

Chapter 1

INTRODUCTION

The use of algorithms to make trading decisions has become a prevalent practice in major stock exchanges of the world. Algorithmic trading, sometimes called high-frequency trading, is the use of automated systems to identify true signals among massive amounts of data that capture the underlying stock market dynamics. Machine Learning has therefore been central to the process of algorithmic trading because it provides powerful tools to extract patterns from the seemingly chaotic market trends. This project, in particular, learns models from Bloomberg stock data to predict stock price changes and aims to make profit over time.

In this project, we examine two separate algorithms and methodologies utilized to investigate Stock Market trends and then iteratively improve the model to achieve higher profitability as well as accuracy via the predictions. Prediction of stock price level movement is thought to be a difficult task of monetary statistic prediction. Associate degree correct prediction of stock worth movement might yield profits for investors. As a result of the quality of exchange information, development of Economical models for predicting is incredibly troublesome. Statistical strategies and neural networks are usually used for statistic prediction. Since stock markets are complicated, nonlinear, dynamic and chaotic.

Artificial neural network is a mathematical model. It has capability to machine learning and pattern matching. Neuron is basic unit of nervous system such as brain. ANN is borrowed from central nervous system. It is inspired by biological technology. Biological neuron stores knowledge in memory bank, while in an artificial neuron the data or information is distributed through the network and stored in the form of weighted interconnection.

Neural networks among varied computing tools are more and more accustomed the monetary prognostication as neural nets are found to be technologically versatile and powerful, ideally suited to perform monetary market research. Many studies have shown that artificial neural networks have the capability to be told the underlying mechanics of stock markets. In fact, artificial neural networks are wide used for prognostication monetary markets.

Chapter 2

About The Company

Inventeron Technologies and Business Solutions LLP, or ITABS, is an Indianbased engineering and electronics company headquartered in Bangalore,Karnataka, India.

It is both product and services oriented software company having its produv=cts in wireless communication technology and provides quality service to its valuable clients in its domain.

Inventeron's core products are embedded components (including IC's, control Boards, Controller, Microprocessors , Fuel systems , Water Level Controllers , Security systems , Biometric Login systems, Wireless Devices etc.), Industrial Products (Network Tower Management Systems) and Apps (E-Commerce Apps , Website Design and Development)

- **PURPOSE** : To be leader in the software industry by providing Enhanced Services , Relationship and Profitability.
- **VISION** : To provide quality services that exceeds the expectations of over esteemed customers.
- **MISION** : To build Long-Term Relationships with our customers and clients and provide exceptional customer services by pursuing business through innovation and advanced technology.

2.1 Services

- Embedded Applications Development.
- Web Design and Development.
- IT Services.
- Server Maintenance.
- PCB Design.
- ERP .
- IT Training.
- Online Web and Multimedia.
- Cloud Computing.
- Business Solutions.

2.2 Products

- Smart Surveillance System.
- Safety and Security Systems.
- Biometrics.
- Smart Traffic Systems.
- Vehicle Tracking Systems.
- Education Management Systems.
- Hotel Management Systems.
- Personal Safety Equipments.
- Wireless Communication Devices.
- Water Controlling Units.

2.3 Domain Expertise

- Education .
- IT.
- Government.
- Food and Beverages.
- Retails.
- Health Care.
- IT Consulting.
- Outsourcing.
- HR Management.

2.4 What It Does

- It does Innovations in Embedded Electronics Providing Quality service and varieties of Engineering Products.
- It deliver the excellence and consequently improve what is essential for human progress by mastering Science and Technology.
- It have an excellent and experienced teams of engineers with atleast and advanced skills who makes the coding process so simple.

- They are enough passionate and dedicated towards their work and they are ready to work anytime regardless of time.

2.5 Training and Placement

2.5.1 Technical Training :

- Domain Expertise.
- Work On Live projects.

2.5.2 Non Technical Training :

- Behavioural Training.
 - a. Soft Skills.
 - b. Personality Development.
- Corporate Culture.
- Off Campus Placement Hacks.
- Entrepreneurship Skills Development.

Chapter 3

OBJECTIVES

Since the tick-by-tick entries retrieved from Bloomberg happen in non-deterministic timestamps, we attempted to standardize the stock data by discretizing the continuous time domain, from 9:00 am to 5:00 pm when the market closes. Specifically, the time domain was separated into 1-minute buckets and we discarded all granularities within each bucket and treated the buckets as the basic units in our learning algorithms. In the past decades, there is an increasing interest in predicting markets among economists, policymakers, academics and market makers. The objective of the proposed work is to study and improve the supervised learning algorithms to predict the stock price.

Technical Objective:

The technical objectives will be implemented in R. The system must be able to access a list of historical prices. It must calculate the estimated price of stock based on the historical data. It must also provide an instantaneous visualization of the market index.

Experimental Objective:

Two versions of prediction system will be implemented; one using Decision trees and other using Support Vector Machines. The experimental objective will be to compare the forecasting ability of SVM with Decision Trees. We will test and evaluate both the systems with same test data to find their prediction accuracy.

Chapter 4

SCOPE OF PROJECT

This project , in particular, learns models from Bloomberg stock data and to predict stock price changes and aims to make profit over time. In this project, we examine two separate algorithms and methodologies utilized to investigate stock Market trends and then iteratively improves the model to achieve higher profitability as well as accuracy via a prediction.

The will be useful for investors to invest in stock market based on the various factors. The project target is to create web application that analyses previous stock data of companies and implement these values in data mining algorithm to determine the value that particular stock will have in near future with suitable accuracy. These predict and analyzed data can be observed by individual to know the financial status of companies and their comparisons.

The main feature of the project is to generate an approximate forecasting output and create a general idea of future values based on the previous data by generating a pattern. The scope of this project does not exceed more than a generalized suggestion tool.

Chapter 5

RELATED WORK AND IMPACT

5.1 Stock Selection

Stock ticker data, relating to prices, volumes, quotes are available to academic institutions through the Bloomberg terminal and Stanford has a easily accessible one in its engineering library.

When collecting stock data for this project we attempted to have a conservative universe selection to ensure that we mined a good universe a priori and avoided stocks that were likely to be outliers to our algorithm to confuse the results.

The criteria we shortlisted by were the following:

1. price between 10-30 dollars
2. membership in the last 300 of SP500
3. average daily volume (ADV) in the middle 33 percentile
4. variety of stock sectors

The data we focussed on was the price and volume movements for each stock throughout the day on a tick-by-tick basis. This data was then further preprocessed to enable interfacing with Matlab and integrate into the machine learning algorithms.

5.2 Preprocessing

Before using the data in the learning algorithms, the following preprocessing steps were taken.

5.2.1 Discretization

Since the tick-by-tick entries retrieved from Bloomberg happen in non-deterministic timestamps, we attempted to standardize the stock data by discretizing the continuous time domain, from 9:00 am to 5:00 pm when the market closes. Specifically, the time domain was separated into 1-minute buckets and we discarded all granularities within each bucket and treated the buckets as the basic units in our learning algorithms.

5.2.2 Bucket Description

For each 1-minute bucket, we attempted to extract 8 identifiers to describe the price and volume change of that minute heuristically. We discussed the identifier selection with experienced veteran in algorithmic trading industry (footnote: Keith).

Based on his suggestions, we chose the following 4 identifiers to describe the price change:

1. open price: price at the beginning of each 1-minute bucket.
2. close price: price at the end of each 1-minute bucket.
3. high price: highest price within each 1-minute bucket.
4. low price: lowest price within each 1-minute bucket.

Similarly, we chose open volume, close volume, high volume and low volume to describe the volume change. With this set of identifiers, we can formulate the algorithms to predict the change in the closing price of each 1-minute bucket given information of the remaining seven identifiers (volume and price) prior to that minute². The identifiers help capture the trend of the data of a given minute.

5.3 Metrics

To evaluate the learning algorithms, we simulate a real-time trading process, on one single day, using the models obtained from each algorithm. Again, we discretize the continuous time domain into 1-minute buckets. For each bucket at time t , each model attempts to invest 1 share in each stock if it predicts an uptrend in price, i.e. $P(t) \text{ close} > P(t) \text{ open}$. If a model invested in a stock at time t , it always sells that stock at the end of that minute(t). To estimate profit, we calculate the price difference $P(t) \text{ close} - P(t) \text{ open}$ to update the rolling profit. If, on the other hand, it predicts a downtrend it does nothing. This rolling profit, denoted concisely as just "profit" in this report, is one of our metrics in evaluating the algorithm's performance.

In addition to profit, we also utilize the standard evaluation metrics: accuracy, precision and recall, to judge the performance of our models. To conclude, each time we evaluate a specific model or algorithm, we take the average precision, average recall and average accuracy and average profit over all 23 stocks in our universe. These are the metrics used for performance in this report.

Chapter 6

ADVANTAGES AND DISADVANTAGES

6.1 Advantages

- Stock market prediction is the act of trying to determine the future value of a company stock or other financial instrument traded on an exchange .
- The successful prediction of a stock's future price could yield significant profit. The efficient-market hypothesis suggests that stock prices reflect all currently available information and any price changes that are not based on newly revealed information thus are inherently unpredictable.
- Allowing you to offer employees extra incentives by granting **share options** - this can encourage and motivate your employees to work towards long-term goals
- Placing a value on your business.
- Increasing your **public profile**, and providing reassurance to your customers and suppliers.
- Allowing you to do business - eg acquisitions - by using **quoted shares** as currency.
- Creating a **market** for the company's shares.

6.2 Disadvantages

Market fluctuations - your business may become vulnerable to market fluctuations beyond your control - including market sentiment, economic conditions or developments in your sector.

Cost - the costs of flotation can be substantial and there are also ongoing costs of being a public company, such as higher professional fees.

Responsibilities to shareholders - in return for their capital, you will have to consider shareholders' interests when running the company - which may differ from your own objectives.

The need for transparency - public companies must comply with a wide range of additional regulatory requirements and meet accepted standards of corporate governance including transparency, and needing to make announcements about new developments.

Demands on the management team - managers could be distracted from running the business during the flotation process and through needing to deal with investors afterwards.

Chapter 7

CODE SNIPPET

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import math
from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import mean_squared_error
from keras.models import Sequential
from keras.layers import Dense, Activation
from keras.layers import LSTM
# IMPORTING DATASET
dataset = pd.read_csv('EOD-AAPL_in.csv', usecols=[1,2,3,4])
dataset = dataset.reindex(index = dataset.index[::-1])
# TAKING DIFFERENT INDICATORS FOR PREDICTION
OHLC_avg = dataset.mean(axis = 1)
HLC_avg = dataset[['High', 'Low', 'Close']].mean(axis = 1)
close_val = dataset[['Close']]
# PLOTTING ALL INDICATORS IN ONE PLOT
plt.plot(obs, OHLC_avg, 'r', label = 'OHLC avg')
plt.plot(obs, HLC_avg, 'b', label = 'HLC avg')
plt.plot(obs, close_val, 'g', label = 'Closing price')
plt.legend(loc = 'upper right')
plt.show()
# TRAIN-TEST SPLIT
train_OHLC = int(len(OHLC_avg) * 0.75)
test_OHLC = len(OHLC_avg) - train_OHLC
```

LSTM MODEL

```
model = Sequential()
```

```
model.add(LSTM(32, input_shape=(1, step_size), return_sequences = True))
```

```
model.add(LSTM(16))
```

```
model.add(Dense(1))
```

```
model.add(Activation('linear'))
```

PLOT OF MAIN OHLC VALUES, TRAIN PREDICTIONS AND TEST PREDICTIONS

```
plt.plot(OHLC_avg, 'g', label = 'original dataset')
```

```
plt.plot(trainPredictPlot, 'r', label = 'training set')
```

```
plt.plot(testPredictPlot, 'b', label = 'predicted stock price/test set')
```

```
plt.legend(loc = 'upper right')
```

```
plt.xlabel('Time in Days')
```

```
plt.ylabel('OHLC Value of Apple Stocks')
```

```
plt.show()
```

```
print("Last Day Value:", np.asscalar(last_val))
```

```
print("Next Day Value:", np.asscalar(last_val*next_val))
```

Chapter 8

SNAPSHOT

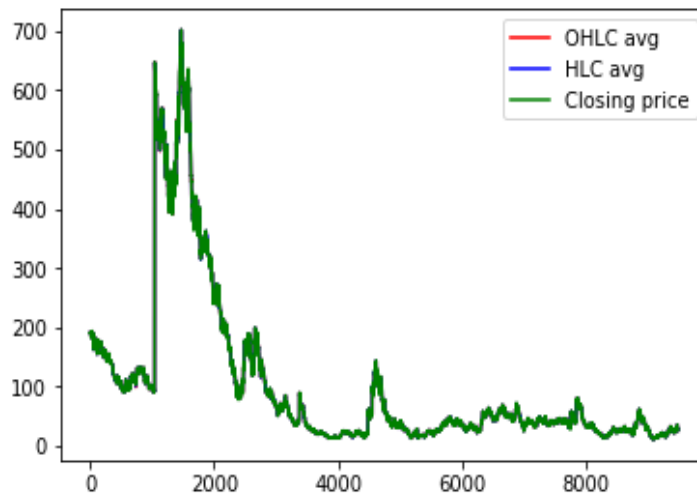


Fig 8.1 Plotting all indicators in one plot

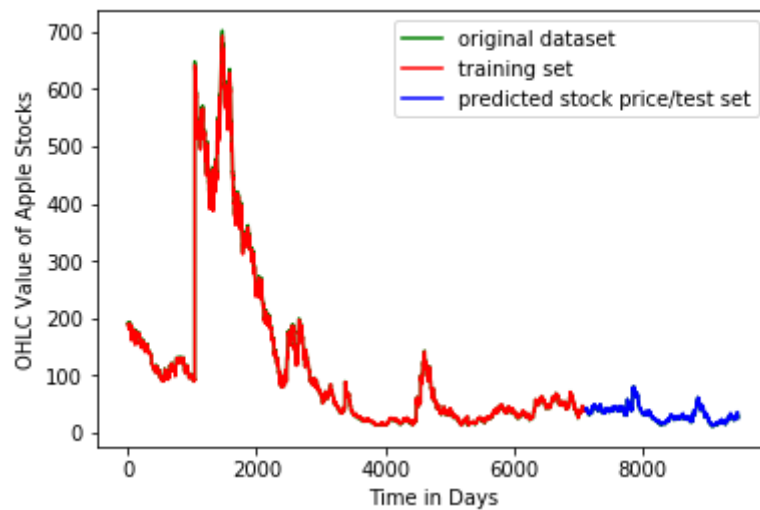


Fig 8.2 Plot of Original dataset , Training Predictions and Testing Predictions.

Chapter 9

REQUIREMENTS

9.1 Hardware Requirements

1. Processor i3 and above.
2. Ram 4gb and above.

9.2 Software Requirements

1. Anaconda navigator as an applications wrapper hub.
2. Spyder as GUI interface for coding.

9.3 Libraries Required

1. Pandas.
2. Matplotlib.
3. Sklearn.linear_model.

Chapter 10

CONCLUSION

Predicting stock market trends using machine learning algorithms is a challenging task due to the trends being masked by various factors such as noise and volatility. In addition, the market operates in various local-modes that change from time to time making it necessary to capture those changes in order to be profitable while trading. our algorithms and models were simplified, we were able to meet our expectation of reaching modest profitability. iterative improve In conclusion, our experience in this project suggests that machine learning has great potential in this field and we hope to continue working on this project further to explore more nuances in improving performance via better algorithms as well as optimizations. iterative improvements achieved through sequential optimizations in the form of discretization, realization of time-locality, smoothing improved results significantly. Cross-validation and forward search were also powerful tools in making the algorithm perform better. would be exploring other international stock markets to find locations where algorithmic trading is able to perform better.

REFERENCES

- [1] Jure Leskovec, TA: Keith Sillats HW 4
- [2] R.Wilson and R.Sharda , “Bankruptcy prediction using neural networks” , Decision Support Systems.
- [3] Neelama Budhani, Dr.C.K.Jha, Sandeep K. Budhani “Application Of Neural Network In Analysis Of Stock Market Prediction”, International Journal Of Computer science And Engineering Technology.
- [4] B. Manjula, S.S.V.N. Sharma, R. Lakshman Naik, G. Shruthi, Stock Prediction using Neural Network, International journal of advantage engineering sciences and technologies.

