



**HAL**  
open science

## Fans' Attachment to Players in the Indian Premier League: Insights from Twitter Analytics

Giridhar B. Kamath, Shirshendu Ganguli, Simon George, Vibha Prabhu

► **To cite this version:**

Giridhar B. Kamath, Shirshendu Ganguli, Simon George, Vibha Prabhu. Fans' Attachment to Players in the Indian Premier League: Insights from Twitter Analytics. International Working Conference on Transfer and Diffusion of IT (TDIT), Dec 2020, Tiruchirappalli, India. pp.451-462, 10.1007/978-3-030-64861-9\_40 . hal-03744777

**HAL Id: hal-03744777**

**<https://inria.hal.science/hal-03744777v1>**

Submitted on 3 Aug 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Fans' Attachment to Players in the Indian Premier League: insights from twitter analytics

Giridhar B Kamath<sup>1</sup>, Shirshendu Ganguli<sup>2</sup>, Simon George<sup>3</sup>, Vibha<sup>4</sup>

<sup>1</sup>Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, 576104, India; Research Scholar, T A Pai Management Institute, Manipal, 576104, India

<sup>2,3</sup>T A Pai Management Institute, Manipal, 576104, India

<sup>4</sup>Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, 576104, India

giridhar.kamath@manipal.edu

**Abstract.** Sports in India are gradually gaining traction, and people are more and more inclined towards various other sports apart from cricket. Even though several sporting leagues are currently running in India, the attention and viewership that the Indian Premier League (IPL) enjoys are unmatched. Social media posts and the growing viewership on TV and Over the Top (OTT) platforms are evidence of this. This research investigates the level of attachment of the fans with the players of their favorite team. We study the opinions of fans through the tweets extracted during the IPL players auctions held in the preseason of the thirteenth edition of the IPL. The fans were very active on social media, airing their views about the various players available in the auction for their team to buy, indicating high team identification. In total, 15,374 tweets that were extracted and processed confirmed the same. Further, using System Dynamics (SD) methodology, we develop a causal loop linking the various parameters of importance in the IPL ecosystem that could help the league and the team managers to understand the importance of player attachment among the fans better. This study is the first in the context of IPL that uses the System Dynamics methodology to justify the findings of twitter analytics.

**Keywords:** Opinion mining, Twitter analytics, Indian Premier League, Player auctions, Causal loop diagram, System dynamics.

## 1 Introduction

Social media has emerged as a medium to express the feelings of the users without any hesitation. Such impulsive postings that people make on social media regarding various events/topics/issues have many implications to various interested parties. Researchers have been studying the social media posts of people and have successfully drawn conclusions about specific topics under concern. E.g., Brand marketers have successfully captured the sentiments/opinions of the people towards their brands based on social media posts. Likewise, social media posts about specific issues help the stakeholders to gauge the sentiments of people like positive, negative, neutral, fear, agony, etc.

In sports, especially in Western countries, the user-generated content on social media has been a prevalent concept. Fans with high attachment to their favorite teams and favorite players have been generating much content on social media, and such contents have gained millions of views and have helped the team or players to maintain their popularity and has garnered much interest from the sponsor and other stakeholders alike.

In India, cricket is watched more than any other sport. With the penetration of the internet happening at a never seen before pace, the viewership of all other sports is on the rise, but cricket stills stand much ahead. The current study deals with the auctions of players in the Indian Premier League (IPL). IPL is a T20 cricket league based in India with a viewership of over 400 million both on TV and over the top (OTT) platforms. IPL is the second most popular sporting league in the world. The first season of IPL was played in 2008 in India, and in 2019 it completed its 12 years. Even though there were several player auctions held from time to time, this research deals with the player auctions held in December 2019 before the 13th season of the IPL. The auctions were live on TV, and also live streaming was available on OTT platforms.

The purpose of this study is to look into the social media posts of the fans during the auctions and assess the involvement of the fans through their posts and their sentiments displayed during the auction process. Fans who were following the auctions have expressed their opinions on social media sites like Facebook and Twitter. Our study deals with posts on Twitter. The tweets were extracted, and then through opinion mining using the software R, we tried to gauge the attachment the fans have towards the various players, both national and international. The word cloud of the frequently used words helped in visualizing the words clearly and helped in highlighting the underlying pattern of the content of the tweets.

Further, we try to highlight the effect of fans' attachment to the players in IPL on the other parameters in and around the IPL ecosystem using a causal loop diagram based on the available literature and highlight the possible linkages between the points of attachment and the fans' attachment to their favorite teams. In this pursuit, we consider the bigger picture with several direct and indirect factors influencing the entire IPL ecosystem.

The rest of the paper is designed as follows: First, we discuss the importance of attachment points in sports. Next, we look at social media analytics literature in the context of sports. Subsequently, we discuss the methodology of this study, followed by the results and discussions leading to the contribution of this study.

## **2 Literature Review**

### **2.1 Attachment Points**

Attachment theory says that human beings tend to maintain psychological ties with particular objects. Attachment points influence indirect sports consumption. Kwon et al. (2005) define attachment points as "a consumer's psychological ties with various

aspects associated with his or her favorite sport team.” Funk et al. (2016) define attachment as the emotional, functional, and symbolic meaning assigned by a fan to a sports team. Heere and James (2007) argue that the American sports team usually carries the name of the city or state in their team name because, according to them, the attachment with a team happens because of several external factors like geographical, ethnic, or several other factors surrounding the team. In the IPL, all the eight teams have their city names included in their team name like Chennai Super Kings (CSK), Delhi Capitals (DC), Kings XI Punjab (KXI Punjab), Kolkata Knight Riders(KKR), Mumbai Indians (MI), Rajasthan Royals (RR), Royal Challengers Bangalore (RCB), and, SunRisers Hyderabad (SRH).

Yoshida et al. (2015) studied attachment points in Japan professional League and established a link between the attachment points and team identification. According to Katz and Heere (2013), attachment points play a role in influencing how the sport is consumed (direct or indirect).

Shapiro et al. (2013) highlighted that multiple attachment points could influence the way the sport is consumed and how the fans would react to the sponsors. Heere et al. (2011) highlighted the importance of attachment towards players in a sport and argued that including players in the marketing plans can be fruitful for the sponsors. Lock et al. (2011) say that attachment towards the sport itself can be a vital factor that influences the formation of team identification in the case of a new team.

Overall, the literature on attachment points indicates that there are three main attachment points with respect to any sport. They are sport attachment, player attachment, and the city attachment. In this research, since we study the sentiments of fans during the pre-season players’ auction, we confine our scope to the players’ attachment with an assumption that fans who are tweeting already have higher sport attachment.

## 2.2 Social Media analytics

According to Pang and Lee (2008) and Liu (2012), Opinion Mining has become an active area of research, and there are various applications such as competitive intelligence (Xu et al., 2011), and market prediction (Liu et al., 2007, and Zhang and Sjena, 2010), product pricing (Zhang and Sjena (2010), Jin et al. (2019), and nation relationship analysis (Chambers et al., 2015). There is a significant growth in user-generated content on the internet. Extracting useful information from the user-generated content is interesting and challenging (Pang and Lee, 2008).

Several prior studies, like Derczynski et al. (2013), Pak and Parouubek (2010), have researched twitter and sentiment analysis. According to Liu (2012) the word opinion can be represented as a quintuple  $(e_i, a_{ij}, s_{ijkl}, h_k, t_\ell)$ , in which  $e_i$  is the  $i^{\text{th}}$  entity,  $a_{ij}$  is the  $j^{\text{th}}$  aspect of the  $i^{\text{th}}$  entity,  $h_k$  is the  $k^{\text{th}}$  opinion holder,  $t_\ell$  is the time when the opinion is expressed,  $s_{ijkl}$  is the opinion or sentiment towards the  $i^{\text{th}}$  aspect of the  $j^{\text{th}}$  entity from opinion holder  $h_k$  at time  $t_\ell$ . As per the definition, for opinion mining, for the given text, the quintuple is determined. The third component is enough for document-level opinion mining instead of the full quintuple. However, for the fine-grained level, more components are required.

The growth of social media platforms, e-commerce, and review sites such as Twitter, Reddit, Yelp, and Amazon provides a sufficient amount of resources as corpora for academic research. Mining the useful information from social media will be valuable and benefits the different applications. Hence social media data can be used as a representative tool of real-time experiences such as for sports events. During these events, people post a large number of posts, thus revealing their expressions, opinions, and suggestions according to the circumstances occurring during the events. Stojanovski et al. (2015) analyzed the views conveyed in the users' posts during football games and could intimate the concerned authorities of possible riots after the match. Fan and Gordon (2014) investigated the differences and similarities between the Chinese NBA fans and western NBA fans with a global presence in terms of their social media usage and points of attachment to a team.

The Social Media Analytics Process (Fan and Gordon, 2014) is mainly a three-stage process that involves capturing relevant data from social media sites by carefully monitoring various social media resources and extracting the required details from those sources. The data extracted is generally in an unstructured format. The second stage involves an understanding of the data and converting the extracted unstructured data in a structured format. This stage is crucial because it involves selecting relevant data for modeling, removing noisy and low-quality data. This process consists of various advanced data analytic strategies to analyze the retained data and gain insights from it. The third stage is the presentation, which becomes a representational way to visualize the relevant information extracted. There are many easy to use built-in analytic tools available in the market for research and academic purposes for Facebook, Twitter, etc. The problem with these tools is that these are "black-box" where a researcher does not know how the data model is built and the methodology is not transparent. The use of programming languages such as Python or R with their associated packages for data science analytics to analyze social media data makes a fully transparent methodology (Yun et al., 2019).

### **3 Methodology**

This research combined two methods. First, opinion mining was done using social media analytics. Next, the importance of attachment points was highlighted using the system dynamics methodology. A causal loop was developed to study the interrelationships between the various constructs and parameters within the scope of this study.

#### **3.1 Twitter analytics using R**

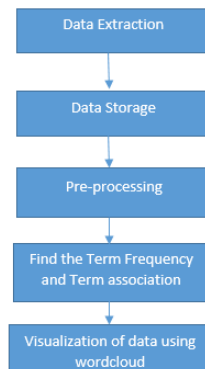
The IPL players' auction was held on 19th December 2019 in Kolkata, India. The tweets were extracted for the following period: one week before the auction date and one week after the auction ended. This research is limited to opinion mining of the tweets during and after the IPL player auction. Fifteen thousand three hundred seventy-four tweets were extracted concerning all the eight teams in the tournament. This study is limited to developing a word cloud using the software R. The word cloud would

enable one to understand the frequently used words in the tweets and analyze the opinions of fans visually, regarding the auctioning and buying of their choicest players by the team management.

### 3.2 Data mining

We have collected data from the popularly used social media platform Twitter. Twitter data is actively used for many purposes like to make recommendations, gather insights, and lodge public complaints. The tweets posted by the users are in the form of hashtags, and the extent that users find valuable information shows its validity. Hence appropriate hashtags used by the concerned eight IPL teams were used to extract the tweets posted. To ingest the social media data from Twitter, Twitter API is used, which requires consumer key, consumer secret key, request token URL, access token URL, and authorize URL. The function called searchTwitter defined in twitter package will pull the tweets associated with the hashtags defined in the function. Data is mined for all eight teams using the teams' hashtags and stored in separate files.

The data thus mined is not well structured and hence requires some preliminary processing. Preliminary processing involves converting all words to lowercase, removing all punctuation, and dropping the stopwords. The pre-processing mainly consists of imposing some structure on the text data and converting it into a document-term matrix (dtm). The necessity of the Document-term matrix is the easiness of analyzing the terms and finding the unique, common, and repeated words from the collection of documents called corpus. Further from this matrix, we can explore the data and find the correlation in the document-term matrix between the term it references and the specific word in the corpus using the term association (i.e., using a function called findAssocs). Data visualization provides an easy way to view and analyze the data. In this paper, from the obtained document-term matrix (dtm), a word cloud is formed, and this is used to visualize the term frequency. Only the top 250 terms were used in this wordcloud. The process used for data mining is depicted in Fig. 1.

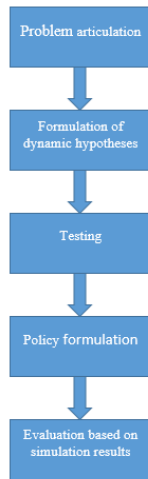


**Fig. 1.** Process of data collection and processing.

### 3.3 System Dynamics methodology

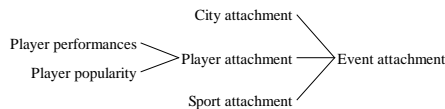
The main objective of using this methodology is to develop a cause and effect relationship between the player attachment and other selected parameters in and around the event. If we consider only the player attachment, it will not be clear how the attachment

fits in with the system. When we look at the bigger picture of the league with player attachment points as one of the parameters, we can see how the whole cause and effect relationship unfolds within the given boundary. The current study takes some links among the variables from the research conducted by Choi et al. (2017), Crescitelli and Figueiredo (2009), Hsiao, Peng, and Huang (2010), and Levine and Call (2003). Hsiao et al. (2010), in their study on a professional baseball game in Taiwan, developed a system dynamics simulation model to understand the system structure and simulate the developing trend. Using System Dynamics methodology (Sterman, 2000), we attempt to model the IPL ecosystem by focussing on the following clusters: attachment points, player attachment, sport attachment, city attachment, team identification, number of fans, team income, and the event income. SD methodology follows the steps as shown in Fig. 2.



**Fig. 2.** System Dynamics methodology (Source: Sterman, 2000).

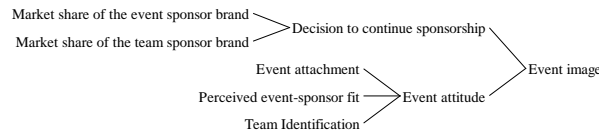
As seen in Fig: 3, we relate the city attachment, player attachment, and sport attachment to the event attachment. The other attachment points drive the attachment towards the event. The players' attachment is also dependent on player performances and the players' popularity. Players who enjoy immense popularity because of their on-field performances and off-field conduct are the most sought after by the teams and the team fans.



**Fig. 3.** Factors affecting event attachment.

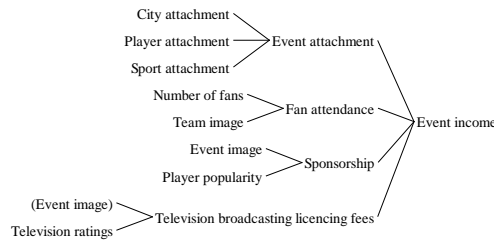
Fans wish that their team buys such popular players who can contribute to the team performances. This is evident from the tweets during the auction period. Fans were

either hoping their teams bought certain players or were seen tweeting positive tweets after their favorite teams bought the anticipated players.



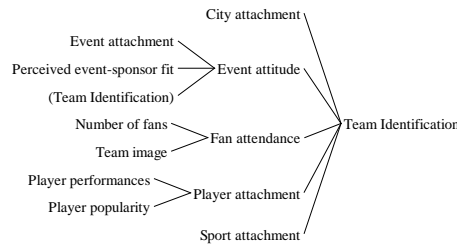
**Fig. 4.** Factors affecting event image.

Fig. 4 shows the links assumed in this study to relate specific parameters with the event image. The market shares of the event sponsors and the team sponsors determine their decision to continue with the sponsorship deals, respectively. This decision could have an impact on the event image. Also, the attitude towards the event that the fans' have influences the event image.



**Fig. 5.** Factors affecting event income.

The number of fans and the image of the team influences fan attendance. The event image and the popularity of the players influence sponsorship. The income that the event generates depends on the factors such as the fan attendance at the games in the stadium, the sponsorship deals with the sponsors, and the television broadcast licensing fees, as shown in Fig. 5.



**Fig. 6.** Factors leading to team identification.

Fig. 6 depicts the connection between the parameters that explain the team identification. As seen in the previous research work, we have equated the attachment points as predictors of team identification. Also, event attitude and fan attendance at the games strengthen team identification. Event attitude is influenced by the event attachment and the perceived fit between the event and the sponsoring brand of the event. We further



assume that the link or the strength of the relationship between the event and the sponsor influences the attitude the fans hold towards the event, which in turn affects the team identification.

#### 4 Results and discussions

The tweets were structured in the pre-processing stage to estimate the word frequency. From these structured tweeted data, we found the term frequency of each term present in the corpus and found the associativity of these terms. From the obtained term frequency matrix, wordcloud was created for data visualization purposes. The wordcloud of the selected three active teams in the auctions are as follows. We can see that the fans have tweeted about the players already existing in their favorite team and also are welcoming the new buys their team management has made. The word cloud helps us visualize the word frequencies as shown in Fig 7.



Fig. 7. Wordcloud for Delhi Capitals.

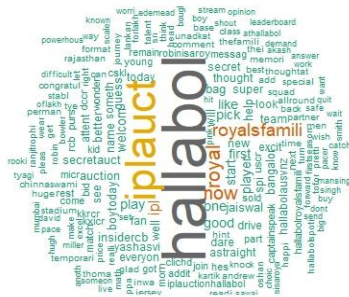
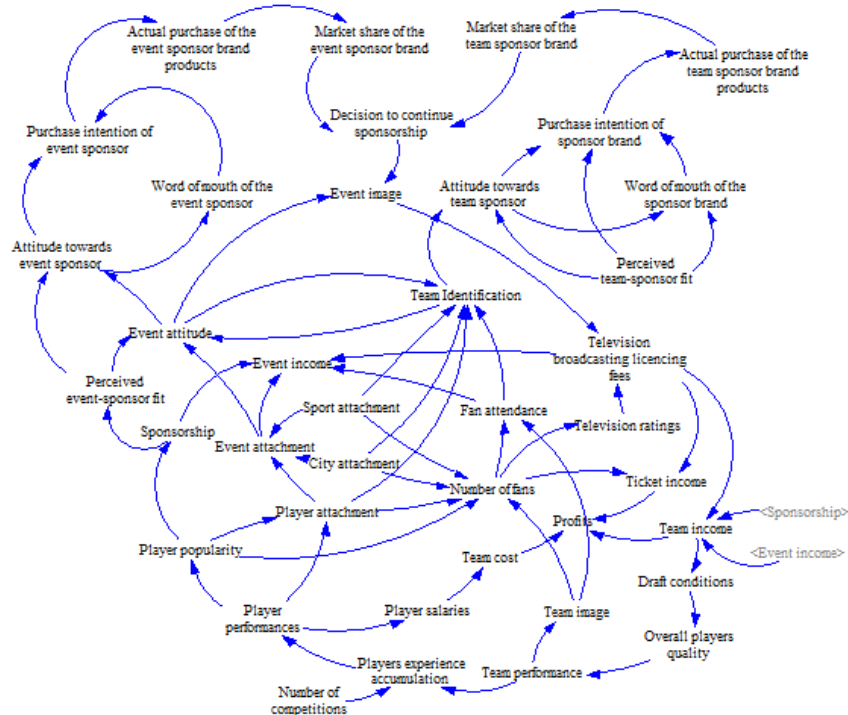


Fig. 8. Wordcloud for Rajasthan Royals.

The wordcloud of the next two teams is shown in figure 8 and 9, respectively. It is very evident that the fans were very concerned about the players of their teams and had a strong attachment, which leads to higher team identification. A clear link between the players' attachment and team identification can be seen here. Fig. 10 depicts the wordcloud after considering the overall tweets of the auction. Here it is visible that apart from the player names, fans have been tweeting about their favorite teams and, in general, positively about the event and the sport as such in excitement.





**Fig. 11.** Causal loop diagram.

The causal loop highlights the benefits of fans' attachment to players in a sporting league. If seen in isolation, the fans' attachment to players may not seem to be of much importance to the sponsors. But when modeled and focused on the interrelations, one can see how strong the attachment to the players is. Attachment to specific players is dependent on the player's past performance and experience too. Hence, teams can target to get players based on their experience in the sport. This could add stability and balance to the side, and a balanced team can always be backed for better performance in the league, thereby satisfying the fans' anticipation.

## 5 Conclusions

Through opinion mining of the 15374 tweets during the players' auction in the IPL, this study highlights the importance of the fans' attachment to players. The context here was the IPL players auctions held in December 2019. The resulting word frequency tables and the wordclouds generated thereby, using the software R indicates that the fans were actively involved in the auctions celebrating the bids won by their respective teams, of their anticipated players. It was clear that the attachment points like sport attachment and city attachment were the driving factors along with the attachment to players that strengthened the team identification of the fans.

Further, a causal loop was developed using the attachment points as the epicentre to model the interrelationship among the various parameters influencing the event income and the team income. The causal loop can help the event organizers to understand the underlying structure of the IPL ecosystem. Further, the respective team management can target the right players to buy for their team so that the fans' team identification caused because of the attachment to players remains intact and favorable to the stakeholders involved, especially the sponsors.

## 6 Limitations and future work

This study has some limitations. Firstly, this study only considered opinion mining of the tweets extracted based on one-off auction in the IPL. Future studies can compare the tweets of two or more auctions and try to draw conclusions about the fluctuating attachment points. Secondly, the current study is limited to drawing the causal loop diagram. Future studies can simulate the model by drawing the stock and flow diagram whereby the system can be simulated using the mathematical equations governing the inter-relationships. The graphs that can be generated as the simulation outputs can highlight the importance of the various causal links. Further, future studies can attempt to model the players' auctions by modeling all the parameters that influence the fans and the team's decision to go for a particular player in the auction.

## References

- 1 Chambers, N., Bowen, V., Genco, E., Tian, X., Young, E., Harihara, G., & Yang, E. (2015, September). Identifying political sentiment between nation states with social media. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing* (pp. 65-75).
- 2 Choi, M., Park, M., Lee, H. S., & Hwang, S. (2017). Dynamic modeling for apartment brand management in the housing market. *International Journal of Strategic Property Management*, 21(4), 357-370.
- 3 Crescitelli, E., & Figueiredo, J. B. (2009). Brand equity evolution: a system dynamics model. *BAR-Brazilian Administration Review*, 6(2), 101-117.
- 4 Derczynski, L., Ritter, A., Clark, S., & Bontcheva, K. (2013, September). Twitter part-of-speech tagging for all: Overcoming sparse and noisy data. In *Proceedings of the International Conference Recent Advances in Natural Language Processing RANLP 2013* (pp. 198-206).
- 5 Fan, W., & Gordon, M. D. (2014). The power of social media analytics. *Communications of the ACM*, 57(6), 74-81.
- 6 Funk, D.C., Alexandris, K. and McDonald, H., 2016. *Sport consumer behaviour: Marketing strategies*. Routledge.
- 7 Heere, B. and James, J.D., 2007. Sports teams and their communities: Examining the influence of external group identities on team identity. *Journal of Sport Management*, 21(3), pp.319-337.
- 8 Heere, B., James, J., Yoshida, M. and Scremin, G., 2011. The effect of associated group identities on team identity. *Journal of Sport Management*, 25(6), pp.606-621.

- 9 Hsiao, C. T., Peng, H. L., & Huang, H. H. (2010). Dynamic simulation for development of the professional baseball game in Taiwan. *International Journal of Electronic Business Management*, 8(3), 173.
- 10 Jin, J., Liu, Y., Ji, P., & Kwong, C. K. (2019). Review on recent advances in information mining from big consumer opinion data for product design. *Journal of Computing and Information Science in Engineering*, 19(1).
- 11 Katz, M. and Heere, B., 2013. Leaders and followers: An exploration of the notion of scale-free networks within a new brand community. *Journal of Sport Management*, 27(4), pp.271-287.
- 12 Kwon, H.H., Trail, G.T. and Anderson, D.S., 2005. Are multiple points of attachment necessary to predict cognitive, affective, conative, or behavioral loyalty?. *Sport Management Review*, 8(3), pp.255-270.
- 13 Levine, R. L., & Call, O. (2003). Models of attitude and belief change from the perspective of system dynamics. In *Proceedings of the 21st International Conference of the System Dynamics Society, New York City, USA, The System Dynamics Society*.
- 14 Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1), 1-167.
- 15 Liu, Y., Huang, X., An, A., & Yu, X. (2007, July). ARSA: a sentiment-aware model for predicting sales performance using blogs. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 607-614).
- 16 Lock, D., Taylor, T. and Darcy, S., 2011. In the absence of achievement: The formation of new team identification. *European Sport Management Quarterly*, 11(2), pp.171-192.
- 17 Pak, A., & Paroubek, P. (2010, May). Twitter as a corpus for sentiment analysis and opinion mining. In *LREc* (Vol. 10, No. 2010, pp. 1320-1326).
- 18 Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, 2(1-2), 1-135.
- 19 Shapiro, S.L., Ridinger, L.L. and Trail, G.T., 2013. An analysis of multiple spectator consumption behaviors, identification, and future behavioral intentions within the context of a new college football program. *Journal of Sport Management*, 27(2), pp.130-145.
- 20 Sterman, J. D. (2000). Business dynamics: systems thinking and modeling for a complex
- 21 Stojanovski, D., Strezoski, G., Madjarov, G., & Dimitrovski, I. (2015, November). Emotion identification in FIFA world cup tweets using convolutional neural network. In *2015 11th International Conference on Innovations in Information Technology (IIT)* (pp. 52-57). IEEE.world (No. HD30. 2 S7835 2000).
- 22 Xu, K., Liao, S. S., Li, J., & Song, Y. (2011). Mining comparative opinions from customer reviews for competitive intelligence. *Decision support systems*, 50(4), 743-754.
- 23 Yoshida, M., Heere, B. and Gordon, B., 2015. Predicting behavioral loyalty through community: Why other fans are more important than our own intentions, our satisfaction, and the team itself. *Journal of Sport Management*, 29(3), pp.318-333.
- 24 Yun, J. T., Vance, N., Wang, C., Marini, L., Troy, J., Donelson, C., & Henderson, M. D. (2019). The Social Media Macroscopic: A science gateway for research using social media data. *Future Generation Computer Systems*.
- 25 Zhang, W., & Skiena, S. (2010, May). Trading strategies to exploit blog and news sentiment. In *Fourth international aAAI conference on weblogs and social media*.