The Uncertainty of Machine Learning Predictions in Asset Pricing

Yuan Liao, Xinjie Ma, Andreas Neuhierl,& Linda Schilling

Discussion at the 2025 WFA by
ROHIT ALLENA
University of Houston

June 23, 2025

Summary of the Paper

Methods for computing forecast confidence intervals (FCI) for stock return forecasts obtained from Neural Networks (NN)

1. Method 1:
   - NN forecasts $\xrightarrow{d}$ Fourier series regression forecasts
   - Standard errors (SEs) of NN forecasts = SEs of Fourier forecasts
   - Thus, closed-form expressions are available for the NN forecast SEs
2. Method 2:
   - k-step Bootstrap method
   - Retrains a previously trained NN for additional k steps
   - Bypasses the computation burden of conventional bootstrap methods, which require repeated training of NNs over many bootstrap samples

Empirically, portfolios that exploit FCIs exhibit improved performance

Comments

Contribution and Related Literature

- The paper's main insights are similar to Allena (Forthcoming, RFS)
- Contributions of Allena (2025):
  1. Estimate FCIs for expected return forecasts obtained from penalied linear (Linear, Lasso and Ridge) and NN models
  2. Dropout procedure for obtaining FCIs for NN forecasts does not involve retraining NNs multiple times
  3. Establish why and how incorporating information from FCIs yield improved portfolios
- Hence, the paper benefits from emphasizing novel insights

My discussion focuses on potential new insights

Discussion on Methodology

1. Deriving closed-form expressions for NN FCIs could be a major advance
2. But need to clearly layout the conditions under which the derivation holds
3. Derivation holds only for carefully-tuned NNs that have
   - Bounded complexity
   - No L_1 or L_2 penalties, early-stopping, and SGD
   - But the empirical section imposes $L2$ and SGD
4. Predictors (x_{it}) are assumed to cross-sectionally independent
   - But the empirical section normalizes predictors based on cross-sectional ranking. So, by construction, x_{it}s are cross-sectionally dependent
   - Even without normalization, x_{it}s (e.g., momentum) are likely to have significant common cross-sectional variation

Discussion on Methodology

Proposed FCIs are asymptotically justified.

- How do the FCIs perform in small samples?
- The paper forecasts returns using neural networks but estimates standard errors using Fourier series regressions.
- If NN estimator $\xrightarrow{d}$ Fourier estimator, why not use Fourier directly for forecasting returns?
- The authors argue that in small samples, Fourier forecasts are more biased—hence the use of NNs for return prediction.
- But wouldn't the same logic apply to standard errors as well?
- Could the Fourier-based SE estimators also exhibit finite-sample bias?
- Simulations suggest no finite-sample bias for the Fourier SEs.
    - What do simulations indicate about the finite-sample bias in Fourier return forecasts?
    - If return forecasts are biased but SEs are not, the paper could emphasize this finding.

Discussion on Empirical Framework

Uncertainty-averse portfolios that incorporate FCIs were previously investigated

- So, the paper benefits from advancing previous applications

Comment on Long-Only FCI Portfolios:

- Among top-K stocks with relatively highest return forecasts, FCI-FDR portfolio consists of the subset that satisfies the FDR threshold.
- The paper argues that FCIs matter by showing that FCI-FDR portfolios outperform the top-K portfolio.
- However, the top-K portfolio may not be the appropriate benchmark. A more suitable comparison would be with the top-S portfolio, where S is the number of stocks selected by the FCI-FDR rule.
- The difference in performance between the top-S and FCI-FDR portfolios can be interpreted as the value of incorporating FCIs.

Conclusion

- This is a nice paper. I learned a lot from reading it.

- The paper would benefit from further sharpening both its theoretical and empirical contributions.