**CS230 Project Midterm Report:   
Crypto Exchange Price Prediction using Limit Order Book**

|  |  |  |
| --- | --- | --- |
| **Ben Gilboa**  **(SUID# - 06278930)** | **Tamal Biswas**  **(SUID# — 05107984)** | **Ashwin Selka Padmanabhan (SUID# — 06246676)** |
| **Stanford University**  **Winter 2018** | | |

**Abstract**

*This midterm report reviews the problem that the project is intended to investigate and solve. It explains the dataset format and acquisition process along with the initial model that we started to build and evaluate. The last section describes the remaining activities, methods and tasks that we plan to do for completing the project*

# Introduction

High frequency trading or Algo trading is gaining significant momentum in stock exchanges. In today’s market sizable portion of the daily traded volume is done by specialized companies using those techniques. In the elaborated stock market it is almost impossible for individuals not using heavy machinery and very fast access to data to gain any advantage as margins and arbitrages are closed in fraction of a second.

The rise of the crypto market and exchanges might reveal opportunities that are long gone in the stock market for small scale algorithmic trading.

In this project we explore and develop deep machine learning model that predict the future price of digital asset such as bitcoin. We intend to build a machine learning RNN (Recurrent Neural Network) that predict the future price of a tradable and volatile digital asset such as the Bitcoin. As input to the model we will use limit order book data along with other historic indications for demand as supply to develop our predictor. Although we chose a digital asset for this project, the principals and methods we develop are transferable to any asset that is tradable in an exchange.

# Prior work

TBD

# Dataset Characteristics and Acquisition

The data that is primarily used in for our predictor is the data from limit order book.

## Limit order book

This is a ledger maintained by the exchange of all limit orders that are pending. The order book has a sorted list of all bid and ask orders with the quantity and associated price for each order. It is usually presented graphically as accumulative plot of all bid and ask orders.

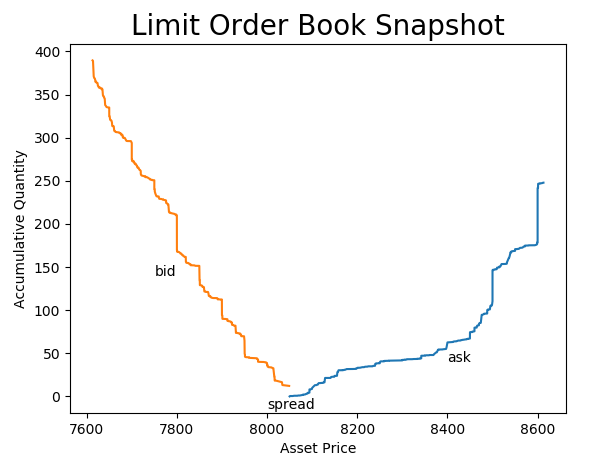
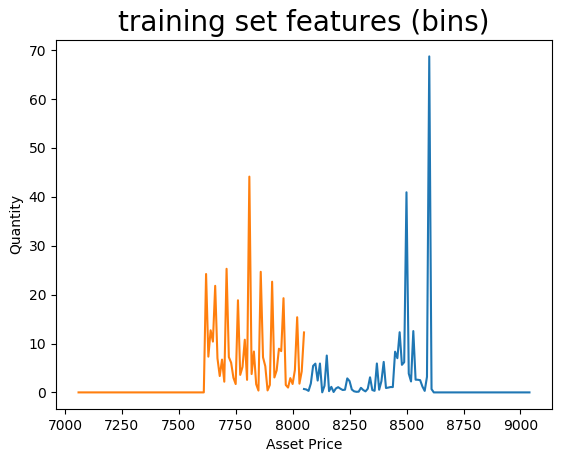


Figure : Limit Order Book Snapshot

The limit order book snapshot represent the demand and supply in the market in a certain point in time. In the above figure, it is clearly seen that the demand is “stronger”. There are much more buyers who are wiling to buy the asset for a price that is lower by 3% from last price than sellers who are willing to sell in a price that is higher by 3% than the last price. This might indicate that the price is about to increase. Our dataset is a sequence of the limit order book along with the last price.

Training examples are snapshots of the order book taken in 1 min intervals. We get 500 highest bid orders and 500 lowest ask orders from the exchange that we use. Since every order has 2 parameters (quantity and price) we cant use it as is. We apply a small modification to the data to extract a training example. We define “bins” of 10$ and we sum the quantities that relate to each bin. From 500 bid orders we create 100 bins that represent the last price down to last price minus 1000$. We can see a plot of one training example (that correlates to figure 1).

Figure : sample of one training example after structured in bins

For the labels we have the last Bitcoin price that corresponds to every training example. We make it a classification problem by comparing the next value of the bitcoin (1 min into the future) to the current price. If the price increased the label is ‘1’ and if decreased or same it is ‘0’