Predicting the outcome of football matches using Convolutional Neural Network

Sahan Dilruk Herath and Siyamalan Manivannan

Department of Computer Science, Faculty of Science, University of Jaffna

sahand.herath@gmail.com , siyam@univ.jfn.ac.lk

Introduction

Football outcome prediction is one of the most challenged area of sports prediction [1]. Due to the larger number of leagues around the world, different people in the world like to know the outcome of football match before it even being played. The outcomes of football matches are important because of the financial assets involved in this. Sports betting is now being established through betting centres and also through sport betting websites such as Bet365, Bet&Win, VCBet, 1xBet, and Betway. Bookmakers and several betters are interested in knowing the result of football games before they have been played. It helps the bookmakers to set up the odds of a win, lose, and draw. In addition, team managers and owners also interested in predicting the results of a game because through that learning they could build strategies and stronger teams against their opponents. Predicting the result of a football game is hard due to a large number of features that affect the outcome of the match. Player attributes, match attributes and team attributes could effect the outcome of a game. Deep learning algorithms like CNN have shown great performance towards predicting the outcomes in difference fields [1,2,3]. In this work we investigate different Convolutional Neural Network (CNN) based deep learning architectures for predicting the outcome of football matches.

Objectives

Objective of this research is to predict the outcome of football matches (win, loss or draw) in European League tournaments using use player features, team features, and match features.

Methodology

Prediction process use three methodologies. Methodologies can be divided according to the type of features used and type of the network that we use for prediction.

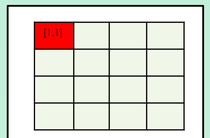
1. Prediction using features of players with CNN
2. Prediction using combination of features of players and team with CNN
3. Prediction using a Siamese neural network

Dataset

For the training, validation and evaluation we used a publicly available dataset of European soccer data which includes match features, player features, and team features of top 10 European leagues from 2009 - 2016. It includes more than 25,000 matches, more than 10,000 player data and more than 300 team data. Player and team attributes of this dataset sourced from EA sports FIFA video game series. Training set includes 16996 matches, and testing set includes 4250 matches. From raining data 46% matches are home wins, 29% matches are home loss and 25% matches are draws.

Preprocessing

For Each match player-player similarity matrix is created for each feature comparing home team players and away team players creating a 11x11 matrix. Combine all these 2D-matrixes and create a 3D-matrix of size 19x11x11 including all 19 features. These pre-processing has used in method 1 and method 2.



Value[1,1]

=

Team 1 player 1 value

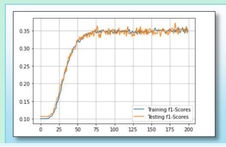
-

Team 2 player 1 value

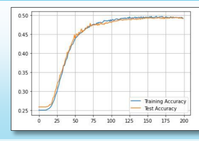
Results and Discussion

Method	Accuracy	F1- score
Method 1	50.28%	0.3465
Method 2	46.09%	0.3623
Method 3	50.96%	0.3730

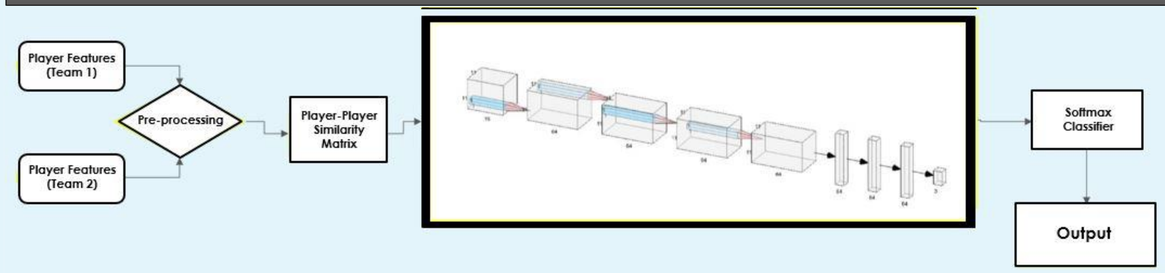
Model trained using SGD optimizer with the learning rate of 0.001 and used cross entropy loss as the loss function. Batch size is set to 128.



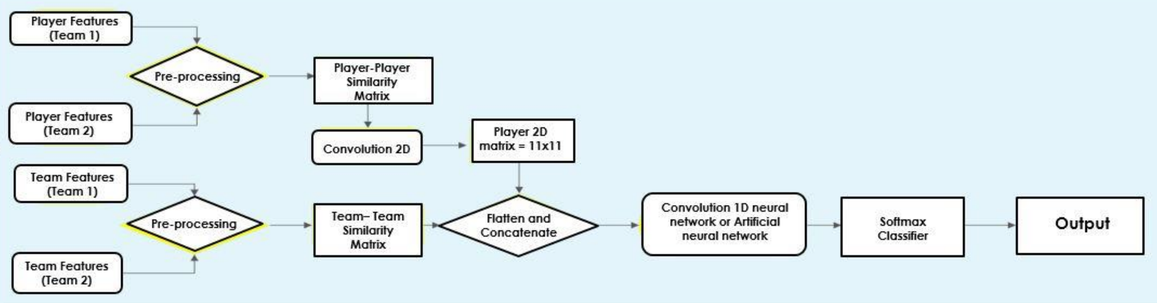
F1-score and accuracy from method 3



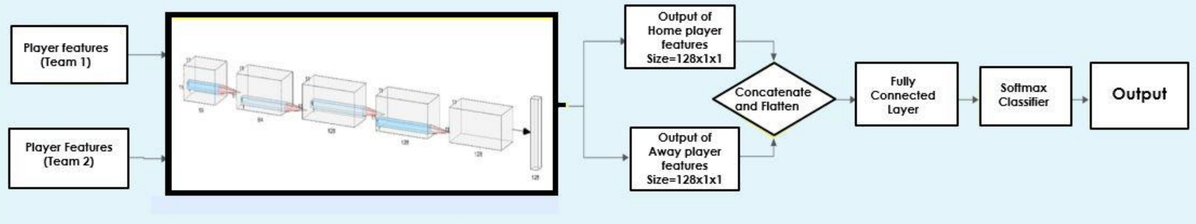
Method 1: Prediction using features of players with CNN



Method 2: Prediction using combination of features of players and team with CNN

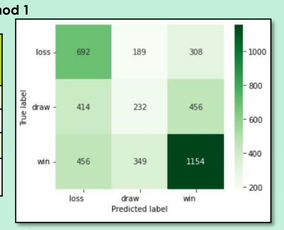


Method 3: Prediction using a Siamese neural network



Method 1

	Precision	Recall	F1-score	Support data
Loss	0.44	0.58	0.50	1189
Draw	0.30	0.21	0.25	1102
Win	0.60	0.59	0.60	1959
accuracy			0.49	4250

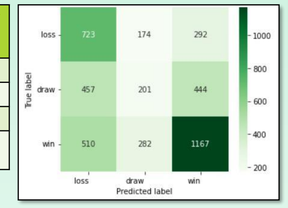


Method 2

	Precision	Recall	F1-score	Support data
Loss	0.34	0.49	0.40	1189
Draw	0.26	0.20	0.23	1102
Win	0.56	0.49	0.52	1959
accuracy			0.41	4250

Method 3

	Precision	Recall	F1-score	Support data
Loss	0.43	0.61	0.50	1189
Draw	0.31	0.18	0.23	1102
Win	0.61	0.60	0.60	1959
accuracy			0.49	4250



All these models show accuracy around 51% and F1-score is around 0.36. But method 2,3 is predicting draws better than method 1.

Conclusion

Predicting the outcome of football matches is challenging in most situations. In this work we proposed three deep learning based approaches for predicting the outcome of football matches using CNN architectures. Our approaches were able to predict the outcome better than chance. We showed that the outcome of the football matches can be predicted with an accuracy of ~50%.

References

1. 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).
2. NA, N., 2021, January. Deep Similarity Learning for Sports Team Ranking. In 2021 Southern African Universities Power Engineering Conference/Robotics and Mechatronics/Pattern Recognition Association of South Africa (SAUPEC/RobMech/PRASA) (pp. 1-6).
3. Rahman, M.A., 2020. A deep learning framework for football match prediction. SN Applied Sciences, 2(2), pp.1-12.
4. Manivannan, S. and Kausik, M., 2019, December. Convolutional Neural Network and Feature Encoding for Predicting the Outcome of Cricket Matches. In 2019 14th Conference on Industrial and Information Systems (ICIIS) (pp. 344-349).