# "How Was the Match?": Semantic Similarity between Electronic Media Commentary and Work Domain Analysis Key Phrases

Gustavo Silva[1][a], Ricardo Ribeiro[1,2][b] and Rui J. Lopes[1,3][c]
[1]*Iscte - Instituto Universitário de Lisboa, Portugal*
[2]*INESC-ID Lisboa, Portugal*
[3]*Instituto de Telecomunicações, Lisboa, Portugal*

Keywords: Football, Work Domain Analysis, Match Annotation, Performance Analysis, Semantic Similarity.

Abstract: Football player's performance can be measured in an objective way (e. g. Goals scored, assists, interceptions), this being seldom a method to compare and rank the best players by categories. Over years of study, many other factors that can influence the players performance were discovered and studied, considering not only objective factors, but also subjective factors. Match commentary from different sources (e.g., social and formal media) also plays an important role on a more subjective performance assessment.
By using semantic similarity analysis, this study aims to contribute to the understanding of the concepts that are used in this commentaries, notably to each extend key phrases associated to match processes are used in commentaries published in social and formal media.

## 1 INTRODUCTION

Team and athlete performance analysis has been an object of study and usage by practitioners (e.g. coaches) for several years. Methodologies, metrics, and studies have been designed to improve the performance of football players and provide a better performance analysis, typically in an objective way using notational analysis to account for several athletes actions (e.g., goals, assists, shoots). As illustrated in Figure 1, these actions are conditioned by many factors, notably the match context is a main factor in sports performance. In fact, there are many human and non-human components working dynamically and constantly changing the environment of a football match (Mclean et al., 2017).

The combination of human factors and football complexity makes performance analysis an extremely challenging task; advances in these studies provide an increasing number of factors that are considered to influence players' and teams' performance. On the other hand, the perceived performance, e.g., by fans or even specialised media, generally does not follow these procedures and metrics and is not expressed via objective metrics.

This study aims to explore the perception of fans and specialised media of the Work Domain Analysis (WDA) structure of football, and if it may exist a relation between objective performance approaches, their metrics and the subjective performance assessment expressed by fans in social media and specialised media on sports websites.

This paper is structured as follows: the next section addresses the related work on performance analysis; Section 3 presents the used information sources; Section 4 describes the methods used to process the data and compute the relation between the perceived performance and the levels of WDA; the achieved results are presented in Section 5; Final Remarks and Future Work close the document.

## 2 LITERATURE REVIEW ON PERFORMANCE ANALYSIS

There are many factors that can influence a match result, over the years, the researchers have been trying to analyse the complexity of this factors. Many aspects of human behaviour can be analysed, consequently, it is important to determine what will be analysed and the reason for this. It is important to

[a] https://orcid.org/0000-0002-4726-4981
[b] https://orcid.org/0000-0002-2058-693X
[c] https://orcid.org/0000-0002-8943-0415



Silva, G., Ribeiro, R. and Lopes, R.
"How Was the Match?": Semantic Similarity between Electronic Media Commentary and Work Domain Analysis Key Phrases.
DOI: 10.5220/0010691100003059
In *Proceedings of the 9th International Conference on Sport Sciences Research and Technology Support (icSPORTS 2021)*, pages 144-150
ISBN: 978-989-758-539-5; ISSN: 2184-3201

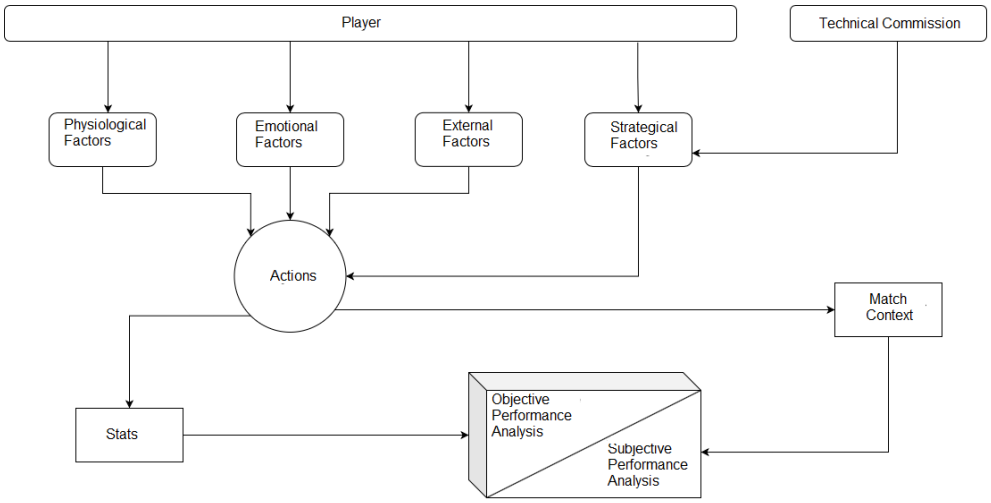



Figure 1: Player performance analysis in a match.

consider the saying: "not everything that counts can be counted, and not everything that can be counted counts" (C. Carling and Reilly, 2005). The sentence above defines what happen in a football match. Players can have a bad performance considering the stats (e.g., goals, assists, shots, interceptions), but still make a good match if we consider the match context, for example, a player who was positioned to avoid counter-attacks or marked a specific opponent player individually and played this role positively, both are hard aspects to measure, but that we cannot ignore when analysing the performance of a player.

## 2.1 Match Annotation

Annotations in football are an important tool to obtain intelligence in a match, even in a general performance of the team or an individual classification of a player. With the sports evolution on the last decades the need of more researches about the complexity of the evaluation of a player performance was found (Barros et al., 2018). In football scenario, even it being a sport very complex, it is possible to analyse the participation of a player in a match through predefined stats (e.g. shots, interceptions, assists, goals).

The match annotations are alternatives to analyse a player performance in a less empiric way, and to compare a player with other player based in a common stat. This form of analysis is used by sports TV channels, sports journalists, and bet sites. According to the Figure 2, the player participation in a match is converted in stats by analysis tools and used by coaches and clubs to assist the decision making.

## 2.2 Work Domain Analysis

Work Domain Analysis (WDA) is a system analysis method that aims to, in a structured mode, associate actors, their fundamental functions and resources used by themselves in a context based in the functional environment that establishes the purposes to be achieved.

In the football scenario, the whole squad has several common functions (e.g., positioning, connect passes, etc.), but each position has specific roles on a football match (e.g., a striker is the responsible for scoring the goals, the central back is the responsible to intercept the opponents and avoid opponents effective attacks). Based on a preview study, the author classified hierarchically a conceptual method to link the functions and purposes of players in a football match (Berber et al., 2020).

The structure is designed from specific components to general components, Each level is linked with the adjacent level based on the relation of the purpose and functions of the player position in a match:

**Functional Purpose.** The main functions of a player in a match (Prevent goals scored, Score goals, Relieve pressure, Create chances). Example: a striker has as main function to score goals.

**Values and Priority Measures.** Criteria used to analyse the progress of a player to achieve the functional purposes(Positioning, Goal Conceded, Saves made, Goals scored). Example: the quantity of goals a striker scored in a match.

**Purpose-related Functions.** Functions that need to be done to achieve the functional purposes(Defend, Attack, Leadership, Adaptability, Communication). Example: a striker has to

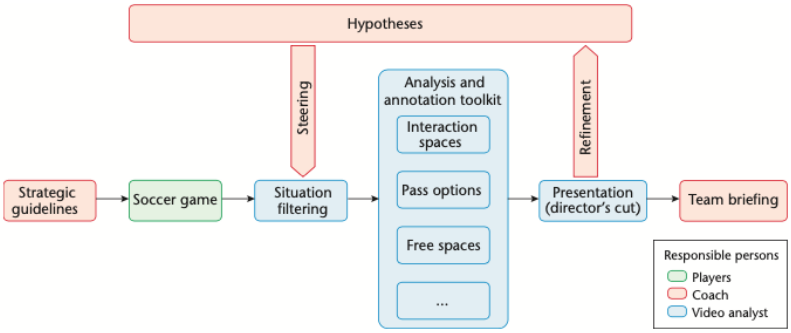



Figure 2: Typical analyse of extraction of Match Annotations in the professional football teams (Stein et al., 2016).

establish communication with the teammates to find the best way to score goals.

**Object-related Processes.** The process used by players to achieve a purpose-related function(Dive, Shooting, Break Lines, Free Kicks, Vision). Example: a striker has to pass, tackle, and kick to perform the purpose-related functions.

## 3 MATERIALS

Three different sources were used in this project as represented in Figure 3. Each source is a different perception of the same match, based on the context of each platform. Reddit comments are in an informal language, where essentially the author has an open space to talk anything he wants about a football match. In Formal Media, there are two sources: Live Match commentary and Player Ratings. The first one, Live Match, is the update, in real time, of the events in a match and the comment about the events as they happen. The Player Ratings comments are analyses about the general participation of a player in a match and the respective rating of the performance. Figure 3 illustrates the relationship between the used information sources and shows their number of items.

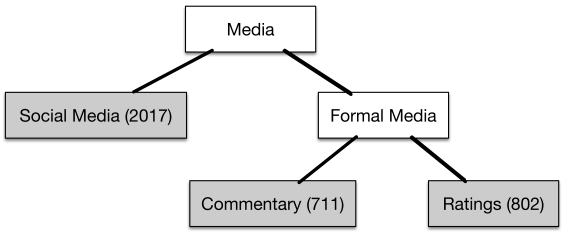



Figure 3: Commentary text sources.

In this present study Reddit was used for obtaining the social media content data. This platform aims connect users by grouping them in communities through the creation of rooms about topics, where users can comment and react to other's comments. In 2020, Reddit had over 52 million daily active users, nearly 303 million posts and two billion comments. Reddit is organised around the following concepts:

**Users.** who interact in Reddit. A user can comment and react in threads, follow other users and join communities.

**Communities.** Group of *Users* with common interests about a topic.

**Threads.** A room where users can interact about a given topic, for example, in a Football Match Thread, the principal objective is talk about football and related subjects.

**Comments.** A space designed for users to interact with other users or just to express opinion. *Comments* must respect the platform policies, but, the user is free to express his opinion using any language, including slang, emojis and hashtags.

Reddit offers an Application Programming Interface (API). By using it, we can retrieve the data through authenticated requests. Then, this data can be filtered and organised. In the end, it can be used to study topics or the sentiment associated to this kind of contents. Also, through Reddit API it is possible to make filtered searches about *Users*, *Communities* and *Threads*. Using the Reddit API, a dataset containing fans' comments and opinions on that match was created.

To obtain formal media contents, it was used web scraping, that consists in collecting data from web pages to obtain data, from different, in this case, sports sites.

As case subject, we used the match between Manchester City F.C. and Chelsea F.C. on the final of the UEFA Champions League 2021, that took place on May 29, 2021. In this match, Chelsea F.C. beat Manchester City F.C. by 1-0, winning the tournament.

# 4 METHOD

To understand the relation between the perceived performance by fans and specialised media, we compute the semantic similarity between the Reddit's posts, live comments, and players' assessments and the levels of WDA. We experimented different approaches and the best results were achieved by using BERT (Devlin et al., 2019) to generate computational representations of the textual data from fans and formal media and of the key phrases corresponding to several levels of WDA and the cosine (Eq. 1) to compute the similarity between these vectorial representations.

$$\text{sim}_{\cos}(\mathbf{x},\mathbf{y}) = \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|} = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2}\sqrt{\sum_{i=1}^{n} y_i^2}} \quad (1)$$

BERT stands for Bidirectional Encoder Representations from Transformers, which hints about its nature. BERT is a language representational model which uses context, left and right, to generate representations for raw text. This model is based on the concept of transformer, which is a neural network architecture that follows the encoder-decoder structure *using stacked self-attention and point-wise, fully connected layers for both the encoder and decoder* (Vaswani et al., 2017). In this work, we used DistilBERT (Sanh et al., 2019), a more efficient version of BERT, that achieves comparable results. As implementation, we used the Python Sent2Vec[1] package.

# 5 RESULTS

The method described in Section 4 was applied to the collected datasets described in Section 3. Specifically, we computed the semantic similarity, $sim_{ij}^{mn}$ between each entry (i.e., sentence), $s_i^m$ and key phrase, $k_j^n$ defined via WDA (here $s_i^m$ corresponds to the $i^{th}$ sentence of dataset $S^{m=1,\ldots,3}$, and $k_j^n$ corresponds to the $j^{th}$ WDA key phrase at level $L^{n=1,\ldots,4}$. This resulted in 12 matrices (three datasets and four levels), with values between 0 and 1, and presented as heat maps in Figure 4, where rows and columns correspond to entries (sentence) and WDA key phrases respectively. These results show a great dispersion of the similarity score across all domains: i.e., between sources, levels, and between entries from the same source at the same WDA level.

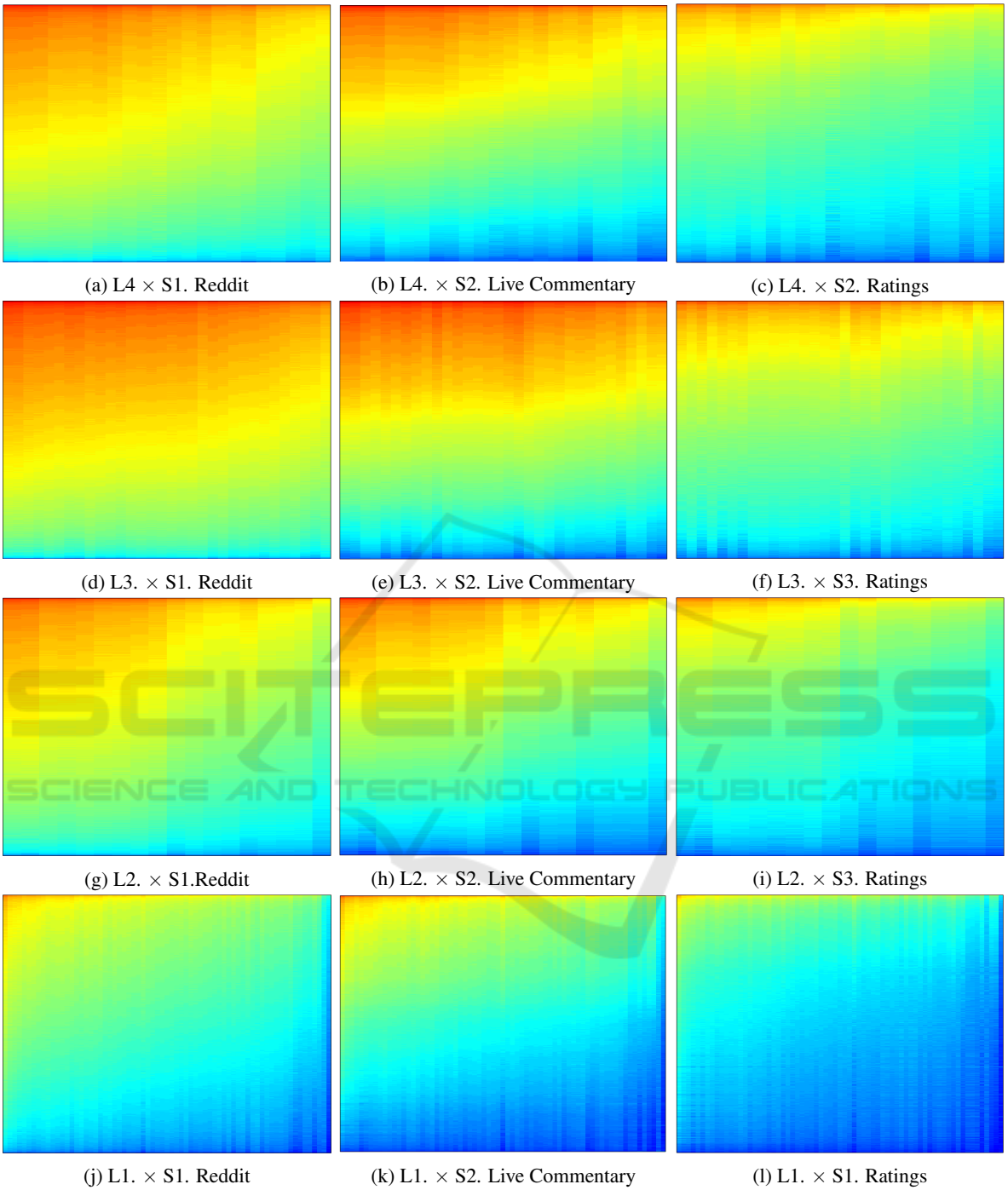
In order to assess how the similarity varied between levels and datasets we computed the similarity mean and standard variation for all the 12 level/dataset combinations. According to Table 1, the general content of all the data sources is more similar with the key phrases from the WDA level *L3. Value & Priority Measures Level*, which means that both, informal and formal media, tend to describe matches using an objective perspective more based on players stats and less based in their participation in more abstract processes (described in level L4). In contrast to this, the WDA level that has less similarity with the content of each information source is *L1. Object-related processes*, which means that in a general context, the comments are not about the secondary (i.e. "means-to-an-end") functions of a player in a match but about the objective performance and the principal functions (e.g., a striker has to score goals). On the other dimension, at all four levels, Reddit entries present the higher similarity values while Ratings present the smaller values. This is contrary to what was expected, that is, that formal media live commentary and player ratings would be more semantically similar to WDA key phrases than fan's comments on social media.

Table 1: Similarity of the different information sources with the WDA levels (mean and standard deviation).

| | S1. Reddit |
|---|---|
| L4.Functional purposes | 0.437±0.078 |
| L3.Value & priority measures | 0.468±0.078 |
| L2.Purpose-related functions | 0.411±0.078 |
| L1.Object-related processes | 0.326±0.075 |
| | S2. Live Commentary |
| L4.Functional purposes | 0.399±0.104 |
| L3.Value & priority measures | 0.427±0.107 |
| L2.Purpose-related functions | 0.378±0.099 |
| L1.Object-related processes | 0.300±0.089 |
| | S3. Ratings |
| L4.Functional purposes | 0.351±0.081 |
| L3.Value & priority measures | 0.378±0.083 |
| L2.Purpose-related functions | 0.330±0.077 |
| L1.Object-related processes | 0.253±0.067 |

We also investigated if the key phrases at each level would or not maintain their similarity rank across the different data sources. According to Table 2, the ranking of most similar key phrases is very similar in the three data sources, i.e., informal and formal media comments typically tend to comment the match based on the same key phrases. (Due to space limitations Table 2 only shows the top five and bottom two key phrases for each level.)

[1] https://github.com/pdrm83/sent2vec



(a) L4 × S1. Reddit (b) L4. × S2. Live Commentary (c) L4. × S2. Ratings

(d) L3. × S1. Reddit (e) L3. × S2. Live Commentary (f) L3. × S3. Ratings

(g) L2. × S1.Reddit (h) L2. × S2. Live Commentary (i) L2. × S3. Ratings

(j) L1. × S1. Reddit (k) L1. × S2. Live Commentary (l) L1. × S1. Ratings

Figure 4: Similarity score between entry and key sentence at different levels (*L4.Functional purposes, L3.Value & priority measures, L2.Purpose-related functions, L1.Object-related processes*).

Table 2: Comparison of rank and mean across layers and entity sources.

| Key ID | Key phrase | Reddit Rank | Reddit Mean | Live Commentary Rank | Live Commentary Mean | Ratings Rank | Ratings Mean |
|---|---|---|---|---|---|---|---|
| L4.22 | Assist in goal scoring | 1 | 0.470 | 3 | 0.426 | 5 | 0.376 |
| L4.8 | Create goal scoring opportunities | 2 | 0.470 | 1 | 0.430 | 1 | 0.381 |
| L4.20 | Bring others into offensive situations | 3 | 0.468 | 2 | 0.429 | 2 | 0.378 |
| L4.7 | Break up opposition attacks | 4 | 0.461 | 5 | 0.423 | 4 | 0.376 |
| L4.3 | Provide a safe passing option | 5 | 0.461 | 6 | 0.423 | 6 | 0.375 |
| ... | | | | | | | |
| L4.2 | Initiate build-up | 21 | 0.394 | 22 | 0.359 | 22 | 0.311 |
| L4.10 | Stretch opposition | 22 | 0.393 | 20 | 0.364 | 21 | 0.318 |
| L3.32 | Goals scored | 1 | 0.493 | 1 | 0.446 | 1 | 0.398 |
| L3.21 | Runs without the ball | 2 | 0.492 | 2 | 0.446 | 3 | 0.394 |
| L3.13 | Effective defensive clearances | 3 | 0.487 | 5 | 0.441 | 5 | 0.391 |
| L3.3 | Goals conceded | 4 | 0.484 | 7 | 0.438 | 7 | 0.389 |
| L3.10 | Effective contests | 5 | 0.484 | 3 | 0.445 | 2 | 0.397 |
| ... | | | | | | | |
| L3.18 | Block shots and crosses | 31 | 0.436 | 32 | 0.394 | 32 | 0.346 |
| L3.7 | Interceptions | 32 | 0.432 | 30 | 0.399 | 30 | 0.352 |
| L2.7 | Maintain position in team structure | 1 | 0.456 | 1 | 0.413 | 1 | 0.365 |
| L2.15 | Play in line with coach ethos | 2 | 0.455 | 2 | 0.411 | 2 | 0.363 |
| L2.13 | Appropriate decision-making | 3 | 0.444 | 3 | 0.407 | 3 | 0.361 |
| L2.18 | Manage own fitness physical condition | 4 | 0.443 | 5 | 0.406 | 5 | 0.358 |
| L2.12 | Maintain resilience | 5 | 0.442 | 4 | 0.406 | 4 | 0.360 |
| ... | | | | | | | |
| L2.5 | Communication | 17 | 0.370 | 17 | 0.343 | 17 | 0.296 |
| L2.1 | Defend | 18 | 0.344 | 18 | 0.320 | 18 | 0.271 |
| L1.62 | Recognise when and how to support team members | 1 | 0.410 | 1 | 0.374 | 1 | 0.330 |
| L1.30 | Recognise/anticipate team member actions | 2 | 0.390 | 2 | 0.355 | 2 | 0.311 |
| L1.15 | Initial distribution of the ball | 3 | 0.385 | 3 | 0.353 | 3 | 0.303 |
| L1.26 | Organise team members at opposition set pieces | 4 | 0.379 | 5 | 0.342 | 4 | 0.294 |
| L1.21 | Provide protection from injury | 5 | 0.373 | 4 | 0.343 | 5 | 0.293 |
| ... | | | | | | | |
| L1.34 | Understand coach's intent | 68 | 0.243 | 68 | 0.229 | 66 | 0.202 |
| L1.65 | Risk-taking | 69 | 0.200 | 69 | 0.188 | 69 | 0.144 |

# 6 FINAL REMARKS AND FUTURE WORK

In the work described in this paper it was possible to explore how key phrases associated to different levels of Work Domain Analysis are used in football matches commentary published electronically by different sources. From this exploratory work the following conclusions could be obtained:

- The similarity score between commentary entries and WDA key phrases shows a great dispersion across all domains (sources, levels, and entries);
- The higher similarity values are obtained at the WDA level *L3. Values & priority measures*. It is worth of note that the key phrases identified at this level have usually a closely related match annotation item (e.g., Goals scored);
- Contrary to what may be expected, comments from users in social media show, for all WDA levels, higher semantic similarity values that commentary entries in formal media.

Concerning future work, we foresee six main ideas on how to increase the potential of this project:

- The informal and formal media have close similarity scores, with higher similarity values being achieved by fans comments - it is important to understand how this conclusion generalise to other matches;
- Perform a more comprehensive study of the different key phrases, notably their relative ranking and their potentially hierarchical structure (e.g., Goals and Goals scored/conceded or Runs and Runs with/without the ball).
- Sentiment polarity of fans perspective can provide unanticipated insights concerning performance analysis of football players. Sentiment analysis captures the subjective part of performance, and analysis based on metrics (stats about passes, goals, and assists, for example) its objective part;
- Apply our method to other social media platforms and sources of formal media commentary notably, comparing how users behaviour in different platforms;
- Try to adapt the used methods to other sports;
- The creation of a specific platform to connect football fans and the Data Department of the Football Teams could lead to an integrated (qualitative and quantitative) perspective on performance analysis.

# ACKNOWLEDGEMENTS

Rui J. Lopes was partly supported by the Fundação para a Ciência e Tecnologia, under Grant Number UIDB/50008/2020 attributed to Instituto de Telecomunicações. Ricardo Ribeiro was partly supported by national funds through Fundação para a Ciência e a Tecnologia (FCT) with reference UIDB/50021/2020.

# REFERENCES

Barros, B., Serrão, C., and Lopes, R. J. (2018). Distributed crowd-based annotation of soccer games using mobile devices. In *Proceedings of the 6th International Congress on Sport Sciences Research and Technology Support - Volume 1: icSPORTS,*, pages 40–48.

Berber, E., McLean, S., Beanland, V., Read, G. J. M., and Salmon, P. M. (2020). Defining the attributes for specific playing positions in football match-play: A complex systems approach. *Journal of Sports Sciences*, 38(11–12):1248–1258.

C. Carling, A. M. W. and Reilly, T. (2005). *Handbook of soccer match analysis: A systematic approach to improving performance*. Routledge, 1st edition.

Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proc. of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4171–4186.

Mclean, S., Salmon, P., Gorman, A., Read, G., and Solomon, C. (2017). What's in a game? A systems approach to enhancing performance analysis in football. *PLOS ONE*, 12:e0172565.

Sanh, V., Debut, L., Chaumond, J., and Wolf, T. (2019). Distilbert, a distilled version of BERT: smaller, faster, cheaper and lighter. *CoRR*, abs/1910.01108.

Stein, M., Janetzko, H., Breitkreutz, T., Seebacher, D., Schreck, T., Grossniklaus, M., Couzin, I., and Keim, D. A. (2016). Director's cut: Analysis and annotation of soccer matches. *IEEE Computer Graphics and Applications*, 36(5):50–60.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, u., and Polosukhin, I. (2017). Attention is all you need. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, page 6000–6010.