BDA501- Applied Statistics

High Frequency Trading Using Limit Order Book Fast Track Project

Kadir Baver Kerimoğlu

11.01.2022

Table of Contents

Table of Figures

Introduction To Project

Algorithmic Trading

Algorithmic trading (also called automated trading, black-box trading) uses a computer program that was a predefined set of instructions in order to place a trade. The trade can produce profits at high speed and frequency that is impossible for a human trader.

The predefined sets of instructions are related to timing, price, quantity, or any mathematical model apart from profit opportunities for the trader. Thanks to Algorithmic-trading, the efficiency of human emotions over the trade markets is getting lower that is also result higher liquid market with more systematic rules.

Benefits of algorithmic trading

Algorithmic-trading provides the following benefits:

- The best possible prices can be used for trades.
- Trade order placement is instant and accurate
- Trades can succeed at a specific time and simultaneously can avoid significant price changes.
- Transaction costs can be reduced.
- Simultaneous automated checks can be described in multiple market conditions.
- Risk of manual errors is decreased to the minimum level when placing trades.
- Can be back-tested using available historical and real-time data to see if it is a viable trading strategy.
- There is no any human-made mistake because of the emotional and psychological factors in this method.

A Short Explanation About High-Frequency Trades

One of the critical types of algorithmic trading is high-frequency trading which works under a low-latency setting. This low-latency setting refers to a system that operates under a single-digit millisecond round-trip time for a packet of data/orders. These data packets, named high-frequency data, provide records of the intraday trading activity. These records contain information such as; Security's ID code, Timestamp, Order Price, Volume of order, Queues (Ask or Bid), Type of Order (Submission, Cancellation, Execution).

It uses complex algorithms to analyze multiple markets and execute orders based on market conditions via using powerful computer programs. The faster execution speeds means, the more profitable traders are.

Benefits of HFT (High- Frequency Trading)

- HFT is complex algorithmic trading in which large numbers of orders are executed within seconds.
- It adds liquidity to the markets and eliminates small bid-ask spreads.

Disadvantageous of HFT

There are some kind of criticism about HFT. For example, being as a replacement of broker-dealers and using mathematical models and algorithms to make decisions in order to take human decision could be seen as an example for criticism. Because decisions happen in milliseconds, and this could result in big market moves without reason.

According to Security and Exchange Commission, HFT is defined by huge numbers of order, proprietary trading and limited holding duration. So, HFT shows kind of challenges which were written as below:

- High speed while giving and cancelling orders
- High level of difficulty about prediction of volatile markets
- Producing the profits from low amount of margins.

Reviewing of The Related Articles

1. Stock Market prediction on High frequency data using Long-Short Term Memory

In this paper there's a model that consists of Long Short Term Memory algorithmic program using High Frequency historical data. The main goal of this model to predict the longer term closing numbers of stocks associated with ten minutes, five minutes and with excellent performance one-minute ahead without the use of Technical Indicators.

I will very shortly introduce in one paragraph what was the steps of research and what are the long term steps of developments of this analysis as well.

After the pre-processing duration (data collection, cleaning, organizing etc.) LSTM model design as built and sub-networks of its consistence were built then whole system began to work according to the equation that defines the relativeness between variables of system.

• Memory cell : stores state

• Front door : controls what to to be told

• Forget door : controls what to forget

• Exit door : controls the quantity of content to change

LSTM unit will attempt to store or not the memory which might be accomplished by doors.

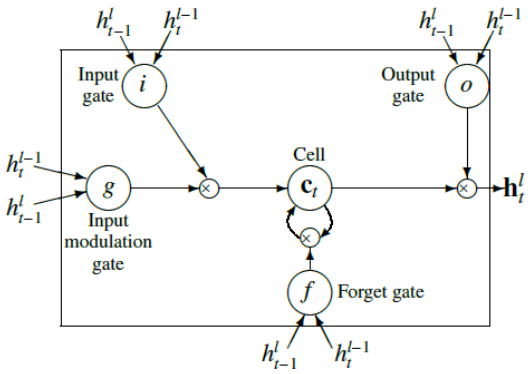


Figure 1 A pair of graphical illustration of LSTM memory cells

In the experimental analysis part author tried to clarify what is the gap of results of every model's prediction to the truth once they became nearer to x-minutes ahead. Then they outlined technical indicators to those results in order to compare the efficiency of every model.

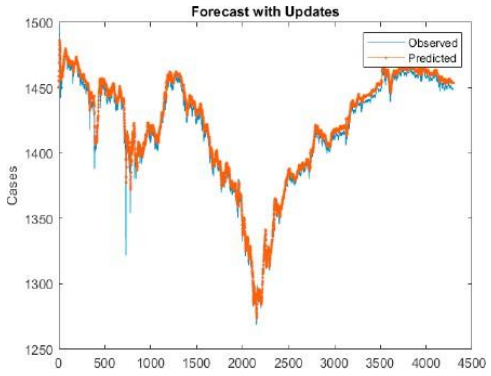


Figure 2 Comparison between Forecast and Observed values

2. Mid-Price Movement Prediction in Limit Order Books Using Feature Engineering and Machine Learning

This thesis was written within the years between 2016-2019 by Adamantios Ntakaris from Tampere University to focus in List of Books dynamics by utilizing machine and deep learning methods.

From thesis especially the parts which is related with High Frequency Data properties and modelling limit order book dynamics were examined in detail.

According to the thesis there are various different properties which can be used in different analysis such as Negative Autocorrelation, Absence of Normality or Log normality with Heavy Tails and Bid-ask bounce.

a. Negative Autocorrelation:

The main idea of negative autocorrelation is that after a short time from begin stock prices are tend to change a bit in negative direction. This concept was proven by some different authors as well. Besides there is an example chart is below which was taken from amazon trading on 9/22/2015. In chart, Correlogram shows the existence of negative first-order autocorrelation with tight confidence intervals (blue lines).

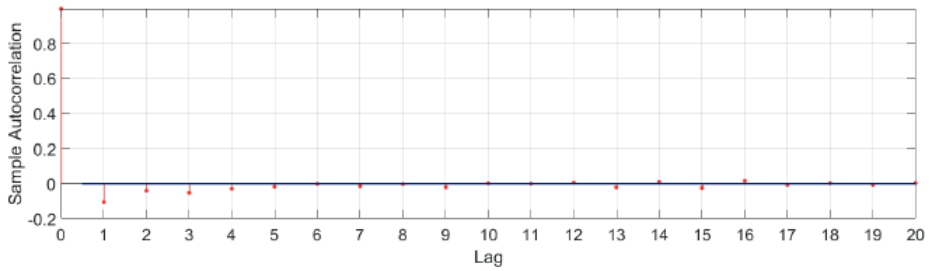


Figure 3 Correlogram up to 20 lags based on Amazon returns for the trading session on 9/22/15.

b. Absence of Normality or Lognormality with Heavy Tails

According to the thesis, stock prices are tend to be a lognormal distribution if ordered value distributed in normally. However, this happens only below the situation with high-frequency data in long duration.

c. Bid- Ask Bounce

Bid-ask spread is one of the most important actors in price forecasting due to the its effect to the liquidity of markets. Higher bid-ask spread means, lower liquidity for market. Bid-ask bounce (e.g., Figure 2.4) is the bouncing of any trade between the bid and ask price. This happens when there are trades on both the bid and asking price, but no real movement in price.

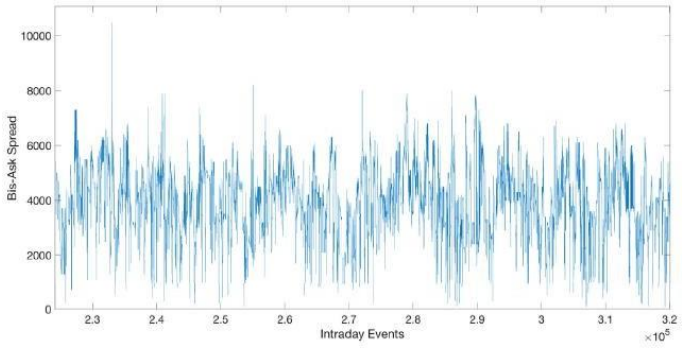


Figure 4 Amazon's intraday (9/22/15) bid-ask spread fluctuations

Details of The Limit Order Book

Modern financial markets operate under a double auction mechanism called limit order book (LOB). LOB accepts two types of orders, limit and market orders.

A limit order is associate order to try and made trade a stock at a value that was determined a previous time past, whereas a market order is traded instantly under the current market price. As a result, LOB acts as a recording mechanism for the unexecuted purchase or sell activity where time and price priority are the primary filters for queues occurrence.

These queues represent, at any time t, the price levels of the bid and ask sides and very according to order arrivals and liquidity levels.

LOB Rules:

Orders, following are outlined as x = (pi, vi, tt) wherever pi , vi , and ti represent the worth (i.e., bid or ask), the quantity and the order submission time at time i, respectively. The set of all active orders, L(t), could be a càdlàg method since for each new limit order x, the subsequent holds:

x ∈ L(ti) and x /∈ limt ↑tx L(t).

This activity defines the depth size for each LOB level in each bid and ask sides. The depth size, particularly within the best bid and ask level, could be a key component for the price formation. The depth of every level in trade stays perpetually because of the limit orders, market orders and cancellations. Limit orders result the scale depth in opposite way if it compares to the market orders and cancellations that take away liquidity from LOB.

When there's a price decrease for the i t h event, (Ri)i≥1 (stationary process) turns into (˜Ri)i≥1. In the case of reduced-form model, for a new order or cancellation arrival, size *Vi* b at time instance ti^b there are two scenarios:

- if $v^b_{i-} + V^b_i \geq 0$ then there is no price change and the order will be satisfied
- if $v^b_{i-} + V^b_i < 0$ then the size of the bid level is reduced together with the price by one 'tick' of size π. Based on the updated $\tilde{R}_i = (\tilde{R}^b_i, \tilde{R}^a_i)$ values for the bid

$$(p^b_i, v^b_i, v^a_i) = (p^b_{i-}, v^b_{i-} + V^b_i, v^a_{i-})\mathbb{1}_{\{v^b_{i-} \geq -V^b_i\}} + (p^b_{i-} - \pi, \tilde{R}^b_i, \tilde{R}^a_i)\mathbb{1}_{\{v^b_{i-} < -V^b_i\}} \quad (2.4)$$

Where 1 is the indicator function, pbi− is the best bid price before the update,vbi− and vai− are the volume sizes for the best bid and ask sides, severally.

In a similar situation, for a new arrival to the ask side, of size V ai , LOB's state will be:

- if $v^a_{i-} + V^a_i \geq 0$ then there is no price change and the order will be satisfied
- if $v^a_{i-} + V^a_i < 0$ then the size of the ask level is reduced together and the price will be increased by one 'tick' of size π. Based on the updated $\tilde{R}_i = (\tilde{R}^b_i, \tilde{R}^a_i)$ values for the bid and ask side, the new LOB state is:

$$(p_i^a, v_i^b, v_i^a) = (p_{i-}^b, v_{i-}^b, v_{i-}^a + V_i^a)\mathbb{1}_{\{v_{i-}^a \geq -V_i^a\}} + (p_{i-}^b + \pi, \tilde{R}_i^b, \tilde{R}_i^a)\mathbb{1}_{\{v_{i-}^a < -V_i^a\}} \quad (2.5)$$

where is that the indicator function, pbi− is that the best bid price before the update,vbi− and vai− are the quantity sizes for the best bid and ask sides before the update, respectively.

HFT Study With “Noname” Stocks

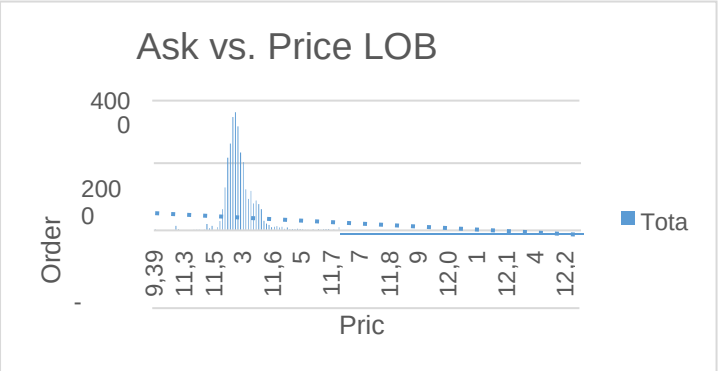
According to the data of LOB of stocks which was named as “noname”, the was a little analysis was done.

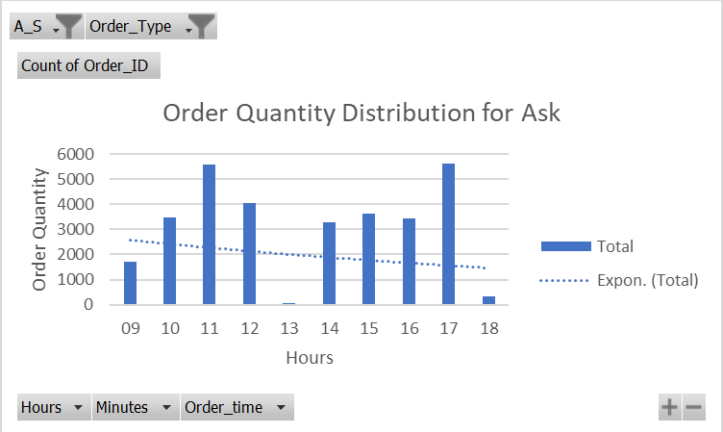
At below graphs it can be seen clearly Bid and Ask orders according to the their quantitys related to different price distributions.

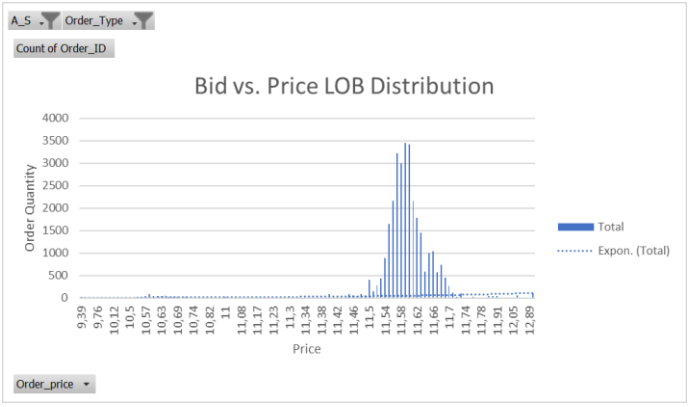
From those graphs it is very obvious that not only buy but also buy orders seem to be concentrated mostly between 11.5 and 11.70.

Looking at the charts with the highest volume of buy and sell orders, two peaked densities are observed in both.(11 am & 17 pm).

Therefore, from these 4 simple graphical works, the best investment commentary I could make for myself that I would bought at 9.00 am with the price of ~11.40 and sell it at 11.30 am with the price 11.60 and then buy it again at 16.00 pm with the price ~11.40 and sell it at 17.30 pm 11.60 for sure.







References

Chen, J. (2021, August 25). High-Frequency Trading (HFT) Definition. Retrieved January 11, 2022, from https://www.investopedia.com/terms/h/high-frequency-trading.asp

C. Lattemann, P. Loos, J. Gomolka, H.-P. Burghof, A. Breuer, P. Gomber, M. Krogmann, J. Nagel, R. Riess, R. Riordan, et al., High Frequency Trading, Business & Information Systems Engineering 04 (2012) 93–108

A. J. Menkveld, High frequency trading and the new market makers, Journal of financial Markets Vol. 16 (2013) 712–740

Adamantios Ntakaris, Oct 25, 2019, Mid-Price Movement Prediction in Limit Order Books Using Feature Engineering and Machine Learning

Lanbouri, Z., & Achchab, S. (2020, August 06). Stock Market prediction on High frequency data using Long-Short Term Memory. Retrieved January 11, 2022, from https://www.sciencedirect.com/science/article/pii/S1877050920317877