Transfer-market Value Predictive Analysis for Football Players using Machine Learning

Abstract: In recent times, there has been substantial increase in the transfer-market operation for soccer (European football) clubs in search of national and especially European glory. As a matter of fact, a lot of money is involved with the transfers of top players in the big European Football leagues. Top clubs in the transfer-market always looks for the big fish and they shell out a lot of money to bag the player as they aim to achieve the unbeatable status in competitions across Europe. Recent years have seen the rise of clubs like Manchester City (England) and Paris Saint-Germain (France) due to huge financial support achieved due to change in ownership while big guns like FC Barcelona, Real Madrid CF(both from Spain), FC Bayern Munich (Germany), Manchester United (England) etc. have always been busy buyers in almost every transfer window. These big guns always aim for the best players that suit their style of play and can take the club to the next level. A player's transfer value is determined by his performance, consistency, skill level as well as temperament and many other factors. Now, with the help of data providing firms, access to historical data is easy and we can do our own predictive analysis according to that. In this work, we analyse the top 100 players' (listed by UEFA) performance and try to do a predictive analysis for the transfer-market value using regression analysis.

1.Introduction

Football is a leisure or a hobby for many people. However, it is also increasingly established that sport is also a business. Majority of the clubs are 100 years old but clubs started to form from 1850s . When football started becoming popular and gathered spectators the business practices developed into the clubs ,till then the clubs were just sporting associations . But since 1980's there has been a major change in the business model of the sport due to the involvement of the corporate's and other business organization. In the last decades the football clubs commercialized largely and the business in football has grown enormously .Transfer market is an essential part of the footballing business world .

Transfer market is the arena where football players are available for transfer to other clubs.

An important decision which the club has to make is in which player the club should invest and how much money has to be invested . It is very important to do research on the players the club wants to invest in , with ease of access to data making the decision process easier .

In this work the transfer value of a player is being predicted on the basis of various parameters .

2.Transfer Market Parameters

In this modern era of football , the transfer of players from clubs to club is becoming more and more complex due to different constraints and environment.The clubs are also allocating more funds to get top players that will both enhance the club's chance of getting any silverware as well as the brand value. It has been seen over the past decade the transfer market parameters have increased hugely , nowadays not only goals and assists matter but huge number of other parameters (including both on-field and off-field behaviour of a player) play a major role .In this work we have chosen 29 such parameters initially and collected data for the 2018-19 and 2016-17 from[2] and [5].

3.DataSet Decription

For this research data containing football players' basic information(name,age),market information(transfer fee for several seasons) and performance information (goals ,assists ,minutes played etc) where required . After going through plenty of useful open source data related to the project work online , the datasets were prepared from [2],[5] .The dataset consists of data for 80 players mentioned in [1] for 2 seasons .

Initially the dataset consisted of about 29 parameters , but after performing correlation check within the parameters the dataset was narrowed down to 15 parameters . We talk about this process in detail in Section 5.

4.Correlation Check and Data Cleansing

After the collection of relevant data the correlation between the parameters were checked and the data was cleaned when , either the correlation co-efficient exceeded +0.5 or fell short of -0.5 .Finally after the data cleansing was done we were left with 15 parameters such as goals ,assists , minutes played ,fouls , clearances etc

5.Linear Regression

In regression problem the goal of the algorithm is to predict real-valued output. linear regression is a widely used statistical learning method which is used to predict a quantitative response Y from the predictor variable X.

Mathematically, we can write a linear relationship as:

$$Y=\beta_0+ \beta_1X_1+ \beta_2X_2+\ldots.+ \beta_nX_n$$

Where:

• y is the response

• β values are called the model coefficients. These values are "learned" during the model fitting/training step.

 • β_0 is the intercept

• β_1 is the coefficient for X_1 (the first feature)

 • β_n is the coefficient for X_n (the nth feature)

The transfer market value prediction analysis is easier when visualised through Linear Regression plots.The function is extensively present in the NumPy library for predictive analysis and the same was used for the analysis part of this work.

The NumPy library has two separate equations for Linear Regression Analysis. When the dataset is overfit an error function is introduced to train the model accordingly .

6.Result and Discussions

The following observations were observed after performing the predictive analysis. Figure (a) represents prediction on training data . Figure(b) represents prediction on test data.
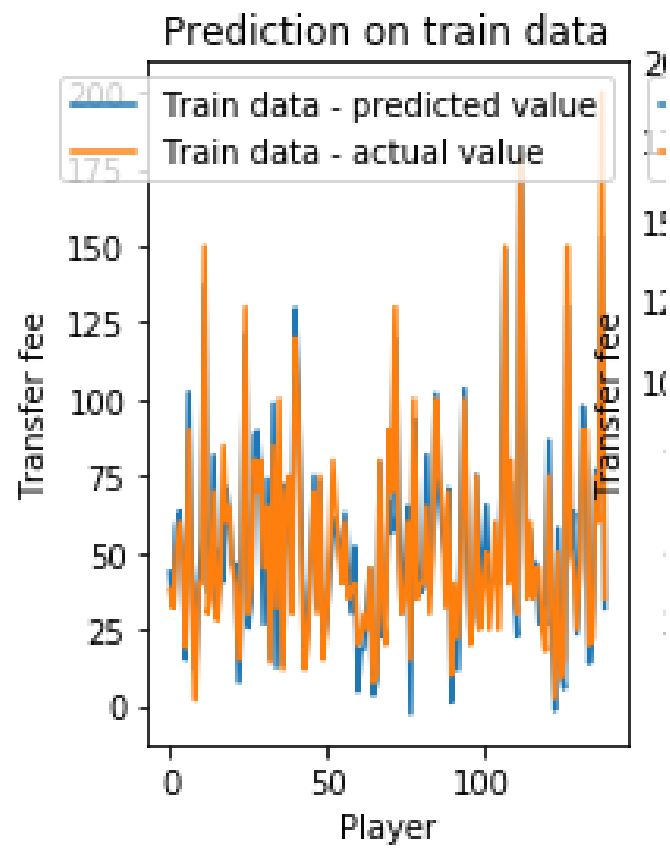
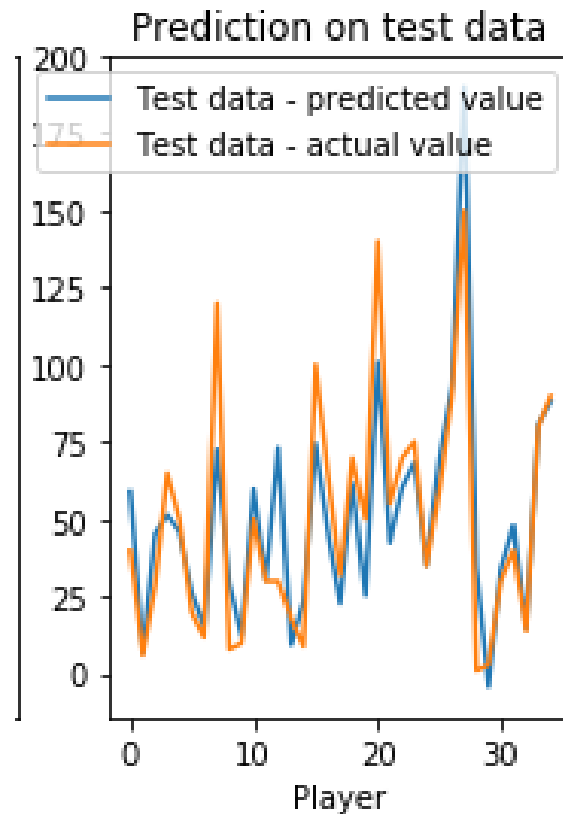Console 1/A ❌

Prediction on train data

Fig.a

Fig.b

7.Conclusions

After the completion of the test run decent prediction was observed where the model predicts correctly for most of the time . In future we would like to use logistic regression for the same and compare the results . In this work we had to omit goalkeepers due to absence of reliable data, which can be tested in future with the corresponding dataset .

8.Acknowledgements

This research is supported by Dr. Prof. Jothi K.R and all the faculties of SCOPE VIT University.

9.References

1. https://bleacherreport.com/articles/2696960-uefa-champions-league-100-ranking-the-top-players-ahead-of-last-8#slide2

2. https://www.whoscored.com/Regions/252/Tournaments/2/England-Premier-League

3. Football player's performance and market value, Miao He, Ricardo Cachucho, and Arno Knobbe, September 2015

4.Kumar, G., Machine Learning for Soccer Analytics, 2013

5. https://www.transfermarkt.co.in/