

Predicting Stock Prices

Dmytro Ryzhkov, Chandler Baggett, Jacob Toomey

Project Description

Is it possible to find patterns in previous stock prices and the history of the stock market in order to predict it's future performance?

People say the stock market is volatile and unpredictable, but high frequency trading firms hire analysts all the time in order to try and predict the stock market.

Can we utilize data mining and machine learning to recognize certain patterns in certain stocks, that can predict how a stock might behave next?

Prior Work

- Genetic Algorithms have been used to try and predict the stock market.
- Banks and High Frequency trading firms hire analysts to predict the stock market, and have talented developers developing algorithms and utilizing machine learning to try and predict the stock market.
- There are many papers written on the subject and a lot of academic research going into utilizing machine learning in order to predict the stock market.

List of Datasets

- NASDAQ provides market data for up to the last ten years
- Yahoo Finance provides market data for the entire history of a certain stock, this is most likely what we will use.

Proposed Work

- Gather data for certain stock or multiple stocks
- Preprocess data and get it into a format that will be used in our work
- Once the data is in an appropriate form, write machine learning program that looks for patterns in stock data
 - Ex. Events leading up to heavy price drops, events leading up to price increases, other market patterns.
- After training the model on enough data sets and learning common patterns, try it on new data sets.
- Monitor real day to day stock prices, and compare them with what our code produces

Tools Intended to Use

- Python
- Scikit-learn (python machine learning library)

Evaluating the results

In order to evaluate our results, we will pick a few stocks and run our code on them to get predictions. We will compare these predictions to the actual stock market data for a number of days, calculating the error between our predicted stock price and the actual stock price on a certain day.