**Data analytics in football sport to identify gaps for the improvement of quality opportunities throughout world-wide teams**

Summary:

Football is a widely known sport. Billions watch and play the game around the world. Data Analytics has assumed a huge role in the world of Football. It has transformed how people approach games, team formation, player selection etc. Data analytics has enabled teams from around the world to understand their game better and perform better. Data analytics is also used to predict the outcomes of games enabling people to make educated guesses while betting. There is no doubt that Football is worldwide sport. However, there are so many teams worldwide who haven’t improved when compared to some of the others. Few teams don’t even manage to make into the main tournaments like FIFA. Some countries lack funding and some teams don’t have the exposure to standard equipment, coaching opportunities etc. It is very important for a Football enthusiast to know that the game keeps evolving towards a point where there are more quality teams around the world. The present paper discusses exactly that and looks to provide an answer to that very question. This paper discusses how a geographical analysis is done to

understand where international football matches are played. It is done also by looking at where the FIFA World Cup matches played. This has helped in realization of the first step that needs to be taken for the growth in popularity. It is also seen how playing friendly matches has changed over the years to understand what teams have felt about it. This will be very important information for upcoming teams.

Football game was always the most famous and popular sport in various European and South American countries that has been played and viewed. With the popularity of global and television broadcasting of the games, football game recently started to popularize in the Indian subcontinent since the early 21st century. India is steadily becoming a global figure in the Football (American: Soccer) world, with more and more official football events happening and also major international stars participating in the new Indian Super League. Two football datasets have been used for the analysis. The main dataset contains the results of international football matches starting from the very first official match played in 1872 - England vs Scotland up to 2018 [16]. The matches range from World Cup to Baltic Cup to regular friendly matches. The other dataset includes the GPS coordinates of each city for geographical analysis. The two datasets have been combined to produce the football dataset and the sample dataset. The dataset has been pre-processed to extract/add new attributes such as year, winners, latitude and longitude. Attributes like year and winners were extracted for convenience. The latitude and longitude coordinates were mapped to each city using another dataset that contains these data. The attributes in the final dataset are match\_date, home\_team, away\_team, home\_score, away\_score, tournament, city, country, year, winner, latitude, longitude.

The matches played over the years were analyzed based on where the matches were held to understand how diverse the organizers have been in extending the range of the venues. The analysis was done to understand the popularity of the game spread across the world. The matches were plotted on a world map based on their venues. The analysis has helped to realize that the football matches have in fact been played around the world in a lot of different countries. But this has been as a result of which team is playing. When it comes to the world cup, the organizers have confined the venues to places wherein there is more popularity or in other words where the revenue generated will be most. It is seen that most world cup matches have been concentrated mostly in Europe and a little in South America. From this analysis it is realized that maybe if the tournaments like the World Cup is conducted in more places, the Global popularity will rise and there’s a good chance of the quality of Football from these countries will improve exponentially.

England and Scotland, being the very first teams to play an international match, started out in 1872 and have had a steady rise to the top. Around 1948 Scotland started lose its momentum and slowed down compared to England. Argentina and Uruguay started playing in the mid-1910s and Argentina has seen a very good rise to the top and Uruguay has been relatively slow. Brazil of course, has also made it to the top with a steady rise having started in the mid-1910s. Of the top 10, Korea Republic started the latest, mid-1950s, and has made it to the top 4 very quickly. It has seen the fastest growth of the rest of the teams.

From the 1920’s there has been a zigzag rise in the number of Friendlies played per year. In mid-2000s the world has seen the most number of Friendly matches played in a year, that‟s slightly 400 matches. Few of the reasons for playing friendlies could be – exposure, tactical or strategical analysis of the opponent teams or to understand the opposition‟s weaknesses. With time teams have realized the importance of playing Friendly matches and that is evident from the analysis .

Conclusion:

We have seen that the most of the top 10 teams have more or less had big tournaments such as the FIFA World Cup in their home country. This has helped us realise the importance of conducting such tournaments in countries that are less productive in the game of Football. This can inspire and motivate more players to dream bigger and perform better.

Why it’s important:

Football may be one of the biggest sports in the world, but majority of the countries in the world don’t have a major following for football. However, several countries, including india, are actively trying to improve the state of football in their country. Papers like these provide valuable insights into the success of football in several nations and provide directions for other countries to emulate this success.

My critiques:

One of the major conclusion drawn from this paper is that hosting the FIFA World Cup has a dramatic affect on the success of football in a country. However, this ignores several other factors.

One such factor is the participation and performance of the home team in the world cup. If there is no excitement in the performance of the home team, then there wouldn’t be as great of a following.

Moreover, several other factors such as practice tournaments and club matches tremendously affect the following to a sport in a country. It’s not just about the world cup, it’s also about what happens when there isn’t a world cup happening that really counts. This paper fails to predict any such factor.

Done by-

Sushanth prabhu (PES2UG20CS359)