Machine Learning Capstone Proposal-Udacity

Hitesh Gulati

hg0111@gmail.com

Project Overview.

Hedge funds are amongst the most successful business for a long time. Stock trading is one of the core activities they employ to generate high returns while taking in minimum amount of risk. This project is aimed to develop an automated stock trader for given risk.

Problem Statement

Stock trading have been around us for long time even before the technology age. Traders initially used pen and papers to draw patterns and identify unique ones as signals to buy a stock, hold it or sell it off. There have been many studies on indicators of stock trading and many indicators also co-exists like candlesticks, moving average, average directional index and so on. These indicators usually take in historical data for stocks (Open, High, Low, Close, Volume and Adjusted Close) and outputs their signal based on what it is looking for like a candlestick would output a signal Buy, Hold or Sell.

These indicators may have been powerful earlier but with advent of time and their overuse are no longer capable enough to predict a buy or sell single handed. The purpose of this project will be to apply these various indicators (or agents) and decide whether to buy, sell or hold on a trade.

Due to limited availability of data only technical indicators are included in this project. However, this project can be expanded to include fundamentals data too.

Datasets and Inputs.

Trading data from Nasdaq of retail industry from year 2006 to 2013(daily prices) have been used as input in this project. This data is then fed into various agents and a signal is generated for every trading day (each record).

Agent	Signal
Hodrick Prescot Trend Agent	Uptrend, Downtrend, Notrend
Moving Average Crossover Agent	Buy, Sell, Hold
Candlestick Agent	Buy, Sell, Hold
Stochastic Agent	Buy, Sell, Hold
Volume Agent	Strong Volume, Weak Volume
Average Directional Index Agent	Strong Trend, Weak Trend
Return to Risk Ratio Agent	Bad, Ok, Good, Excellent

These signals are considered as input to the learner.

Solution Statement

The solution to build an effective model is to combine all these agents and provide it as an input to the learner. The learner will provide an output based on the signals from various agents.

Benchmark Model

Outputs can be obtained by looking ahead seven trading days that what would have been the correct decision to take given the stock traded in this manner.

For eg: If the stock had very less (less than pre-defined value) return to risk ratio in next seven days, it would have been better to hold the stock (or not to be involved in the trade). If the ratio is good enough, we should have traded the stock. If in next seven days stock gave positive return(more than transaction cost=0 for convenience) we should have bought (Buy) the stock, and if it gave negative return we should have sold (Sell) it. These signals Hold, Buy and Sell are considered as outputs for the learner.

Evaluation Metrics

The learner will be evaluated based on it effectiveness of taking the right decision using various models as compared to single model. In other is it better to use various agents to predict the output than is single agent. For Eg: If Stochastic Agent predicts Buy, and out model predicts Hold, than on an average which performs better.

Project Design

Once we have a set of inputs and outputs the data will be split in roughly 80-20 ratio for training and testing. The split will be chronological i.e. training data will be from earlier period and testing data will be from later period, rather than random to avoid forward looking bias