

Soccer Team Management System FIFA-19

Jeemisaba Chudasama (Jac786), Ankitha Gavinolla(ag1527), Ajinkya Patankar(ap256)

*Department of Computer Science, Rutgers University– New Brunswick, NJ, USA

{jeemisaba.chudasama, ag1527, ap256}@rutgers.edu

Abstract—FIFA 19 is a football simulation video game in the FIFA series, developed and published by Electronic Arts it is played by people around the world, in order to help them we try to build a system which builds the best possible team using different Constraints, find the best Substitute of a player, compare the performance of player. The main idea of this system is to enable the manager, individual players or business analysts to visualize this and also have some statistical analysis about the game.

Index Terms—Data Pre-Processing, Visualisation, algorithm, D3, plotly, Soccer, team management, statistical analysis .

I. PROJECT DESCRIPTION

Overview: We aim to provide an application that allows users to build their own possible best team using a large dataset containing information on over 18 thousand players with multiple attributes and features. The data can be used to build a more coordinated team within a given budget. This shall allow users to see which players are perfect fit for a position. Our project can be useful for soccer fans to follow their favorite players and for managers to manage their team. The Soccer Team Management System recommends the manager of a team to select players depending on the combined score(Score is given for each player when playing at different position out of 100 based on their past performance at that position) for the team selected. The score of teams thus formed is calculated by an algorithm. The dataset contains many attributes which can be analyzed to draw meaningful insights.

A. Stage1 - The Requirement Gathering Stage.

System description: Given a large dataset of soccer players from the game FIFA 19, we would aim to build the best team. Types of users are as follows:

1. Players of Fan (Naive)

Analyze the performance of other teams and player. Improve his game by seeing the past performance of the player having similar skills and potentials as him.

2. Club Manager and Investor(Intermediate):

want to build the best team based on different formation and constraints or one who wants to Substitute a player.

3. Game Developer(Professional):

Find the complex relation like between performance and potential of the player with respect to their

value.

Real world scenario:

• - Scenario 1:

Description: A user wishes to know the best possible they can get from a list of all professional, active players.

System Data Input: User selects a formation (4-4-2, 4-3-3, etc).

Input Data Types: The formation is read as a string.

System Data Output: It Visualize the player on the football ground at the position they are intended to play The system returns a table

containing a list of positions (corresponding to formation) along with the best player for that position plus some extra information.

Output Data Types: Json file of the player along with their photo and skill parameter.

System input data type: Player Name: String, Player Code: Integer, Country: String, Club: String, Rankings: Integer.

System data output: Team according to the algorithm displayed with positions on the field.

System output data type: Players List: JSON object ,Combined Team Score: Integer

- Scenario 2: Team Building with Constraints

Description: A user wishes to know the best possible team based for a specific club or nation along with the budget constraint let say budget is 100000000 Euros. System Input: Name of the club/ country they want to make team for and budget. System output: Player will be arranged in the field as per their position and we will have the total score of the team.

- Scenario 3: Description: User want to find the substitute of a player in team.

System Input: select the player you want to substitute from the field or from the table. System output: A new team will be created we will have a JSON file of the new players list. The player will be visualized on the field and will have a box around it(so that we can easily identify it).

- Scenario 4: Description: A user want to compare the skills of different player. System Input : Name of the

player whose skills they want to compare as string in the search box. System output: A Radar chart showing the comparison of different skills. output data type: a jason file containig player name and its skill score as a list.

- Time line for completion of the major implementation stages:

Stage 1: Requirement Gathering - by Feb 15

Stage 2: Design - by March 17

Stage 3: Infrastructure Implementation - by April 10

Stage 4: User Interface - by April 15

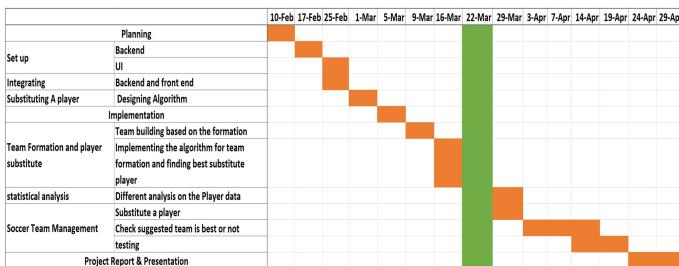


Fig. 1: Detailed Time line

B. The Design Stage

We will have all the data stored in data base than we will clean and pre-process it. We will have a backed in python which is connected to this database and Our front-end would respond to user input and accordingly call the respective functions in Python Django and display the output on the web page. Our flow diagarm is show in the fig below.

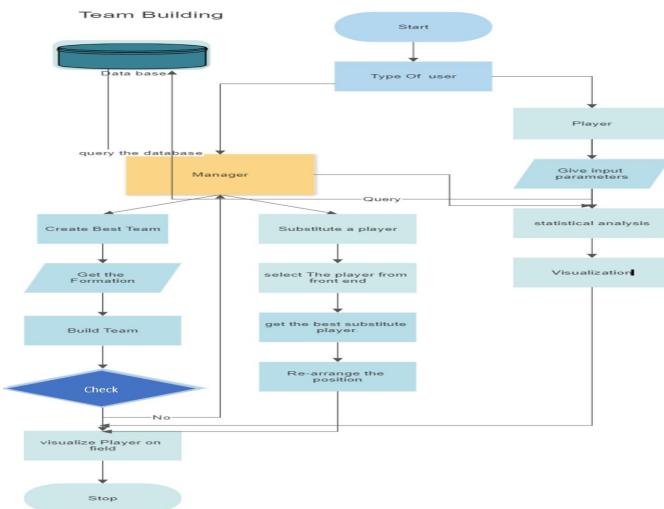


Fig. 2: Workflow Diagram

1) Data Cleaning and Pre-processing: The data is from the website by extracting the Player personal data, followed by Player IDs and their playing and style statistics. It contains information about players and skills, personal information, photo etc.

- format: CSV File
 - Numbers of player in Data set: 18k+
 - number features: 88 per player
 - source <https://sofifa.com>

For skill attribute, we have an integer from 0 to 100 that measures how good a player is at that attribute. Examples of attributes are: dribbling, aggression, vision, marking and ball control. Normalizing: Player Value in Our dataset contains values in different units, some in terms of million while others in thousands. When we read the data we would convert all Values to in terms of million.

Cleaning: We need to do some pre-processing for variables like Height, Weight and monetary variables like Value, Wage so that they're converted to a numeric format(example: removing the currency symbol) show in the fig below.

Handling Missing Values: Missing values were not removed directly as we thought we might need it for the future, so what we did it for all the column that had missing values we created one more column where we replace the missing value with standard deviation.

ID	Name	Age	Photo	Nationality	Flag	Overall	Potential	Club	Club Logo	Value	Wage	Special	Preferred F
158023 L. Messi	Lionel Messi	31	https://cldr Argentina		https://cldr	94	94	FC Barcelo		€110.5M	€55K	2202	Left
20801 Cristiano R	Cristiano Ronaldo	33	https://cldr Portugal		https://cldr	94	94	Juventus		€77M	€405K	2228	Right
190871 Neymar Jr	Neymar Jr.	26	https://cldr Brazil		https://cldr	92	93	Paris Saint-Germain		€118.5M	€290K	2143	Rights
193080 De Gea	David De Gea	27	https://cldr Spain		https://cldr	91	91	Manchester United		€72M	€260K	1471	Right
192985 K. De Bruy	Kevin De Bruyne	27	https://cldr Belgium		https://cldr	91	92	Manchester City		€102M	€355K	2281	Right
183277 E. Hazard	Eden Hazard	27	https://cldr Belgium		https://cldr	91	91	Chelsea		€93M	€340K	2142	Right
177003 L. Modrić	Luka Modrić	32	https://cldr Croatia		https://cldr	91	91	Real Madrid		€67M	€240K	2280	Right
176580 L. Suárez	Luis Suárez	31	https://cldr Uruguay		https://cldr	91	91	FC Barcelona		€80M	€455K	2346	Right
155862 Sergio Rarr	Sergio Ramos	32	https://cldr Spain		https://cldr	91	91	Real Madrid		€51M	€380K	2201	Right
200389 J. Oblak	José Sá	25	https://cldr Slovenia		https://cldr	90	93	Atletico Madrid		€68M	€94K	1331	Rights
188545 R. Lewandowski	Robert Lewandowski	29	https://cldr Poland		https://cldr	90	90	FC Bayern		€77M	€205K	2152	Right
182521 T. Kroos	Toni Kroos	28	https://cldr Germany		https://cldr	90	90	Real Madrid		€76.5M	€355K	2190	Right
182493 D. Godin	Daniel Giménez	32	https://cldr Uruguay		https://cldr	90	90	Atlético Madrid		€44M	€125K	1946	Right
168542 David Silva	David Silva	32	https://cldr Spain		https://cldr	90	90	Manchester City		€60M	€285K	2115	Left
215914 N. Kanté	Nicolas Kanté	27	https://cldr France		https://cldr	89	90	Chelsea		€63M	€225K	2189	Right
211110 P. Dybala	Paulo Dybala	24	https://cldr Argentina		https://cldr	89	94	Juventus		€89M	€205K	2092	Left
202126 H. Kane	Harry Kane	24	https://cldr England		https://cldr	89	91	Tottenham Hotspur		€83.5M	€205K	2165	Right
194765 A. Griezmann	Antoine Griezmann	27	https://cldr France		https://cldr	89	90	Atlético Madrid		€78M	€145K	2246	Left
192448 M. ter Steg	Martin Ter Stegen	26	https://cldr Germany		https://cldr	89	92	FC Barcelona		€58M	€240K	1328	Right
192119 T. Courtois	Thibaut Courtois	26	https://cldr Belgium		https://cldr	89	90	Real Madrid		€53.5M	€240K	1311	Left
189511 Sergio Bus	Sergio Busquets	29	https://cldr Spain		https://cldr	89	89	FC Barcelona		€51.5M	€15K	2065	Right
179813 E. Cavani	Eduardo Cavani	31	https://cldr Uruguay		https://cldr	89	89	Paris Saint-Germain		€60M	€200K	2161	Rights
167495 M. Neuer	Manuel Neuer	32	https://cldr Germany		https://cldr	89	89	FC Bayern		€38M	€130K	1473	Right
153079 S. Agüero	Sergio Agüero	30	https://cldr Argentina		https://cldr	89	89	Manchester City		€64.5M	€300K	2107	Right

Fig. 3: Uncleaned data Snippet

II. - THE IMPLEMENTATION STAGE

A. Technology used

Following are the technology we used in our project to make it fast and more interactive.

- Front-end: HTML, CSS, JAVA-Script, Django, Ajax, D3.js
 - Back-end: Python Django, SQLite
 - Programming Language: Python
 - Libraries: plotly, Matplotlib, Pandas, numpy

1) Statistical Analysis:

- The maximum Number of Player belong to which Country?

we found that maximum number of player belong to Europe which was around 1600 followed by Germany and Spain. We visualize it using a categorical bar chart which represents the total number player from each country

and also along with the position they play (Goalkeeper, Forward, Midfielder, Defender). The fig below is the output graph of that.

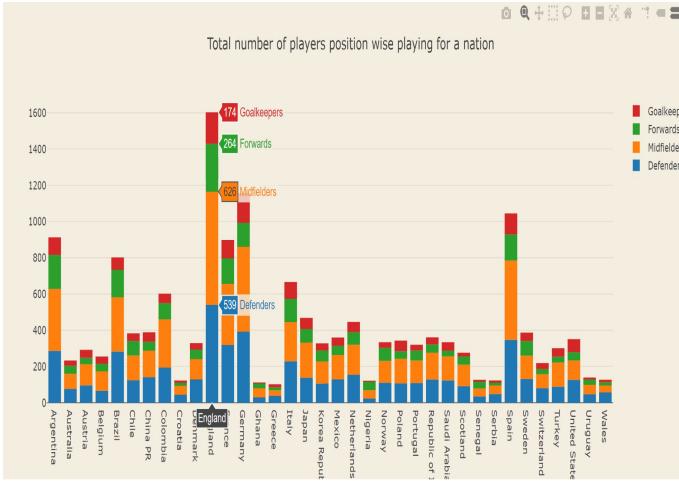


Fig. 4: Number of Player based on country

Interactivity : When user click on any country he can get number of player from country based on the position. User can also analyze the distribution of player for any one or more position by clicking on the position defined in the right of the graph.

But if you ask any football fan he will say that this analysis is false as they will say the maximum number of player should belong to Brazil so we further carried out one more analysis to see what is the average potential of the the player based on the country and we found that Brazil was top on that list, we concluded that reason for this will be as many player from Europe are dont play in the big matches as they don't have the potential for that people don't know about them, we also analyze it using the graph which is there in the our system.

- Potential VS Overall Power of a player. We tried to analyze what are the player who play at the top of there potential and who under perform based on there overall power (which is calculated using top 25 %).

Motivation: The reason behind doing this is to help the buyer to find the best player based on his performance and if he ready to invest in improving in performance he can get it for cheaper price. below fig represents our analysis: **Interactivity :** For the convenience of user we represent the player whose potential is same as the overall by white line. When ever user clicks on the percentage of potential he wants or overall score he will get the name of player who have that value.

- Relation between Value and Wage of the player

Value of a player is base on his performance and popularity of among the fan we all know that value of the good player are in millions but when it comes to wages it is not the same. so wanted to analyze is there a direct relation between value and the wage of a player? for example say

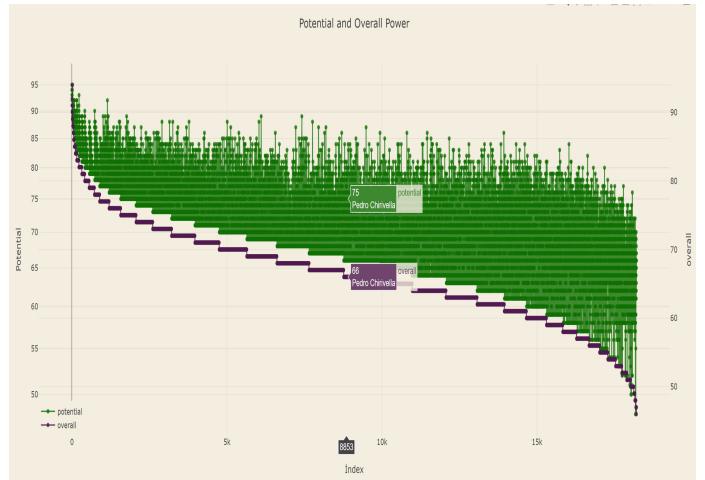


Fig. 5: Potential VS Overall Power of a player

if a player have highest value is it likely that he will be receiving the highest wage so manager can analyze the wage of a player based on its value. here is the below graph for that.

Interactivity: User can define the number of player he want to analyze we have divided player in 3 categories for this well payed, average payed, low payed and the category and will have an output graph like the one shown in the fig 6.

After this the user can check is the player really worth the price or is there any other player which he can get at the lower price so for this a another graph will be created which will as show in the fig 7 which plots the value of the player based on its potential.

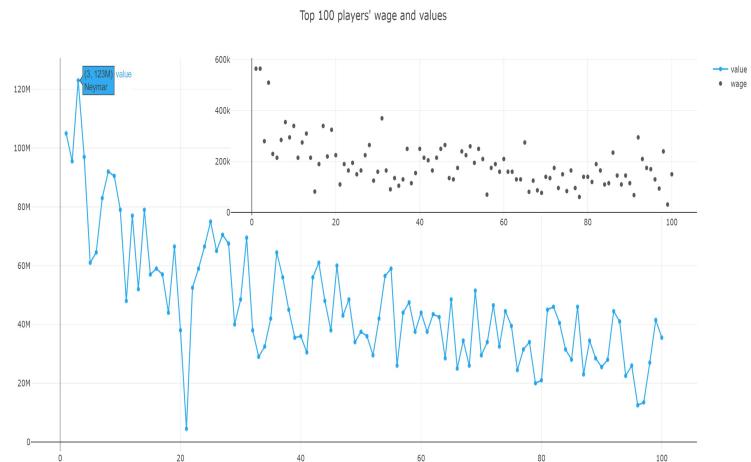


Fig. 6: Relation between Value and Wage of the player

Fig 7 Helps to find the user should he really invest in a player for this it takes into consideration three things Potential overall score of the player and its value.

Interactivity: All bar represents different player which

is labeled with its name when a user hover over it it will get it value potential and difference between overall and potential. There are few more interesting statistical



Fig. 7: Is the player Worth it

analysis that we did and you can find them in the image file (st4.JPG, st5.JPG, st6.JPG,)

III. STAGE4 - USER INTERFACE

Our project has 4 main pages.

- Home page has an auto slider which shows functions we have namely statistics, build own team, manage club.
- In build team page user can build his own team according to constraints specified by user like only formation strategy in which there is unlimited budget or both formation and budget. After building a team the players icon in their respective positions are visualized in field along with team score. Simultaneously, a table will be populated with the players from the newly built team displaying details about players like name, nationality, overall potential, position and user can also search for player information in search box. In addition, when user hovers over player icon his name would be displayed.
- In manage club page user can build team either for a club or nation. Other constraints include formation strategy and algorithm (whether greedy or improved) to build team. After building a team, user can also choose one player from drop down box to substitute, the new player will be highlighted in red colored square.

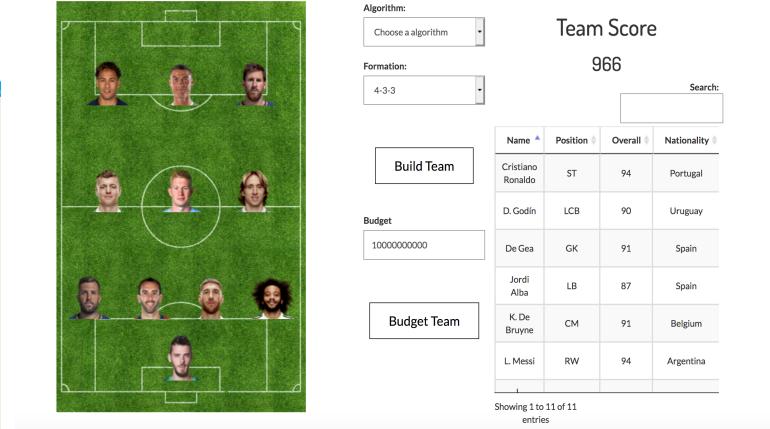


Fig. 8: Budget constrain team with budget of 1000M

IV. ALGORITHM

Algorithm used:

A. Greedy

For a given formation like (4-3-3 or 4-3-2-1) fix best player(i.e player with highest score in corresponding position) at each position at once starting from striker,followed by attacking, mid field and defending positions. But by following greedy approach we may not build best team.

Alternative Try all possible combinations and pick up best one (time complexity is $(n_{C11}) * 11!$)

B. Improved algorithm

Once user selects a formation, we get the positions related to that formation (ST,LW,RW etc) and then get list Of attacking, midfield, defender positions like [[ST,RW,LW,CAM], [LCM,RCM],[RCB,LCB,LB,RB]] for 4-3-2-1A and then first form best attackers team by picking top players(=number of attacking positions) for each position permute them to get best attacking team later midfielders and defenders are fixed in similar fashion.

To choose best one for each configuration, we calculate two scores:

Total score= sum of all players score in corresponding assigned position
Pref Score= for each position in configuration (if preferred position==Assigned position add $2 * (\text{preferred position-score})/100$ to pref score)

Time complexity= $(n,a)*a! * (n-a,m)*m!*(n-a-d,d)*d!$
n=no.of players

a=no of attacking positions m=no of mid field positions d=no of defending positions



Fig. 9: Liverpool team with formation strategy 4-3-2-1A, greedy algorithm

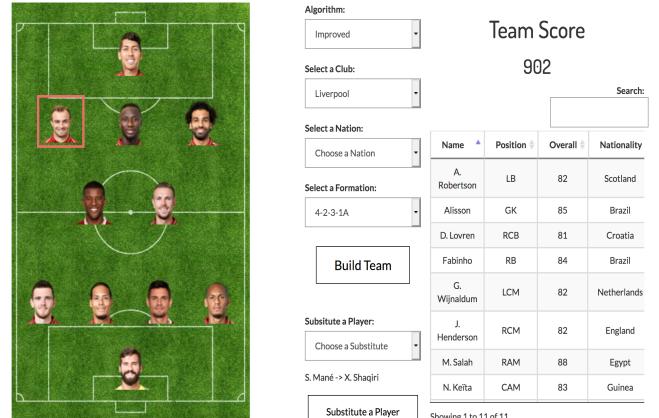


Fig. 11: Liverpool team with formation strategy 4-3-2-1A after substituting player('S.Mane')

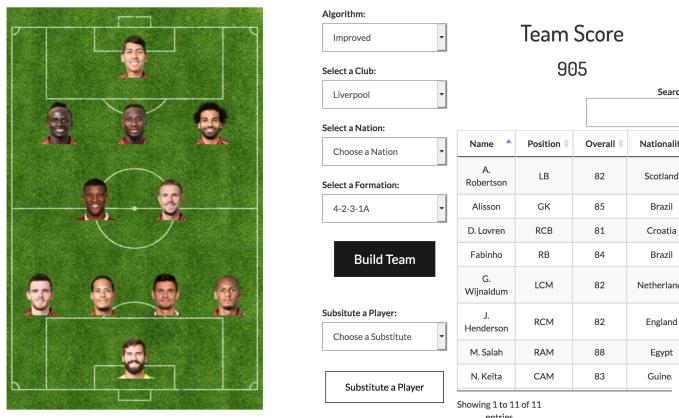


Fig. 10: Liverpool team with formation strategy 4-3-2-1A, improved algorithm

C. substitute player,budget team

Once the user selects a player to substitute, we add each player from the bench and build different teams accordingly. We choose the best team according to the above approach.

Under budget constraint, find the best player for position but with a market value = 1/11 of budget.

On our stats page, there is a world country map for user to navigate our data, each country has a color which means the total number of players. User can click certain country to zoom in and hover mouse to get a tool tip of players table in this country and get each players inform

D. Comparing Player

User can compare the performance of any number of player by adding adding name of the player in search box and then it can compare it with the other player and user will have a radar char of the player plotted as well

as they will also have a table of compression. this can be used by the player to improve their skills and see what are the strong skill of the other player who players at the same position. Fig below represents one such chart radar chart comparison of the two player but user can add up to 20 player at time and visualize all them together.

The reason we used radar chart they are easy to read for the user the far the value from the center better the performance is. In this Radar chart we are comparing the performance of the player for the 12 major skills like speed Acceleration, Passing, Dribbling, Finishing etc .

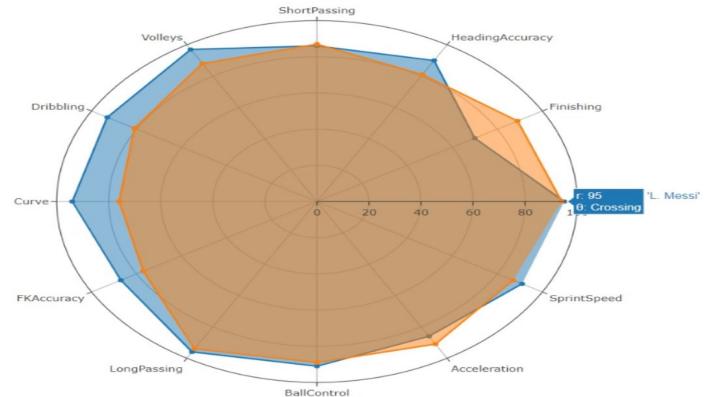


Fig. 12: Radar Chart of player Performance

V. REFERENCES:

- <https://www.degruyter.com/downloadpdf/j/hukin.2012.31.issue-1/v10078-012-0015-7/v10078-012-0015-7.pdf>
- Andre, D. and Teller, A., 1998, July. Evolving team darwin united. In Robot Soccer World Cup (pp. 346-351). Springer, Berlin, Heidelberg.

- Van Den Broucke, L. and Baert, S., 2019. And at the end, the Germans always win, dont they? An evaluation of country-specific scoring behaviour in the dying seconds of international club soccer games. *PloS one*, 14(4), p.e0202852.
- <https://arxiv.org/abs/1506.07768>
- <https://oaktrust.library.tamu.edu/handle/1969.1/173614>
- <https://link.springer.com/article/10.1007/s40279-017-0836-6>
- <https://www.kaggle.com/jsaguiar/complete-eda-time-series-with-plotly>
- <https://github.com/amanthedorkknight/fifa18-all-player-statistics>
-