

Bundesliga Predictor

Deliverable 3

Note: I am sorry that this deliverable is somewhat short and brief, yet I am trying to maximize the time on actually coding the project so that my accuracy improves. Please excuse any mistakes or typos in this document and feel free to reach out to me if you have any questions. At the moment, I have a very clear vision about where I want this project to go, it is only about finding the time to complete everything on time. Finally, and this goes without saying, the last deliverable will be of much higher quality. Thank you very much.

1. Final Training Results

In my two predictor files (Bundesliga_Predictor_1819 and Premier_League_Predictor_1819), which are almost the same, I have reached an approximate 60% accuracy. I would recommend looking at the Premier League file. In the end, I have decided to take a slightly different approach and compare multiple classifiers and their performance.

- 1) The linear SVC is rather unstable and predicted an accuracy from around 40% - 60% with 49% being the mean.
- 2) The Random Forest Classifier predicted with an accuracy score of 55%, yet the data might be overfitted since the training accuracy score was around 98%.
- 3) The XGBoost that was my initial decision as the main classifier performed with an accuracy score of around 58%.
- 4) Surprisingly enough, the multinomial naive Bayes classifier performed with an even better accuracy of cca 60%.

I am currently trying to achieve even better accuracies (at least close to 75%) with improvements including adding past k Home/Away goals, corners, shots, where k is the number of preceding matches. I am currently programming a model that will figure out the best value for k. Moreover, I am planning on adding more features (such as shots on target, bookings...) to "X" if the accuracies achieved are still not satisfactory.

Todo list & questions:

- 1) Find the confusion matrix. Should I do this?
- 2) Find the most important hyper-parameters in my set after additional features are added. Should I do this?
- 3) Figure out how to save my weights.
- 4) Integrate my project into a Web App using Flask probably. Is it okay that the application will not be interactive and only show results?

2. Final Project Demonstration

I am planning on uploading the upcoming games to this page and the predictions of each of my models. I would furthermore like to analyze the upcoming and their precisions. I am hoping to achieve a result of 7/10 in Bundesliga and Premier League. This app will not be interactive but rather provide a list of upcoming games in each league and the given predictions.

Note 2: There are no upcoming football games until the 29th of March, therefore that will be the first time I will be able to put my models to test with the true unpredictability of this beautiful game.