Pattern Recognition
2022-2

Replacing players
Minimum Viable Product
Requirements

Acosta Imandt Daniel
Castillo del Valle Pablo David
Garduño Martínez Eduardo

09/05/2022

The MVP is at the end of the document
The new parts are in blue

The problem

We want to study the players of FIFA to find similarities between them and find the ones that have similar skills with distinct techniques of clustering like KNN and K-means. Also with the help of genetic algorithms we want to create different teams that play well and have a lot of potential.
There are a lot of parties interested in software of this kind, especially the different national teams, to find the best lineups for the players or find others that can replace some existing ones.
Currently, there exist some recruiters of players, but they need some help, especially in these times. So we can provide them software that will help them to make better and easier decisions.
So we will achieve success if our software is making good decisions and putting to test is showing good results.
To create our model, we will use the database of the players in FIFA, that we can find in Kaggle. We want to use this because it's very complete and gives us a lot of information about individual players.

Methodology

1. First we need to explore the database, and do some cleaning, conserving the columns that we will use as our dimensions.
2. Then we will map every player in this space and implement K nearest neighbors to find the players that have a lot of similarities. This can be used when an old player is retiring and the team is looking for some newer player that can take his place, take the attributes of that player some time before the prime of his career and compare it with the new players.
3. Also we will use K-means,to make some clusters of different players by their attributes, see if this clusters correspond to the positions that they play like defense or goalkeepers.
4. Then we want to implement some genetic algorithm to find different combinations of players that achieve the best results in games.
5. We will present our findings to the clients, with some feedback and conclusions, that they can implement to improve their teams.
6. Finally we will gather more data and see the results that we get, to improve our software and get better results, for future consults.

It is very important that we evaluate our model and that it is giving good results, in the case of our database, it is easy because we can see the position of every player so we can see if k-means is giving good results, for the part of genetic algorithm we can use our algorithm with the database of some prior

years, and then compare it with the actual year and see if with our observations the overall of the team improves.

We need to do a lot of different things that our competition and other prior works,like for example https://www.kaggle.com/code/stpeteishii/fifa20-player-value-important-factor-clustering and https://www.kaggle.com/code/vishnudasbharathraj/fifa-clustering-preprocessing-pca-scaling-cluster but use some interesting ideas, both our goals are very different of theirs, they are more interested in find the features that are more important for the potential of players and find the correlations between some attributes like the age of the player and the overall potential, using clusterings, but their are not interested in find if those clusters correspond to different positions like defense or goalkeepers, like we want to do.Also we want to find similarities between players to supplant those.We could use some of their ideas to improve our work, but at the end the goals are completely different.

The Minimum Viable Product should reach the following key aspects

- ~~Our model gives good results at finding similar players.~~
- ~~Our algorithm is well optimized giving fast an reliable results.~~
- ~~It should be possible for our clients to get automatically results about the players they are looking. Get the attribute clusters of every player in the league.~~

- We will present a python notebook, where there will be a lot of options so the user can find the best combination of players for their teams, and change an specific player.
- We will have our data base of the players with their stats. And with K-means and KNN we can find similar players.
- The users will get automatically what they are looking for.