

Survey on Predicting the Winning Football Team using Machine Learning Algorithms

F. Allwin Edilbert¹, Iwin Thanakumar Joseph²

¹⁻²Department of Computer Science and Engineering, Karunya Institute of Technology and Sciences, Coimbatore, Tamil Nadu, India.

Abstract : The survey's emphasis is on analysing and contrasting football's winning team prediction, the advent of Data Science has provided us with the potential to construct predictive systems with incredible precision. Machine learning is, owing to its incredible efficiency, used in almost all environments in one direction or another. One field where forecasting methods have achieved more attention is the estimation of the outcomes of football matches. This description shows our focus on developing a common predictive model to forecast winning team outcomes. Using machine learning algorithms predicting the results of a football match, we obtained and create set of features, thus developing with high accurate predictive method using machine learning techniques.

1. Introduction:

In over 200 countries, football is played by 250 million players and the world's most common sport. The world's most successful domestic team in the English premier league is several studies and experiments prediction is going on football matches prediction. From analyzing player performance, player chemistry with teammates coaches and etc... And also analyzing team performance how the team perform in wet pitch, dry pitch and weather condition and many more.

By many ways we can analyze and predict the football matches outcome by using twitter we can able to predict the football outcomes. By applying machine algorithms, by applying artificial neural networks data mining, statistical models, deep learning.

For gambling, forecasting football matches is used. The annual football match wager is thought to be legally worth \$10 billion and unlawfully between \$40-50 billion, powered by the global economy as a whole. Mostly, many models are useful for bookmakers. A lot of research is going on in the area of bookmakers.

2. Related Works

Here the study has been done in the area of prediction of winning sports team.

In [1] they have generalized a predictive model in English premier league (EPL), (i.e.. Football league) using exploratory data analysis and feature engineering they have used Machine learning approaches such as Support Vector Machine, Gradient boosting (GB). If 0.2156 on the Ranked Probability Score, the GB produced output (RPS). And they also consider a betting organization where the value obtained by the Graded Probability Score (RPS) is 0.2012 for the same duration.

In [2] they developed a Framework predicting sports result using Machine Learning. The artificial neural network was on (ANN) with that learning methodologies are utilized, data- sources, appropriate-means of evaluation model are used for the results in sport prediction. This architecture (SRP-CRISP-DM) is a Novel sports prediction frame work and it can be used in machine learning technique.

In [3] they proposed that using a public data it is possible to predict football competition with the help of Machine learning approach. Methods used are naive Bayes (NB), multilayer perceptron classifier and principal component analysis (PCA). Here they have achieved an accuracy of 54.702%. This accuracy have achieved by the naive Bayes multilayer perceptron classifier was used in the combination with principal component analysis (PCA).According to bookmaker odds the accuracy was 55.297%.

In [4] they suggested the use of big data and machine learning has developed an expert framework for game prediction. This expert system collect data and analyze the data and the performance and good at identifying the important features for winning and predicting the game outcomes. For collecting the game records from a data source a software module was

developed. The prediction accuracy exceeds above 90% from this it confirms that the Support Vector Machine and Ensemble Machine Learning Algorithm are the most useful for predicting the best results.

In [5] they proposed a conceptual framework to predict the win or lose team in the score game by using study of social networks and gradient boosting (GB). According to experimental result gradient boosting is the best classifier on performance. And considering and comparing the accuracy of two different type prediction and win-draw-lose prediction. Win-draw-lose accuracy is way lower than the win-lose prediction accuracy of win-draw-lose is 33% and win-lose is 50%.

In [6] they proposed football betting and prediction of football matches using machine learning with help of player characteristics based on the results. A Machine learning framework was delivered for achieving excess return amount in betting and for forecasting the football matches. The realized profit is around 30%.

In [7] they proposed predicting the outcomes of pro-league football matches in Iran using an artificial neural network (ANN). Artificial neural networks, maps of self-organization (SOM) and approaches to machine learning. To predict the outcomes of the artificial neural network football matches used (ANN). The outcome shows that five out of six were precisely expected for the AFC Champions League or the AFC Champions League.

In [8] they proposed an innovative approach on Prediction of football matches using the player attributes. And the models used are LSTM regression model, LSTM classification model, and dense model without (LSTM). This approach produce 3 main models structure in overall. From that the LSTM regression model prediction accuracy of 52.479%. Comparatively it is a good result.

In [9] they proposed a predictive model for Prediction of the result of the football match using the twitter data sentiment analysis. By applying Support vector machine (SVM) we can extract tweets and clean the tweets and make it structured for sentiment analysis. Tweets are classified into positive negative and neutral. Implementation of text mining, sentiment analysis and machine learning to predict the team's win-draw-loss ratio. According to their predictive model the accuracy is 67%.

In [10] the proposed use of Neural Network analysis can be used to predict win lose with past games result. Here they collect the past finished matches' attributes and their records to predict the outcome of the match. Where they use artificial neural network for design, train, test. Collected past finished matches attributes for sensitive analysis. For optimization sensitive analysis is done. Neural performance models showed a right win percentage of 83 percent and a loss of 72.7 percent.

In [11] they also created a prolog probabilistic logic software that helps us with the context information of the soccer domain. This method forecasts a team game with the aid of historical data and domain awareness. This probabilistic logic model has better performance than the predictive machine learning algorithm. Observation has been achieved with more parameters and the precision of the model prediction laws.

In [12] they proposed, prediction can be done in international football tournament matches using random forest. Methods used random forest, covariate-based approaches and regression based approaches. They compared two fundamentally different approaches for modeling and prediction of matches in international football tournaments. Random forest and regression are the two different models. Over all analysis shows regression based approach perform lesser than random forest and with regression based approach the prediction performance has variety of measures. Random forest provides a satisfactory result.

In [13] they proposed in sports data predictive analysis is done by using goggle prediction API the overall study of this paper is the usage of Google prediction API to analyses the data. Linear regression, Naïve Bayes, Data mining is used here. Data mining is very essential here. Google prediction API choose the best classifier for predicting the values. This proposed model will give results based on the previously provided data. The more the data better the model and result will be better.

In [14] a novel approach to measuring the on-the-ball contribution of football players to games has been proposed. Here, the passes in games are measured and the possession of the ball by each player is valued. There's a good sense to the Pass process. This approach incorporates dynamic time wrapping with K-nearest neighbor search.

In [15] they introduced a model called DOLORES that predicts the results of football matches around the world. The Dolores model is a combination of complex and hybrid Bayesian networks. The Dolores model will make a good prediction when matching between a and b. With the aid of historical evidence, it is also possible to predict if the opponent's team is not involved in the game. Predictive error is 0.94 per cent higher and 116.78 per cent lower than the bottom participant. Dolores created 20% + ROI on the basis of approximately seven seasons of the English Premier League.

In [16] they suggested that in nearly two-thirds of situations, the winner of the English county's twenty cricket games could be expected. The classification methods used are Naïve Bayes Logistic regression, random forest and gradient enhanced tress. This classification the most effective one is naive Bayes. At last comparison was done in betting odds benchmarks and naïve Bayes player naive Bayes performance was way better than the old benchmarks.

In [17] they proposed using machine learning classifier we can predict the outcome of Dutch football. There are some machine learning algorithms are used generalized boost model (GBM) naive Bayes K-nearest neighbor (KNN). They have used multiple data mining techniques for analysis and prediction. At last the generalized boost model (GBM) made an accuracy of 82.12% correctly. GBM often produce predictive accuracy that cannot be beaten.

In [18] Work is a system that predicts the winner in the sports game. They used greedy logic, data acquisition, data processing, and predictive decision-making approaches. The outcome can differ if the data features are changed. The result indicates that the best classifier is 5 percent more accurate than the referent classifier, which favors the team with a better ranking.

In [19] they suggested a comparison between the result-based on ratings and the goal-based ODM (offense defense model). They applied models to the open national soccer database. They considered as covariates all the combination of result-and globally dependent regression with rating and goal-based ODM rating difference. If it was found that Elo based ordered logit model outperforms all the other models. Including a more nuanced literary model.

In [20] they proposed using machine learning player valuation can be done in European football. Data mining is used. This paper shows that it compares and contrasts the strengths of players, the best statistical abilities in European football. This paper includes the Q function and the effect of a deep reinforcement learning model on the assessment of the player's behavior. This minimized model analyzes players by characterizing the value and in part dependency story. Valuation accuracy F1 top 10 percent, F1 top 25 percent, F1 top 50 percent.

In [21] they Suggested use of long-term memory recurring neural network prediction can be achieved at a glance to train a soccer player. The problem they addressed and they tackled that with a solution. The data has been collected using PMSys system. The dilemma of deriving peaks in the ability of a soccer player to succeed from a subjective self-reported wellness they train recurrent neural networks with long-term memory using data from two professional soccer teams. Their model predicts with precision positive and negative peaks and recalls over 90 percent.

In [22] they proposed two fundamental case of sport analysis one is team performance and player performance prediction. They followed two approaches. And first approach gave them a satisfactory result. And the second approach for team performance gave a remarkable result. They used machine learning algorithms for this. Finally SVM with polynomial kernel observed clearly and achieved a good result in every league.

In [23] they proposed that based on internet information results will be forecasted in football matches. They used some methods based on the weighted sum of indicators, method based on poison distribution, method based on forecasting rules. Method based on weighted sum of indicators predicts match result accuracy is 77.6%. The poison distribution uses another sets and predict match accuracy is 83.5%. Method based on rules of forecasting predict match accuracy of 70.4%. They propose that accuracy is depend upon the forecasting factors. Better the forecasting factor choice better the accuracy.

In [24] they suggested that a prediction could be made to find the winning team using decision tree. And decision tree algorithm is so simple to understand it and it is primarily used for data interpretation and forecasting. Their experiment showed that two teams were playing in recent matches. Considering both the home and the field. The highest accuracy is 57.7%.

In [25] they proposed that predicting DE commitments in college football recruiting using twitter data. Here they used logistic regression. By using this model the result ere promising in first step of predicting DE commitments in college

football. Based on their performance and popularity and skills players will be recruited. Linear combination of parameters and the input vector will be incredibly easy to compute.

In [26] they suggested a novel way of forecasting a soccer match. In this paper, a novel method was developed to forecast soccer matches. Here they demonstrate an alternative sources for curated data. Which can be done through very intensive data collection. This will help the prediction process and result will be very accurate. The main challenge of machine learning project is to collect data by this approach collecting the data will be easy and effective.

In [27] they proposed an open international soccer database for machine learning. Collecting data will be a difficult part in machine learning projects. The more clear data you collect the more accuracy will be appeared. For collecting good data they created a data base and stored about 52 leagues from 35 countries. And also they predicted the result.

In [28] they suggested using Machine Learning to predict football league winners of bookies. Methods used Bayesian Learning, Lazy Learning, Decision Tree, intelligent machine architecture. The result obtained is accurate and established. Finally, it can be concluded that the study requires more practical data and adequate expertise to make the model developed more advanced and accurate.

In [29] using machine learning techniques predicting football. Problem is while predicting the outcome the factors become difficult in processing it so for that they developed a software solution in order to try and solve this problem the solution will be an optimized one. The result shows a satisfactory capability of prediction which I superior to one of the reference method. The accuracy is around 60%.

In [30] they proposed player pairs valuation in ice hockey. This paper covers all about finding the players chemistry level with their teammates in the match. And predicting the best player’s pairs by using data mining analysis and Q function. From coaches data will be collected mainly and they also did hypothesize. Innovative approach has done.

3. Performance Metrics, Methodology, Merits and Demerits of an existing system are:

REF NO	METHODOLOGY	PERFORMANCE METRICS	MERITS	DEMERITS
1	Support vector machine(SVM) Gradient boosting Ranked Probability Score(RPS)	Gradient Boosting-(0.2156) Ranked probability score - (0.2012)	Lower the RPS to better predictive accuracy.	In spite of encouraging results, The model was unable to surpass the bookmaker's predictions.
2	Artificial neural network(ANN) SRP-CRISP-DM frame work	This model predictive accuracy IS effective.	Framework focuses on the prediction of outcomes for team sport rather than individual sport.	According to industry methods this model still lags
3	Naive Bayes or Multilayer perceptron Principal Component Analysis (PCA)	Naive Bayes or a Multilayer classifier of perceptron’s. Was used in conjunction with the key component analysis- (PCA) 54.702 Precision.	It increases the combined accuracy of bookmakers by over 0.7% with the help of public data into the model	This was not carried out for German bundesliga instead of Dutch eredivise.
4	Principal Component Analysis Support vector machine Ensemble Machine learning Algorithm	Ensemble model’s predictive accuracy exceeds 90	Support vector machine(SVM) Ensemble Machine Learning algorithm are useful tool that can reliably predict	Owing to the transfer of teams, coaches and team managers, this model cannot be applied for a single season.

			games outcomes.	
5	Gradient Boosting. Support Vector Machine. Neural Network. Decision Tree. Case-Based reasoning. Logistic Regression.	The result Gradient Boosting(GB) Is the best classifier in terms of efficiency.		
6	Random Forest Boosting. Support Vector Machine. Linear Regression.	Cumulative return 40% from 1993 to 1997 The realized benefit is about 30% for 200 matches in 2013/2014.	This Machine learning framework achieves excess return through appropriate betting.	Baseline methods have not been able to yield any income
7	Radial Basis Function. Neural Network.	A valid win and loss percentage of 83.3 percent and 72.7 percent respectively showed the success of the Neural model. Error testing of 0.37 percent Root is the mean sq. Error if 2.9 percent is needed	The neural network can be used to adjust preparation, strategies and opposition to improve results.	
8	Support vector machine(SVM) Naive Bayes(NB) Text mining Sentimental analysis Predictive analysis	This model predictive accuracy exceeds 67%	SVM is a non-probabilistic binary linear classifier algorithm which takes input and predicts for each given input	Ensemble model can be used
9	LSTM Regression model LSTM classification model Dense model without (LSTM)	LSTM regression model prediction accuracy of 52.479%	Innovative approach of processing data and predict result	Average book makers accuracy
10	Artificial Neural Network. Self-Organizing Map. Machine Learning Approach.	The result indicates that five out of six were correctly predicted to play in the AFC Champions League or the Azadegan League.	ANN is used for complex patterns and prediction problems.	Genetic Algorithm (GA) is advised to set the ANNs parameters optimally.
11	Ada Boost Decision tree. K-Nearest Neighbours. Logistic regression Random Forest Support Vector Machine Probabilistic logic model	Optimized performance from probabilistic logic model	They developed Probabilistic logic program in probLog	It remains an issue to be addressed to construct probabilistic logic models with a large dataset.
12	Random forest(RF) Covariate-based approaches Regression based approach	Random forest provided very satisfactory results.	Random forest is based on a baggage algorithm and uses an ensemble learning technique.	Complexity and long training period.

13	Linear regression(LR) Naive Bayes Data Mining	The proposed model will generate results on the basis of the data previously given. The better the data model, the better the results.	Linear regression has the lower time complexity Naive Bayes is easy to implement.	Under fitting. Main imitations of Naive Bayes is the assumption predictor.
14	XGBoost algorithm K-nearest neighbour(KNN)	This model predictive accuracy was very satisfactory	Paper proposed a novel approach to assessing on-the-ball football players' contributions to games	XGBoost is sensitive to outliers. KNN Accuracy depends on quality of the data.
15	Dynamic ratings Hybrid Bayesian network	Predictive error of 0.94 percent Higher and 116.78 percent lower. Dolores produced 20% + ROI on the basis of approximately seven EPL seasons.	Dolores provides empirical proof that a model for a match may make a good prediction.	In this paper, the model presented had to be limited to the data on the target score.
16	Naïve Bayes Logistic regression Random forest Gradient boosted tress	The most effective method was Naive Bayes	The prediction has done almost with two thirds of instances	External data's can be added to predict the wining.
17	Generalized Boost models(GBM) Naïve Bayes(NB) K-nearest neighbour(KNN)	Using GBM model made is with an accuracy of 82.15% Correctly.	GBM often produce predictive accuracy that cannot be beat.	GBM can overemphasize the outliers and cause over fitting.
18	Greedy logic Data Acquisition. Data Processing. Decision Making prediction Methods.	Modification could lead to better outcomes, as the selection of features will be eliminated.	Easy and effective way to select features and model building.	Few things that can be done to boost the accuracy of the forecast
19	Effect and Goal-Based Regression model. Rating and target is based on ODM ranking. Logit Model Ordered	The ordered logit model based on ELo outperforms all other models.	Effective method are used here	Can include more complex models
20	Data mining. Bayesnet. Naive Bayes. Liner Logistic regression model K-nearest neighbour Decision tree	Accuracy of assessment F1 from the FW F1 from the MD DF's F1 F1 from the GK	Among other items, this prediction result indicates a consistent order in the position based on prediction scores.	Explicitly must take into account.
21	Machine learning algorithms. Recurrent Neural Network(RNN)	The overall performance model is promising.	RNN can model sequence of each sample can be assumed to be dependent on previous one.	Disadvantage is RNN are vanishing gradient and gradient exploding problem.
22	Naive Bayes.	SVM poly kernel is	The risk of over	Larger the

	Decision Tree. Random Forest. K-Nearest Neighbour. SVM-rbf kernel, ploy kernel XGBoost.	observed to steadily achieve good results in every league.	fitting is less in SVM.	datasets target cases are overlapping.
23	Poison distribution. Weighted sum. Forecasting rules.	Poison distribution Predicted results accuracy is of 83.5%.	Calculations are very straight forward than those binomial.	The method based rules of forecasting predicts the results and in which takes a little bit more time.
24	Decision tree algorithm	The highest accuracy is 57.7%.	Decision tree algorithm is easy to understand and to implement.	Small changes in data set can change the structure of decision tree.
25	Logistic regression.	Promising result in first step of predicting DE commitments in college football. Using online social network	Linear combination of parameters and the input vector will be incredibly easy to compute.	Logistic regression can be done with more than two, but not many.
26	Principal component analysis (PCA)	Good alternative source for curated data.	Improves algorithm performance. Reduce over fitting.	Data standardization is must for PCA.
27	Statistical approaches. Poison model.	This has a relatively good database for prediction process.	Easy to collect and implement.	Not available in a wide variety of countries and leagues.
28	Bayesian learning. Lazy learning. Decision tree. Design of intelligent system.	The results obtained are precise and define.	Bayesian learning Provides natural and principal way of combining prior information with data.	Lazy learning it takes large space to store entire training dataset.
29	Naive Bayes. Bayesian Networks. Logit Boost. K-Nearest Neighbour. Random forest. Artificial Neural Network.	Artificial neural network (ANN) with predictive accuracy up to 68% of the network is known to have 5 hidden layers.	ANN is used to model complex and prediction problems.	KNN with large data the prediction stage is slow.
30	Data mining. Analysis. Q-function.	This analysis shows the good chemistry that performed acrossed the seasons	Innovative approach has done.	Could have added more features.

3. Conclusion

Survey on prediction of winning football team is still evolving, due to the difficulty of some data collection. Different scholars and researchers are still working hard to find a full prediction model using multiple techniques. The result from this study will support the same with an individual and team performance. This has been accomplished by various pre-processing, selection of features, statistical model, classification as classifiers, machine learning techniques and exceptional precision in prediction of football.

4. References

- [1]. Baboota, R. and Kaur, H., 2019. Predictive analysis and modelling football results using machine learning approach for English Premier League. *International Journal of Forecasting*, 35(2), pp.741-755.
- [2]. Bunker, R.P. and Thabtah, F., 2019. A machine learning framework for sport result prediction. *Applied computing and informatics*, 15(1), pp.27-33.
- [3]. Tax, N. and Joustra, Y., 2015. Predicting the Dutch football competition using public data: A machine learning approach. *Trans. Knowl. Data Eng.*, 10(10), pp.1-13.
- [4]. Gu, W., Foster, K., Shang, J. and Wei, L., 2019. A game-predicting expert system using big data and machine learning. *Expert Systems with Applications*, 130, pp.293-305.
- [5]. Cho, Y., Yoon, J. and Lee, S., 2018. Using social network analysis and gradient boosting to develop a soccer win-lose prediction model. *Engineering Applications of Artificial Intelligence*, 72, pp.228-240.
- [6]. Stübinger, J., Mangold, B. and Knoll, J., 2020. Machine Learning in Football Betting: Prediction of Match Results Based on Player Characteristics. *Applied Sciences*, 10(1), p.46.
- [7]. Arabzad, S.M., Tayebi Araghi, M.E., Sadi-Nezhad, S. and Ghofrani, N., 2014. Football match results prediction using artificial neural networks; the case of Iran Pro League. *Journal of Applied Research on Industrial Engineering*, 1(3), pp.159-179.
- [8]. Danisik, N., Lacko, P. and Farkas, M., 2018, August. Football match prediction using players attributes. In *2018 World Symposium on Digital Intelligence for Systems and Machines (DISA)* (pp. 201-206). IEEE.
- [9]. Rathan, M., Deepthi, R.N., Anupriya, S. and Vishnu, V., 2018. Football Match Outcome Prediction Using Sentiment Analysis of Twitter Data. *International Journal of Advanced Research in Computer Science*, 9(Special Issue 3), p.78.
- [10]. Hassan, A., Akl, A.R., Hassan, I. and Sunderland, C., 2020. Predicting Wins, Losses and Attributes' Sensitivities in the Soccer World Cup 2018 Using Neural Network Analysis. *Sensors*, 20(11), p.3213.
- [11]. Saritha, M. and Milton, R.S., 2020. A probabilistic logic approach to outcome prediction in team games using historical data and domain knowledge. *JOURNAL OF AMBIENT INTELLIGENCE AND HUMANIZED COMPUTING*.
- [12]. Schauburger, G. and Groll, A., 2018. Predicting matches in international football tournaments with random forests. *Statistical Modelling*, 18(5-6), pp.460-482.
- [13]. Ujwal, U.J., Antony, P.J. and Sachin, D.N., 2018. Predictive analysis of sports data using google prediction API. *International Journal of Applied Engineering Research*, 13(5), pp.2814-2816.
- [14]. Bransen, L. and Van Haaren, J., 2018, September. Measuring football players' on-the-ball contributions from passes during games. In *International Workshop on Machine Learning and Data Mining for Sports Analytics* (pp. 3-15). Springer, Cham.
- [15]. Constantinou, A.C., 2019. Dolores: a model that predicts football match outcomes from all over the world. *Machine Learning*, 108(1), pp.49-75.
- [16]. Kampakis, S. and Thomas, W., 2015. Using machine learning to predict the outcome of english county twenty over cricket matches. *arXiv preprint arXiv:1511.05837*.

- [17]. Hijmans, A. and Bhulai, S., Dutch football prediction using machine learning classifiers.
- [18]. Zdravevski, E. and Kulakov, A., 2009, September. System for Prediction of the Winner in a Sports Game. In International Conference on ICT Innovations (pp. 55-63). Springer, Berlin, Heidelberg.
- [19]. Robberechts, P. and Davis, J., 2018, September. Forecasting the FIFA World Cup–Combining result-and goal-based team ability parameters. In International Workshop on Machine Learning and Data Mining for Sports Analytics (pp. 16-30). Springer, Cham.
- [20]. Nsolo, E., Lambrix, P. and Carlsson, N., 2018, September. Player valuation in European football. In International Workshop on Machine Learning and Data Mining for Sports Analytics (pp. 42-54). Springer, Cham.
- [21]. Wiik, T., Johansen, H.D., Pettersen, S.A., Baptista, I., Kupka, T., Johansen, D., Riegler, M. and Halvorsen, P., 2019, September. Predicting Peek Readiness-to-Train of Soccer Players Using Long Short-Term Memory Recurrent Neural Networks. In 2019 International Conference on Content-Based Multimedia Indexing (CBMI) (pp. 1-6). IEEE.
- [22]. Chazan–Pantzalis, V. and Tjortjis, C., Sports Analytics for Football League Table and Player Performance Prediction.
- [23]. Klyuchka, Y.A., Cherednichenko, O.Y., Vasylenko, A.V. and Yakovleva, O.V., 2017. Forecasting the results of football matches on the Internet based information. Вісник Національного технічного університету ХПІ. Серія: Системний аналіз, управління та інформаційні технології, (55), pp.51-59.
- [24]. Tang, X., Liu, Z., Li, T., Wu, W. and Wei, Z., 2018, August. The Application of Decision Tree in the Prediction of Winning Team. In 2018 International Conference on Virtual Reality and Intelligent Systems (ICVRIS) (pp. 239-242). IEEE.
- [25]. Bigsby, K.G., Ohlmann, J.W. and Zhao, K., 2019. The turf is always greener: Predicting decommitments in college football recruiting using Twitter data. Decision Support Systems, 116, pp.1-12.
- [26]. Shin, J. and Gasparyan, R., 2014. A novel way to soccer match prediction. Stanford University: Department of Computer Science.
- [27]. Dubitzky, W., Lopes, P., Davis, J. and Berrar, D., 2019. The open international soccer database for machine learning. Machine Learning, 108(1), pp.9-28.
- [28]. Esumeh, E.O., 2015. Using machine learning to predict winners of football league for bookies. Int. J. Artif. Intell, 5, p.22.
- [29]. Hucaljuk, J. and Rakipović, A., 2011, May. Predicting football scores using machine learning techniques. In 2011 Proceedings of the 34th International Convention MIPRO (pp. 1623-1627). IEEE.
- [30]. Ljung, D., Carlsson, N. and Lambrix, P., 2018, September. Player pairs valuation in ice hockey. In International Workshop on Machine Learning and Data Mining for Sports Analytics (pp. 82-92). Springer, Cham.