2019

Applying Machine Learning to Day Trading of Stock

UG Final Year Project Interim Report

INTISAR HAQUE 150370413 SUPERVISOR: PAULO OLIVA

Abstract

Require results and conclusion to effectively write.

Contents

Introduction

A user looking into a Stock Market realises ample opportunities for growth and profit. A successful buying and selling of a security generates profit using the basic and unitary concept of "buy low, sell high". This concept can be applied to a myriad of different trading examples. Bulk buy printers at £300 a unit and sell at £500 a unit and one can see himself with £200 profit per item sold.
The same concept is applied to the Stock Market. Hundreds of companies go public to open their shares to investors in hopes of gaining easy access to a large capital. This opens up a relationship with market participants ranging from individuals, investment banks, and insurance companies all with their own applied methods of when would be the right time to purchase and sell stock.

However, the multiple system's put in place to make a profit out of the market are not necessarily reliable. Different schools of thought dictate the market to be random and a gamble outside of economic analysis, while some parties claim to be making a profit out of these gambles.

This report breaks down the different methods used to make a profit from market movement and narrows down on technical time series analyses. Different methods of this analysis are evaluated and the way machine learning can have an effect on this is discussed.

Problem Statement

Day trading is a newer approach to solving the stock market. Fundamental analysis is ignored and technical analysis is applied instead. Traders opt for short term profits (within the day) whereas stock-brokers look to invest over long term periods.
However, many experts and specialists claim that day-trading is a form of gambling and that stock movement outside of economic/business factors are entirely random [1]. However, day-trading has been a tactic for decades and there have been made many rules and regulations around it [2].
This report will attempt to understand whether different day-trading tactics have merit and if it's the results are reproducible. If so my program will aim to create algorithms of the analysis techniques and check whether using historical data these techniques can repeatedly output a profit.
Further from that I will aim to analyse the different analysis algorithms and compete them to be able to tier the different tactics.
Room exists for AI and automated trading to have worthwhile implementation. Day-trading can take up long schedules in individual's days, and the rule-systems that they have in place can translate well into machine. Trader psychology has a large effect on reliable trading, implementing automated rule sets could eliminate ill behaviours caused from mind sets such as fear of missing out or revenge trading.

Finance Background

Before we break down how entities predict stock price movement, we must first discuss the value of the stock itself. Share price multiplied by amount of shares available for purchase dictate a company's market cap.

But how is share price determined? How is Tesla's stock currently $333.04 at the time of writing? When a company decides to become public it goes through an Initial Public Offering (IPO). This process has many steps, one of the main ones being the pricing of each unit. Here, large banks work with the company to determine what they believe the company to be worth and how many shares to split up the worth by. The price offering is then released to the public where potentially thousands of investors queue up to purchase the stock.

From here it's a game of supply and demand. The amount of stock that a company puts out is limited. A trillionaire can't just off the whim decide to purchase all the stocks of Facebook. This static supply causes the demand of the stock to be the main, if not sole factor, in deciding the worth of the company. Everything is succeeded by the demand.

Therefor to capitalise on the idea of buying low and selling high, investors must use techniques to predict when a company's value is expected to be increased by a surplus in demand or expected to decrease by a withdrawal of demands.

Understanding these behaviours, despite being quite nuanced, is simple for the everyday person to see. Apple is a simple example. In 2007, despite being public for decades their share price was only $12.15. As the popularity of their products such as the iPod and the iPhone started to increase, in turn raising their revenue, more and more stake holders started taking notice.

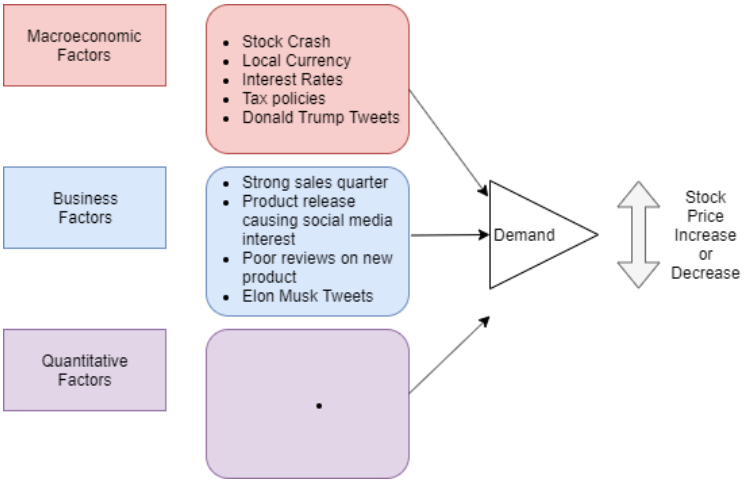


Fig 1.1 Quantitative Features to be added.

Fig 1.2

As Apple became more prolific in the public eye different investors around the world started speculating that Apple would only get bigger. They wanted to be a part of the company itself, as a shareholder. This resulted in enormous amounts of investors flocking to Apple to purchase their stock with the hopes of receiving dividend checks but more importantly in cashing out once the stock price reaches a huge level.

More closely, let's consider an individual who purchased 100 shares of Apple on 5th January 2007 costing him $1215, reasons of doing so being dismissed. If he were to sell all 100 stocks on 25th November 2019 (in time of writing), he'd receive $26454. That's over $25000 of profit (dismissing dividend payments) from only putting an initial capital dump of $1200.

This is the investor's ideal and dream scenario. Of course for every Apple there are thousands of other companies who do not perform the same, Apple being unicorn like in their rise. And for every investor who bought in on January 2007 there are thousands of investors who bought in to Apple in 2018 when their stock price was already into the hundreds.
Considering the above example, that particular investor made a 2000% profit over 12 years. If we consider the amount of money that investment banks and hedge funds are willing to spend in investments, a million dollar buy in soon becomes a billion.
With this simple flow of money it becomes quite clear on why stock investment has so many minds and methods behind it. The money to be made is astounding.

Aim

The aim of the project is to establish different stock-market prediction techniques and algorithms and how trading can be used to predict seemingly random movement.
To then understand how this can be aided by data mining on historical stock price datasets.
The program made will be able to match chart patterns with market movement and use data mining to rank the different chart patterns.

Objectives

- Develop a clear explanation to the limitations of day trading on stock prices and thereby the limitations of AI analysis on day-trading profit making.
- Create algorithms that suitably represent different rules used by day traders.
- Read in historical stock-price data using an API and match algorithms to the chart patterns.
- Be able to match how chart patterns link to market movement and then tier the observed rules.
- Finish the product to be able to function live and use an assortment of the chart patterns in decision making for buying stock.

Literature Review

Approaches to the Stock Market

Fundamental Analysis (Expert Systems and external factors)

Fundamental analysis is the idea of understanding a company's history combined with the industry they lie in and which economic region they operate from and within. All these things considered, investors can make decisions on their purchases. Such things to consider are: a company's quarterly records, them performing higher than expected; industry rivals, a competitor potentially stealing the company's market share; state of the nation's economics, if FTSE 100 have been declining for a long period of time this is an indication of a UK company's stock potentially not rising either.

Fundamental analysts spend their whole lives learning about the relationship between various financial factors and a company's stock price. Banks hire analysts by the thousands in a single city alone.

However many questions have been raised on the worth of stock market predictions. Investors such as Warren Buffet and more notably Jack Bogle have made their worth on not trying to beat the market, but to invest in the market itself and gaining profit from the increase in the market. This is obviously backed up by the fact that stock markets, despite the numerous crashes continue to grow.[3]

These 2 investors advise to spread your investments, with the latter even coming up with the idea of index funds. So with this tried and tested method why do companies and individuals attempt to forecast growth? Entities insist on beating the market for a faster rate of return, essentially people want more money and they want it sooner.

Trader *psychology is a huge factor on why index investors* don't believe you can beat the market.[4]

This is more so backed up by a research paper conducted in 1932 [5]. In this Cowles makes a discussion comparing how insurance companies and financial services perform in forecasting stock prices. Results from 45 organisations were gathered, mainly from magazine publication with their recommendations.

The results of the organisations come with the underline of how the predictions scored related to market averages. That is to say if the market increased by 5% and a particular prediction increased by 2.5% this would result in a score of -2.5%. Minus scores is what is seen across this paper. Table 4 of this paper shows us 24 randomly selected results. 10/24 (42%) of the forecasts beat the market with the remaining 58% having lost to the market.

This low score of beating the market is reiterated in multiple results of this paper with only 12/36 (33%) financial services and insurance companies beating the market. If specialists could not fare to beat the market then it makes sense to "bet" on the market itself. Again, investors don't simply want to beat the market, they want to beat everybody else beating the market. This point of winners and losers tie on later.

Importantly, the results of this paper must be analysed. Results were taken in USA in 1932. If one were to look at historical data of large indices, NASDAQ, FTSE, Dow Jones etc there can almost be a guaranteed visual of market prices increasing year on year.

Market crashes result in huge slumps of stock prices, investor confidence being at the low. Low confidence means low demand meaning rapid selling of stock leading to stock prices plummeting, leading to crashes. However, if you take another step back and plot in a straight line how different indices have performed since the start of prices ever being recorder, you will see a line with a positive gradient. The market always increases.

This leads onto the results of the paper being using results from 1932. Only a few years after the great depression, though households struggle and companies hit historic lows, the direction the

market takes following a crash is quite consistent throughout numerous crashes. That direction is up.
I therefor argue that beating the market in 1932 or other rising economy periods is a difficult task, regardless of whether it is an expert or a novice investor attempting to do so. *None the less, these companies could not beat the market.*
The sentiment I received from this paper was quite anti-investment. The creation of index funds, such as those offered by The Vanguard Group, from 1960's onwards however can tackle the issues presented by the paper. Instead of aiming to beat the market, individuals can bet with the market and receive returns that are overall average. This is quite commonplace, with banks and pensions funds opting to use different indices to receive returns that overtime looks to be generous.
This of course is not enough; individuals and banks look towards methods to beat the average, to increase their returns. When given the option of taking £10 off of an individual and promising them either £50 in 30 years' time versus risking it to £20 tomorrow, different entities will have different psychologies towards this.

Technical Analysis

The report now moves in to compare different ways of technical analysis, that is to partially or totally dismiss qualitative factors that go into prices changing (a new CEO with a proven track record could increase stock price) and apply different time-series techniques.

Traditional Quantitative Analysis vs Day Traders

Technical analysis can be split up into 2 main ideas, Quantitative Analysis versus Trading. The former term is often used in conjunction with many similar terms often leading to confusion. Algorithmic trading, time series analysis, mathematical modelling etc. For simplicity I will umbrella these terms into Quantitative Analysis.
To make it clearer, quantitative analysis is what is performed in large banks and institutions. Strategies often involve creating mathematical models that take in various inputs about a stock before outputting the value of a stock purchase. For example Black-Scholes formula which estimates pricing of certain options (options can be understood as an abstracted form of buying/selling stock), uses a series combination of integration and partial differentiation and takes in values such as current stock price, strike price of the option, interest rate, portfolio value, and drift rate among many others. The Black-Scholes model has been adapted to creating various other models, and is itself widely employed within hedge-funds and investment banks.
Specifically, Pinnacle Advisory Group's S&P 500 model to predict movement of that index is a dynamic model which takes in 18 factors to forecast the index over a 6-month time frame. The model gives an output of whether to buy, hold, or sell.[Fig 1.3]
There are also many micro-techniques and trends that are used by quants and day traders, these include mean reversion, market timing, trend following, contrarian investing. Financial models that are used in quantitative analysis often have a form of these linear trends integrated into these models where as day-traders often use some of these patterns as their primary signals.
Finally, there exist certain trading tactics that are often used by day traders. Trading applications often include many of these within their graphs for the day-trader to easily be able to visualise their own predictions. [Fig 1.4].

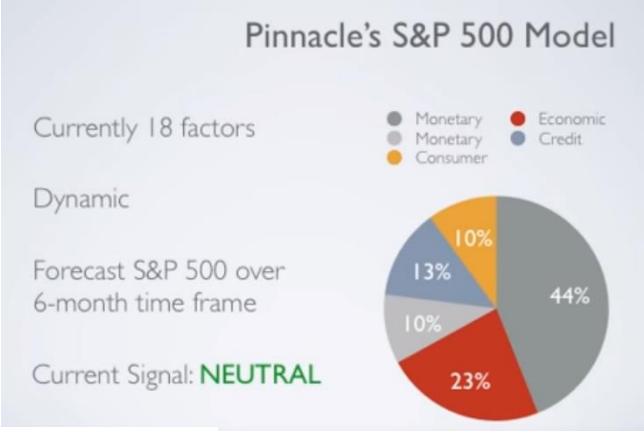


Fig 1.3 [6]

Fig 1.4 [7]

Quantitative Analysts usually take on the role of formally attired bankers and investors who work in an office in a financial capital of the world. To secure these roles individuals must show excellent academic skills and be able to work under high-pressure environments.
On the other hand, it is commonly seen for Day Traders to be in a more relaxed individual environment, not needing a background in finance or maths, often working from home or remotely, as long as they can get access to the internet.

Day trading

Whilst the idea of becoming a quant seems difficult and farfetched for many, the passion to work within the prestige of finance and with the rush of money is not lost on millions of traders around the world.
Brokernotes reported that between 2017 and 2018 there was in increase of 4.3 million people who trade online from 9.6 to 13.9 million. The same report shows us that 58% of all online traders are millennials with younger people increasing their presence in online trading with elder people reducing.
There are various trailing factors that could explain this but a leading factor is the rapid rise of the worth of crypto-currency and the social media coverage of it. This can be shown from the significant increase in "what do females trade" with cryptocurrency increasing from 47.3% to 59.1%. [8]]
The over glorification of traders on social media was covered by Symeon Brown for The Guardian. In this it is shown how certain young self-titled traders receive eyes on their pages by featuring him "buying £250,000 cars, boarding private jets". The individual, Elijah Oyefeso, belong to a group of many hundreds of social media traders who promote an extravagant lifestyle whilst not being clear on the actual wealth received. The article explains how Oyefeso's registered companies did not in fact declare any income.
The purpose of these individuals is to attract unbeknownst individuals to, as the article states, join their team. This usually involves referring the new joiner to start trading on certain trading platforms for which the social media influencers receive a commission on their referrals. [9]
Despite the claims made by social media traders, Brokernotes shows us that the average household income for UK traders is £35,742 in 2018 (almost £3k lower than quoted in 2018).
These figures back up what has been dictated on the report of why many specialists dictate day trading a gamble. Of course, governments have recognised this and binary options (a process by which a user can purchase options and result in gaining a profit or losing his initial purchase input

entirely) which is a service available in many trading platforms and even stock exchanges have been banned in the US and now regulated by the Gambling commission in the UK. [9]
Despite the report having come to the conclusion that day-trading does seem to be quite volatile and not as nuanced as specialist fundamental/quantitative analysis done in banks, there is still profit to be made. Brokernotes report applies a negative connotation to the fact that <1% of traders have an income of about £75,000. The report also fails to outline the manners in which they have collected the data and thereby fail to define what they consider as a "trader".
On page 7 it cannot be established whether they have reported on the "earnings" of the trader to be pure net-profit through trading or just the salary that the individuals lie in (possibly implying that trading profits occur outside of their main income).

Day Trading Techniques

In this section there will be a breakdown of certain Day Trading techniques and how Machine Learning can provide use for the manual tasks.

Fibonacci Retracement

Without delving too deep about Fibonacci theorem it can be said that there are various patterns that occur that follow Fibonacci number ratios. Certain traders believe that this can be applied to finance as well. Percentages that we will consider are 61.8, 38.2, 50, and 23.6. The idea of Fibonacci retracement is that after there is a major price movement, such as between April 09-April-10 in Fig 1.5 we expect a pullback or a rally at certain percentages. These percentages indicate how much the current price has moved back to the initial starting price. Fig 1.5 points at the price maintaining itself at 50% indicating a potential uptrend.
Two issues within separate journals tackle the idea of Fibonacci Retracement within different markets, BRVM (an African stock market) and Forex (a currency trading market). [11][12] Both journal issues discuss how the Fibonacci ratio does indeed have prevalence in the chart graphs despite both set of results being under different time frames. Otake and Fallou's journal draws trends over daily measurements of stock whilst Gaucan's journal draws trends and uses datasets from prices that are mere hours apart.

Fig 1.5 [10]

Both journals conclude with similar sentiments. It can therefor be concluded that Fibonacci ratio can be seen to have a role to play in price charts, however it cannot be used as a unitary measurement signal. That is to say both journals recommend using it in combination with other tools to create a better ruleset. One of which Gaucan recommends is Candlestick.

Candlestick

At the time of writing the report, searching “trading” in google images and observing the top 2 rows of pictures, 9 of the 13 show candlestick. Candlesticks are a very popular way to sum up purchasing results from a period of time. If we were to analyse stock prices in a frequency of once every 2 minutes. Within each 2-minute interval there is potentially hundreds of trades occurring which alter the price of the observed stock. The candlestick tool is a useful way of displaying the major changes that occurred within these 2 minutes.

To explain candlestick, it is best to be used with an image. [Fig 1.6]. Bear(ish) indicated a downward market trend and Bull(ish) indicates an upward market trend. The thick parts of the candle dictate the starting and closing price of each 2-minute interval. The wicks on the top and bottom indicate the maximum and minimum that the stock reached within those 2 minutes. Variances in the candlestick is often used to dictate forthcoming patterns. [App 1.1].

Day-traders by default will have candlesticks showing when analysing their graphs. When observing certain candlesticks, this is a signal for traders to then buy or sell stock.

In Prado et al research paper [14] the effectiveness of candlestick chart analysis is measured. This is being done in the Brazilian stock market. The method of testing the effectiveness of each candlestick pattern is shown in Table 1 of their paper where they apply simple up or down predictions to each pattern. The paper states that unlike the book that they were referring to about Candlesticks in a US market, they could not produce strong results for applying Candlestick method to the Brazilian market. It is however stated that there were certain successes, specifically with Harami-Bullish for certain days. They therefor nullify the usefulness of these patterns and recommend development of the technique for different markets and times.

However, their use of the candlestick patterns may be criticised. Their methodology doesn’t state the context behind how they read in the occurrences of each candlestick pattern and its corresponding price movement. Traders strongly recommend using candlesticks in a greater context, for example if the prices have generally been bullish for long periods of time, do not then use a candlestick indicator such as “Pin” to buy stock presuming the market will become bearish.

The report however does bring into consideration which Stock Market would be best to train and test the model’s data on.

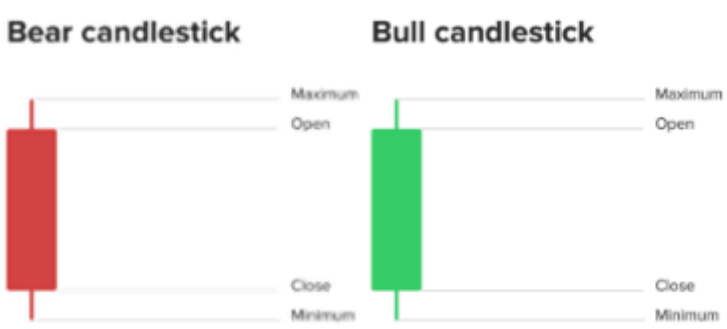


Fig 1.6 [13]

Design Approach

Issues with current solutions

Various AI systems currently implement a hybrid system combining a rule based Expert System with prior fundamental analysis knowledge, combined with a neural network that also takes in quantitative analysis [15].
Current AI solutions, as they are not focused on day-trading, makes decisions that last between days to months. My datasets will need to have x values on a time iteration of possibly seconds to minutes. Datasets that are used in most solutions only access the closing price for each time period. To be able to implement candlestick algorithms I will need data sets that contain not only the closing price for each time period but also the opening price, max price, and min price.
I have identified "https://www.tradingview.com/" to access URL's that contain graphs containing candlesticks, alongside "https://www.viewex.io/" to extract a CSV file with the relevant data.

Proposed solution

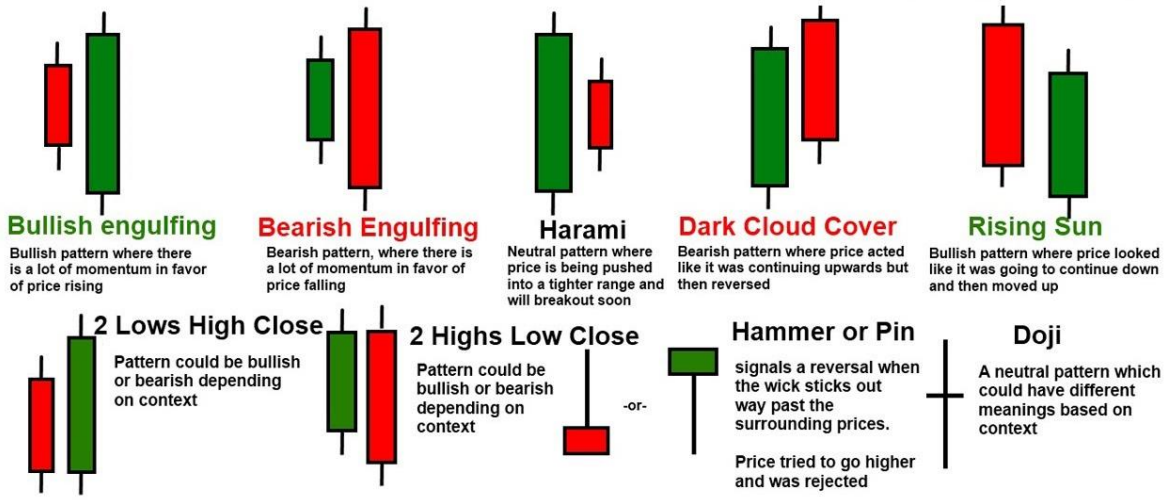
My proposed solution as opposed to the hybrid solutions is to implement a purely algorithmic trader only using time-analysis. However this algorithmic trader will be backed up by a data-mining model. This model would have a spectrum of classes between -1 to 1. The closer the classification comes to 1 the more confidence it will provide the trading bot to buy. This confidence meter pass a certain thresh-hold could be to purchase more stock. The inverse applied to a classification of -1.
This model will also have multiple groups of multiple attributes.

- Group 1 – time interval between data: 5s, 10s, 20, 40s, 80s, 160s, 320s
- Group 2 – candlestick patterns: Bullish engulfing, harami, pin, doji, etc.
- Group 3 – indicator: RSI, MACD, Stochastic
- Group 4 – potential Quantitative Analysis models

Within this model, per iteration of back testing there can only be:
1 attribute from Group 1,
1-All attributes from Group 2,
1 attribute from Group 3,
1-Many attributes form group 4.

This model is what will be used to feed the trading bot with relevant information. Expanding on this, the model will behave differently for when it is tested in different stock markets, and different climates.
As expressed in the literature review, beating the market on a steadily increasing market is difficult, therefor a relatively stable market will be required. FTSE has been floating in the min 7000's for months, therefor UK stock options are possible. However, UK stock options are expensive and possibly out of budget for when it comes to deploying.
Denmark's OMX Copenhagen 20 has been operating at around 1000 for multiple years therefor that is an indication for a cheaper stock market entry whilst being accessible as it is in Europe.

Appendix



App 1.1

Table 1. Patterns to be studied and their hypothesis

Patern		Subsequent patterns (Hypothesis)
Prior Trend	Name	
Bull	Bearish Engulfing	Go down
Bear	Bullish Engulfing	Go up
Bull	Harami -Bearish (Bearish Pregnant Woman)	Go down
Bear	Harami-Bullish (Bullish Pregnant Woman)	Go up
Bull	Evening Star	Go down
Bear	Morning Star	Go up
Bull	Bearish Abandoned Baby	Go down
Bear	Bullish Abandoned Baby	Go up
Bull	Hanging Man	Go down
Bear	Hammer	Go up
Bull	Shooting Star	Go down
Bear	Inverted Hammer	Go up
Bull	Dark Cloud Cover	Go down
Bear	Piercing Pattern	Go up
Bull	One Black Crow	Go down
Bear	One White Soldier	Go up

App 1.2

References

[1] Fama, E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2), p.383.

[2] Finra.org. (2019). *Day-Trading Margin Requirements: Know the Rules | FINRA.org*. [online] Available at: https://www.finra.org/investors/learn-to-invest/advanced-investing/day-trading-margin-requirements-know-rules [Accessed 29 Nov. 2019].

[3] Lindauer, M., Larimore, T. and LeBoeuf, M. (2014). *The Bogleheads' guide to investing*. Hoboken: Wiley.

[4] Investopedia. (2019). *Can Anybody Beat the Market?*. [online] Available at: https://www.investopedia.com/ask/answers/12/beating-the-market.asp [Accessed 29 Nov. 2019].

[5] Cowles, A. (1933). Can Stock Market Forecasters Forecast?. *Econometrica*, 1(3), p.309.

[6] Locatelli, S. (2012). *YouTube*. [online] Youtube.com. Available at: https://www.youtube.com/watch?v=X-GcLmE5YoI [Accessed 29 Nov. 2019].

[7] Stock Members. (2019). *RSI Indicator – I love it!*. [online] Available at: https://www.stockmembers.com/rsi-indicator-i-love-it/ [Accessed 29 Nov. 2019].

[8] Brokernotes (2018). *THE MODERN TRADER*. [online] Available at: https://brokernotes.co/wp-content/uploads/2017/08/BN-research-report_2018-FINAL.pdf [Accessed 29 Nov. 2019].

[9] Brown, S. (2018). *Fake it till you make it: meet the wolves of Instagram*. [online] The Guardian. Available at: https://www.theguardian.com/news/2018/apr/19/wolves-of-instagram-jordan-belmont-social-media-traders [Accessed 29 Nov. 2019].

[10] Fidelity.com. (n.d.). *What Is A Fibonacci Retracement? - Fidelity*. [online] Available at: https://www.fidelity.com/learning-center/trading-investing/technical-analysis/technical-indicator-guide/fibonacci-retracement [Accessed 29 Nov. 2019].

[11] Gaucan, V. (n.d.). How to use Fibonacci retracement to predict forex market. *Journal of Knowledge Management, Economics, and Information Technology*, 2069(5934).

[12] Otake, T. and Fallou, F. (2013). Can we apply Fibonacci retracement in the African market?. *academicJournals*, 7(24).

[13] Xena Exchange Blog. (2019). *Japanese Candlestick Charting Techniques | Xena Exchange Blog*. [online] Available at: https://xena.exchange/blog/japanese-candlesticks-the-history-of-the-eternal-battle-between-the-bulls-and-the-bears/ [Accessed 29 Nov. 2019].

[14] Prado, H., Ferneda, E., Morais, L., Luiz, A. and Matsura, E. (2013). On the Effectiveness of Candlestick Chart Analysis for the Brazilian Stock Market. *Procedia Computer Science*, 22, pp.1136-1145.

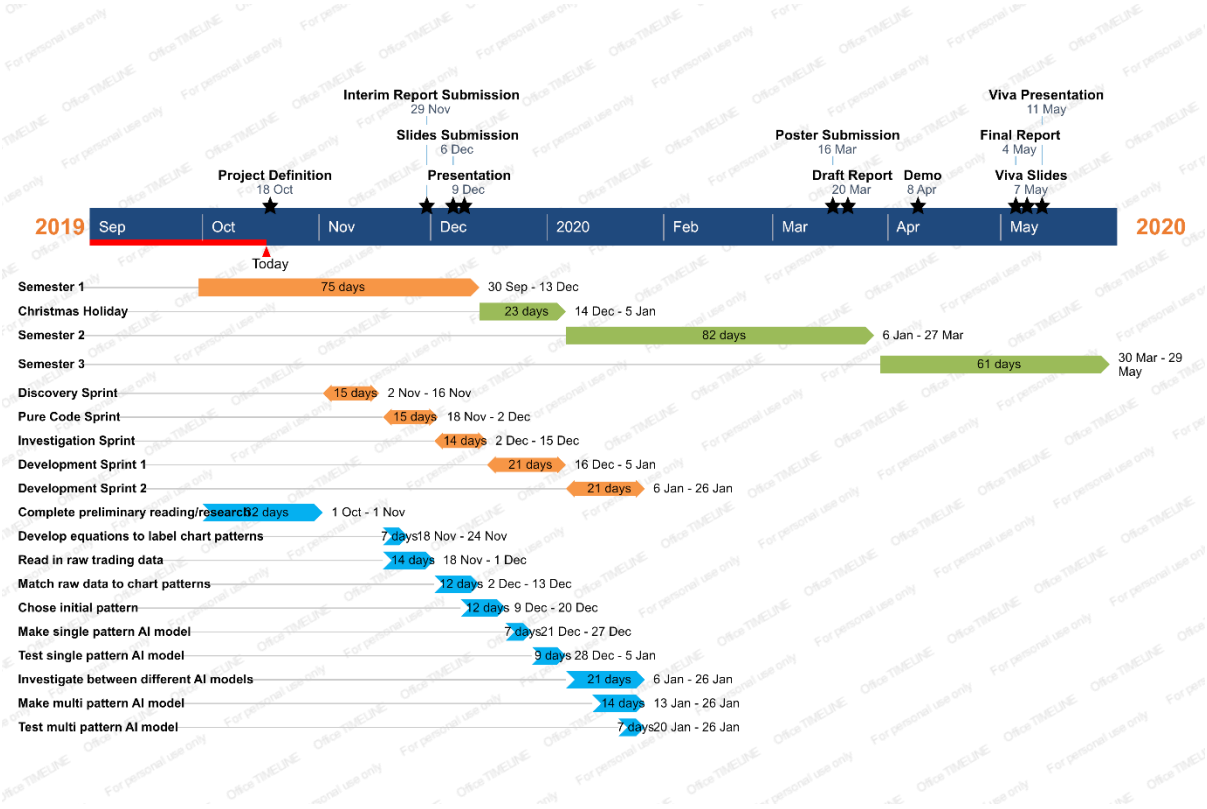
[15] Tsaih, R., Hsu, Y. and Lai, C. (1996). *Designing a hybrid AI system as a forex trading decision support tool - IEEE Conference Publication*. [online] Ieeexplore.ieee.org. Available at: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1562920 [Accessed 29 Nov. 2019].

Risk Management

Risk Description	Impact Description	Likelihood	Impact Rating	Preventative Actions
Budget Limitation	Not being able to purchase a worthy amount of stock to test deployment. Not being able to test deployment on certain companies with higher value stock price.	0.5	9	Location markets that are accessible and cheaper.
Latency	Day Trading requires low ping connections to trading server. Lag can caused delays purchase/selling of stock, potential profit hinderance especially if accessing local stock exchange at busy periods.	0.3	6	UK internet speeds are quite stable. Use wired connection from Home/Uni desktops, avoid laptops.
Market Access	The right market may have certain issues: Lack of accessible dataset, lack of trading platform that operates on a potentially niche/international market, law's that prevent international trading.	0.5	8	A perfect market is required to make a model and train with as much data as possible and rapid access to the country's stock exchange is required. Asian stock exchanges may be inaccessible compared to a European one.
Trading costs	With the limited budget it means that a trading cost of perhaps £1 sinks 1% of the £100 budget. With a budget of perhaps £10000, the £1 trading cost is only 0.001% of the budget.	0.7	7	Choosing the right market. Perhaps raising more capital.
Model not definitive	With the pseudoscience and speculation behind day-trading techniques, it's possible that the model may not produce accurate decisions for the trading bot	0.8	9	Extensive research to find various different chart patterns.

Gant chart

Old:



New: