Goethe-Universität Frankfurt am Main
Fachbereich Wirtschaftswissenschaften

Professur für Betriebswirtschaftslehre,
insbesondere e-Finance

Prof. Dr. Peter Gomber

Theodor-W.-Adorno Platz 4
RuW, Postfach 69
D-60323 Frankfurt am Main

Telefon +49-69-798-34683
Telefax +49-69-798-35007
E-Mail gomber@wiwi.uni-frankfurt.de

http://www.efinance.wiwi.uni-frankfurt.de

Master Thesis: Using Generative Deep Learning to Simulate Market Impact in Backtesting

In the realm of high-frequency trading (HFT), strategy backtesting on limit order book (LOB) data requires complex mathematical models and careful consideration of market microstructural dynamics to estimate the effect that, in hindsight, some order would have had on historical data. More precisely, a sophisticated backtesting engine uses a dedicated transaction cost model to account not only for commissions and fees, but also for latency-bound slippage and liquidity-bound market impact. The case of market impact is particularly challenging because some order that never happened would influence the subsequent order stream indefinitely. Consequently, a perfect solution to this problem cannot possibly exist, yet any innovation in approximate models has the potential to better predict live trading performance.

Goodfellow et al. (2014) kick-started the advent of generative deep learning research with the introduction of generative adversarial networks (GAN), a game-theoretical approach to neural network training that teaches a generator network to produce realistic imitations of image data. Since then, a surge of updates to the original design has refined the model to support virtually any form of unstructured data, even high-dimensional medical time-series data (Esteban et al., 2018). In finance, GAN have so far been used to generate stock market order streams (Li et al., 2020), similar to real market data in terms of all price, quantity and inter-arrival time distribution as well as intensity and best bid/ask evolution. Moreover, Lezmi et al. (2020) used the same approach to improve the robustness of trading strategy backtesting, providing a starting point for this thesis.

Using generative deep learning, the aim of this thesis is to design and build a backtesting engine that is able to simulate non-parametrically the market impact of both market and limit orders. Learning from historical market data, the underlying model should be used to internalise the distribution of a universal price formation process as observed by Sirignano & Cont (2019) that, in response to some simulated order, would manipulate subsequent LOB snapshots to reflect the incurred market impact. Put differently, the student is expected to (1) explore design options regarding both data and model engineering, (2) implement a working prototype and (3) analyze the degree to which the simulated price formation process is similar to historical LOB data. Level 2 market data is provided, including high-resolution depth and trade information required to observe the market impact of historical orders, covering the entire DAX 30 universe between 2014 and 2016. Given the technical complexity, the student should have experience with python programming and, ideally, bring to the table a strong interest in applied deep learning research.

Supervisor: Jonas De Paolis

Literature:

- Esteban, C., Hyland, S., & Rätsch, G. (2018). *Real-valued (Medical) Time Series Generation with Recurrent Conditional GANs*. arXiv:1706.02633 [stat.ML].
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). *Generative Adversarial Networks*. arXiv:1406.2661 [stat.ML].
- Lezmi, E., Roche, J., Roncalli, T., & Xu, J. (2020). *Improving the Robustness of Trading Strategy Backtesting with Boltzmann Machines and Generative Adversarial Networks*. arXiv:2007.04838 [cs.LG].
- Li, J., Wang, X., Lin, Y., Sinha, A., & Wellman, M. (2020). *Generating Realistic Stock Market Order Streams*. arXiv:2006.04212 [q-fin.ST].
- Sirignano, J., & Cont, R. (2019). *Universal Features of Price Formation in Financial Markets: Perspectives From Deep Learning*. In: Quantitative Finance, Vol. 19, No. 9, pp. 1449–1459.