

Analysis of Limit Order Book Data with Deep Neural Networks

Patrik Reizinger^{1*}, Bálint Gyires-Tóth¹

¹Budapest University of Technology and Economics, Department of Telecommunications and Media Informatics

*E-mail: rpatrik96@sch.bme.hu

Keywords: limit order book, neural networks, deep learning

In recent years, the advancements of information technology made it possible to collect, store and analyze data on a large scale. However, the vast amount of information originating from diverse sources is not simple to explore. Having the means does not imply a full understanding of the underlying process, as in some cases, it has extremely complex dynamics, such financial markets do. Nonetheless, being able to develop a better understanding of those complex phenomena could be of enormous advantage, as it may help to detect anomalies or even to prevent financial emergencies, as witnessed in the last decade.

Recently, significant research interest was devoted to the topic in the machine learning/deep learning community as well, thus the new generation of methods was introduced into finance besides traditional statistics, operating generally with linear models like ARMA or ARIMA. Besides regression, the other common task is to try to predict the direction of the change in the price of an asset (increasing/decreasing/stagnating). [1] investigated price arbitrage on the S&P500 dataset using random forests, gradient-boosted trees, and deep neural networks. [2] carried out a large-scale study with 500 GPUs to extract information from stock prices. The authors of [3] used Variational Autoencoders to detect anomalies in time series, i.e. local sequences which cannot be well explained by the process generating the majority of the sequence. Besides that, several researchers have focused on a specific segment of financial data, i.e. onto LOB (Limit Order Book). LOB is a multilevel representation of the available bid (buy) and ask (sell) orders at given prices and timesteps. LOB is suitable for capturing high-frequency information, thus its gain of popularity (see [4], [5], [6], [7]). Besides that, some authors have used another approach: they – emphasizing the information captured by the visualized LOBs – utilized 2D convolutional neural networks, which are common in image processing ([8], [9]).

Despite the significant progress in the field, a common bottleneck is hard to overcome as the phenomena being modeled are rather complex, a huge amount of data should be used. However, researchers do not have access to high-quality datasets in general. Nonetheless, [10] introduced an open-source, anonymized dataset from Nasdaq Nordic, which can be utilized for benchmarking. The same research group tested several models from linear regression to MLP or recurrent neural networks ([10], [7], [6]), but they excluded from their research the newest advancements regarding recurrent architectures.

Thus, this work aims to extend the benchmarks on the dataset of [10] to novel architectures. As the benchmark dataset mentioned above consists of multiple variants of the same data, we opted for the version excluding the auction period. Furthermore, our experiments did not equalize the classes. The focus of our research interest was, besides multilayer LSTM [11] and 1D convolutional VAE, the Transformer architecture of [12], which utilizes the attention mechanism. Besides using their vanilla variants, both the Convolutional VAE and the Transformer have undergone some architectural tailoring. In case of the former, attention mechanism was introduced to 1) make the training cycle of the VAE simpler by eliminating the need of a separate training phase to minimize MSE reconstruction error and 2) to utilize the selectivity of attention during prediction to capture the most important information and neglect less important. Regarding the Transformer architecture, as it was proposed for Natural Language Processing tasks, more significant changes were made. First, as in [13], the decoder part was removed due to the different nature of the task – this former our baseline for the Transformer. Furthermore, we created a downscaling variant of this Transformer, as we argue that it is not computationally efficient to make the downscaling in one step with the classifier head.

As a result of our experiments on the database of [10] utilizing the architecture mentioned above, we report a significant increase in the F1-score compared to the results of [7] and [6]. Furthermore, we even achieved the performance of [6] on balanced data (where the authors have used technical features for their best results), although we restricted ourselves to the unbalanced case, and used only LOB data (which includes the raw prices and volumes for 10 levels and other features calculated from those, for details see [10]). Comparing different architectures, the Transformer baseline performed the best (0.583), followed by its downscaled variant (0.556) – note that as the latter had 25% fewer parameters, the change is not that significant. The Convolutional VAEs performed quite the same (vanilla: 0.535, attention variant: 0.513), while the LSTM was the worst (but had higher F1 than the networks in the benchmarks) by achieving 0.466.

Acknowledgments

The research considered here has been supported by the European Union, co-financed by the European Social Fund (EFOP-3.6.2-16-2017-00013, Thematic Fundamental Research Collaborations Grounding Innovation in Informatics and Infocommunications).

References

- [1] C. Krauss, X. A. Do, and N. Huck, “Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500,” *European Journal of Operational Research*, vol. 259, no. 2, pp. 689–702, 2017.
- [2] J. Sirignano and R. Cont, *Universal features of price formation in financial markets: perspectives from Deep Learning*. Available: <https://arxiv.org/pdf/1803.06917>.
- [3] C. Zhang and Y. Chen, *Time Series Anomaly Detection with Variational Autoencoders*. Available: <https://arxiv.org/pdf/1907.01702>.
- [4] Z. Zhang, S. Zohren, and S. Roberts, “DeepLOB: Deep Convolutional Neural Networks for Limit Order Books,” <https://arxiv.org/pdf/1808.03668.pdf>.
- [5] J. A. Sirignano, “Deep learning for limit order books,” *Quantitative Finance*, vol. 1, no. 1, pp. 1–22, 2018.
- [6] P. Nousi *et al.*, “Machine Learning for Forecasting Mid-Price Movements Using Limit Order Book Data,” *IEEE Access*, vol. 7, pp. 64722–64736, 2019.
- [7] A. Ntakaris, G. Mirone, J. Kannianen, M. Gabbouj, and A. Iosifidis, “Feature Engineering for Mid-Price Prediction With Deep Learning,” *IEEE Access*, vol. 7, pp. 82390–82412, 2019.
- [8] A. Tsantekidis *et al.*, “Forecasting stock prices from the limit order book using convolutional neural networks,” in *2017 IEEE 19th Conference on Business Informatics (CBI)*, 2017, pp. 7–12.
- [9] J. Niño, G. Hernandez, A. Arévalo, D. Leon, and J. Sandoval, “CNN with Limit Order Book Data for Stock Price Prediction,” in *Advances in intelligent systems and computing, 2194-5365, volume 880, Proceedings of the Future Technologies Conference (FTC) 2018. Volume 1*, K. Arai, R. Bhatia, and S. Kapoor, Eds., Cham, Switzerland: Springer, 2019, pp. 444–457.
- [10] A. Ntakaris, M. Magris, J. Kannianen, M. Gabbouj, and A. Iosifidis, “Benchmark dataset for mid-price forecasting of limit order book data with machine learning methods,” (en), *Journal of Forecasting*, vol. 37, no. 8, pp. 852–866, <https://onlinelibrary.wiley.com/doi/full/10.1002/for.2543>, 2018.

- [11] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [12] A. Vaswani *et al.*, *Attention Is All You Need*. Available: <https://arxiv.org/pdf/1706.03762>.
- [13] H. Song, D. Rajan, J. J. Thiagarajan, and Spanias Andreas, “Attend and Diagnose: Clinical Time Series Analysis using Attention Models,” <https://arxiv.org/pdf/1711.03905.pdf>.