Project Proposal

CFRM421: Machine Learning for Quantitative Finance
Anderson Lee
Using Ensemble of MLPs for Stock Return Prediction

Group

1. Anderson Lee

Introduction

The stock market has been a popular field where previous works has intensive and extensive research on predicting prices and classifying trends using multiple fundamental machine learning techniques. Commonly used techniques include SVM[1, 6], Logistic Regression[6], [2], and Transformer[4]. In addition, multimodal machine learning that aims to include textual information from financial news is also invested.[3] Although RNNs, LSTM, and the modern transformers are believed to be effective in time-series data like stock market, it is recently shown that the traditional linear layers might exceed the performance of these modern transformers.[5] As motivated by such work, the goal of this project is to predict mid-term and long-term stock return by an ensemble of multiple linear models that look back the previous data in different time window sizes.

Data

The raw data comes from Yahoo Finance. It includes daily trading information such as volume, low price, high price, open, close, and so on. In addition, it includes monthly evaluation measures such as PeRatio, ForwardPeRatio, EnterpriseValue, and so on. In a larger scale, it includes quarterly data like Revenue, Profit, BasicEPS, and so on. The preprocessing might involve creating additional valuable features like MCAD and dimensionality reduction with PCA or AutoEncoder.

Concern

The linear models proposed in [5] is really simple. It's basically a single linear layer with a slight adjustment. Alternatively, I can also replace the linear models with LSTM and perform the same ensemble techniques.

References

[1] Noguer i Alonso, M., Batres-Estrada, G., & Moulin, A. (2020). Deep Learning for Equity Time Series Prediction. Gilberto and Moulin, Aymeric, Deep Learning for Equity Time Series Prediction (November 23, 2020).

[2] Daiya, D., & Lin, C. (2021, June). Stock movement prediction and portfolio management via multimodal learning with transformer. In ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 3305-3309). IEEE.

[3] Li, Q., Tan, J., Wang, J., & Chen, H. (2020). A multimodal event-driven lstm model for stock prediction using online news. IEEE Transactions on Knowledge and Data Engineering, 33(10), 3323-3337.

[4] Liu, J., Lin, H., Liu, X., Xu, B., Ren, Y., Diao, Y., & Yang, L. (2019, August). Transformer-based capsule network for stock movement prediction. In Proceedings of the first workshop on financial technology and natural language processing (pp. 66-73).

[5] Zeng, A., Chen, M., Zhang, L., & Xu, Q. (2022). Are transformers effective for time series forecasting?. arXiv preprint arXiv:2205.13504.

[6] Dai, Y., & Zhang, Y. (2013). Machine learning in stock price trend forecasting. Stanford University, http://cs229.stanford.edu/proj2013/DaiZhang-MachineLearningInStockPriceTrendForecasting.pdf. Erişim Tarihi, 21, 2021.

[7] Das, A., Kong, W., Leach, A., Sen, R., & Yu, R. (2023). Long-term Forecasting with TiDE: Time-series Dense Encoder. arXiv preprint arXiv:2304.08424.