Ledgers have been the backbone of economics since ancient times. They are a record of contracts and transactions, for the buying and selling of goods. They were originally physically maintained, but now are all digital, making complex global economics viable.

Distributed ledger technologies (DLT) are a new revolution where the ledger is stored on a global decentralised network, which is secure and fast. This network is a decentralised system built on trusted cryptography, not people or a central authority. To access the network, cryptographic keys and signatures are needed and if any changes are made to the DLT, every member in the network is instantly updated. Because of this, there is a full audit trail of information available that can't be changed. There is no master copy that can be lost as all information is shared.

Digital Currency is a term that refers to all types of monetary assets in digital form. It can be both regulated like a countries central bank's currency, or unregulated like virtual or cryptocurrency, which is issued and controlled by the software developers. An example of a digital currency that has benefitted from a DLT is Bitcoin that uses a blockchain as the DLT.

When developing a digital currency, it is vital that people feel comfortable using it and trust that their money will be in safe hands. The benefits of a DLT are that firstly there is no central authority controlling the network and abusing their power. The DLT is open and transparent as any changes to the ledger are shown to anyone in the network. This means that no dodgy transactions can be made by the owners of the digital currency, therefore reducing corruption. Another benefit is that a cryptographic key and signature is needed to make transactions on the network, so users of the digital currency will feel safe knowing that the network is secure from criminals attempting fraudulent spending.

When developing a digital currency, there is usually a lot of overhead involved when using a centralised system. There needs to be back-ups of the centralised server as there would be a single point of failure. As DLTs are a decentralised network there is no single point of failure. If a digital currency used a DLT network, if one node failed it would be simple to recover the network from a neighbouring node.

Users will only want to use digital currencies if they provide some benefit compared to current solutions. If a centralised system is used, transactions often take a long time to be approved. If a digital currency was to adopt a DLT approach, a benefit would be that transactions would be much faster due to the cryptographic trust built in the network that doesn't need to be approved by an authoritative member.

Another benefit of DLT is that there are different types, suitable for any digital currency. These are: public where the network open to anyone on the chain; private where the network is maintained by those who are subscripted to it; and permissioned where the creators determine who may act as transaction validators on their network.

A smart contract is a self-executing contract with the terms of the agreement between buyer and seller being directly written into lines of code. The code and the agreements contained within the smart contract exist across a distributed, decentralized blockchain network. The code controls the execution, and transactions are trackable and irreversible. Using a smart contract allows for the algorithm to be public so that no corruption can happen. Smart contracts are automatically executed once the conditions of the agreement are met. Smart contracts permit trusted transactions and agreements to be carried out among anonymous parties without the need for a central authority, legal system, or external enforcement mechanism.

There are many problems with the user experience of smart contracts and cryptocurrency. Firstly, there is no consistent design of a digital wallet. Across all cryptocurrencies a user must learn a unique way to use each individual digital wallet. Because smart contracts require paying with cryptocurrency, there is a user experience problem trying to use smart contracts.

Following on from this point, smart contracts also require the user to have extensive knowledge of cryptocurrency in general. A user would need to know their public and private keys. The concept of managing keys can be complex and can require a high learning curve. Each user essentially becomes a bank, so needs to understand the transaction costs that go along with every transaction.

Smart contracts are required to use the cryptocurrencies' network, which is often an extremely poor payment network. Transactions are almost always more expensive than centralised banks and the on-chain transactions never have been well suited for in-store purchases. The network is also much more unreliable when compared to PayPal for example.

Finally, as smart contracts sit on an unregularized network, there can be problems with security. Hackers lure uneducated users to send their info such as the private key to take all their money or even steal their identity. Uneducated users can do all kind of mistakes like sending their money to the wrong address, so it disappears. Instead of creating the future of commercial systems, some people are using the system poorly or fraudulently. Smart contract's UX designers would have to find, using the given technology, solutions that would ensure and secure it users.

Solving the poor user experience behind smart contracts would require a lot of work as without careful design a smart contract could be set up incorrectly and money could be lost. A single interface for all cryptocurrencies would need to be implemented, and users would need to be educated on this single interface. The people that are working to solve this kind of problems today are developers, cryptographers and mostly people with a technical background. UX designers are really what is needed to make things simple and clear for the average user.

MODULE CODE:	COMP6212		Student ID:		29466423
Question number:	A1c	This is page number:		3	of my answers

Smart contracts are the code for automatically paying and distributing money across a blockchain network, based on a pre-determined set of rules. If this code is made more complicated than it needs to be, there is more room for error so issues may occur. Smart contracts are at the heart of the network, but not a complete solution meaning that they should not be used if they are not needed. Anything that can be done off the blockchain should be. If issues do occur, money may be lost or sent to the wrong place.

Knowing that a wide range of valid concerns will be attended to by off-chain processes greatly reduces the scope of functions that need to be coded inside a contract. A minimalist contract should only address concerns that cannot be addressed any other way. Minimalism leads to clarity about exactly what the contract needs to prove, and the minimum logic and data required to construct the proof. Minimalism is the first defence against defects that might have non-trivial consequences. The ideal contract is something so simple that there are obviously no defects.

Smart contracts run automatically, which is good for an automation standpoint, but bad for an error control standpoint. One deployed there is no execution control, so the code needs to be highly secure. This is difficult to obtain because blockchain technology is constantly evolving, which means bugs and faults are bound to occur. Because of this future failure needs to be planned for.

As with any well written code, a smart contract should include several testing functions. All code must be well document using clear English. Test cases should be implemented to fully understand how the code is working with different inputs. When writing the methods involved in the contract, constructors must be used with all input arguments being validated. Exceptions must be thrown when the arguments do not comply with the expected input. Finally, code needs to be written to change the state of the contract when issues do occur.

The features within the contract that are extremely important to future proof are interactions, emergency stopping, rate limiting, delayed actions and balance limiting.

Future proofing interactions is important as when interacting with another contract we are handing over control. We need to make sure we have finished with the original contract before handing over to another one. This can be done by using a boolean variable as a flag to represent completion.

Future proofing an emergency stop is important, as sometimes we need to stop a deployed contract if something goes wrong. This can be done by a trusted third party named in the contract aka the Contract Admin, or by pre-determined programmed rules that trigger the stop.

Future proofing rate limiting aims to stop a method in a contract being called so often that it effects the operation of the contract. An example of this would be a DDoS attack. It can also be used to limit the amount of currency that an owner can withdraw, or the number of tokens issued over a period to prevent a rapid drain in funds.

Future proofing delayed actions is needed as we want to slow things down to give contract owners time to react to a malicious event. It is Often combined with a emergency stop to get money out of the contract.

Future proofing balance limiting allows for the amount of currency to be monitored. Doing this means that payments that exceed a predefined maximum can be rejected quickly. This would help users stop accidently spending too much, or to stop a criminal from spending too much at once.

Coloured coins, until recently were a theoretical concept that stemmed from the need to create tokens for transferring assets on the Bitcoin blockchain. The ability to create tokens that are associated with things in the real world, backed by a blockchain network is a useful tool. This is because the blockchain can register and protect the real-world assets. The tokens were name coloured coins and become a certificate of ownership of asset commodity. With these coins it is possible to represent anything in the world, from stocks, commodities, real estate, fiat currencies and even other cryptocurrencies.

Creating a coloured coin involves embedding metadata to identify the asset represented by the coin, onto an extremely small amount of Bitcoin that has no real monitory value. It is marked so it can be tracked as a coloured coin, not just a small amount of Bitcoin. If the owner of the coloured coin wanted to sell parts of the asset, they can now sell parts of the coloured coin and all the transactions are recorded on the blockchain. The only issue is that coloured coins require a special digital wallet to store.

An example of using coloured coin is in a restaurant setting. You can imagine that in the modern age, stamping a card for a rewards program is outdated. If after each meal a customer was given a coloured coin as a reward, it could be stored on the blockchain. If after 10 meals, the customer collected 10 coins, they could be redeemed for a certain amount of gold, which equates in value to a free meal. Everything is recorded on the blockchain, and the process is a lot quicker and doesn't involve a physical stamp card.

Fundamental analysis evaluates a stock by attempting to measure its intrinsic value. Everything from the overall economy to the financial strength and management of the individual company is evaluated. Doing this aims to identify opportunities where the value of a share varies from its current market price. It relies on the assumption that there is some sort of delay in influencing the share prices by these fundamentals. This is how decisions to buy, or sell are made using fundamental analysis.

Technical analysis attempts to identify opportunities to buy or sell a stock by looking at statistical trends, such as movements in a stock's price and volume. It is assumed that any information related to the fundamental analysis is factored into price, meaning that there is no need to pay close attention to it. Technical analysis does not attempt to measure a stock's intrinsic value. Instead, it uses stock data to identify patterns and trends that suggest what a stock will do in the future. Common techniques for this are simple moving averages, support and resistance, trend lines and momentum-based indicators. This is how decisions to buy, or sell are made using technical analysis.

Both analysis methods aim to find the stock's future value and generate buy and sell signals. However, when comparing the differences between fundamental and technical analysis, it shows that they come from opposite schools of thought in the way they predict the future value of a stock and what kind of investments they are used for.

In fundamental analysis, decision making is based on the qualitative and quantitative information available from reports. However, in technical analysis, decision making is based on mathematical market trends of the stock. Following on from this, in fundamental analysis, both past and present data are considered, whereas, in technical analysis, only past numerical data is considered.

In fundamental analysis a stock is examined to identify its current intrinsic value for a long-term investment opportunity. In contrast, technical analysis a stock is evaluated, and its price is forecasted, for a short-term investment.

In fundamental analysis, longer periods are used to evaluate stocks when compared to technical analysis. Because of this, fundamental analysis is used by investors who want to invest in stocks whose value will increase in several years as an investment. On the other hand, technical analysis is used when the trade is for short term only to make short term profits.

A portfolio is a collection of financial investments that can involve stocks, bonds, commodities, cash, and cash equivalents. Stocks and bonds are generally considered the main assets within a portfolio; however, a portfolio can also include many different types of assets that all have a certain level of risk involved. A client's tolerance for risk, investment objectives, and time horizon are all critical factors when assembling and adjusting the investments within a portfolio.

To measure the risk of a single asset in the portfolio you use the variance. If the returns generated by the asset are spread out over the last few years, this asset has a high variance, so the risk is high. The variance is the measure of volatility of the return. The variance is measured as the average squared deviation from the mean. The standard deviation (σ) of the return on the assets is the square root of the variance.

$$\sigma^2 = \frac{\sum_{i-1}^{N}(r_i - \mu)^2}{N}$$

r = return, N = number years, μ is the average

It is important to note that the variance of an asset can be broken down into two parts. The total variance is a combination of systemic risk and non-systematic risk. Systematic risk is something all assets face such as inflation, natural disasters, and political uncertainty. Non-systematic risk is felt by the particular asset or industry based on market movements. We can use different techniques to reduce this risk.

Diversification is a key concept in portfolio management used to reduce the non-systematic risk. This simply states not to put all your eggs in one basket. Diversification tries to reduce risk by selecting investments among differing financial industries, and markets. By doing this, the aim is to maximise returns by investing in different areas that would each react differently to the same event.

In portfolio we are concerned about how different assets vary in relation to each other. To quantify the relationship between different investments within the portfolio, the covariance is calculated. Covariance is a measure of how two assets move together. The value can be between any real number, so the result can be difficult to interpret. What the covariance does tell you is the direction of movement of two assets. If positive the assets show a similar behaviour. If negative the assets show a differing behaviour.

$$\mathrm{Cov}\,(X, Y) = \frac{\sum(X_i - \bar{X})(Y_j - \bar{Y})}{n}$$

Ideally the portfolio would include assets that have a negative covariance, as this shows the best diversification for the lowest risk. If on asset starts to perform badly, we would hope a partner asset would start to perform very well minimising the overall risk.

Even though covariance does not give an insight into the amount of direction between several assets, the correlation can. Correlation is a standard measure of the linear relationship between two assets. The correlation can produce values between the ranges of minus and plus one [-1,+1]. A score of 1 shows a perfect positive correlation between two assets, where the returns of the two assets move together. A score of 0, shows that there is no correlation between the two assets.

A score of -1 shows that there is perfect negative correlation between the two assets, where the returns of the two assets move in opposite directions.

$$\rho(R_i , R_j) = \frac{\text{Cov (Ri , Rj)}}{\sigma R_i * s\ Rj}$$

The relationship between covariance and correlation is given by a standardising factor of the standard deviations from both assets. Covariance and correlation both primarily assess the relationship between variables. Both correlation and covariance are indicators of the relationship between two variables. They indicate whether the variables are positively or negatively related, defining how they move together.

When a portfolio has more than two assets, the variance of the portfolio is given by the average variance and covariance over all assets in the portfolio. Having lower risk assets will reduce the overall risk, at the detriment of reducing the expected return from the investments.

$$\sigma_p^2 = \frac{\sigma^2}{N} + \left(\frac{N-1}{N} * Cov\right)$$

σ_p^2= portfolio variance (we also call this risk)

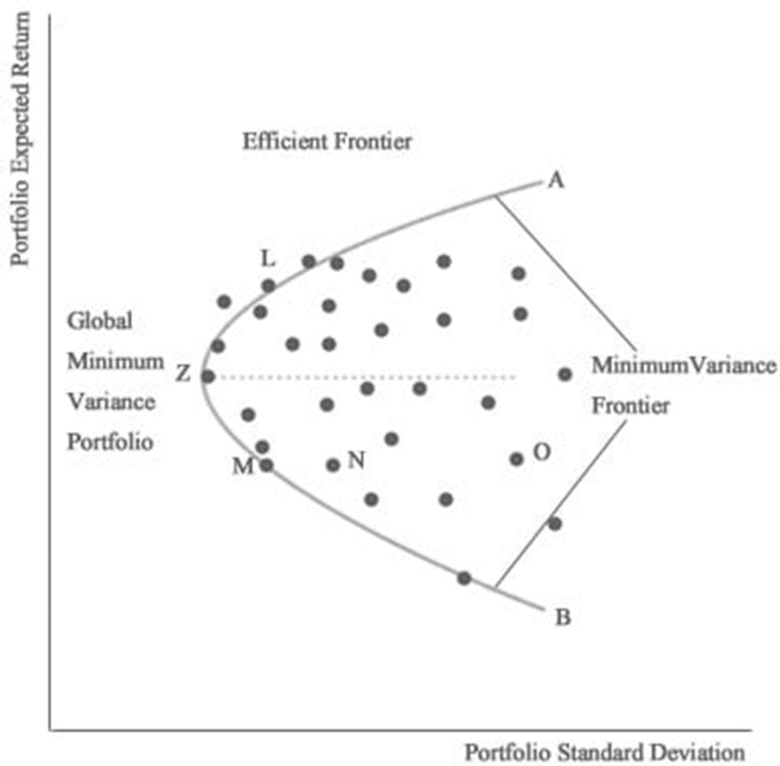
σ^2= average variance

Cov = average covariance

N= number of assets

MODULE CODE:	COMP6212		Student ID:	29466423
Question number:	B1c	This is page number:	8/9	of my answers

A portfolio includes many different types of assets that all have a certain level of risk involved. A client's tolerance for risk, investment objectives, and time horizon are all critical factors when assembling and adjusting the investments within a portfolio. A standard deviation for the portfolio can be calculated along with the portfolio's expected return. This can then be plot on a graph comparing the portfolio's standard deviation against the portfolio's expected return. The figure below shows many different portfolios plotted on one graph.

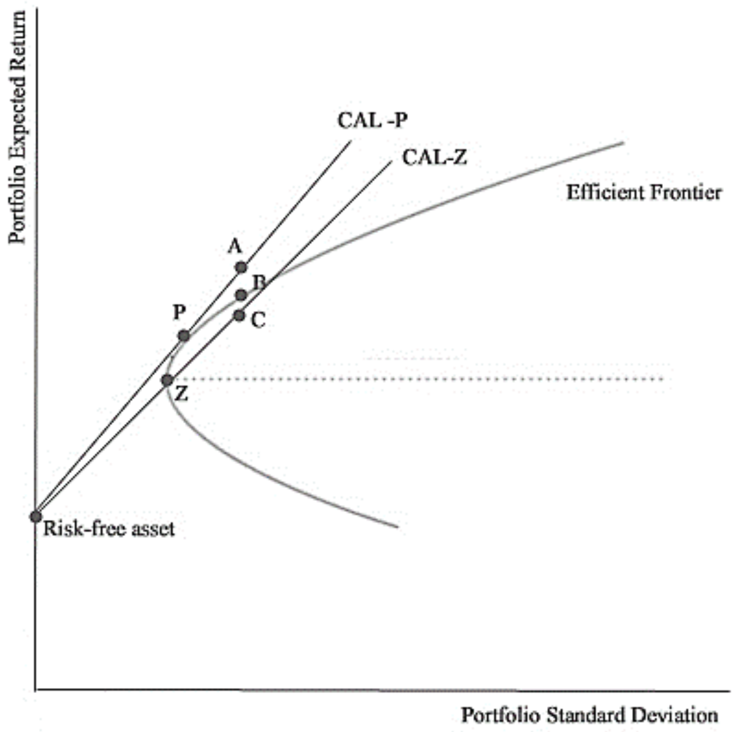


Along with the portfolios, several important lines are plotted on the graph. The first of these is the Minimum Variance Frontier. This line connects all the portfolios that have the minimum risk for a given expected return. A good way to understand what this means is to compare the portfolios M, N and O. All the portfolios have the same expected return; however, portfolio O is much riskier than N, which is also riskier than M. As M lies on the Minimum Variance Frontier, it would be the ideal portfolio for its expected return.

Portfolio Z is the Global Minimum Variance Portfolio. This means that is has the least risk when compared to all other portfolios on the graph. It would not make sense to pick any portfolio with an expected return less than this, as the portfolio would be riskier, with less expected return. A good way to demonstrate this would be to compare portfolios L and M. Both L and M have the same risk, however portfolio L has a much greater expected return so would always be chosen over portfolio M.

The Effective Frontier is the portion of the graph above the global minimum variance. This represents the portfolios that will give the highest returns for a given risk level. A portfolio should be selected in this region, dependent on how much risk is allowed. Conservative investors will aim for a spot on the left side of the efficient (low return, low risk) while aggressive investors will aim for the right side (high return, high risk). If a portfolio falls below the efficient frontier, then it is inefficient. This means that it exposes the investor to too much risk for the specified return or, conversely, provides too low a return for the specified risk.

The effective frontier is useful for deciding which risky portfolio is best, however a complete portfolio is the one that has a mixture of risk free and risky assets. This can be added to the effective frontier with the use of Capital Allocation Lines (CAL). The CAL is a line created on a graph of all possible combinations of risk-free and risky assets. As the figure below shows, many CALs can be plot on the graph.



The slope of the CAL is known as the reward-to-variability ratio. The steeper the slop the better the portfolio, that is the higher the return with the lower risk. The graph displays the return investors might possibly earn by assuming a certain level of risk with their investment. This means that investors receive a higher expected return in exchange for taking on more risk

A Kalman Filter is a technique used to filter out noise from data. It is a minimizing algorithm which updates the state estimate when information arrives, which enables for processing the information in real-time. Kalman Filters are widely applicable to applications that make linearity assumptions; however, they can be extended to non-linear applications using the Extended Kalman Filter. It infers that at every step the posterior density is Gaussian and a parametrized covariance.

Finance data, especially the price of a stock is said to be extremely noisy. Statistical methods try to cope with this; however, the noise would make many other methods of stock price prediction obsolete. The Kalman Filter can be used a data pre-processing step, to remove as much noise from the data as possible.

The Kalman Filter can also be used in a real-time feedback system of models that aim to predict the stock market price. When the model makes a prediction, the filter can take a portion of predicted data, along with the associated errors and feed it back to the input of the system to improve the model's accuracy.

A Kalman filter can also be used for to estimate the parameters of a model when the model relies on non-observable data. This can be extremely useful for statistical models that aim to predict the stock price, where the underlying distribution needs to be extracted.

The Kalman filter is also very fast so can analyse large amounts of data without too much delay to the system. This would be extremely useful when aiding in predicting stock prices, as it is an extremely fast pasted environment.

Machine learning is a branch of computer science, and a subset of Artificial Intelligence that deals with the science of programming computers so that they can learn from data. A machine learning model will aim to generate a new insight from the data and continuously improve, given more data and tuning. Machine learning is split into a model and algorithm aspect.

Machine learning techniques fall into several categories, supervised and unsupervised. Supervised learning involves training the model on labelled data and reducing the error of new decisions the model makes. It is mainly used for classification and prediction tasks. Unsupervised learning involves placing unlabelled data into different groups based on some distance or similarity metric. It is mainly used for clustering tasks.

Due to the affordability of high-power computing, many machine learning methods can be used in the finance industry, especially for stock price prediction. The price of a stock can be influenced by a large variety of complex independent features, so machine learning can be used as a tool to predict the price of a stock without human emotion and its risk.

When just considering the time series elements of a stock, supervised regression techniques such as SVM Regression, Polynomial Regression with regularised terms and Bayesian Inference can be used to predict the future value of a stock by fitting a trend line to the data. The models can be trained over several sliding windows to capture trend, cycle, seasonality, and irregularity of the stock price.

A stock price is also affected by other factors such as the current state of the company; whether the market is bullish or bearish; how similar stocks are doing; and potential random features. Using supervised classification techniques such as SVM Classification, Random Forrest and Neural Networks, these factors can be used to help classify if a stock will increase or decrease in value or pass a set boundary. This can aid the regression techniques using an ensemble of methods.

The stock price is also heavily influenced by the people who own and trade the stocks. Using sentiment analysis, the feelings of these people could potentially be captured. This could help predict the movement of the stock prices.

Although most stock price predictions involve supervised techniques, unsupervised techniques can also be of use. Using techniques such as PCA, the very high dimensional feature space of the stock market can be reduced into a lower dimensional space. This will lower computational power and allow for a less complex system, helping prevent over-fitting. Matrix decomposition and factorisation techniques can also be used to fill in missing data should it arise.

To predict stock market prices, both statistical and machine learning methods aim to predict the same thing, the future price of a stock. However, they both follow different processes which work in different situations. To be put simply; statistics draws population inferences from a sample, and machine learning finds generalisable predictive patterns.

Statistical methods focus on inference, which aims to extract the underlying probability distribution of the input training dataset. The model parameters that are found from this inferred distribution allows a quantitative measure of confidence that a discovered relationship describes a true effect that is unlikely to result from noise.

Machine learning concentrates on using general-purpose learning algorithms to find patterns in the input training dataset. The model makes minimal assumptions about the data-generating systems. They often are effective even when the data is gathered without a carefully controlled experimental design and in the presence of complicated nonlinear interactions.

When predicting the price of a stock, it may be useful to understand why a model made a certain correct or incorrect decision. If a statistical method is used the effects of predictor variables and distributional parameters about the outcome variable are known. We can understand what caused the model to be successful or to fail. However, if machine learning is used it usually does not attempt to isolate the effect of any single variable. It is considered a black box model where its decisions are not well known. We cannot understand what caused the model to be successful or to fail.

Often the stock market is a fast-moving environment, where prices need to be calculated quickly and accurately. For a machine learning model, it can be used on a dataset with little pre-processing. This means time is not lost in this step. However, most machine learning models need certain hyper parameters to be tuned based on the specific situation which slows down the time it takes to implement. In contrast the statistic model does not require tuning, so no time is lost there. However, statistical models often require a time-consuming data reduction step to understand the underlying distribution to increase the model's accuracy. Both models have their problems when it comes to implementation time, so careful consideration needs to be taken when using them in a fast-paced stock price prediction environment.

The data from the stock market can have a varying signal to noise ratio which can affect the accuracy of a model's prediction. Due to statistical methods inferring the underlying probability distribution, of the training data, it can account for the variance within the test data. When the signal to noise ratio is small a statistical method will be more accurate in predicting a stock price. When the signal to noise ratio is large, a machine learning method will have a much better accuracy as is does not need to model the underlying randomness of the stock price.

Depending on how long a stock has been public, there may be a difference in the amount of data available for a certain stock. This difference in the amount of available data will affect which technique should be chosen. A statistical model will work well with a small dataset, typically 20 events per candidate predictor. In contrast a machine learning model usually requires a much larger dataset, which typically requires 200 events per candidate predictor. Thus, machine learning can sometimes create a demand for "big data" when small-moderate sized datasets will do. As most stocks will have enough data points over a period, machine learning is the best option. However, when looking at very short-term data or a new stock, a statistical method will be more appropriate.

When the stock market price is predicted, we may simply be interested in just the price or want to know the uncertainty behind the prediction. Machine learning methods will simply give the predicted value, so useful if not very interested in estimating uncertainty in forecasts. In contrast statistical methods will give the predicted mean value, along with the uncertainty of the forecast.

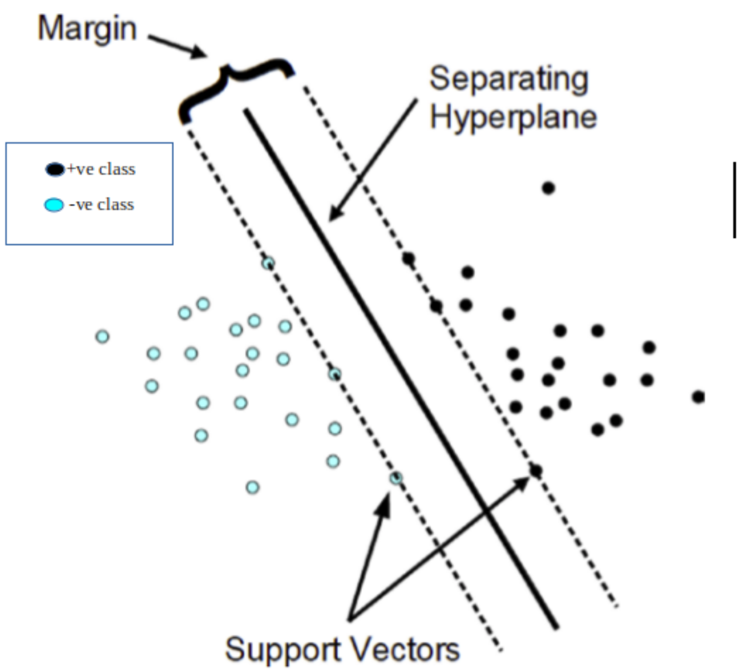
Both machine learning and statistical methods have their benefits and faults, so combing the best aspects into a hybrid method will be the future. For example, the features for a machine learning model could generated by a statistical model, to get a high accuracy and to understand which features make the model work well.

A Support Vector Machine (SVM) is a supervised machine learning algorithm which can be used for both classification and regression tasks. SVMs work by finding a hyperplane and support vectors that either separate classes well or fit a series of datapoints well. As SVMs use a linear hyperplane and margins, it is often useful to transform the data into a high-dimensional space using a kernel transformation, so the data can be linearly separated or best fit to. Slack variables can be introduced to stop the SVM from over/under-fitting with noisy datasets filled with outliers and overlapping classes. These slack variables can be tuned using a hyper-parameter when training the model.

A SVM Classifier is implemented by finding the optimal linear separating hyperplane, that maximises the distance between the margins. This is done to reduce the misclassification of new data. The data points at the margins are used to create support vectors. Only these data points defining the support vector are important and so we can ignore other training data points.

The problem with this method is that hard margins do not work well when the data is noisy; there are outlier datapoints; and when the classes overlap. To combat this, soft margins are introduced where vectors allowed to violate the margin, but the number of violations are minimized. There is a trade-off between how accurate the model is and how well the model works on unseen data. This is known as the bias-variance trade-off. Tuning the parameters that control the slack variable that define the soft margin combats this issue.

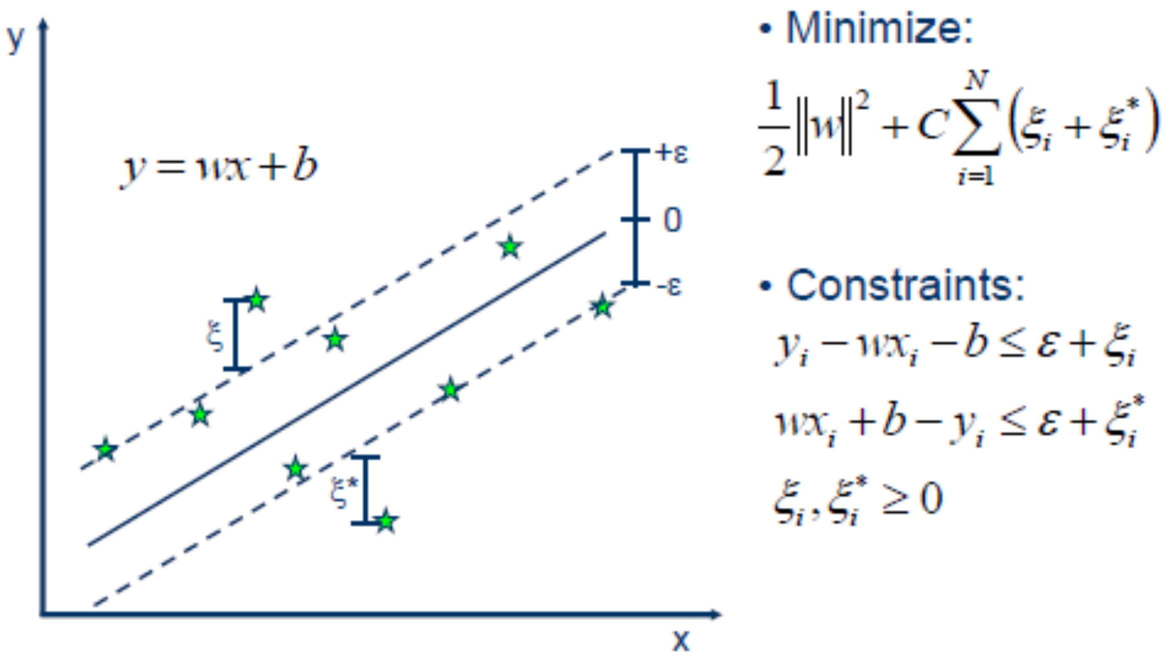
The SVM classifier can be used to help predict stock prices by classifying whether a stock will increase or decrease in price, based on factors such as the current state of the company; whether the market is bullish or bearish; how similar stocks are doing; and potential random features. It could also be used to classify if the stock prices go below a certain threshold or not.



A SVM Regressor is implemented by trying to find a dividing plane so that all points lie within a margin, the exact opposite of what the classification task was trying to do. Compared to normal regression SVR gives us the flexibility to define how much error is acceptable in the model and will find an appropriate hyperplane in higher dimensions to fit the data. The aim is to minimize error, individualising the hyperplane which maximizes the margin.

Much like the classification model, having a hard margin will determent the model. Slack variables will need to be introduced to allow for data still outside the margins to be accounted for. This is because there is the possibility of errors that are larger than the set error. Doing this will help the regression model generalise well and prevent any over/under-fitting.

Much like regular regression, the SVM regressor can be used to predict the future value of a stock by fitting a trend line to the data. The SVM allows for a maximum error to be defined, which is extremely useful for predicting a stock's price. As stocks are extremely variable having the improved performance from the SVM regressor is very useful in predicting its future value.



An exogenous variable is any independent variable whose value is not determined by the model. An exogenous variable is not affected by the model, but still may determine the value of one or more dependent variables. In other words, one variable within the formula does not dictate or directly correlate to a change in another. Exogenous variables are fixed when they enter the model; are taken as a given in the model; influence endogenous variables in the model; are not determined by the model; and are not explained by the model.

Examples of exogenous variables for stock price prediction are GDP (Market Prices), Unemployment, Interest Rate, Imports Goods & Services, Exports Oil Imports, Gross National Income, M1 Money Supply, Productivity, British Pounds/ US Dollar, Contribution to CPI, Oil price, Oil Invest, Government Gross Reserve. These features were selected using prior knowledge from macro-economic literature.

Whilst none of the features appear to have an impact when using machine learning to predict stock prices, it turns out that these exogenous variables can explain the variance in the predicted stock prices. These variables were able to help understand the black box machine learning model. This shows that by using prior knowledge of a system, exogenous variables can be selected to both improve the stock price prediction and help understand the black-box model.