Implementation of the game model of the Polish Ekstraklasa team using machine learning techniques

Jakub Pogodziński
Faculty of Mathematics and Computer Science
Adam Mickiewicz University
Uniwersytetu Poznańskiego 4 Street, 61-614 Poznań, Poland
Email: jakpog1@st.amu.edu.pl

Abstract—**The aim of the thesis is to create a model defining the style of play of a team playing in the Polish Ekstraklasa**

I. Introduction

THE AIM of the thesis is to create a model defining the style of play of a team playing in the Polish Ekstraklasa. The limitation to the highest Polish league class is dictated by the differences in the style of play depending on the league. The model is to be created on the basis of data about the team's game (e.g. XGoals, xAssists, distance run). The data will mainly come from the StatsBomb portal. To build the model, supervised and unsupervised learning techniques will be used and compared to find the relationship between the team's statistics and the determination of its playing style. Supervised learning requires expert knowledge to determine the relationship that is being sought. Unsupervised learning can allow to find new features that are not yet named and the relationship between them has not yet been noticed.

II. Potential applications

Building a model of the team's game will allow you to identify the strengths and weaknesses of the team, to prepare tactically for a particular opponent and to find players who match the style of play of the team.

III. State of research

In the article "Introducing Similar Team Search" [1]. the possibility to search for similar clubs based on statistics is shown. There is only the possibility to compare clubs. There are no other correlations between them. In the article Balla M. Which teams are the most similar to Bayern Munich? [2]. (https://totalfootballanalysis.com/article/teams-similar-bayern-munich) in the magazine Total Football Analysis there is a comparison of clubs that are the most similar to Bayern Munich. The author writes that 25 metrics from WyScout were taken for analysis, without specifying which metrics, and then using the K-means algorithm the data were grouped. The distance between the sets was then counted. The article by Marton Ball throws an interesting perspective on comparing ensembles. Unfortunately, it is not given in sufficient detail how the conclusions presented were reached.

IV. Demonstration of the need for own research

The articles presented above are very poor and only present a concrete way of comparing the teams' playing styles. What is missing is a more universal approach, a model that will describe the differences and similarities between teams in a more global way.

V. Discuss the planned experiment - e.g. what the training data will be, what the test data will be, what methods / algorithms will be used

The data will come from the StatsBomb portal. These will be individual player metrics and aggregate metrics for the whole team in each league game over the past few seasons. They will be split into a training and testing set. The data will be grouped and a comparison of clustering algorithms (Ward, K-means) and principal component analysis algorithms will take place. The data will be compared with supervised learning algorithms, where labels will be assigned to the data based on the expert knowledge gathered.

VI. Summary - what we expect to gain from the project and why our experiment is worth doing

In conclusion, football is increasingly based on statistics and analysis. New tools allow a better understanding of the game. Better understanding translates into more effective work on tactics and better sporting, financial and social (fan satisfaction) results. This project aims to create a model to determine a team's playing pattern, highlighting its strengths and weaknesses. Football teams will be able to adapt better to matches, thus increasing their chances of achieving victory both from a one-match perspective and by winning more matches, they will increase their chances of achieving the best possible final result.

References

[1] Statsbomb, "Introducing Similar Team Search" *https://statsbomb.com/2020/12/introducing-similar-team-search*

[2] Balla M., "Which teams are the most similar to Bayern Munich?" *https://totalfootballanalysis.com/article/teams-similar-bayern-munich,*