Vrije Universiteit Amsterdam
Case Study in Econometrics and Data Science, 2023

Comparing football match forecasting performance between neural networks and score-driven time series models

Research proposal
Group 3 - Topic 6
Jan Koolbergen - 2667333 - j.j.koolbergen@student.vu.nl
Romnick Mozes Evangelista - 2709596 - r.m.evangelista@student.vu.nl
John Mossaad - 2792444 - j.mossaad@student.vu.nl

12-01-2023

Motivation

Football match betting is a hugely popular activity, due to the thrill and excitement it adds to the viewing experience, with the potential for financial gains. It is then natural that there have been several attempts to predict football matches using econometrics models. One such example is Koopman and Lit (2017), where the authors used score-driven time series for forecasting football match results. Another example is Karlis and Ntzoufras (2008), where the authors assume a Skellam distribution for the difference in goals. These approaches however are not able to capture complex non-linear relationships, which might prove useful when trying to forecast the match outcomes. We therefore consider using a neural network, as these may be able to capture those intricate non-linear relations and complex patterns in the data. In this research we aim to compare the score-driven time series models of Koopman and Lit (2017) to neural network methods in terms of prediction performance.

Research Question

How does the performance of artificial neural networks compare with that of score-driven time series models in forecasting football match results?

Plan

1. We initiate our research by delving into the existing literature to gain knowledge about different neural networks, specifically, neural networks with some form of short term memory. Short-term memory in football statistic is crucial in capturing recent patterns and dynamic changes, enabling more accurate forecasts.

2. Second, we train the found neural network(s) on the same database used by Koopman and Lit (2017) (can be found here[1]). We also implement the methods from the paper, so that we can compare them.

3. Finally, we compare the different forecasts of our neural network method and that of Koopman and Lit (2017) using multiple out-of-sample measures and statistics such as the Diebold and Mariano (1994) test, succes ratio, and directional accuracy.

References

Diebold, F. X. and R. S. Mariano (Nov. 1994). *Comparing Predictive Accuracy*. Working Paper 169. National Bureau of Economic Research. DOI: 10.3386/t0169.

Karlis, D. and I. Ntzoufras (Aug. 2008). "Bayesian modelling of football outcomes: Using the Skellam's distribution for the goal difference". In: *Ima Journal of Management Mathematics - IMA J MANAG MATH* 20. DOI: 10.1093/imaman/dpn026.

Koopman, S. J. and R. Lit (2017). "Forecasting Football Match Results in National League Competitions Using Score-Driven Time Series Models". In.

[1] https://www.football-data.co.uk