2017-9

Can Deep Learning Techniques Improve the Risk Adjusted Returns from Enhanced Indexing Investment Strategies

Anthony Grace

Technological University Dublin

Can Deep Learning Techniques Improve the Risk Adjusted Returns from Enhanced Indexing Investment Strategies

Anthony T. Grace

D13123578

A dissertation submitted in partial fulfilment of the requirements of
Dublin Institute of Technology for the degree of
M.Sc. in Computing (Data Analytics)

2017

I certify that this dissertation which I now submit for examination for the award of MSc in Computing (Data Analytics), is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

This dissertation was prepared according to the regulations for postgraduate study of the Dublin Institute of Technology and has not been submitted in whole or part for an award in any other Institute or University.

The work reported on in this dissertation conforms to the principles and requirements of the Institute's guidelines for ethics in research.

Signed: _________________________________

Date: **29 June 2017**

ABSTRACT

Deep learning techniques have been widely applied in the field of stock market prediction particularly with respect to the implementation of active trading strategies. However, the area of portfolio management and passive portfolio management in particular has been much less well served by research to date. This research project conducts an investigation into the science underlying the implementation of portfolio management strategies in practice focusing on enhanced indexing strategies. Enhanced indexing is a passive management approach which introduces an element of active management with the aim of achieving a level of active return through small adjustments to the portfolio weights. It then proceeds to investigate current applications of deep learning techniques in the field of financial market predictions and also in the specific area of portfolio management. A series of successively deeper neural network models were then developed and assessed in terms of their ability to accurately predict whether a sample of stocks would either outperform or underperform the selected benchmark index. The predictions generated by these models were then used to guide the adjustment of portfolio weightings to implement and forward test an enhanced indexing strategy on a hypothetical stock portfolio.

Key words: deep learning, portfolio management, enhanced indexing, neural networks, portfolio optimisation

ACKNOWLEDGEMENTS

Firstly, I would like to thank Almighty God for granting me the strength and perseverance to complete this research project.

I would like to thank my supervisor, Brian Leahy, for his expert support and guidance and for keeping me on track throughout this project.

I would also like to thank all of the lecturers who have thought me so much throughout the course of this MSc and without whose support this task would have been impossible.

Finally, I would like to thank my family and friends for their motivational support throughout.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

1. INTRODUCTION

Overview of Project Area

The scale of the publicly listed securities market is vast with the World Bank estimating the total market capitalisation of domestically listed stocks at US$ 61.8[1] trillion or approximately 96.8[2] per cent of global GDP in 2015. The World bank figures also show that the United states accounted for US$ 25.1 trillion[3] of this representing approximately 139.0[4] per cent of the US GDP. In addition to the colossal size of the public markets in capitalisation terms the complexity of the investment allocation decisions facing market participants is compounded by the sheer number of listed companies to select from with the World Bank numbers showing a total of 43,539[5] publicly listed companies in 2015 of which 4,381[6] were listed in the United States.

The requirement to generate sufficient returns to meet an individual's financial needs throughout their working life and into retirement combined with the obvious potential of public stock markets to contribute towards the achievement of these goals has contributed to a significant level of interest in stock market investment. However, the aforementioned scale of global public markets presents significant challenges to the individual investor who does not necessarily possess the financial skills necessary to make sound investment decisions. In response to this problem a sophisticated investment management industry has developed to provide investors with advice and access to the markets. The level of assets under management (AUM) across asset classes and not just restricted to public equity markets was estimated by Boston Consulting Group to have been US$ 71.4 trillion[7] in 2015.

---

[1] Available online at: http://data.worldbank.org/indicator/CM.MKT.LCAP.CD

[2] Available online at: http://data.worldbank.org/indicator/CM.MKT.LCAP.GD.ZS

[3] Available online at: http://data.worldbank.org/indicator/CM.MKT.LCAP.CD?locations=US

[4] Available online at: http://data.worldbank.org/indicator/CM.MKT.LCAP.GD.ZS?locations=US

[5] Available online at: http://data.worldbank.org/indicator/CM.MKT.LDOM.NO

[6] Available online at: http://data.worldbank.org/indicator/CM.MKT.LDOM.NO?locations=US

[7] Available online at: https://www.bcgperspectives.com/content/articles/financial-institutions-global-asset-management-2016-doubling-down-on-data/?chapter=2#chapter2

Investment managers construct a portfolio of stocks on behalf of their clients in accordance with an investment policy. In addition, the investment management industry offer a range of investment vehicles in which investors can acquire units including Mutual Funds, Hedge Funds and Exchange Traded Funds (ETFs). In the context of stock market investment, the strategies adopted by investment managers constructing portfolios on behalf of individual clients or on behalf of a managed fund can be broadly categorised as either active or passive strategies. Active strategies seek to outperform the market through expert stock selection. That is by purchasing stocks which they predict will offer superior returns than the general market and potentially by shorting stocks which they expect to underperform. Passive strategies in contrast seek to match the return earned by a benchmark index and have as their investment management objective the minimisation of deviations from index performance. A variation of the passive strategy referred to as Enhanced Indexing seeks to achieve improvements in risk adjusted returns over that generated by the index by making small adjustments to the weights in which stocks are held in the investment portfolio compared to their weight in the index.

The large scale of global stock markets combined with large daily trading volumes and the attention of legions of analysts leads to the production of large volumes a data. The availability of this data and the potential for significant financial gain has made stock market investment decision making an area in which deep learning techniques have been widely researched and implemented in practice.

1.1 Background

As further discussed in the literature review, the greater part of academic research in the context of the application of deep learning techniques to stock market investment has been in the field of stock selection and stock trading more generally. This has led to the development of quantitative trading strategies and the launch of quantitative funds by hedge fund managers such as Renaissance Technologies and D.E. Shaw which made strong returns for 2015 (Vardi, 2016).

However, as evidenced from the review of the literature the potential portfolio management applications of deep learning techniques has been less widely researched. While active portfolio management strategies which rely on stock selection to achieve their investment strategies are served to an extent by the research towards predicting stock returns there is limited material of applicability in the passive management context.

Passive management is of growing importance with the share of US Mutual Funds subject to passive management increasing from a quarter to a third in three years (Marriage, 2016). The increasing size of the passive management industry places additional importance on increasing the level of research into the application of deep learning techniques to this segment of the market.

Due to the presence of some similarities to active management strategies and the existence of extensive research in the area of the prediction of future stock prices and stock returns the area of Enhanced Indexing has been selected as a field which offers promise in terms of the potential for developing improvements in techniques for delivering superior risk adjusted returns.

1.2 Research Project

The research project will investigate whether deep learning techniques can improve the risk adjusted returns from enhanced indexing strategies by building deep neural networks to predict whether the selected sample of stocks will outperform the index and then using this information to determine the adjustments to be made to index weights.

1.3 Research Objectives

The research objective can be framed in terms of the following hypothesis:

H_0 Deep learning techniques cannot improve the risk adjusted returns from enhanced indexing strategies

H_1 Deep learning techniques can improve the risk adjusted returns from enhanced indexing strategies

The overriding objective of the study is to design and execute an experiment which will seek to reject the null hypothesis. In order to achieve the overriding objective a number of subsidiary objectives have been established.

Firstly, the study will investigate and document the current state of the art in terms of both the relevant aspects of portfolio management and the current applications of deep learning techniques to stock market decision making. Having established a theoretically sound testing period the study will then seek to perform an exploratory investigation of the raw data and select a suitable benchmark index to form the basis for the experiment.

A further objective of the study will be to select a suitable sample of index components to re-weight to test the hypothesis and to develop a list of candidate features to predict the performance of these stocks relative to the index.

Finally, the study will seek to evaluate and refine the models developed to thoroughly test the hypothesis.

1.4 Research Methodologies

The research methodology will consist initially of secondary research into the current applications of deep learning techniques in the field of investment research followed by in depth primary research to test the research hypothesis.

The primary research shall consist of the construction of deep neural network models the results of which will be utilised to make decisions as to the appropriate re-weightings required to deliver enhanced risk adjusted returns.

Conducting the experiment requires both the selection of a suitable benchmark index and a sample of index components to re-weight. A set of criteria shall be established to facilitate the selection of both the index and the index components.

The selection criteria will seek to ensure that the data is adequate to test the hypothesis and is suitably homogenous to preclude the results from being impacted by factors which do not need to be considered to effectively test the hypothesis and which as such are outside the scope of this study.

As the objective of the research is to assess the impact of deep learning the initial models developed will contain a single hidden layer with the number of hidden layers then being increased to assess if an improvement in performance is achieved.

The model development will utilise a two-stage approach. In the first-stage deep neural network models will be developed to predict whether the sample stocks selected will outperform or underperform the index. The outputs from the first-stage will provide the inputs to the second-stage. The second stage will adjust the weights of the selected stocks increasing the weights on stocks which are predicted to outperform the index and reducing the weights of those stocks which are predicted to underperform.

The deep neural network models developed in stage-one will be assessed in terms of their accuracy rate. The model outputs from phase one will be either one of two possible outputs; outperform or underperform. The re-weighted index developed in stage-two will be assessed in terms of its risk adjusted returns. The risk adjusted returns will be measured as the returns per unit of risk (i.e. returns divided by standard deviation). These results will then be compared to performance of the index and if the results for the model exceed those of the index the null hypothesis will be rejected.

1.5 Scope and Limitations

The hypothesis shall be tested by reference to a single stock market benchmark. In order to ensure that the index is suitable and that any contributions generated can be extrapolated to other similar indices strict selection criteria have been established to ensure that the specific hypothesis can be thoroughly tested and that the test results will not be influenced by factors outside the scope of the study. Secondly, as it is not feasible or indeed desirable to test the hypothesis using all index components a sample of stocks have been selected from the index. Again, strict selection criteria have been applied in order to allow extrapolation to the components of the index more generally.

The potential candidate features for predicting outperformance or underperformance have been restricted to raw stock data and measures derived therefrom. Macro-economic data has been excluded from the scope of the study as stocks from different industry sectors exhibit differing sensitivities to such factors and as such pose the risk of distorting the results.

In addition, the study period used to provide the data to train and then validate and test the models developed is limited to a five-year period. This period was selected to cover a period in which a broadly similar macro-economic environment prevailed. While the period could be extended, a significant extension would pose the risk of introducing external influences on the stock movements which do not persist through time.

1.6 Document Outline

The remining chapters of this dissertation will be organised in the manner described below:

Chapter 2 – Literature Review: This chapter will provide a comprehensive discussion of the current state of the art. The review will cover both the relevant aspects of portfolio management and the stock selection process. It will then proceed to discuss deep learning techniques more generally and then move towards an assessment of the existing research in the area of the application of deep learning techniques in securities selection and quantitative trading. Next consideration will be given to the approaches to implementing neural networks considered in the literature. Finally, it will provide an evaluation of the current state of the art and identify gaps in the literature.

Chapter 3 – Design and Methodology: This chapter will discuss the data requirements for the conduct of the experiment along with the sources which will be utilised to obtain this data. It will next provide a detailed analysis of the selection criteria for identifying a suitable benchmark index to test the hypothesis and specify the index selected. It will then provide a similarly detailed analysis of the selection criteria for selecting the index components to utilise for the experiment and specify the sample selected. The chapter will then provide details of candidate features and of the model to be implemented. Finally, the criteria to be used to evaluate the model will be detailed.

Chapter 4 – Implementation and Results: The chapter will provide the detailed implementation and refinement steps of both the single-layer and multilayer models. It will also provide the detailed results of the experiments conducted.

Chapter 5 – Evaluation and Analysis: This chapter will apply the evaluation criteria elaborated in Chapter 3 to the results generated in Chapter 4. It will state whether the null hypothesis can be rejected.

Chapter 6 – Conclusions: This chapter provides an overview of the results of the study detailing the contributions made to the existing body of knowledge and will provide an outline of additional work to be performed in the future.

2.0 LITERATURE REVIEW

2.1 Introduction

This chapter will undertake a review of the current state-of-the art employing a multi-pronged approach. An investigation will be undertaken into the application of deep learning techniques to a specific aspect of portfolio management. Portfolio management and deep learning are both extremely broad and important fields in their own right and as such have attracted a considerable degree of academic interest.

This study will first undertake an analysis of the concept of portfolio management and will then proceed to perform a more in-depth investigation into the mechanisms used to implement a portfolio management strategy in practice. While the broad portfolio management theories discussed in the academic literature relate to the *asset allocation* decision this study will narrow the focus onto the investment decisions to be made in the context of the management of stock portfolios.

As securities selection is an integral part of the management of a stock portfolio this area warrants particular attention and will be addressed next. This section will discuss the concepts of factor investing and the categorisation of stocks into standardised categories based on their attributes. These factors and attributes will later form the basis for the model feature selection decisions.

The chapter will then proceed to address the mathematical aspects of the investigation. This section will commence with a discussion of the existing research on the application of deep learning techniques to the prediction of stock market performance and to the identification of factors influencing stock prices. It will then proceed to discuss some of the current applications of deep learning techniques to the implementation of portfolio management strategies.

The chapter will then discuss some pertinent technical aspects of building neural network models and identify the measures employed to evaluate the results of these models.

In the final section of the chapter a critical evaluation of the current state-of the art will be undertaken identifying the gaps which currently exist.

2.2 Portfolio management approaches

2.2.1 Modern Portfolio Theory

In his seminal paper ‘Portfolio Selection’ Markowitz (1952) described an approach to portfolio construction which forms the foundation of what has now become known as Modern Portfolio Theory (MPT). The rational investor seeks to maximise returns while minimising risk. As this is a forward-looking assessment returns in this context are expected future returns. Risk in this context means the possibility that the expected returns will not be achieved and in mathematical terms is quantified as *variance*. Both expected returns and variance are computed by reference to historical prices. Markowitz (1952) demonstrates that risk-reward assessments should not be made on a security by security basis and that instead a diversified portfolio should be created considering the correlations between the securities. The author demonstrated that by adding securities which are less than perfectly correlated to the portfolio the variance (i.e. risk) of the portfolio can be decreased and introduced the concept of a combination of securities which has become known as an efficient portfolio. A portfolio is assessed as being efficient if no further reduction in risk can be achieved without causing a reduction in return. As such an efficient portfolio is one which maximises return for a given level of risk. The methodology utilised by Markowitz (1952) to derive the sets of efficient portfolios has become known as *mean-variance analysis*.

The second major component of Modern Portfolio Theory is what has become known as the Capital Asset Pricing Model (CAPM) proposed by Sharpe (1964) in his influential paper ‘Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk’. In his paper the author proposes that the expected return on a security is comprised of a time value of money component equivalent to the return on a risk-free asset plus a risk premium.

In practical application, the return on the risk-free asset is taken as the return on the appropriate government debt. Sharpe (1964) further considers the risk premium in terms of the efficient portfolio concept. Arising from this the author presents the concepts of *systematic risk* and *unsystematic risk*. In this context, Sharpe (1964) ascribes the term *systematic risk* to the portion of a risky securities total risk which is attributable to its correlations with the efficient portfolio with the remainder of its risk being termed *unsystematic risk*. Having established that all forms of risk other than *systematic risk* can be eliminated through diversification Sharpe (1964) argues that the only aspect of a securities risk requiring consideration is what he describes as its responsiveness to economic events. This measure of responsiveness has become known as the securities *Beta* and will be considered further in the model development phase of this study.

As the above analysis is quite academic the concepts proposed Sharpe (1964) will now be explained in more pragmatic terms without direct synopsis of the literature. In summary Sharpe (1964) considers portfolios which may consist of risk-free assets, risky assets or some combination of the two. The risk-free asset is generally deemed to be the government debt and the risky assets are stocks. An investor who invests in a broad market index such as the S&P 500 is fully diversified. An individual stock is influenced by broad market trends and by factors specific to that stock. The effect of stock specific factors can be reduced by diversification (i.e. holding a portfolio comprised of many stocks rather than just a few). Since a stocks return is determined by broad market factors and factors specific to this stocks return can be assessed in terms of a single measure being its sensitivity to the market (i.e. 'Beta').

2.2.2 Portfolio management approaches

Portfolio management approaches are categorised as being either passive or active or a hybrid of the two. Passive management approaches seek to replicate the return on a benchmark index whereas active approaches attempt to outperform the market through stock selection. The excess return sought through active management is referred to as *alpha*.

The relative merits of these portfolio management approaches is an intensely contested topic with numerous writers and researchers making contributions to the debate. Poppick (2016) notes the generally lower costs and superior performance as grounds for opting for passive rather than active management and notes research produced by

Morningstar which illustrated that a mere 12 per cent of US large-cap active managers outperformed their passive management counterparts over the preceding ten years. The writer however refers to further commentary from Morningstar supporting the view that active management can perform better in certain areas such as mid-cap and less efficient foreign markets but that even then this performance may be offset by the extra fees charged. Earlier research conducted by Frino and Gallagher (2001) on S&P 500 index funds also confirms the hypothesis that passive funds provide superior performance compared to the net of costs returns from actively managed funds.

Enhanced indexing passively tracks a benchmark index but adds a small element of active management to achieve an outperformance of the index. Wu *et al.* (2007) describe enhanced indexing strategies as an approach employing risk-management strategies to achieve outperformance over a closely tracked index. The authors highlight that the traditional approach to the implementation of this strategy involves subjective decision making and describe an alternative goal programming based approach which seeks to replace this subjectivity with an objective optimisation function in which the conflicting objectives inherent in this process are collectively optimised.

The authors note the observation that the portfolio optimisation under enhanced indexing in effect constitutes the optimisation of the *information ratio* which quantifies the active return (i.e. *alpha)* relative to the *tracking error*. They further note a contrast between passive management strategies and enhanced indexing strategies in that optimisation under passive management requires only that consideration be given to the minimisation of *tracking error* whereas enhanced indexing requires the optimisation to be constrained by both the requirement to minimise *tracking error* and to maximise *alpha* thereby introducing the need for trade-offs.

Wu *et al.* (2007) refer to research demonstrating that determining an appropriate level for *tracking error* results in the optimisation of both *alpha* and the *information ratio.* They note that the best level for tracking error to achieve this optimisation has been shown by various researchers to be in the range of 1.75 per cent to 4.00 per cent. The authors proceed to demonstrate that their application of the goal programming approach in the context of the Taiwan Stock Exchange achieves improved returns and lower costs.

The implementation of passive index tracking and enhanced index tracking is further considered by Dose and Cincotti (2005). Similar to Wu *et al.* (2007) the authors frame the optimisation problem in terms of the trade-offs between the minimisation of *tracking error* and the maximisation of *alpha.* The authors consider the optimisation in the context of constraints imposed on both the number of stocks to be included and the portion of funds which may be assigned to these individual stocks.

The approach taken by Dose and Cincotti (2005) differs in scope to that being investigated in this research project in that they additionally seek to replicate the return on the index from a subset of its components whereas as elaborated in Chapter 3 and Chapter 4 this research project is focused solely on the enhancement to the returns to the index rather than on the index itself. Nevertheless, the insights provided by Dose and Cincotti (2005) are of relevance in the context of the enhancements to the index and are considered in this context and in the context of future development in this study in a broader scope. The authors utilise time-series clustering analysis to group the stocks into homogenous groups and then test the results using a stochastic optimisation approach. The authors concluded that the use of the clustering approach for stock selection offers improvements over a random selection approach, particularly in an enhanced indexing context. They attribute this improvement to the noise reduction effects of clustering.

2.2.3 Portfolio risk management measures

The comparison of the performance of different investment strategies requires that appropriate risk measures be utilised to accurately measure the expected returns in the context of the expected risks.

The goal of the rational investor is the maximisation of the returns from their investment portfolio. However, future returns are uncertain and can only be estimated probabilistically. The probability weighted estimate of future returns is referred to as *expected returns.* Risk in a portfolio management context is concerned with the potential for deviation in the actual returns from expected returns.

Treynor and Black (1973) introduced what was to become known as the *Information Ratio.* This ratio expresses the *excess return* (i.e. the return on the portfolio minus the return on the benchmark) relative to the standard deviation of the excess returns (i.e.

the *tracking error*) and is a measure of risk adjusted returns. The higher the excess returns earned relative to the variability of those returns (i.e. risk) the higher the information ratio and vice versa.

There are other common risk measures described in the literature such as the measure which has become known as the *Sharpe Ratio* as proposed by Sharpe (1966) and later revised by the same researcher (Sharpe, 1994). The *Sharpe Ratio* expresses the return in excess of the risk-free rate relative to the standard deviation of returns. There are also measures such as the ratio known as the *Sortino Ratio* which was proposed by Sortino and Price (1994) which considers only downside risk.

Given that the objective of enhanced indexing is to achieve small improvements in return over and above the returns generated by the index the *information ratio* is the risk measure most directly relevant to this task and as such will be the primary measure used in the evaluation phase.

2.3 Securities selection

2.3.1 Classification of equity securities

The classification of equity securities is concerned with segregating the securities into categories which are expected to have similar future performance. At a basis level the process involves placing stocks which are expected to perform well in one basket and those which are expected to perform poorly in another. There is however an almost infinite number of metrics and characteristics which could be used to group stocks into these different investment categories. The one which is commonly applied in practice is the ‘value’ versus ‘growth’ continuum. Capaul *et al.* (1993) considered the effectiveness of this distinction in terms of portfolio construction in an international context and observed that there was a significant difference in the relative performance of the portfolios formed in this way over their testing period and that this difference in performance could not be attributed to random factors.

For the purposes of their study the researchers used a single metric, ‘Price to Book’ (P/B) to distinguish growth stocks from value stocks with growth stocks being those which exhibited high P/B ratios and value stocks being those which had low P/B ratios. While Capaul *et al.* (1993) observed that value stocks outperformed growth stocks over their testing period other researchers have observed different results over different periods. Pisani (2015) observed that the relative performance of value versus growth stocks has varied over time citing research conducted by, Gerstein Fisher, an investment firm, which does show however, that historically there have been lengthy periods in which one category has continuously outperformed the other.

While Capaul *et al.* (1993) consider only the P/B Ratio in defining a classifier to distinguish growth and value stocks more generally in the financial press and in the investment profession both the P/B Ratio and the P/E Ratio (price to earnings ratio) are factors used in making the classification. Penman (1996) considers the relative effectiveness of the P/B ratio and the P/E ratio in predicting future growth and concludes that the P/B Ratio is a useful metric in this context as it is linked to future profitability whereas the P/E Ratio is not.

Having considered the evidence presented in the literature supporting the existence of a growth value effect it is necessary to consider the composition of the returns achieved under each category in order to facilitate the formulation of predictions. Fama and French (2007) perform such a de-composition of the sources of returns for both value and growth stocks.

The researchers analyse the returns for the portfolios into income and capital gains components. The income component is comprised of dividends and the capital gains component is further segregated into three sub-components namely; increases in book equity which arise mainly from profit retention, P/B Ratio convergence attributable to mean reversion in performance and expectations, and thirdly the general market trend towards higher P/B Ratios over the testing period.

The researchers observe different relative impacts for growth and value stocks from these components. They note that convergence in P/B Ratios provides most of the capital gains from value stocks but is negative for growth stocks. The researchers further note that increases in book equity is a significant source of capital gains for growth stocks but is negligible or negative in the case of value stocks. The impact of the increase in market P/B ratios over the test period is assessed as small compared to the convergence effects for both value and growth stocks.

2.3.2 Technical analysis

Technical analysis centres around the identification of trends and inflection points based on ratios and historic data for the purposes of making predictions regarding the future performance of individual stocks or the performance of the overall market.

One aspect of this is the concept of ‘mean reversion’ meaning that while prices and ratios may fluctuate in the short to medium term there is a tendency to return to an underlying long term average. Becker *et al.* (2012) investigate mean reversion in the context of the P/E Ratios on the S&P 500. The authors note that it has been suggested by numerous researchers that future stock prices changes can be predicted using the P/E Ratio on the basis of mean reversion.

The authors perform unit root tests to demonstrate stationarity in the time-series data with the implication that the P/E Ratio exhibits mean reversion. On the basis of their test results Becket *et al.* (2012) express support for the view that a P/E Ratio above its long-term average is predictive of a future period of lower stock price expansion and or higher growth in earnings.

Another aspect of technical analysis is the study of stock price momentum. Momentum refers to the rate of change in stock prices over a specified period and is commonly measured by means of moving averages. Chiarella *et al.* (2006a) note that despite the existence of considerable research demonstrating that markets exhibit at least *weak-form efficiency* which by extension implies that technical trading rules would be ineffective in achieving market outperformance their use remains common in practice.

The researchers undertake an investigation into the effectiveness of moving averages using a dynamic technique employing both technical and fundamental concepts. Under their proposed model the buying of securities is driven by both technical and fundamental factors with the technical component being determined by the differential between current prices and the long term moving average. In their model the fundamental component is driven by mean reversion to a fundamental price which has been derived based on analysis of factors such as earnings and other economic factors. It follows that the traders in their experiment make decisions based on these factors.

For the purposes of their analysis the researchers further assume that all market participants operate as either technical traders or fundamentalists which they note are the most common trader types in practice. The authors note the intuitive logic that the actions of the fundamentalists and technical traders would be expected to exert stabilising and destabilising effects on prices respectively. The researchers note as a significant finding the observed impact of changing the length of the period used in the moving-average rule noting that increasing it can create instability in a system which was otherwise stable.

On a related note Chiarella *et al.* (2006b) demonstrate in another paper that the use of moving averages combined with market noise generates instability causing prices to deviate from fundamentals for lengthy periods. The implication of the above for this study is that the evidence for the existence of market inefficiencies moves the balance in favour of technical measures and away from fundamental measures as model features for shorter term stock selection decisions.

As technical trading rules depend on the existence of predictable trends it would be impossible for such strategies to work if stocks followed a random walk pattern. Brock *et al* (1992) investigate the effectiveness of moving averages and the randomness of stocks. The researchers produce evidence in favour of the effectiveness of technical trading and confirm that the results they produce are not attributable to random factors by performing simulations using a number of models including; random walk, general autoregressive conditional heteroskedasticity in mean model (GARCH-M) and the Exponential GARCH model.

The authors note that the trading signal produced by moving average rules is provided by the interaction between two moving averages which are calculated over long and short periods respectively. They note that a buy signal is generated when the short term moving average moves higher than the long term moving average and that conversely a sell signal is generated when the movements are in the opposite direction. They further note that a band is often imposed around the moving averages to remove the effect of conflicting signals which might arise when the short and long term averages are running close together.

The authors note that other implementations of this trading rule require the behaviour of the short term moving average to be assessed over a number of days following a cross over with the long term line. In such implementations a trading signal will only be generated if the required pattern is observed over those days. They also note that still other implementations give consideration to trading volume.

Brock *et al.* (1992) also consider a further technical trading rule referred to as the 'trading break out range' which is based on the concept of 'resistance levels'. In this context, the researchers note that technical traders hold the view that investors are willing to sell when a stock reaches its previous peak and buy when it reaches its previous trough thereby establishing a trading range with the resulting buying/selling pressure generating resistance to moves outside this range.

The author notes that a sell signal is generated under this rule when the stock price falls below the lower resistance level and conversely a buy signal is generated when it moves above the upper resistance level. It is noteworthy that Brock *et al.* (1992) identify an asymmetry in the returns generated from buy signals compared to those generated from sell signals under these technical trading rules with buy signals producing consistently higher returns and lower volatility.

The Relative Strength Index (RSI) developed by Wilder (1978) provides a further momentum measure in common usage. The RSI is the ratio of the average of up moves to the average of the down moves over a specified period. The Up and Down moves are calculated based on the movement in the closing price compared to the previous day. If the price has increased from the prior day the Up move is calculated as the closing price minus the previous day closing price and the Down move is assigned a value of zero. If the price has decreased from the prior day the Down move is calculated as the prior day closing price minus the current day closing price and the Up move is assigned a value of zero.

The index produces values between 1 and 100 with values towards 0 being indicative of an oversold market and values towards 100 being indicative of an overbought market. Wong *et al.* (2003) perform an assessment of the effectiveness of technical trading in the context of the Singapore stock market utilising the RSI as the counter trend indicator and moving averages for trend following demonstrating that both these measures produce positive results. In their description of the operation of the RSI they note that the index reading required to produce an accurate signal varies depending on the length of the period covered by the index with the values moving toward the mid-point as the length of the period increases. Of particular relevance is the observation that in trending markets the RSI can become misleading as it becomes stuck at one end of the scale.

Wong *et al.* also describe various trading rules derived from the RSI such as ‘Peak’ strategies which requires the index to cross the threshold and return back inside and a ‘Touch’ strategy which merely requires the index the reach one of the bounds.

While the Relative Strength Index (RSI) can be used to assess the performance of an individual stock a further measure is useful in terms of comparing the performance of one stock to another or in the context of index benchmarking the performance of an individual stock relative to the performance of the benchmark.

One such measure is the Relative Strength (RS) ratio which takes the price of the stock as the numerator and the level of the index as the denominator. Although multiple references are found to this ratio or variations of it in practice it appears to have no specific coverage in the academic literature. Nevertheless, given its practical usage and relevance to the topic of this study it will be used as a candidate feature for model construction.

In addition to changes in the price of a security technical analysis also requires consideration of the volume of trading driving these price changes. The intuition behind this is that a price movement on a large trading volume carries more information about investor thinking than one driven by lower volume. Scott *et al.* (2003) observe that previous research has indeed found that trading volume is useful in forecasting future US stock market returns and that this momentum effect has been shown to be more evident in high volume than in low volume securities. The authors contend however that what appears to be a response to volume is in fact a result of investors under-reacting to earnings news.

The researchers structure their argument by observing that earnings news impacts trading volume and price with the impact being larger for growth stocks. They further argue that investors exhibit over confidence and that as a result there is a time lag between the news release and the adjustment to investor expectations. They contend therefore that when differing growth rates are factored in the full momentum-volume relationship is explained by delayed reactions to earnings news. Despite the mixed views contained in the literature volume will be included as a candidate feature for initial model development.

In addition to the individual and combinatory technical indicators discussed above there is a broader concept referred to as investor sentiment representing the general attitudes of investors and is ranked on a scale from *bearish* to *bullish*.

Baker and Wurgler (2007) propose a definition of investor sentiment in which investors are said to hold views unsupported by the facts. The writers mention one approach to the measurement of investor sentiment as being based on the consideration of investor biases such as over confidence and conservatism. The writers proceed to present an alternative macroeconomic top down approach to sentiment measurement and use it to demonstrate the important finding that hard to value stocks are more susceptible to sentiment effects.

Bandopadhyaya and Jones (2008) note a recent trend in the literature towards non-economic influences on asset prices such as investor sentiment and refer to research which suggests that changes in investor sentiment may be a superior explanatory factor for short term price movements that the fundamentals. The authors consider the usefulness of two measures of investor sentiment, namely the VIX and the Put-Call Ratio (PCR) published by the Chicago Board Options Exchange (CBOE) and conclude that the VIX provides the superior measure.

2.4 Quantitative investment approaches

2.4.1 Identifying profit opportunities using deep learning techniques

Shen *et al.* (2011) note that predicting stock market indices is a challenging task and one which has generated a lot of interest. The authors categorise existing stock index prediction models into one of two categories comprising of: those based on statistical theories and those based on artificial intelligence which might alternatively termed as traditional machine learning approaches and deep learning approaches. They observe that the weight of current research demonstrates that the later outperforms the former particularly in short term forecasting.

The authors discuss a neural network based approach to stock market prediction to forecast the Shanghai Security Exchange Index. The approach they adopt employs a Radial Basis Function (RBF) neural network optimised using an Artificial Fish Swarm Algorithm (AFSA). They note that the RBF neural network is a form of feedforward neural network which has had broad application in the field of short-term prediction owing to its self-learning and self-adaption characteristics.

With respect to this model Shen *et al.* (2011) note that it exhibits the typical limitations of artificial intelligence models and discuss some of the optimisation approaches that have been adopted in practice to counter this. They note the use of both Support Vector Machines (SVMs) and Genetic Algorithms along with the use of AFSA approaches in other non-stock market related contexts. The authors note that the AFSA approach is analogous to the feeding patterns of shoals (aka swarms) of fish whereby with each individual fish passing on the knowledge it has gained to the rest of the swarm such that the behaviour of the swarm as a whole is optimised. The researchers contrast the AFSA with standard k-means clustering which they note is prone to convergence to a local minimum. They observe that the AFSA approach in being parallel and operating independently of initial values counters this problem. As such in their approach the researchers use an AFSA approach to modify the centre parameters for the k-means clustering algorithm which they employ in the neural network optimisation.

To extract appropriate model inputs the researchers utilise data mining techniques to identify the factors which exert a strong effect over the performance of the stock index. Shen *et al.* (2011) identify twelve such factors which they categorise into three groups; technical indicators, yield indicators and closing price indicators. The technical group is comprised of volume and moving average measures which as discussed elsewhere in this literature review are popular measures with research in the literature supporting their effectiveness but with some dissenting views being noted with respect to volume. The factors shown to be the best predictors by Shen *et al.* (2011) ultimately did not include this volume measure.

The second group concerning *stock yield* calculates the average for this measure calculated over various numbers of days prior to the current date. While Shen *et al.* (2011) do demonstrate the effectiveness of the stock yield measure this researcher has a preference for an alternative measure relating to this concept utilising the P/E Ratio.

Unlike stock yield the P/E Ratio is not impacted by differing dividend policies and as such this alternative measure is utilised in the modelling experiment.

Finally, the closing price group utilised by Shen *et al.* (2011) considers the closing price on the days prior to the current day which although not explicitly stated by the writers appears to be geared towards measuring the impact of short term momentum.

The researchers test the effectiveness of their proposed model in predicting the Shanghai index using combinations of the twelve model features and find that the best results were achieved using a combination of technical features comprising specific moving average and stock yield measures.

Kara *et al.* (2011) undertake a comparative investigation into the relative effectiveness of support vector machines (SVMs) and neural networks in predicting the direction of daily movements in stock exchange indices utilising the Istanbul Stock Exchange National 100. The authors reference the observed success of both SVM and neural network models in financial prediction but also point to their known limitations in the context of stock market prediction due to the levels of noise and complex dimensionality present in these markets.

From this perspective the study undertaken by Kara *et al.* (2011) is of interest as it is conducted by reference to an emerging market stock market which, as noted by the authors, has experienced high levels of historical volatility and hence provides significant exposure to factors which have been observed to inhibit the training of neural networks. The researchers demonstrate that the neural network model which they develop outperforms their SVM model with the neural network model obtaining almost 76 per cent predictive accuracy.

As noted above and elsewhere in this literature review the evidence regarding the usefulness of volume measures in making financial market predictions points both ways. The use of volume measures in the context of neural network based prediction models is specifically considered by Chavaranakul and Enke (2008).

The researchers outline the origins of the concept of equivolume charting which is a framework for considering how stock prices move in the context of volume levels as opposed to a purely temporal context. This equivolume charting concept modifies the standard bar chart showing price on the vertical axis and time on the horizontal access by using the width of the bars on the chart to measure the volume for the measurement period. Also, rather than being fixed to the horizontal axis the bottom of the bar is set to equal the low price for the measurement period as shown on the vertical axis and similarly the top of the bar is set to equal the high price for the day.

As such the equivolume chart is comprised of a series of floating bars (referred to as 'equivolume boxes') of varying heights and widths arranged chronologically across the horizontal plane. The larger the volume during the measurement period the wider the bar. Similarly the height of the bar increases the greater the gap between the high and low price for the measurement period. Chavaranakul and Enke (2008) use days as the measurement period for the purposes of their experiment.

The researchers utilise two technical indicators derived from equivolume charting in their model namely the Volume Adjusted Moving Average (VAMA) and the Ease of Movement (EMV). The researchers implement their model using a Generalised Regression Neural Network (GRNN) testing its effectiveness by reference to the S&P 500. As justification for the use of a neural network model the authors cite their established recognition, classification and forecasting strengthens in the context of financial data along with their ability to handle incomplete or ambiguous data which is subject to extreme short term volatility. With respect to the specific choice of a GRNN the authors describe it as a one-pass parallel algorithm which can operate smoothly in the context of multi-dimensional problems even with limited data and does not require assumptions as to the form of the function to be made in advance.

The researchers objective is to assess whether the application of a neural network approach can provide enhancements which supplement the trading signals provided by the combination of the VAMA and EMV measures to improve trading profits. In this context the researchers implement GRNN models to generate predictions for the future values of both the VAMA and EMV measures. To generate the predictions for the future VAMA values the researchers utilise the GRNN model to formulate forecasts of future stocks prices and the parameters of the associated equivolume boxes.

With respect to the testing period Chavaranakul and Enke (2008) select a period in which the market was trending noting the fact that moving average based measures have been shown to perform well in such markets and to be ineffective in range bound markets. The researchers demonstrate that their neural network based enhancement produces superior results to trading strategies based purely on trading signals provided by VAMA and EMV measures by providing the trading signal earlier.

Examining another perspective on the utility of neural networks in stock market prediction Zhang and Lenan (2009) evaluate the performance of back propagation neural networks optimised using an improved 'bacterial chemotaxis optimization (BCO)' technique in prediction stock market indices.

The researchers consider both next day index predictions and a longer term prediction for the index fifteen days hence. By way of background the researchers observe that BCO is a biologically inspired concept based on the feeding practices of bacteria. In this context the information sharing between the bacteria leads initially to an improvement in the feeding environment but then causes the bacteria colony to coalesce into 'local nutrients' which the writers observe is analogous to the algorithm become caught in a local minimum trap. The BCO improvement outlined by the writers is analogised by reference to the bacteria considering other nutrient sources in their surroundings and as such the bacteria will move to the new location negating the tendency towards local minimum traps. In terms of experiment results the researchers contend that their model provides improvements in both predictive accuracy along with reduced training time and complexity as compared to the back propagation neural network model not optimised in this way.

2.4.2 Implementing deep learning techniques in quantitative trading

The terms '*algorithmic trading*' and '*quantitative trading*' appear at times to be used interchangeably with no robust definition of either being found in the literature. However, rather than debating the technical distinction between the terms this researcher will focus in line with the scope of this study on a review of the research aimed towards generating profits from deep learning approaches rather than on pure trade execution applications.

Vanstone and Finnie (2009) present a framework for the development of algorithmic trading models using neural networks which they refer to by the synonymous term 'mechanical trading systems'. Their proposed definition refers to it as operating on the basis of fixed rules with no discretionary elements. The authors note the absence of a formal methodology in the literature addressing the processes involved in designing such a system and note the problem this presents for researchers in developing functioning neural network trading models.

The first step in the framework proposed by Vanstone and Finnie (2009) involves identifying possible model inputs through fundamental and technical analysis which has already been covered by this literature review. Rather than considering the precise steps in model development covered by the authors this literature review is concerned with evaluating the insights into model development and the nature of the constraints which are considered relevant by the literature. Vanstone and Finnie (2009) refer to a 1997 text by Tushar S. Chande '*Beyond Technical Analysis: How to Develop and implement a Winning Trading System*' in which that author proposed that a trading system comprised of three principal components being trading rules, risk control and money management. The development of trading rules has already been addressed by this literature review. In the context of risk management the author proposes that risk management may be articulated in the context of a trading system as the establishment of pre-determined exit rules for open trades, that is the market developments which would prompt the investor to close the position such as incurring loses of a defined magnitude. The author describes the money management decision as determining the capital allocations to specific trades.

With respect to the neural network architecture Vanstone and Finnie (2009) consider different perspectives on the determination of the appropriate number of hidden layers and note a considerable diversity of decision approaches including some based on theory and others based on experimentation. The most pertinent observation from this aspect of Vanstone and Finnie (2009) is that there appears to be no consensus on this aspect of neural network design.

Vanstone and Finnie (2009) also note a number of practical constraints which the trading system must consider in order to be realistic including avoiding using information which is not available in the market at that time, taking account of restrictions on short selling and slippage which refers to the fact that trades cannot necessarily be executed at market open price.

2.4.3 Implementing portfolio management using deep learning

Bahrammirzaee (2010) provides an overview of the use of artificial intelligence techniques in the finance area specifically considering the factors giving rise to increasing interest in this area and concludes that artificial intelligence approaches offer improved accuracy as compared to the traditional alternatives although the author notes that this improved performance is not universal.

While the writer considers a broader financial industry context including credit evaluation, portfolio management and financial prediction only the latter two are reviewed in the context of this study. The researcher notes that various approaches have been applied in this context and segregates these into three categories; namely parametric approaches such as logistic regression, non-parametric approaches such as k-nearest neighbour, and soft computing approaches such as artificial intelligence and observes that the latter category, and in particular neural network approaches have attracted the most interest.

While the author considers three areas of artificial intelligence including expert systems and hybrid intelligence in addition to neural networks due to the focus in this study on deep learning only the neural networks aspect of the researchers work will be considered. In the context of portfolio management, Bahrammirzaee (2010) undertakes a review of the work of a number of researchers and notes that all of the research reviewed points to superior results being achieved using neural networks, and in particular backpropagation neural networks, as compared to the results achievable using traditional methods.

In the context of financial prediction, the author notes the complex non-linear nature of the problems presented as a reason for the broad use of neural networks in this area and observes that a large volume of research has been performed in this area including work on asset price and exchange rate prediction. However, in the area of financial prediction the author noted that while there was a large volume of research which generally supported the superiority of neural network approaches some researchers reported no consistent outperformance of neural network techniques as compared to other approaches.

Fernandez and Gomez (2007) consider the use of neural networks in the context of generating Markowitz efficient portfolios and propose a specific neural network model application to achieve this and compare the results to that achieved with heuristic algorithms.

As discussed earlier in this chapter the Markowitz approach to portfolio management utilises mean-variance analysis to construct efficient portfolios which seek to maximise return for a given level of risk. Fernandez and Gomez (2007) utilise a generalised form of the Markowitz model with additional constraints pertaining to both the number of securities to be included and the capital to be allocated to each security. As the researchers point out these constraints are useful in practical applications of portfolio optimisation.

The researchers note that resolving the portfolio selection problem in the context of the standard Markowitz model constitutes a quadratic programming problem but is transformed into a mixed quadratic programming and integer programming problem when generalised to include the additional constraints imposed by the researcher as described above.

It is noted that there is no specific algorithm in existence to solve the portfolio optimisation problem when configured in this form which has led to the use of heuristic approaches based predominantly on evolutionary algorithms, simulated annealing and tabu search. The researchers observe that a neural network model known as the Hopfield network has been applied in other optimisation contexts and perform an experiment to assess its effectiveness in the context of portfolio selection.

The standard Markowitz approach to portfolio selection involves mapping out the *efficient frontier* using the mean-variance approach. In their experiment, Fernandez and Gomez (2007) adopt a neural network approach to generate what they describe as the general efficient frontier which unlike the standard *efficient frontier* takes account of the additional constraints of the number of securities to be included and the capital invested in each.

In their implementation of the neural network model the researchers adopt a two-step approach involving ‘neuron pruning’ and ‘objective function minimisation’ which is performed iteratively until the cardinality constraint is satisfied with a further iterative process being performed to balance the allocation of funds between the securities to ensure that the capital allocation constraint is satisfied. The researchers apply their model along with models based on genetic algorithms, simulated annealing and tabu search to several stock indices and compare the results. They observe that while no model dominated across all the investment strategies their neural network model outperformed for low risk portfolios.

Providing another perspective on the topic Zhu *et al*. (2011) note the issue arising with respect to portfolio optimisation given the non-linear constraints which apply and the difficulties of addressing these using conventional techniques. To address these issues the authors present a 'Particle Swarm Optimisation Technique (PSO)' which they benchmark against a genetic algorithm approach. Zhu *et al.* (2011) acknowledge the results of research undertaken by Giovanis (2009) which the authors note demonstrated the benefits provided by a hybrid model using artificial neural networks combined with PSO algorithms in a portfolio optimisation context.

However, the author points to further research which highlight problems arising in a portfolio management context when using neural networks such as local minimum traps and overfitting. In response to this the researcher develops a PSO model with constraints derived from the Markowitz approach and the *Sharpe Ratio.* The researcher demonstrates superior results for the PSO model developed compared to a Genetic Algorithm based approach.

The author does not present a comparison of the results generated by the PSO model to those derived from a neural network or neural network hybrid model however the insights highlighting potential problems with such models are nevertheless useful in the context of this study.

Ko and Lin (2008) also consider the specific area of resource allocation which in the context of portfolio management refers to the allocation of funds between the various components of the portfolio. The writers acknowledge the power of neural networks in dealing with such resource allocation problems but point to the problem arising from the inability of regular neural network implementations to ensure that the asset weights calculated by the model sum to 100 per cent.

The writers state that most neural networks utilise correction procedures such as dividing the output vector by its sum to handle this problem which in their view poses the risk of compromising the model. The researchers propose an alternative solution to the problem using a neural network model which they term a 'Resource Allocation Neural Network (RANN)' which introduces a feature to ensure that the asset weights sum to 100 per cent at each learning epoch rather than adjusting the final model after training has been completed.

Yu *et al.* (2008) address the application of neural networks to the determination of the mean-variance-skewness trade-off in portfolio optimisation and demonstrate positive results for the neural network based model under evaluation. As discussed earlier the standard Markowitz model adopts a mean-variance approach to optimisation which by implication assumes a normally distributed population. Yu *et al.* (2008) discuss the importance of considering skewness when dealing with non-normal populations and refer to some extensions of the Markowitz model to perform portfolio optimisations using a mean-variance-skewness approach.

The researchers note the challenges presented by this multi-faceted optimisation problem and highlight the restricted nature of the methodologies available to solve this problem. In response to this problem the researchers propose a neural network based model know as a Radial Basis Function (RBF) network to perform the required optimisation. The researchers test their model empirically using a number of prominent stock market indices and currency pairs and conclude that the RBF network model is an efficient tool for mean-variance-skewness optimisation.

Freitas *et al.* (2009) utilise a neural network model to develop a different approach to portfolio construction implementing what they refer to as a 'prediction-based portfolio optimisation'. All of the portfolio optimisation approaches discussed so far in this section have employed derivations of the Markowitz model.

The authors provide an overview of their model inputs contrasting it with the standard Markowitz mean-variance approach. Freitas *et al.* (2009) utilise predicted returns in place of the mean return calculated from historic data. In place of a risk measure based on variance and covariance calculated from the historic returns the researchers use measures of variance and covariance computed from the time-series of errors of prediction. Finally, with respect to the normality assumption the model proposed by

Freitas *et al.* (2009) relies on a normal distribution of the error of prediction as opposed to normality of the distribution of returns.

The authors note the challenges with respect to accurately predicting stock returns but point to the potential for error reduction in a portfolio context as the errors in the individual stock predictions partially offset when combined in a portfolio.

The researchers use an 'auto-regressive moving reference neural network (AR-MRNN)' to implement their model. Model testing was performed using Brazilian stock market data with the model results being compared to the performance of the IBOVESPA index and results which would be achieved under a mean-variance approach and conclude that the prediction-based model which they developed has the potential to produce superior results to mean-variance models when used for short-term trading.

2.5 Development considerations and evaluation approaches

A detailed investigation and discussion of the construction of neural network models or the mathematical logic underlying the detailed implementation decisions is beyond the scope of this study. However, in order to establish a theoretical sound basis for model development consideration is required to be given to certain concepts pertinent to the design and evaluation decisions which will be implemented in this study.

Lawrence *et al.* (1997) note the two important measures of machine learning model usefulness as being their ability to generalise and scale with complexity. They further note the tendency towards overfitting which these models exhibit. With respect to overfitting the writers discuss the concept of degrees of freedom in the context of neural networks. As described by the authors this is essentially concerned with the relationship between the number of samples and the number of model inputs. The researchers note general rules of thumb such as one stating that the number of inputs should be considerably less than the number of samples but note that other factors also play a role and as such in their view rules of this nature are not wholly reliable. They also demonstrate that the specific rule of thumb mentioned above, which they note is widely held to be true with respect to common datasets, does not necessarily hold.

In terms of evaluating the effectiveness of the neural network model the use of a number of different measures has been observed in the course of this literature review.

Ko and Lin (2008) utilise Root Mean Squared Error (RMSE) and Expected Return as the evaluation metrics in their portfolio selection model tested on Taiwanese stock market data. Zhang and Lenan (2009) and Chavaranakul and Enke (2008) utilise Mean Squared Error. Kara *et al.* (2011) utilised the misclassification rate for their study into the use of neural networks and support vector machines to predict the direction of the daily movements in the Istanbul Stock Exchange National 100 Index.

Evans *et al.* (2013) provide a useful discussion of evaluation measures in the context of a foreign currency algorithmic trading model which they developed using neural networks and genetic algorithms. The authors note that the selection of the appropriate evaluation metric is critical to both model development and the assessment of differing strategies. The authors segregate the universe of available evaluation metrics into two categories being the more traditional measures deriving from the field of statistics and those relating directly to the measurement of the model objectives. In the former category they include measures such as Mean Squared Error (MSE) and Mean Absolute Error (MAE). In the latter category they include measures such as Cumulative Investment Return, Annualised Return and the Sharpe Ratio. Evans *et al.* (2013) also note that in the context of market predictions the more critical metric relates to correctly predicting the direction of the market with a second important metric being the prediction error.

Fernandez and Gomez (2007) utilise metrics falling within the second category discussed by Evans *et al.* (2013) in their portfolio optimisation study, that is, metrics relating to the performance of the model objectives. The specific measures used by Fernandez and Gomez (2007) include Mean Return Error and Variance of Return Error.

2.6 Summary

2.6.1 Summary of the current state of the art

The review of the literature has identified a considerable volume of research in both the application domain and the technical specification. There has been extensive research into the factors driving the prices of both individual securities and index values. The origins of the current approach to both securities selection and portfolio management extends back decades and has continued to grow with the assistance of advances in mathematics and computer science.

The bedrock for current portfolio construction approaches (referred to as 'Modern Portfolio Theory') was established by Markowitz and Sharpe commencing in the 1950's. The guiding principal behind portfolio management is that there is a trade of between risk and reward and that portfolios should be structured so as to maximise the return for a given level of risk.

A powerful attribute of efficient portfolios is that risk can be reduced without necessarily reducing returns by combining assets which are less than perfectly correlated in the portfolio. In this way a diversification effect is achieved and if the market portfolio is held the investor is no longer exposed to security specific risk. The market portfolio is generally approximated by a broad market index such as the S&P 500 Index.

Portfolio management strategies may be categorised as either passive or active strategies. Passive strategies track benchmark indices such as the S&P 500. Active strategies attempt to outperform the market through securities selection. Hybrid approaches such as Enhanced Indexing combine elements of both strategies.

The objective of passive management strategies may be stated as the minimisation of tracking error whereas that of active strategies may be stated as the maximisation of alpha. Hybrid strategies must balance these two objectives and require a dual optimisation approach.

Securities selection is the key to successful active management. A high level classification of stocks into discrete categories can assist in this task. A commonly applied approach is to split stocks into 'value' and 'growth' categories. The P/B Ratio and P/E Ratios are used to achieve this. A low ratio denotes a value stock and high ratio a growth stock.

At a more granular level securities selection techniques may be broadly categories as either technical analysis or fundamental analysis. Both approaches have proponents and critics. Market participants who consider markets to be at least *weak form efficient* reason that technical analysis which is based on historic pattern analysis cannot provide predictive power and instead utilise fundamental analysis. Other researchers report that empirical research supports the use of technical analysis. This researcher has considered the weight of the research on each side, particularly in the context of short term predictions and has concluded in favour of technical measures for model development.

The research has identifying numerous popular technical measures including various forms of moving averages, trading ranges, volume and momentum measures which will be considered in the feature selection process.

Machine learning approaches more generally and neural network approaches in particular have been widely researched with the weight of research being supportive of their effectiveness in both a price prediction and portfolio management context. A key feature of the research is that various types of neural network models including GRNN, RBF networks and BP neural networks have been tested and demonstrated to be effective in various contexts but with no overall implementation standing out as dominant.

In addition to the selection of neural network type a considerable volume of research has been devoted to the testing of various optimisation techniques. Neural networks have been shown to be prone to certain problems such as overfitting which is known to be a particular problem when the dataset contains a lot of noise which is the case for stock market data. Various optimisation techniques have been developed to counter this problem particularly *metaheuristic* techniques. Again the various optimisation techniques have been shown to be effective in particular contexts but with no clear best approach evident from the literature.

2.6.2 Gaps in the current state of the art

Overall, there has been a considerable volume of research which covers various aspects of security selection and portfolio optimisation. However, there appears to be limited research specifically addressing the area of portfolio management in the context of passive investment or more particularly in the field of enhanced indexing.

Given the importance of the passive management area and the potential for enhanced indexing in this area this appears to be an area which would benefit from further research.

3.0 DESIGN AND METHODOLOGY

3.1 Introduction

This chapter will discuss the data requirements for the testing of the hypothesis along with an overview of the sample selection and data compilation process. It will also provide an exploratory summary of the data characteristics along with the details of the input selection process and provide a table of the final selection of candidate inputs which will be used for model implementation in the subsequent chapter.

3.2 Data compilation

Equities trading has a long history with the traces of modern stock markets reaching back more than a century. Throughout this modern history, the variations in the prices of both individual stocks and the performance of the overall market has generated considerable interest with the result that there is a vast archive of media reporting from both the financial and general press. The growth of equities trading following the 1980's deregulation and the arrival of the internet in the 1990's has combined to drive the creation of an almost infinite number of online sources for equities related research.

The potential sources of information may be segregated into two categories; subscription based services used primarily by financial institutions and investment firms and freely available online data repositories provided by various organisations. The general-purpose subscriptions based market is dominated by two firms, Bloomberg and Thomson Reuters, with a number of other specialist providers offering other products. The providers of free securities information are primarily tech companies such as Google Finance and Yahoo Finance and media firms such as The Financial Times and CNBC.

The nature of the study is such that the primary inputs are raw financial data and measures derived from this data. As such all of the required data is available online free of charge. Having given due consideration to reliability and ease of use Yahoo Finance has been selected as the primary data source.

3.3 Study Period

Financial markets are impacted by a combination of factors including economic cycles, investor sentiment, the political environment, regulatory changes and in the very long run demographic changes and the balance of global economic power.

From the perspective of the investment profession the primary impact of these factors is that investment strategies which work during one period may cease to work or even reverse in a subsequent period.

Deep learning techniques require significant amounts of data to train algorithms capable of finding relationships between the model features and the target variable. This presents a challenge in that the study period must cover a period of sufficient length to provide the data required to train the model while simultaneously not extending so far into the past as to introduce discontinuities in the data. The recent past included a sharp drop in equity markets from 2007 to 2009 with a sustained recovery with a general upward trend from 2012 onwards.

Having due regard to the above it has been determined that the study period should cover a five-year period with the models being trained using market data for the four-year period from 1 January 2012 to 31 December 2015 with testing being performed using the date for the year from 1 January 2016 to 31 December 2016.

3.4 Sample selection

As discussed earlier enhanced indexing strategies require small changes to the weights of the individual stocks held in the portfolio compared to the weights for these individual stocks in the benchmark index. As such, the conduct of the experiment requires both the selection of a suitable stock market index and selection of suitable stocks within that index in order to assess the impact of the changed weightings.

3.4.1 Benchmark index selection

There are a vast number of stock market indices in existence covering both developed and emerging markets with indices tracking both broad market performance and individual industry sectors. For the purposes of this study the selected benchmark must satisfy a number of criteria.

3.4.1.1 Media coverage

As an initial proxy for research coverage the selected index must be of such prominence that it is regularly reported on by the international financial press. This criterion limits the potential candidates to the broad market indices connected with stocks traded in the larger developed markets plus China. China alone among emerging economies has both the size and geo-political importance to exert influence on global stock market sentiment thereby ensuring international media attention. Coverage of indices reflecting the performance of stock markets in the smaller advanced economies rarely extends beyond the national press and as such do not satisfy this criterion.

3.4.1.2 Index construction

The methodologies used by the various index providers vary considerably. Some indices are value weighted (i.e. weighted by reference to the market capitalization of the constituent companies) whereas other are price weighted (i.e. calculated by reference to an average of the share price of the component companies without regard for their respect market capitalisations). Further indices may be price only indices or may be calculated on a total return basis and therefore include the impact of reinvested dividends. This study requires the use of a price only index. Most prominent indices are price only and hence this aspect of the criterion is less restrictive. Since the objective of the study is to assess the impact of varying index weights on performance a value weighted index is required thereby excluding price weighted indices such as the Dow Jones Industrial Average (DJIA) and the Nikkei 225.

3.4.1.3 Geographic scope

Prominent indices may be categorised as national, regional or global in scope. National indices include only stocks traded within an individual country whereas regional and global indices select their constituents from the population of eligible stocks traded within the specific region or globally respectively. To limit the impact exposures to multiple macro-economic environments global and regional indices are excluded as candidates. While the experiment could be extended to model the impact of such varying global factors this is beyond the scope of the study as it is not intrinsic to the research objective.

3.4.1.4 Index diversification

The objective of the experiment is to assess the impact of changing the weighting of individual stocks within the index. As such it is necessary to select an index which is not concentrated within a small number of industry sectors as adjusting the weightings on such an index gives rise to the risk of introducing sector bias. In line with the experiment objective any improvement in risk adjusted returns should be attributable to stock selection rather than sector tilts. This criterion would point to selecting an index with a large number of diversified components such as the S& P 500 and away from indices such as the FTSE 100 which has high concentrations of mining and financial stocks.

3.4.1.5 Domestic Focus

To limit the impact of varying exposures to macro-economic and geo-political risk applicable to individual index components the selected index should limit or exclude international indices. Also, while most large capitalisation companies have at least some international exposure an index reflecting the performance of home grown companies within a large economy will have less exposure than one in a smaller more international economy.

3.4.1.6 Liquidity and indexing

As discussed above the growth of passive investment management has led to the increased availability of Exchange Traded Funds (ETFs) tracking major indices. The volume of funds indexed to a fund is a good measure of its importance to investors. Since the objective of the experiment is to assess the ability of deep learning techniques to provide improved risk adjusted returns from enhanced indexing it is important that the selected index be one which is in practice widely indexed against by the passive investment management industry.

The S&P 500 is among the most widely tracked stock market indices. In addition, as a large portion of the indexes value derives from large capitalisation US stocks with high liquidity the transaction costs associated with trading in these stocks is relatively low thereby reducing portfolio re-balancing costs. This fact combined with high levels of industry diversification and a large number of constituents domiciled in a single large domestic economy point to its suitability as a candidate for experimentation.

Based on the criterion evaluated above the S&P 500 has been selected for use in this study.

3.4.2 Individual stock selection

The S&P 500 consists of the 500 largest US stocks. For the purposes of this study the sample of S&P 500 constituents selected must satisfy a number of criteria.

3.4.2.1 Liquidity

As large cap stocks trade more heavily than small cap stocks the sample shall be selected from the largest index constituents.

The twenty largest S&P 500 constituents are shown in Table 3.1 below. Sector information was obtained from Bloomberg.com While the precise composition of the Top 20 inevitably evolves over time as the market capitalisations of the constituents change such changes in composition are not relevant to this project as the objective is to define a sample from which to select a group of large liquid stocks rather than to track a specific section of the index.

Ticker	Company	GICS Sector
AAPL	Apple Inc.	Information Technology
AMZN	Amazon.com Inc.	Consumer Discretionary
BAC	Bank of America Corporation	Financials
BRK.B	Berkshire Hathaway Inc.	Financials
CVX	Chevron Corporation	Energy
FB	Facebook, Inc.	Information Technology
GE	General Electric Company	Industrials
GOOG	Alphabet Inc.	Information Technology
GOOGL	Alphabet Inc.	Information Technology
HD	The Home Depot, Inc.	Consumer Discretionary
JNJ	Johnson & Johnson	Health Care
JPM	JP Morgan Chase & Co.	Financials
MSFT	Microsoft Corporation	Information Technology
PFE	Pfizer Inc.	Health Care
PG	The Proctor & Gamble Company	Consumer Staples
T	AT&T Inc.	Telecommunications Services
VZ	Verizon Communications Inc.	Telecommunications Services
WFC	Wells Fargo & Company	Financials
XOM	Exxon Mobil Corporation	Energy

Table 3.1 Top 20 S&P 500 constituents

3.4.2.2 Sector concentrations

Stocks within specific industry sectors are impacted by macro-economic factors in different ways to stocks in other industry sectors. For instance, financial stocks exhibit higher sensitivities to changes in the interest rate environment than non-financial stocks and energy stocks tend to be impacted more strongly by oil price movements than health care stocks for example. Therefore, to test the hypothesis across a broader range of stocks and to ensure that results are the result of stock selection rather than sector bias the sample selected must be diversified across the industry sectors.

3.4.2.3 Continuity

The selected stocks must have been publicly traded and included within the large cap category throughout the study period. This criterion eliminates Facebook (FB) whose IPO was in 2012. The selected company's must have been in existence in their current corporate form throughout the study period. Large cap company's regularly make acquisitions and engage in restructurings and as such to test the hypothesis under realistic conditions such activities do not constitute ground for exclusion under this criterion. However, to preserve data continuity company's which have engaged in transformative mergers and acquisitions activities during the study period are not suitable for selection.

Table 3.2 shows the sample selected based on the criteria outlined above.

Ticker	Company	GICS Sector
AAPL	Apple Inc.	Information Technology
MSFT	Microsoft Corporation	Information Technology
JNJ	Johnson & Johnson	Health Care
XOM	Exxon Mobil Corporation	Energy
AMZN	Amazon.com Inc.	Consumer Discretionary
JPM	JP Morgan Chase & Co.	Financials
WFC	Wells Fargo & Company	Financials
GE	General Electric Company	Industrials
T	AT&T Inc.	Telecommunications Services
PG	The Proctor & Gamble Company	Consumer Staples
PFE	Pfizer Inc.	Health Care
CVX	Chevron Corporation	Energy

Table 3.2 Sample stocks selected

3.5 Data analysis and feature extraction

Having selected a suitable index and selected suitable constituent stocks from that index it is necessary to explore the characteristics of the data with a view to feature extraction.

3.5.1 Trend analysis

Trend analysis is an important component of technical analysis. Looking at stock prices on a day to day basis is likely to show the stock alternate between gains and losses however over time the stock price is likely to either trend upwards or downwards or perhaps remain flat. To filter out the noise generated by day to day fluctuations technical traders calculate moving averages for trend detection. The moving average is a straight average of the daily closing price over a specified number of trading days. While the moving average may be calculated over any number of days the most common in practice as indicated in the literature are the 50-day moving average and the 200-day moving average used to detect short term and longer term trends respectively with a rising 200-Day moving average being indicative of a long term upward trend and similarly, a rising 50-Day moving average is indicative of a short term upward trend. However, as observed in the literature review it is the signalling provided by the interaction between these two moving averages which is of importance in detecting a turning point. That is detecting the point when an upward trend becomes negative or vice versa.

As the objective of the study is to implement adjustments to index weights in order to outperform a pure indexing strategy the key point is to reflect the timing of the turning point of the selected stock relative to that for the overall index. As such the mentioned moving averages for both the selected stocks and the index will be included as candidate features.

3.5.2 Volatility measures

Since an integral part of the objective of the study is to increase risk adjusted returns the risk profile of the selected stocks is of critical importance. The benchmark measure of risk with respect to stocks is the standard deviation. The standard deviation may be calculated using either the security price or the returns on the security. However, given the large dispersion in the prices of the selected stocks it has been determined that the standard deviation of returns would provide a better measure by permitting direct comparability.

As volatility is a measure of the variability of returns around the trend it is calculated over shorter periods than moving averages. To provide a good measure of volatility over the recent past. 10-day, 20-day and 30-day volatility measures will be calculated. The equivalent volatility measures for the benchmark index will also be calculated to permit an assessment of the volatility of the selected stock versus the index.

3.5.3 Momentum measures

In technical trading the concept of momentum is founded on the belief supported by research discussed in the literature review that past increases in a stock's price tend to predict future increases. Momentum is calculated as the percentage change in the stock's price over a given number of days. Momentum can be measured over any period. However, for the purposes of this study three measures shall be used calculated over 1-day, 5-day and 20-day trading periods respectively in order to adequately capture various aspects of momentum.

Relative Strength (RS) ratios provide a related measure of momentum of particular significance to this study as it assists in identifying stocks which have higher or lower momentum relative to the index. The RS ratio is calculated as the stock's closing price divided by the closing price of the index.

3.5.4 Relative performance measures

As a key part of the study objective is to increase the weighting on stocks which outperform the index and reduce the weighting on those stocks which under perform a direct measure of this relative performance is required.

To achieve this the price performance of the stock and the index will be assessed over 1-day, 5-day and 20-day periods.

3.5.5 Volume measures

As observed in the literature review a price movement driven by high volume clearly provides a stronger signal than one driven by low volume. As such consideration of trading volume is important in assessing the value of signals provided by other indicators.

To reduce the impact of noise arising from the day to day fluctuations in trading volumes the current day's volume expressed as a percentage of the previous 10 trading day's volume will be used in addition to the day's volume as a candidate feature.

3.5.6 Earnings based measures

The Price Earnings (P/E) ratio ranks among the most common measure of value used in stock analysis. The P/E ratio divides the company's stock price by its earnings per share and effectively returns the number of year's earnings it would take to recover the stock's price (i.e. the earnings multiple).

The level of the P/E ratio of a stock or of an index relative to its historic average is a common proxy for assessing whether the stock or index is over or under valued.

However, as the numerator in the P/E ratio is a relatively slowly changing feature and concerned more with fundamental value rather than technical analysis it will be excluded as a candidate feature.

Feature	Calculation methodology
Closing Price (Adjusted)	The Closing Price on the Day adjusted for the impact of corporate actions
50-Day Moving Average (Stock)	Arithmetic average of the Closing Prices over the preceding 50 trading days
50-Day Moving Average (Index)	Arithmetic average of the Closing Prices over the preceding 50 trading days
200-Day Moving Average (Stock)	Arithmetic average of the Closing Prices over the preceding 200 trading days
200-Day Moving Average (Index)	Arithmetic average of the Closing Prices over the preceding 200 trading days

Feature	Calculation methodology
10-Day Volatility (Stock)	Standard deviation of the percentage daily returns over the preceding 10 trading days
10-Day Volatility (Index)	Standard deviation of the percentage daily returns over the preceding 10 trading days
20-Day Volatility (Stock)	Standard deviation of the percentage daily returns over the preceding 20 trading days
20-Day Volatility (Index)	Standard deviation of the percentage daily returns over the preceding 20 trading days
30-Day Volatility (Stock)	Standard deviation of the percentage daily returns over the preceding 30 trading days
30-Day Volatility (Index)	Standard deviation of the percentage daily returns over the preceding 30 trading days
1-Day Price Change (Stock)	Percentage change in price over preceding trading day
5-Day Price Change (Stock)	Percentage change in price over preceding 5 trading days
20-Day Price Change (Stock)	Percentage change in price over preceding 20 trading days
Relative Strength (RS) Ratio	Closing Stock price divided by Closing Index Level
14-Day Relative Strength Index	Ratio of daily increases to daily decreases over preceding 14 trading days
Daily Volume	Number of shares traded on the trading day
10-Day Average Volume	Average of number of shares traded over the preceding 10 trading days

Table 3.3 List of candidate features

3.6 Model Development

Having identified the candidate features it is now necessary to consider the specifics of the model to be developed to assess the hypothesis.

3.6.1 Model Overview

As discussed in the literature review above artificial neural network (ANN) with multiply hidden layers (i.e. deep neural networks (DNN)) are the most common approach to deep learning.

As the objective is to specifically assess the effectiveness of deep learning the models will initially be developed with a single hidden layer and then re-run with increasing numbers of layers to assess the impact of adding depth on the results. The model will use all of the candidate features listed in Table 3.3 as input variables.

The first phase of the experiment will consist of the training the neural network against the target variable using the data for the four-year period from 1 January 2012 to 31 December 2015. The experiment shall be established as a classification problem. As such, the target variable will take one of two possible outcomes; ‘outperform’ or ‘underperform’. The neural network model shall be trained separately for each of the selected stocks listed in Table 3.2 above. In this context, the target variable ‘outperform’ shall be applied if the stocks return for the day exceeds the return provided by the selected index. Conversely, the target variable ‘underperform’ shall be applied if the stocks return for the day is less than the return provided by the index.

The neural networks developed will then be validated using the 2016 data for the selected stocks.

The second phase of the experiment will involve applying the results of phase one to determine appropriate adjustments to the index weights of each stock. In order to minimise tracking error and in line with standard investment management rules a constraint shall be imposed to limit the amount by which each individual stock’s weight may be varied from its index weight. The limit imposed shall be plus or minus 1 per cent of the portfolio value.

Finally, the risk and return of the adjusted index shall be computed and compared to the performance of the (unadjusted) index. Risk in this context shall be assessed by means of the standard deviation of returns. Risk adjusted returns shall be compared by comparing the return per unit of risk (i.e. return divided by standard deviation).

3.6.2 Evaluation approach

The evaluation approach shall be two-fold. Firstly, the effectiveness of the neural networks developed will be assessed by reference to the accuracy rate. This accuracy rate shall be calculated for both the single hidden layer networks and the deep networks.

Secondly, a conclusion shall be drawn with respect to the effectiveness of incorporating the outputs of the neural networks into the adjusted weight index to deliver improved risk adjusted returns.

4.0 IMPLEMENTATION AND RESULTS

4.1 Introduction

This chapter will commence with an exploration of the pertinent characteristics of the raw data comparing and contrasting the metrics of the individual stocks to those of the selected benchmark index. It will then proceed to discuss the data preparation considerations and the approach adopted to compute the derived features and the creation of the data files for use in the developed model.

The detailed model development and testing process will then be presented for the single layer models first followed by a detailed discussed of the construction and testing of the deeper models. Both the single layer and multi-layer models will be subjected to forward testing with the results from this stage then being used to assess the construction of the optimised portfolio which will be discussed next. The final section of the chapter will include a presentation and discussion of the results of the portfolio optimisation phase which will contribute to the answer to the research question which will fully evaluated in Chapter 5 of this thesis.

4.2 Analytics tool kit

The conduct of effective analysis requires that suitable analytics software be selected for each individual tasks to be performed as part of the investigation. The principal components of this research project can in practical terms be categorised as; data consolidation, data exploration, derived feature computation, model construction and model evaluation.

The nature of this study necessitates the sourcing of raw data from different sources. All of this data is available to download in csv format and is obtained from a single source. The raw data consists entirely of numeric data presented in tabular format with the date field providing the primary key. While the data source offers a Python API it has been determined that since only a single static data file for each security will be required with no requirement for data updates direct csv download is the most efficient approach.

The computation of derived features requires the calculation of various arithmetic means, mathematical inequalities, ratios and counts. Since these are all relatively simple in mathematical terms the use of specialised mathematical libraries is not required. Since the same calculations must be performed for each of the twelve selected stocks for both the training and test data effective automation of the repetitive task is an important consideration. In this context an Excel model has been utilised to both compute the derived features and output the data in the required csv format.

The data exploration task involves both a computational and a visualisation element. As the statistical calculations are more mathematically intensive the Python scipy library has been utilised for this task. Pythons matplotlib has been used for the visualisations for this task and all other visualisations throughout this study.

The model construction and evaluation tasks are implemented using Anacondas scientific distribution of the Python programming language. In particular the Python numpy and scikit-learn libraries and the Keras neural network wrapper with Theano backend have been employed to implement the neural networks.

4.3 Data exploration

The overall objective of enhanced indexing is to achieve superior risk adjusted returns by making small deviations from index weights through judicious stock selection. In order to access the impact of changing the weights it is necessary to consider how the metrics for the individual sample stocks differ from those for the index. This has been accessed on a historical basis over the four year period to 31 December 2015.

4.3.1 Relative price performance

The data exploration commenced with a high level review of the price performance of the stocks over time versus the price performance of the index.

The performance of a portfolio is ultimately determined by the relative price performance of its constituents compared to the benchmark. Over time the cumulative impact of diverging performance between the index and the stock can have a substantial impact.

Line charts have been produced for the index and each individual stock to visually illustrate this impact. In order to achieve comparability the price movements for both the index and the sample stocks have been expressed taking a base of 100 at 1 January 2012 rather than using the actual stock prices. For the purposes of this analysis the actual stock price is not important. What is of relevance is how the price movements evolve over time.

The line charts for the three largest sample stocks (by market capitalisation) are shown below. The full suite of charts for all of the sample stocks are contained in Appendix I.

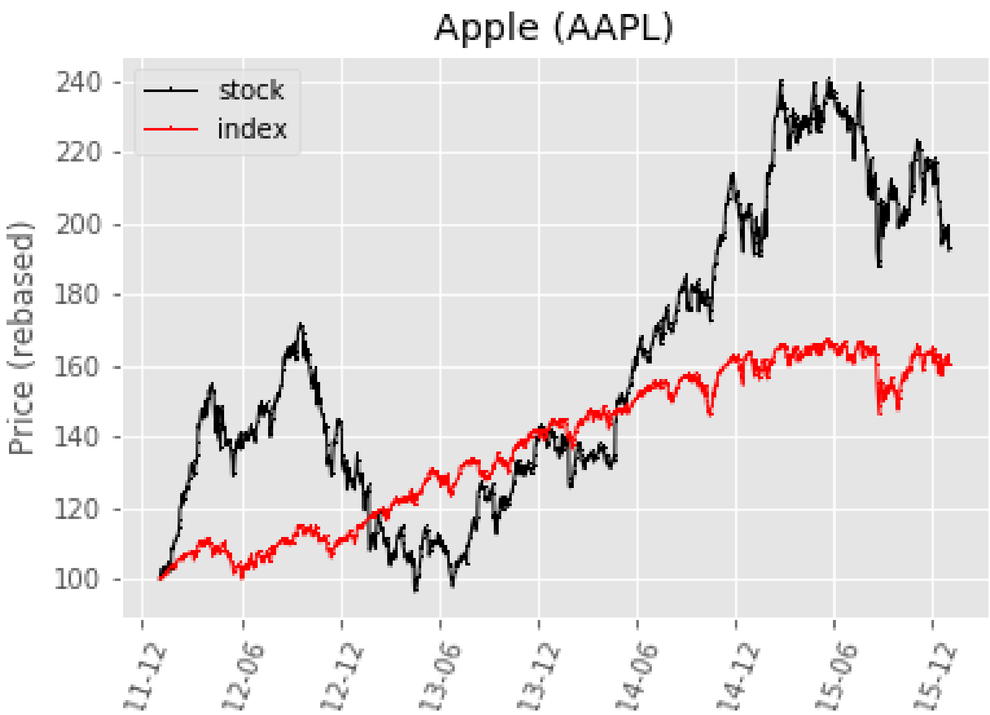


Figure 4.1 Daily prices for Apple relative to the S&P 500

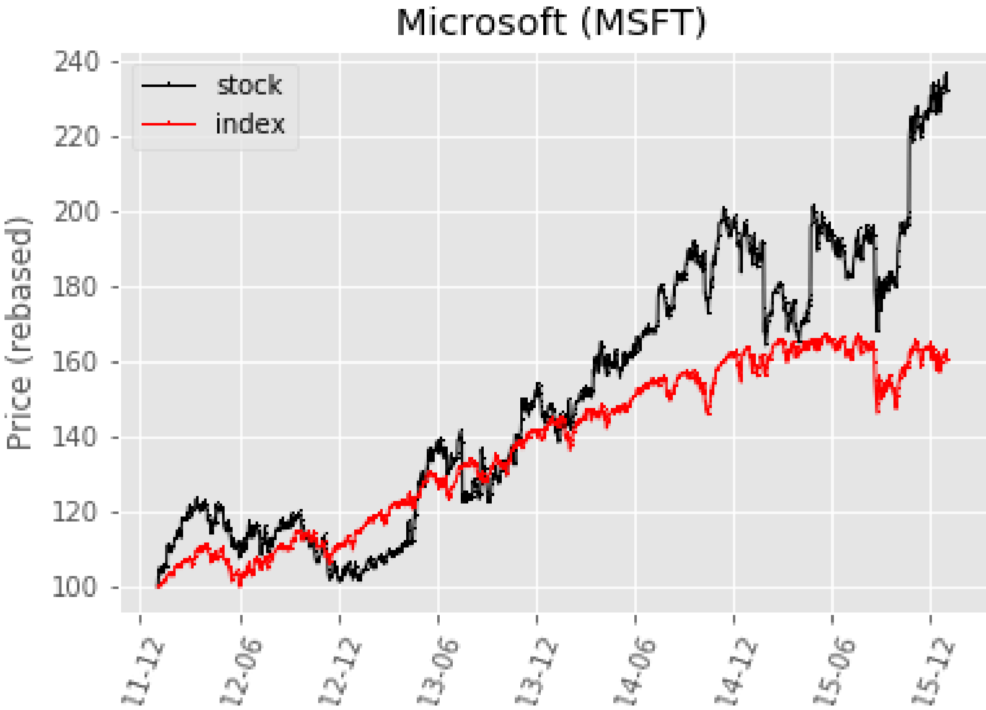


Figure 4.2 Daily prices for Microsoft relative to the S&P 500

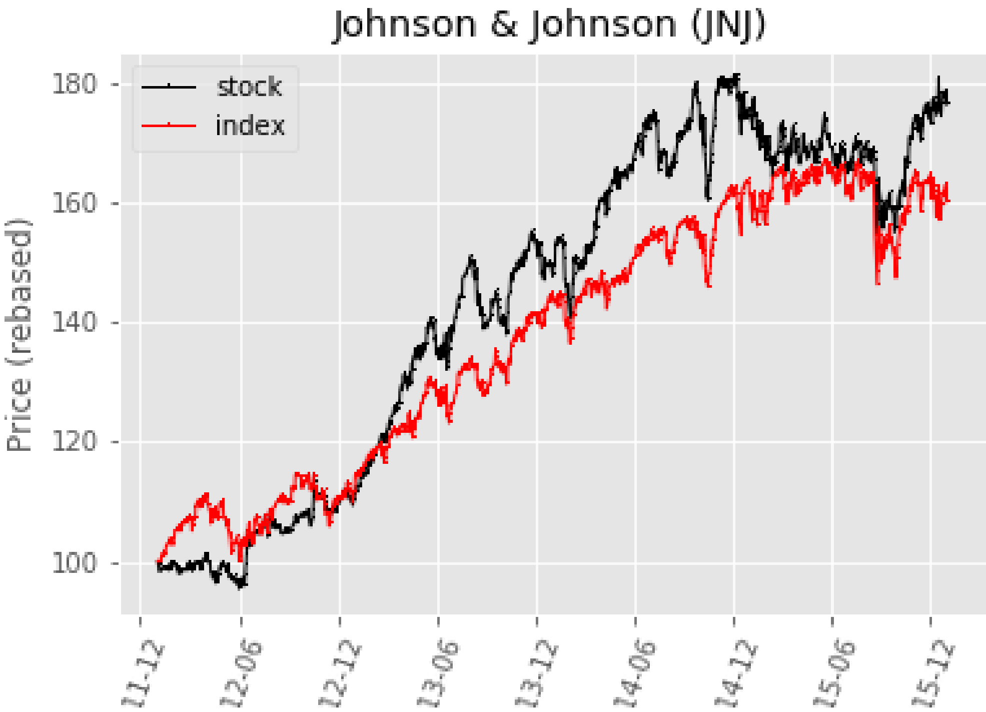


Figure 4.3 Daily prices for Johnson & Johnson relative to the S&P 500

It is clear from these charts that there has been a significant difference in the price performance of the individual stocks over time.

Apple for instance has deviated significantly from the index for lengthy periods with periods of both sustained underperformance and outperformance. Microsoft tracked the index more closely in the earlier part of the period before then outperforming for a time then reverting back towards the index performance (on a cumulative basis) and then entering a phase of significant outperformance. Johnson & Johnson by contrast shows less deviation from the index over the exploratory period. Pfizer shows a similar pattern to Johnson & Johnson, as does General Electric until late in the exploratory period where it begins to significantly outperform. Other stocks such as Exxon, and to a lesser extent Wells Fargo, show a consistent trend of outperformance over the period. Chevron is the only stock in the sample selected which shows a relatively consistent trend of underperformance.

The development objective is to create a deep neural network model which achieves outperformance but low volatility. The analysis above shows graphically that certain stocks show considerable variation in price and deviation from the index whereas others show less variation. Intuitively the stocks that provide consistent performance versus the index would be more suitable for model development however the line charts only provide an initial assessment of the problem and other factors are at play the metrics for which will be discussed further below.

4.3.2 Distribution of returns

A key consideration in minimising *tracking error* is the volatility of the daily returns of the stocks which may be re-weighted versus the index. As such an initial visual investigation into the distribution of returns was performed. To facilitate this histograms showing the frequency of the daily returns of each magnitude (25 basis point intervals) observed over the exploratory period have been created.

The histograms for the index and the three largest sample stocks (by market capitalisation) are shown below. The full suite of histograms for the all of the sample stocks are contained in Appendix II.

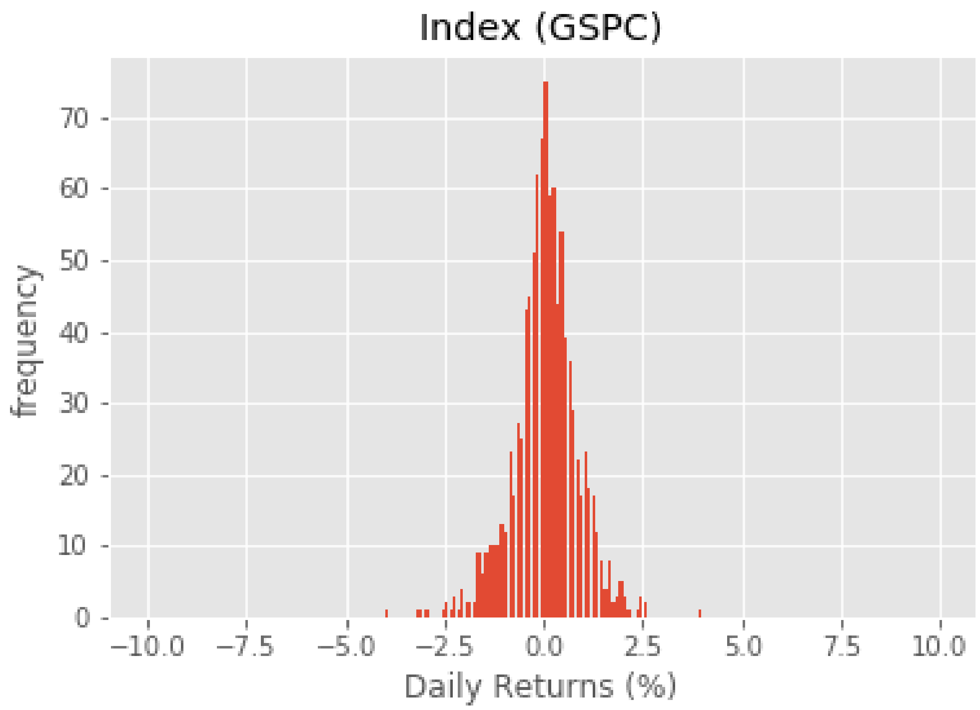


Figure 4.4 Distribution of daily returns for the S&P 500

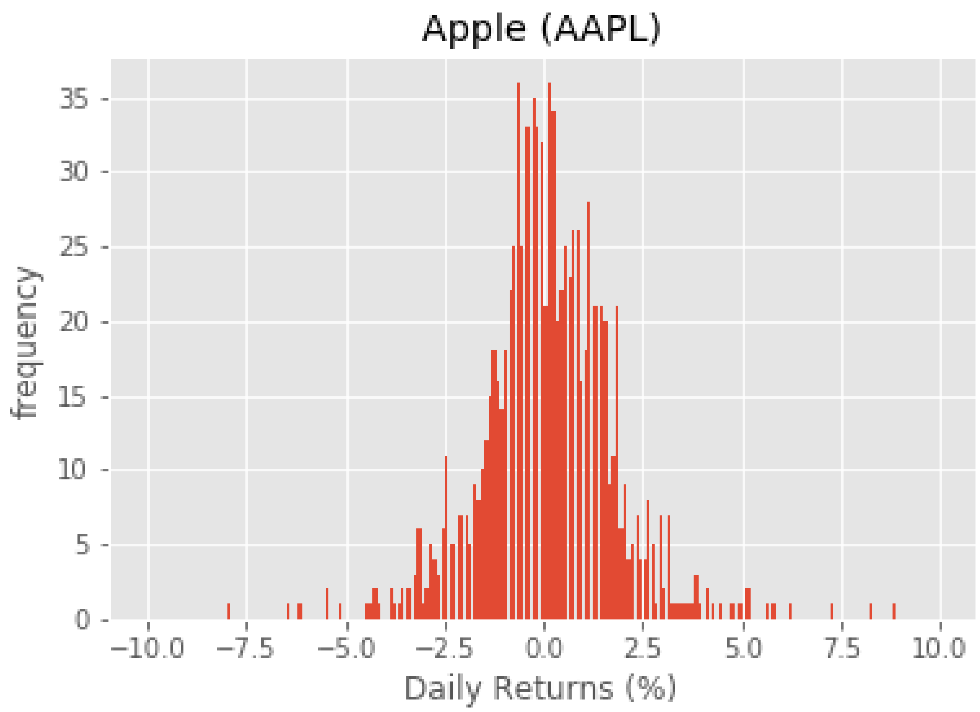


Figure 4.5 Distribution of daily returns for Apple

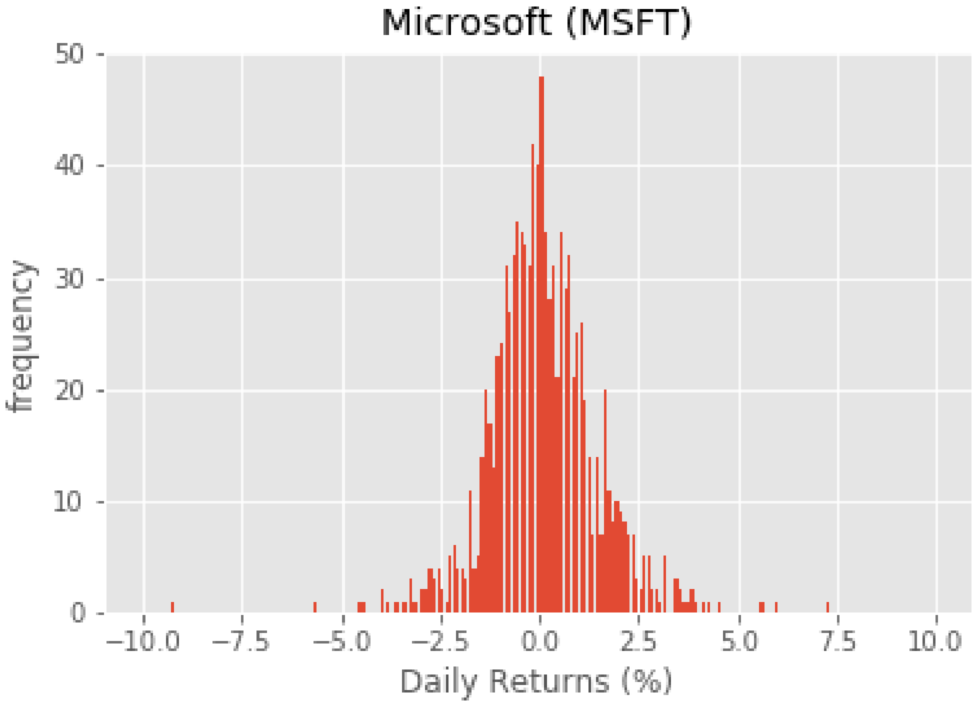


Figure 4.6 Distribution of daily returns for Microsoft

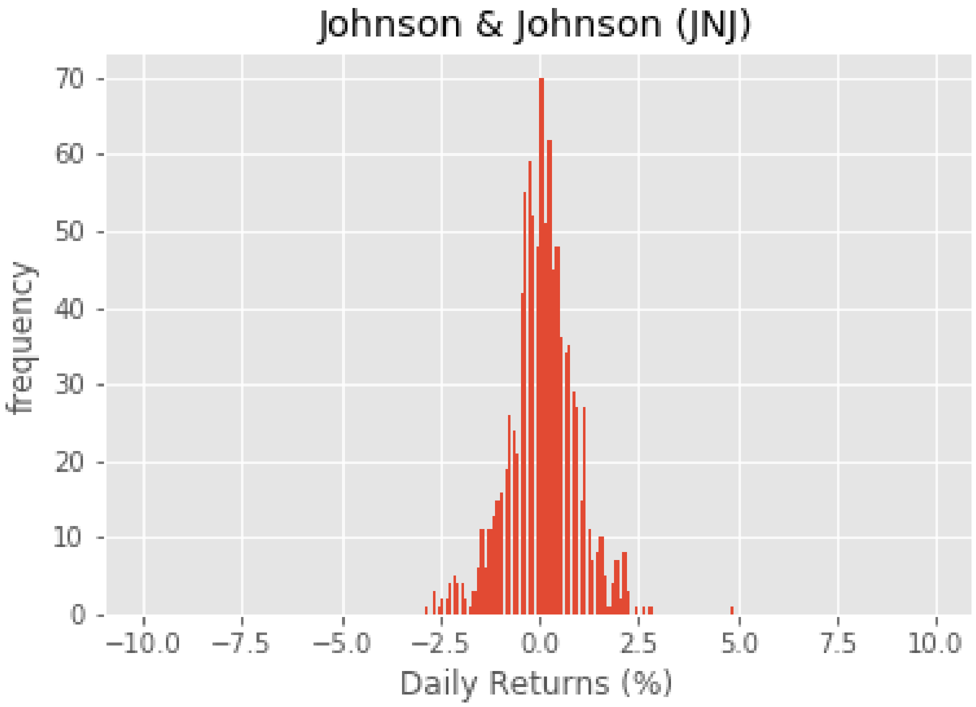


Figure 4.7 Distribution of daily returns for Johnson & Johnson

The most notable feature of the above distributions and those for the remaining stocks is that the daily returns are tightly clustered and rarely move outside a range of plus/minus 2.50 percent. Indeed all of the distributions visually resemble the normal distribution and appear to have minimal or slight negative skew meaning that the frequency of positive returns outweigh the frequency of negative returns. The distribution statistics will be tested in more detail in the section below.

The normal distribution type supports the use of measures such as *tacking error* and the *information ratio* and *share ratio* which will be used in the evaluation. These measures all utilise standard deviation as an input and as such require a reasonable degree of symmetry.

The findings with respect to the distribution of the returns is significant to the modelling exercise as the occurrence of frequent large daily movements would increase the risk of deviating significantly from the index. The fact that the returns are tightly clustered is a positive observation from the perspective of this study.

4.3.3 Daily return statistics

Table 4.1 and Table 4.2 below details the significant measures in terms of the magnitudes and distributions of the daily returns for both the index and the sample stocks.

Ticker	Minimum	Maximum	Mean
GSPC	-3.94	3.90	0.05
AAPL	-12.36	8.87	0.08
MSFT	-11.4	10.45	0.1
JNJ	-2.87	4.83	0.06
XOM	-4.73	5.52	0.01
AMZN	-11.00	15.75	0.15
JPM	-9.28	7.03	0.09
WFC	-5.90	5.78	0.08
GE	-4.06	10.80	0.08
T	-5.03	4.17	0.04
PG	-5.88	4.04	0.03
PFE	-4.47	4.20	0.06
CVX	-5.42	6.23	0.01

Table 4.1 Range and mean for daily returns over the exploration period

Ticker	Standard Deviation	Excess Kurtosis	Skew	Sharpe Ratio	Losses > 10 %	Gains > 10 %
GSPC	0.81	1.84	-0.21	0.06	0	0
AAPL	1.68	5.41	-0.21	0.05	1	0
MSFT	1.48	10.17	0.18	0.07	1	2
JNJ	0.85	1.7	-0.01	0.07	0	0
XOM	1.08	2.63	-0.02	0.01	0	0
AMZN	1.97	11.81	0.93	0.08	1	3
JPM	1.41	3.42	-0.18	0.06	0	0
WFC	1.14	2.39	-0.05	0.07	0	0
GE	1.16	8.09	0.96	0.06	0	1
T	0.93	2.50	-0.16	0.04	0	0
PG	0.90	4.27	-0.15	0.04	0	0
PFE	1.04	1.67	0.06	0.06	0	0
CVX	1.25	2.45	0.04	0.0	0	0

Table 4.2 Statistics for daily returns over the exploration period

The metrics presented above confirm the observations drawn from the charting exercise and provide a more precise quantitative view on the more important aspects of the data. Of particular interest is the absence of large daily gains and losses exceeding 10 percent from eight of the twelve sample stocks. Indeed of the four which do exhibit large moves the highest count was three out of over one thousand trading days studied.

As previously noted from the histograms the data is subject to only minimal skew with only two of the twelves stocks (Amazon and General Electric) displaying measures on this metric of significance. All of the stocks and the index display excess (positive) kurtosis measures indicating that larger and smaller than usual returns are more common than would be expected from a normally distributed variable. Three of the stocks (Amazon, Microsoft and General Electric) display particularly higher measures on this metric.

Of particular significance for this study is the noted differentials in terms of the *Sharpe ratio* for the index and the selected stocks. These differentials provide positive support for the potential to extract superior risk adjusted returns from appropriate stock selection.

4.4 Data preparation

The primary inputs for the investigation are raw financial data which are readily available in flat file format and as such require minimal or no preparation to facilitate analysis.

As discussed earlier in this study the inputs consist of measures derived from the raw data plus the raw data itself. The primary data preparation task consisted of computing the derived measures and consolidating all of these measures plus the non-derived inputs into a single csv format file which could then be imported to the selected analytics tools.

4.5 Model construction and validation

4.5.1 Single-layer model

4.5.1.1 Model parameters and training methodology

Building neural network models with strong predictive abilities requires effective hyperparameter optimisation. There are numerous parameters which must be set when creating a neural network including; selecting the appropriate optimization algorithm, learning rate, weight initialisation, the batch size, the number of epochs, the number of neurons in each layer and determining the optimal number of layers.

Hyperparameter optimisation is an area in which a considerable amount of research has been performed with numerous variants of scientific approaches being proposed including grid search, random search and Bayesian optimisation in addition to the manual judgemental approach. A sample of the literature noted included Bergstra and Bengio (2012) who consider the relative merits of random search versus grid search concluding in favour of the former. Bergstra *et al.* (2011) consider the topic of hyperparameter optimisation more generally noting that it has historically been the domain of human judgment due to the relative efficiency of this approach in computationally constrained environments but show that algorithmic approaches can produce better results. Dahl *et al.* (2013) utilise a Bayesian parameter optimization approach.

However, while developing an accurate model requires that the models hyperparameters be optimised the scope of this research does not extend to ensuring that all parameters are set to the best possible values. Rather it is required that the parameters be set to reasonable values and that these then be held constant in order to access the impact of increasing the models depth.

As such this development exercise will adopt a manual judgement approach selecting a finite set of parameter combinations and selecting from among these parameters by running the model against the data or the largest stock in the sample (Apple (AAPL)). This approach is effectively a limited grid search approach where human judgement has been applied to limit the size of the grid. Human judgment could be removed by utilising a wider range of values for the parameters which would have the effect of significantly increasing the computational resource requirements and as discussed above is beyond the scope of this project. The grid search parameter optimisation has been performed by reference to a single sample stock rather than optimising the parameters for each of the twelve stocks individually as the objective is to promote generalisation whereby the same model could be applied to any other out of sample stock selected from the index. Optimisation at the stock level is possible but is beyond the scope of this project and will be evaluated for possible future research.

This project utilises the Keras neural networks wrapper to train the models. While there are a large number of parameters which could be varied this researcher has deemed the batch size and the number of epochs to be the parameters for which effective parameter optimisation is significant to the project results. As such these two parameters are considered in the grid search. All other parameters have been maintained at the defaults contained in the Keras module or other settings deemed reasonable for the purpose of this research project. The optimisation of the number of layers and number of neurons per layer will be considered separately in the context of the development of the multi-layer models.

For the purposes of the single-layer model the number of neurons in the single hidden layer has been set to equal the number of inputs. This has been done to avoid introducing subjectivity into the model development process by permitting a one to one mapping.

In addition to these numeric hyperparameters key consideration relates to the selection of the optimisation algorithm to use. As can be seen from the literature review this modelling decision has received a lot of attention with various researchers concluding in favour of different models. However, no overall approach could be determined as dominant from the literature review. Given the absence of evidence in favour of an overall best optimisation algorithm this researcher has adopted the baseline Back Propagation model using standard Stochastic Gradient Descent (SGD) model. Evaluation using other more specialised or tailor made algorithms is possible but again is beyond the scope of this research.

A more general concern noted from the literature review was that the ratio of the number inputs to the number of samples in the training data should not be two high.

The imported data set for this study consists of 25 inputs and 1,006 samples and as such no issues are noted in this regard.

The model is being trained using daily data over the four year period to 31 December 2015. As discussed in Chapter 3 this period was selected to cover a period in which the market was trending and therefore provide an environment which supports the use of technical trading rules. As noted from the literature review such rules are less effective in more volatile environments. As such efficient use of this data is of critical importance. In order to avoid partitioning the data into training and validation sets this researcher has adopted a k-fold cross validation approach.

4.5.1.2 Data pre-processing

In addition to the selection of model parameters and training methodology consideration is required to be given to the level of data pre-processing if any to be performed. In this context the options were to either use the raw data without any pre-processing, to use data which has be standardised or to use data which has been both standardised and normalised.

The intuition is that since the neural network optimisation algorithm expects the input data to conform to the standard normal distribution it would not work well without data standardisation. However, applying the standardisation formula to change the data to conform to the parameters of the standard normal distribution (i.e. mean of zero and standard deviation of one) may have a distorting effect if the raw data itself is not itself at least approximately normally distributed. To this end the distribution of returns were measured and visualised during the data exploration phase of this project discussed above and found to be approximately normally distributed. Given the intuitive logic to using standardised data as opposed to raw data limited testing was deemed necessary to affirm the superiority of the standardisation approach. The 1-NN models for the largest stock in the sample (Apple (AAPL)) were run using both the raw data and standardised data affirming the superiority of the standardisation approach.

While the appropriateness of standardising the data was clear the decision as to whether to also normalise the data was more ambiguous. Normalising the data has the effect of expressing all of the input variables on a same scale (in the case on a scale from zero to one) and has as its objective the reduction of the risk that inputs which are measured on a larger scale could dominate the model but also reduces the impact of outliers in the data. Since the effect of the trade-off between these effects was difficult to assess intuitively and the ex-ante view was that the choice could potentially have a significant effect on the results it was determined appropriate to assess the 1-NN models across the full sample of stocks using data which has been standardised and also using data which has been both standardised and normalised.

4.5.1.3 Model training and validation

The validation results for the single-layer neural network models developed for each stock are shown in Table 4.3 below.

	Standardised		Standardised & Normalised	
Ticker	**Accuracy**	**Standard Deviation**	**Accuracy**	**Standard Deviation**
AAPL	49.81%	4.84%	47.82%	3.90%
MSFT	50.79%	4.41%	51.22%	5.01%
JNJ	53.89%	5.52%	52.67%	8.54%
XOM	49.90%	4.15%	50.90%	4.22%
AMZN	50.49%	4.07%	46.83%	4.61%
JPM	48.91%	4.17%	48.72%	3.82%
WFC	51.50%	3.84%	50.51%	6.09%
GE	53.78%	3.82%	54.08%	3.38%
T	49.31%	3.66%	49.80%	4.40%
PG	49.89%	6.32%	49.60%	3.38%
PFE	51.29%	6.35%	50.73%	5.80%
CVX	51.29%	5.21%	49.40%	5.97%

Table 4.3 Validation results for the single-layer model

The accuracy measure indicates the percentage of the samples in the validation set which the model correctly labelled as being ‘outperform’ or ‘underperform’ respectively and is the key measure of the predictive power of the model. The standard deviation measure quantifies the distribution of the dispersion of the accuracy figures calculated across each of the folds in the k-fold cross validation. In terms of interpretation of the results the higher the accuracy figure and the lower the standard deviation measure the stronger the predictive power of the model all else equal.

Since this is a classification task the risk that the results may be distorted by an imbalance in the data set must be considered and if necessary accounted for before commencing detailed evaluation of the results. Since the daily returns of both the index and the sample stocks are approximately normally distributed the number of positive and negative daily returns are well balances. Also the number of days when the individual stocks outperform or underperform has been assessed and found to be approximately balances and therefore no distortion effect in this regard is expected.

There are many factors influencing the performance of the overall stock market and the performance of individual stocks. The model developed utilises a combination of raw data and technical indicators derived from this raw data. These factors cannot quantity the precise relationship between the relative performance of the sample stocks and the selected index. Therefore accuracy measures approaching 100 per cent are not expected.

As a stock will either outperform or underperform the index on any given trading day this is a binary classification problem. To be better than random guessing and demonstrate predictive value the model must be correct more often than it is incorrect. Therefore, the models will be assessed as having predictive value if the accuracy measure exceeds 50 percent. However, to satisfy the research question the excess over the 50 per cent threshold must be sufficient to offset the trading costs associated with portfolio re-balancing. This second aspect of the assessment will be covered in the portfolio optimisation section at the end of this chapter.

Table 4.4 below illustrates the relative performance of the twelve 1-NN models trained using the standardised data compared to the same models trained using data which had been both standardised and normalised. For the purposes of this study the hold out set in the K-fold cross validation shall be referred to as the ‘validation data’ and the unseen data used for forward testing shall be referred to as the ‘test data’.

	Standardised	Standardised & Normalised
Accuracy > 50%	7	6
Accuracy < 50%	5	6

Table 4.4 Relative performance of 1-NN models on validation data

As can be seen from the table above the model trained on the standardised data produced models with predictive power in a majority of cases whereas the model which also utilised normalisation of inputs achieved predictive power in half the samples.

However, since the 1-NN model is the baseline model and will form the basis of the deeper models it is important to consider the relative number of instances in which each model achieved the higher level of predictive accuracy and also how well each model generalises to the unseen data. Table 4.5 below illustrates the same metrics as shown in Table 4.4 but in this case using the trained model to make predictions on the unseen test data.

	Standardised	Standardised & Normalised
Accuracy > 50%	10	6
Accuracy < 50%	2	6

Table 4.5 Relative performance of 1-NN models on test data

Both models achieve at least the same level of predictive accuracy when applied to the unseen data as achieved in training with the model trained on the standardised data again achieving superior predictive results over the model trained on data which was both standardised and normalised.

Finally, in terms of predictive accuracy the matrix in Table 4.6 below shows the frequency with which the respective models record the higher predictive accuracy across the twelve sample stocks. The model trained on the standardised data again demonstrates superior results to the model trained on data which had been both standardised and normalised.

	Validation Data	Test Data
Standardised	8	7
Standardised & Normalised	4	5

Table 4.6 Relative performance of 1-NN models on test data

As normalisation did not produce superior models the data pre-processing shall be restricted to standardisation. The 1-NN model trained on the standardised data will form the baseline model for the remainder of this study and the outputs from this model are analysed further below.

4.5.1.4 Evaluation of 1-NN model

Table 4.7 below summarises the predictive success of the models with respect to the individual stocks.

Ticker	Company	GICS Sector	Success?
AAPL	Apple Inc.	Information Technology	✗
MSFT	Microsoft Corporation	Information Technology	✓
JNJ	Johnson & Johnson	Health Care	✓
XOM	Exxon Mobil Corporation	Energy	✗
AMZN	Amazon.com Inc.	Consumer Discretionary	✓
JPM	JP Morgan Chase & Co.	Financials	✗
WFC	Wells Fargo & Company	Financials	✓
GE	General Electric Company	Industrials	✓
T	AT&T Inc.	Telecommunications Services	✗
PG	The Proctor & Gamble Company	Consumer Staples	✗
PFE	Pfizer Inc.	Health Care	✓
CVX	Chevron Corporation	Energy	✓

Table 4.7 Predictive success of 1-NN models by stock

For the purposes of this study a model is assessed as ‘success’ in terms of predictive power only if this predictive power is demonstrated on both the validation and test data. On the basis of this definition the 1-NN model development process was successful in achieving predictive power in the case of seven of the twelve stocks sampled.

It is interesting to observe that the predictive success of the model is not concentrated in a particular industry but rather diversified across all industries contained in the sample. In addition predictive power is not concentrated towards either the larger or smaller end of the capitalisation scale but spread across the continuum. This lends support to the potential to successfully develop a single baseline model across the entire sample and not to seek to refine the model hyperparameters for each stock individually.

4.5.2 Multi-layer model

The multi-layer model development process takes the final tuned single layer model and further develops the model by adding additional layers. The multi-layer model development process will assess the impact of adding increasing numbers of additional hidden layers.

The number of neurons will be decreased in each successive additional layer. The 2-NN, 3-NN and 5-NN models will reduce the neuron count by 5 in each successive layer, while the 10-NN reduces the neuron count at the rate of 2 per additional layer and the 15-NN models at the rate of 1 neuron by additional layer.

The validation results (percentage accuracy) are shown in Table 4.8 below for each of the multi-layer models along with the single layer models for ease of comparison followed by metrics on the impact of the number of hidden layers on the relative predictive power across the twelve samples in Table 4.9. Next Table 4.10 illustrates the extent to which the neural networks displayed predictive power across the varying depths by stock.

Ticker	1-NN	2-NN	3-NN	5-NN	10-NN	15-NN
AAPL	49.81%	50.90%	54.59%	52.19%	49.69%	51.39%
MSFT	50.79%	50.68%	52.97%	51.28%	49.51%	53.98%
JNJ	53.89%	53.69%	51.40%	51.29%	51.20%	48.71%
XOM	49.90%	49.51%	50.01%	49.71%	50.80%	50.40%
AMZN	50.49%	49.50%	52.09%	49.01%	47.91%	48.91%
JPM	48.91%	47.60%	48.42%	47.01%	47.62%	50.90%
WFC	51.50%	50.29%	53.19%	51.60%	51.42%	50.49%
GE	53.78%	52.68%	54.08%	48.51%	51.01%	49.28%
T	49.31%	51.20%	51.50%	51.00%	49.90%	49.00%
PG	49.89%	51.58%	50.87%	50.72%	49.11%	49.00%
PFE	51.29%	51.50%	50.80%	48.03%	50.02%	50.12%
CVX	51.29%	52.19%	50.19%	52.17%	49.61%	51.59%

Table 4.8 Validation results for the multi-layer models

In Table 4.8 above the green cells indicate the existence of predictive power whereas the red cells indicate the absence of predictive power.

Model	Predictive Power	No Predictive Power
1-NN	7	5
2-NN	9	3
3-NN	11	1
5-NN	7	5
10-NN	5	7
15-NN	7	5

Table 4.9 Impact of number of hidden layers on predictive power on validation data (by depth)

It is clear from Table 4.9 that the 3-NN neural network is the one with the most consistent results across the samples demonstrating predictive power in the case of eleven of the 12 samples. The deeper 10-NN model demonstrates the poorest results achieving predictive power in the case of only one of the twelve samples. The deepest 15-NN model achieves no improvement in predictive power as compared to the 1-NN model in terms of the number of samples correctly predicted however the relative strength of the predictions generated must also be considered before drawing definitive conclusions.

Ticker	Predictive Power	No Predictive Power
AAPL	4	2
MSFT	5	1
JNJ	5	1
XOM	3	3
AMZN	2	4
JPM	1	5
WFC	6	0
GE	4	2
T	3	3
PG	3	3
PFE	5	1
CVX	5	1

Table 4.10 Impact of number of hidden layers on predictive power on validation data (by stock)

It is clear from Table 4.10 that the predictive power of the models in validation vary considerably over the twelve stocks in th sample. For Wells Fargo all six models demonstrated predictive power whereas for JP Morgan only one of the models demonstrated predictive power. Broader industry factors would not appear to be the cause of the divergence across the stocks in the sample. Note that both JP Morgan and Wells Fargo fall within the same industry grouping. Also relative market capitalisation does not seem to be a determining factor. This would seem to indicate that additional factors specific to the individual stocks are exerting a large influence on the relative performance of the stock versus the index and are been driven to a lesser extent by global factors than assumed in the model inputs.

4.6 Model forward testing

While the results on the validation data were positive in that at least one of the six models with varying numbers of hidden layers displayed predictive power for each of the twelve sample stocks it is also necessary to use these models to make predictions on unseen test data before conclusions can be drawn as to their effectiveness. A model will be assessed as being successful only if it can demonstrate predictive power on both the validation data and the test data.

Firstly, Table 4.11 below illustrates the results obtained from the predictions made on the unseen data with green cells indicating the existence of predictive power and red cells indicating the absence of predictive power.

Ticker	1-NN	2-NN	3-NN	5-NN	10-NN	15-NN
AAPL	53.78%	52.59%	47.81%	45.02%	45.02%	47.81%
MSFT	51.39%	45.42%	51.79%	49.80%	49.40%	53.78%
JNJ	51.79%	56.57%	51.79%	49.80%	49.40%	55.38%
XOM	48.21%	43.82%	49.00%	49.40%	46.22%	44.22%
AMZN	54.58%	50.60%	50.60%	51.00%	51.00%	52.99%
JPM	50.20%	50.60%	51.00%	47.01%	51.79%	50.20%
WFC	51.79%	47.81%	53.39%	49.40%	52.59%	47.41%
GE	52.19%	46.61%	50.20%	48.61%	45.02%	52.19%
T	52.59%	48.21%	50.20%	52.59%	52.59%	48.21%
PG	45.42%	53.78%	52.19%	52.99%	47.81%	48.61%
PFE	52.59%	46.61%	44.62%	45.42%	51.00%	53.39%
CVX	50.20%	48.21%	41.43%	48.21%	48.61%	53.39%

Table 4.11 Testing results for the multi-layer models

As can be seen from Table 4.11 the models also performed well when applied to the testing data with thirty-eight of the seventy-two models demonstrating predictive power. However, the primary focus is on the models which demonstrated predictive power in both validation and testing and the results of this analysis is presented in Table 4.12 below. The results in Table 4.12 form the basis for the remainder of the analysis undertaken in this section.

Ticker	**1-NN**	**2-NN**	**3-NN**	**5-NN**	**10-NN**	**15-NN**
AAPL						
MSFT						
JNJ						
XOM						
AMZN						
JPM						
WFC						
GE						
T						
PG						
PFE						
CVX						

Table 4.12 Matrix of predictive success

As before the green cells in Table 4.12 above indicate the existence of predictive power whereas the red cells indicate the absence of predictive power.

Ticker	**1-NN**	**2-NN**	**3-NN**	**5-NN**	**10-NN**	**15-NN**
Validation	7	9	11	7	5	7
Testing	10	5	8	3	5	7
Both	7	3	7	2	2	4

Table 4.13 Number of models demonstrating predictive power

Table 4.13 provides an additional analysis to summarise how well the respective models trained and generalised to unseen data and achieved predictive success across the twelve sample stocks.

The most pertinent observation from the summary in Table 4.13 is that changing the number of hidden layers had a significant effect on the predictive success achieved by the models. Similarly, it is noted that there was no linear or directional trend observed in that the results did not gradually improve or deteriorate as each additional layer was added. This observation highlights the complexity of the process required to train a neural network and the difficulty of isolating the sources of the models predictive power.

The general observation however which can be drawn from the above results is that deeper models do not necessarily offer improved performance compared to the simpler models with the best results in terms of number of accurate models being achieved by the 1-NN and 3-NN models and with the worst results produced by the 10-NN model. The deepest 15-NN model developed in this study while outperforming the 10-NN model produced inferior results to the simplest 1-NN model.

Ticker	Predictive Power	No Predictive Power
AAPL	1	5
MSFT	3	3
JNJ	3	3
XOM	0	6
AMZN	2	4
JPM	1	5
WFC	3	3
GE	2	4
T	2	4
PG	3	3
PFE	3	3
CVX	2	4

Table 4.14 Impact of number of hidden layers on achieving predictive success on both validation data and test data

Similar to the results achieved in validation it is clear from Table 4.14 that the predictive power of the models in validation vary considerably over the twelve stocks in th sample. It is also noted that number of models achieving predictive success overall is lower than the number demonstrating predictive power in validation indicating that over-fitting may be an issue particularly in the deeper models.

4.7 Portfolio optimisation and results

The results of testing the single-layer and multi-layer models did not generate evidence that deeper models were capable of generating superior results when compared to the simpler single-layer model. While the 3-NN model matched the performance of the 1-NN model in the high level metric counting the number of sample stocks for which it displayed predictive power all other models underperformed the 1-NN model on this metric. Since no other model demonstrated superiority to the baseline 1-NN model the 1-NN model output shall be utilised for the evaluation performed in this section.

The first consideration in terms of the implementation of the optimisation strategy is to assess the associated trading costs. While there are other incidental costs the only cost considered impactful for the purposes of this analysis is the *bid-ask spread.* This information has been sourced from Yahoo Finance for a sample of three stocks provided an average spread of approximately 0.14% which will be used as an estimate for all stocks in the sample. The bid-ask spread tends to be low for liquid high-cap stocks and as the entire sample are large cap stocks this is deemed to be a reasonable basis. While 0.14% of the stock price appears low when considered for a single trade it acts as a significant drag on returns if traded frequently.

As evidenced in the literature review the area of portfolio optimisation is a well-established field and has been thoroughly researched in terms of both its theoretical underpinnings and its practical implementation. As such the objective of this research is not to ascertain the precise portfolio re-allocations to be made under the imposed constraints. Rather the objective is to assess the impact of adjusting the index weights based on the predictions made with respect to the sample stocks using the neural network models developed in the preceding sections of this chapter on the performance of the portfolio.

The neural network models can serve as reliable tools to effect change to the portfolio allocations only if they have been demonstrated to display predictive power. As such the portfolio optimisation exercise will proceed by reference to the seven stocks for which the 1-NN model displayed predictive power.

Each of the seven stocks will be considered individually with respect to a hypothetical USD 100 million portfolio re-balanced daily. To implement the enhancement to the index 1 per cent of the portfolio will be re-allocated to the sample stock if the neural network predicts outperformance such that the portfolio will comprise of USD 9.9 million tracking the index and USD 1 million tracking the sample stock. Conversely, if the neural network predicts underperformance a short position of USD 1 million in the sample stock will be taken with USD 101 million then tracking the index.

The performance of the hypothetical portfolio has been modelled using the data from the test period and assessed in both absolute terms using the US Dollar amount of the excess returns and in risk adjusted terms using the *information ratio* which considers the excess return generated per unit of risk (i.e. t*racking error*) incurred.

The result from the hypothetical portfolio simulation are detailed in Table 4.15 below. Note that the returns presented are the arithmetic mean of the daily returns over the simulation period. It would be more theoretically correct to utilise compounded returns and incorporate the impact of the previous days performance on the invested balance of the portfolio in the subsequent days calculations however given the small re-balancing percentage and the relatively small average daily price movements it is not necessary to add this additional level of complexity to the simulation to assess the aspects of the experiment which are relevance to the research objective**.**

Ticker	Excess Return (USD ‘000)	Active Return	Information Ratio
MSFT	(582)	-0.58%	-0.23
JNJ	(527)	-0.53%	-0.27
AMZN	290	0.29%	0.07
WFC	(100)	-0.10%	-0.04
GE	(291)	-0.29%	-0.17
PFE	137	0.14%	0.05
CVX	(37)	-0.04%	-0.01

Table 4.15 Portfolio simulation results

The objective is to achieve a positive excess return with higher information (IR) ratios been superior to lower ratios. While only two of the seven stocks generated excess returns in the simulation exercise it is important to emphasise the constraints in terms of model tuning which apply in the context of this research and which will be evaluated in Chapter 5. While the precise causes of the underperformance of the simulated portfolios versus the predictive power of neural network models ascertained in the model testing section above will be discussed in Chapter 5 the simple causal factor is that the models consider a binary classification of performance (i.e. ‘outperform’ or ‘underperform’) and do not specifically consider the level of outperformance or underperformance. While generally the stock movements are small and approximately normally distributed misclassifications on larger movements will cause the performance of the simulated portfolio to deviate and to be inferior to the that implied by the predictive power of the neural networks.

Clearly, given the small excess returns generated the models developed would not provide positive returns from implementing the enhanced indexing strategy. However, the predictive power of the models is likely to be understated due to under tuning of the model parameters. With more extensive tuning the predictive power of the models is anticipated to improve and if trading frequency is restricted excess returns after trading costs are considered likely to be achievable.

5.0 EVALUATION AND ANALYSIS

5.1 Introduction

This chapter will build on the detailed analysis and evaluation of the experimentation results documented in the previous chapter and evaluate and interpret these results in the light of what is already known and will discuss the impact of the scope limitations on the experiment results. It will then specifically examine the extent to which the research findings confirm or refute existing knowledge. Finally it will outline how the this study extends the existing body of knowledge.

5.2 Answer to the research question

The experimentation task was conducted using a rigorous process undertaking a series of sequentially ordered steps each of which utilised and built upon the results from the prior phase. As such the results produced are robust and provide an adequate and theoretically sound basis to both answer the research question and to form a foundation for future work.

The primary limitation of this research was with respect to the tuning of the neural network models. While consideration was given to the parameters which were deemed essential to produce results of sufficient quality to address the specific question there are many other aspects to the model development which were heuristically set to parameters deemed reasonable but which would require extensive fine tuning to achieve the maximum predictive power from the baseline model architecture. While the writer is satisfied that the results adequately support the contention that the baseline model is capable of generating reliable predictions further work would be required to fine tune the model parameters before it could be deployed in an investment management context.

With respect to the model development the key operational objective was to produce models which could accurately predict the outperformance or underperformance of select stocks relative to the benchmark index such that their weights could be moderately adjusted to improve the returns from the enhanced indexing strategy. The contention in the research question was that deep learning could be applied to achieve this. To this end neural networks with increasing numbers of hidden layers were developed. While the results are to an extent impact by under tuning in the baseline model it appears clear from the experimentation results that a large level of depth is not required to achieve the objective. In fact the simpler models outperformed the deeper models which would lead this researcher to believe that fine tuning and feature selection may play a greater part in generating greater predictive power than further increasing the number of hidden layers.

This leads to the second major limitation of this research project. While the model inputs were rigorously and systematically identified and cover the major aspects of technical trading they inevitable can only form a subset of the entire universe of possible technical features. The range of potential features include both global features common to all stocks along with features specific to individual industries or even to individual corporations. While the features selected were considered sufficient to generate a model with predictive power it became apparent during experimentation that they did not explain the full relationship and further analysis would be required in order to identify additional features to increase predictive power.

As observed in the preceding chapter the best neural network model developed was successful in demonstrating predictive power in the case of seven of the twelve stocks when the standard for success was set at a requirement to exceed 50 per cent accuracy in both validation and testing. In addition, at least one of the multi-layer models successfully demonstrated predictive power for eleven of the twelve stocks. A constraint imposed on this research in terms of time meant that the development of models with varying depth for individual stocks was not feasible but is an area for possible future work.

Of the seven stocks for which predictive power was achieved by the best performing neural network only two successfully translated into superior investment returns when utilised to guide the adjustment of the index weights. Further in the case of these two stocks the excess returns generated were insufficient to cover trading costs.

In this context it must first be understood that the neural network models are designed to assess if an individual stock will outperform or underperform the index but not to quantify the amount of such outperformance/underperformance. Since the stock returns are approximately normally distributed with larger than normal gains or losses being rare the fact that the outperformance/underperformance is not quantified would not impact the usefulness of the model so long as the predictive power reaches a level sufficient to negate the effect of outliers and additional monitoring was implemented to mitigate the risk of market shocks. While the limitations imposed on this research did not allow for the level of model tuning required to improve the predictive power to this extent the results of the experiments undertaken and insights from the literature review would indicate that it is feasible to do this.

The impact of trading costs significantly curtails the returns achievable from any form of active trading including enhanced indexing. The predictive power of the models developed was not sufficient to generate excess returns after trading costs however this can be partially addressed through improvements achieved through more extensive model tuning. However, achieving excess returns would nevertheless only be achievable if the frequency of portfolio re-balancing is restricted.

5.3 Confirmations and refutations of prior research

The literature review identified a considerable volume of research supporting the usefulness of neural network models in making financial market predictions and the results of this research confirm this finding. This research also confirmed the difficulties in configuring an training neural network models to achieve strong predictive power. This is impacted by both the complexity of the financial markets themselves and of the mathematics underlying the neural network based approach to deep learning. It is instructive that no single best neural network architecture or optimisation technique was observed from the review of the literature and this research would similarly indicate that perhaps no such best approach can exist. Instead achieving strong predictive power requires the identification of a combination of model architecture and optimisation techniques along with selection of the correct combination of input features. While the existing research provides valuable guidance on this process it appears that what is required is extensive computing resources to search through the entire range of possible combinations in order to find the optimal model for the portfolio optimisation.

The literature review also identified significant research and analysis focusing on the challenges in achieving excess returns from active portfolio management when trading costs are considered. The findings of this research did not succeed in producing evidence to refute the contention that trading costs eliminate much or all of the gains from active portfolio management. However, the experimentation results achieved provide some support for the contention that with further fine tuning of the models after cost excess returns can be achieved if trading frequency is restricted.

5.4 Extensions to existing knowledge

While there is considerable research into the effectiveness of neural networks and deep learning in the area of stock price prediction and related areas there is limited research relating to the application of these models to portfolio management and in particular limited research in the area of passive portfolio management.

This research project has investigated the feasibility of using such deep learning methods in an area of portfolio management where there is a degree of overlap with active trading allowing research performed in that area to be evaluated with respect to its potential application to passive portfolio management.

While the constraints imposed on this research project inevitable limit the level of model tuning which can be performed and hence the predictive strength of the models developed the results achieved are positive in terms of establishing that deep learning models can be of value in a passive portfolio management context and with further development have the potential to achieve improved results.

Overall this research project has established a framework in which deep learning techniques can be adapted to the requirements of passive portfolio management. It has also designed baseline neural network models which have been deployed in order to guide the re-weighting of stocks in an enhanced indexing strategy and which with further fine tuning offer a viable deep learning based approach with practical applications.

6.0 CONCLUSION

6.1 Introduction

This chapter will provide an overview of the work performed, the results achieved and the conclusions reached. It will address each of these in turn in the same order and using the same structure as employed in executing this research. It will then proceed to assess the contributions and impact made by this research and outline areas for future work and recommendations arising from the research findings.

6.2 Problem definition

This research project set out to ascertain as to whether deep learning techniques could be utilised to improve the risk adjusted returns from enhanced indexing strategies. This required the use of a multi-step sequential model development process in which candidate features were initially identified, baseline neural network model built and then further developed and finally applied to adjust the index weights to implement the enhanced indexing strategy.

6.3 Design, experimentation, evaluation and results

The initial stage of the research project consisted of a review of the literature to ascertain the current knowledge with respect to both portfolio management and the implementation of portfolio management strategies. The literature review also investigated current research in the area of securities selection and the categorisation of stocks based on their attributes. It also evaluated the body of existing research into the use of deep learning techniques in financial market predictions and in portfolio management. The literature review then proceeded to investigate the available research into the technical aspects of neural network implementations.

Having complete the review of the literature the focus of the study moved to primary research to design the model and implement the models which would be used to test the hypothesis. This initially involved the identification of suitable sources from which to obtain the required raw data. Identifying suitable time periods in which to train and forward test the models was an important consideration. The objective was to select a period in which the market was trending and which was sufficient homogenous as to avoid the necessity to accommodate factors not directly relevant to the research objective but also of adequate length to provide sufficient data to train the model.

The next consideration was to select a suitable benchmark index and a suitable sample of stocks from this index to use as the basis for the model development. Selection criteria to achieve this were developed and on this basis the S&P 500 was selected as it was a highly liquid well researched index based in a large developed domestic economy. Using similar reasoning a sample of twelve stocks were selected from among the twenty largest index constituents. A list of candidate features were then derived for use in model development. This feature selection process was guided by the findings from the literature review and comprised of a combination of raw financial markets data and technical features derived from this data.

Following on from the determination and documentation of the model features and design work commenced on the construction of the neural network models. An initial baseline single-hidden layer model was developed and used as the basis for the construction of successively deeper models. Work was performed to ensure that the parameter selections and data pre-processing decisions deemed critical to the development of models with successful predictive power were appropriately implemented. The full optimisation of all model parameters was beyond the scope of this research and as such all other parameters were heuristically established at settings deemed reasonable.

The neural network models developed achieved a level of predictive power sufficient to support the conclusion that such models are capable of accurately the outperformance or underperformance of the selected stock relative to the benchmark index. The analysis of the results supported the contention that improved predictive results could be achieved from these models on completion of further fine tuning of model parameters.

The under tuning of the neural network models was found to limit the potential of these models in successfully guiding the index re-weighting decisions to support the implementation of the enhanced indexing strategy. However, the level of success achieved in the experimentation phase was considered sufficient to conclude that such models with further tuning are capable of guiding this process. The challenges with respect to generating excess returns from active trading after accounting for trading costs was noted from the literature review. The predictive power of the models developed was not sufficient to refute the contention that trading costs would eliminate most or all of the excess returns generated. However, the results achieved are considered to support the view that excess returns net of trading costs can be generated provided that trading frequency is restricted.

6.4 Contributions and impact

As noted from the literature review there is limited research into the application of deep learning techniques to portfolio management and in particular to passive portfolio management.

This research project has developed and implemented a model for enhanced indexing taking and has demonstrated that deep learning techniques have the potential to improve the implementation of such portfolio management strategies. The research also affirmed the effectiveness of neural network models in financial market predictions which has been demonstrated by other researchers in other areas of financial market investing as discussed in the literature review.

This research has also established a framework and set of baseline models which can be used as the basis of future research and practical deployment.

6.5 Future work and recommendations

As discussed above the most significant limitation on this research project was the with respect to the tuning of the neural network models. Future work will be undertaken to optimise all of the model parameters and thereby maximise the predictive power of the models developed. As this is a computationally intensive task this evaluation will be performed using GPUs rather than the CPU approach employed in this research project.

A second areas for future work will be in relation to the evaluation of a more extensive range of model inputs. There is an almost infinite combination of inputs and derived inputs which could be used to train the models. As an initial step in guiding this aspect of the future work an exercise will be undertaken to ascertain the relative impact of the features in the current model with a view to determining which inputs provide the largest contribution to the models predictive power.

This research project was based on a sample of twelve stocks from an single benchmark index. In order to assist in the identification of features which explain a larger part of the relationship between historic stock market data and measures derived from this data the scope of the research can be extended to cover a wider range of stocks and also perform similar analysis with respect to other major indices.

In addition model refinement would benefit from training and testing on additional time periods with evaluation performed over longer periods with additional complexity being incorporated into the models to take account of the differing macro-economic environment and market conditions which this step would introduce. The model will also be evaluated against shorter sub-periods to investigate the impact of seasonality and use this information to add further sophistication which will facilitate the strengthening of predictive power.

7.0 BIBLIOGRAPHY

- Bahrammirzaee, A. (2010). A comparative study of artificial intelligence applications in finance: artificial neural networks, expert system and hybrid intelligent systems. *Neural Computing and Applications, Volume 19 Issue 8, November 2010 Pages 1165-1195*
- Baker, M. Wurgler, J. (2007). Investor Sentiment in the Stock Market. *The Journal of Economic Perspectives, Volume 21 Issue 2 Spring 2007, Pages 129-151*
- Bandopadhyaya, A. Jones, A. (2008). Measures of Investor Sentiment: A Comparative Analysis Put-Call Ratio Vs. Volatility Index. *Journal of Business & Economics Research, Volume 6 Issue 8, August 2008 Pages 27 – 34*
- Becker, R. Junsoo, L. Benton, G. An empirical analysis of man reversion of the S&P 500's P/E ratios. *Journal of Economics & Finance, Volume 36 Issue 3, July 2012 Pages 675-690*
- Bergstra, J. Bardenet, R. Bengio, Y. Kegl, B. (2011). Algorithms for Hyperparameter optimisation. *Proceedings of Advances in Neural Information Processing Systems 24 (NIPS 2011).* Retrieved from: http://papers.nips.cc/paper/4443-algorithms-for-hyper-parameter-optimization
- Bergstra, J. Bengio, Y. (2012) . Random Search for Hyper-Parameter Optimization. *Journal of Machine Learning Research, Volume 13 3/1/12, Pages 281-305*.
- Capaul, C. Rowley, I, Sharpe, W. (1993). International Value and Growth Stock Returns. *Financial Analysts Journal, Volume 49 Issue 1 Jan-Feb 1993 Pages 27 – 36*
- Chavaranakul, T. Enke, D. (2008). Intelligent technical analysis based equivolume charting for stock trading using neural networks. *Expert Systems with Applications, Volume 34 Issue 2, February 2008 Pages 1004-1017*
- Chen, Y. Ohkawa, E. Mabu, S. Shimada, K, Hirasawa, K. (2009). A portfolio optimization model using Genetic Network Programming with control nodes. *Expert Systems and Applications, Volume 36 Issue 7, September 2009 Pages 10735-10745*
- Chiarella, C. He, X. Hommes, C. (2006a). A dynamic analysis of moving average rules. *Journal of Economic Dynamics and Control. Volume 30 Issues 9-10, September-October 2006 Pages 1729-1753*
- Chiarella, C. He, X. Hommes, C. (2006b). Moving average rules as a source of market instability. *Physica A: Statistical Mechanics and its Applications, Volume 370 Issue 1, 1 October 2006 Pages 12-17*

- Dahl, G. Sainath, T. Hinton. G. (2013). Improving Deep Neural Networks for LVCSR using Rectified Linear Units and Dropout. *2013 Conference on Acoustics, Speech and Signal Processing (ICASSP), Pages 8609 -8613*
- Dose, C. Cincotti, S. (2005). Clustering of financial time series with application to index and enhanced index tracking portfolio. *Physica A: Statistical Mechanics and its Applications, Volume 355 Issue 1, 1 September 2015 Pages 141-151*
- Edirisinghe, N. Zhang, X. (2007). Generalized DEA model of fundamental analysis and its application to portfolio optimization. *Journal of Banking & Finance, Volume 31 Issue 11, November 2007 Pages 3311 – 3335*
- Evans, C. Pappas, K. Xhafa, F. (2013). Utilizing artificial neural networks and genetic algorithms to build an algo-trading model for intra-day foreign currency speculation. *Mathematical and Computer Modelling, Volume 58 Issue 5, September 2013 Pages 1249 – 1266*
- Fama, E. (1965). The Behavior of Stock-Market Prices. *The Journal of Business, Volume 38 Issue 1, January 1965, Pages 34-105*
- Fama, E. French K. (2007). The Anatomy of Value and Growth Stock Returns. *Financial Analysts Journal, Volume 63 Issue 6 November-December 2007 Pages 44 – 54*
- Fernandez, A. Gomez, S. (2007). Portfolio selection using neural networks. *Computers & Operations Research, Volume 34, Issue 4, April 2007 Pages 1177-1191*
- Freitas, F. De Souza, A. de Almeida, A. (2009). Prediction-based portfolio optimization model using neural networks. *Neurocomputing, Volume 72 Issue 10, June 2009 Pages 2155 – 2170*
- Frino, A. Gallagher, D. (2001). Tracking S & P 500 Index Funds. *The Journal of Portfolio Management, Volume 28 Issue 1, Fall 2001 Pages 44-55*
- Kara, Y. Boyacioglu, M. Baykan, O. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange. *Expert Systems with Applications, Volume 38 Issue 5 Ma 2011 Pages 5311-5319*
- Ko, P. Lin, P. (2008). Resource allocation neural network in portfolio selection. *Expert Systems and Applications, Volume 35 Issue 1, July-August 2008 Pages 330-337*
- Lawrence, S. Lee Giles, C. Tsoi, A. (1997). Lessons in Neural Network Training: Overfitting May be Harder than Expected. *Proceedings of the Fourteenth National Conference on Artificial Intelligence, AAAI97 AAAI Press 1997 Pages 540-545*
- Liu, Q. Guo, Z. Wang, J. (2012). A one-layer recurrent neural network for constrained pseudoconvex optimization and its applications for dynamic

portfolio optimization. *Neural Networks, Volume 26, February 2012 Pages 99-109*

- Markowitz, H. (1952). Portfolio Selection. The *Journal of Finance, Volume 7 Issue 1, March 1952 Pages 77-91*
- Marriage, M. (2016, September 11). Passive funds take third of US market. *Financial Times.* Retrieved from: https://www.ft.com/content/4cdf2f88-7695-11e6-b60a-de4532d5ea35?mhq5j=e1
- Penman, S. (1996). The Articulation of Price-Earnings-Ratios and Market-to-Book Ratios and the Evaluation of Growth. *Journal of Accounting Research, Volume 34 Issue 2 Autumn 1996 Pages 235 – 259*
- Pisani, B. (2015). Trader Talk: Value versus growth is the big investment debate for 2016. Retrieved from: http://www.cnbc.com/2015/12/23/value-versus-growth-is-the-big-investment-debate-for-2016.html
- Poppick, S (2016). The new passive-active management compromise. Retrieved from: http://www.cnbc.com/2016/05/31/active-versus-passive-investing-this-is-better.html
- Scott, J. Stumpp, M. Xu, P. (2003). News, Not Trading Volume, Builds Momentum. *Financial Analysts Journal, Volume 59 Issue 2, March-April 2003 Pages 45 – 54*
- Sharpe, W. (1964). Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance, Volume 19 Issue 3, September 1964 Pages 425-442*
- Sharpe, W. (1966). Mutual Fund Performance. *The Journal of Business, Volume 39 Issue 1 Part 2: Supplement on Security Prices, January 1966 Pages 119-138*
- Sharpe, W. (1994). The Sharpe Ratio. *The Journal of Portfolio Management, Volume 21 Issue 1 Pages 49-58*
- Shen, W, Guo, X. Wu, C. Wu, D. (2011). Forecasting stock indices using radial basis function neural networks optimized by artificial fish swarm algorithm. *Knowledge-Based Systems, Volume 24 Issue 3, April 2011 Pages 378 – 385*
- Sortino, F. Price, L. (1994). Performance Measurement in a Downside Risk Framework. *The Journal of Investing, Volume 3 Issue 3, Fall 1996 Pages 59 – 64*
- Treynor, J. Black, F. (1973). How to Use Security Analysis to Improve Portfolio Selection. *The Journal of Business, Volume 46 Issue 1, January 1973 Pages 66-86*
- Vanstone, B. Finnie, G. (2009). An empirical methodology for developing stockmarket trading systems using artificial neural networks. *Expert Systems with Applications, Volume 36 Issue 3, April 2009 Pages 6668 – 6680*

- Vardi, N. (2016, January 13). Top Quant Funds Stand Out with Good 2015. *Forbes*. Retrieved from: https://www.forbes.com/sites/nathanvardi/2016/01/13/top-quant-hedge-funds-stand-out-with-good-2015/#24e3361d5c60

- Wilder, J. (1978). New Concepts in Technical Trading Systems. Trend Research

- Wong, W. Manzur, M. Chew, B. (2003). How rewarding is technical analysis? Evidence from Singapore stock market. *Applied Financial Economics. Volume 13 Issue 7 Pages 543 – 551*

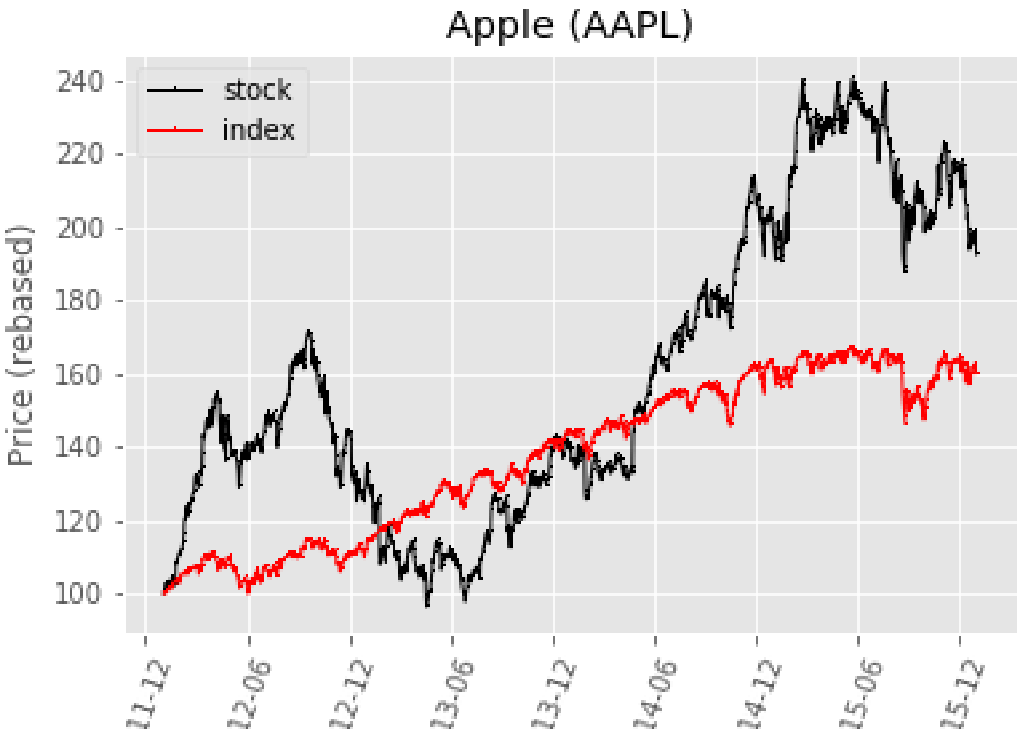
- Wu, L. Chou, S. Yang. C. Ong, C. (2007). Enhanced Index Investing Based on Goal Programming. *The Journal of Portfolio Management, Volume 33 Issue 3, Spring 2007 Pages 49-56*

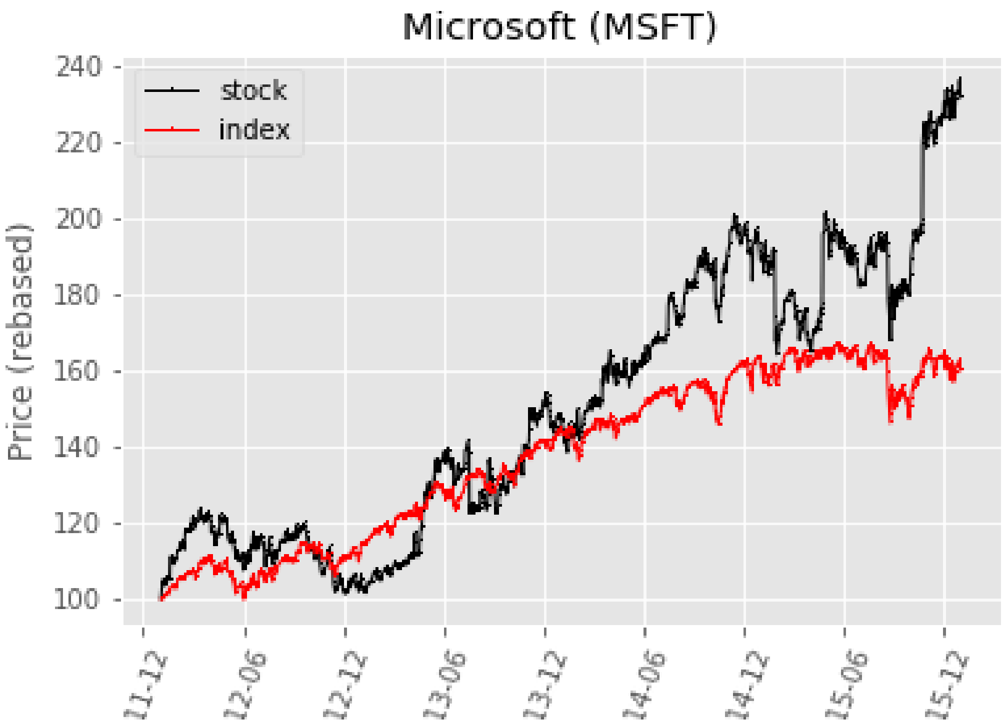
- Yu, L. Wang, S. Lai, K. (2008). Neural network-based mean-variance-skewness model for portfolio selection. *Computers & Operations Research, Volume 35 Issue 1, January 2008 Pages 34-46*

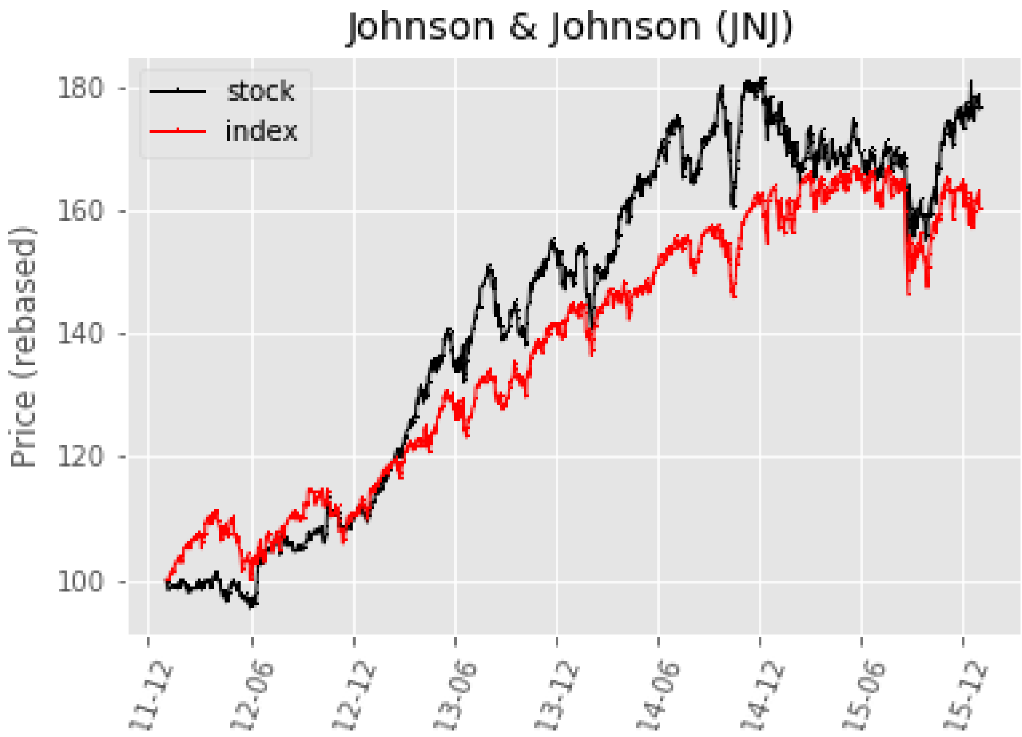
- Zhang, Y. Wu, L. (2009). Stock market prediction of S&P 500 via combination of improved BCO approach and BP neural network. *Expert Systems with Applications, Volume 36 Issue 5, July 2009 Pages 8849 – 8854*

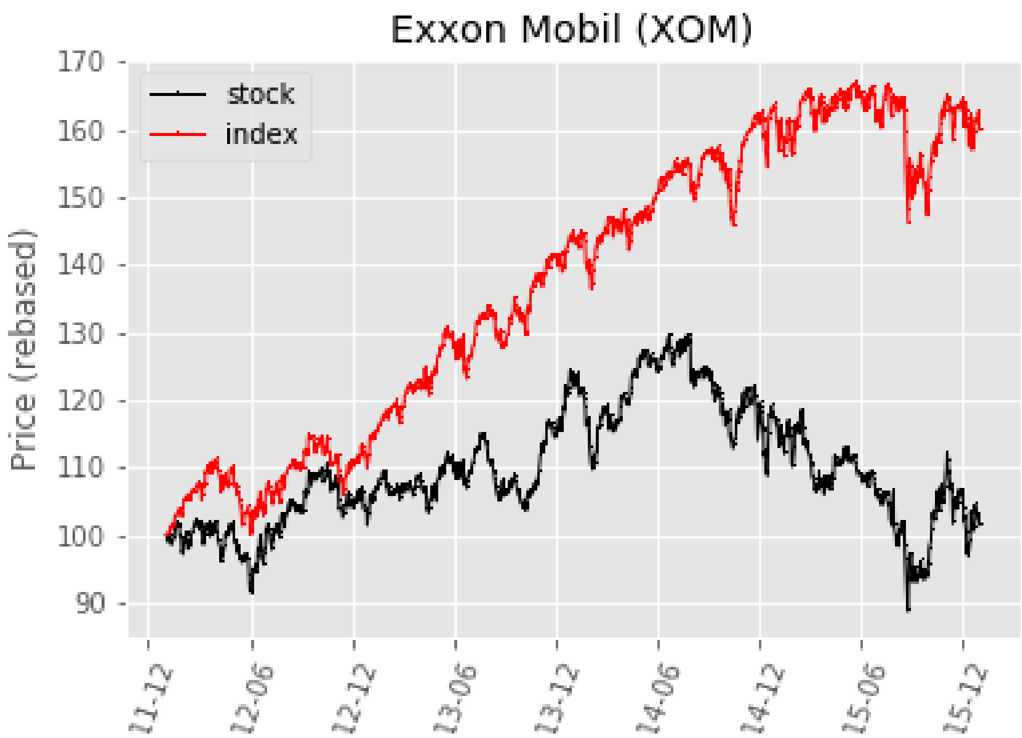
- Zhu, H. Wang, Y. Wang, K. Chen, Y. (2011). Particle Swarm Optimization (PSO) for the constrained portfolio optimization problem. *Expert Systems with Applications, Volume 38 Issue 8, August 2011 Pages 10161-10169*

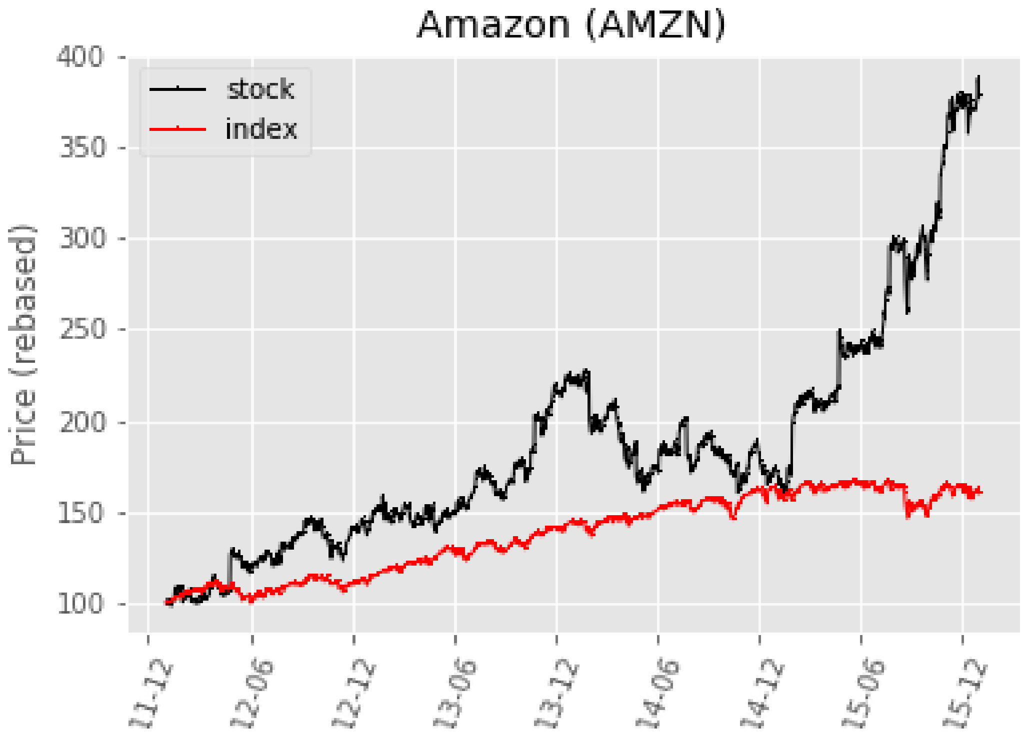
8.0 APPENDIX I – PRICE PERFORMANCE OF SAMPLE STOCKS

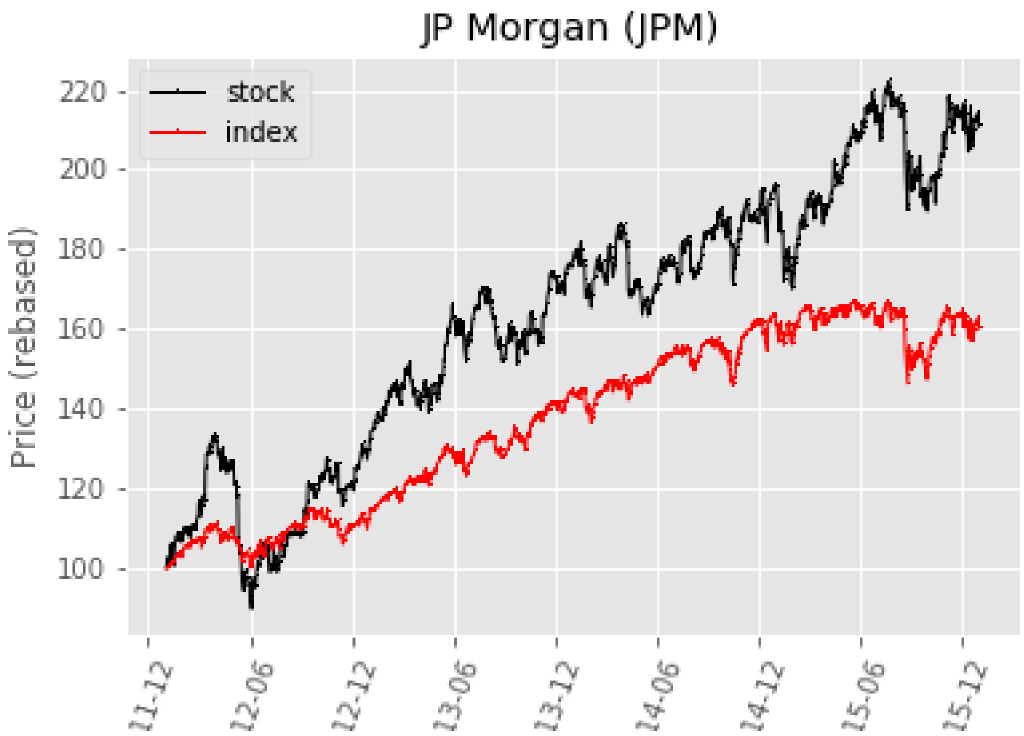


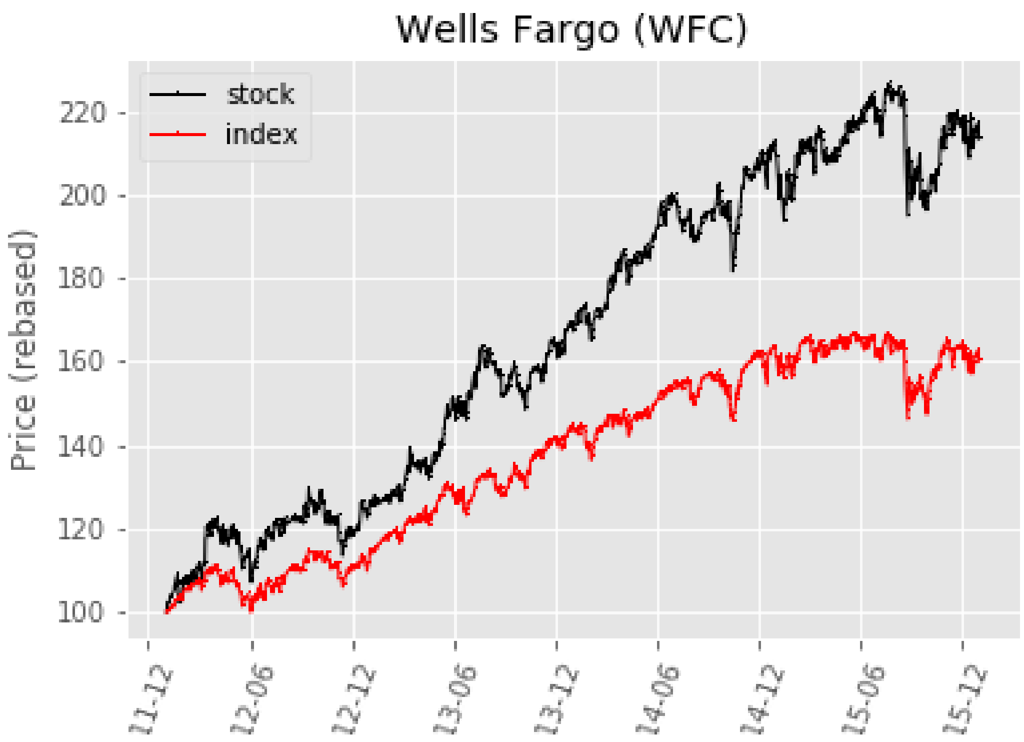



Johnson & Johnson (JNJ)
stock
index
Price (rebased)
180
160
140
120
100
11-12
12-06
12-12
13-06
13-12
14-06
14-12
15-06
15-12

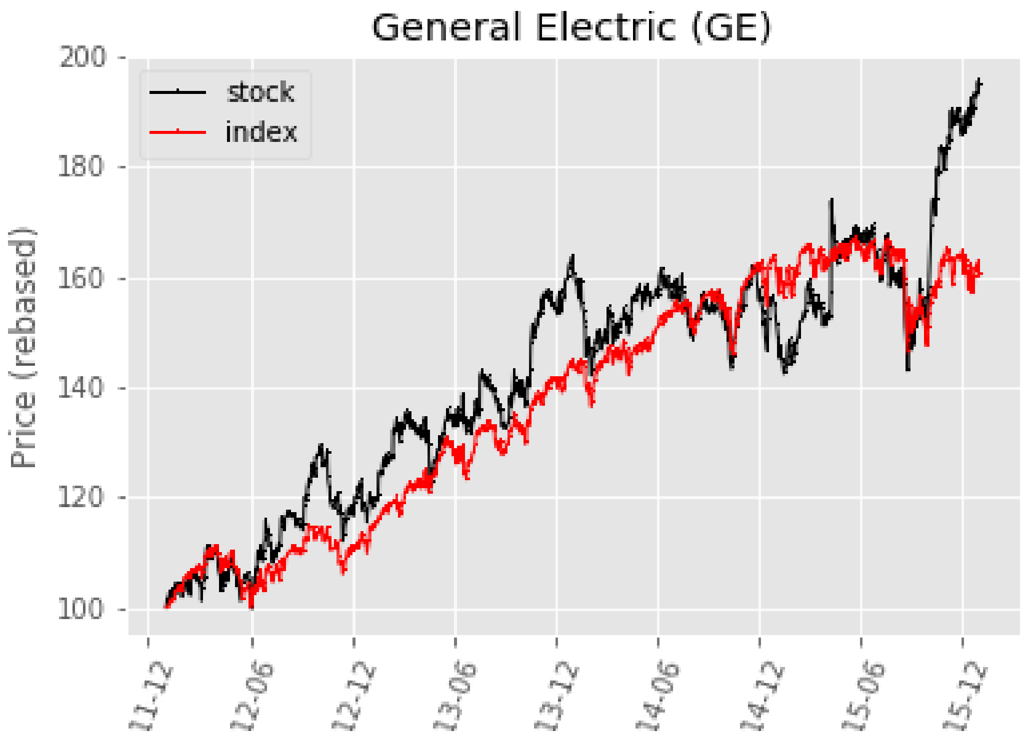

Exxon Mobil (XOM)
stock
index
Price (rebased)
170
160
150
140
130
120
110
100
90
11-12
12-06
12-12
13-06
13-12
14-06
14-12
15-06
15-12


Amazon (AMZN)
Price (rebased)
stock
index
400
350
300
250
200
150
100
11-12
12-06
12-12
13-06
13-12
14-06
14-12
15-06
15-12

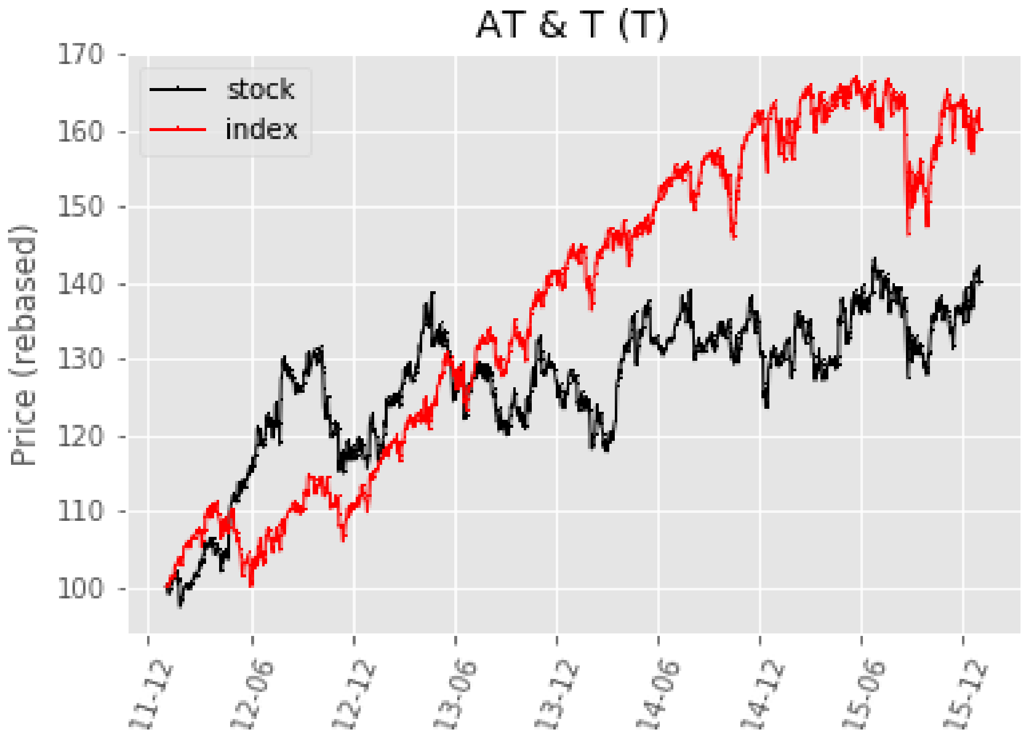

JP Morgan (JPM)
Price (rebased)
stock
index
220
200
180
160
140
120
100
11-12
12-06
12-12
13-06
13-12
14-06
14-12
15-06
15-12

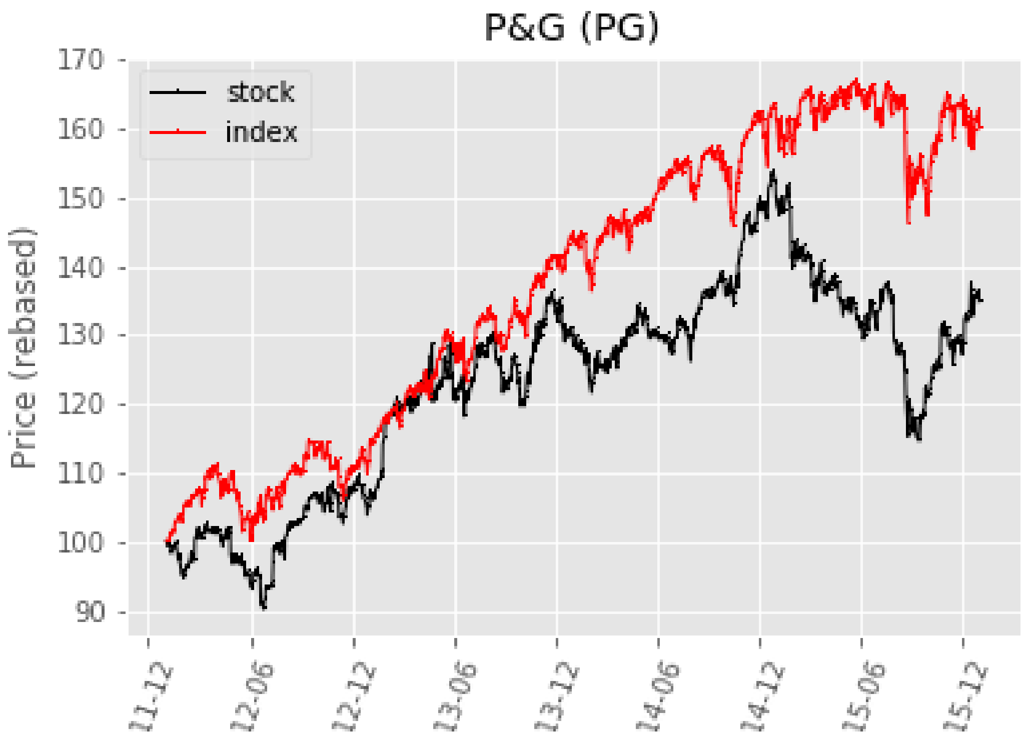


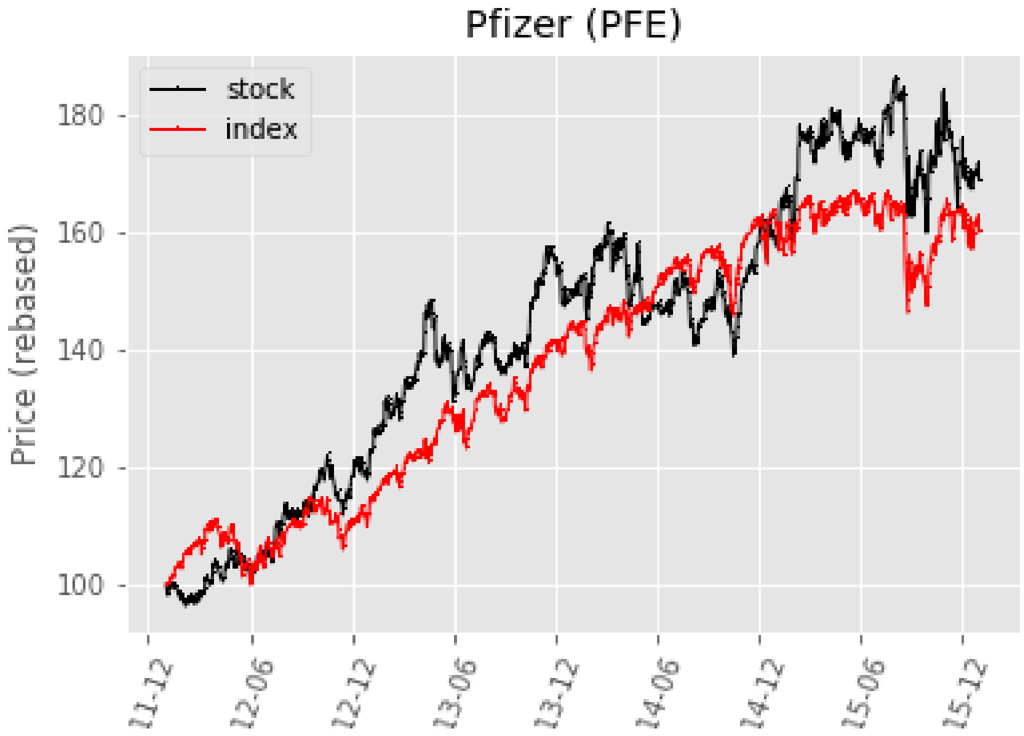
Wells Fargo (WFC)
stock
index
Price (rebased)
220
200
180
160
140
120
100
11-12
12-06
12-12
13-06
13-12
14-06
14-12
15-06
15-12

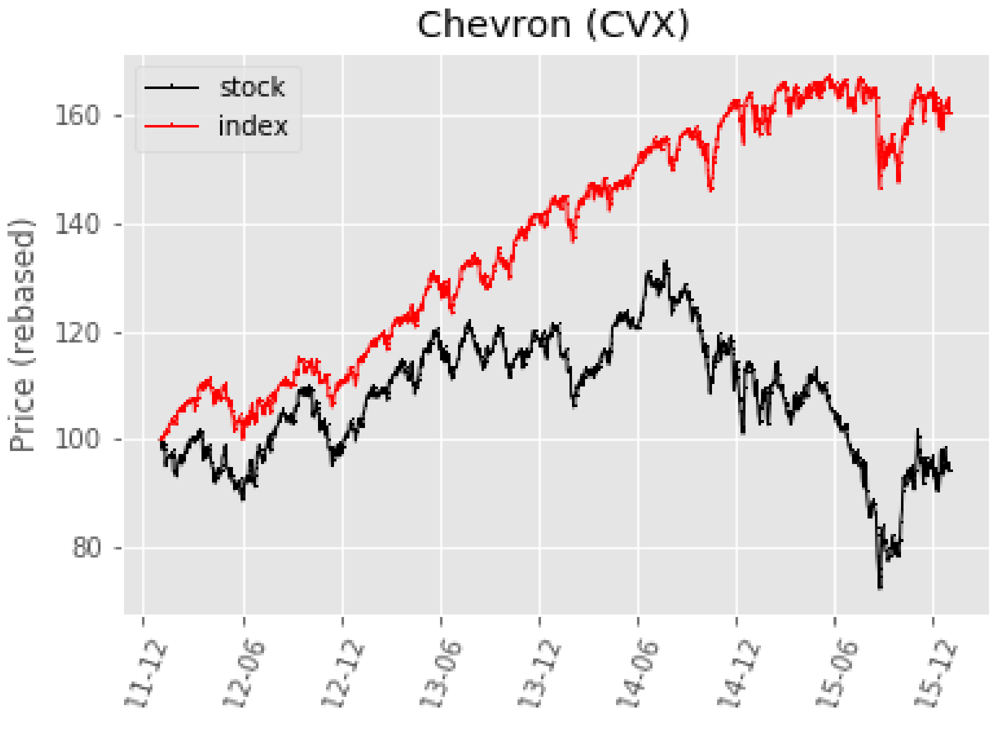


General Electric (GE)
stock
index
Price (rebased)
200
180
160
140
120
100
11-12
12-06
12-12
13-06
13-12
14-06
14-12
15-06
15-12

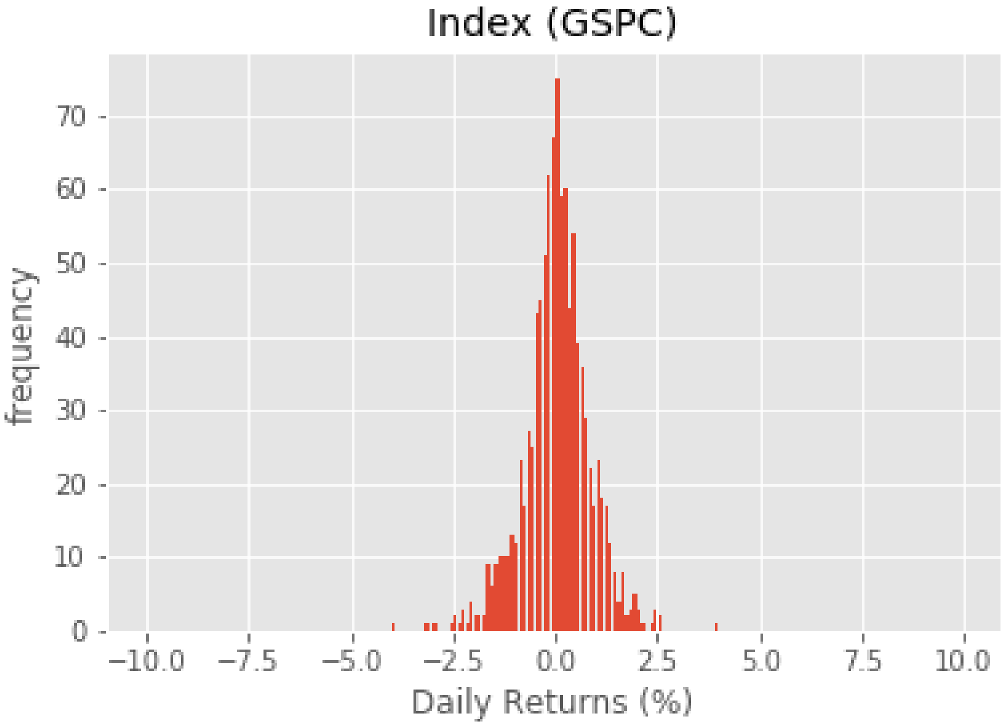

AT & T (T)
stock
index
Price (rebased)
170
160
150
140
130
120
110
100
11-12
12-06
12-12
13-06
13-12
14-06
14-12
15-06
15-12

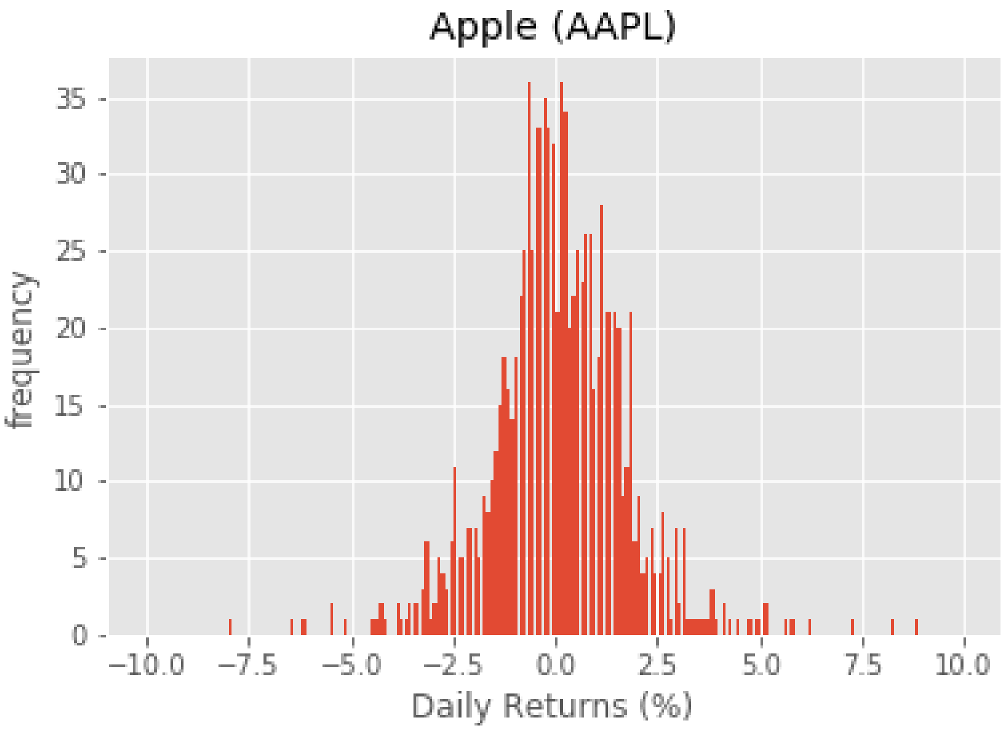

P&G (PG)
stock
index
Price (rebased)
170
160
150
140
130
120
110
100
90
11-12
12-06
12-12
13-06
13-12
14-06
14-12
15-06
15-12

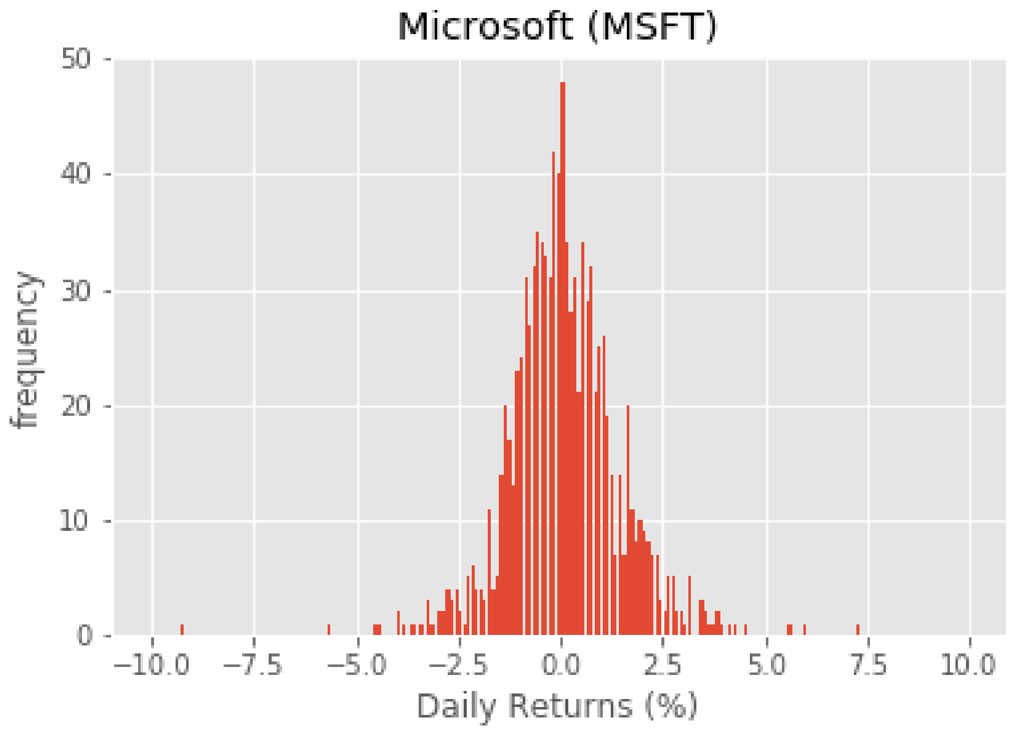

Pfizer (PFE)
stock
index
Price (rebased)
180
160
140
120
100
11-12
12-06
12-12
13-06
13-12
14-06
14-12
15-06
15-12

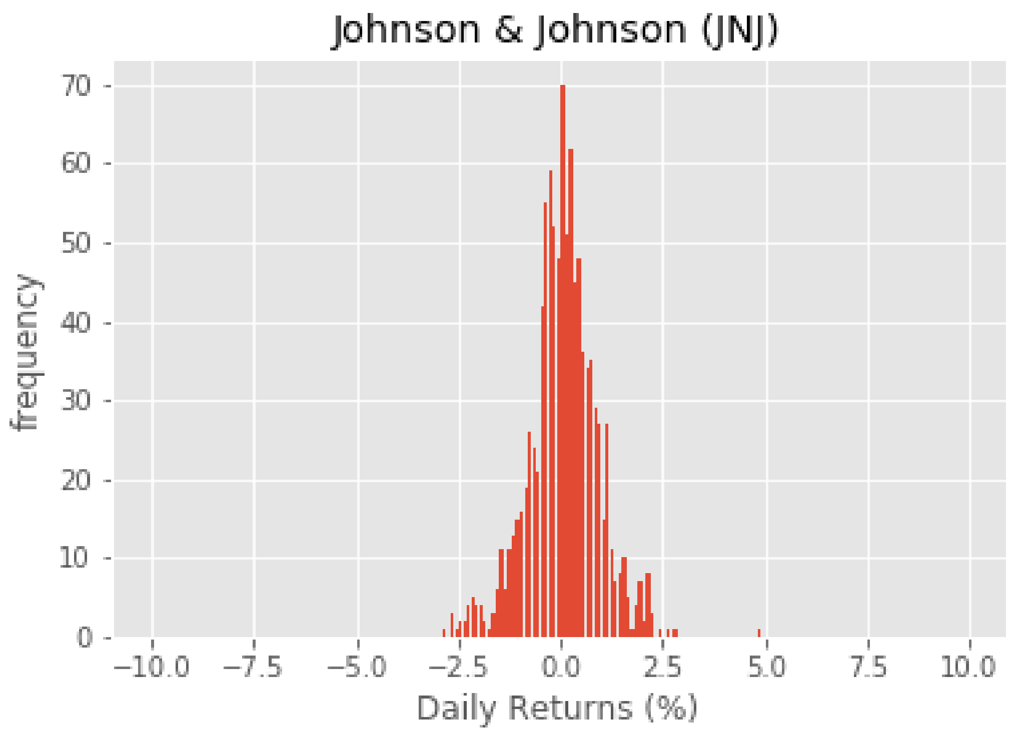

Chevron (CVX)
stock
index
Price (rebased)
160
140
120
100
80
11-12
12-06
12-12
13-06
13-12
14-06
14-12
15-06
15-12

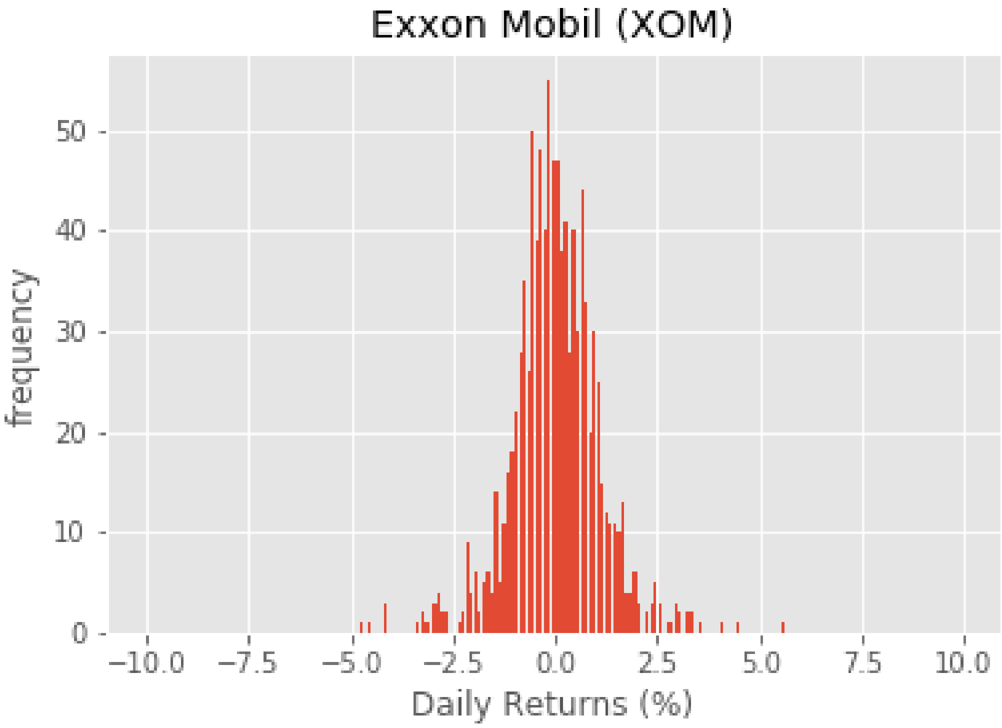
9.0 APPENDIX II – DISTRIBUTION OF DAILY RETURNS



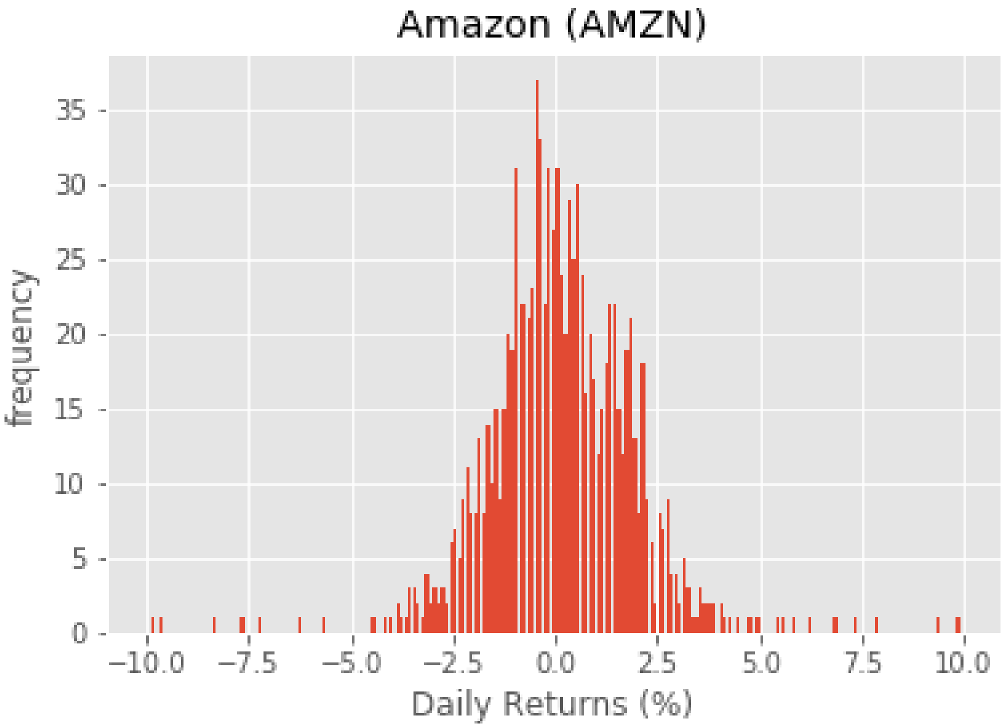



Microsoft (MSFT)
frequency
Daily Returns (%)
50
40
30
20
10
0
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0

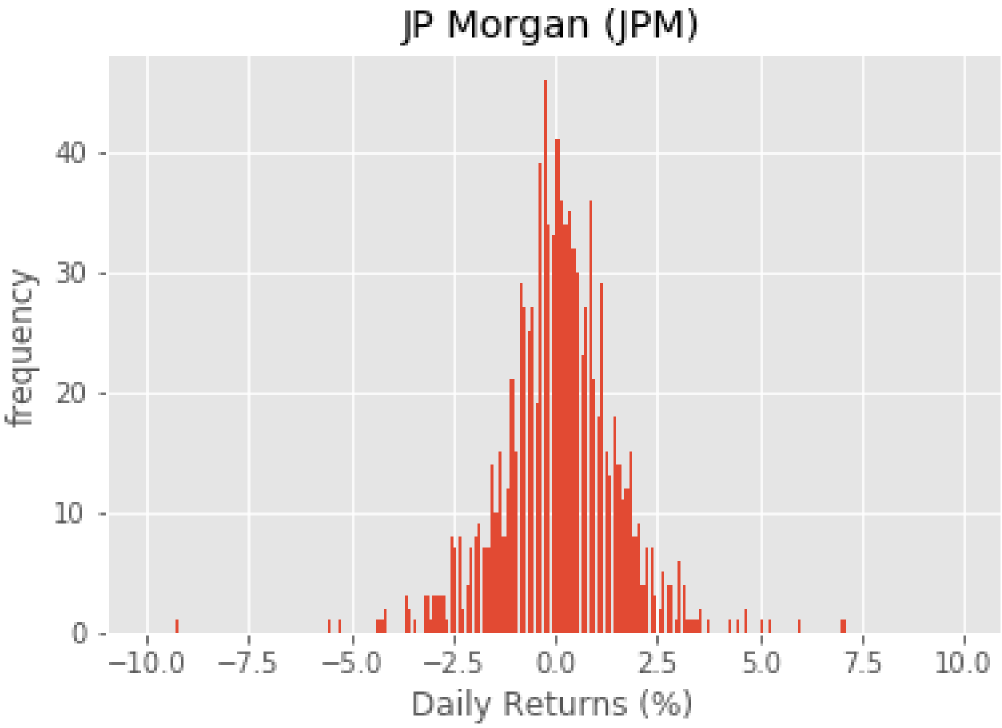

Johnson & Johnson (JNJ)
frequency
Daily Returns (%)
70
60
50
40
30
20
10
0
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0



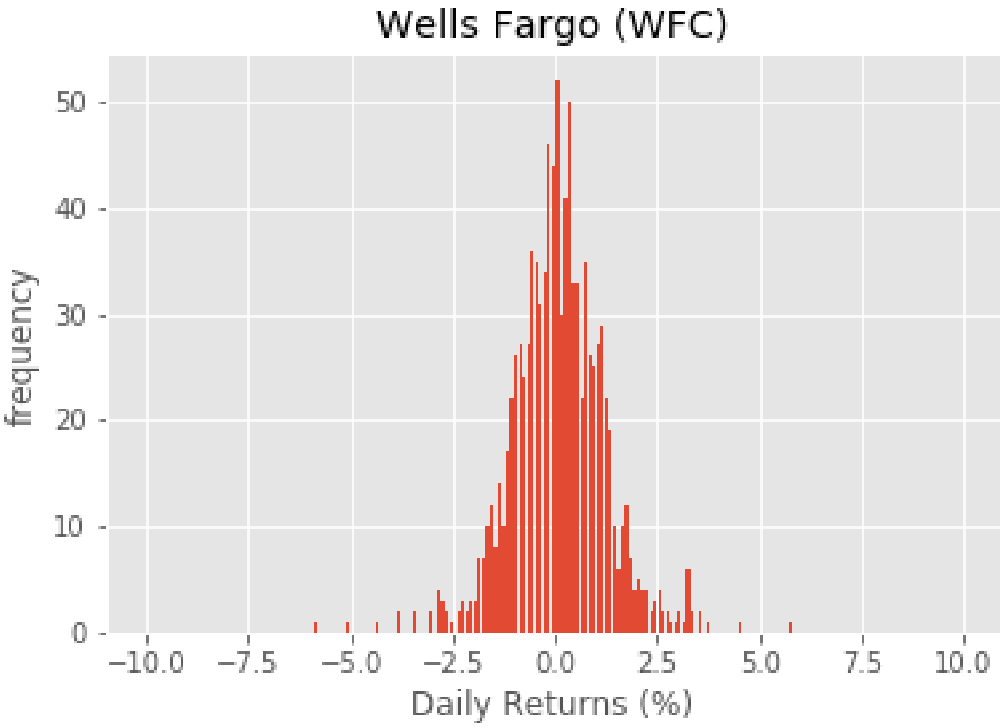
Exxon Mobil (XOM)
frequency
50
40
30
20
10
0
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0
Daily Returns (%)



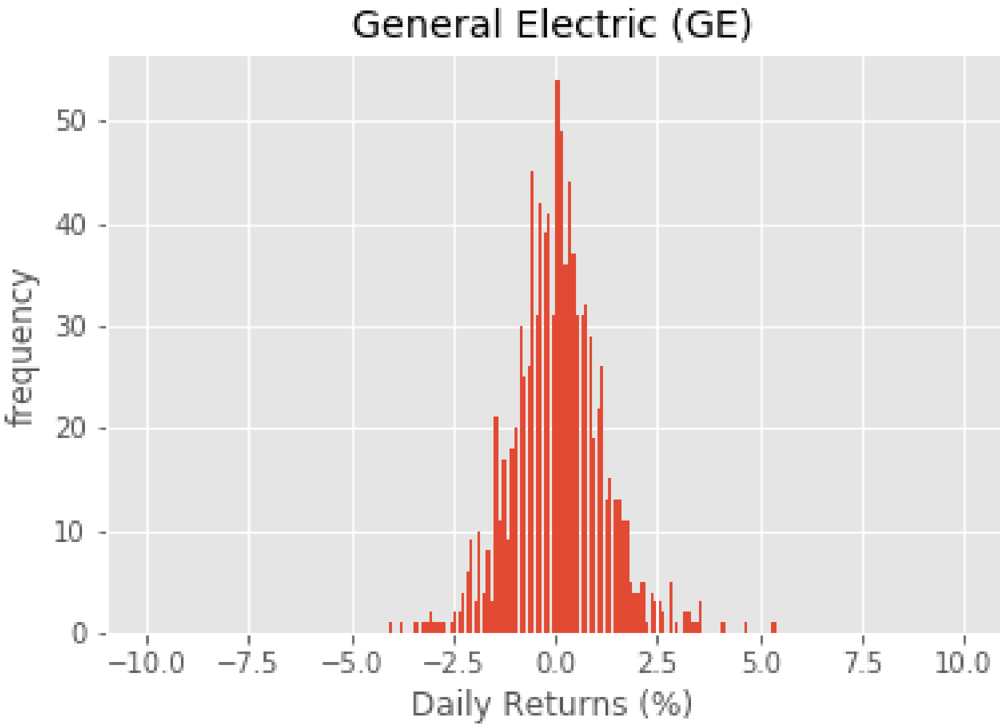
Amazon (AMZN)
frequency
35
30
25
20
15
10
5
0
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0
Daily Returns (%)

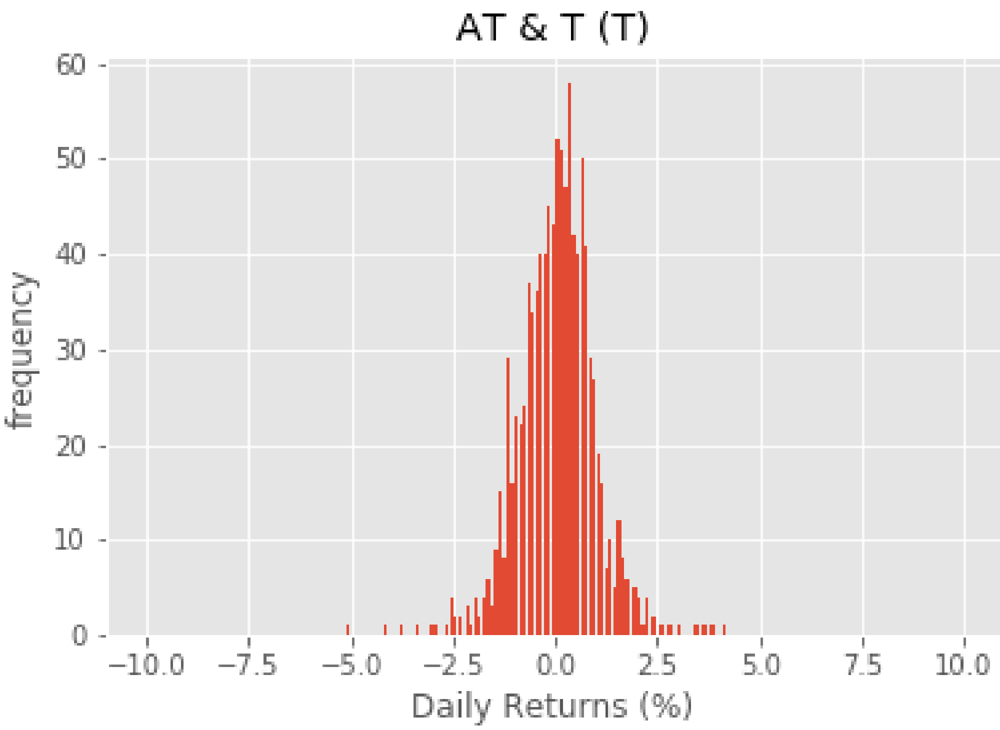


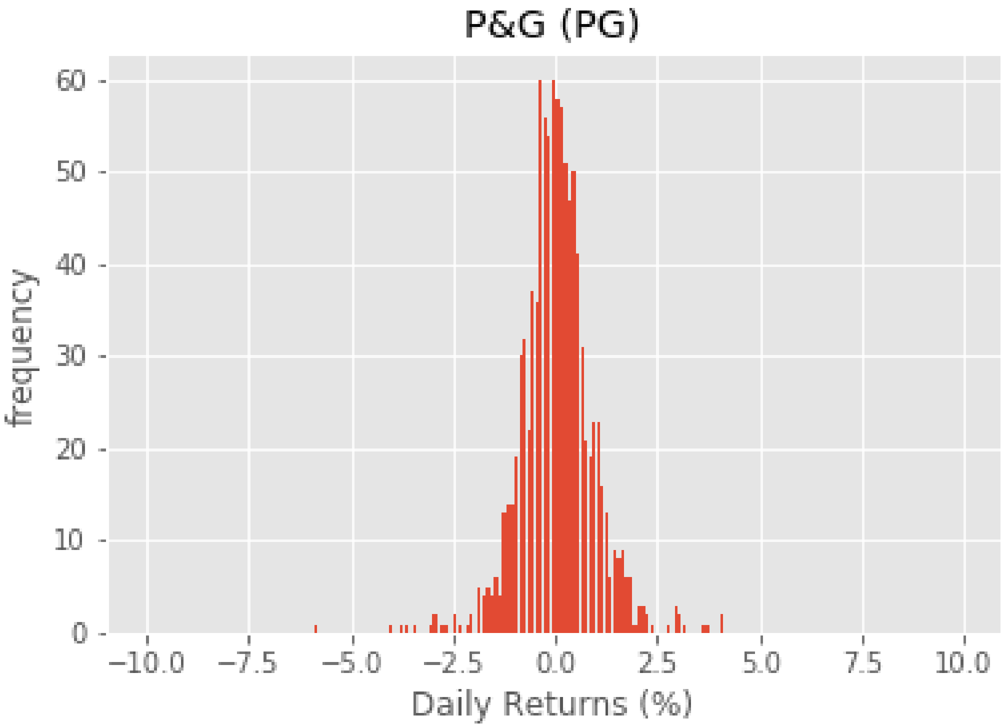
JP Morgan (JPM)
frequency
40
30
20
10
0
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0
Daily Returns (%)

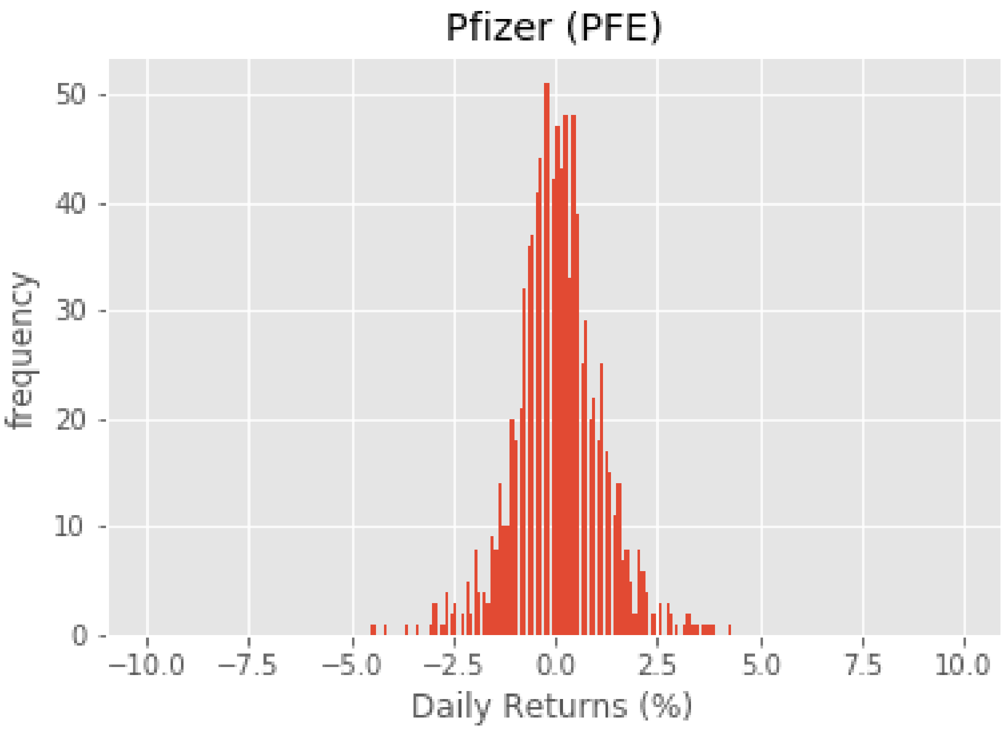


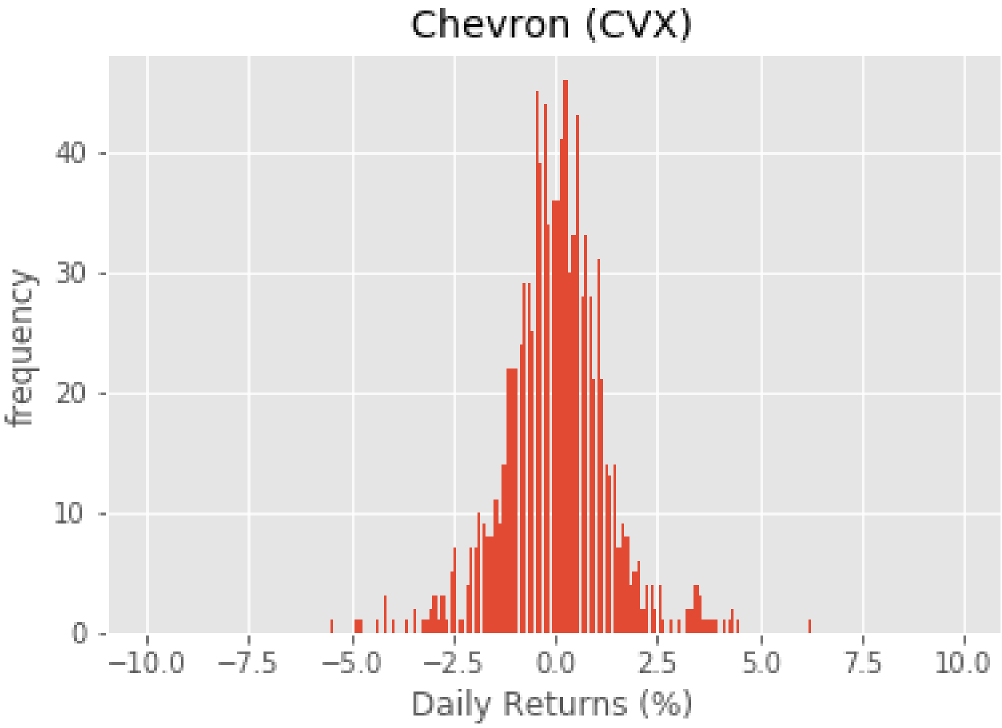
Wells Fargo (WFC)
frequency
50
40
30
20
10
0
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0
Daily Returns (%)


General Electric (GE)
frequency
50
40
30
20
10
0
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0
Daily Returns (%)


AT & T (T)
frequency
60
50
40
30
20
10
0
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0
Daily Returns (%)


P&G (PG)
frequency
60
50
40
30
20
10
0
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0
Daily Returns (%)


Pfizer (PFE)
frequency
50
40
30
20
10
0
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0
Daily Returns (%)


Chevron (CVX)
frequency
0
10
20
30
40
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0
Daily Returns (%)

10.0 APPENDIX III – GLOSSARY OF FINANCIAL TERMS

Term	Definition
Algorithmic trading	The use of algorithms for the purposes of trade execution with the focus being on efficient execution in terms of minimisation of trading costs and market impact (i.e. price changes caused by placing a large trade)
Alpha	The return over and above the market return achieved through selection of stocks which perform better than the overall market
Asset allocation	The decision to be made in portfolio management in respect to the amount to invest in each category of investments (e.g. stocks, bonds, property).
Bearish	A market is said to be 'bearish' when a negative view on prices prevails with the expectations that prices will fall
Beta	In the context of *mean-variance analysis* as proposed by Markovitz (1952) Beta measures the responsiveness to the market. Stocks with Betas of less than one are less volatile (i.e. less risky) than the market and stocks with high Betas are more volatile (i.e. more risky)
Bullish	A market is said to be 'bullish' when an optimistic view on prices prevails with the expectations that prices will rise
Efficient frontier	The full set of *efficient portfolios* which can be constructed from the available stocks and risk-free assets.
Efficient Market Hypothesis	A theory proposed by Fama (1965) which contends that stock prices fully incorporate the impact of all available market information
Efficient Portfolio	In the context of *mean-variance analysis* as proposed by Markovitz (1952) a portfolio is said to be efficient if no further increase in return can be achieved without taking on an increased level of risk.

Term	Definition
Excess return	The difference between the return earned on the portfolio and the return earned on the benchmark index and is also referred to as the active return.
Expected return	The probability weighted estimate of future returns (i.e. sum of the range of projected future returns multiplied by their individual probabilities)
Information ratio	A measure of the risk adjusted returns which expresses the *excess returns* earned relative to the *tracking error* and is computed as the excess return divided by tracking error
Mean-variance analysis	The process used to derive *efficient portfolios* as proposed by Markovitz (1952) and which seeks to optimise the level of return ('mean') with respect to the level of risk ('variance') i.e. achieve the maximum level of return for the minimum level of risk
Metaheuristic	A high level conceptual framework using analogies with simplifying assumptions used to guide a search for a solution
Quantitative trading	The use of machine learning and deep learning techniques to develop identify opportunities to make profits from identified mispricing and market trends
Sharpe Ratio	A measure of risk adjusted returns which expresses the return from the portfolio in excess of the *risk free rate* relative to the standard deviation of the portfolio returns.
Sortino Ratio	A measure similar to the *Sharpe Ratio* but which only considers downside risk
Stock Yield	Earnings per share as a percentage of the share price
Systematic risk	In the context of *mean-variance analysis* as proposed by Markovitz (1952) systematic risk is the risk arising from the overall market and which therefore cannot be reduced by holding a diversified portfolio of stocks.
Tracking error	The standard deviation of the *excess returns* (i.e. the standard deviation of the differences between the portfolio return and the index return over the assessment period)

Term	Definition
Unsystematic risk	In the context of *mean-variance analysis* as proposed by Markovitz (1952) unsystematic risk is the risk associated with holding specific stocks and this risk can be reduced or eliminated by holding a more diversified portfolio of stocks.
Variance	The sum of the squared differences between the mean return and the individual observed returns and is used to measure the dispersion of returns which is the key risk measure in *mean-variance analysis*
Weak-form efficiency	A component of the *Efficient Market Hypothesis* proposed by Fama (1965) which contends that stock prices already incorporate the impact of all historic publicly available information.