Return predictability with machine learning methods: a short summary

Fredrik Leonardo Nilsson

In my master's thesis, I evaluate the predictive performance of machine learning methods against traditional asset pricing models in forecasting stock returns. In addition, I assess the economic relevance of these predictions by constructing high-minus-low portfolios, to explore whether machine learning methods can be used to profitably extract signals from the market.

I evaluate nine methods, which include OLS, the Fama-French 5-factor model, variable selection methods such as Lasso and Elastic Net, dimensionality reduction techniques such as Principal components regression and Partial least squares, as well as non-linear methods based on gradient boosting (Lgbm and XGBoost) and feed-forward neural networks. I used both a fixed and recursive estimation scheme. The data spans 65 years, from 1957 to 2021, with monthly frequency and over 4 million stock observations with more than a hundred regressors, which include both firm characteristics and macroeconomic variables. For each method, I assess its predictive performance using out-of-sample R^2. The fixed estimation method shows the predominance of variable selection methods, which achieve a R^2_{oos} of 0.34. When switching to the computationally-intensive recursive window, R^2_{oos} generally increases and non-linear methods become the best-performing ones, topped by neural networks with R^2_{oos} equal to 0.49. The results are coherent and statistically meaningful.

Next, I construct zero-cost, decile-based portfolios with equal and volatility weights, and compute informative portfolio statistics. The predictive performance of machine learning methods is translated into significant economic gains. The equal-weighted XGBoost portfolio achieves an annualized Sharpe ratio of 1.19, yielding cumulative returns that are three times higher than those of the S&P 500. The volatility-weighted neural network portfolio achieves a Sharpe ratio of 0.84. Interestingly, the worst-performing portfolios are those with the lowest turnover, suggesting that the performance of machine learning methods comes in large part from frequent portfolio adjustments. Machine learning portfolios seem to work well especially in periods of crisis, during which they outperform the market index the most.

Portfolio performance significantly improves when constructing centile-based portfolios, to better exploit the large dimensions of the dataset. The XGBoost portfolio achieves a Sharpe ratio of 1.49, and the neural network portfolio of 1.40. The higher concentration of high-quality signals in the extreme centiles allows centile-based portfolios to deliver better risk-adjusted performance, justifying the move towards more granular portfolio construction.