Department of Informatics,

University of Leicester

CO4015 Computer Science Project

Dissertation

for

Stock Market Prediction using python

Author: Osato Osagie (1690 11647)

# Introduction

The term ‘stock market’ does not refer to a single market. It refers to several stock exchanges dispersed around the world. In these stock exchanges, traders and investors can purchase and sell shares of companies that are publicly traded. The share prices of these companies constantly fluctuate in response to the law of supply and demand.

A share is a small possession stake in a publicly traded company. The price of a stock portrays the expectation of stock investors and market analysts on the company’s future earnings.

When traders believe a company will perform well, they bid the price up by creating demand for stocks in that company. On the other hand, those traders who don’t believe in a company’s future will bid the price down by selling their holdings, thus creating an excess supply. Sellers aim to receive a high price for each share hopefully more than their initial payment, whilst buyers aim to acquire shares at their lowest price possible so that they can make a profit when selling the stock.

Investing in stocks is deemed as a reliable method to achieve profits that beat inflation over time. The returns, on average do better than those of other investments, such as bonds and commodities. According to research, as of February 2021, the total market value of all companies trading on the London Stock Exchange stood at 3.67 trillion British pounds [1].

There are two ways to make a profit on the stock market. Investors can either trade stocks or decide to hold them. Trading requires investors to buy and sell stocks frequently, taking advantage of little ticks in cost. Investors who purchase and lean toward to let their stocks appreciate in esteem over time. In some cases, investors who hold shares get rewarded with regular payments of dividends.

Investors have previously found ways to obtain insight about the businesses listed on the market for as long as markets have existed in order to increase their investment returns. However, owning to the size market and the pace at which transactions are conducted, this is not possible today. Simple statistical analysis of financial data can reveal certain trends, but in recent years, investment firms have increasingly turned to Artificial Intelligence (AI) systems to search for patterns in vast quantities of real-time equity and economic data.

AI is the ability of machines to behave like humans and learn autonomously. For instance, a machine might display learning and problem-solving abilities without the use of hard-coded software containing detailed instructions [1].

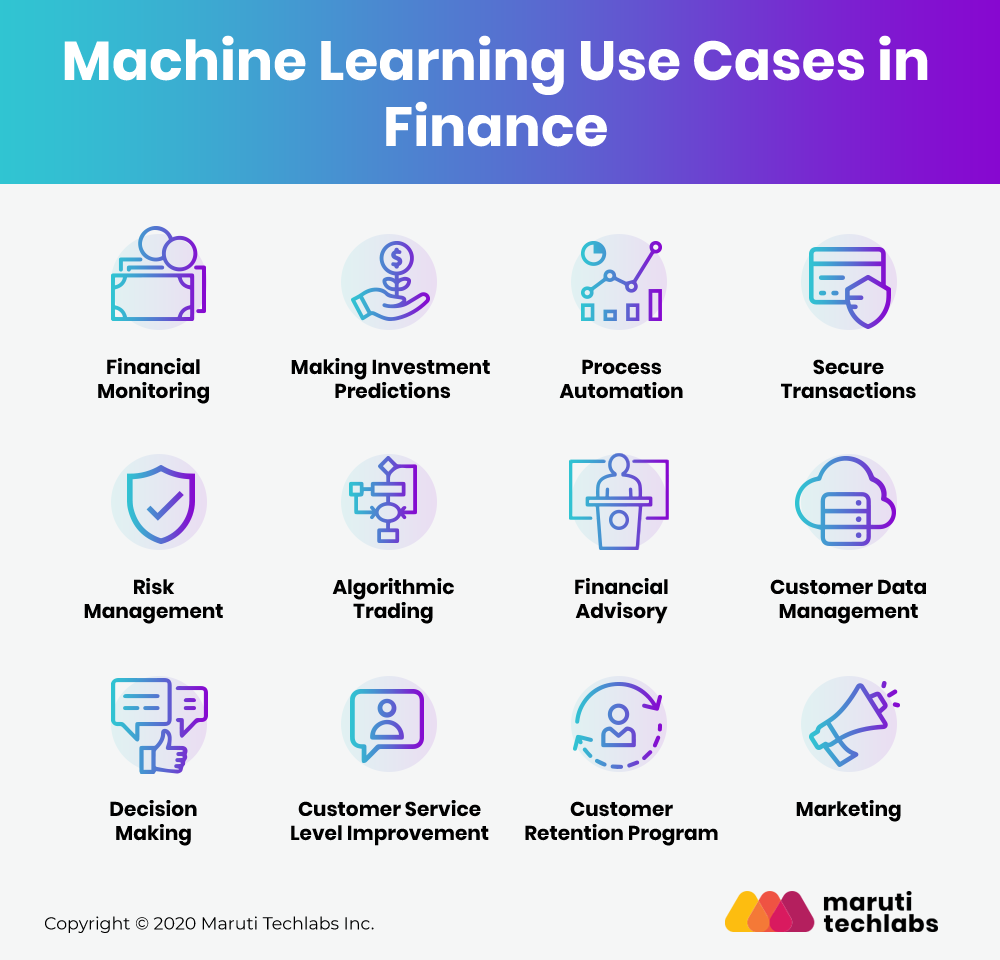


Figure 1- An image showing the main use cases of Machine Learning in Finance [2]

Machine Learning (ML) is a subfield of AI, that enables machines to learn from historic data or experiences without being explicitly programmed. Figure 1 shows the different use cases for Machine Learning in finance. The project I will be building focusing on using ML to make investment predictions. Using ML to make investment predictions is advantageous as it can lead to better predictions of stock prices, fewer errors, and greater efficiency for the investor. To do this, ML algorithms extract key insights from the dataset, learns from it then apply several techniques to accurately predict the result.

However, it is important to take into consideration the other factors that might affect the price of a company’s stock. The stock market is very volatile, thus meaning no system can accurately predict it.

Deep Learning (DL) is a subfield of ML. It teaches computes to learn by example in the same way that humans do. A prominent example of DL is self-driving vehicles, allowing them to identify a stop sign or differentiate between a pedestrian and a lamppost. DL models are able to achieve cutting-edge precision, often even outperforming humans. A wide collection of labelled data and neural network architectures with several layers are used to train models.

## 1.1 Aims

Predicting markets has become an increasing priority for investors. The primary goal of an investor is to buy a stock when its value is low and sell when the value of the stock is high. However, this can be daunting for financial investors as they are unaware of the stocks that will return maximum profits. Using Machine learning to predict the long-term value of a stock makes this process somewhat easier. This project aims to predict the stock market price of a company using supervised machine learning algorithms.

## 1.2 Objectives

To meet the aims this project sets out to achieve, I have broken it down into several tasks. These objectives are as follows:

* Obtain real-time and historical equity data from Alpha Vantage API.
* Clean data and form data sets with the obtained data.
* Build python functions to calculate technical indicators from the obtained dataset.
* Train a Decision Tree and SVM model to predict the S&P 500.
* Test various models to find which one works best for predicting Standard and Poor’s 500 (S&P 500).
* Fine tune model parameters to have as low bias as possible while also having low variance on the training data.
* Split datasets into training and test data, and train models with data.
* Tests the models with data and measure accuracy using metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared (R2) .
* Evaluate which features have the most impact on stock prices.
* Visualise results using a line chart showing the predicted prices versus the actual prices of the stock.
* Implement a user-friendly interface for the prediction tool.

## 1.3. Changed objectives

Following the feedback from my interim report to ‘add more up to date technologies, some objectives have been changed. These include:

* Train a Ridge Regression, Random Forest and Long Short-Term Memory models to predict the S&P 500.

## 1.4. Resources & Tools

This project is built using Python programming language. I chose to use Python because it is easy and flexible to use. Python is also versatile and has a robust collection of libraries that make machine learning tools easily available to use. Also, Python has a diverse pack of visualisation options available which makes it ideal for creating graphs and charts.

The Python code will be written using Jupiter notebook. It is a web application that allows users to create documents containing live code and visualizations. In addition, I have used various Python libraries to access tools that have enabled me to build my project. These include:

* SKlean
* Matplotlib
* Pandas
* Numpy

# Literature Review

This chapter will discuss current literature that will be used to set the stage for this project. Such literature are based on subjects surrounding the current technology development of AI in trading and investing,

## 2.1. Impact of AI on Trading and Investing

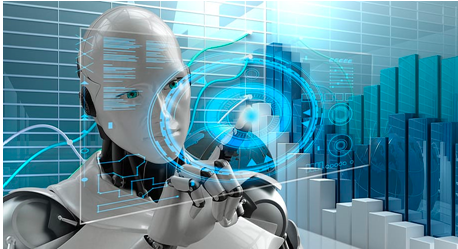


Figure -Photo credit: [3]

AI, ML, and DL have been transforming finance and investing. Although humans remain an important part of the trading equation, AI is becoming increasingly important. Electronic trades account for approximately 45 percent of cash equities trading revenues, according to a new report by Coalition [4], a U.K research company. Though hedge funds are wary of automation, many of them use AI-powered analysis to generate investment ideas and build portfolios.

“Machine Learning is progressing much faster, and financial institutions are among the early adopters.” Intelenet Global Service's vice president of global business growth, Anthony Antenucci, recently said [5]. Globally companies are developing new products and technology that use AI to make trading and investing more data-driven and effective. Bloomberg announced in September 2017 that Japan’s third largest lender will use AI in the equities sector through algorithm-based services for institutional clients [6]. In the United States, Merrill Lynch is currently testing a stock platform to find value in small-cap stocks that analysts would otherwise ignore [7].

Currently, there are companies already using AI for smarter trading. Aquan, a UK data science company [8] uses its platform to democratize trading by encouraging data scientists of all backgrounds to build algorithmic trading strategies that assist in the resolution of investment problems. As a result, investment clients will benefit from data science without having to invest in costly in-house expertise. Aquan has had significant industry impact, they graduated from Techstars in 2018 and was recently named the 2019 Europa Awards’ Hottest Fintech n Europe.

Similarly, EquBot’s proprietary investment technology, which is affiliated with IBM, blends AI with an active exchange-traded fund (ETF). The business centralizes the investment process by gathering and processing data from different sources (news articles, social media postings, financial statements) from around the world to “build a cause-and-effect understanding of economies, businesses, and management”. The impact EquBot has had in the industry has also been significant, recently they launched the AI-powered Foreign Equity ETF, which aims to invest in established international markets outside of the United States.

## 2.3 Alpha Vantage Stock Application Programming Interface (API)

In this project, I will be using an API called Alpha Vantage. An API is where a website provides a set of structured Hypertext Transfer Protocol (HTTP) requests that return JavaScript Object Notation (JSON) or Extensible Markup language (XML) files.

In this project, the Alpha Vantage API provides access to historical and real-time data for a variety of markets. The API allows me to access the data directly in python, from there I can manipulate the data or store it for later use. Alpha Vantage provides its service at no fee. They permit 5 requests per minute and 500 requests per day.

If a higher rate cap is needed, there are many premium plans available. Premium plans vary in price from $29.99 per month for 30 requests per minute to $249.99 per month for 1200 requests per minute. For stocks, Alpha Vantage provides historical and real-time info. There are many time frames to choose from, ranging from 1-minute bars to weekly. The most important benefit is that it is absolutely free. Furthermore, the information is comprehensive. For stocks, we find price data going back 20 years.

# References

[[1] https://www.statista.com/statistics/324578/market-value-of-companies-on-the-london-stock-exchange/](%5b1%5d%20https://www.statista.com/statistics/324578/market-value-of-companies-on-the-london-stock-exchange/)

[2] Marutitech.com (-) 12 Use Cases of AI and Machine Learning In Finance [ONLINE]. Available at: <https://marutitech.com/ai-and-ml-in-finance/>

[3] <https://www.pxfuel.com/en/free-photo-qualj>

[4] <https://www.crisil.com/en/home/our-businesses/crisil-coalition.html>

[5] <https://www.itprotoday.com/machine-learning/how-ai-trading-systems-will-shake-wall-street>

[6] <https://www.bloomberg.com/news/articles/2017-09-20/mizuho-is-said-to-offer-ai-trading-service-before-mifid-overhaul>

[7] <https://www.cybersecobservatory.com/2017/06/06/artificial-intelligence-transforming-investment-strategies/>

[8] <https://www.auquan.com/>