

Use dynamic Bayesian networks to analyze limit order books and commodity markets

The goal of this project is to capture information from limit order books to predict market direction, which can enable repricing of orders and more efficient market making. Such an approach allows the market maker to provide liquidity while making profits at the same time. To know the markets, we need to consider regime switching and hidden orders. We wish to develop online modeling and reach high performance computing. I will compare the result of my method with the results from different machine learning models like random forest classifier, extra trees classifier, Ada boost, gradient boosting and SVM....

In order to efficiently predict directional price movements, we need to develop features that permit reliable predictions from certain states.

- Bid-Ask Spread: A positive value indicating the current difference between the bid and ask prices in the current order books.
- Mid Price: The average of the bid and ask prices is weighted by their inverse volume.
- Trade Probability: Measuring the probability that buyers or sellers crossed the spread in recent executions.
- Bid-Ask Volume Imbalance: A signed quantity indicating the number of shares at the bid minus the number of shares at the ask in the current order books.
- Signed Transaction Volume: A signed quantity indicating the number of shares bought in the last 20 seconds minus the number of shares sold in the last 20 seconds.