FBA | QUANTITATIVE FINANCE RESEARCH GROUP

Financial Machine Learning

Homework 11

Due at 07:00 pm (Korea Standard Time) on Sunday, December 4.

Submit one file: written solutions with executable Python code

Problem 1. Text book: Hands-on Machine Learning. Submit .ipynb file.

(a) Practice all the codes in the Text book Chapter 15. And show that they work well.

Problem 2. New York Stock Exchange

Context

This dataset is a playground for fundamental and technical analysis. It is said that 30% of traffic on stocks is already generated by machines, can trading be fully automated? If not, there is still a lot to learn from historical data.

Here is couple of things one could try out with this data. Choose one subject you want to do.

(1) One day ahead prediction: Rolling Linear Regression, ARIMA, Neural Networks, LSTM

Acknowledgements.

Stock price prediction using LSTM RNN and CNN-sliding window model

A_LSTM-based method for stock returns prediction

Stock markets price movement prediction with LSTM neural networks

(2) Security clustering, portfolio construction/hedging

Acknowledgements.

RNN-DBSCAN A Density Based Clustering Algorithm

Stock2Vec A Hybrid Deep Learning Framework

Deep Hedging

Content

Dataset consists of following files:

prices.csv: raw, as-is daily prices. Most of data spans from 2010 to the end 2016, for companies new on stock market date range is shorter. There have been approx. 140 stock splits in that time, this set doesn't account for that.

prices-split-adjusted.csv: same as prices, but there have been added adjustments for splits.

securities.csv: general description of each company with division on sectors

fundamentals.csv: metrics extracted from annual SEC 10K fillings (2012-2016), should be enough to derive most of popular fundamental indicators.