124	Later Trigger
405	SOUNEW	24	40	-0.050	7	Before Trigger
406	SOUNEW	30	28	-0.214	18	Trigger
407	SOUNEW	34	91	-0.165	46	After Trigger
408	SOUNEW	39	65	-0.231	14	Later Trigger

Appendix O – Computer Specifications

Operation System:	Windows 10 64-bit, x64-based processor
Application:	RStudio Version 1.2.5042
Processor:	AMD Ryzen 9 3900X 12-Core Processor 3.80 GHz
Installed memory (RAM):	CORSAIR Vengeance LPX 128GB (4x32GB) DDR4 2666
Storage:	Samsung 960 PRO NVMe M.2 1TB SSD
Graphics Card:	MSI Gaming X GeForce GTX 1660 Super (Was not used)
Motherboard:	MSI Prestige x570 Creation
Power Supply:	Cooler Master MWE Gold 650 Full Modular