Project Proposal

Team members: Student A

Project title: Modeling limit order book price dynamics

Description of the problem context: A limit order book (LOB) is a pricing mechanism for a stock/commodity that is used in most of the exchanges worldwide. Every participant places their buy or sell orders with specific prices and quantities. If a match is found then the transaction occurs and the order book is updated. Records of high-frequency trading activity are organized into a database with two components: the message book and the order book. The message book contains information about each trading event, including the time and the type of transaction (bid or ask). The order book keeps unexecuted limit orders for bid and ask. The mid-price is defined as the mean of the best ask price and the best bid price at time t. The mid-price movements can indicate a potential profit.

The objective is to capture the dynamics in high-frequency limit order book data and forecast movements of the mid-price. Real data from NASDAQ will be used to build the model. A LOB contains a massive amount of rapidly evolving data, with the possibility of important patterns forming and disappearing within splits of seconds making their observations impossible to humans. Therefore machine learning can be used.

Type of the problem: Supervised learning. Training data can be labeled by class in advance and a model can be used to assign new data to those classes. Methods including logistic regression and multi-class support vector machines can be investigated.

Initial thoughts on techniques that might be used: Preprocess the LOB data to make it suitable for the model. Construct and validate the learning model using performance measures. The objective is to construct a learning model that can forecast the mid-price (possibly in real time).

Hyperlink of the problem description: https://lobster.wiwi.hu-berlin.de/index.php