**STOCK MARKET ANALYSIS**

**RELATED WORK: LITERATURE REVIEW AND REFERENCE**

The ‘Related Work’ section will be synthesizing our literature review and reference. We cite and discuss references gained from at least three (3) different sources. These could include technical or scholarly journals, papers, books, market reports, theses/dissertations or websites. Our review and references for this project include:

• Data reading from various resources in hardcopies and softcopies that captured from the summaries of existing researches, journals and market reports;

• Market talk about trend and modelling of stocks and their measurement using Technical Analysis;

• Reliability of an algorithm as a prediction tool;

• The goal of this project is to find a small enough problem (stock performance) that can be modelled by using less number of parameters (closing price of a stock).

**Related WORK SECTION**

According to a famous fictional science in [[1]](#footnote-1)Isaac Asimov’s Foundation namely Psychohistory motivation of historical and current events which combines history, sociology, and mathematical statistics of making general predictions about the future behaviour of very large groups of people saying that *it would not work* to see the past and predict the future algorithmically. [[2]](#footnote-2)Asimov presents a device called Prime Radiant, a device that stored psychohistorical equations showing the future development of humanity. The device;

Radiant, was designed by Hari Seldon; operates through the power of the mind allowing the user to zoom in to details of the equations and change them. As it being interactive, employs a type of colour-coding to equations within itself for ready comprehension by psychohistorians. the colours name Seldon Black, Speaker Red, Deviation Blue and other colours such as Notation Green and Projection Purple. The Radiant later proceeded to the 2nd series of foundation called Seldon Plan. Through this Seldon Plan, Seldon himself doubt the ability of pure mathematical approach to predict historical processes, and that he recognized that the development of psychic skills had the ability to invalidate the assumptions underlying his models. The foundation’s series has developed psychohistory to predict the actions of large groups of humans even robots technically falls under umbrella of psychohistory because humans built them and thus represent more or less human “action” or at least possess a thought-framework similar of their human creators that psychohistory can predict their actions. However, [[3]](#footnote-3)psychohistory *cannot predict* the actions of a sentient alien race; their psychology differed so much from that of humans which normal psychohistory cannot understand or predict their actions. In the end of the series, Asimov admitted that he is not able to help on a science in which things could be predicted by statistical basis or probabilistic. This was proven by a e-book writer namely Bruce Mazlish[[4]](#footnote-4) who wrote that anyone working in the field of psychoanalysis and history, in relation to either individual or group behaviour, must learn not only his own historical discipline but the very involved theory and data of combined studies of sociology, economics, political science, psychoanalysis and history.

However, one famous movie fiction had offered better view. Euclid, the computer predicting stock market behaviour in[[5]](#footnote-5)[Pi (film)](http://en.wikipedia.org/wiki/Pi_(film)) proved that human can see the past and predict the future algorithmically (but not exactly what the movie shown which one person very rich and destroy the markets the next day). The film, about a[mathematician](https://en.wikipedia.org/wiki/Mathematician)with an obsession to find underlying[complete order](https://en.wikipedia.org/wiki/Complete_order)in the real world, contrasts two seemingly irreconcilable entities: the imperfect,[irrational](https://en.wikipedia.org/wiki/Irrational)humanity and the[rigor](https://en.wikipedia.org/wiki/Rigour#Mathematical_proof)and regularity of mathematics, specifically[number theory](https://en.wikipedia.org/wiki/Number_theory). The film was well received with 88% approval rating based on 56 reviews with an average rating of 7.3/10. The film however contrasts with the Machine Learning technique such as [[6]](#footnote-6)Adaptive Neuro Fuzzy Inference System (ANFIS). This technique can work as predictor for arbitrary time series, predicting the next value based on previous values, even though the relationship between samples is often hard to define. A deep learning in mathematical area, ANFIS is possible to identify two parts in the network structure, namely premise and consequence parts. Each layer has its own function that the final layer is able to produce the final output which back to the space of initial parameters.

Algorithm predicts the future is one of popular topic where we could see many studies on it being done for instance Massachusetts Institute of Technology (MIT) through The Computer Science and Artificial Intelligence Laboratory has made the breakthrough in predictive vision by training an algorithm using 600 hundred hours of YouTube videos. By searching for patterns and recognizable objects like hands and faces, the algorithm was able to predict human interactions such as hugging, kissing, shaking hands or high fiving. Tests proved the algorithm to be correct 43 percent of the time when shown a still frame taken one second before the action happens. By way of comparison, human subjects were able to correctly predict the action 71 percent of the time. Nonetheless, a system program called [[7]](#footnote-7)Halting Problem has shown otherwise. Surely algorithms can converge 100% true prediction but not 100% accuracy due to flipping around the problem whether the algorithm can compute the running times of given code without running which lies under Halting Problem as claimed by [computability theory](https://en.wikipedia.org/wiki/Computability_theory_(computer_science)). The halting problem is the problem of determining, from a description of an arbitrary[computer program](https://en.wikipedia.org/wiki/Computer_program)and an input, whether the program will finish running, or continue to run forever.[[8]](#footnote-8)[Alan Turing](https://en.wikipedia.org/wiki/Alan_Turing)proved in 1936 that a general[algorithm](https://en.wikipedia.org/wiki/Algorithm)to solve the halting problem for all possible program-input pairs *cannot exist*. For any program f that might determine if programs halt, a "pathological" program g called with an input can pass its own source and its input to f and then specifically do the opposite of what f predicts**g**will do. No**f**can exist that handles this case. A key part of the proof was a mathematical definition of a computer and program, which became known as[Turing machine](https://en.wikipedia.org/wiki/Turing_machine); the halting problem is[undecidable](https://en.wikipedia.org/wiki/Undecidable_problem)over Turing machines. Turing's proof is one of the first cases of[decision problems](https://en.wikipedia.org/wiki/Decision_problem)to be concluded. The theoretical conclusion that it is not solvable is significant to practical computing efforts, defining a class of applications which *no programming invention can possibly perform perfectly*.

Some questions mingle around on why machine learning, neural networks and other AI-approaches are not widely used in stock market predictions as lots market practitioners have tried to treat stock market as a time series and use mathematical models to predict future. And there are several firms that specialize in high frequency like algorithm trading by the most famous fund manager specialized in technology; Renaissance Technologies but only able to have US$ 65 Billion assets under management (AUM). This contrast with other AUMs which not technology adapted firms like Black Rock in the Scientific Active Equities, SAE; an investment group with Barclays Global Investors, which BlackRock acquired in year 2009 were far too large for that. BlackRock has US$ 4.5 Trillion on AUM. However, according to an [[9]](#footnote-9)online article, Machine Learning automates the discovery of predictive algorithms that are able to **continuously improve as they get access to more data**. The focus has been on automating many of the tasks traditionally performed by data scientists, including data cleaning, model selection, data clustering, automatic feature generation and dimensionality reduction.

There are also several views on analysis tools practised by market traders for instance Technical Analysis is claimed to be looed down by market practitioners.  Lewis in his [[10]](#footnote-10)book Flash Boys highlights the rise of Algorithmic trading and also hints the fact it is doomed to fail, [[11]](#footnote-11)BlackRock investments lost $1.5b on algorithms trading. The problem with the issue of wall street pros hating technical analysis is about being able to package technical analysis and reproduce it, the moment it is packaged, it becomes an Algo and loses its edge. There are successful ideas like Eliott Wave Analysis, that cannot really be reproduced because the algorithms are all about market making not buy side. Some buy side traders claim, no point in wasting time on Algorithms, etc. You are not going to make a fortune because you have a robot. The guys on the sell side need jobs and the algorithms are just their pitch to the fund firms who employ them.

Active traders have over the past years experiencing large losses as a result of volatile markets and recent market bubbles. Many claimed human error is the most reasons for losses. Therefore, can machine learning algorithms or models predict the stock prices? According to a [[12]](#footnote-12)published article by *I Know First*, an investor does due diligence, by performing extensive research regarding a firm, they are still susceptible to errors mostly due to fear or greed. Machine-learning algorithms make far fewer mistakes than humans would. Although many say that a big problem to using algorithms is they find it difficult to react to systematic events, due to new financial uncertainties being seen on the macro level, especially from Central Banks, individual investors lack this quality as well[[13]](#footnote-13). Additionally, unlike most AI-based algorithms, machine learning has much more capabilities that have allowed it to react better and give better result than human do. As well, the algorithm is able to trade forex, which is a valuable instrument used strongly by hedge funds across the globe. The cost the reduction and the higher premiums offered by these new types of algorithms should convince investors to utilize these new tools today.

The most influential study on this would be [[14]](#footnote-14)a paper in 2000 by Andrew Lo, et al., “Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation”. They concluded that “over the 31-year sample period, several technical indicators include geometric shapes in historical price charts do provide incremental information and may have some practical value.

Relying on Machine Learning is a phenomenon. In this technology era, the evaluation of machine learning (ML) models is a crucial step before deployment. It is essential to assess how well a model will behave for every single case. In many real applications, along with mean error of the model, it is also important to know how this error is distributed and how well probability estimations are made. It is claimed that many current ML techniques are good in overall results but have a bad distribution assessment of the error. It is therefore a calibration techniques and calibration measures are important element in the ML practice as from the scientific context, the primary goal of ML methods is to build a hypothesis (model) from a given data set while Calibration is a post-processing technique to improve error distribution of a predictive model. Calibration and ML complete each other. According to Philip Tetlock and Dan Gardner through their book titled Super Forecasting-The Art & Science of Prediction (2016) a study between two variables; forecast and percent correct show how calibration and resolution capture distinct facets of good judgement. The result shows the calibration and resolution are superb because the combination of calibration and resolution which available in the ML make good forecast.

Although these literature presents various opinions and views on prediction and forecasting by algorithms or Machine Learning, this project paper will primarily focus on their application in stock market analysis.

1. Book ‘Foundation Trilogy’ by Isaac Asimov [↑](#footnote-ref-1)
2. Isaac Asimov <https://en.wikipedia.org/wiki/Isaac_Asimov> [↑](#footnote-ref-2)
3. Psychohistory (fictional) <https://en.wikipedia.org/wiki/Psychohistory_(fictional)> [↑](#footnote-ref-3)
4. e-Book ‘Psychoanalysis and History’ by Bruce Mazlish [↑](#footnote-ref-4)
5. Pi (film) <https://en.wikipedia.org/wiki/Pi_(film)> [↑](#footnote-ref-5)
6. Research Paper ‘Fuzzy Modelling Using Generalized Neural Networks and Kalman Filter Algorithm’ by Jyh-Shing R. Jang of University of California. [↑](#footnote-ref-6)
7. <https://en.wikipedia.org/wiki/Halting_problem> [↑](#footnote-ref-7)
8. <http://www.newsweek.com/artificial-intelligence-algorithm-predicts-future-473118> [↑](#footnote-ref-8)
9. https://www.pionline.com/article/20170320/ONLINE/170319958/machine-learning-will-transform-investment-management [↑](#footnote-ref-9)
10. Book ‘Flash Boys’ by Michael Lewis [↑](#footnote-ref-10)
11. <https://www.zerohedge.com/news/2017-01-11/blackrocks-robo-quants-are-pace-post-r>ecord-losses [↑](#footnote-ref-11)
12. https://iknowfirst.com/algar-ai-algorithmic-based-hedge-fund-why-hedge-funds-should-adapt-with-technology-to-continue [↑](#footnote-ref-12)
13. <https://iknowfirst.com/algar-algorithmic-trading-strategies-for-european-stocks-returns-up-to-193> [↑](#footnote-ref-13)
14. https://[scholar.google.com/scholar](https://scholar.google.com/scholar?hl=en&as_sdt=0%2C32&q=Foundations%2Bof%2Btechnical%2Banalysis%3A%2BComputational%2Balgorithms%2C%2Bstatistical%2Binference%2C%2Band%2Bempirical%2Bimplementation&btnG=) [↑](#footnote-ref-14)