Prediction of German Bundesliga football match results using ensembles of classifiers

Szymon Głowania[1][0000−0003−1137−4350]

Department of Machine Learning, University of Economics in Katowice, 1 Maja, 40-287 Katowice, Poland. szymon.glowania@uekat.pl
http://www.ue.katowice.pl/kum

Abstract. Prediction of sports performance, in particular football, is becoming an increasingly interesting and popular topic for both research and business applications. Available results are mainly based on data from English premier leagues. The current state of knowledge provides opportunities related to improvement of classification accuracy and exploration of other leagues very popular among fans and business environment (e.g. German Bundesliga).
The main objective of the paper is to verify the hypothesis whether the proposed approaches related to the use of heterogeneous ensembles of classifiers will allow obtaining predictions at a higher level of classification accuracy than classical machine learning algorithms, in particular artificial neural networks.
The proposed approach was tested on real data. Appropriate models were trained, and their accuracy was evaluated based on the results of the football matches of the German Bundesliga. The methods selected based on the literature review were experimented. The achieved results of the classification quality assessment measure confirm the realization of the set objective and hypothesis.

Keywords: heterogeneous classifiers · ensembles of classifier · sport result prediction · neural networks.

1 Prediction of sport results using machine learning

Artificial intelligence and machine learning are finding more and more applications in fields related to economy, politics, science, but also entertainment. Researchers, specialists, and enthusiasts alike are finding more and more interesting applications for the available algorithms.

The high availability of data and the growing use of sports analytics is leading to the popularization of these applications and the application of ever newer tools to them. The amount of investment by companies both in the teams or players themselves, and in parts of the sports entertainment industry, is leading to the increasing use of machine learning algorithms in this area. The use of machine learning can be found in a wide variety of sports, such as football [1, 13], basketball [11], rugby [9]. Regarding number of publications (based on to

scopus.com, as of 05.03.2022), football remains the most popular sport in which machine learning is applied. The most analyzed European league is the English Premier League [1], however, in the literature you can also find studies based on other leagues that are very popular among fans, such as the German Bundesliga [13]. This paper focuses on the analysis of the German Bundesliga due to its popularity and continuation of previous research [8].

2 Machine learning algorithms selected for experimentation

The most common approach to predicting sports performance is classification. In this approach, using data and selected algorithms, it is possible to predict the membership of a new object to one of the predefined classes. In the analyzed example, three classical classes possible for the end of the match are selected: home team win (1), visiting team win (2) and draw (0).

Classical machine learning algorithms, which are also the most popular in sports performance prediction, were used in the study. Two original approaches based on heterogeneous ensembles of classifiers are also presented to improve classification performance.

The first of the selected algorithms is decision trees. The tree presented in this paper was created based on [3, 7]. The maximum depth was set to 3, and the algorithm chosen was CART.

Another frequently used approach is the support vector machine. The linear classifier is chosen, and its implementation is consistent with [4].

Artificial neural networks are some of the most popular algorithms that are now widely used in both science and business due to the increase in computing power of computers. For the experiment, networks were created with the structure: 6–3–2–1 and 6–3–1. L-BFGS was chosen as the solver, and the maximum number of iterations is 100. The solution is implemented according to the publication [12].

Classifier ensembles are an approach based on building a classifier using a collection of individual classifiers. This approach can significantly improve the efficiency of the solution, as presented in, e.g. [6, 10].

Boosting and AdaBoost is an algorithm whose idea of operation is based on the creation of many simple classifiers, but through their number on the improvement of representativeness and consequently quality. The fathers of this approach are considered to be Y. Freund and R. E. Schapiro, and the algorithm used in this paper was implemented according to their publications [5].

Bagging is a method developed by L. Breiman to improve classification and performance of regression models due to accuracy and stability by reducing variance. The experiment uses an implementation that is consistent with the publication [2].

Random forests are a method of building an ensemble of classifiers proposed by L. Breiman. This approach is based on creating multiple decision trees and then creating a single classifier from them. A random forest with a maximum

tree depth of 3 and a number of estimators equal to 100 was chosen for the experiment, the entire operation of the algorithm is as published [3].

In this paper, two original approaches based on heterogeneous ensembles of classifiers [8] are used. The first ,,all" approach is to use previously trained classifiers: decision tree, support vector machine, AdaBoost, bagging and random forests to build a new classifier. The decision class was selected by simple voting. In the subsequent ,,selected" approach, the same algorithms were used, but the class decision had to be made unanimously by the algorithms used. This approach does not guarantee complete coverage of all decision cases, but can achieve high accuracy.

3 Conducted experiments

The necessary data set was retrieved using a solution prepared by the author. The data concerns the results of individual games and league tables in the seasons from 2011/2012 to 2020/2021 of the German Bundesliga. They cover a total of 340 matches from 10 years, resulting in 3060 records for analysis.

The experiments were conducted using the train and test method. The data was divided into a training set and a test set. The training set contained data up to and including the 18th round of the 2020/2021 season, while the test set contained data from the 19th to the 34th round of the 2020/2021 season. In accordance with previous research and the literature review, observations from the first 5 rounds of each season were removed from the set [8].

The data available about the meeting and the league table after the meeting were used in the experiment. The selected attributes are: ,,round", ,,home team's place in the league table", ,,total points scored by the home team", ,,visiting team's place in the league table", ,,total points scored by the visiting team" and ,,difference of points scored by the home team and visiting team".

The tested models were evaluated for quality using classification accuracy. The measure was calculated according to the formula $accuracy = \frac{(TP+TN)}{(TP+TN+FP+FN)}$ and using the confusion matrix. The results of accuracy for algorithm: Random=33,33%, Artificial neural networks 6–3–1 = 49,67%, Bagging = 49,67%, Support vector machine = 54,90%, Decision tree = 56,21%, Artificial neural networks 6–3–2–1 = 56,21%, Random forests = 56,86%, AdaBoost = **59,48%**, Approach (all) = **59,48%**, Approach (selected) = **61,54%**.

Each of the algorithms used achieved higher accuracy than the random approach. Regarding classification accuracy, the best result was obtained by the ,,Approach (selected)". In this solution, there is no guarantee of complete coverage of the test set. The next two approaches ,,Approach (all)" and ,,AdaBoost" will perform well when a decision is needed for each observation and their prediction accuracy is also at a high level.

The experiments show that the use of both classical machine learning algorithms and the proposed approaches based on heterogeneous ensembles of classifiers allows obtaining satisfactory results in terms of prediction accuracy.

4 Summary and further work

The main objective of the paper was to verify whether the proposed ,,all" and ,,selected" approaches using heterogeneous ensembles of classifiers will allow obtaining predictions at a higher level of classification accuracy than classical machine learning algorithms, in particular artificial neural networks. The goal was achieved, and the research hypothesis was confirmed by the conducted experiments. The best results in terms of classification accuracy were obtained by ,,Approach (selected)", and the most versatile in application are: ,,Approach (all)" and ,,AdaBoost".

Further work is planned to experiment with heterogeneous ensembles of classifiers by incorporating additional voting methods and extending the set with additional features.

References

1. Baboota, R., Kaur, H.: Predictive analysis and modelling football results using machine learning approach for english premier league. International Journal of Forecasting **35**(2), 741–755 (2019)
2. Breiman, L.: Bagging predictors. Machine learning **24**(2), 123–140 (1996)
3. Breiman, L.: Random forests. Machine learning **45**(1), 5–32 (2001)
4. Fan, R.E., Chang, K.W., Hsieh, C.J., Wang, X.R., Lin, C.J.: Liblinear: A library for large linear classification. the Journal of machine Learning research **9**, 1871–1874 (2008)
5. Freund, Y., Schapire, R.E., et al.: Experiments with a new boosting algorithm. In: icml. vol. 96, pp. 148–156. Citeseer (1996)
6. Hansen, L.K., Salamon, P.: Neural network ensembles. IEEE transactions on pattern analysis and machine intelligence **12**(10), 993–1001 (1990)
7. Hastie, T., Tibshirani, R., Friedman, J.H., Friedman, J.H.: The elements of statistical learning: data mining, inference, and prediction, vol. 2. Springer (2009)
8. Kozak, J., Głowania, S.: Heterogeneous ensembles of classifiers in predicting bundesliga football results. Procedia Computer Science **192**, 1573–1582 (2021)
9. McCabe, A., Trevathan, J.: Artificial intelligence in sports prediction. In: Fifth International Conference on Information Technology: New Generations (itng 2008). pp. 1194–1197. IEEE (2008)
10. Opitz, D., Maclin, R.: Popular ensemble methods: An empirical study. Journal of artificial intelligence research **11**, 169–198 (1999)
11. Pai, P.F., ChangLiao, L.H., Lin, K.P.: Analyzing basketball games by a support vector machines with decision tree model. Neural Computing and Applications **28**(12), 4159–4167 (2017)
12. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning representations by back-propagating errors. nature **323**(6088), 533–536 (1986)
13. Schauberger, G., Groll, A., Tutz, G.: Modeling football results in the german bundesliga using match-specific covariates (2016)