PERCEPTION MODEL TO ANALYZE FOOTBALL PLAYERS' PERFORMANCES

Usman M. Khan, School of Electrical Engineering and Computer Sciences, National University of Science and Technology, Islamabad, Pakistan, usmanwardag@gmail.com

Zain Kabir, School of Electrical Engineering and Computer Sciences, National University of Science and Technology, Islamabad, Pakistan, 13beezkabir@seecs.edu.pk

Asad W. Malik, School of Electrical Engineering and Computer Sciences, National University of Science and Technology, Islamabad, Pakistan, asad.malik@seecs.edu.pk

Abstract

In this paper, we presented a technique to determine player perception (a performance measure) through their online presence (media feedback), which is far more accurate than one determined by traditional means of looking at player statistics. The proposed model is a research-in-progress that takes into account the factors such as toughness of the fixtures, toughness of opponents, the character of players, and et cetera. Feedback from multiple sources such as blogs, newspaper articles and editorials are analyzed to assign a rating to each player that is indicative of their performance on and off the field. A comprehensive perception model is designed for this purpose to assign scores to each feature of performance to not only get an idea of the overall impact of player but also get a sense of how well the player performed in each aspect of the game.

Keywords: Perception, Football, Online, Player, Data

1 BACKGROUND

The recent advancement of football analytics tools is down mainly to two factors: 1) the advancement of data analysis techniques, and 2) in-depth record of football statistics. These analytics tools are vastly employed by clubs (both amateur and professional) to measure players' performances, analyze the over-valued and under-valued skills (as in the case of Money Ball), and help club management buy the most effective players in a particular budget range (Severini 2014). Various past and present studies continue to focus on developing algorithms and techniques to better assess players' performances. However, despite the advancement in football analytics techniques, this approach has some key shortcomings: 1) Coaches are not always interested in statistical figures but often in the character and mental toughness shown by a player. For instance, one renowned football coach argues that it is the psychology of a player that is far more important than his actual skill level (Lewis, 2014).

Things like how a player reacts to a goal, how his team-mates see him, how he celebrates, are significant, 2) The statistical techniques fail to take into account the factors such as quality of opposition, toughness of fixtures and context of games, and 3) Data are not widely available. One, there is only a handful of companies who keep track of this data and second, for the analysis to be anywhere near successful, the dataset has to be really complex. *Opta*, a leading provider of football data, covers as many as 1500 events across every football match.

2 INTRODUCTION

The idea behind the model developed is to assess player perception through subjective feedback of media. In all of the top-flight leagues across Europe, players get regular feedback from journalists, bloggers, coaches and performance analysts. The feedback is either positive or negative depending on the performance of players. The model uses that feedback to (1) determine if the performance of a particular player was positive, negative or neutral, (2) determine the performance of player in individual aspects of the game such as defending, attacking, passing, creativity and the like, and (3) classify players based on a set of qualities such as tough, physical, creative, incisive and a set of other adjectives.

2.1 Model Rationale

The rationale of this model is to look at player performance from a novel perspective of subjective analysis. The challenge of such a model is to deal with the bias that is inherent in subjective analysis. Person (a), for example, might have a liking for a player and he might rate him high despite a sub-standard performance. Another person (b) might dislike that player and consequently might rate him low despite an exceptional performance - such bias can never be completely eliminated but can be minimized to a considerable extent by picking a large dataset that contains well over hundreds of feedback articles.

As such, there are a lot of conflicts and irregularities in algorithms currently used to measure performance. Castrol Index, Capello Index and WhoScored index are some of the widely used rating systems that have not made their algorithms public (Kumar 2012). As an example of the irregularities of these systems, some algorithms correlate 'distance covered' directly to good performance while others correlate it to nothing significant at all. Then, there are some that say that pure distance does not predict anything but distance covered while sprinting might be a good predictor of player performance (Frencken et al. 2012).

The success of these traditional models is highly *situation-specific*. Each team has a different style of play and therefore, one performance metric cannot fit all of them. A team with a defensive mentality would stress more on conceding less than scoring more. Another team with an attacking approach would do the opposite. It is, therefore, impossible to compare the players (strikers, for instance) of these two teams without considering their contexts.

Finally, there is a widespread consensus that *football is an art* (Briles et al. 2014). Although statistics can help quantify the performance of a player to some extent, they do not present a complete picture. Guus Hiddinck argues "*I am more interested in the actual route the player runs and less in the pure distance"*. Therefore, an objective analysis of a player performance without relevant context is often meaningless.

2.2 Model Edge

First, the feedback in the proposed model is not based solely on some statistical figures such as goals, assists or the like but on overall impact the player had on the passage of the game. This is much in line with how man of the match performance (MOM) is decided. A group of analysts give a subjective opinion on the impact each player had on the proceedings of the game and that decides the award. The current model collects the opinions of analysts, bloggers, coaches, experts from all over the world and combines them to assess the overall impact.

Second, the model takes into consideration the performance context. In general, when assessing the performance of midfielders, pass accuracy is one of the most important metrics. Consider a scenario where a player (a) attempts 100 passes and get 90 of them right. Another player (b) also attempts 100 passes and get only 79 of them right. Statistical models would rate player (a) higher because of a greater passing accuracy. In many cases, however, the player (b) should receive better rating because he delivered more adventurous passes. This is just one of the downsides of using statistics to assess player performances. There are other cases where factors like the toughness of fixture, quality of opposition, and et cetera are not factored in at all. However, a subjective analysis of the game would take these factors into account and thus, give a better assessment of player performance.

Third, the model also supports classification of players into different groups such as creative players, physical players, attack-minded players and et cetera. This allows teams to select players who would fit club's philosophy and combine well with other members of the team. For instance, creative midfielders who specialize in delivering *through* balls require a striker up front who is willing to make long runs. A striker who has a lot of pace would fit perfectly in this system. On the contrary, teams that play a lot of short passes and dominate possession would require more of a physical striker.

Fourth, the model keeps track of both on and off-the-field profiles of players and is, therefore, a handy indicator of players' personality. For example, if a player is involved in off-the-field controversy, it would get noticed by media and would, therefore, affect his rating. So, the player performance is measured both in terms of how well he performs on the pitch and behaves off-the-field.

Finally, data required for this model are simple information on the web which is accessible to everyone. The current model uses *Google News* as a channel to get access to a vast array of information.

2.3 Examples of Subjective Feedback in Football

Even with the advance of football analytics tools, players' performances are still largely measured by subjective means. For example, Man of the Match (MOM) awards are decided by a group of experts who consider the context of the match and then assess the impact of each player.

Similarly, in tournaments such as FIFA world-cup, the awards of best player, best goal keeper and et cetera are decided through subjective feedback by a group of people that includes coaches and experts from all over the world. Finally, Ballon d'Or, often deemed as the biggest individual award in football, is decided by football coaches, players and media personnel from all football-playing nations.

All these examples use subjective means to assess player performance and indicate the shortcoming of statistics when it comes to measuring the impact of individual players.

3 RELATED WORK

A lot of work has been done in the field of perception analysis, most of which revolves around analyzing the sentiments of users towards products/brands. For instance, Tumasjan et al (2010) developed a model that predicted election results and popularity of candidates based on the activity of users on social media. The key idea behind such models is to mine the opinion of users and determine their political affiliation. The perception model is also popular with brands such as Apple and Microsoft who seek user feedback on a regular basis (Kao 2007).

Sentiment analysis itself is a complex domain that has attracted a lot of attention over the past few decades. Researchers have devised various natural language processing algorithms (Liu 2012) but efforts are still being made to increase the accuracy of those models.

4 PROPOSED MODEL DESCRIPTION

When a query about a player is made, the model begins by accumulating data related to that specific player. The data is then mined and processed to indicate either positive or negative feedback, which is indicative of player's performance. The model works in following steps:

4.1 Data Collection

The data is scraped from *Google News* which contains a vast array of media reports linking to a query. Whenever a search is made about a player, the top articles linked to that player are retrieved. The number of articles retrieved may vary depending on how recent the data should be. For instance, if the requirement is to analyze player performance over the past day only, fewer articles should be collected. However, a greater number of articles will give a better model in terms of accuracy of results.

A Google News search of a player may, however, show results other than the performance of player. Therefore, the player query is modified by adding the keyword *"performance"* at the end of it. For instance, when a search of Ronaldo was performed on Google News, the first article was titled "*Ronaldo wants to stay with Real Madrid*". However, when the query was modified by adding performance term at the end, the first article was titled "*Breaking down Christiano Ronaldo's performance for Real Madrid vs. Espanyol*". Finally, Google News is a good choice since it contains the most recent information and, therefore, gives an accurate assessment of the player's current performances.

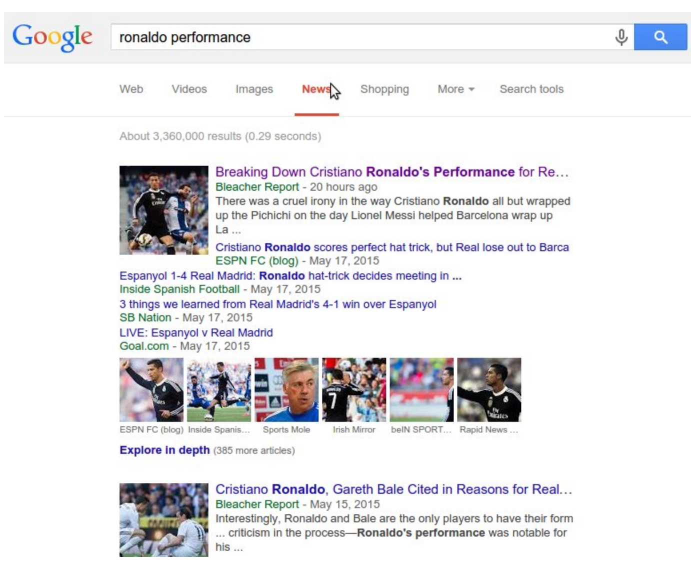

Figure 1. Example player page of "Ronaldo Performance" on Google News that contains links of latest stories and blogs

4.2 Selection of Relevant Data

When a query of a specific player is performed on Google News, it may contain news items on other players' performances as well. In order to address this problem, the text between the <title> tags of HTML of the news page is parsed. There are two criterions to ensure the selection of relevant data:

i- The title text must contain the name of player searched.

ii- The title text must not contain the name of any player other than the player searched. To achieve this objective, all proper nouns in the title text are extracted and matched against a database of player names. If any match is found, the news item is ignored.

4.3 Pre-Processing

Before the data can be used for any reasonable analysis, some pre-processing operations have to be carried out. Multiple cleaning operations such as removing punctuations, lowercasing the text and et cetera are performed in this step.

4.4 POS Tagging

We identify POS (parts of speech) in each sentence. This is done in R using *openNLP* package which assigns POS tags to each word in the sentence to a very good accuracy.

For example, the sentence "Ronaldo plays football" will return "Ronaldo._*NN* plays_*VBP* football_*VBP*".

4.5 Performance Feature Identification

In this step, the frequency of each feature is calculated to identify performance features, which are defined as follows. *A noun is a performance feature if it has a frequency of at least 1% in the complete text.*
In football texts, the performance features could be a variety of different nouns such as attack, dribble, defence, strength, physicality and et cetera.

4.6 Opinion Sentence Identification

After identification of performance features, the model searches for opinion sentences which are defined as follows. *A sentence is an opinion sentence if it contains at least one feature and one or more adjectives.* As an example of opinion sentence, if "attack" is a performance feature, then the following will be an opinion sentence.
"*Real Madrid's attack was lethal all evening".* The sentence meets the criteria of opinion sentence as it contains at least one performance feature (attack) and one adjective (lethal).

4.7 Keyword Identification

After the text is narrowed down to opinion sentences only, keywords are identified. To perform this step, we scan through all the adjectives in the sentence. *An adjective is a keyword if it is a part of an opinion sentence.*
In the above example of opinion sentence "*Real Madrid's attack was lethal all evening",* lethal is the only keyword present.

4.8 Weight-age Assignment

In this step, weight-ages are assigned to all adjectives in opinion sentences. The adjective closest to feature gets a weight-age of 1. All other adjectives get a weight-age decrementing by a value of 0.1, depending upon their location in the sentence.

4.9 Score Calculation for Features

In this step, keywords are matched against positive and negative words' librariers to determine the emotion they convey. Each feature score is then calculated based on weight-ages of keywords. If scores of adjectives are given by S_1, S_2, ... up to S_n and weight-ages by W_1, W_2, ... up to W_n, then score for each individual feature is calculated as:

```
for k from 0 to n
        product = S_k * W_k ;
        score = score + product;
end
```

4.10 Overall Score Calculation

The scores of all performance features are aggregated in this step to calculate overall player performance. If there are 'n' features in the text whose scores are represented by Sn, then overall score is calculated by taking the average of n feature scores as follows:

Overall score = $\sum(Sn) / n$

4.11 Scores in Independent Sections

To identify performance in independent sections such as 'attack', 'defense' and 'midfield', a synonym search (example in table 1) is performed to put similar features in categories. This is done by constructing a synonym table using chained hash tables in which pre-defined keywords are linked with their synonyms. Suppose, there are N independent features in which player's performance needs to be analyzed, then the chained hash table is given as follows:

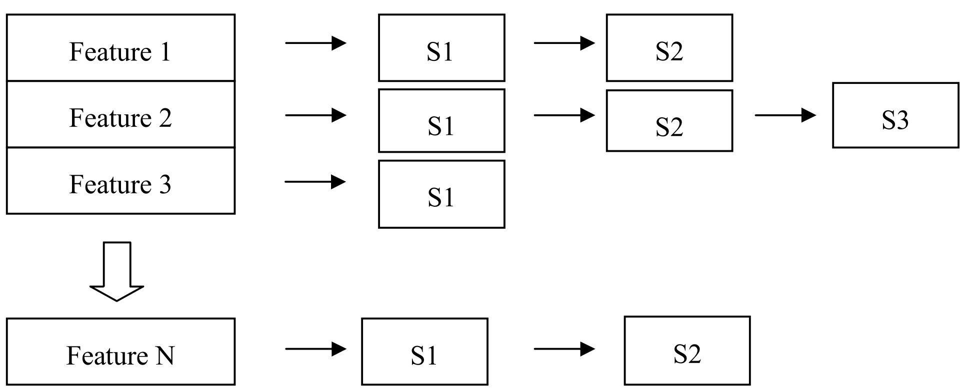

Figure 2. Chained Hash Table Model to represent synonyms for each performance feature which are used to find players' aggregate score in specific categories

Feature	Synonyms
Pace	Speed, Stride
Defence	Protection, Guarding
Attack	Onslaught, Offensive, Charge
Duel	Contest, Encounter

Table 1. *Random Entries from Performance Features Synonym Table*

Afterwards, the scores of all synonyms for respective features are aggregated to get an expert analysis of player performance in independent sections of the game.

4.12 Visualizations

Multiple visualizations can be performed in order to better able to make sense of the data. For the current model, Google Motion is used which gives real-time coverage of the events. In addition, word clouds are generated to get an idea of keywords a player is associated with.

Figure 3. *Word Cloud of Player showing the different adjectives a player (Steven Gerrard) is associated with on web media*

5 ALGORITHM EVALUATION

Evaluating the effectiveness of the proposed model is a challenging problem. A part of the problem is derived from the fact that current ranking models and performance indices are based largely on statistical events. For example, *Premier League Performance Index* is based on six different indices including winning performance, player performance, appearances, goals scored, assists and clean sheets. Similarly, other rating systems such as Castrol Index, Capello Index and WhoScored index are based solely on player statistics (Kumar 2012).

Some of the notable football events where a subjective ranking system is used are FIFA Ballon d'Or award, FIFA Puskas Award and MOM (Man of the Match) performances. The former two are awarded once in a year and are, therefore, not feasible reference points to determine the effectiveness of proposed model. However, MOM performances can serve as a benchmark to compare the proposed subjective method with the traditional statistical ones (such as Hiscock, Dawson et al. 2012).

In future studies, best performers for each competitive match across the different leagues in Europe can be determined through both the proposed scheme and existing statistical models and can then be compared with the actual MOM performances awarded. This would help draw a comparison between the two performance measurement approaches.

6 APPLICATIONS

While the model can be used as an effective means to measure player performance, it finds applications in a lot of other domains as well. For example, brands tend to do surveys and background checks to see if a player or group of players meet their ethos and are in line with their product vision. The model can help narrowing down such players.

The model can also be used to make links between players and explore chemistry of a group of players. This can be useful in learning how well a particular player will interact with a set of other players. Teams can, thus, go in the transfer market and select players that are likely to form a good combination with the already existing members of the team.

Finally, since each aspect of a player's game is analyzed, coaches can learn if a player would meet the kind of system they play in. For instance, some teams have a counter attacking approach that requires players to defend well and then release bursts of pace when ball is recovered. Accordingly, the selection staff can select players who fit the criteria of counter attacking football.

7 CONCLUSION

In this paper, we have proposed an efficient and accurate way of measuring player performance, which is through the online media presence of players. The objective of this model, which is a research-in-progress, is to adopt a subjective approach that takes the context in light while measuring player performance. Conventional models which use advanced statistics to evaluate player performance have three key shortcomings. First, they do not throw significant light on mentality and character of a player. Second, they ignore some key information such as context of game, quality of opposition and the like. Last, but not the least, they rely on expensive set of data. In proposed model, we presented a subjective approach that amasses opinions of football coaches, experts, journalists and bloggers and uses natural language processing techniques to make a sense of feedback and classify it as positive or negative. A detailed algorithm is then employed to calculate both overall scores and scores in individual departments such as attack, defense, creativity and the like.

References

Briles, A. And Yeager, D. (2014). Beating Goliath: My Story of Football and Faith. St. Martin’s Press.

Frencken, W. et al. (2012). Variability of inter-team distances associated with match events in elite-standard soccer. Journal of Sports Science. 30(12): 1207-1213

Hiscock, D., Dawson, B., et al. (2012). Game movements and player performance in the Australian Football League. International Journal of Performance Analysis in Sport 12(3): 531-545.

Kao, A. (2007). Natural Language Processing and Text Mining. USA: Springer Science & Business Media

Kumar, G. (2012). Machine Learning for Soccer Analytics. KU Leuven

Lewis, T. (2014). How Computer Analysts took over British Top Football Clubs. [online]. Available at http://www.theguardian.com/football/2014/mar/09/premier-league-football-clubs-computer-analysts-managers-data-winning

Lin, B. (2012). Sentiment Analysis and Opinion Mining. Morgan & Claypool Publisher

Opta. http://www.optasports.com/

Package openNLP. http://cran.r-project.org/web/packages/openNLP/openNLP.pdf

R: The R project for statistical computing. http://www.r-project.org

Severini, T.A. (2014). Analytic Methods in Sports: Using Mathematics and Statistics to Understand Data from Baseball, Football, Basketball, and Other Sports. CRC Press

Tumasjan, A. Sprenger, T. O. Sandner, P. G. And Welpe, I. M. (2010). Predicting Elections with Twitter: What 140 Characters Reveal about Political Sentiment. Technische Universität München, Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media