Decision Support System for Soccer

MSIS 637 – Decision Support System

December 1, 2020

Jon Constant & Shashank Kala

Table of Contents

Abstract

Soccer is getting more competitive every day, and this has led to the rise of analytics in this sport. Many major clubs in the world use sophisticated softwares to analyze the performance of the team and develop the strategies to have a competitive edge over their opponents. The aim of this research paper is to study the decision support system being used by various teams all over the world, comparing these methods and also suggesting the improvements which may prove beneficial to the team. Decision making is a basic component of any game, particularly open, quick, unique group activities, for example, volleyball, football, soccer etc. To prevail with regards to dominating any match and rivalry at public and worldwide level in all individual and group activities there is a need to reexamine all achievement factors to settle on a superior choice to win. At top level mentors and competitors appear to reliably use sound judgment in circumstances that are profoundly transiently compelled.

Introduction

Soccer is the most popular sport which is played in nearly 200 countries all over the world, hence it's important that teams should play with some strategy, hence there has been a rise in the use of analytics.[1] The starting points of soccer analytics can be traced back to Charles Reep, an Englishman, war veteran who did the initial analysis during one game in 1950. Reep got resentful about a group's disgraceful scoring endeavors and started recording perceptions and patterns he found in a journal. In that game, he noticed the huge proportion among attacks and goals, presuming that even an unassuming lift in scoring productivity would bring about an extra goal for each game. Reep built up a framework to document spatial data for each play, rapidly jotting as the match went on, supposedly going through 80 hours after each game.

Whether people knew about it or not, data analytics in sports has always been a component of every game. Although, what is sports analytics? According to Masoud Nikravesh of *Experfy*, sports analytics is [2] the processes that identify and acquire the knowledge and insight about potential players' performances based on the use of a variety of data sources such as game data and individual player performance data. These advanced and sophisticated types of analytics should be able to extract valuable actionable insights for the coaches and managers to utilize. With the addition of sports analytics in today's sports, it makes it easier for coaches and club managers to make adjustments. There are a number of domains in which sports analytics can be utilized throughout sports. First and foremost, they are used to predict scores of the game before it occurs. In soccer, this is the hardest task asked for someone to complete because there are many factors that go into soccer that can deter this prediction. Secondly, sports analytics are utilized to predict the performance of a club or individual player. Depending on a team's opponent, a club could be projected to outscore their opponent 4 to 1 just because they are purely

better. A third reason they are used in soccer is to come up with new strategies and formations for different opponents because success does not come in the exact same way every game. A coach or club manager utilizes sports analytics to analyze where their team's strengths and weaknesses are in order to practice more on those weak points, as well as to sharpen those strong points. That's how good teams take the necessary steps to becoming a great team. A fourth way soccer clubs use sports analytics is to gauge a players "plus-minus" in order to find out how much they really contribute to the team's success, which then makes it an easier decision for managers to sign/resign or release a player. There are many other reasons as to how sports analytics are beneficial to clubs, as well as spectators.

Sports Analytics: An overview

After learning about some basics as to what sports analytics can be used for, who in fact uses this information? The top three members aside from superfans who dive deep into this data are club managers, coaches and players. Club managers and coaches use this data to assess what players play efficiently and effective. This is a key success factor for clubs to always attempt to improve because all it takes is a club manager or a coaches decision to begin the game with a certain player or in a specific formation for the team to react to that change as a whole. Although, believe it or not, players are honed in on this data more than people imagine. One might ponder, why do players look into this data the most? This is because there are approximately 250 million soccer players across the globe that would kill to be at the professional level in their cleats, so they put a lot of pressure on themselves to perform to the best of their abilities because if they don't, they will get their spot taken. Each club has a pipeline of recruited players ready to step up to the play in the major league because these players earn millions of dollars and this sport is not taken lightly worldwide. The United States is not as big for soccer players because they have a

number of other professional sports that are prospering just like soccer is in foregin countries. This is why soccer is played throughout the world because outside of the United States, it is without a doubt the most successful sport in terms of stability and opportunity to make a name for yourself and hopefully make millions of dollars. That is why players look into sports analytics the most while club managers and coaches take a secondary role in this information because again, their job is not directly on the line. The players are the ones who are on the field fighting for the club while the coaches need to implement formations and strategies.

Although, club managers, coaches and players are not the only ones who look into sports analytics. News outlets, sports analysts, fans and gamblers also use this data. News outlets and other sports analysts utilize these statistics for obvious reasons. News outlets use them to update people that an upcoming game is about to occur or rely on the results of the game. Sports analysts use this data to predict scores and call games. It's important for announcers to be educated on the game because no person wants to watch a game with an announcer who is mispronouncing names and using the wrong terminology. Announcing a soccer game is one of the hardest sports to call because it's the fastest game on two feet, so the ball is non stop moving along with the fact that there are 22 players on the field at once, which they should know how to pronounce their names and statistics about them. Fans and gamblers use sports analytics for the same reasons. They both want to understand and grasp an idea of which team has the best chance of winning the match prior to it starting, as well as while it's occurring. Some fans are just plain old fans for many reasons, but "superfans" and gamblers know the statistics behind players and how specific clubs match up against each other. Many people would like to know for example, how many goals Ronaldo has scored against Italy in his career during the World Cup compared to other teams to find out if he has had success against them in the past. That's simply one of

thousands of scenarios in which superfans and gamblers love to know prior to games as a sense of knowledge. These types of statistics are why sports analytics is essential knowledge in today's sports.

Soccer Tactics

In soccer there are many different statistics that are kept track of by people and softwares. The reason soccer needs to use software to keep track of statistics unlike many other sports is because of how fast it is. The ball is constantly moving and so are the players, unless there is a foul. There are team statistics and individual player statistics. The first and most important statistic in soccer is a club's win/loss record and whether they are on the winning streak or a losing streak. Another statistic that is piggybacked off of that is a club's record when playing home or away and what is their winning percentage when home or away. The second most crucial stat is how many goals a team scores and lets in because in order to win a soccer game you have to put the ball in the net. In addition to that stat, a third important statistic is how many shots a club is taking and how many they are giving up. More often than not, the team who takes more shots on goal will win the soccer game unless a team's shooting accuracy is really high. A fourth important stat is how many yellow or red cards a team gets per game.[3] A yellow card is issued as a caution, giving a player a second chance to remain on the pitch for the rest of the game, but being shown a red card results in immediate ejection from the field of play. If a player is cautioned twice in a game with two yellow cards they are then shown an automatic red card. And in that case, the player is disqualified from the match and forces their team to play "man down" for the remainder of the game.

From an individual standpoint, there are rather tedious statistics that are recorded more often by software systems due to the game being too fast for human capabilities. In addition,

statistics are more important than others depending on the position a player plays on the field. There are four main positions on the field which include goalkeeper, defenders, midfielders and attackers. Each of which have top priorities that are different in terms of how they approach the game. Goalkeeper is like many other sports where their main priority is to stop the opposing team from scoring in any way possible. As a goalkeeper in soccer, they can use any part of their body to deflect the ball from going into the goal. That is a goalies number one priority and how they get evaluated. Their save percentage tells a lot about them. In addition, they have to be effective at communicating things like where the ball is and if an opposing player is open because they must organize the defense. Due to the fact that goalkeepers make saves and have the ball in their hands it's crucial that they give good outlet passes and take care of the ball. The ways in which a goalkeeper is statistically evaluated is based on how many goals they let in, shots saved, passes on target and fouls. Goalkeeper is an extremely hard position to play because the goal is very big and without defenders who are going to lay it all out on the line, it's tough to be a superior goalkeeper.

With that being said, the hardest position on the field is being a defender because of the skill people in today's game have with the soccer ball. Defenders keys to being effective are preventing as many shots as possible, stealing the ball from opposing players, transferring the ball to midfielders and marking up opponents. Without those habits as a defender, no one will succeed at the professional level. There are a number of components to the game that defenders get statistically evaluated on which includes balls stolen from opponents, successful challenges from opposing players, passes on target, fouls, throw-ins, shots and how many times they got beat one-on-one. Tying zero-zero is better than giving up a goal and losing one-nill. In fact, it's most likely a defenders fault if a team gives up a goal.

As mentioned earlier, soccer is the fastest game on two and is played on [8] a large field, with a fast moving ball and rare substitutions meaning soccer players can expect to log some heavy mileage over 90-plus minutes. Midfielders tend to run the most, sometimes reaching nearly 9.5 miles, according to *SportVU*. Midfielders do the most running because they play both offense and defense. Therefore, these players are generally extremely athletic and have a high soccer IQ. Four keys responsibilities in which midfielders must be effective in order to be elite at their position is stealing the ball and redirecting it to the attackers, taking shots on goal, accurate passes and pressuring the ball. More often than not, midfielders will not score the ball, but will assist on a goal because they do the dirty work in the middle of the field, hence why they are called mid-fielders. Therefore, midfielders performance gets evaluated on many major statistics which include passes on target/corner kicks/cross-balls, dribbles without losing the ball, assists, goals, takeaways, throw-ins, fouls and how many times they got beaten. There are a lot of aspects of the game that midfielders need knowledge of especially because decisions with and without the ball need to be made within seconds due to players surrounding them at all times during the game.

Last but not least, attackers who are usually well known players that get paid the most on a team because they are the most skilled cause without goals, a club can never win a game. Attackers are the ones who finish the season with the most goals and assists because that is their main priority, which is to create offensive opportunities. These players are generally the most popular because no fans get more excited than when their club scores a goal. With that being said, attackers key responsibilities are scoring goals, assisting teammates, stealing the ball from opponents in the attacking region and dribbling with poise. Therefore, effective attackers are evaluated on the statistics of goals/shots/shots on goal/shots off target, assists, dribbles without

losing the ball, takeaways, fouls, and how many times they got the ball taken from them. Soccer games are normally low scoring games while it's rare to see a team score 4 or more which is why it's vital to start the most efficient and effective players.[9] The most common setup is known as a 4-4-2. This is four defenders, four midfielders, and two attackers," which means those two attackers must be the most talented players with the ball and can deal with pressure.

Technical Directors

In today's society, every sport across the globe utilizes some sort of analytics team to enhance performance. These analysts will breakdown the results and work to provide insights to coaches and managers. Although, some sports must use different approaches than others in order to record accurate data. Therefore, what's the process for soccer? Soccer uses databases and decision support systems to evaluate statistics and performance. First, clubs must program and develop databases to collect data and statistics from a game. Then, that information is analyzed by data analysts and broken down to make it straightforward to understand. Next, those data analysts will pass that data along to the most important person in this chain who is the technical director. In fact, [10] the role of a technical director is to manage and oversee all player and coaching development within the club. In addition, the technical director makes sure that all coaches are following the club curriculum, as well as implementing the club philosophy and playing style across the entire club." This process helps technical directors clearly see the performance of a team as well as individual players. In addition, data collectors assist technical directors decide on new tactics, formations or substitutions as a way to change up the flow of games. At the end of the day, the technical directors have the final say as to what goes on across an entire organization because they analyze the data, which is ever so important with every sport today.

21st Club

The decision to early recruit young talent happens more so in soccer than in any other sport in the world. The best soccer players in the world are generally not physical specimens as they are rather more skilled individuals. That's why there are many young talented players between the ages of 19-23 playing at the professional level because soccer is dependent on speed and skill, and less about strength. This is where the 21st Club comes into play. One of the greatest tactics for improving a club is to scout for players who are still growing but already have the talent and speed. 21st club in an organization worldwide that actively recruits young prodigies for clubs. Once clubs have their eyes set on a 14-18 year old, they will attempt to sign them on a small salary to incentivise that individual to continue to work hard and develop their soccer skills. Due to their signed salary, no other club can poach that player and once they are good enough to compete at a high level, the organization "calls" them up to the major league. Club 21st evaluates the performance of young players to gauge whether or not that player will be good a few years down the road. Some players never pan out, but that's the risk clubs are willing to take because they would rather take a chance on a player than let their rival club sign the young talent. From a management standpoint, young talent is worth the investment because one or two young players can assist in turning around a franchise which as a result will likely increase ticket sales, hence is more revenue in the clubs pocket.

Performance & Match Analysis in Soccer

Soccer analytics can be broadly classified into two types which is match analysis and performance analysis. Both of these techniques play a vital role to improve the gameplay of the team. Match analysis centers everything about the matches, for example post-match analysis, various strategies and tactics whereas performance analysis has wider scope which includes various other disciplines. In general, performance analysis involves notational and motion analysis. Notational analysis refers to the recording of all actions or events in a "what," "where" and "when" manner (Carling et al., 2005). It allows for key elements of the performance to be quantified in a valid and consistent way (Nevill et al., 2008) to ensure an accurate and objective representation of the game (Carling et al., 2005). There are four major purposes of notational analysis: (1) analysis of movement; (2) tactical evaluation; (3) technical evaluation; (4) database developing and (5) evaluation and immediate feedback by coaches (Hughes and Franks, 2004). In preparation for the upcoming game, coaches need to make several important decisions, primarily decide on the roster and the strategy. This depends on several factors like the current status in the competition, level of the opponent and the overall fitness of the players, or perhaps even the weather (Rein and Memmert, 2016). Notational analysis is a means of recording events so that there is an accurate and objective record of what actually took place (Carling *et al.* 2005). There should be at least five elements which should be recorded: the position (where?), the players involved (who?), the action concerned (what?), the time (when?) and the outcome of the activity. Another method of match analysis is motion analysis which uses the various positions of players using the cameras. The various things that are necessary for match analysis include:

1. Touches by player
2. Total possession

3. Shots on target

4. Tackles won

5. Fouls

[5] The transformation that has happened in soccer, would not be accessible with-out cutting edge advancements in technology. FIFA characterizes these as Electronic Performance and Tracking Systems (EPTS) which are utilized to screen and improve player execution by following the player and ball positions, and can be utilized in blend with different gadgets like pulse screens, accelerometers, spinners or different gadgets for estimating physiological boundaries (FIFA, 2017b). FIFA likewise characterizes three significant kinds of gadgets: (1) Optical-based camera frameworks; (2) Local posi-tioning frameworks (LPS), and (3) GPS/GNSS satellite frameworks.

To understand the importance of Match analysis let us analyze the recent game between Spain and Germany.

Analyzing Spain Vs Germany Game:

[4] Spain and Germany both went into the game with 4-3-3 formation, but Spain had made several changes for this game against Germany. During build-up phases, Spain split their centre-backs midfielders wide and had two midfielders drop in to support. As a result, winger midfielders were tasked with marking the full-backs so it was up to midfielders to man-mark Spain's dropping midfielders. This allowed Spain's defender to cover the space in between the midfield and defensive line and mark forward if he was to drop into midfield to create the 4 v 3 overload. This strategy clearly meant that Spain was trying to create more space by using their

defenders. If we look at the past statistics of Spain, they like to have more possession in the game, for example in one of the matches in Euro 2019 against Malta, [6] they had nearly 83 percent of possession. Similarly in the game against Germany, Spain had 69 percent of possession, this forced Germany to change their normal strategy at half time, but it was too late as they were already 3 goals down.

There are many companies such as Opta Sky Sports which are official data partners for English Premier league. [7] They provide services such as: Advanced Metrics, Augmented data and predictive analytics. The kind of data provided by the Opta Sky Sports is as shown below:
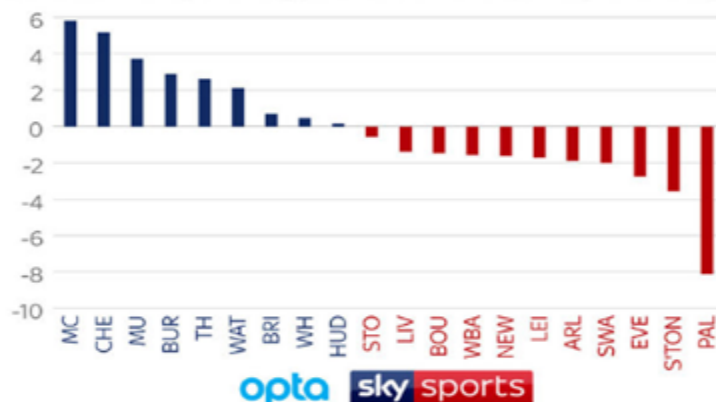
Manchester City are the most clinical Premier League team this season, while Crystal Palace should have scored eight goals

The current systems are too expensive and sophisticated to be used at the lower levels, hence these are generally used by the tier 1 club and National teams. Therefore, we provide an alternative approach which will be more feasible for smaller teams.

Alternative Approach

An alternative approach could be implemented using a knowledge based decision support system. A knowledge-based decision-making (KBDM) is a decision-making process which uses predetermined criteria to measure and ensure the optimal outcome for a particular subject. By setting up a thinking process and rationale behind a decision, it is used to make successful and strategic decisions. To collectively improve awareness of an issue or accepted criteria, it collects critical context essentials. To give a rough idea we can see the flow diagram below where information such as, number of goals, game type, Player statistics through API if it is fetched, but it can also be fed manually at a smaller level. The probabilities of the outcomes are calculated using Decision tree approach, but it requires cleansing of data from the expert such as when the system is fed with the data, it will generate the possible decisions that could be made along with the probabilities based on the current situation such as weather forecast, strength, weakness, etc.

Pros:

1. Simple and less expensive

2. Gives multiple decisions output with the probabilities

Cons:

1. It will require a new implementation of the system for team use.

2. Professional support team may not be available

3. All sophisticated features are not available.

DATA:Player Stats,Away or
Home,Strength,Weakness etc

KNOWLEDGE D/B

Using DT for outcomes

Decision and probabilities
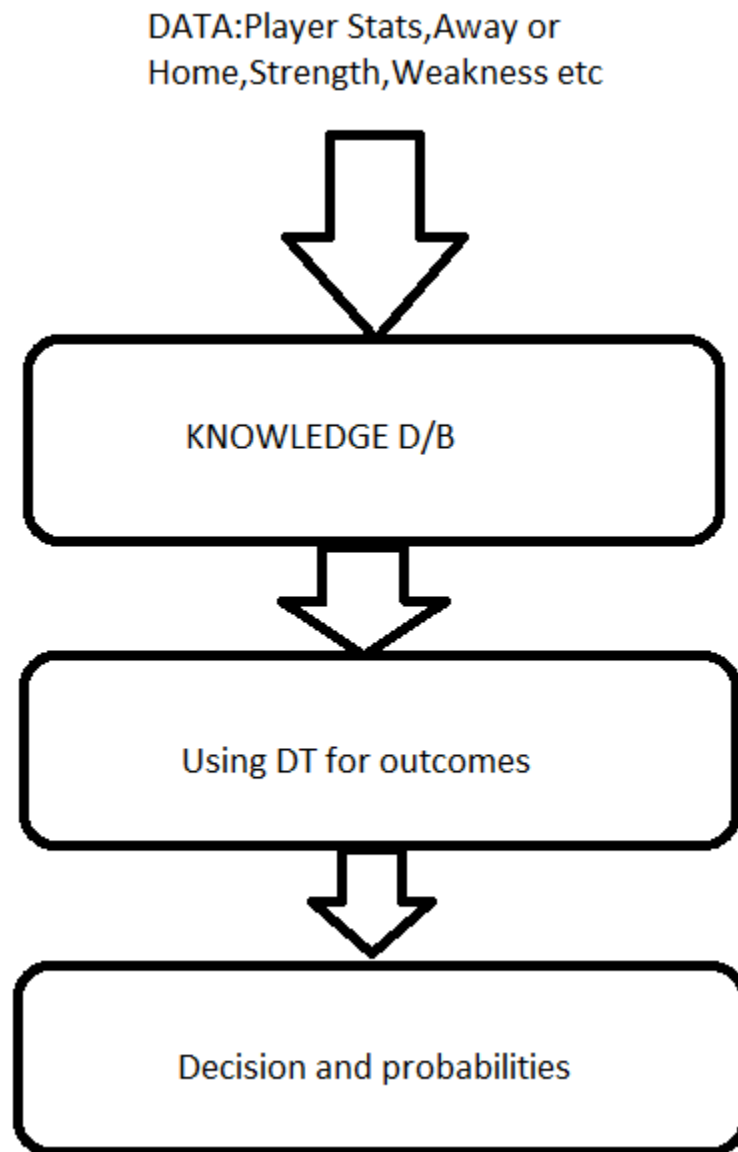
Conclusion

In this research, we conducted various researches on soccer analytics and how the system has evolved over the time. Although, there are many sophisticated analytical softwares available for club managers. We dissected how exactly these analytical tools work and provide data to the soccer coaches and managers. However, these tools are only accessible for the teams playing at a higher level as they are expensive to buy. Hence, we tried to showcase a simple model which could be developed by the smaller teams for improving their game through analytics. As a suggestion, instead of purchasing softwares to collect data, lower level organizations can hire one or two sports analysts to keep track of the important stats rather than taking away no data at all. This gives teams a competitive edge over other low level teams that may not be contracting any data from their soccer games. Overall, any sort of data from games is better than none because it's essentially feedback on their performance of a game that is utilized to benefit from, which every sport today is implementing in their games because it's necessary if a team wants to take the next step in improving their club.

References

[1]https://sites.duke.edu/socceranalyticsrevolution/history-and-background/

[2]https://www.experfy.com/blog/bigdata-cloud/moneyball-some-insights-to-soccer-analytics/

[3]https://www.goal.com/en-us/news/premier-league-suspensions-how-do-bookings-red-card-bans-work-in-/1e6awd5os0a031pri0kcq3eu0c

[4]https://www.espn.com/soccer/report?gameId=570646

[5]Big Data Event Analytics in Football for Tactical Decision Support, Palvina Krockel

[6]https://www.marca.com/en/football/spanish-football/2019/03/27/5c9aac28e5fdeadb588b45da.html

[7]https://www.optasports.com/services/analytics/

[8]https://www.runnersworld.com/runners-stories/a20805366/the-distance-run-per-game-in-various-sports/

[9]https://howtheyplay.com/team-sports/Positions-in-Soccer-and-Their-Roles

[10]http://assets.ngin.com/attachments/document/0039/2519/CamptonNation_MAY2013_PRINT.pdf