

# Machine Learning for Football Betting

Alexander Kostavelis\*

June 2025

## Abstract

This thesis investigates the performance of two supervised machine learning models, XGBoost and a feedforward neural network, to predict whether a football match will have more or fewer than 2.5 goals. Before the analysis, we present the mathematical background of the models. Using historical data from the top five European leagues downloaded from Football-Data.co.uk, relevant features are engineered and used to train the two models. Their performance is evaluated in terms of prediction accuracy as well as profitability when simulating bets on the over/under 2.5 goals market in the current 2024/2025 season. Although the results are not entirely satisfactory in terms of classification performance, the models show signs of profitability and should be investigated further. We propose several directions for further improvement, with the main suggestion being to incorporate more representative features that better capture the different aspects of match tactics and play styles, as a model is only as good as the data it learns from.

---

\*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden.  
E-mail: [akostavelis3@gmail.com](mailto:akostavelis3@gmail.com). Supervisor: Ola Hössjer and Johannes Heiny.