

The background of the slide is a light gray image of bare, intricate tree branches, creating a complex, web-like pattern.

Random Forest Trading Strategy with Preprocessed Order Book Data

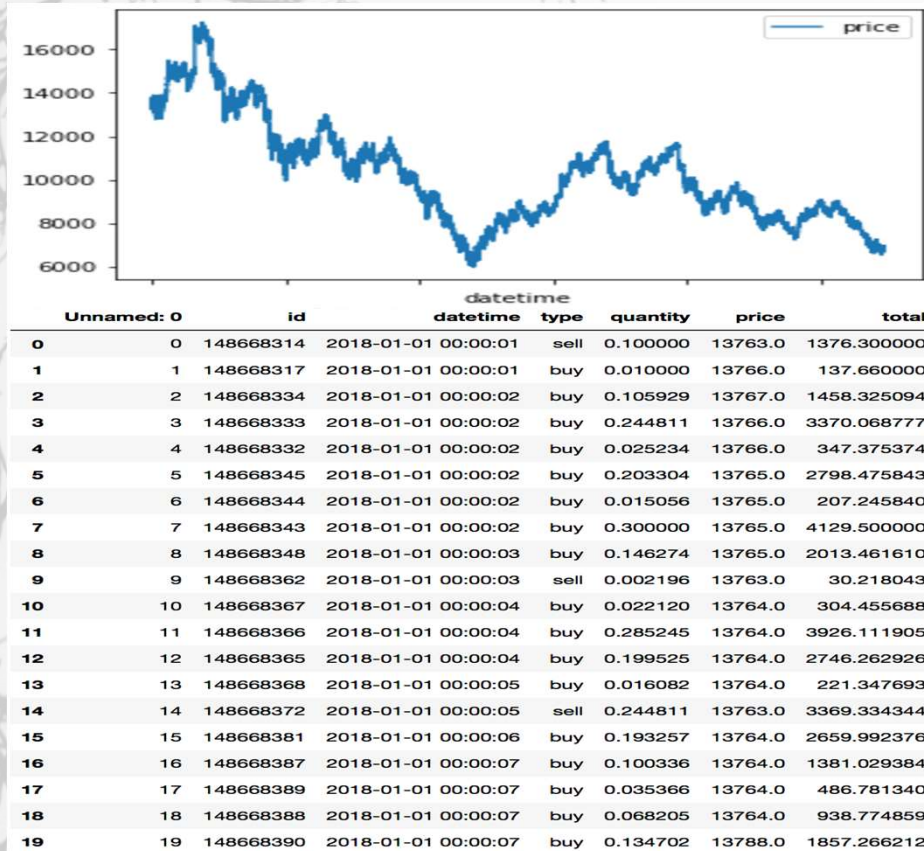
R. Ding, D. Jiao, X. Li

M.S. Financial Engineering, Columbia University

Sep. 30, 2018

I. Data Exploration

What we got as raw data



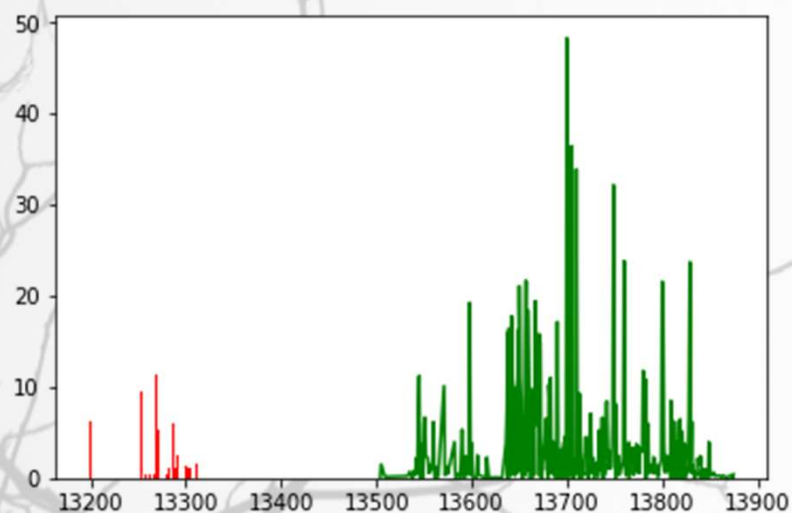
VS

What we want as order book



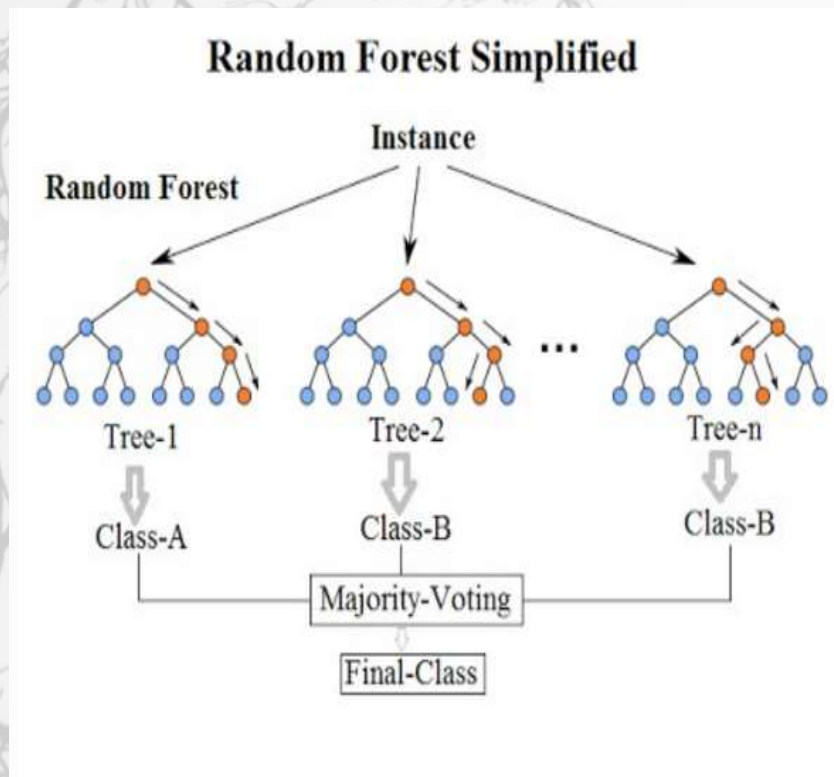
II. Order Book Construction

- Object-Oriented Programming
- Major variables: self.book, self.prices, self.best_bid, self.best_ask
- Process the order: price, volume, direction



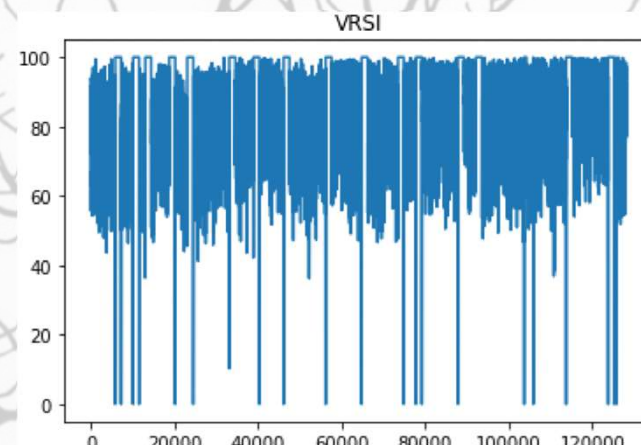
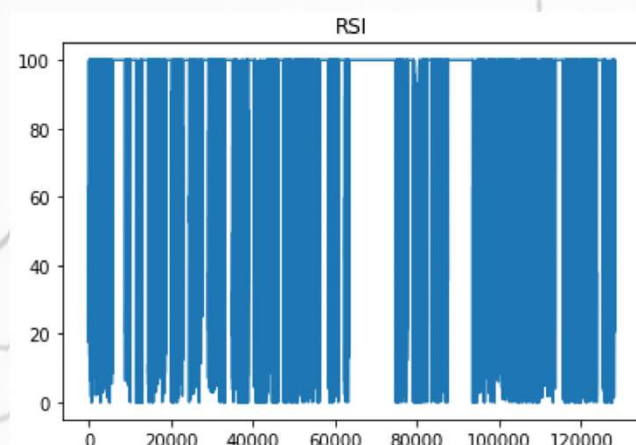
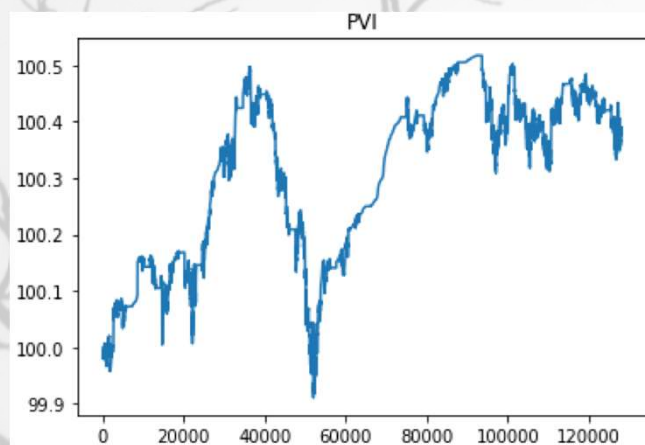
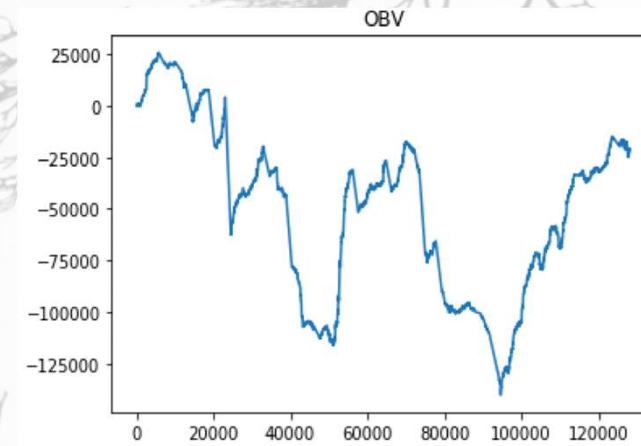
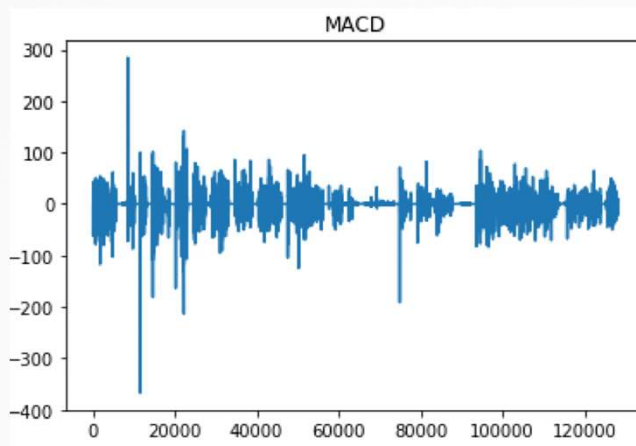
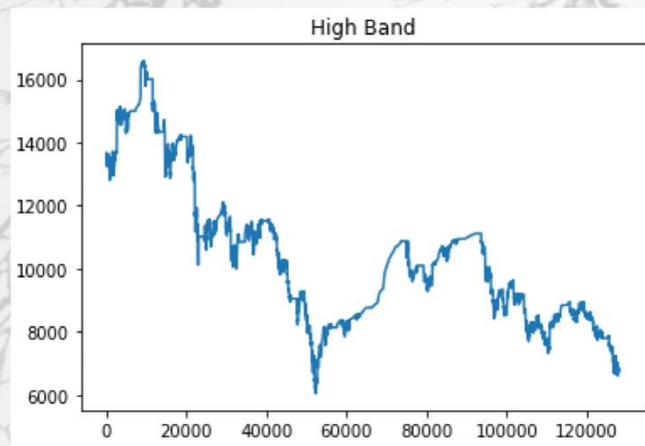
	open	close	high	low	vol
2018/1/1 0:00	13763	13765	13788	13763	1.89
2018/1/1 0:01	13765	13746	13765	13741	5.71
2018/1/1 0:02	13746	13741	13750	13741	3.74
2018/1/1 0:03	13741	13735	13741	13735	3.80
2018/1/1 0:04	13735	13734	13735	13734	7.09
2018/1/1 0:05	13734	13564	13734	13559	30.76
2018/1/1 0:06	13559	13570	13570	13559	6.22
2018/1/1 0:07	13570	13581	13581	13570	13.68
2018/1/1 0:08	13581	13615	13615	13581	11.86
2018/1/1 0:09	13615	13634	13636	13607	25.70

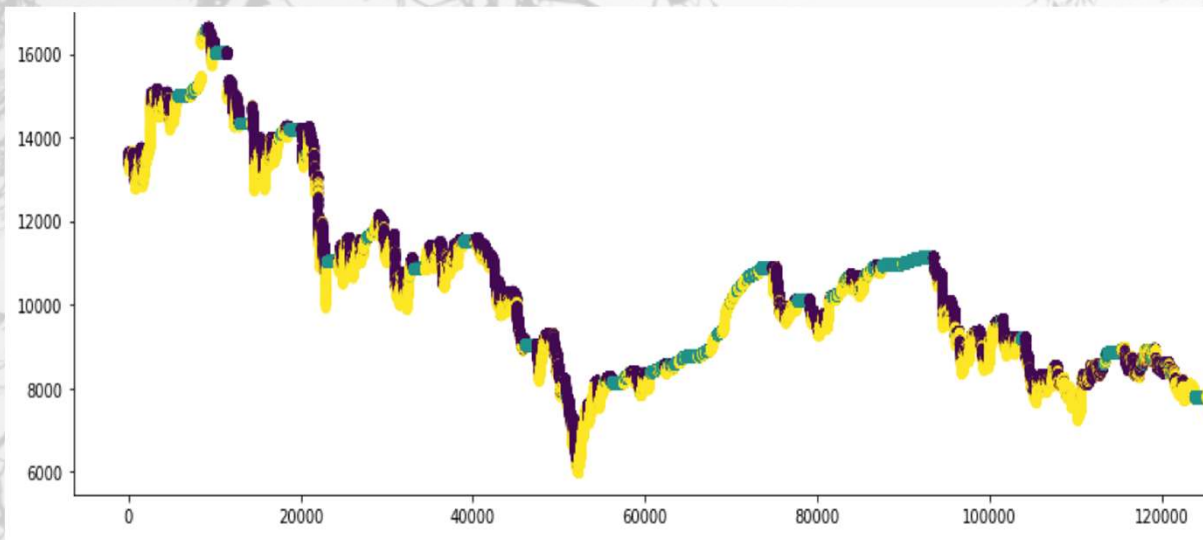
III. Random Forest Prediction Model Based Trading Strategy



- We have the minute candlestick data from the order book data processing phase
- We can construct various technical indicators based on volume, high, low, open, and close
- Using a random forest classifier to label the target, which can be thought of as some expected return in a future period
- Choose features from technical indicators based on their prediction accuracy performance
- Reach a test accuracy of 90% using several indicators involving both the price and volume aspects

Indicators in Our Strategy





```
'2018-01-03 19:36:01 end position Return: 0.155%',  
'2018-01-03 19:38:01 enter long',  
'2018-01-03 19:39:01 end long and short Return: 0.149%',  
'2018-01-03 19:46:01 end short and long Return: 0.216%',  
'2018-01-03 22:06:01 end position Return: 0.635%',  
'2018-01-03 22:08:01 enter long',  
'2018-01-03 22:10:01 end position Return: 0.0%',  
'2018-01-03 22:11:01 enter long',  
'2018-01-03 22:12:01 end position Return: 0.0%',  
'2018-01-03 22:14:01 enter long',  
'2018-01-03 22:15:01 end position Return: 0.007%',  
'2018-01-03 22:17:01 enter long',  
'2018-01-04 02:20:01 end position Return: 1.089%',  
'2018-01-04 02:50:01 enter long',
```

Trading Results based on the prediction model:

- Use the predicted labels to generate buy/sell (long/short) signals.
- The plot on top shows all predicted labels/signals and the plot on the bottom shows a small part of the trading records.