A study of Prediction models for football player valuations by quantifying statistical and economic attributes for the global transfer market

Conference Paper · March 2019

DOI: 10.1109/ICSCAN.2019.8878843

Dibyanshu Patnaik
Narsee Monjee Institute of Management Studies

Kartikeya Prakash

Krishna Samdani
Narsee Monjee Institute of Management Studies

A study of Prediction models for football player valuations by quantifying statistical and economic attributes for the global transfer market

Dibyanshu Patnaik
B.Tech Computers
Mukesh Patel School of Technology Management and Engineering,
NMIMS University
Mumbai, India
dibyanshupatnaik@gmail.com

Harsh Praharaj
B.Tech Computers
Mukesh Patel School of Technology Management and Engineering,
NMIMS University
Mumbai, India
praharajhp@gmail.com

Kartikeya Prakash
B.Tech Computers
Mukesh Patel School of Technology Management and Engineering,
NMIMS University
Mumbai, India
kartikeya1899@gmail.com

Prof. Krishna Samdani
Assistant Professor
Mukesh Patel School of Technology Management and Engineering,
NMIMS University
Mumbai, India
krishna.samdani@nmims.edu

Abstract— **The global transfer market in the field of football has long been devoid of technology. Over the last two decades, the economic dynamics have completely been changed, and the amount of money which has been pumped has increased exponentially. Valuations for any physical entity with a well-defined set of parameters is a challenge in itself and depends on numerous other things. And doing the same for an individual with varying parameters, based on the source of data, is even more challenging. With this paper, we attempt to find the most optimum way of fetching data of football players and applying the right model on it so as to extract meaningful information, thus reducing the gap between the estimated prices and the Final price. We do so by crowdsourcing the data and applying regression techniques along with Opta index. We also try to find the results using Neural Networks and conclude with a comparison between our models**

Keywords— Moneyball; Data-Driven; Regression; Neural Networks;

I. Introduction

According to FIFA [1], 6.37 Billion dollars were spent in the global transfer market in 2017 alone. While in areas like arts, talent can be hard to identify due to it being subjective in nature. Whereas in sports, individuals can be assigned verifiable definite economic value by reviewing the vast amount of numerical data - both statistical and financial. Players and their performances can be averaged out over a relatively larger time period thus equating their capabilities to regress to the mean, causing luck to balance out in the process. This paper is set out to examine the various approaches which can be taken for aptly predicting values of football players who are potential targets for given clubs.

Football Clubs can buy players from the transfer market period which is split into the summer and the Winter Transfer Windows. Most major signings happen during the summer, and the winter is reserved for getting in reinforcements for the final half of the season.

Clubs traditionally spend 20% of their revenue per season in the Transfer market. Although the club finances are governed by the FIFA Fair Play Rules, the income and spending of every club vary based on external parameters like merchandise sale, ownership, and corporate sponsorships.

The competitiveness and the intensity of the game vary from region to region based on the participation of the people, media coverage, and club budget. This, in turn, brings out the contrast in the level of players and also fluctuates the market value and the skill level based on the parameters such as country, hype generated by the media, competitiveness of the league in which they play and their experience. The bigger the role the player plays in his team, the more likely they may be valued in the market like being the finest penalty taker, or spot free kick specialist, or other roles such as being a playmaker, chance creator, having excellent speed, etc.

Every ardent sports fan has heard of the term Moneyball—the concept that statistics can demonstrate a side that perhaps the customary way of looking at performances may not. It was this idea that drove Oakland Athletics, an MLB team, to assemble an impressive team with a high attribute-to-cost ratio using underpriced players with good performance statistics; thus setting a record of 20 consecutive wins in the year 2002. The team management advocated a statistics-based analysis as part of their transfer strategy [2]

@IEEE 978-1-7281-1524-5

II. METHODOLOGY

A. Data Collection and Parameter Setting

For each player, we go through the list of games he played in, and extract the features as we are going to see. The statistical technique applied in this study is regression analysis. Regression analysis predicts the effect of the independent variable on the dependent variable. Since there are several independent variables, a multiple regression model can be employed.

Firstly, we see that data can either be crowdsourced [3] or be taken from sofifa.com that derives its data from the EA Sports' FIFA game franchise [4]. Moreover, a data aggregators like Transfermarkt can also be used to fetch basic statistic details for every player, but the exact process of doing so is not known [5]; apart from the fact that they involve a few empowered individuals, who base their opinions on the fans' rating and the statistical parameters involved. [3]

Not all the parameters mentioned will be applicable to all the players involved; hence we divide the player attributes into two categories: Attacking and Defending. The attributes like goals and assist define the attacking characteristics of a player, while tackles and challenges define the defending characteristics of that player. Based on the position of play of the player, one of the above-mentioned categories is given higher weight as compared to the other. Surprisingly, in [15] it was found that age doesn't have a direct impact on the performance of a player.

In his study, [16] states about Superstar players ie, to see the effect quantile regression used for seeing how a player who is in the top 5%, change the value of an incremental goal of goals scored compared to a player who scores the median number of goals. The study concludes with evidence of the superstar effect stating that the goals scored by forward players which are 'superstars' are worth more to their market value than an average player.

In the thesis by Louivion and Pettersson [2], they have concluded that even after using 12 covariates for their multiple linear regression model, performance features like the 2/3 point shots had a much higher impact on the final estimate than other features like age, assists, steals etc.

In [3], 3 broad parameters. Player Characteristics, Player Performance, and Player Popularity are considered. Each of these has sub-parameters as shown. Player Characteristics in terms of quadratic of age (age-squared), footedness, league, and player position is considered. Age reflects both potential and experience. It is a nonlinear relation considering that players' values increase in mid-twenties and decrease thereafter. Two footedness reflects the flexibility of a player as that would help them adapt to various positions on the pitch. The league in which the player plays is also important a shown in [6]. Also [3] concludes that because of the lack of flexibility offered by goalkeepers, they earn significantly lower than midfielders.

Player Performance is based on playing time in different league games. Other parameters like goals and assists indicate a players' scoring skills and overall impact on the game. For example, while forwards are supposed to score goals, defenders should win tackles, and midfielders are expected to defend and attack equally well. Hence aggregated indices have also been used by some researchers so as to equate the parameters irrespective of the player position.

Player Popularity based on not only their talent but also their crowd-pulling ability and brand value, help influence the transfer fee. The emergence of "Superstar" Players [7], whose transfer fee is much higher compared to the rest, can be explained by this index. [6] have found that even Google search hits have a significant effect on the player valuations.

Like above, [5] follows a similar approach but takes in just the performance characteristics. The median difference between their estimated value and Transfermarkt's estimate turns out to be 34% and the mean difference around 60%. This happens because their model doesn't incorporate the commercial aspect to the pricing. Hence the model prescribed by [3] gives optimum results.

In another approach, for preprocessing the collected data, dimensionality reduction with Principal Component Analysis is used as suggested in [4]. In their model, they selected the best model based on the desired number of dimensions they set to check and compare the model accuracy. It chooses parameters like relevant performance-related features like Dribbling, Shot Power, Attacking, Finishing, Head Accuracy, Acceleration, Crossing, Skill, Curve, and Ball Control, which have been normalized to an equitable scale for comparison.

Linear regression was used for building the model of market value. It gave a very low accuracy because the pattern of the market value is exponential in nature. Hence the log converted market value gave a better prediction. An accuracy of 84.34% and a SD of 0.84 was observed in the model. Using log based value improved the model and gave an accuracy of 91%.

In [7], a comparative study of the Big 5 European clubs are made and is concluded that the football clubs are essentially success maximizers and not profit maximizers. Hence the demand for extremely high-quality players is much higher than the demand of just high-quality players. Their work finds evidence that player performance has the highest coefficient of importance for determining transfer fees. Also, the time remaining in the contract of the player has a positive influence. For player positions, this paper uses 5 categories instead of the 3 seen earlier. And each of the player performance parameters having a varying degree of influence on the player position.

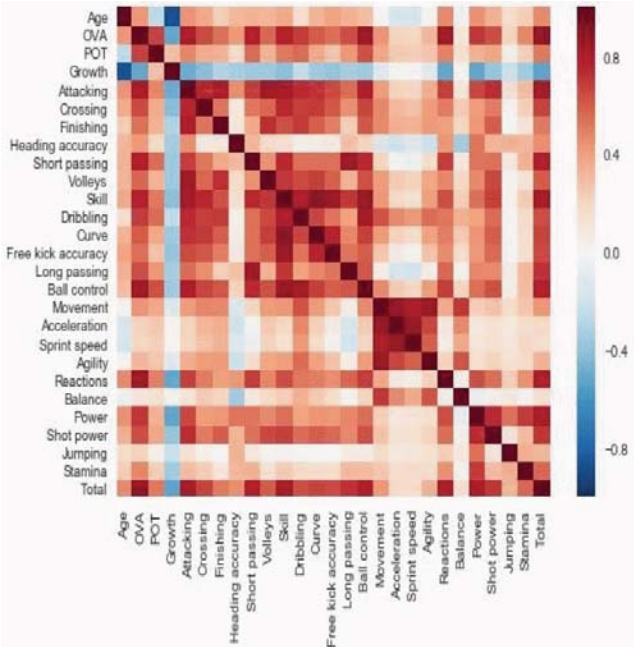

Fig. 1. Heat map of the dataset showing the relationship amongst the attributes [4]

One important point to note is that clubs are the workplaces for the football players, and like in any other individual's professional life, players to prefer to move to better clubs. Upward and downward job mobility for a football player can be defined as his move from a lower rated club to a higher rated club and vice versa respectively. The market value of a football player tends to increase with an increasing pattern in upward job mobility. The negative impacts of downward job mobility can be reduced since players sometimes take up crucial roles in lower rated teams which doesn't hamper their market value.

Job mobility patterns are considered as an important parameter since the willingness of football clubs to pay a certain amount of transfer fee for a player can be a deterministic factor on his market value. The two main contextual factors while observing patterns in job mobility are the time of mobility events in player's careers and the squad role of players for their new club [8].

B. Modelling the Data

1) Crowd based Estimation models

As discussed above, Transfermarkt gives their own valuations of players based on parameters they think fit. Assuming y arbitrary parameters taken by the crowd to assess a player, Transfermarkt then uses the concept of "judge", where all the valuations for a player are evaluated to form a single value, along with the judge adding an additional x parameters which are relevant for determining a player's market value (for eg: Age, Position) and other random parameters, thereby giving out a final estimate of the market value for a particular player. Hence its an optimum source for the data, but the estimates can still be improved upon by setting more realistic parameters and by applying other operations (like taking log based values).

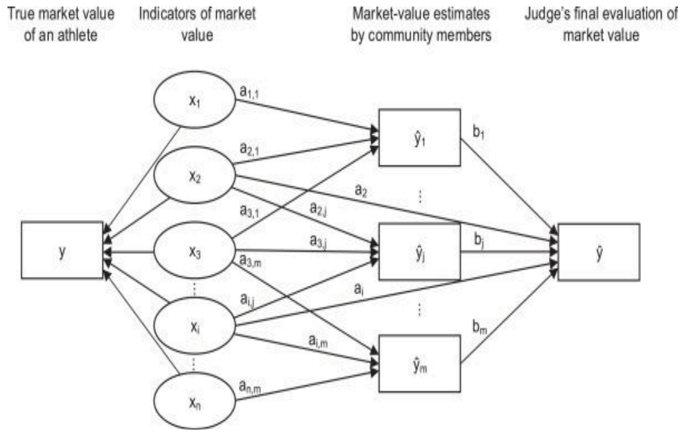

Fig. 2. Conceptualization of market value estimation at Transfermarkt [3]

2) Multilevel Regression Modelling

There are certain problems with crowd-based estimation model due to which the results obtained are not considered as the standard valuation of a player [3]. Since Transfermarkt is an open forum, all values are public, which directly eliminates the competitive edge in clubs for buying a player, and hence hampers the market value. Secondly, crowd estimation can only be made for players who are widely popular, which does not cover small leagues and lesser-known players, in turn not helping scouts screen through the first layer of a humongous pool of footballers throughout the globe. Therefore, data-driven models for judges is preferred, since data-driven models prove accuracy of 10% more than that of human analysis.

For the multilevel regression analysis, a data of 4217 players on 146 teams from the top five European leagues comprising of statistics across 6 seasons (2009/10 to 2014/15) was taken and 4 models were built which were classified on the basis of Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC).

The baseline model took the log of the player's previous market value as the significant parameter along with an intercept to calculate the current market value.

Improving upon the baseline model, Model 2 took previous market value and Age squared as the significant variables, introducing player characteristics and physical demographics which include Age: which denotes his experience vs potential, footedness: since a double-footed player could be worth more as it is rare, and Position: where players are broken down into three main categories. Attackers, who are meant to score more goals and remain in the forward half of the field, Defenders, who are meant to win tackles and stop the opposition from scoring, hence being in the backward half, and Midfielders, who have to strike the balance between their gameplay in the forward and backward halves.

Model 3 considered player performance as their significant parameter along with the other significant parameters in Model 1 and 2. Player performance could be defined by various factors dependent upon position. For attackers, Goals scored, which denote a player's scoring ability and Assists, which denote the player's contribution towards a scored goal are the key points to predict the worth of the player. For midfielders, Dribbles, Passing, and Assists could be considered as the significant factors, and for defenders, we have Tackles and Dribbles. Some factors remain independent of the player's position such as Red Cards and Yellow Cards, which inversely impact the market value of a player, as the repercussions of multiple cards are not desirable.

Variable	Measurement	Mean	Median	St. Dev.	Min.	Max.
Player valuation						
Transfermarkt's market value	EUR	5588,529	3000,000	8208,470	50,000	120,000,000
Player characteristics						
Age	Years	26.51	26.00	4.08	17.00	40.00
Height	Centimeters	181.49	182.00	6.15	161.00	203.00
Player performance						
Minutes played	total p.s.	1612.39	1612.00	884.85	90.00	3420.00
Goals	total p.s.	2.39	1.00	3.85	.00	50.00
Assists	total p.s.	1.64	1.00	2.25	.00	20.00
Passes	total p.g.	29.45	28.48	13.36	1.55	110.03
Successful passes	percent p.g.	.78	.78	.07	.43	1.00
Dribbles	total p.g.	1.21	.90	1.12	.00	9.58
Successful dribbles	percent p.g.	.51	.50	.24	.00	1.00
Aerial duels	total p.g.	2.22	1.79	1.71	.00	15.50
Successful aerial duels	percent p.g.	.47	.48	.18	.00	1.00
Tackles	total p.g.	2.21	2.09	1.21	.00	9.00
Successful tackles	percent p.g.	.71	.72	.14	.00	1.00
Interceptions	total p.g.	1.35	1.25	.92	.00	7.13
Clearances	total p.g.	2.09	1.07	2.35	.00	13.44
Fouls	total p.g.	1.10	1.03	.53	.00	4.27
Yellow cards	total p.s.	3.48	3.00	2.89	.00	18.00
Red cards	total p.s.	.20	.00	.46	.00	3.00
Player popularity						
Wikipedia page views	total p.s.	104,509.30	23,944.00	319,022.80	.00	8786,701.00
Google Trends search index	average index p.s.	13.36	13.21	12.38	.00	91.83
Reddit posts	total p.s.	15.42	2.00	38.79	.00	789.00
YouTube videos	total p.s.	36,075.46	918.50	141,882.30	.00	1000,000.00

Notes: p.s. = per season; p.g. = per game; $N = 10,350$

Fig. 3. Player valuations and relevant attributes [3]

Model 4 took into account a player's popularity as well because since we are talking about the market value of a player, the popularity of a player can help sell jerseys and tickets which results in a direct monetary benefit to a club. For calculating a player's popularity, the mention of the player on the soccer subreddit, the number of visits on his Wikipedia page, Google searches and YouTube video mentions were taken as significant parameters.

Based on the Akaike Information Criterion (AIC), the models were classified on the goodness of fit. Relatively comparing all the models, it was derived that with increasing significant blocks added, the goodness of fit was increasing for all the models, denoted by a significant drop in the AIC. For Model 2, the AIC dropped from 17416.2 to 14983.2. For Model 3 and 4, the AIC dropped from 14983.2 to 10172.0 and from 10172.0 to 10035.9 respectively.

Multilevel regression models.

Dependent variable: Log of market value	Model 1	Model 2	Model 3	Model 4
Fixed effects				
Intercept	6.789***	6.492***	7.432***	7.272***
	(.132)	(.219)	(.203)	(.200)
Log of previous market value	.543***	.610***	.495***	.486***
	(.006)	(.005)	(.005)	(.005)
Age^2		−.002***	−.002***	−.002***
		(.000)	(.000)	(.000)
Height		.002	.001	.001
		(.001)	(.001)	(.001)
Footedness		−.003	−.006	−.007
		(.022)	(.017)	(.017)
Minutes played			.000***	.000***
			(.000)	(.000)
Goals			.026***	.024***
			(.002)	(.002)
Assists			.016***	.015***
			(.002)	(.002)
Passes			.006***	.005***
			(.001)	(.001)
Successful passes			.301***	.286***
			(.083)	(.083)
Dribbles			.030***	.028***
			(.005)	(.005)
Successful dribbles			.035	.034
			(.019)	(.018)
Aerial duels			.013***	.014***
			(.004)	(.004)
Successful aerial duels			−.005	−.006
			(.028)	(.027)
Tackles			−.021***	−.018***
			(.005)	(.005)
Successful tackles			.049	.050
			(.030)	(.030)
Interceptions			−.013	−.010
			(.008)	(.008)
Clearances			.003	.003
			(.003)	(.003)
Fouls			.002	.004
			(.010)	(.010)
Yellow cards			−.004*	−.004*
			(.002)	(.002)
Red cards			.007	.007
			(.009)	(.008)
Log of Wikipedia page views				.016***
				(.002)
Log of Google Trends search index				.006
				(.004)
Log of Reddit posts				.026***
				(.005)
Log of YouTube videos				.007**
				(.002)
Random effects				
σ_1 (Player/Team/League)	.444	.298	.179	.185
σ_2 (Team/League)	.280	.217	.237	.219
σ_3 (League)	.138	.137	.150	.120
σ_4 (Position)	.083	.052	.056	.050
σ_5 (Continent of origin)	.057	.053	.034	.029
σ_6 (Season)	.107	.089	.089	.098
σ_7 (Residual)	.409	.411	.347	.343
Log Likelihood	−8699.1	−7479.6	−5058.0	−4986.0
AIC	17,416.2	14,983.2	10,172.0	10,035.9
BIC	17,481.4	15,070.1	10,374.9	10,267.8

Notes: * $p < .05$ ** $p < .01$ *** $p < .001$; standard errors are in parentheses. Number of observations: 10,350. Number of groups: Players, 4217; Teams, 146; Co tinents of origin, 6; Seasons, 6; Leagues, 5; Positions, 3.

Fig. 4. Multilevel Regression model for defined attributes [3]

3) Option based pricing

Another data-driven strategy to evaluate the transfer value of a football player can be the Opta Index modeling approach. While statistically determining a player's value, the parameters taken into consideration are goals, assists, dribbles, tackles, form and other performance related points, as seen in previous models. In the Opta Index Model, we make use of "Opta Index", an index calculated by carefully examining almost all the performance determining factors, thus generating a single value which can be said to be a differentiating factor among footballers. Opta, a company set in the 90s, is responsible for generating the Opta Index and are dedicatedly assessing video evidence via an independent assessor for calculating their parameters. A player earns his Game Score through his performance on the field, and the Index score is calculated from the Game score of the last 6 games, with the minutes of gameplay kept as an exposure factor.[18]

The value of an Opta point is given by:

X= V/S

Where V is the turnover of the club and S is the number of Opta points which are satisfying the following criteria:

$$dS / S = \mu_s \, dt + \sigma_s \, dW_s \quad \ldots i$$

Where μs, σs are constants, μs is the expectation return rate, σs denotes volatility, Ws is the standard Wiener process.

If the number of a player's Opta Index points is given by N

Y= NX is the value of the player based on his Opta index points.

For calculating the value of a player based on Opta Index, we apply Ito's Lemma to the above equations before proceeding to the final calculations of results [19]

This Opta Index can further aid the crowd-sourced data based multilevel log-based linear regression model and thus improve on the deficiency of linear regression.

C. Neural Networks

Neural Networks are extremely versatile machine learning tools that can learn features and use the knowledge it has learned to make predictions. Neural Nets have demonstrated their capability by solving many relevant real-world problems ranging from Detection of Bombs in suitcases to predicting myocardial infarction, pertaining to the problem statement, Pricing Football players in the transfer market, which is one of the leading controversial issues among managers, agents, owners and but of course the fans, the author has made use of Neural Nets which utilises several machine learning models such as regularisation, annealing and momentum descent, and predicts the price of the player within 6.32% of his actual price. The model does not take into account goalkeepers as well as outlying players such as Lionel Messi, who has a price tag estimated to be well in excess of $100 million [10].

A multilayer perceptron neural network is designed which is fed a data of more than 15,000 players (gathered from the famous football simulation game FIFA 17), which eventually predicts the price of a football player [11]. The model is optimized by experimenting with different activation functions (such as ReLu, Softmax, Sigmoid etc), number of neutrons and layers, learning rate and its decay etc. The final model achieved an accuracy of 87.2% among 119 pricing categories and places any player within 6.2% of his actual price.

As mentioned, to construct a pricing model for players, data is gathered from FIFA 17, which has official licenses for all the major teams and players. The initial data for every individual player is obtained from a team of 9000 data reviewers comprising of managers, scouts and season ticket holders [12]. Each person in the team rates every player in about 300 categories which are fed to the weighting algorithm to compute the final attributes for each player. The Neural Network model inputs these features of the players as input parameters and computes a price.41 input features are considered from which 37 are on a scale of 0-99,3 on a scale of 1-5 and age on a scale of 16-43[13]. Similar to the split ratio of MNIST dataset, 10,914 players are used as a part of the training set, 1926 for validation and 2500 for the test set.[14]

1) Hyper Parameters:

Activation Functions: A neural network makes use of an activation function to learn Non-linear complex functional mappings between inputs and the response variable. The ideal output should be one-hot encoded in 119 categories, therefore, the activation function for the output layer needs to be between 0 and 1 which gives us a choice between sigmoid function and softmax probability density. On experimentation, ReLU was chosen as the activation function for all hidden layers except for the output layer, for which Softmax was chosen.

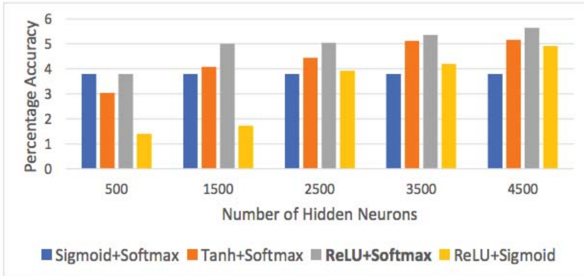

Fig. 5. Accuracy for different Activation Functions when number of Hidden Neurons is varied [17]

Number of Neurons in 1st Hidden Layer:

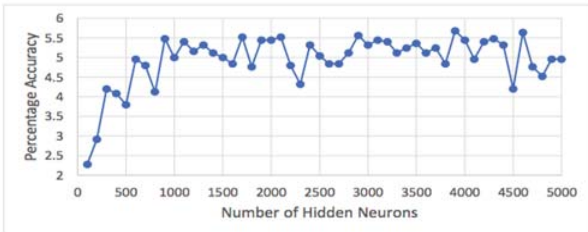

Fig. 6. Percentage accuracy when the number of Hidden Neurons are varied [17]

A proper trend is not recognizable. Maximum accuracy is obtained for 3900 neutrons but once the other parameters are varied, it is observed that 2000 neutrons will be a better choice

2) Learning Rate:

Fig. 7. Variation of Learning Rate. [17]

Initially the learning rate is varied logarithmically but then is switched to linear variations to fine tune it.An optimal value of 0.01 is chosen.

3) Annealing:

The network learns with time and simultaneously cost function gets minimized as per the gradient descent algorithm. If we choose a slow learning rate, the neural network might take forever but at the same time if we select a large learning rate, there is a hazard of oscillating about the minimum point instead of settling in.

We follow the below rule:

$$\eta t = \frac{\eta 0}{1+kt} \quad \ldots \text{ii}$$

Here, *n(0)* is the initial learning rate, *n(t)* is the learning rate after *t* epochs and *k* is the annealing coefficient(0.001).

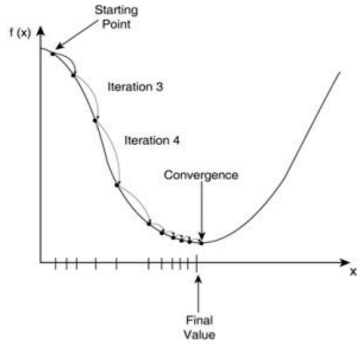

Fig. 8. Gradient Descent [17]

4) Regularization:

There is a good chance that a certain weight gets out of hand and adversely affect the network. To avoid this and maintain the performance of the network, we penalize high value of weights.

III. Results

Based on the study of the three different models as discussed above, we arrive on definite inferences that have been summarized in the table below.

This table summarizes the pros and cons of each model while also stating how accurate the results are. We see that the Multilevel Linear Regression approach provides the most relevant results as compared to the other approaches.

Model	Dataset	Advantages	Disadvantages	Recommended
Crowd-based Estimation	Transfermarkt	The backend of the system has a large number of people calculating the value, thereby increasing the number of inputs for each parameter	1. Since a lot of parameters are involved, there is a high probability of arbitrary parameters not required for the transfer value 2. Making all estimations public reduces the competitive edge of the market, hampering the market value	Low
Multilevel Regression	Transfermarkt	The system has high tending accuracy, and the accuracy can be improved by adding more blocks	The parameters defined for the blocks(player performance, player popularity) do not provide the best outcome.	Highest
Option based	Opta	The system has the best set of defined parameters for player performance, keeping in mind the recent forms and performance, with all analysis done from video footage of the matches	The market value of a player is not solely dependent on a player's performance, and the Option based pricing model ignores economic trends, hence making the value predicted rigid	High

IV. Conclusion

For an application such as pricing an item, it is more important to predict a price close to the actual value instead of getting a exact match which was achieved equally accurately in both the models. We further conclude that the characteristics of greatest importance in determining the transfer price are not only performance related but also depend on other parameters like contract length and popularity. Job mobility also plays a decisive part in the pricing.

Top clubs generally pay more money than the market estimate for attracting top talent; whereas at the same time, a club lacking a player in a particular position often end up paying a few quid more than the estimate, to fill the void. Finally, the amount of games a player plays or the goal scoring opportunities they create increase the transfer worthiness. The impact on the price of a player playing in any of the top clubs increases proportionately more than players with similar attributes playing for smaller teams. Moreover, as discussed about 'Superstar' players, these individuals, although might be just marginally better than a corresponding player, will see a higher positive impact on their transfer fee, in comparison to a similar performance by a corresponding player, while also commanding a much higher transfer fee.

As discussed about top clubs, this research concludes that it is likely that the Premier League clubs, as well as clubs influenced by large investors, pay larger transfer sums, because they are generally wealthier. A case in point is, Real Madrid, FC Barcelona and Atlético Madrid, the top tier clubs in Spain and respectively the first, second and fourth strongest team according to the Euro Club Index, attract players for a lower transfer price, likely due to their popularity. Next, sub-top teams, such as AS Roma, Borussia Dortmund, AS Monaco, Atalanta and Lazio, receive higher transfer prices for players than expected, likely because they sell players to the strongest teams throughout Europe and multiple teams are interested in the same player.

Acknowledgment

We would like to thank Assistant Professor at NMIMS MPSTME, Prof. Krishna Samdani for his constant motivation that helped us achieve our set targets with the highest possible standards. His theoretical based approach toward Research helped us in clearing our fundamentals and culminated in us working and successfully completing this paper.

We would also like to thank the Computer Department of our esteemed institution Mukesh Patel School of Technology Management and Engineering for giving us the opportunity to do research work as part of the curriculum

References

[1] "FIFA Global Transfer Market Report: a new dimension in transfer spending after USD 6.37 billion spent in 2017", FIFA.com, 2018. [Online].Available:http://www.fifa.com/governance/news/y=2018/m=1/news=fifa-global-transfer-market-report-a-new-dimension-in-transfer-spendin-2927376.html [Accessed: 15-Sep-2018]

[2] "Analysis of Performance Measures That Affect NBA Salaries." [Online]. Available: www.diva-portal.org/smash/get/diva2:1114463/FULLTEXT01.pdf. [Accessed: 15-Sep-2018]

[3] Oliver Müller, Alexander Simons, Markus Weinmann, "Beyond crowd judgments: Data-driven estimation of market value in association football", European Journal of Operational Research, Volume 263, Issue 2, pp- 611-624. doi: https://doi.org/10.1016/j.ejor.2017.05.005

[4] R. Pariath, S. Shah, A. Surve and J. Mittal, "Player Performance Prediction in Football Game," 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), COIMBATORE, India, 2018, pp. 1148-1153. doi: 10.1109/ICECA.2018.8474750

[5] R. Stanojevic and L. Gyarmati, "Towards Data-Driven Football Player Assessment,"IEEE 16th International Conference on Data Mining Workshops (ICDMW), Barcelona, 2016, pp. 167-172. doi: 10.1109/ICDMW.2016.003

[6] Garcia-del-Barrio, P. and Pujol, F. (2007), "Hidden monopsony rents in winner-take-all markets—sport and economic contribution of Spanish soccer players" Manage. Decis. Econ., Vol 28, pp- 57:70. doi:10.1002/mde.1313

[7] Football players' transfer price determination based on performance in the Big 5 European leagues. [Online]. Available: https://run.unl.pt/bitstream/10362/18646/1/Geurts_2016.pdf [Accessed: 17-Sep-2018]

[8] Thijs A. Velema, "Upward and downward job mobility and player market values in contemporary European professional football", Sport Management Review, 2018

doi: https://doi.org/10.1016/j.smr.2018.02.004

[9] "Football players' transfer price determination based on performance in the Big 5 Europeanleagues" [Online]. Available: //https://run.unl.pt/bitstream/10362/18646/1/Geurts_2016.pdf [Accessed: 24-Sep-2018]

[10] "Lionel Messi to Chelsea: Barcelona star could cost £500m in total - but only Manchester United and Real Madrid could afford him," [Online]. Available: http://www.independent.co.uk/sport/football/transfers/lionel-messi-to-chelsea-barcelona-star-could-cost500m-in-total-and-only-manchester-united-and-real-9967992.html [Accessed: 1-Oct-2018]

[11] E. Sports, "FIFA 17 - ALL LEAGUES & TEAMS IN FIFA 17".[Online]. Available: https://www.easports.com/fifa/news/2016/fifa-17-leagues-and-teams. [Accessed: 18-Sep-2018]

[12] "FIFA 17's player ratings system blends advanced stats and subjective scouting," ESPN FC, [Online]. Available: http://www.espnfc.us/blog/espn-fc-unitedblog/68/post/2959703/fifa-17-player-ratings-system-blends-advanced-stats-and-subjective-scouting. [Accessed: 27-Sep-2018]

[13] "Players FIFA 19 Oct 23, SoFIFA", Sofifa.com, 2018. [Online]. Available: http://sofifa.com/players/top. [Accessed: 25- Oct- 2018].

[14] Y. LeCun, C. Cortes and C. J. C. Burges, "THE MNIST DATABASE of handwrittendigits," [Online].Available: http://yann.lecun.com/exdb/mnist/[Accessed: 5-Oct-2018]

[15] "Predicting Market Value of Soccer Players Using Linear Modeling Techniques."[Online]. Available: https://Www.stat.berkeley.edu/~, www.stat.berkeley.edu/~aldous/Research/Ugrad/Yuan_He.pdf.[Accessed: 12-Oct-2018]

[16] "Predicting Transfer Values in the English Premier League.[Online]. Available: "http://Dialog.econ.duke.edu/,2016,dialog.econ.duke.edu/honors/wp-content/uploads/2016/09/Dylan-Newman.pdf. [Accessed: 13-Oct-2018]

[17] Dey, Sourya. "Pricing Football Players Using Neural Networks",2017

[18] Tunaru, Clark, Viney "An option pricing framework for valuation of football players"

[19] Qi, Zhijun, Zhiqiang "An Analysis of Football Player Transfer Problems Based on Real Options" [Online]. Available: https://ieeexplore.ieee.org/document/5576988