



Deliverable D3.4

MSCA-ITN Training for Big Data in Financial Research and Risk Management “BigDataFinance”

Grant Agreement: 675044

This is Deliverable D3.4 of the Work Package 3 (WP3) in “Training for Big Data in Financial Research and Risk Management” (BigDataFinance) Innovative Training Network Marie Skłodowska-Curie project 2015-2019.

Name of the deliverable: “A report of statistical analyses on the order book dynamics with ultra-high frequency data”

Description

The modeling of the limit order book (LOB) is certainly one of the major challenges for the contemporary econometrics. Electronic exchange platforms has constantly grown in the last decades and the trading ruled by sophisticated and competing algorithms led to a market characterized by high complexity. The data recorded from the limit order book platforms is multidimensional, is on a one-millisecond time stamp precision (ultra-high frequency) and shows stochastics behavior. (i) In the first research mentioned in the bottom, the authors focus on the randomness in certain types of order book events, in particular they seek for evidence that there is a “long term behavior” in the processes, e.g. the noise can be identified of being of a certain type and labelled according. This research shed light that the LOB are not totally chaotic but, although the considerable amount of randomness in the occurrence of its events, theses’ a “logic in the chaos” with certains patterns that the method employed uncovers. (ii) Papers two and three both tackle the problem of LOB predictability. In particular, the problem is about the prediction of the mid-price. The goal is clearly relevant for practitioners as well as for academics. As a further contribution, in the second paper we create and openly share a new benchmark dataset with LOB data extracted and processed from a number of securities traded and Nasdaq Nordic, then by applying a number of ML techniques we show that it’s indeed possible to forecast mid-price movement with good precision. (iii) The third paper, which relies on the same data as the second paper, also aims at the prediction the mid-price movements, but the method proposed is a variant end extension (in terms of estimation, parameter tuning and model validation) of the common discriminant analysis. This models is shown to outperform other competing ones in terms of some specific measure associated to the quality of forecasts.

Date, place: June 18th, 2018, Tampere, Finland

Name, position: Martin Magris, Marie Skłodowska-Curie Fellow, Tampere University of Technology



Title: Long-range Auto-correlations in Limit Order Book Markets: Inter- and Cross-event Analysis

Journal: Paper submitted for the IEEE SSCI 2017 conference, available in conference proceedings

Authors: Martin Magris, Jiyeong Kim, Esa Rasanen, Juho Kannianen

Abstract: Long-range correlation in financial time series reflects the complex dynamics of the stock markets driven by algorithms and human decisions. Our analysis exploits ultra-high frequency order book data from NASDAQ Nordic over a period of three years to numerically estimate the power-law scaling exponents using detrended fluctuation analysis (DFA). We address inter-event durations (order to order, trade to trade, cancel to cancel) as well as cross-event durations (time from order submission to its trade or cancel). We find strong evidence of long-range correlation, which is consistent across different stocks and variables. However, given the crossovers in the DFA fluctuation functions, our results indicate that the long-range correlation in inter-event durations becomes stronger over a longer time scale, i.e., when moving from a range of hours to days and further to months. We also observe interesting associations between the scaling exponent and a number of economic variables, in particular, in the inter-trade time series.

Paper available at: <https://arxiv.org/abs/1711.03534>

Title: Tensor Representation in High-Frequency Financial Data for Price Change Prediction

Journal: Paper submitted for the IEEE SSCI 2017 conference, available in conference proceedings

Authors: Dat Thanh Tran, Martin Magris, Juho Kannianen, Moncef Gabbouj, Alexandros Iosifidis

Abstract: Nowadays, with the availability of massive amount of trade data collected, the dynamics of the financial markets pose both a challenge and an opportunity for high frequency traders. In order to take advantage of the rapid, subtle movement of assets in High Frequency Trading (HFT), an automatic algorithm to analyze and detect patterns of price change based on transaction records must be available. The multichannel, time-series representation of financial data naturally suggests tensor-based learning algorithms. In this work, we investigate the effectiveness of two multilinear methods for the mid-price prediction problem against other existing methods. The experiments in a large scale dataset which contains more than 4 million limit orders show that by utilizing tensor representation, multilinear models outperform vector-based approaches and other competing ones.

Paper available at <https://arxiv.org/abs/1709.01268>



Title: Benchmark Dataset for Mid-Price Prediction of Limit Order Book data

Journal: Submitted to Journal of forecasting

Authors: Adamantios Ntakaris, Martin Magris, Juho Kannianen, Moncef Gabbouj, Alexandros Iosifidis

Abstract: Presently, managing prediction of metrics in high frequency financial markets is a challenging task. An efficient way to do it is by monitoring the dynamics of a limit order book and try to identify the information edge. This paper describes a new benchmark dataset of high-frequency limit order markets for mid-price prediction. We make publicly available normalized representations of high frequency data for five stocks extracted from the NASDAQ Nordic stock market. Furthermore, we define an experimental protocol that can be used in order to evaluate the performance of related research methods. Baseline results based on linear and nonlinear regression models are also provided and show the potential that these methods have for mid-price prediction.

Paper available at <https://arxiv.org/abs/1705.03233>