Proposal: Investment and Trading Capstone Project
Build a Stock Price Indicator

March 7, 2017

Soyoung Park

1. Domain background

There are many fundamental indicators that stock analysts use to predict whether a certain company's stock price will go up or down regardless of the stock market trend.[1] In this project, a stock price predictor will be built in which training set is Intel Corporation's data pair of "fundamental indicators"-"Momentum of daily trading data" over a certain date range, and outputs will be predicted momentum of how much price would go up or not for the given query fundamental indicators. A simple Python script will be mainly used, and suggestions will be made on whether/what trades(Buy/Sell/Nothing) is worth to be made.

2. Problem statement

Whenever a company made a quarterly earning announcement, the stock holders wonder whether to hold the stock or a certain transaction is to be made. There are many sources of forecast on how stock price would behave in the near future, but it is not clear how those estimate is drawn and how reliable it is. Machine learning/Deep learning can be very good alternative solutions to convince an investor to take a historical dataset and understand how much one company's public announcement makes influence in stock price. Also, it would be very critical to understand for a company which fundamental indicators drive the stock price most sensitively.

3. Datasets and inputs

Various US fundamental indicators will be pulled from Quandl using Python API [2], and the followings are the query codes to be explored:

Quarterly data:

SF1/INTC_CAPEX_MRQ
SF1/INTC_ASSETS_MRQ
SF1/INTC_REVENUE_MRQ
SF1/INTC_EQUITY_MRQ
SF1/INTC_GP_MRQ
SF1/INTC_NCFO_MRQ
SF1/INTC_DPS_MRQ
SF1/INTC_EPS_MRQ

SF1/INTC_EBIT_MRQ

Annual data:

SF0/INTC_DE_MRY
SF0/INTC_RND_MRY
SF0/INTC_TANGIBLES_MRY
SF0/INTC_SHARESWA_MRY
SF0/INTC_TAXEXP_MRY
SF0/INTC_PB_MRY
SF0/INTC_NETINC_MRY
SF0/INTC_CASHNEQUSD_MRY
SF0/INTC_DEBTUSD_MRY
SF0/INTC_COR_MRY
SF0/INTC_CURRENTRATIO_MRY
SF0/INTC_DEBT_MRY
SF0/INTC_DILUTIONRATIO_MRY
SF0/INTC_LIABILITIESNC_MRY
SF0/INTC_BVPS_MRY
SF0/INTC_DEPAMOR_MRY
SF0/INTC_FCFPS_MRY
SF0/INTC_EBITDA_MRY
SF0/INTC_EPSDIL_MRY
SF0/INTC_NCFDEBT_MRY
SF0/INTC_NCFCOMMON_MRY

Their abbreviation and definition of each indicators will be discussed in detail in the main project document. The daily price data will be also pulled from quandl database with query code “WIKI/INTC”. Additional indicators can be attempted to be created with the pulled dataset according to the approach discussed in [4].

4. Solution statement

For training set, a pair of multi-dimensional fundamental indicators-momentum of adjusted close price for 30 days will be used to build a historical model. For test set, a list of fundamental indicators will be given to predict the momentum.

5. Benchmark model

SVM, random forest, and adaboost will be used to predict the momentum of adjusted close price for 30days. Also, feature importance will be studied for the cases that are trained with random forest and adaboost [4][5]. Once key features are determined and selected, the dataset will be re-trained to improve accuracy.[6]

6. **Evaluation metrics**

RMSE of momentum will be used as evaluation metrics of each model.

7. **Project design**

A training interface will accept the pair of fundamental indicators-price data(momentum of adjusted close price for 30 days) of range (start_date= Q1 1980, end_date=Q4 2016 every quarter) for only Intel (ticker symbol INTC) and a model of stock behavior will be built based on several classifiers including SVM, random forest, and adaboost. A query interface will accept a list of of fundamental indicators from the testset, and outputs the momentum value indicating how much stock prices will change.

A basic run of the core system would involve one call to the training interface, and one or more calls to the query interface. At the end, a suggestion will be made on whether/what trades(Buy/Sell/Nothing) is worth to be made on a certain earning announcement/set of fundamental indicators.

8. **References**

[1]http://money.cnn.com/quote/forecast/forecast.html?symb=INTC
[2]https://www.quandl.com/data/SF1-Core-US-Fundamentals-Data?keyword=INTC
[3]Definition of indicators:

Quandl: https://www.quandl.com/data/SF1-Core-US-Fundamentals-Data?keyword=INTC
Quantopian : https://www.quantopian.com/help/fundamentals#valuation
Investopia:http://www.investopedia.com/articles/active-trading/041814/four-most-commonlyused-indicators-trend-trading.asp

[4]https://www.quantopian.com/posts/machine-learning-on-quantopian

[5]http://cs229.stanford.edu/proj2013/Lin-Feature%20Investigation%20for%20Stock%20Market%20Prediction.pdf

[6]http://cs229.stanford.edu/proj2013/DaiZhang-MachineLearningInStockPriceTrendForecasting.pdf