Project Proposal
Forecasting Stock Indexes Trends

ECEN 765 - Machine Learning with Networks

October 11, 2017

Skipper af Rosenborg **UIN** 526006676

ELECTRICAL & COMPUTER ENGINEERING
TEXAS A&M UNIVERSITY

Contents

1 Introduction and Motivation

Prediction of stock trends is a difficult task as it involves a lot of unknown variables. There are generally two approaches to predicting stock trends, fundamental analysis and technical analysis. Fundamental analysis predicts the stock trend by looking on a companies indices such as profit, debts, cash flow, suppliers, market demand and more. Technical analysis analyse historical prices of a stock in order to predict future trend.

The value in accurately predict the trend of a stock index is priceless for investors, as they based on this can time when to sell and buy their stocks, and thereby optimise their profits.

Previous research has investigated whether Artificial Neural Networks (ANN), Support Vector Machine (SVM), Random forest or Naïve-Bayes is the best machine learning algorithm to predict index trends, in this case binary, hence either upwards or downwards. The research is based on using technical metrics calculated from the stocks closing, low and high price[1]. Previous studies has also studied whether it is better or worse to do the prediction based on the raw data or filtered data. The studies shows that filtered data is a better predictor[2]. Previous research has also investigated which technical measures and in which combination are the most important for predicting stock prices[3].

My project proposal is to perform a similar comparison of machine learning algorithms but with noise filtered data. I will furthermore see if I can derive new technical measures based on the dominating frequencies and their phase observed from a Fourier transformation. I have not yet found any articles that try to derive technical measures based on this, I do however have knowledge about a technical stock analyst that supposedly uses these to predict stock trends. From personal experience I have found this to be a good predictor.

2 Deliverable

I will confirm that noise filtered data is better to use for predicting stock trends for ANN, SVM, Random forest and Naive Bayes. I will also compare the algorithms based on packages in scikit-learn to find the best one. Furthermore I will investigate if a new technical measure can be successfully establish based on the dominating frequencies and phase from a Fourier transformation.

3 Methodology

I will develop the project in four phases.

1. I will do a more detailed literature study to further investigate what is best practice.
2. I will develop a filtering algorithm that can de-noise the data.
3. I will develop algorithms using scikit-learn as a foundation to compare filtered and unfiltered data, as well as, the four different algorithms.
4. I will try to develop a new metric and incorporate it in the previously developed algorithms.

4 Resources

The historical price for stocks and indexes for my project is public available from Yahoo Finance. I will furthermore need to study Fourier transformation and the theory and

implementation of the four mentioned algorithms. I will use the books in the course as well as online resoursces for this.

5 Milestones

- October 6th - Finish literature study
- October 20th - Finish data smoothing
- November 3rd - Finish machine learning algorithm
- November 17th - Finish development of new metrics
- November 24th - Finish testing
- December 5th - Class presentation
- December 15th - Hand in report

References

[1] Patel J, Shah S, Thakkar P, Kotecha K. *Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques* Expert Systems with Applications, 2015

[2] AHuang H, Zhang W, Deng G, Chen J. *Predicting Stock Trend Using Fourier Transform And Support Vector Regression.* IEEE 17th International Conference on Computational Science and Engineering, 2014.

[3] Oriani F, Coelho G. *Evaluating the impact of technical indicators on stock forecasting.* IEEE Symposium Series on Computational Intelligence, 2016