# EFFECTIVE STOCK PRICE FORECASTING USING MACHINE LEARNING TECHNIQUES WHILST ACCOUNTING FOR THE STATE OF THE MARKET

A thesis submitted to The University of Manchester for the degree of

Doctor of Philosophy

in the Faculty of Science and Engineering

**2019**


**ERHAN BEYAZ**

**School of Computer Science**

# Table of Contents

Final Word Count: 34,443 words

# List of Figures

# List of Tables

# Glossary of Abbreviations

| | |
|---|---|
| AD | Chaikin Accumulation / Distribution |
| ADX | Directional Movement Index |
| AMH | Adaptive Market Hypothesis |
| ANN | Artificial Neural Network |
| ATR | Average True Range |
| AUC | Area Under the Curve |
| BVPS | Book Value per share |
| CCI | Commodity Channel Index |
| CRF | Conditional Random Forest |
| DT | Decision Tree |
| EMH | Efficient Market Hypothesis |
| EPS | Earnings Per Share |
| EY | Earnings Yield |
| FFNN | FeedForward Neural Network |
| GA | Genetic Algorithms |
| GBM | Geometric Brownian Motion |
| LR | Linear Regression |
| LSTM | Long Short Term Memory |
| MACD | Moving Average Convergance / Divergence |
| MAPE | Mean Absolute Percentage Error |
| MFI | Money Flow Index |
| MSE | Mean Square Error |
| PCA | Principal Component Analysis |
| RMSE | Root Mean Square Error |
| RNN | Recurrent Neural Network |
| ROC | Rate of Change |
| RSI | Relative Strength Index |
| STOCH | Stochastic Oscillator |
| SVM | Support Vector Machines |
| SVR | Support Vector Regression |
| VIX | Volatility Index |

# Abstract

Machine learning methods have been successfully applied to stock price forecasting. Although finance practitioners and academics have advocated for the benefits of using fundamental and technical analyses together, the machine learning research has been focused on using the technical analysis based indicators almost exclusively. The main target for prediction by machine learning researchers have been forecasting of next day's price for a market index or a firm's stock. Another challenge presented in stock price forecasting is the impact of the overall stock market volatility on the individual stock prices. The aim of this thesis was to investigate into the impact on machine learning-based stock price forecasting by using various inputs (technical, fundamental, and combined) and also by accounting for the states of stock market. A framework is proposed which enables the selection of the best performing model with relevant inputs and which can also factor insensitivity of the stock's price to various states of the market. The initial simulations were run for 147 companies with 252 days out stock price forecasting, and further simulations were undertaken for 85 companies with 126 days out stock price forecasting. We show the importance and relevance of using the fundamental indicators and combination of the technical and fundamental indicators when forecasting financial time-series into the horizons of 126 and 252 days. The proposed approach for integrating the moods exhibited by the stock market is embedded into the forecasting process. The explicit identification and inclusion of the market states were more effective for 126 days than for 252 days, but also when the combined indicator set was not being used as the input. Another contribution of the thesis was the framework that provided an improved structured approach for conducting financial time series forecasting (RMSE of 0.0614 vs. 0.3175 of the Random Walk model).

# Declaration

I hereby declare that no portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

# Copyright Statement

http://www.library.manchester.ac.uk/about/regulations/) and in The University's policy on Presentation of Theses

# Acknowledgements

I would like to thank my supervisor John A. Keane for his never-ending support, patience, guidance and infinite amounts of wisdom. I am also very thankful for the guidance and support that Firat Tekiner has provided who has kept me motivated and on track when I seemed to start taking my eyes off the ball. I would like to thank Xiao-jun Zeng for his ideas and feedback. My friend and confidant Nihat Yilmaz provided technical assistance, emotional support, and served as a sounding board for my ideas whenever I have needed. My friend Yeliz Yesilada has generously provided much-needed clarification at the most crucial tipping points. My mother and brother have been a dependable source of encouragement at every single step of the way. I can honestly say that my journey would not have been possible without the encouragement, support, and sacrifices that my wife Nedret had shown over the years, for which I am and will be eternally very grateful. Our two little girls (Ayka and Masal) have also been patient, kind and understanding as I had to spend time away from them whilst working on my PhD. My sincerest thanks goes to all the above from the bottom of my heart.

# Chapter 1. Introduction

Can stock prices be predicted and thereby profited upon? The Efficient Market Hypothesis (EMH) essentially believes that this is not possible in markets that are efficient (where information is disseminated quickly) (Campanella et al., 2016; Urquhart and Hudson, 2013; Shonkwiler, 2013; Malkiel, 2003). However, if this holds true, the investment field in Finance would not exist. The Efficient Market Hypothesis adhered to by many Finance researchers, assumes that markets are made up of rational investors who have factored in all available information into the current price of the stock (i.e. the markets are efficient), and the future prices of the stocks cannot be predicted and profited from (Manahov and Hudson, 2014; Malkiel, 2003). The EMH states that the future prices will only be determined by events that are not yet known, but as they do take place they will be reflected in the prices of the stocks (Urquhart and Hudson, 2013; Malkiel, 2003). Therefore, followers of the EMH believe that stock prices cannot be predicted but rather that stocks tend to follow a Random Walk (RW) model (Shonkwiler, 2013; Hull, 2009; Malkiel, 2003). Believing in the EMH and consequently that stock prices follow a random walk would mean that it is not possible for investors to predict and therefore profit from these stock price predictions in a sustained way (i.e. any profit made is due to sheer chance) (Urquhart and Hudson, 2013; Malkiel, 2003). Although EMH and random walk model has been an influential theory in the Finance world where it went from being a "theory to a doctrine" (Manahov and Hudson, 2014), Adaptive Market Hypothesis (AMH) has been proposed as an alternative (Lo, 2017). AMH believes that investors as human beings show both rational and irrational behaviour, and also that investors adapt to their changing environment (economy, technology, etc.) (Lo, 2017). This creates times where the market is efficient, as assumed by EMH, but also times when the markets are not following the random walk model and are predictable (i.e. can be profited from) (Lo, 2017). Therefore, AMH allows for the possibility that methods can be used to predict stock prices, and EMH does not (i.e. random walk).

In general, stock prices are determined by forces of supply and demand in the stock market which in turn is driven by traders' decisions to buy or sell a company's stock (Thomsett, 2015). technical analysis and fundamental analysis are the two main schools of thought that finance practitioners subscribe to when making trading decisions and predicting stock prices (Thomsett, 2015; Rockefeller 2011). Technical analysis relies on past stock price and trading volume information (Rockefeller, 2011; Lorenzo, 2013) whilst fundamental analysis relies on measuring the potential ability of the company in question to generate economic value (e.g. profitability, long term growth potential, etc.) (Thomsett, 2015). Regardless of whether the analyst belongs to the technical analysis or fundamental analysis group, they generate a forecast on the price of a company's stock, and make a trading decision (buy, sell, hold). For example, in its simplest form, if analysts expect the price to be going up in the future, they would consider buying the stock; if the prices are expected to be declining, then the analysts would consider selling them (Thomsett, 2015; Lorenzo 2013). Historically, analysts tended to belong exclusively to one camp and hence could hold "polar-opposite views regarding the efficacy of fundamental versus technical analysis" (Schwager and Turner, 1995). Although these two approaches have developed as competitive methods, finance practitioners have shifted to a combined approach where technical and fundamental analysis could be used in tandem (Thomsett, 2015; Rockefeller, 2011). The impetus behind this is that an analyst can benefit from taking into consideration both technical and fundamental indicators in their analysis and decision making, regardless with which school of thought an analyst identifies (Thomsett 2015; Schwager and Turner, 1995). The researchers in the Finance and Economics fields have pointed to the benefits of using these two schools of thought together and have used them together in stock price generation, stock selection, and foreign currency trading (Hong and Wu, 2016; Chen et al., 2015; Amini et al., 2015; Wafi et al., 2015b).

The recent advances in hardware and software has resulted in computing playing a more central role in stock markets and trading (Lo, 2017; Tkac and Verner, 2016; Ibidapo et al., 2017). Stock price forecasting based on machine learning methods has proven to be both popular and successful (Tkac and Verner, 2016; Ibidapo et al., 2017; Lo 2017) where machine learning methods have been widely used to learn from past movements of a company's stock price and to generate future forecasts (Cavalcante et al., 2016; Nassirtoussi et al., 2014; Atsalakis and Valavanis, 2009). Some of the popular machine learning-based methods used in financial forecasting are Artificial Neural Network (ANN) and Support Vector Regression (SVR) (Cavalcante et al., 2016; Nassirtoussi et al., 2014; Atsalakis and Valavanis, 2009).

However, surveys have shown that machine learning research has mainly focused on applying machine learning to stock price forecasting utilizing technical indicators mainly, and has downplayed fundamental indicators (Ibidapo et al., 2017; Cavalcante et al., 2016; Atsalakis and Valavanis, 2009). This has been attributed to the fact that the indicators based on technical analysis (technical indicators) are more readily available than the indicators based on fundamental analysis (fundamental indicators) (Cavalcante et al., 2016; Atsalakis and Valavanis, 2009). This is in stark contrast to the position taken by finance practitioners and researchers who as previously stated argue for the benefits of using technical and fundamental analysis together. This has raised some questions that needed investigating:

*Q1.1 What are the potential consequences of this tendency by machine learning researchers to use technical indicators over fundamental ones?*

*Q1.2 Is one type of analysis significantly better than the other? If so can the effect of each be isolated?*

*Q1.3 Would combining the different types of analysis together yield better performance than using them separately?*

In addition to the concern with predominantly focusing on technical indicators, another complication with respect to forecasting models arises from treating the relationship between the drivers and stock prices as static, whereas this is, in actuality, a dynamic relationship (Cavalcante et al., 2016; Cavalcante and Oliveria, 2015). Financial time series data (such as stock prices) have been shown to be non-stationary in nature, and “Concept Drift” does happen over time where “the relationship between input data and the target variable” change and the learning methods should be able to handle this concept drift (Cavalcante and Oliveria, 2015). It is important to take this into account as otherwise the forecasting model trained on a set piece of training data is less applicable and effective at forecasting. Changes in the overall environment (political, economic, and regulatory, etc.) have been shown as candidates for causing the relationship to change over time (Cavalcante et al., 2016). Markets tend to go through various states (such as trending, non-trending, chaotic, bullish, bearish, in a recession, etc.), and these fluctuations do have an impact on the stock price. According to Rockefeller (2011), two main states of the market are bullish (i.e. stock prices are rising) and bearish (i.e. stock prices are in a slump). Therefore, the dynamic nature of the overall stock market could be impacting the machine learning-based forecasting models (Cavalcante et al., 2016). In terms of measuring the states of the market, various market sentiment indicators are used (e.g. Volatility Index (VIX) measures the expected fear in the market) (Achelis, 2000; Rockefeller, 2011). VIX, also known as the fear index, is widely used to represent the expectations of the stock market, where a high VIX means an expectation of high volatility in the stock prices, and a low VIX means an expectation of a low level of volatility in the stock prices (Rockefeller 2011). Clustering techniques have been applied to financial time series data for forecasting tasks (Aghabozorgi et al., 2015; D'Urso et al., 2013; Fu, 2011). Furthermore, clustering approaches on time series data have been utilized to develop localized forecasting models (Tsinaslanidis and Kugiumtzis, 2014; Cherif et al., 2011; Wu and Lee, 2015). However, these approaches have been limited to targeting the fluctuations exhibited within the stock price forecasted, rather than targeting the sensitivity of the stock in question to the various states of the market. These have raised the additional following questions that needed investigating:

*Q2.1 What would the effect be of adapting the forecasting models to account for different market states?*

*Q2.2 How would it be best to go about achieving and evaluating this approach?*

*Q2.3 Can clustering methods be used to capture the states of the stock market and integrating this within the stock price prediction process?*

## 1.1 Research aims and objectives

The aim of this research is to investigate both

1. the impact of using various input sets, and accounting for the sensitivity of a particular stock's price movements to the moods of the stock market, and

2. to evaluate the forecasting performance of machine learning-based methods for stock price forecasting.

The following research questions were posed

- Research Question 1 (RQ1): Which set of indicators, technical or fundamental, would result in better stock price forecasting performance? Investigation into this area should provide insight into whether technical or fundamental analysis is better from the machine learning-based approaches point of view.
- Research Question 2 (RQ2): Does forecasting performance improve when the technical and fundamental indicators are used together as sources for machine learning and not in isolation? Given that finance practitioners are benefiting from using these two approaches together, it is hypothesized that machine learning methods should similarly benefit from a combined use of indicators.

- Research Question 3 (RQ3): When forecasting the stock price of an individual company, would considering the states/moods of the overall stock market improve the forecasting performance? Given the dynamic nature of the stock market, it is hypothesized that identifying and accounting for the various states of the stock market within the stock price forecasting process should improve accuracy.

- Research Question 4 (RQ4): What would be the effect of using a framework which is able to identify relevant indicators and can account for the states of the overall stock market on the forecasting performance of machine learning-based methods applied to financial time series forecasting?

In order to be able to investigate these research questions, a framework to identify relevant inputs for the forecasted stock and which can account for the moods of the stock market has been designed and developed. The framework was implemented to run experiments for 147 companies from S&P 500 (http://www.spindices.com/indices/equity/sp-500) forecasting stock price movement 252 days out using models (ANN, SVR, DT, LR) which were provided with technical, fundamental, and combined input sets. For analysis purposes, further simulations were run with a subset (85) of companies for 126 days as the forecasting horizon. Furthermore, in order to capture market moods, the framework was implemented for the same scenarios but with various state layer definitions (VIX, RSI of S&P500, and Put-Call Ratio). Root Mean Square Error (RMSE) was used to capture the predictive accuracy of the models. The predictive performances of the models were compared against the performance of the Random Walk method.

## 1.2 Investigation results and thesis contributions

The following are the main contributions of the thesis:

- An approach to capturing and embedding the mood of the overall stock market into the stock price forecasting has been proposed and implemented (Section 3.3.5.6). Existing approaches tend to apply clustering methods directly to the time series data of the company whose stock price is being forecast. The proposed approach differs from these by first applying clustering methods to the stock market time series data representing the mood of the market and for each mood of the stock market identified builds individual forecasting models predicting the stock price of the company of interest.

- Experimentation covering 85 and 147 companies for forecasting horizons of 126 and 252 days respectively, enabling for the comparison of the impact of the technical and fundamental indicators and their combination in forecasting stock prices, as well as the impact of integrating market mood into stock price forecasting. Although a large portion of the existing research in this topic involves very short term (next day's) forecasting using the technical analysis based indicators, this work focuses on 6 months to a 1 year out forecasting which can be relevant to finance practitioners. This thesis also demonstrates that the fundamental analysis based indicators are impactful for such forecasting horizons and should not be ignored. The currently observed over-reliance on the technical indicators by machine learning researchers are partly explained by the difficulties encountered in obtaining and constructing the fundamental analysis based indicators. To that end, this thesis provides a description of how to retrieve and generate fundamental analysis based indicators that can be utilized for financial time series forecasting.

- A framework has been designed/developed (Section 3.2) to identify relevant inputs for the forecasted stock, to pick out the best performing machine learning method(lowest RMSE); further, the framework can incorporate the dynamic nature of the overall market within the forecasting process (by considering various states of the market as described in Section 3.3.5.6). The framework‘s main contribution is that it is able to bring together existing solutions under one umbrella with a goal of improving forecasting accuracy and providing insight. Its modular design provides a robust yet flexible way to compare and contrast the forecasting performance of models under various scenarios using combinations of different input sets and machine learning methods. The framework can easily be implemented for different forecasting horizons and sets of companies, to generate robust (cross-validated appropriately for time series as described in Section 3.2.5.2) forecasting performances. The framework is also able to go through various combination of input and machine learning methods selected, and identify a customized model (machine learning method, input set, market mood sensitivity) description that can be put forth as the most likely model to achieve successful forecasting of the stock price for the company and horizon selected, which could be of interest to the finance practitioners.

Results from the investigations:

The analysis of the results from the experiments has provided the following observations:

- Based on a 252 days out forecasting horizon, models using the fundamental indicators outperformed models using the technical indicators in 66% of the cases, versus 24% by models with technical indicators. A review based on the sectors of the companies has also revealed that the fundamental indicator-based models on average outperform models based only on the technical indicators, though the outperformance was more pronounced in certain industries than others. A similar observation was made when the forecasting horizon was set for 126 days. This confirms that the fundamental indicators are relevant in the stock price forecasting (252 and 126 days out) problem domain and suggests that using the technical indicators only and ignoring the fundamental indicators can result in sub-optimal results.

- It was observed that when only the technical indicators were used as the input, Decision Tree (DT) models outperformed the rest in 99% of the cases. When the fundamental indicators were used, SVR and DT performed equally well. However, when the combined indicator sets was used, SVR outperformed other methods in 99.5% of the cases respectively. This suggests that using only the technical indicators as input sets may provide a partial representation and as such result in incomplete interpretations.

- The models using the combined indicators were able to outperform (statistical significance of p=0.05) models using both the technical and fundamental indicators in isolation in 78% of the cases, when forecasting horizon was 252 days. The experiments conducted confirmed the view of the finance researchers and practitioners that using the technical and fundamental together results in more accurate forecasts than when using them in isolation.

- The models which used the forward-looking market mood indicators (VIX and Put-Call ratio) outperformed the models using the backwards-looking indicator (the RSI of SP500), and were, therefore, more effective at capturing the states of the market. With regards to the effective number of states exhibited by the stock market, the models using cluster size of 3 for market states outperformed the others (5, and 7).

- The models accounting for the states of the stock market (with State layer) were outperformed by the models not accounting for the states of the stock market (without

State layer) in 82% of cases when the forecasting horizon was 252 days. However, when the forecasting horizon was 126 days, this outperformance was reduced to around 50%. Furthermore, when the technical or fundamental indicators were being used as inputs, the models with the state layer were able to outperform the models without the state layer even if the forecasting horizon was 252 days. However, when the combined indicators were used the models with the state layer did not outperform the models without the state layer. In addition to being influenced by the inputs, another factor which made a difference between the models with and without the state layer was the machine learning method. When using the combined indicators and SVR as the machine learning method, the models without the state layer outperformed the ones with the state layer.

- The framework was able to outperform (statistical significance, p=0.05) the Random Walk model in all the cases considered. The framework also outperformed the ANN model using technical indicators. The outperformance of the framework was consistently observed across various industries. Thus the framework demonstrated its ability to add value by being able to pick out the best performing machine learning method per input set. Furthermore, it generated the results that were used for conducting comparison and analysis.

The research presented in this thesis lead to the following two papers:

- Erhan Beyaz, Firat Tekiner, Xiao-jun Zeng, and John A. Keane. Comparing technical and fundamental indicators in stock price forecasting. In Proceedings of the IEEE

DSS 2018 (4th IEEE International Conference on Data Science and Systems), Exeter, UK, 2018.

- Erhan Beyaz, Firat Tekiner, Xiao-jun Zeng, and John A. Keane. Stock price forecasting incorporating market state. In Proceedings of the IEEE DSS 2018 (4th IEEE International Conference on Data Science and Systems), Exeter, UK, 2018.

## 1.3 Organisation of Thesis

The thesis is divided into seven chapters. Following the introduction, an overview of trading and investment analysis, as well as a review of machine learning-based stock price forecasting is provided in Chapter 2. This chapter describes the Efficient Market Hypothesis as well as the technical and fundamental analysis methods used for stock valuation and investment decisions by finance practitioners. In Chapter 2 a review of application of machine learning methods in the domain of stock price forecasting is covered from the point of view of inputs, data processing, machine learning methods, benchmarks used and applicable performance measures. The inputs used by machine learning researchers are predominantly based on technical analysis as these inputs are relatively easier to acquire than fundamental analysis-based ones. Furthermore, the need to account for the volatile stock market movements in stock price forecasting with machine learning methods is discussed as a potential area of investigation.

Chapter 3 formalizes the opportunities identified in Chapter 2 into research objectives. This chapter introduces the proposed framework through which the experiments will be carried out and describes the data, assumptions and the approach taken.

Chapter 4 presents the results of the experiments conducted with respect to technical, fundamental and combined input sets. Reviews of features are provided followed by a comparison of the relative performance of models using the different input sets.

Chapter 5 focuses on analysing the results from the implementation of the state layer. The effectiveness of the market mood indicators and level of granularity used for representing the moods are reviewed. Furthermore, the overall performances of the models with the state layer are compared to ones without the state layer.

Chapter 6 provides an analysis of the proposed framework's performance against the Random Walk Model. Furthermore, given the reliance of the machine learners on technical indicators and also ANN being implemented so widely, ANN model using technical indicators were also used as an additional benchmark. Following this, a review of the contribution of the various layers (input, machine learning, and state) of the framework is undertaken.

Finally, conclusions and suggestions for future work are discussed in Chapter 7.

# Chapter 2. Background

The goal of this chapter is to provide the proper context for the main task at hand namely stock price forecasting using machine learning methods. Section 2.1 briefly overviews Trading and investment analysis related concepts. Specifically, the Efficient Market Hypothesis and Adaptive Market Hypothesis are covered, followed by an overview of technical analysis and fundamental analysis to give an understanding of the underlying assumptions and data used by each approach. This is followed by a discussion of whether these approaches can be substitutes or complements to each other.

Section 2.2 reviews the use of machine learning methods in forecasting stock prices. The review looks at various components that make up the models: Typical Inputs, Data Pre-processing and Feature Selection, Machine Learning Methods and model makeup, Benchmark Methods, Performance Measures, and Real-world applicability. This section ends with a review of the potential opportunities for research. Finally, Section 2.3 summarizes the main points from this chapter.

## 2.1 Trading and Investment analysis

Although historically the EMH has been prevalent, recent research shows that markets follow the Adaptive Market Hypothesis (AMH), (Lo, 2017; Urquhart and Hudson, 2013; Manahov and Hudson, 2014). Section 2.1.1 covers the main tenets of the EMH and AMH. There are two schools of thought that predominantly are used for investing/trading: fundamental analysis and technical analysis. Although these schools of thought have different underlying assumptions, they do have points of convergence. Sections 2.1.2 and 2.1.3 provide a review of technical analysis and fundamental analysis respectively, which is followed by a discussion in Section 2.1.4 on whether they might be complementary.

### 2.1.1 Efficient Market Hypothesis vs. Adaptive Market Hypothesis

The EMH assumes that markets are made up of rational investors who have factored in all the available information into the current price of the stock. In other words, the current price of the stock reflects all the information that is contained in the past prices, as well as any currently held expectations about the future (Shonkwiler, 2013). Thus, the only change to the stock price can be introduced by new developments which are not known at the moment, and are therefore random in nature (Malkiel, 2016). Efficiency refers to both the information and the speed with which it is disseminated. There are different levels of market efficiency:

1. The *weak* form of efficiency states that past prices cannot be used to predict future prices (i.e. technical analysis cannot be used to make profit) (Shonkwiler, 2013; Hull, 2009).

2. The *semi-strong* form of market efficiency states that no publicly available information can be used in predicting future prices (fundamental analysis cannot be used to make profit) (Shonkwiler, 2013).

3. The *strong* form of market efficiency states that not even insider information can be used in predicting future prices (Shonkwiler, 2013).

Thus, the EMH would dictate that stock prices are random and cannot be predicted and profited upon. Based on the EMH, the stock prices are believed to follow a random walk method, where "today's stock returns would have no statistical relationship to tomorrow's stock returns” (Lo, 2017). The EMH and therefore the random walk method have been deeply engrained in economic and finance research (Lo, 2017). However, validity of the assumptions of EMH has been challenged (Lo 2017; Manahov and Hudson, 2014; Lo 2004) and Adaptive Market Hypothesis (AMH) has been proposed as an alternative “based on evolutionary principles”. AMH states that the market prices are a result of interaction of a mixture of participants and environmental factors, and “market efficiency is dynamic and context-dependent” (Manahov and Hudson, 2014). According to the AMH, it is possible to forecast and profit from the stock market, and that the market goes through times of efficiency and inefficiency (Lo, 2017). AMH does not necessarily discredit EMH completely, but rather extends it by stating that investors exhibit both rational and irrational behaviour and they do adapt to their changing environment (such as political and technical changes). As the investors adapt to these changes, the opportunities for being able to predict the stock market come about but that also there exist periods where the stock prices do follow a random walk method.

### 2.1.2 Technical analysis

Technical analysis is founded on the belief that history will repeat itself because investors are humans who tend to “act similarly under similar conditions” (Tsinaslanidis and Zapranis, 2016). Thus, technical analysis is formed of a set of tools and approaches which “try to classify these repetitive investment/trading behaviours and their corresponding impacts to the market prices” (Tsinaslanidis and Zapranis, 2016). More specifically, technical analysis relies on the use of security price, and trading volume information as the most significant aspects in deciding what the security price will be in the future, with the goal of making profit and “beating the market” (Rockefeller, 2011). Rockefeller (2011) distinguishes between a trader and an investor such that the trader is usually working with a shorter timeframe (defined as “anywhere from a minute to a year”), whereas the investor is working with a longer timeframe (defined as “months to forever”). In addition to the timeframe, the traders seek to achieve profit by buying/selling the securities at appropriate times, whereas the investor is willing to make profit via dividend payments/bond coupon payments, in addition to buying/selling the securities. Having made this distinction, Rockefeller (2011) goes on to claim that investors and traders can equally use technical analysis to achieve their goals.

According to Tsianlanidis and Zapranis (2016), technical analysis is a chart-based approach and the tools used by technical analysts can be broadly grouped into the following categories: technical indicators, patterns, candlesticks and filter rules. technical indicators are derived from the application of mathematical formulas to the historical price (Open, High, Low, Close) and volume (number of shares traded for the stock) data. These indicators are in turn plotted along with a price chart and help guide the decision (buy, hold, sell, etc.) making by the technical analysts. Table 2-1 displays some of the popular technical indicators and provides a brief description for them (Patel et al., 2015; Krollner et al., 2010).

| Indicator Name | Brief Definition |
|---|---|
| **Average True Range (ATR)** | As described by Achelis (2000), this technical indicator shows the volatility in the stock price and is derived from a moving average of "True Ranges" for the stock. True range is calculated by using the financial time series of daily open, high, low and close prices for a stock and is the highest of the "distance from today's high to today's low, or the distance from yesterday's close to today's high , or the distance from yesterday's close to today's low" Achelis (2000). |
| **Moving Average Convergence / Divergence (MACD)** | MACD is calculated by taking the difference of the two moving averages (Short term - Long term) of the stock price (Achelis, 2000). Furthermore, a signal line which is a smoothed average of this difference is calculated. When MACD is above zero this indicates that price is likely to be on the rise (i.e. a bullish expectation for the price) and vice versa. Typically 26 days is used for the long-term, 12 days is used for short term, and 9 days is used for the signal line (Achelis, 2000). |
| **Stochastic Oscillator (STOCH)** | According to Achelis (2000), Stochastic Oscillator "compares where a security's price closed relative to its price range over a given time period". The stochastic oscillator has two main parts; K% (number of time periods used in the stochastic calculation) and D% (number periods for which a moving average of K% is calculated). |
| **Rate of Change (ROC)** | ROC is the percentage change in the stock price between two time periods (today versus 14 days ahead). |
| **Money Flow Index (MFI)** | MFI represents the money inflow/outflow from the stock. Money flow is defined as the volume of trading multiplied by the typical price (average of high, low, and close price) of the stock for the day. If the money flow is higher than the previous day's money flow this is considered to be a positive money flow and the opposite is considered to be negative money flow. Money flow index is derived from ratio of positive money flows to negative money flows over a defined period of time and is a value between 0 and 100. |
| **Commodity Channel Index (CCI)** | CCI measures how far the stock price is from its usual average (Achelis, 200). It is a ratio ranging from 0 to 100 and is used as an indication of overbought or oversold securities. |
| **Relative Strength Index (RSI)** | RSI provides a measure of the internal strength of the security by factoring average upward and downward price changes observed over selected time periods (Achelis 2000). It is a ratio that ranges from 0 to 100. |
| **Directional Movement Index (ADX)** | Directional Movement Index is used for assessing if a stock is trending. It is derived by comparing the positive directional index over a certain period with the negative directional index over the same period (e.g. 14 days). |

| | |
|---|---|
| **Chaikin AD (AD)** | According to Ulrich (2014), the "Chaikin Accumulation / Distribution (AD) line is a measure of the money flowing into or out of a security". |

**Table 2-1 Description of Popular technical indicators**

### 2.1.3 Fundamental analysis

Fundamental analysis (Krantz, 2016) is used by stock analysts to determine the expected value of a company's stock and its intrinsic value, based on a study of the underlying business drivers related to the company and its products. In its simplest form, fundamental analysis relies on reviewing the operational ability, financial performance, strategic initiatives of the company and the overall economic environment to determine the company's future expected profits over the long-term. As shown in Table 2-2, once the intrinsic value of the stock is calculated, it is compared against its current market prices (what it is currently trading for), and a trading decision is made accordingly:

| Trading Signal | Intrinsic value vs. Current Market Price |
|---|---|
| BUY | Intrinsic value > Current Market Price |
| SELL | Intrinsic value < Current Market Price |

**Table 2-2: Trading Decision for fundamental Analyst**

Krantz (2016) provides a thorough discussion of how fundamental analysis can be used to identify and invest in stocks. The accuracy of the valuation (i.e. the intrinsic value of the stock) provided by fundamental analysis is as good as the inputs and assumptions used as part of the valuation. In making the necessary assumptions, the analyst will have to use information provided in the company financial statements, company announcements, competitor information, and industry trends. fundamental analysis is a relatively comprehensive valuation technique that involves generating the expected underlying value created by the firm from various angles (Krantz, 2016):

- **Individual Company perspective:** A valuation for the stock price is generated by gauging the expected sustainable profitability of the company operations by reviewing financial metrics, operational efficiencies, and managerial expertise, and coupling this information with future expected growth and management strategies.

- **Industry and Macroeconomic perspective:** The risks or opportunities to a company are identified by reviewing its performance and positioning in relation to other competitors, against its sector/industry, and with respect to the state of the overall economy.

According to Wafi et al. (2015a), the fundamental analysis based stock valuation approaches can be categorized as follows:

- ➢ Dividend Discount Models (DDM), which uses the expected future dividend pay-out to generate the valuation
- ➢ Price Multiples, which uses the ratio of the stock price to a fundamental driver (e.g. Price / Earnings, Price / Sales, Price / Book Value, etc.).
- ➢ Discounted Cashflow model (DCF), which the firm's ability to generate future cash flows and uses these to generate the valuation
- ➢ Residual Income Model (RI), relates the assets of the company (measured by the book value) and its ability to generate earnings (Earnings per share (EPS)) in calculating the value delivered by the firm.

### 2.1.4 Technical vs. fundamental or technical and fundamental

Krantz (2016) states that from a philosophical perspective "technical analysts and fundamental analysts are diametrically opposed to one another." Technical analysis believes the driver of the price of the stock to be the momentum of the market and its past historical price, and that there is no additional information (that can be discovered by doing fundamental analysis) that is not factored into the price of the stock already. Krantz (2016) points out the fundamental analysts are concerned that technical analysts tend to consider momentum and market dynamics which can result in "groupthink" pushing prices away from their intrinsic value. Schwager and Turner (1995) recount that these two schools of thought have enjoyed varying levels of popularity and success over time. Up to the 1970s, fundamental analysis was preferred over technical analysis. However, this trend reversed throughout the 1970s and 1980s as the market experienced much commodity inflation, as technical analysis became the preferred method, and yet in the 1990s the trend reversed back in favour of fundamental analysis (Schwager and Turner, 1995). However, according to Rockefeller (2011) technical analysis has made yet another comeback and has finally been accepted as a viable approach (Nazario et al., 2017; Rockefeller, 2011).

Schwager and Turner (1995) have conducted numerous interviews with prominent traders from both schools of thought who expressed extreme suspicion against successful use of the other methodology; it was also noted that successful traders always make some use of the fundamental indicators. Thus, Schwager and Turner (1995) also find that the two methods are not mutually exclusive and can and have been used together by traders where fundamental analysis is relied on when determining <u>which stocks</u> to trade in and technical analysis is relied on to determine the <u>timing</u> of the trade. Similarly, Krantz (2016) points out that fundamental analysis can help in identifying which stock might be currently undervalued, but it does not mean that the market is actually going to eventually move to the intrinsic value at least in the short term. Thus, Krantz (2016) believes the largest potential weakness of using fundamental analysis in isolation is that it might help find which stock to buy/sell, but not necessarily the best time to buy/sell. On the other hand, Krantz (2016) points out that fundamental analysis is a valuation technique that is firmly grounded on analysis of business drivers, and therefore is less likely to be overly impacted by emotional swings of the market. According to Krantz (2016), this can protect long-term investors from making bad decisions based on short term volatility. Admitting that these two approaches come from different angles, Krantz (2016) believes that they can be used synergistically. Krantz (2016) further suggests that a fundamental analyst can monitor technical indicators to compensate for the lack of timing, as well as serving as an “early warning system” to spot changes in the market that necessitates exit and formation of bubbles. Rockefeller (2011) also supports the view that technical and fundamental can be used together and notes that many technical traders do use fundamental analysis in aiding them to make trading decisions.

Thomsett (2015) believes that fundamental and technical analysis can and should be used together; the reason being that it does not make sense to lose sight of information provided by either school of thought, regardless to which the trader might subscribe. Thomsett (2006) further states that the technical indicators help confirm/question the trends/assumptions that fundamental analysis is generating and is therefore useful in completing the picture for the trader in making an informed trading decision. Vanstone and Finnie (2009) point out that both approaches have merit and can be used in a complementary manner to forecast stock prices. Vanstone and Tan (2003) noted that the traders have grown more accustomed to seeing these two schools of thought as complementary.

Finance researchers have shown interest in comparing the performance of technical and the fundamental indicators, and investigated the benefits of using them together. Chen et al. (2016) show that using the fundamental indicators (FSCORE formed by addition of numerous financial ratios to indicate the financial strength of the company) to complement technical indicators has outperformed (information ratio of 0.1845 versus 0.1335) a momentum strategy using technical indicators only, when the investment horizon is 6 months. Wafi et al. (2015b) compared the predictive performance in one day ahead forecasting in Egyptian stock market of technical indicators (lagged prices) and the fundamental indicators (Book Value per share (BVPS) and EPS). When the one day price was being forecast, technical indicators outperformed the fundamental indicators (RMSE of 69.9 vs. 82.5). However, when one day ahead return was being forecast, the fundamental indicators outperformed (1.30 vs. 1.38) the technical ones. Similarly, Bettman, Sault, and Schultz (2009) tested whether combining fundamental and technical analysis were statistically significantly better predictors for stock values compared to using only fundamental analysis factors (Book value and Earnings Per Share) or technical analysis indicators (momentum strategies) in isolation. The tests found that fundamental analysis (Adjusted r-square of 0.7629) and technical analysis (Adjusted r-square of 0.7546) were both effective methods of share price valuation and that the combination (Adjusted r-square of 0.7686) of the two was the best predictor of the three approaches. A similar shift has been observed to have taken place in another part of the finance industry: Bettman et al. (2009) state that there have been numerous documented cases in the foreign exchange market of combined use of technical and fundamental analysis by market participants (Lui and Mole, 1998; Oberlechner, 2001; Zwart et al., 2009).

In summary, although the debate on whether fundamental analysis and technical analysis can be substitutes, or which is “better”, has not been resolved, what has been agreed is that use of information provided by both approaches together may be extremely valuable.

## 2.2 Use of Machine Learning methods in stock price forecasting

According to Cavalcante et al. (2016), stock price forecasting problem has traditionally been approached from two camps: Statistical Techniques and Machine Learning Techniques. Statistical techniques operate from the assumption that the underlying relationship between the stock prices and their drivers are of a linear nature. However, financial time-series (such as stock price data) have been shown to be non-linear and noisy, therefore making machine learning methods, which can handle such data characteristics, the better forecaster of the two approaches (Cavalcante et al., 2016; Hsu et al., 2016). Fuelled by advancements in computing power over recent years, the use of computing in making trading decisions has increased commensurately (Cavalcante et al., 2016; Atsalakis and Valavanis, 2009). Although machine learning techniques have been “widely accepted to studying and evaluating stock market behaviour” (Atsalakis and Valavanis, 2009), there is not a clearly identified set of indicators and methodology that can be used to consistently forecast stock prices effectively (Cavalcante et al., 2016).

Cavalcante et al. (2016) provide a financial trading framework (Figure 2-1) with forecasting, which is a customized version of the systemic approach for forecasting captured by Palit and Popovic (2006); it shows the steps involved in financial time-series forecasting with machine learning methods.

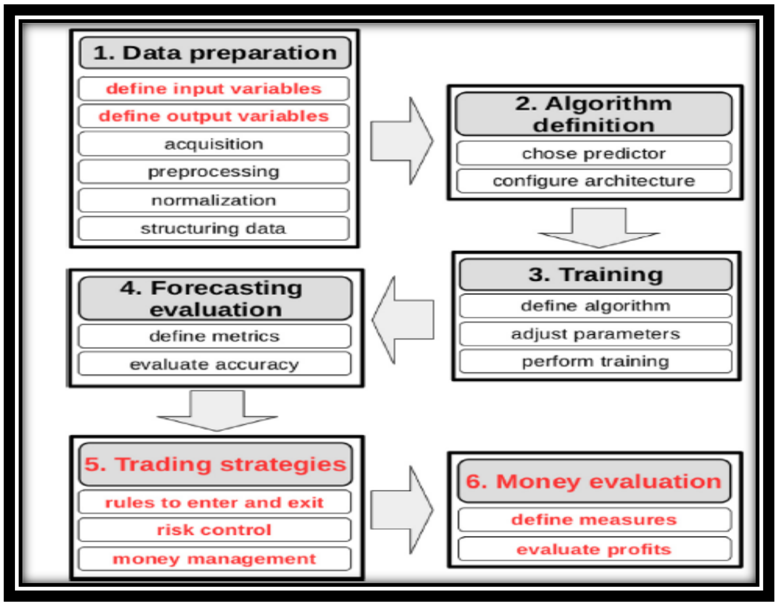


**Figure 2-1 Financial trading framework with forecasting (Cavalcante et al., 2016)**

The following sub-sections review the application of machine learning to stock forecasting from the perspective of various different aspects and considerations that make up the models: Typical Inputs, Data Pre-processing and Feature Selection, Machine Learning Method and model makeup, Benchmark Methods, Performance Measures, and Real-world applicability.

### 2.2.1 Typical Inputs

The type of inputs used in forecasting stock prices are, to a great extent, dependent on the underlying investment analysis approach: fundamental or technical analysis. Historically, work using machine learning methods has shown a clear preference for using technical indicators as inputs (Cavalcante et al., 2016). Cavalcante et al. (2016) did a recent survey of the application of machine learning methods on Financial Markets, where they reviewed research articles from 2009 through 2015 and found that "technical analysis is the most used approach in the study of financial markets in the surveyed papers". Out of the 56 papers reviewed, 47 of them used technical indicators and 9 of them used fundamental indicators. Furthermore, other older surveys show a very similar trend. The preference by researchers to use technical indicators was noted in the survey by Atsalakis and Valavanis (2009) which focused on use of Neural Networks (NNs) and Neuro-fuzzy models in forecasting the stock market, whereas the majority of the articles reviewed utilized technical indicators as inputs. Similarly, Krollner et al. (2010) conducted a survey of 46 research papers on machine learning models used for stock index forecasting and focused their review on aspects of technologies used, the forecasting timeframe, input variables and evaluation methods. Krollner et al. (2010) noted that over 75% utilized technical indicators and of these indicators simple moving average (SMA), exponential moving average, relative strength index (RSI), rate of change (ROC), moving average convergence/divergence (MACD), and Stochastic oscillator and average true range (ATR) were used most often. Vanstone and Tan (2003) did a similar survey in 2003 where they looked at use of machine learning in investment and financial trading. These authors noted a similar trend of over-reliance on using technical indicators as inputs and explained this by pointing out that machine learning approaches are dependent on large volumes of data and that technical indicators were more available and easily accessible, especially on a daily basis, whereas fundamental indicators became available on a less frequent basis (quarterly or yearly) and as such were much less preferred. However, Cavalcante et al. (2016) points out that the fundamental indicators are making their way into input sets in the form of news and sentiment analysis, especially for next day stock price prediction.

According to the surveys conducted by Cavalcante et al. (2016), Krollner et al. (2010), Atsalakis and Valavanis (2009), and Vanstone and Finnie (2003), most studies reviewed used technical indicators only, a few used fundamental indicators only, and a handful looked at combining the two schools of thought to a degree. Chandwani and Saluja (2014) used technical indicators, fundamental indicators and their combination as inputs for forecasting the stock direction for companies in the Indian stock market. ANN, SVR models which were optimized through Genetic Algorithms (GA), as well as the plain (without the optimization) versions of ANN and SVR models were used to generate the forecasts. The results indicated that ANN model optimized via Genetic Algorithm which was using a combined indicator set achieved the highest accuracy rate (80.51%) which was followed by the plain SVM model with technical indicators which achieved an accuracy rate of 79.4%. The study included only 25 companies from an emerging market and did not necessarily provide a review of the results on a company level. In predicting the direction of Apple's stock price for the next day forecasting models using various data sources were utilized by Weng et al. (2017). The data sources included Market data (Financial time series data and P/E ratio), technical indicators, Wikipedia Traffic, and Google news counts. The performance of the models (ANN, SVR, DT) using each data source individually and in a combined manner were compared. The best performing model used all the data sources and achieved a hit ratio of 85% with an Area Under the Curve (AUC) of 0.874. Even though the study was focused on predicting the one day ahead direction change in the stock price of only one company, it showed that putting various data sources improved the prediction accuracy.

As covered in Section 2.1.3 there has been a significant shift within the trader community and academic finance circles towards complimentary use of technical and fundamental analysis, yet as can be seen from the above, adoption of this change have not been at the same rate by researchers who use machine learning techniques for stock price forecasting. This has been attributed to the fact that technical indicators are easier to gather and are more available (Cavalcante et al., 2016).

### 2.2.2 Data pre-processing and feature selection

Data pre-processing involves cleansing of the data (removal of missing values, etc.) and also standardizing the input data which allows machine learning methods to learn more effectively from the inputs. Atsalakis and Valavanis (2009) emphasize the importance of data pre-processing/sampling in affecting forecasting performance and noted that "all articles referring to data pre-processing find it useful and necessary." Vanstone and Finnie (2009) recommends using ratios and not necessarily the actual values as inputs especially when inputs will be provided to neural networks, in order to enhance the generalization ability of the neural networks in finding solutions. Another important step in data pre-processing is the addressing of the missing values. Although missing values can be imputed via multiple approaches (using an overall average, etc.), for financial time series data, according to Romero and Balch (2014), "the common approach is to fill forward: to treat missing values as the same level as the last known value recommends" and, where missing values are at the beginning of a series, to fill backward.

In addition to data pre-processing, the predictive performance of the forecasting model is also influenced by the application of feature selection methods. Feature selection is defined as the process carried out "to discard attributes that appear to be irrelevant" (Russell and Norvig, 2010). It has long been known that when using machine learning techniques, it is important to address the issue of feature selection to reduce the set of inputs into relevant ones thereby minimizing the effects of the "curse of dimensionality" (Bellman, 1961).

According to Torgo (2017), feature selection methods come in as filter or wrapper methods. Filter methods reviews the features and remove the less relevant ones based on defined metrics. Filter methods do not take into account the forecasting approach, whereas wrapper methods do take these into account. For example, Torgo (2017) applies Principal Component Analysis (PCA) and Random Forest (RF) as filter methods to remove less relevant or redundant input data. Zhong and Enke (2017) applied PCA, Fuzzy robust principal component analysis (FRPCA) and kernel-based principal component analysis (KPCA) as feature selection methods to provide the data to an ANN in forecasting the S&P 500 Index ETF. Zhong and Enke (2017) compared performance of the models using the various feature selection methods to each other as well as to the model without feature selection and found that "PCA and ANNs gain significantly higher risk-adjusted profits than the comparison benchmark" and outperformed the other feature selection alternatives.

### 2.2.3 Machine learning Method and model makeup

Another important factor in the modelling decision is the machine learning technique(s) that will be deployed. In terms of a broad categorization, the majority of the machine learning approaches tend to fall into three main groups (Cavalcante et al., 2016):

- models that use a single machine learning technique,

- models that use a hybrid combination of machine learning techniques with optimization techniques, and
- models that are an ensemble of various single models.

With respect to the time series forecasting models that use a single machine learning technique, although there are a large number of alternatives (e.g. ANN, SVR, DT, GA, Hidden Markov Models (HMM)) that have been included in the research studies, Cavalcante et al. (2016) point out ANN and SVR to be "the more common soft computing techniques applied in forecasting financial time series." One of the reasons for the success of ANN and SVR in the realm of financial forecasting has to do with the nature of the financial time series data (Cavalcante et al., 2016):

- Non-linear, uncertain and notoriously noisy
- No identical statistical properties may be observed at each point in time
- Highly volatile

The success of ANN and SVR models have made them popular and widely implemented in research circles, as evidenced by their prominence in the relevant surveys (Cavalcante et al., 2016; Krollner et al., 2010; Atsalakis and Valavanis, 2009). Cavalcante et al. (2016) includes 19 studies applying ANNs and 10 other studies applying SVR to financial forecasting problems. Krollner et al. (2010) carried out an extensive survey of publications (46) covering machine learning techniques used in financial time series forecasting and found that single ANN-based methods were by far the most frequently applied methodology (21 out of the 46 reviewed publications). Cavalcante et al. (2016) notes that "ANNs have become very popular in the context of financial market forecasting", and furthermore include a reference to a study which states that "the majority of work which proposed the use of ANNs for solving the financial forecasting problem have used a multi-layer feed-forward neural network (MLP) trained with backpropagation algorithm with great success." The review carried out by Cavalcante et al. (2016) reveal that " 'forecasting', 'technical', and 'MLP' concepts co-occur several times in combination with other concepts on primary studies." Atsalakis and Valavanis (2009) noted that roughly "60% of the surveyed articles use feed-forward neural networks (FFNN) and recurrent networks". Cavalcante et al. (2016) point out that SVR models are widely used as "alternatives to ANN" models. Cavalcante et al. (2016) further point out that "the solution of SVM may be global optimum, while conventional ANNs tend to produce just local optimum solution." Cavalcante et al. (2016) point out that recently deep learning methods have been applied successfully and that they can be used for financial time series forecasting as well. Deep Learning methods are exceptionally powerful at being able to do feature extraction where unknown relationships can be identified without necessitating experts to define these.

According to Cavalcante et al. (2016) there are two broad types of approaches to hybrid combination of machine learning techniques used for stock market forecasting: (1) a machine learning method is used to optimize the parameters of another machine learning method that is used for actual forecasting; (2) the forecasting task is divided amongst the machine learning methods. Figure 2-2 demonstrates an example belonging to the first category based on approach by the work of Hadavandi et al. (2010a) which compared performance of their evolutionary ANN with a static ANN model and the model proposed by Esfahanipour and Aghamiri (2010). After using stepwise regression for feature selection, Hadavandi et al. (2010a) used a Genetic Algorithm (GA) (truncation, crossover, mutation) to optimize the parameters (transfer function types, number of hidden layers, number of nodes for each layer) of their ANN-based forecasting model. The performance of the models was compared using MAPE. The evolutionary model outperformed both the static ANN and the model proposed by Esfahanipour and Aghamiri.

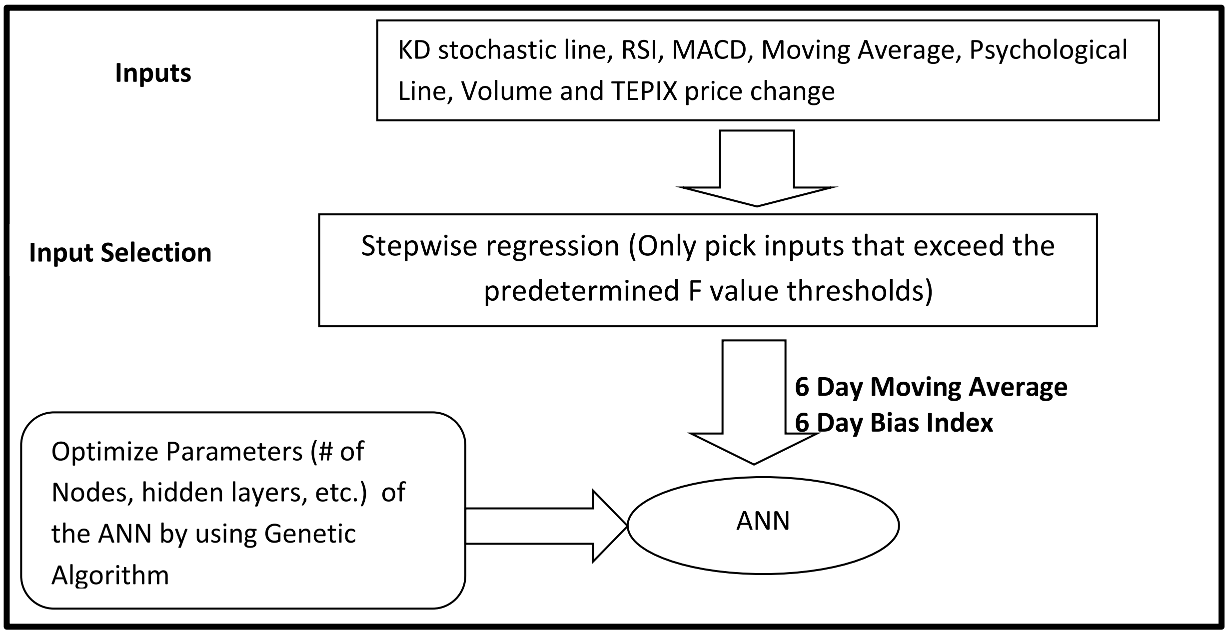


**Figure 2-2 Forecasting Approach by Hadavandi et al. (2010a)**

Another example in hybridizing two forecasting techniques can be seen in Patel et al. (2015), who use the SVR as an initial layer to transform raw inputs into forecasted technical indicators and then provide these indicators into SVR, RF and ANN for actual forecasting task of 1,10,15, and 30 days for CNX Nifty and the S&P Bombay Stock Exchange. These authors compared the predictive performance of the hybrid approach with the single SVR, RF, and ANN models and found that the hybrid approach outperformed models that are an ensemble of various single models.

In addition to single and hybrid models, there are also the ensemble of models. Xu et al. (2010) compared the predictive performance (measured by MAPE, MAE, and RMSE) of a single ANN, an ensemble forecast using the bagging method and an ensemble forecast using the constraint bagging method. The authors found that the constraint bagging approach performed relatively better than a traditional bagging approach which in turn performed better than any of the single neural networks in the experiments to forecast the next day's closing price for Dow Jones Index (DJIA). Cavalcante et al. (2016) points out that ensembles "allows exploring additional information and the consensus among individuals that compose the ensemble with the goal of improving the generalization performance when compared with an individual learning method."

### 2.2.4 Benchmark Methods

Another component of the model is the evaluation/benchmark methods against which the performance of a proposed model is compared. The survey by Krollner et al. (2010) pointed out that from the 68 articles reviewed, 41% used other machine learning techniques as the benchmark; this was followed by statistical methods in 26% (e.g. ARIMA) of the cases, Buy and Hold in 13% of the cases, Random Walk in 9% of the cases, and no benchmark in remaining 10% of the cases. With respect to the other machine learning techniques in the cases where the research paper was proposing an improvement over an established machine learning technique, the performance of the established machine learning technique and that of the proposed model were compared. Similarly, in the 101 articles surveyed by Atsalakis and Valavanis (2009), the distribution of the benchmark models were as follows: ANN (23%), Buy and Hold (23%), Linear / Multivariate regression (18%), ARIMA (7%), Random Walk (5%), Genetic Algorithms (2%), others (23%).

### 2.2.5 Performance Measures

Performance measures enable evaluation of the success of machine learning methods in forecasting. Atsalakis and Valavanis (2009) found that generally, the performance measures used were either statistical (such as RMSE, MAE, MSPE, etc.) or economic / profit-oriented (such as Hit Rate, Average Annual Profits, Annual rate of return, etc.). Of the 72 studies reviewed, 26 used statistical measures only, 26 used economic/ profit-oriented measures only and 20 used a combination of both statistical and economic/profit-oriented measures. Krollner et al. (2010) found that 31 out of 46 studies used forecast error as an evaluation metric, but that this was not necessarily reflective of the real-world as "a smaller forecast error does not necessarily translate into increased trading profits", as this is also influenced by trading decisions. Hyndman and Athanasopoulos (2014) suggests using "the MAE or RMSE if all your forecasts are on the same scale" and alternatively to use "the MAPE if you need to compare forecast accuracy on several series with different scales, unless the data contain zeros or small values, or are not measuring a quantity".

### 2.2.6 Real-world applicability

#### *2.2.6.1 Forecasting Horizon*

There might be some limitations with regards to being able to put to use the findings from the research on machine learning-based financial time series forecasting. One such limitation might be coming from the forecasting horizon that is being considered in these studies. 31 out of the 46 studies reviewed by Krollner et al. (2010) in their survey were focused on doing one day ahead forecasting of stock indices but that this in itself "does not necessarily mean that an investor can take advantage of this information in terms of trading profit, especially since the index itself cannot be traded". Only 9 studies focused on multiple forecasting timeframes.

#### *2.2.6.2 Volatility in stock prices introduced by external factors*

Another limitation with respect to real word applicability might be due to the relationship between the stock price and its drivers changing over time (Cavalcante et al., 2016; Hsu et al., 2009), as the financial markets evolve over time. For example, according to Cavalcante et al. (2016), "The time series of stock prices of a company may change its behaviour due to changes in political and economic factors or due to changes in the investor psychology or expectations." Thus, identifying the drivers of stock valuation and developing static models and weights based on these inputs is inefficient, as the dynamic nature of the market makes it difficult to find one approach/model that is valid at all times. The stock market is known to exhibit Bull and Bear states, which are periods of "upward and downward trends of stock index or positive and negative stock index returns over a period of time" (Jiang and Fang, 2015). According to Jiang and Fang (2015), the methods applied to identify the states typically fall under either non-parametric methods or parametric methods. The non-parametric methods involve using the peaks and troughs of the time series data, whereas the parametric methods involve developing "econometric models to quantitatively study the time series" (Jiang and Fang, 2015). Applying Markov switching model to monthly S&P500 returns, Jiang and Fang (2015) identified 4 distinct states of the market to be exhibited from 1926 to 2011:

- **State 1:** Extreme Bear Market characterized by low mean return and very high volatility

- **State 2:** General Bear Market characterized by negative mean return
- **State 3:** Volatile Bull Market characterized by high mean return and high volatility
- **State 4:** Steady Bull Market characterized by high mean return

Munnix et al. (2011) calculated on a daily and intra-day basis Pearson correlation coefficients for stocks making up the S&P 500 for 19 years (1992-2010). These correlation coefficients were cross-checked against known financial crises timings (e.g. 2008-2009 Financial crises). Market states were defined by using the correlation between different industry branches and intra-branch correlations. High correlation was observed during market crisis moments, and low correlation was observed during stable and calm periods. Top-down clustering were applied using k-means to create clusters which resulted in 8 market states being exhibited between the years of 1992 and 2010. The financial market was shown to go back and forth among these 8 states and exhibit those states at varying lengths of time.

Cavalcante and Oliveira (2015) state that such concept drift does take place in financial time series data; one approach to addressing this is to recalibrate the models on a pre-determined basis (i.e. implicit), and the other approach is to have a trigger (i.e. explicit) which “monitor some statistics of the data stream in order to detect concept drifts” (Cavalcante, Minku, Oliveira, 2016). Cavalcante and Oliveira (2015) simulated an approach where the online sequential extreme learning machine is updated with such an explicit drift detection which resulted in improved speed whilst maintaining accuracy. Perceptually Important Points (PIP) or Turning Points (TP) have been used for such segmentation of time series data to serve as triggers for concept drift (Cavalcante et al., 2016). Tsinaslanidis and Kugiumtzis (2014) used PIP to segment time series into groups of sub-sequences and

Dynamic Time Warping (DTW) was used to match the current instance with the previous similar sub-sequences. They used this approach to show predictability for 18 major financial market indices and GBP/USD exchange rate. Hsu et al. (2009) used a two-layered “divide and conquer” approach, where Self-Organizing Maps (SOMs) are used as an initial filter to break the historical price data into groupings that show similar characteristics. These groupings in effect represented the various states/conditions/ moods of the market. Once these groupings were created, the authors then applied SVR to each grouping in order to establish the relationship between the independent variables (lagged closing prices) and the dependent variable (relative price change 5 days into the future). The authors discovered that the two-layered approach of SOM + SVR was superior to using SVR by itself, consistently across the 7 major market indices.

Hadavandi et al. (2010b) proposed a stock forecasting system, named clustering-genetic fuzzy system, where they leveraged the strengths of SOM for clustering and fuzzy logic for rule extraction and a genetic algorithm for optimization. The authors applied their model to predict the next day’s closing price for IBM, Dell, British Airways (BA), and Ryanair (RA). The basic approach of the authors’ model is shown in Figure 2-3.

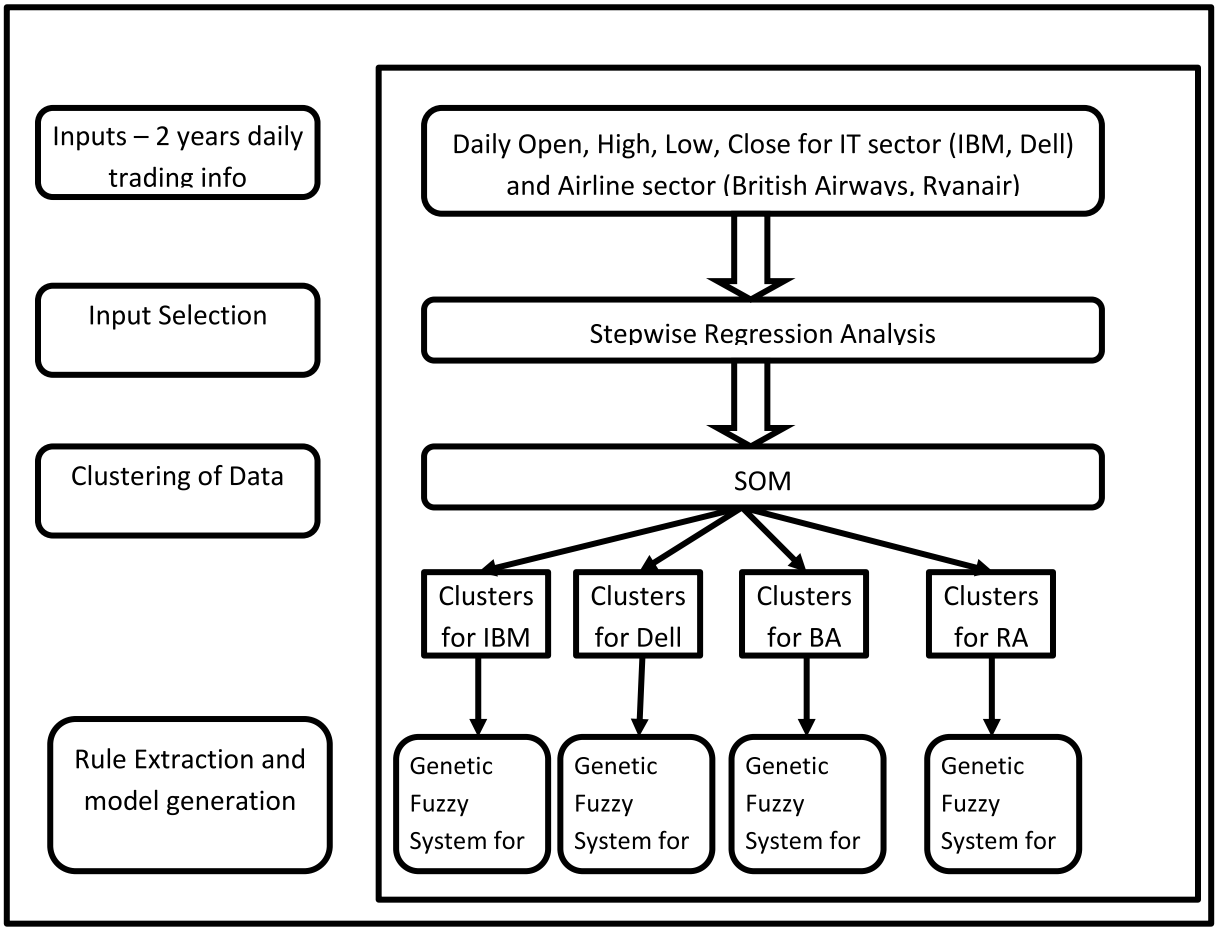


**Figure 2-3 Forecasting Model by Hadavandi et al. (2010b).**

Hadavandi et al. (2010b) used stepwise regression in selecting the relevant variables for each company. These selected inputs were then fed into a SOM so that the time-series data can be grouped into clusters. Then for each cluster, a genetic fuzzy system was developed. The fuzzy logic approach was deployed to extract the initial set of rules automatically, which were then passed onto a genetic algorithm for further refinement through crossover and mutation, eventually generating a database of rules. These rules were then further refined to generate the best rules where the top 10% was preserved and the remaining was further improved through crossover and mutation. The authors compared the performance of their clustering-genetic fuzzy system with other forecasting models (Hidden Markov Model (HMM) by Hassan and Nath (2005), a fusion model (HMM, ANN, GA) by Hassan, Nath, and Kirley (2007), and a combination model of HMM and Fuzzy Logic by Hassan (2009)) that were used to predict the same stock data and the authors' model was able to outperform other models based on MAPE values.

Another example demonstrating that including the state / condition / moods of the market is a more robust approach was exhibited by Khoa et al. (2006) who compared the price forecasting capabilities of a Feed-Forward Neural Network (FFNN) with a Recurrent Neural Network (RNN) (which accounts for a state layer in its functioning) and found the RNN to be superior (more profitable by up to 25%). Thus, there appears to be an incremental informational value to be extracted from identifying the various states of the market and then developing optimized forecasting models that are successful for each of those states.

Based on the review presented in Section 2.2, the following conclusions emerge:

- The majority of machine learning researchers use technical indicators; Fundamental indicators are starting to be used by more such researchers but not directly from fundamental analysis but rather indirectly as news or sentiment indicators. There are only a handful of machine learning implementations that have considered using a combined set of indicators even though, as covered in Section 2.1, combined use is advocated in finance circles.
- The studies on machine learning-based financial time series forecasting mainly focus on next-day forecasting

- ANNs and SVRs are widely implemented and have been successfully used towards stock price forecasting
- Markets tend to go through various states and the dynamic nature of the markets pose a challenge in forecasting the stock price and should be considered.

## 2.3 Conclusion

Section 2.1 provided an overview of trading-related concepts such as Efficient Market Hypothesis, Adaptive Market Hypothesis, technical and fundamental analysis. It was stated that fundamental analysis and technical analysis are the two main methods used by traders as part of their trading decision. Although it has been widely thought that they are mutually exclusive, finance practitioners are evidencing that they are complementary and that traders can benefit from using technical as well as fundamental indicators regardless of which school of thought they most identify with. Machine learning methods have been applied widely to stock price forecasting successfully.

Section 2.2 provided an overview of the application of the machine learning methods to stock price forecasting. It is also shown that the studies regarding these machine learning approaches have relied on the use of technical indicators almost exclusively. It is highlighted that when developing these models, the technical indicators, as well as fundamental indicators should be considered and at least included in the initial data set. Furthermore, it is highlighted that when building machine learning-based stock forecasting/trading models, the dynamic nature of the market, should be taken into account. Chapter 3 places these identified gaps as research questions to be investigated, proposes a framework aimed at facilitating such investigations, and discusses the details of the experiments conducted with respect to these investigations.

# Chapter 3. Proposed Framework and Experiment Setup

Chapter 3 introduces a framework that is used to carry out the investigations along the research questions stated in Section 1.1 and describes the set-up of the experiments conducted to investigate the research questions posed. The chapter is organized as follows: Section 3.1 describes the proposed framework, Section 3.2 provides the details of the experiments and framework implementation, and Section 3.3 concludes the chapter.

## 3.1 Proposed Framework

Based on Chapter 2, the below-listed dimensions/factors were identified as critical aspects when attempting to forecast financial time series through machine learning methods:

- What input(s) to provide to the machine learning method(s)?
- Which machine learning method(s) to use, and how to determine the best performing architecture?
- How to identify and account for the states of the overall stock market?
- What output to forecast?

As these key questions were being answered and experiment design was taking shape for the simulations, a framework came about which combined all the investigations under one roof and provided an ability to compare various scenarios to each to other.  Specifically, the framework:

- identifies and selects the best performing forecasting model through identification of the relevant inputs affecting the stock price
- accounts for the sensitivity of the stock price to various states of the market

Figure 3-1 shows the various layers making up the proposed framework: Input, Market State, Model and Output.

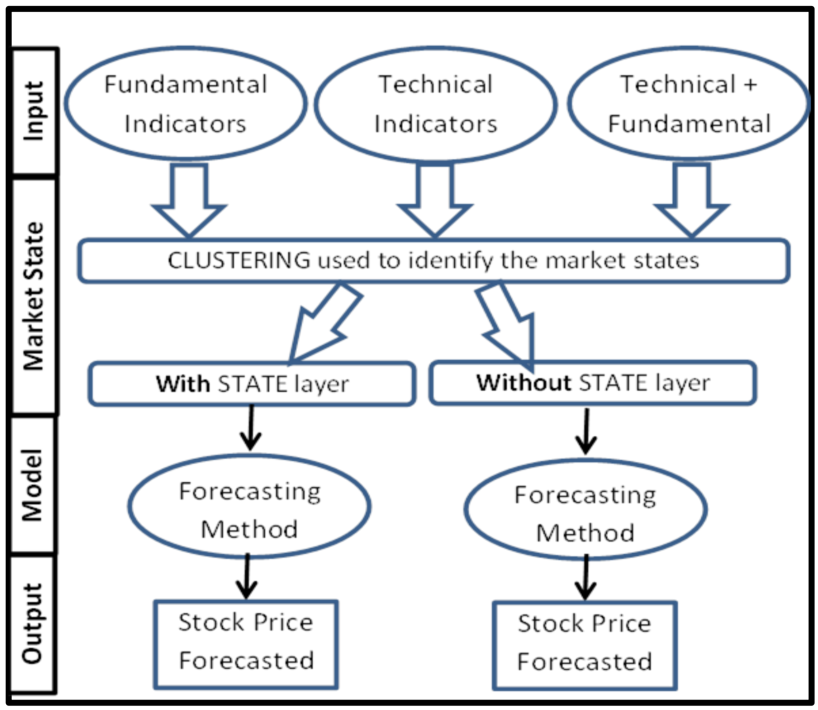


**Figure 3-1 Layers of the proposed Framework**

The Input layer determines the type of financial indicator(s) to be provided to the model layer. It is categorized into technical indicators (based on technical analysis), fundamental indicators (financial variables based on fundamental analysis) and their concatenated combination. This layer will help determine the relevant input set for the stock in question. This layer will be used to answer research questions Q1 and Q2. The Market State layer will be utilized to identify the various moods/states of the stock market, and to determine if the stock in question is sensitive to the various states of the market. During the inclusion of the market state in the forecasting process, a clustering algorithm is used to identify the various states of the overall stock market, and then associated forecasting models for the target company will be developed for each state of the market. Furthermore, the framework also accommodates stocks that might not be affected by the different states of the market. This layer will be used to research question Q3. The Model layer includes machine learning method(s) (e.g. ANN, SVM, etc.) utilized in forecasting the output, and the output layer represents the forecast generated. In order to answer research question Q4, the framework was implemented as part of the experiments. Section 3.2 discusses the details of the experiments conducted to answer the research questions posed and the implementation of the framework.

## 3.2 Experiment Setup

### 3.2.1 Overview

The framework has been implemented for 147 companies to predict percentage change in a selected company's stock price in the next year (252 trading days out), using various machine learning methods exposed to technical indicators, fundamental indicators and their combination, and to states of the stock market. Predictive performance of the framework, as measured by Root Mean Square Error (RMSE) has been compared against the random walk model as well as the ANN model using technical indicators. The following sections provide details with regards to the implementation of the framework: Section 3.2.2 describes the process of determining the companies covered, Section 3.2.3 describes the various inputs used at the input layer, Section 3.2.4 describes the various machine learning methods used at the model layer, and Section 3.2.5 describes the approach used towards implementation of the framework and experiments.

### 3.2.2 Companies Covered

The decision of which companies were included in the study was a result of trying to find a balance of these two factors: maximizing the number of companies which had sufficient input data present and also maximizing the duration of timeline available on which to conduct the experiments. S&P 500 index is a list of companies which holds the largest companies operating in the US. These companies make up roughly 80% of the overall market value in the US (https://us.spindices.com/indices/equity/sp-500) and the S&P 500 index serves as an indicator of the overall market. This list was used as a starting point and financial information (both technical and fundamental) on all the companies in this index were retrieved. As expected, the fundamental data, specifically the Analyst estimates from the IBES database, did not go as far back as the technical data did. Furthermore, in order to robustly test out the research questions stated in Section 1.1, a sufficiently long enough time period was selected so as to ensure that there was some market turbulence. Table 3-1, for example, shows the number of companies which had sufficient "fundamental indicator" data available on the company being forecast based on various timeline starting dates.

| Starting Date | Jan 1991 | Jan 1994 | Jan 1996 | Jan 2005 |
|---|---|---|---|---|
| Number of Companies | 275 | 329 | 342 | 440 |

Table 3-1 Number of companies with available data

As a result of the review of the data, companies and various significant market turbulences experienced over years, 147 companies[1] were selected that which had information available from the beginning of 1996 to the end of 2015.

### 3.2.3 Inputs

Once the companies and the timeline for the experiments were selected, input data was collected from various sources, cleansed, and transformed as needed for the experiments. Every company in the study had the following unique input sets defined: technical indicators, fundamental indicators, combined indicators, technical indicators with feature selection, fundamental indicators with feature selection, and combined indicators with feature selection. The following subsections detail the various steps followed, and assumptions made in creating these input sets.

#### *3.2.3.1 Technical indicators*

For each company in the study, end of day stock price data (Open, High, Low, Close, and Volume) was retrieved from Quandl (Quandl.com,2016) for the time period needed for the study. As described in Romero and Balch (2014), from the financial time series data retrieved, instead of the raw data, an adjusted[2] set of prices was preferred since the adjusted set of prices puts the whole of time series data on the same level by factoring in any corporate actions (such as dividend payments, stock splits, etc.). It was necessary that there

[1] Table A.1 shows the distribution of companies per industry, as well as the tickers of the companies.

[2] https://www.investopedia.com/terms/c/corporateaction.asp

was a data point for each date that was selected for the simulations. For the cases with missing data, the average of the data from the closest available trading days was used (Romero and Balch, 2014). Functions available in the TTR library (Ulrich, 2015) were used to calculate the needed technical indicators. Table 3-2 shows the list of technical indicators picked based on the coverage of technical analysis in Patel et al. (2015), and Thomsett (2015) and also the parameters used (mainly the defaults in TTR library (Ulrich, 2015) in generating them where relevant.

| |
|---|
| Average True Range (ATR) over a period of 14 days. |
| Moving Average Convergence Divergence (MACD) with simple moving average method and 26 days & 12 days for the slow and fast periods respectively. |
| Money Flow Index (MFI) over a period of 14 days. |
| FastK and FastD values of Stochastic Oscillator using 14,3, and 3 days for FastK, FastD, SlowD respectively. |
| Directional Movement Index (DMI) using 14 days |
| Commodity Channel Index (CCI) using 20 days, and 0.015 as the constant to apply to the mean deviation. |
| Relative Strength Index (RSI) using 14 days and weighted moving average. |
| Price Rate of Change (ROC) over 252 trading days. |
| The Chaikin Accumulation / Distribution (AD) line. |

**Table 3-2 List of technical indicators**

#### 3.2.3.2 Fundamental indicators

The fundamental indicators used in the experiments can be categorized into various groups:

- those dealing with the performance of the company in question,
- those related to direct competitors,
- those related to the industry the company in question belongs to,
- and macroeconomic indicators.

*Company related data*

For each company, IBES (2016) estimates were used to calculate the daily values for the Earning's Yield Ratio, short term and long-term EPS growth rates. As a high-level overview, the steps that were carried out were as follows, which are described in further detail in the following paragraph:

- Retrieve the median EPS estimates (1yr, 2yr, and long-term growth%) from the IBES database (2016)
- Convert these from monthly to daily frequency
- Convert EPS 1 year out estimate into an Earnings Yield Ratio
- Calculate a proxy for short-term EPS growth rate (EPS 2 year out / EPS 1 year out)

In order to help guide investors and provide insight, financial analysts provide on a monthly basis their recommendations on buying/holding/selling stocks of a select set of companies (mainly companies which can be influential for the economy). The analysts base these recommendations on their expectations of how in the short and long-term fundamental performance drivers of the company (e.g. Sales, Expenses, Book Value, Earnings Per Share, Dividend payout ratio, etc.) will fare in general and with respect to their competitors and their industry. The IBES database (2016) contains these monthly announced forecasts by financial analysts on companies as well as their recommendations on buying/holding/selling the stock. Although there are many fundamental indicators (e.g. Price to Earnings Ratio, Price to Book Ratio, Sales, Financial Ratios such as ROE, ROA, etc.) which have been linked to performing fundamental analysis (Thomsett 2015), the majority of the data was sparsely available for these fundamental indicators with the exception of Earnings Per Share (EPS) related forecasts from the analysts. For the companies selected and the study period, EPS is available in a relatively consistent manner. Therefore, the median of the monthly estimates by financial analysts for EPS 1 year and 2 years out, as well as long-term expected growth percentage in EPS, were retrieved over the study period for each company. As these estimates were only available on a monthly basis, their frequency of occurrence was converted to daily by using "the last observation carry forward" method (Ryan and Ulrich, 2014), where until the next release of the monthly estimates became available, the last available estimate was used for each trading day in between the estimate announcements. Once the data was retrieved, cleansed and put in the form of a financial time series, the data was transformed as explained further.

The first transformation was applied to the EPS 1 year out estimate figure by dividing it by the prior day's stock closing price data P, effectively providing an Earnings Yield (EY) ratio EPS/P (http://www.investopedia.com/terms/e/earningsyield.asp) in a daily format. The transformation had two purposes: (1) to convert a monthly estimate into a daily ratio, and (2) to generate a metric that can be used for stock price valuation. In the initial tests, Price to

Earnings (P/E) ratio which is the inverse of EY (EPS/P) was used. Even though P/E is a popular and widely mentioned multiplier, it proved to be an unreliable data set for the tests. This was due to the EPS figure being zero (or close to it) in numerous cases which resulted in very large spikes in ratios or invalid numbers (i.e. division by zero). As stated in Thomsett (2015), the growth rate is of particular concern for the stock valuation. In order to get a view on the short term expected growth in EPS by the analyst, the second transformation was done by dividing EPS 2 years out with EPS 1 year out. The long term EPS growth rate was not further transformed and was used as available.

Usually, public companies disclose their financial figures (Sales, Earnings, etc.) on a quarterly basis. Thus, the company-related data becomes available every 3 months, and at the end of the fiscal year, the values for the whole past year are consolidated. In order to reflect this delay in the arrival of information, Vanstone and Finnie (2009) had suggested that fundamental data, in general, should be displaced by a certain amount of time (e.g. by about 3-6 months) and provided to the machine learning method at the later time, so that data that was unavailable at the time are not provided to the model. The company-related data included in the experiments were based on expectations of the analysts historically at that particular point of announcement and therefore the time displacement suggested by Vanstone and Finnie (2009) has not been deemed to be necessary.

*Competitor-related data*

In addition to collecting the analyst estimates for the target company, the analyst estimates were also collected for the competitors of the target company. The first step was to define who the competitors of the companies included in the experiments were. During the initial phases of data collection for simulations, the competitor lists as provided by Yahoo Finance (2016) were utilized; however, this functionality has been discontinued in recent years. The Thomson One (2016) database provides a large amount of financial summary information for companies including data from financial filings. For each company, the Thomson One (2016) database also provides a comparative table showing how each company is doing with respect to its competitors that are identified via their proprietary algorithm. For each company in the study the top two competitors (largest market capitalization) were retrieved from the Thomson One database. For each competitor retrieved, the EY ratio was calculated as described in "Company related data".

*Industry-related data*

The industry designation for the company was determined using the industry classification available on the Yahoo Finance website (2016). The daily index price data for each corresponding industry was retrieved from the MSCI USA IMI SECTOR INDEXES website (2016). In order to smooth out the data, the TTR library (Ulrich, 2015) was used to transform the raw price data into moving average convergence and divergence (MACD) indicators for short term (with 26 days and 12 days) and medium-term (with 126 days and 12 days).

*Macroeconomic indicators*

Unlike the company-specific data, the macroeconomic indicators represent movements in the overall economy and are the same for all companies in the study. Given the global world economy, one such indicator was based on the foreign currency data, where the daily value of "Trade Weighted U.S. Dollar Index against Major Currencies" was transformed with MACD (126 days and 12 days) functionality available from the TTR library (Ulrich, 2015). Another macroeconomic indicator selected was the “S&P 500 futures data” whose daily value was transformed further using MACD (26 days and 12 days) functionality available from the TTR library (Ulrich, 2015). The final macroeconomic indicator used was derived from the ratio of the 10 year to 2-year constant maturity rate which was transformed using MACD (26 days and 12 days) functionality available from the TTR library (Ulrich, 2015). Quandl (2016) was used as the data source for foreign currency and futures data, whereas the FRED website (2016) was used to retrieve the treasury rate information. Table 3-3 shows the list of fundamental indicators used.

| |
|---|
| Earnings Per Share (EPS) 1 year out for the company / Price |
| (Earnings Per Share (EPS) 2 years out ) / (Earnings Per Share (EPS) 1 year out) |
| EPS long term growth rate percentage |
| Earnings Per Share (EPS) 1 year out for competitor 1 / Price for competitor 1 |
| Earnings Per Share (EPS) 1 year out for competitor 2 / Price for competitor 2 |
| Daily MSCI industry index prices (MACD, 252 days, 12 days) |
| Daily MSCI industry index prices (MACD, 26 days, 12 days) |
| S&P 500 Futures prices (MACD, 252 days, 12 days) |
| Daily Trade Weighted U.S. Dollar Index against Major Currencies (MACD, 252 days, 12 days) |
| 10 year to 2-year constant maturity rate (MACD, 26 days, 12 days) |

**Table 3-3 List of fundamental indicators**

#### *3.2.3.3 Combined indicators*

The combined indicator set was created by concatenating the technical and fundamental indicators described in Sections 3.2.3.1 and 3.2.3.2 respectively. The resulting indicator set consisted of 20 indicators (10 technical + 10 fundamental), without any type of modification to the original data sets such as giving more weight to some variables over others. This combined indicator set is meant to represent the case where the technical analysis based indicators and the fundamental analysis based indicators are used together and at the same time without giving preference to any one of the schools of thought over the other. Thus, it was necessary to keep the indicators that became part of the combined indicator set to be the same as the ones making up the technical and fundamental indicators described in Sections 3.2.3.1 and 3.2.3.2 respectively, so that research question Q2 can be investigated. As demonstrated in Figure 3-2, the technical and fundamental input sets were joined together into a larger input set by aligning the dates of the two daily financial time series data described in Sections 3.2.3.1 and 3.2.3.2.

| | Technical Indicators (10) |
|---|---|
| Feb 5 1996 | **T1, T2, T3, T4, T5, T6, T7, T8, T9, T10** |
| Feb 6 1996 | **T1, T2, T3, T4, T5, T6, T7, T8, T9, T10** |
| Feb 7 1996 | **T1, T2, T3, T4, T5, T6, T7, T8, T9, T10** |
| Feb 8 1996 | **T1, T2, T3, T4, T5, T6, T7, T8, T9, T10** |
| … | |

| | Fundamental Indicators (10) |
|---|---|
| Feb 5 1996 | **F1, F2, F3, F4, F5, F6, F7, F8, F9, F10** |
| Feb 6 1996 | **F1, F2, F3, F4, F5, F6, F7, F8, F9, F10** |
| Feb 7 1996 | **F1, F2, F3, F4, F5, F6, F7, F8, F9, F10** |
| Feb 8 1996 | **F1, F2, F3, F4, F5, F6, F7, F8, F9, F10** |
| … | |

| | Combined Indicators (20) |
|---|---|
| Feb 5 1996 | **T1, T2, T3, T4, T5, T6, T7, T8, T9, T10,F1, F2, F3, F4, F5, F6, F7, F8, F9, F10** |
| Feb 6 1996 | **T1, T2, T3, T4, T5, T6, T7, T8, T9, T10,F1, F2, F3, F4, F5, F6, F7, F8, F9, F11** |
| Feb 7 1996 | **T1, T2, T3, T4, T5, T6, T7, T8, T9, T10,F1, F2, F3, F4, F5, F6, F7, F8, F9, F12** |
| Feb 8 1996 | **T1, T2, T3, T4, T5, T6, T7, T8, T9, T10,F1, F2, F3, F4, F5, F6, F7, F8, F9, F13** |
| … | **T1, T2, T3, T4, T5, T6, T7, T8, T9, T10,F1, F2, F3, F4, F5, F6, F7, F8, F9, F14** |

**Figure 3-2 The technical indicator and the fundamental indicator sets are concatenated using the dates to make up the combined indicator set**

#### *3.2.3.4 Feature Selection methods*

As covered in Chapter 2, Principal Component Analysis (PCA) has been successfully used for dimension reduction in machine learning. PCA was applied to each of the data sets described in Sections 3.2.3.1, 3.2.3.2, and 3.2.3.3 in order to create additional data sets containing inputs that are most relevant. PCA reduced dimensionality by creating linear combination of the inputs, retaining only the portion necessary to explain a selected percentage of variance in the input data. In this case 90%, which was the default value for the “pre-process” function (Kuhn, 2014) in R, was used.

### 3.2.4 Machine Learning Methods

As stated in Section 2.2.3, ANN and SVR methods have been successfully utilized for similar type of problems and they are “the more common soft computing techniques applied in forecasting financial time series” (Cavalcante et al., 2016). Thus, among the many alternatives available (ANN, SVR, HMM, GA etc.), ANN and SVR were selected for implementation as machine learning methods due to their demonstrated success on this task and wide implementation. Furthermore, Decision Trees (DT) and Linear Regression (LR) were also implemented as simpler (with few or no parameters to set) alternative forecasting methodologies to compare against. As stated in Section 2.2.3, machine learning approaches can typically be classified into single, hybrid, or ensemble of models. Only single models were considered in this study as the focus of the research questions were on investigating the impact of the technical versus fundamental indicators and their combination, and also the impact of the states of the stock market on the forecasting process. Single models were deemed to be sufficient for investigating the research questions, and inclusion of hybrid or ensemble of models would have increased the model complexity (see Section 3.2.6) further. As stated in Section 2.2.3, deep learning methods were also being applied to financial forecasting, whereby the data is provided in relatively raw nature (e.g. Daily OHLC data per company) and through the various layers these models are able to learn relationships between the input features and the output(s) and extract features (Cavalcante et al., 2016). Given that the focus of research questions was as stated above, the shallow models which were provided with expert (technical and fundamental analysts) defined features were being utilized and compared to each other. Thus, the deep learning models were not implemented as part of this thesis. The models were implemented using R (R Core Team (2013)) and the libraries available in the open-source data mining tool WEKA (Hall et al., 2009).

With regards to the implementation of a neural network, a Multi-Layer-Perceptron (MLP), a feed-forward neural network using backpropagation, has been deployed. The choice for using MLP was based on its popularity and success as stated in Section 2.2.3. The neural network was built with sigmoid activation function and squared error as the loss function. Regarding the parameters of the neural network, the WEKA default parameters for the MLP were left in place with the exception of the following adjustments:

- Both the input and output were normalized to take on values between -1 and 1
- 20 % of the data was used as the validation set size during training
- Default value of 500 was used for number of epochs of training, but to prevent overfitting, early stopping is used where training is stopped if the validation set error gets worse for more than 10 instances in a row
- The decay option was set to True so that the learning rate would be decreased at each epoch

In addition to the above parameters, the ideal number of hidden neurons, the learning rate and momentum rate were decided based on several tests during the parameter optimization phase, which is described in more detail in Section 3.2.5.3. Table 3-4 shows the various parameters used in determining the architecture:

| | |
|---|---|
| # of Hidden Layers | 3,5,7 |
| Learning Rate | 0.05, 0.3, 0.6 |
| Momentum | 0.1, 0.3,0.7 |

**Table 3-4 Parameters Tested for ANN architecture**

For the SVR, the C and gamma values, as shown in Table 3-5, have been tested over several scenarios, which are described in more detail in Section 3.2.5.3, to determine the optimum model calibration:

| | |
|---|---|
| C Values | 0.125, 0.5 , 2 |
| Gamma Value | 0.01953, 0.125 ,0.5 , 1 |

**Table 3-5 Parameters Tested for SVR architecture**

In the case of DT and LR, the default parameters available with the WEKA libraries were used.

## 3.2.5 Implementation details of the scenarios and framework

### *3.2.5.1 Overview*

Using the inputs and machine learning models, forecasting scenarios per each company were run. Figure 3-3 shows the various phases that are carried out as part of the framework.

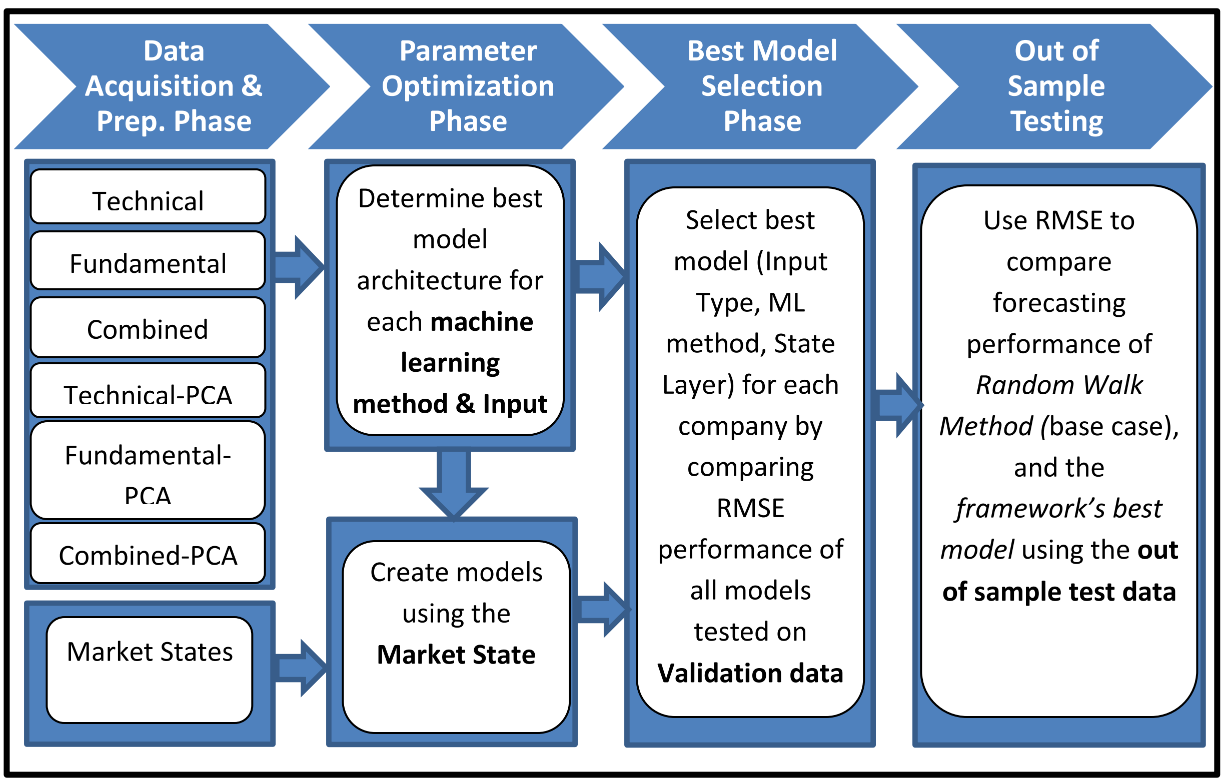


**Figure 3-3 High-Level Process Flow Overview for Framework**

The first phase involves collecting and preparing the relevant input sets, details of which have been covered in Section 3.2. The “Parameter Optimization Phase” is next; its goal is to identify the best performing architecture for each machine learning method and input set (e.g. technical, fundamental, combined) combination. Models using these architectures are trained and used to forecast on Validation data, with and without the inclusion of the state layer. The performance of these models is compared by the framework to identify the best performing model (that is, the one with the lowest RMSE on validation data set) per each company. During the “Out of Sample Testing” phase, this selected best performing model is again trained and tested on the reserved out-of-sample test data, whose performance is compared to that of the benchmark (Random Walk method).

#### *3.2.5.2 Training and Test Sets*

One of the crucial questions during the experiment design is to determine the number of data points needed, which in return determines the sizes of the training and test sets. Vanstone and Finnie (2009) suggests that in picking the number of data points “the main principle is to capture as much diverse market activity as possible (with a long training window), whilst keeping as long a testing window as possible (to increase shelf life and model confidence)” and recommends “sourcing at least 10 years data for each security, and then performing an 80:20 split”.  Similarly, Hyndman and Athanasopoulos (2014) stated that the typical size of a test set is 20%. Based on this guidance, 2,365 (roughly 10 years of data) was set as the data size for the simulations and the data was split into 80% training data (1,892 data points) and 20% testing data (473 data points).

One of the ways to ensure robustness with the experiments carried out is to use K-fold cross-validation (Kuhn and Johnson, 2013) repeated a few times in order to ensure that the results observed would be applicable under different conditions or trials. The classical K-fold cross-validation approach randomly splits the data into training and test sets based on the number of folds chosen. For example, 10-fold cross-validation would repeat 10 times of randomly splitting the data into training set (90%), and test sets (10%) to train and generate forecasts from the model. The performance of the model would be the average of the collective performance over these 10 repetitions. However, this random generation of the training and test sets means that the "classical" cross-validation approach cannot be used in modelling of financial time series forecasting. When doing financial time-series forecasting, it is important to separate testing data from training data such that the chronological order of the data is preserved (Torgo, 2017). The reasoning behind this is to ensure that the model is not prematurely exposed to information in the training phase (look-ahead bias), potentially producing unrealistically good performance. Given that "classical" cross-validation approach was not usable for time series, a modified version of the cross-validation is typically used (Torgo, 2017) which captures the essence of cross-validation whilst upholding the principle of making sure that the testing data was always used chronologically after the training data. Modelled after the approach described in Torgo (2017), a set of random starting points (10 in the case of these simulations) were generated, and from each starting point available data was split into training and test sets such that testing data was chronically after the training data. As stated at the beginning of Section 3.3.5.2, the Test Set size was set at 473 and the Training Set size was set at 1,892. This resulted in 10 training sets and 10 associated test sets which were randomly picked to generate data points under different market conditions and essentially emulate the principle of the robustness of 10 fold cross-validation.

The first step is to generate 10 random starting points. For each company, daily financial time series data from January 1996 through December 31st 2015 for the variables stated in Section 3.3.3 were collected. This resulted in 5,037 data points (20 years data x 252 trading days per year) per company, which represented the complete timeline of data points that were gathered. As explained in Torgo (2017) the data points shown in red in Figure 3-3 were ineligible to be considered as random start points, so that training and test data set sizes chosen, 473 and 1,892 respectively, can be adhered to once a random starting point was generated. If the ineligible data points were not excluded and random starting points were to be one of these points then there would not be enough points for constructing either the training or the test sets. Subtracting these ineligible points (473 + 1,892) from the 5,037 data points resulted in 2,672 trading days (shown in green in Figure 3-4) from which 10 random starting points were generated for the experiments.

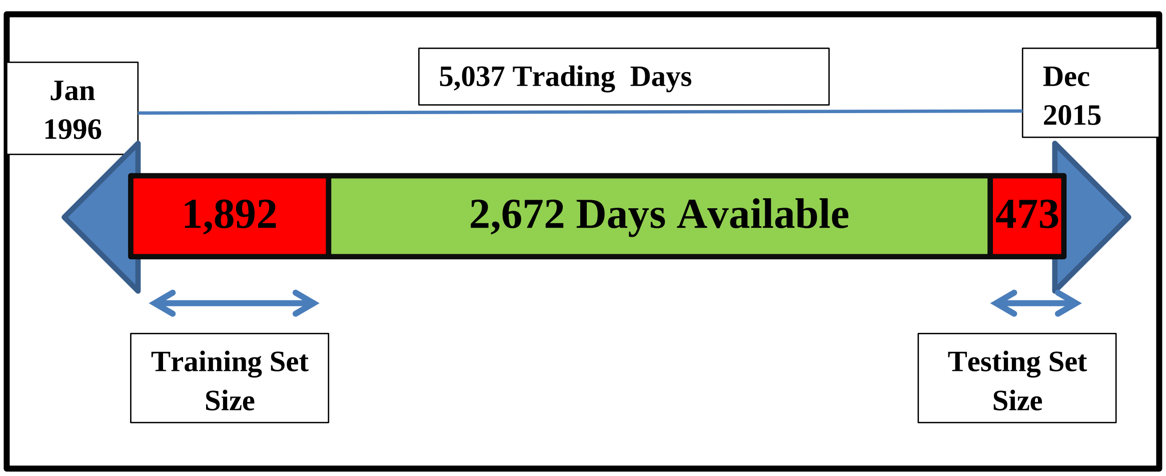


**Figure 3-4 Available timeline for random start point generation**

Once a random number generator was used to generate a random starting point (i.e. pick a number from 1,893 through 4,564 inclusive) from within the 2,672 available trading days, the data was split into training and test sets for the various phases (explained in beginning of Section 3.3.5.1 and shown in Figure 3-3) of the experiments. Figure 3-5 shows the approach followed for one such random starting point, which is denoted by "x" in the diagram. Once "x" is picked, 1892 instances chronologically prior to it would be designated as the training set for the Out-of-Sample Testing phase, and the point "x" and the 472 instances chronologically after it would be designated as the testing set for the Out-of-Sample Testing phase. As shown in figure 3-5, the training set for the Out-of-Sample Testing phase is itself further split using the recommended 80:20 (as described at the beginning of this section) ratio into the training and testing sets for Parameter Optimization & Best Model Selection phases.

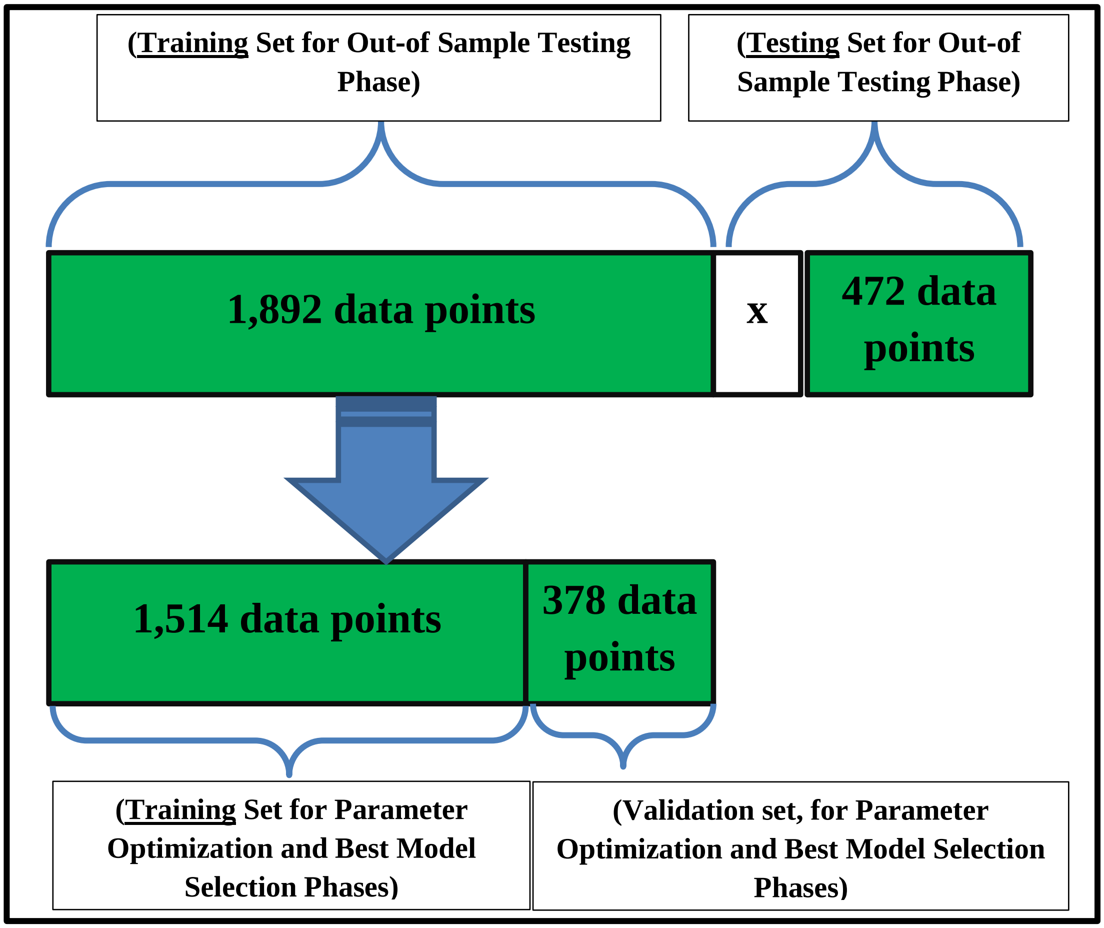


**Figure 3-5 Random start points for creating Test Sets**

Figure 3-6 overlays the 10 random points against the overall stock market performance (using Russell 2000 performance), where it can be seen that some test start points (marked in x) fall during market up-swings, and other start points occur during market downturns. This further illustrates that the approach taken did result in the models being tested under different market conditions.

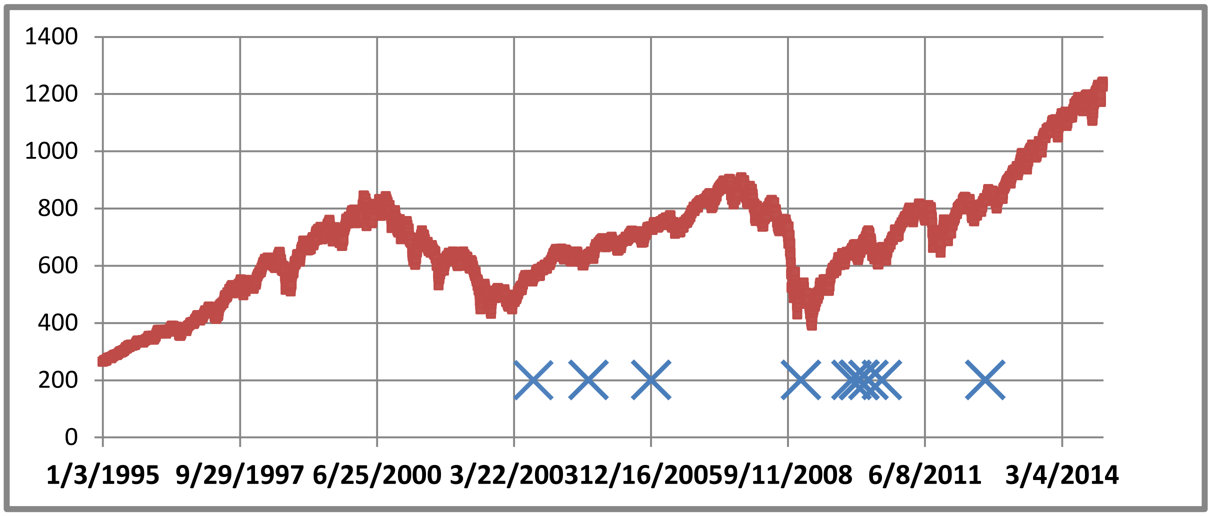


**Figure 3-6 Random test start points versus the Russell 2000 index**

#### *3.2.5.3 Parameter Optimization Phase*

The goal of this phase is to iterate through various model architecture options available for the machine learning models and determine the best parameter set for each machine learning method and input set (e.g. technical, fundamental, combined) combination. Thus, during the parameter optimization phase various parameters for the machine learning methods are used in training the models on the input data, and once the model has been trained, it is used to predict (i.e. to generate forecasts) on an unseen test data set (the Validation set, as shown in Figure 3-5). The error metric of RMSE was calculated by comparing the forecast values against the actuals from the unseen test data set. The architectures which yielded the lowest RMSE were selected as the best.

Figure 3-7 illustrates the “sliding window” approach used in model training and testing during the parameter optimization phase; this was implemented by Targo (2017). The size of the test sets was selected as 10, whereby a model is recalibrated with more recent data every 10 test instances. This was set at 10 to alleviate some of the computational cost of running through the large combination of parameter selection for each machine learning method.

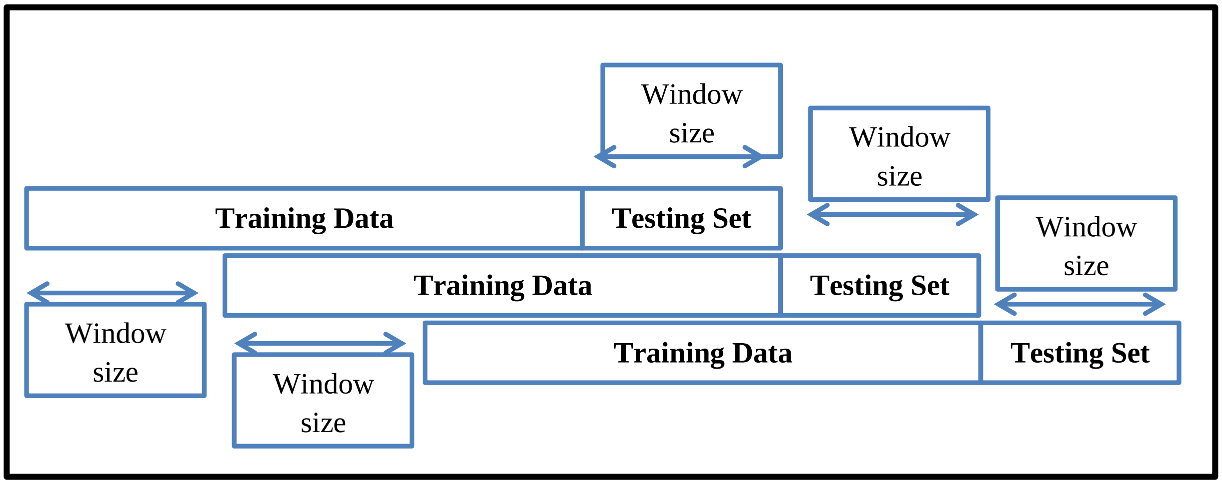


**Figure 3-7 Rolling Window approach for training and testing**

The output of the parameter optimization phase is a list of the best performing architectures per each unique triple combination of company, input type, and machine learning method.

### *3.2.5.4 Best Model Selection Phase*

After having defined the best performing model architecture for each machine learning method and input combination, these model architectures are used in the "Best Model Selection Phase". This phase covers two groups of scenarios: (1) models without the state layer, and (2) models including the state layer. Ultimately, the purpose of this phase is so that the Framework can determine the single best performing model per company, as defined by the input type, machine learning method, and state layer definition.

#### 3.2.5.5 Model without the state layer

In the case where the market state layer is not included in the model, the training set simply consists of the most recent observations from the company input sets (i.e. technical, fundamental, and combined) prior to the testing instance. This is similar to the approach utilized in the parameter optimization phase; however, in this case, the model recalibration was done on every single testing instance (i.e. window = 1 day).

#### 3.2.5.6 Model including the state layer

The inclusion of the state layer takes a different approach towards how the composition of the training set is determined. Similar to the approach taken in building the model without the state layer, the beginning point is again the testing instance at hand. However, instead of using the most recent company information as the training data, the training set is formed by taking into account the state of the market, where only the training data from dates which exhibit a similar market mood are used. Among the market sentiment indicators described by Achelis (2000) VIX, SP500 index (Relative Strength Index), and Put to Call ratio were selected as alternative market sentiment indicators for the experiments. The values for these indicators were collected for each date that is part of the study to form a time series of each market mood indicator. To account for the state of the market, firstly the date of the testing instance is used to retrieve the values of the market sentiment indicator (e.g. VIX) on that particular date and back to the first available date (i.e. January 1st 1996). The values of the market sentiment indicator on the dates prior to the testing date are provided to a clustering algorithm to create the various moods that the overall stock market exhibited. With respect to the clustering algorithms, the “kmeans++” clustering algorithm (Arthur and

Vassilvitskii, 2007) which address the initialization related issues that plain vanilla “kmeans” clustering algorithm has. Figure 3-8 illustrates this step of clustering the market mood indicator values, the output of which is a set of clusters and all the trading days that fell under each cluster.

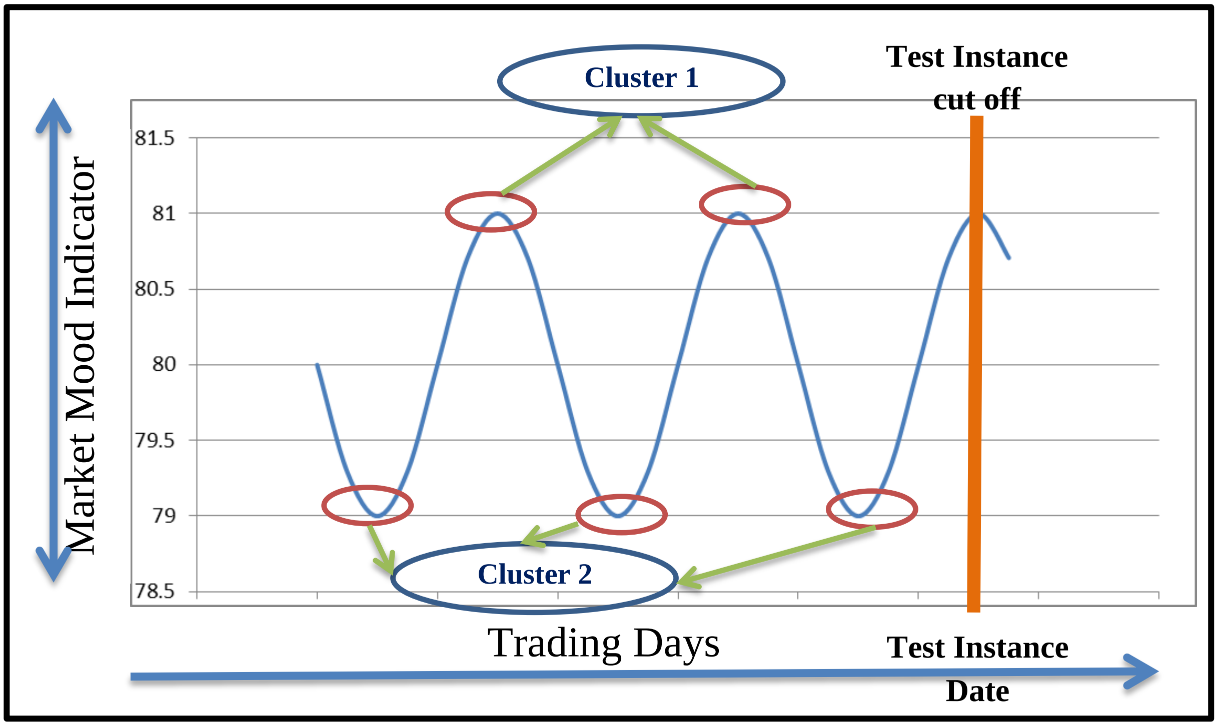


**Figure 3-8 Use Market Mood indicator values to create distinct clusters to represent moods of stock market**

The next step is to identify the cluster to which the market sentiment indicator from the testing date belongs to. The distance (Euclidian) between the value of the market sentiment indicator and centroids of each clusters are calculated, and the testing date is assigned to the clusters where this distance is minimized (i.e. it is closest to). This step is illustrated in Figure 3-9 and it identifies the “active” state or the mood of the market exhibited on the testing date.

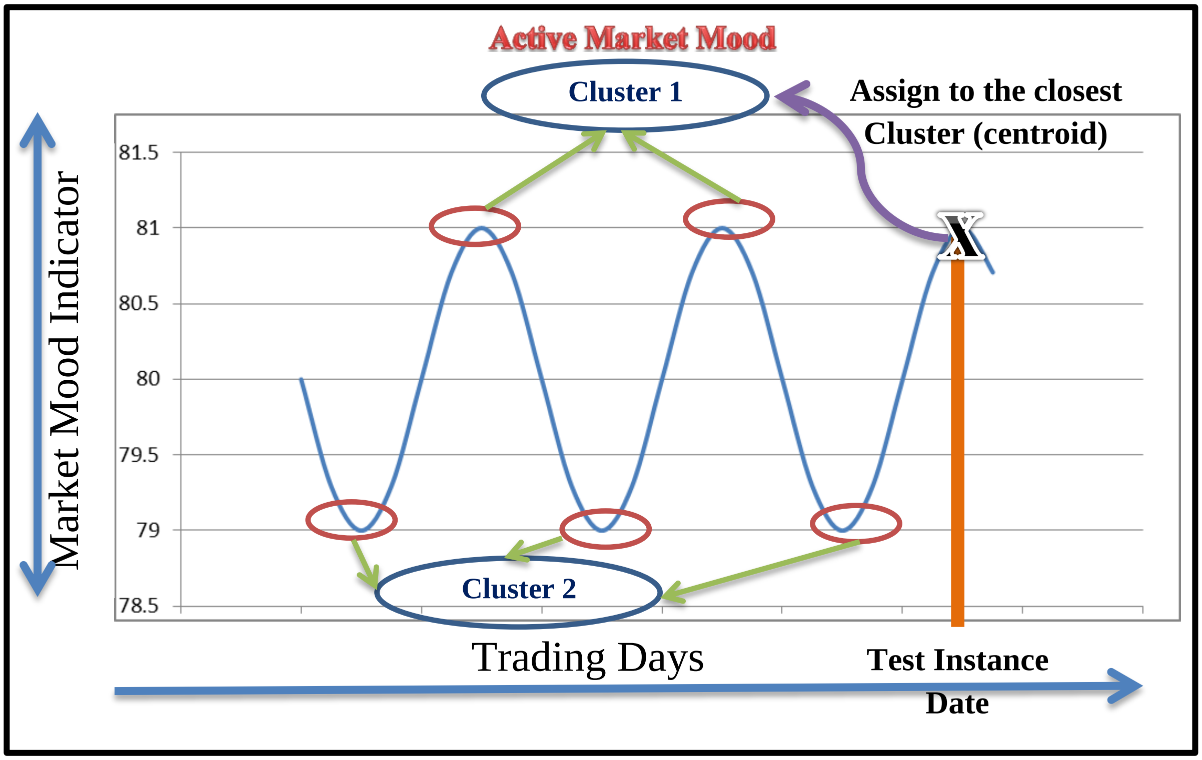


**Figure 3-9 Assign the market mood indicator value to the closest centroid to identify the market mood exhibited (i.e. 'active') on test date**

The following step is to explicitly associate the mood of the market with the training instances that are made available to the forecasting model. Having identified the 'active' market mood, all the previously captured trading days that fell under this market mood (see Figure 3-8) are used to retrieve the instances from the training set of the company's input set (technical, fundamental, or combined). This effectively filters the training set to contain only the instances where the market mood exhibited is the same (as represented by the 'active' market mood). Figure 3-10 shows the approach taken to identify the training instances that will be used to train the forecasting model.

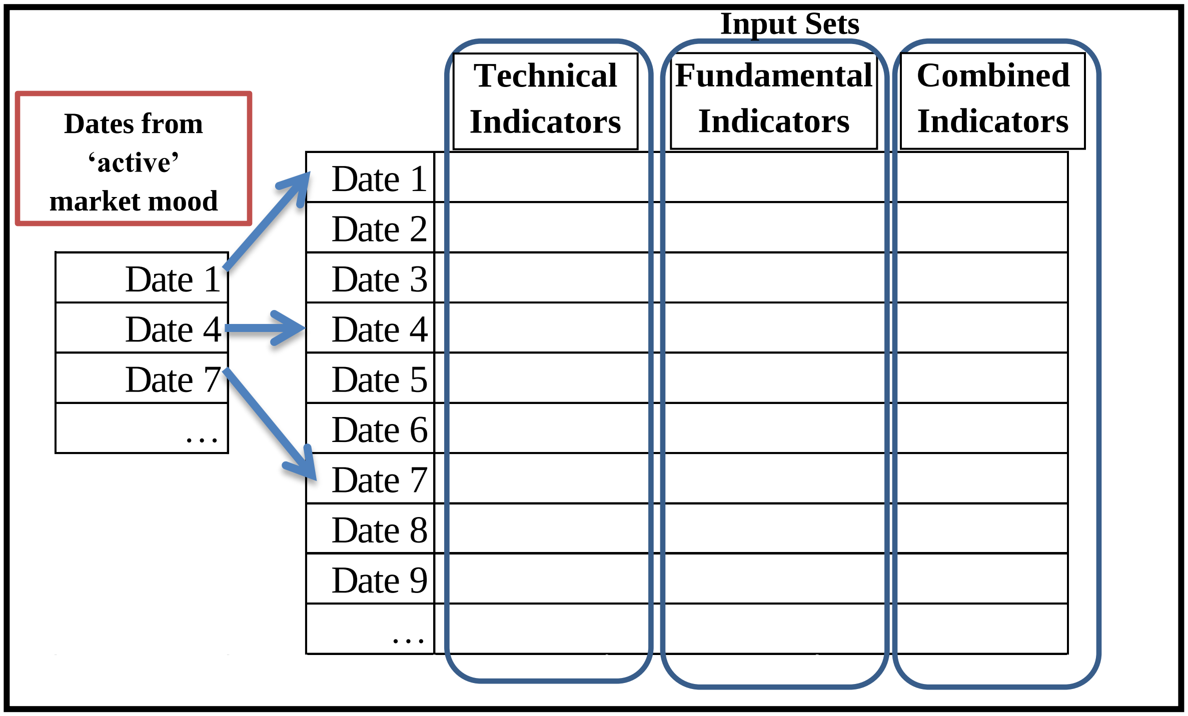


**Figure 3-9 State Mapping the dates from active market mood to the input features using the trading days**

This approach does result in a non-uniform set of instances in terms of size of the training set, depending on the number of trading days that made up the 'active' market mood. As shown in Figure 3-5, depending on the phase of the implementation, the training set sizes were 1,892 or 1,514. In order to keep the training set size the same (i.e. to have comparable results) throughout the experiments, the training sets generated during the implementation of the state layer had to be further adjusted (increased or decreased) to match these two training set sizes (depending on the phase of the implementation). In the cases where the size of the training set matching the active market state was larger than the training set used in the remainder of the experiments, the suggested training set was reduced to the training set size used in the remainder of the experiments where only the most recent portion of the data was kept. Thus, the instances identified by the active market state were first sorted by date (e.g.

from oldest to the newest) and only desired number (e.g. 1,892 or 1,514 depending on the stage) of the newest instances were selected to make up the training set. The remaining instances were ignored. Figure 3-11 demonstrates this approach.

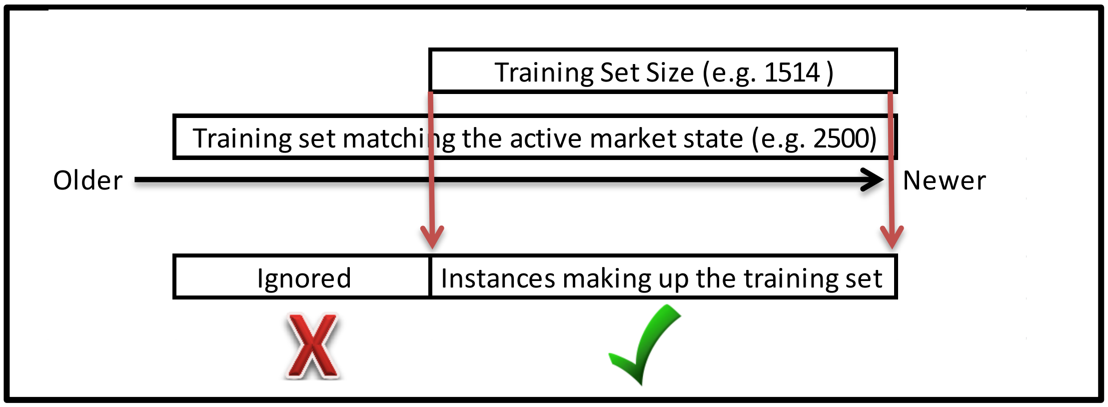


**Figure 3-10 Reducing the training set suggested by the active state to the size used throughout the experiments**

In the cases where the reverse is true, the training data matching the active market state was first sorted by date (oldest to newest) and the data set was replicated until the training set size reached the same number of instances as used in the remainder of the experiments. Figure 3-12 demonstrates the approach taken.

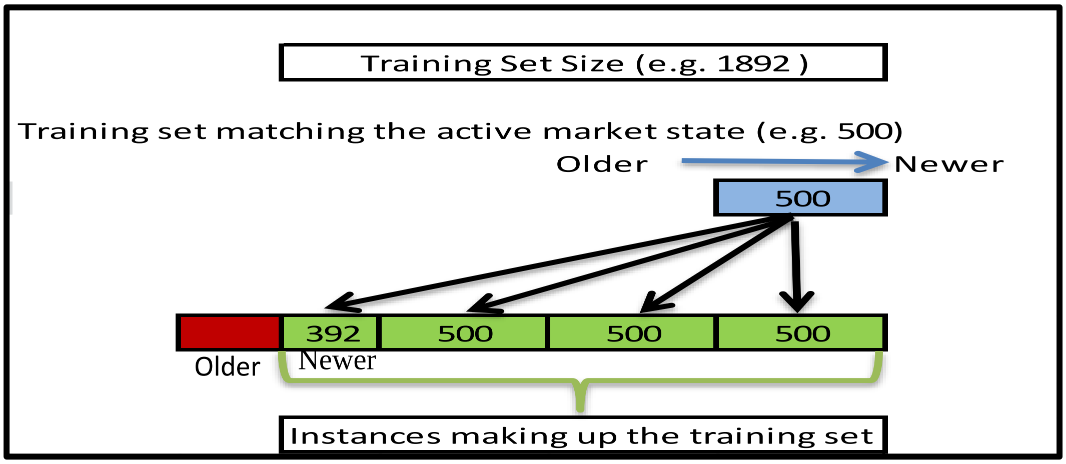


**Figure 3-11 Increasing the training set suggested by the active state to the size used throughout the experiments**

One of the parameter decisions to be made was the number of clusters, which would effectively represent the number of states assumed by the market. As mentioned in Chapter 2, stock markets can generally be considered as exhibiting an upward trend (e.g. "Bull Market"), or downwards trend (e.g. "Bear Market"), or stationary (i.e. side-way movements). Based on this view of the stock markets, 3 was selected as a starting point for the number of clusters (i.e. minimum number of clusters). As stated in Chapter 2, Munnix et al. (2011) observed the stock market to exhibit 8 distinct states during the time period overlapping with the one used in the experiments. The choice of the maximum number of clusters was based on this study. Based on these (3-8) reference points for the potential number of states, the numbers of clusters selected were 3, 5 and 7, as exemplar examples with progressive step change of 2. For each company input and machine learning combination, forecasting models with state layer using cluster sizes of 3, 5, and 7 were implemented. The following is a pseudo-code representation providing a high-level summary of the approach taken with the state layer implementation algorithm:

FOR each test point of the company data

{

- Retrieve all available timeseries data for the market mood indicator (e.g VIX) prior to date of the test instance
- Generate clusters for the market mood indicator based on the predefined number (3,5,7)
- Assign value of market mood indicator on the test point date to closest cluster (relevant market mood) using Euclidean distance to cluster centers
- Determine the historical dates included in the relevant market mood cluster
- Expand or Filter down these dates to the desired training set size
- Train the forecasting model only on the historical data points of company timeseries data belonging to the relevant market mood
- Generate a forecast for the test instance, and capture the forecasting error

}

**Figure 3-12 Pseudo-code for implementation approach of model's state Layer**

### *3.2.5.7 Out of Sample Testing Phase*

The "Best Model Selection Phase" allows the framework to compare and pick the best performing model per company (defined through input type, machine learning method, and state layer definition). The "Out Of Sample Testing Phase" provides the ability to run these best models on previously unseen data so that forecasting performance can be robustly assessed. The performance (RMSE) of the models in this phase will be utilized to answer the research questions stated in Section 3.2. From an implementation point of view, the approach used for this phase is the same as the cross-validation approach that is used as part of the "Best Model Selection Phase." In addition, in the "Out Of Sample Testing Phase," the models were recalibrated at each testing instance (i.e. the rolling window with 1 day based on the approach defined by Torgo, 2017). However, the data used for training and testing is different in this phase, as outlined in Section 3.3.5.2. In essence, the approach taken during this phase follows the same lines of training the models on training data and using these models to predict forecasts on "out-of-sample" test data (Figure 3-5). Finally, the performance (RMSE) of the models on this test data is captured for analysis and compared against a benchmark.

#### *3.2.5.8 Benchmark Models*

The benchmark models utilized throughout the analysis derived from the related research questions posed in Section 1.1 and can be broadly categorized as: the random walk method and the machine learning base model. A benchmark used across many of the research questions is the random walk method. The use of the random walk method as a benchmark was based on two main factors. Firstly, the random walk method has wide implementation and acceptance in the Finance domain as described in Section 2.1.1, and secondly, the random walk method has been utilized as a benchmark by machine learning researchers as stated in Section 2.2.4. As described in Section 2.1.1, EMH is ubiquitous in the world of Finance and supports the view that stock prices follow a random walk model. The random walk method is engrained in Finance to such as extent that Hull (2009) described it as "the most widely used model of stock price behaviour." It, therefore, made sense to use random walk as a benchmark from the point of the research questions, particularly with respect to RQ1 & RQ2. As stated in Section 2.1.1, if the random walk method:

- outperformed the machine learning-based models using the technical, fundamental, or combined set of indicators, this would support the strong form of EMH,

- outperformed the machine learning-based models using the technical indicators, but could not outperform the models using the fundamental and combined set of indicators, this would indicate a semi-strong form of EMH,
- could NOT outperform the machine learning-based models using the technical, fundamental, or combined set of indicators, this would support the weak form of EMH.

For the purposes of the experiments, the random walk model has been implemented as described by Hull (2009) for each company, where 16,384 Monte Carlo iterations were used to generate sample price paths to forecast each test instance from the out of sample test set. According to Hull (2009), Equation 1 (also known as Geometric Brownian Motion (GBM)) can be used to model the expected change in the stock price over a small time interval.

$$\Delta S = \mu S \Delta t + \sigma S \epsilon \sqrt{\Delta t}$$

| | |
|---|---|
| **ΔS:** | Change in the Stock Price |
| **S :** | Current stock price |
| **$\Delta t$:** | Small time interval (in years) |
| **μ:** | Annual expected Rate of Return from Stock (%) |
| **σ:** | Annual Volatility of the Stock |
| **$\epsilon$ :** | Random drawing from standardized normal distribution Ø(0,1). |

**Equation 1 - GBM formula**

As pointed out by Hull (2009), the risk-free rate can be used for the expected rate of return for stocks (μ) when the investors are assumed to be risk-neutral - this assumption was used in the experiments. In estimating volatility, for a pre-determined time period, the stock prices are observed at regular intervals (daily in this case) and the standard deviation is captured from this sample and annualized (assuming 252 trading days) as described by Hull

(2009). For a given stock price, the equation for the GBM model can be used to calculate expected stock price (stock price today + Expected change in price) at a future point in time T (T= $\Delta t$). For the purposes of the experiments, this was set at daily increments ($\Delta t=1$) until reaching 252 days (T).

With respect to RQ3, comparisons of machine learning-based models were used as the benchmark and not the random walk method. This was done following the approaches mainly taken by machine learning researchers as described in Section 2.2.4. For example, in ascertaining whether implementation of the state layer was of added value, models not using the state layer (base) were compared to models using the state layer (proposed enhanced version) whilst keeping other factors (e.g. inputs, machine learning methods, etc.) the same.

With regards to RQ4, two main benchmarks were utilized. The first one is the random walk method as already described. The second benchmark utilized were ANN model which used technical indicators only. The reasoning for this to be selected as one of the benchmarks to compare is based on how widely and successfully this approach has been implemented by machine learning researchers in financial forecasting domain. Specifically, the reasoning for this is based on the extension of the trends of technical indicators being the most often used by machine learning researchers (as stated in Section 2.2.1) and ANN models being the most often deployed methods (as described in Section 2.2.3).

#### *3.2.5.9 Error metric*

When choosing the error measurement metric to be used, the main consideration given was the nature of the task at hand. Specifically, given that the output being forecast (% change in stock price) was of continuous scale (i.e. regression task) and not a categorical one (i.e. classification task), only error metrics suitable for a regression problem were considered and ones suitable for classification tasks were ignored. As mentioned in Section 2.2.5, statistical error measures (e.g. MAPE, MSE, RMSE) or economic / profit-oriented (such as Hit Rate, Average Annual Profits) error measures are typically put to use by researchers in financial time series forecasting research. As far as the scope of the experiments was concerned, the focus was limited on the forecasting performance and not necessarily on the trading performance. Expanding the scope to include the trading performance would require taking the generated forecast and making further assumptions on (trade entry/exit, slippage costs, etc.), as stated in Section 2.2.5. Given that the focus of the experiments were set on forecasting performance only, the economic / profit-oriented measures were not undertaken, and the statistical error measures were utilized. From the statistical error measures, MAPE was not used because the actual output values (% change in the stock price) could take on zero or near-zero numbers, and such cases would make the MAPE values be very large or undefined (Hyndman and Athanasopoulos, 2014). RMSE was selected as it has the advantage of being measured on the same scale as the output being forecast. Equation 2 shows how RMSE is calculated. For each testing instance (i), the difference of the actual output ($\boldsymbol{y_i}$) and the predicted value ($\boldsymbol{\hat{y}_i}$) is squared, and the summation of these squared differences for the test set instances (of size N) are averaged before finally taking the square root of this average.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N}}$$

Equation 2 - RMSE formulae

### 3.2.6 Summary of Experiment Setup

147 companies were selected to be included in the experiments. During the parameter optimization phase, for each company, forecasting models with the set of scenarios shown in Table 3-6 were built and used for forecasting, where RMSE were captured:

| **Model Layer** | **Scenarios** | **# of unique cases** |
|---|---|---|
| Input | technical, fundamental, combined | **3** |
| Machine Learning | | |
| | ANN (3 Hidden nodes x 3 Learning rate x 3 momentum) | **27** |
| | SVR (3 C values x 4 Gamma values) | **12** |
| | DT | **1** |
| | LR | **1** |

Table 3-6 Scenarios parameter optimization phase

For each company, this resulted in 123 (3 unique inputs x 41 different machine learning method options) unique models being trained and tested on 3780 test instances. As described in Section 3.2.5.3, the model was retrained (recalibrated) every 10 testing instances. Thus, during the parameter optimization phase, for each company models were trained for 46,494 (123 x 378) times. The goal of this phase is to identify the parameters that yielded the lowest RMSE on the validation data (see Section 3.2.5.2) and provided the model make up (i.e. parameters and architecture) for each input (technical, fundamental, combined) and machine learning method (ANN, SVR, DT, LR) combination per company. Thus, each company had 12 unique model definitions (3 inputs x 4 machine learning methods). In terms of complexity, each machine learning method and input combination took varying amounts of time to process. Table 3-7 displays the length of time (hours) it took to train and test for the 3780 instances for one company per each machine learning and input combination.

| **ML Method** | **Technical** | **Fundamental** | **Combined** |
|---|---|---|---|
| ANN | 2.2 | 4.4 | 17.6 |
| SVR | 2.7 | 5.4 | 21.6 |
| DT | 0.8 | 1.6 | 6.4 |
| LR | 0.3 | 0.6 | 2.4 |

**Table 3-7 Model training times by ML method and input type during the parameter optimization phase**

These 12 unique definitions were then used in the Best model selection phase, during which the set of scenarios in Table 3-8 were built and used for forecasting:

| Model Layer | Scenarios | # of unique cases |
|---|---|---|
| Input | technical, technical with PCA, fundamental, fundamental with PCA, combined, combined with PCA | **6** |
| Machine Learning | ANN, SVR, DT, LR | **4** |
| State Layer | No State Layer, With State Layer (VIX_3, VIX_5, VIX_7, S&P500_RSI_3, S&P500_RSI _5, S&P500_RSI _7, Put_to_Call_3, Put_to_Call_5, Put_to_Call_7) | **10** |

**Table 3-8 Scenarios for Best model selection phase**

For each company, this resulted in 240 (6 unique inputs x 4 different machine learning method options x 10 state layer options) unique models being trained and tested on 3780 test instances. Due to the way the state layer was implemented (see Section 3.3.5.6) the model was retrained (recalibrated) for every testing instance. Thus, during the Best model selection phase, for each company models were trained for 907,200 (240 x 3780) times. Based on the comparison of the predictive performance (i.e. RMSE) of the 240 models, the framework identified best performing (i.e. picked the model with the lowest RMSE across the 240 models on the validation set) model for each company. The output of this phase is to define (input + machine learning method and parameters + state layer option) one unique model that is picked by the framework and put forth as having the highest likelihood at being able to predict the percentage stock price change of that company. With regards to model complexity, the models without the state layer implementation took the same amount of time to complete as shown in table 3-7. There was not a significant difference in run times between models with and without the state layer.

Finally, these best performing models per company were trained and tested on 4730 test instances during the out-of-sample testing phase. Again, since the models were recalibrated at each testing instance as per the approach taken with the implementation of the state layer, this resulted in 167 models being trained and test 4730 times for a total of 789, 910 (167 x 4730). As a baseline comparison, the random walk method for the same testing instances was implemented, as described in Section 3.3.5.8. For each testing instance of a company a price path is recursively generated for the forecast horizon chosen (252 days or 126 days), where each data point on the price path would take the previous day's forecasted value as its input and run 16,384 Monte Carlo simulations to generate the forecast the next day's forecasted value. Thus, for each test point of each company 4,128, 768 (16,384  x 252) parallel Monte Carlo simulations were being calculated.

## 3.3 Summary of Proposed Framework and Experiment Setup

Section 3.1 provided a description of the proposed framework. Section 3.2 supplied the implementation approach taken for the experiments and framework. In Chapters 4, 5, and 6, the results of these experiments are used in investigating the hypothesis set forth in Section 1.1. Chapter 4 provides a review of the results from the input set point of view, specifically targeting RQ1 and RQ2. Chapter 5 provides a review of the results with respect to the implementation of the state layer, specifically targeting RQ3. Chapter 6 looks at the performance of the proposed framework versus benchmarks, and investigates into the contributions of the various layers, specifically targeting RQ4.

# Chapter 4. Analysis of the input feature sets

Chapter 3 put forth the research questions and described the setup of experiments conducted to investigate the research questions posed. Chapter 4 analyses the results of the experiments conducted from an input/feature set point of view. In Section 4.1, the technical, fundamental and combined indicator sets are analysed to ascertain relevance of these indicators as inputs to the forecasting models. In addition, the predictive performance of models using each input set is reviewed compared with the random walk method. In Section 4.2 a comparative review of the predictive performance of the models using these indicator sets are undertaken. Section 4.3 provides a summary of the investigation into the technical, fundamental and combined input sets and concludes the chapter.

## 4.1 Review of technical, fundamental, and combined input sets

In reviewing the features comprising the input sets, the main approach taken was to ascertain their relevance to the forecasting task at hand, and also look at the relative value added by the features. Random Forest methods are able to provide a measure of variable importance which can be used to rank individual features. Random forests are implementations of the ensemble models where the underlying assumption is that a better performing learner can be constructed through a combination of a set of learners which might be weak learners (such as decision trees) on their own. The ensemble is put together through the method known as bagging where multiple random samples (with replacement) from the training data are drawn and these random samples are used to generate multiple decision trees. The forecasts of these trees are averaged to generate the final forecast by the ensemble method. In building a random forest in addition to using bootstrapped training samples (i.e. bagging), not all the features are made available to the decision tree to be used as a splitting criterion. This is done to safeguard against the randomly built decision trees all being dominated by a few of the features that might be more important than the remainder, whereby the trees built will mostly have the same make up. Through permuting the values of each feature in the multiple trees that are built and capturing the decrease in accuracy, random forests can output a variable importance measurement. However, as described by Strobl et al. (2008), in using the mean decrease in accuracy measures, it is important to have trees that are unbiased and also that relationship between the features are also taken into account. Strobl et al. (2008) recommends using conditional random forest where the permuting of the variables is carried by taking into account the interrelationships among the features. Thus, in terms of understanding the relative value added by the features, each indicator set (technical, fundamental, combined) was run through Conditional Random Forrest (CRF) using cforest (Strobl et al., 2008; Strobl et al., 2007; Hothorn et al., 2006) to generate a ranking based on the mean decrease in accuracy that each indicator generated.

As far as measuring the relevance of the features, the Boruta algorithm, described in Kursa and Rudnicki (2010), was used to identify which indicators were relevant or not in the forecasting task. The Boruta algorithm is a wrapper feature selection method where it is using random forests as the underlying forecaster, and utilizing the variable importance statistic generated by the random forest (mean decrease in accuracy) to ascertain whether a feature is “relevant” or not. The Boruta algorithm provides an approach for determining which features are important/relevant and which ones are not by introducing random features (called "shadows") into the data set, measuring the mean decrease in accuracy by each feature in the data set (including the shadow ones), and classifying the original features as important (if they have a higher mean decrease in accuracy than the maximum mean decrease in accuracy achieved from any of the shadow variables. As a result of repeating this process over many times, the original features from the input set are classified as being relevant or irrelevant. Initial set of analysis was conducted on the simulations with a forecasting horizon of 252 days. Given that the focus of this section is on the inputs, for this part of the analysis models with the state layer were not included.

### 4.1.1 Technical indicators

Running the 10 technical indicators through the Boruta algorithm (see Section 4.1) indicated that all the inputs were relevant to the forecasting of the output. For each company, the technical indicators were run through CRF (see Section 4.1) and ranked (1 being more influential, 10 being least influential). Figure 4-1 shows the ranking per each variable averaged across all the companies.

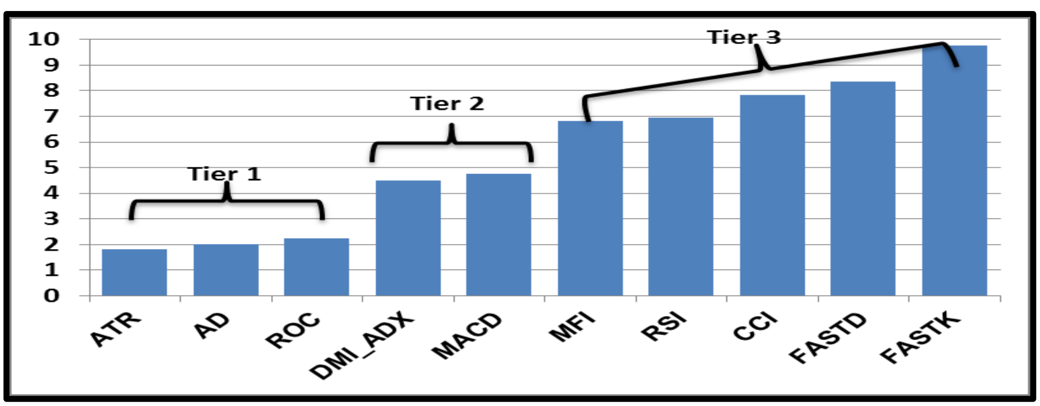


**Figure 4-1 Average ranking of technical indicators, 252 Days Out Forecasting (1 = influential, 10 least influential)**

The analyses of these rankings indicate that ATR, AD, and ROC consistently ranked as being influential (marked as Tier 1, in Figure 4-1). Furthermore, DMI_ADX and MACD consistently were in the middle of the pack (marked as Tier 2, Figure 4-1) in terms of their ability to influence the forecasting task. Finally, the remaining technical indicators (MFI, RSI, CCI, FASTD, and FAST K) consistently ranked at the bottom (Tier 3, Figure 4-1). As these rankings were across all companies in the study, a further review on the sector level was carried out. Table 4-1 summarizes the average rankings across the industry sector in which the target company is in.

| Sector | # of Companies | ATR | AD | ROC | DMI_ADX | MACD | MFI | RSI | CCI | FASTD | FASTK |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Consumer Discretionary | 20 | 1.65 | 2.15 | 2.30 | 4.30 | 4.65 | 6.95 | 6.85 | 7.85 | 8.50 | 9.80 |
| Consumer Staples | 16 | 2.06 | 2.00 | 2.00 | 4.56 | 4.81 | 6.88 | 6.75 | 8.31 | 7.94 | 9.69 |
| Energy | 13 | 1.54 | 1.85 | 2.62 | 4.46 | 4.77 | 6.23 | 7.31 | 8.08 | 8.38 | 9.77 |
| Financials | 14 | 2.14 | 1.86 | 2.07 | 4.64 | 4.43 | 6.50 | 7.07 | 7.93 | 8.50 | 9.86 |
| Health Care | 18 | 1.94 | 1.61 | 2.44 | 4.56 | 4.94 | 6.78 | 6.67 | 7.94 | 8.44 | 9.67 |
| Industrials | 31 | 1.65 | 2.26 | 2.10 | 4.55 | 4.74 | 7.13 | 6.87 | 7.58 | 8.29 | 9.84 |
| Information Technology | 13 | 1.62 | 1.85 | 2.77 | 4.46 | 4.85 | 6.54 | 7.00 | 7.62 | 8.69 | 9.62 |
| Materials | 11 | 2.00 | 2.27 | 1.73 | 4.45 | 4.91 | 7.18 | 6.73 | 7.73 | 8.27 | 9.73 |
| Utilities | 11 | 1.82 | 1.91 | 2.27 | 4.27 | 4.73 | 6.91 | 7.55 | 7.55 | 8.18 | 9.82 |
| **Overall** | **147** | **1.80** | **2.00** | **2.24** | **4.48** | **4.76** | **6.83** | **6.94** | **7.83** | **8.35** | **9.76** |
| **Overall Ranking** | | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
| | | Tier 1 | | | Tier 2 | | Tier 3 | | | | |

**Table 4-1 Average ranking of technical indicators for 252 Days Forecasting based on mean decrease in accuracy (1= influential, 10 least influential)**

Table 4-1 indicates that the tiered structure that emerged in Figure 4-1 is also exhibited at the industry sector level. With the exception of a few sectors (Materials and Health Care), in majority of the cases, the sector level ranking shown for the features remained the same as the overall ranking. Even though the indicators might swap places with their neighbours as to their rankings for a particular business sector, the indicators making up the different tiers stayed the same across all sectors. For example, for the Materials sector, ROC is ranked as more influential (1.73) than ATR (2.00), whereas ROC is ranked less influential (2.30) compared to ATR (1.65) for the Consumer Discretionary sector.

Next, the predictive performances of the models using the technical indicators were reviewed. Table 4-2 summarizes the performance of the models in terms of RMSE averages in the overall and per sector.

| Sector | # of Companies | ANN | SVR | DT | LR | Avg. RMSE | BestPerformer_RMSE | Random Walk Model | Best vs RW statistical Outperformance |
|---|---|---|---|---|---|---|---|---|---|
| Consumer Discretionary | 20 | 0.3178 | 0.2381 | 0.1905 | 0.3762 | 0.2807 | 0.1853 | 0.4032 | 100% |
| Consumer Staples | 16 | 0.1379 | 0.1157 | 0.0897 | 0.1654 | 0.1272 | 0.0887 | 0.2060 | 100% |
| Energy | 13 | 0.2686 | 0.2187 | 0.1544 | 0.3201 | 0.2405 | 0.1544 | 0.4008 | 100% |
| Financials | 14 | 0.2123 | 0.1749 | 0.1283 | 0.2804 | 0.1990 | 0.1275 | 0.2920 | 100% |
| Health Care | 18 | 0.2875 | 0.2306 | 0.1582 | 0.3667 | 0.2608 | 0.1582 | 0.3672 | 100% |
| Industrials | 31 | 0.2108 | 0.1672 | 0.1165 | 0.2536 | 0.1870 | 0.1165 | 0.2826 | 100% |
| Information Technology | 13 | 0.3180 | 0.2411 | 0.1766 | 0.3833 | 0.2798 | 0.1758 | 0.3451 | 92% |
| Materials | 11 | 0.2074 | 0.1635 | 0.1151 | 0.2532 | 0.1848 | 0.1147 | 0.3034 | 100% |
| Utilities | 11 | 0.1642 | 0.1125 | 0.0840 | 0.1935 | 0.1386 | 0.0829 | 0.2128 | 100% |
| **Overall** | **147** | **0.2378** | **0.1865** | **0.1360** | **0.2899** | **0.2126** | **0.1349** | **0.3142** | **99%** |

**Table 4-2 Average RMSE for models using technical indicators for 252 Days Forecasting**

Table 4-2 includes the average performances obtained from the machine learning methods (ANN, SVR, DT, LR) implemented, as well as the average for the best performer (lowest RMSE) of the 4 machine learning methods. Furthermore, the performance of the random walk method is included, in order to assess how the models using technical indicators performed against it. In all sectors, the best model as well as the average of the models (ANN, SVR, DT, LR) achieved lower RMSE. For each company in the study, the forecasting errors of the models using technical indicators were compared to that of the random walk method and the percentage of cases with **statistically significant (p=0.05)** better performance was noted. With the exception of the Information Technology sector, all the models significantly outperformed the random walk method. Table 4-3 shows the performance of the models using technical indicators for the Information Technology companies in the study, where the only company where the random walk model was not outperformed was Linear Technology Corporation (LLTC). However, the industry sector of the company does not appear to be a factor, as for Microchip Technology Corporation (MCHP) the best model with the technical indicators was able to statistically outperform the random walk method.

| Ticker | Company Name | Industry | ANN | SVR | DT | LR | Avg. RMSE | BestPerformer_RMSE | BestPerformer_ML | Random Walk Model |
|---|---|---|---|---|---|---|---|---|---|---|
| ADBE | Adobe Systems Incorporated | Application Software | 0.3064 | 0.2232 | 0.1367 | 0.3638 | 0.2576 | 0.1367 | DT | 0.2573 |
| HRS | Harris Corporation | Communication Equipment | 0.2035 | 0.1997 | 0.1388 | 0.2278 | 0.1924 | 0.1388 | DT | 0.2863 |
| WDC | Western Digital Corporation | Data Storage Devices | 0.3215 | 0.2466 | 0.2162 | 0.4639 | 0.3120 | 0.2162 | DT | 0.4514 |
| AAPL | Apple Inc. | Electronic Equipment | 0.4263 | 0.3231 | 0.2317 | 0.5741 | 0.3888 | 0.2317 | DT | 0.7451 |
| INTC | Intel Corporation | Semiconductor - Broad Line | 0.2026 | 0.2316 | 0.1246 | 0.2748 | 0.2084 | 0.1246 | DT | 0.2364 |
| TXN | Texas Instruments Incorporated | Semiconductor - Broad Line | 0.3330 | 0.2667 | 0.1772 | 0.4480 | 0.3062 | 0.1772 | DT | 0.2473 |
| SWKS | Skyworks Solutions, Inc. | Semiconductor - Integrated Circuits | 0.4917 | 0.3480 | 0.2539 | 0.6788 | 0.4431 | 0.2539 | DT | 0.5851 |
| XLNX | Xilinx, Inc. | Semiconductor - Integrated Circuits | 0.2445 | 0.2004 | 0.1175 | 0.3405 | 0.2257 | 0.1175 | DT | 0.2104 |
| LLTC | Linear Technology Corporation | Semiconductor - Specialized | 0.2605 | 0.1935 | 0.1833 | 0.2598 | 0.2243 | 0.1833 | DT | 0.1544 |
| MCHP | Microchip Technology Incorporat | Semiconductor - Specialized | 0.2272 | 0.1288 | 0.1397 | 0.2557 | 0.1879 | 0.1288 | SVR | 0.2189 |
| AMAT | Applied Materials, Inc. | Semiconductor Equipment & Materials | 0.2103 | 0.1653 | 0.0977 | 0.2549 | 0.1820 | 0.0977 | DT | 0.2025 |
| LRCX | Lam Research Corporation | Semiconductor Equipment & Materials | 0.3587 | 0.2266 | 0.1967 | 0.3453 | 0.2818 | 0.1967 | DT | 0.3072 |
| MU | Micron Technology, Inc. | Semiconductor- Memory Chips | 0.5485 | 0.3807 | 0.2823 | 0.4953 | 0.4267 | 0.2823 | DT | 0.5836 |

**Table 4-3 Performance of models using technical indicators for Information Technology companies**

The analysis of the performance of the models obtained by the various machine learning methods used indicates that Decision Trees outperformed the rest in 95% of the cases, where SVR was the best in the remaining 5%.

### 4.1.2 Fundamental indicators

Running the 10 fundamental indicators through the Boruta algorithm (see Section 4.1) indicated that all the inputs were relevant to the forecasting of the output. For each company, the fundamental indicators were run through CRF (see Section 4.1) and ranked (1 being more influential, 10 being least influential). Figure 4-2 shows the ranking per each variable averaged across all companies.

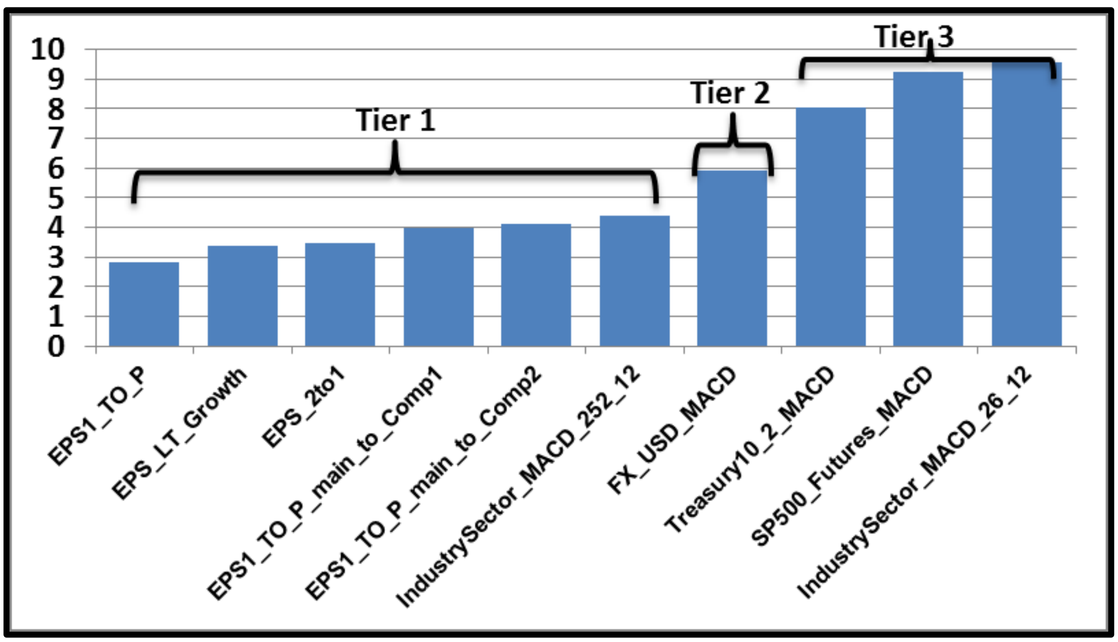


**Figure 4-2 Average ranking of fundamental indicators (1 = influential, 10 least influential)**

Analysis of these rankings indicate that company-related indicators (EPS_1_TO_P, EPS_LT_Growth, and EPS_2to1) consistently ranked as being influential (marked as part of

Tier 1, Figure 4-2). These company-related indicators were expected to be highly influential as they form one of the core components of the fundamental analysis and relate directly to the company that is being forecast. The company-related information was followed by competitor-related data (EPS1_TO_P_main_to_Comp1, EPS1_TO_P_main_to_Comp2) and long-term (252 days) portion of the industry-related indicators (IndustrySector_MACD_252_12), which are shown as part of Tier 1, Figure 4-2. The long-term portion of the industry-related indicator ranked influential whereas the short term portion of the industry-related indicator ranked least influential. Given that the forecasting horizon is 252 days, this is sensible to observe. The other features which ranked consistently as less influential are the macroeconomic indicators of SP500_Futures_MACD and the Treasury_10_2. The average rankings for the Tier 1 indicators for fundamental indicators are higher than the Tier 1 indicators for technical indicators, indicating that the features in the set exert influence of varying levels for different companies. The foreign exchange (FX_USD_MACD) related indicator from the macro-economic indicator group (marked as Tier 2, Figure 4-2), was ranked on average more influential than the rest of the microeconomic indicators (marked as Tier 3, Figure 4-2). The Tier 3 indicators were consistently ranked as being less influential across all companies in the study. As these rankings were across all companies in the study, a further review on the sector level was carried out. Table 4-4 summarizes the average rankings for the fundamental indicators across the industry sector in which the target company is in.

| Sector | # of Companies | EPS1_TO_P | EPS_LT_Growth | EPS_2to1 | EPS1_TO_P_main_to_Comp1 | EPS1_TO_P_main_to_Comp2 | Industry Sector_MACD_252_12 | FX_USD_MACD | Treasury 10_2_MACD | SP500_Futures_MACD | IndustrySector_MACD_26_12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Consumer Discretionary** | 20 | 2.55 | 3.70 | 3.65 | 4.45 | 4.05 | 4.10 | 5.80 | 8.10 | 9.30 | 9.30 |
| **Consumer Staples** | 16 | 2.63 | 2.81 | 3.56 | 4.69 | 4.13 | 4.31 | 5.94 | 8.19 | 9.13 | 9.63 |
| **Energy** | 13 | 2.62 | 2.92 | 3.15 | 4.31 | 4.46 | 5.31 | 5.54 | 7.69 | 9.15 | 9.85 |
| **Financials** | 14 | 2.07 | 4.00 | 3.86 | 4.64 | 3.86 | 3.93 | 5.86 | 8.00 | 9.43 | 9.36 |
| **Health Care** | 18 | 2.83 | 3.28 | 3.56 | 3.67 | 4.06 | 4.72 | 6.17 | 7.94 | 9.17 | 9.61 |
| **Industrials** | 31 | 3.52 | 4.29 | 3.52 | 3.13 | 3.84 | 4.10 | 5.68 | 8.03 | 9.13 | 9.77 |
| **Information Technology** | 13 | 2.92 | 2.92 | 3.38 | 3.92 | 4.31 | 4.31 | 6.23 | 8.23 | 9.62 | 9.15 |
| **Materials** | 11 | 3.36 | 2.82 | 3.00 | 4.00 | 4.55 | 4.73 | 5.82 | 8.09 | 9.00 | 9.64 |
| **Utilities** | 11 | 2.55 | 2.36 | 3.09 | 4.00 | 4.45 | 4.73 | 6.91 | 8.00 | 9.45 | 9.45 |
| **Overall Avg Ranking** | **147** | **2.85** | **3.40** | **3.46** | **3.99** | **4.12** | **4.40** | **5.94** | **8.03** | **9.24** | **9.55** |
| **Overall Ranking** | | **1.00** | **2.00** | **3.00** | **4.00** | **5.00** | **6.00** | **7.00** | **8.00** | **9.00** | **10.00** |
| | | Tier 1 | | | | | | Tier 2 | Tier 3 | | |

**Table 4-4 Average ranking of fundamental indicators based on mean decrease in accuracy (1= influential, 10 least influential)**

Table 4-4 indicates that compared to the technical indicators, the average ranking of the fundamental indicators show more variance at the business sector level. For example, for the Financials and the Industrials sectors, the EPS_LT_Growth is ranked as being less influential relative to the other sectors. Table 4-5 shows the rankings for fundamental indicators for the companies in the Financial sector. For example, for the companies in the "Accident and Health Insurance" industry, the long term industry-related indicator (IndustrySector_MACD_252_12) is more influential than the long term EPS growth expectation for the individual companies (EPS_LT_Growth). Although it is possible to see trends like this on an industry level, there are individual companies which do not follow such groupings on an industry level. For example, companies in the "Regional – Midwest Banks" industry grouping tend to place a higher influence on EPS_LT_Growth versus IndustrySector_MACD_252_12, but Fifth Third Bancorp does not fit into this general categorization.

| Company Name | Industry | EPS1_TO_P | EPS_LT_Growth | EPS_2to1 | EPS1_TO_P_main_to_Comp1 | EPS1_TO_P_main_to_Comp2 | IndustrySector_MACD_252_12 | FX_USD_MACD | Treasury10_2_MACD | SP500_Futures_MACD | IndustrySector_MACD_26_12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| AFLAC Incorporated | Accident & Health Insurance | 2 | 7 | 5 | 4 | 3 | 1 | 6 | 9 | 8 | 10 |
| Unum Group | Accident & Health Insurance | 1 | 5 | 4 | 6 | 3 | 2 | 7 | 8 | 9 | 10 |
| Weyerhaeuser Company | Real Estate Investment | 3 | 7 | 2 | 5 | 1 | 6 | 4 | 8 | 10 | 9 |
| PNC Financial Services Group | Money Center Banks | 1 | 2 | 3 | 6 | 5 | 4 | 7 | 8 | 9 | 10 |
| SunTrust Banks, Inc. | Money Center Banks | 1 | 7 | 4 | 3 | 5 | 2 | 6 | 8 | 10 | 9 |
| H&R Block, Inc. | Personal Services | 4 | 2 | 6 | 3 | 1 | 8 | 5 | 7 | 9 | 10 |
| Allstate Corporation | Property & Casualty Insurance | 3 | 2 | 1 | 6 | 4 | 5 | 7 | 8 | 9 | 10 |
| Cincinnati Financial Corporation | Property & Casualty Insurance | 2 | 6 | 1 | 4 | 7 | 5 | 3 | 8 | 10 | 9 |
| Loews Corporation | Property & Casualty Insurance | 1 | 4 | 7 | 8 | 2 | 3 | 5 | 6 | 9 | 10 |
| Progressive Corporation | Property & Casualty Insurance | 3 | 2 | 1 | 4 | 6 | 5 | 7 | 8 | 9 | 10 |
| Fifth Third Bancorp | Regional - Midwest Banks | 1 | 7 | 4 | 3 | 6 | 2 | 5 | 9 | 10 | 8 |
| Huntington Bancshares | Regional - Midwest Banks | 1 | 2 | 3 | 5 | 6 | 4 | 7 | 8 | 10 | 9 |
| KeyCorp | Regional - Midwest Banks | 2 | 1 | 7 | 5 | 4 | 3 | 6 | 9 | 10 | 8 |
| Zions Bancorporation | Regional - Pacific Banks | 4 | 2 | 6 | 3 | 1 | 5 | 7 | 8 | 10 | 9 |

### Table 4-5 Rankings for fundamental indicators for the companies in the Financial sector

Table 4-6 summarizes the performance of the models using the fundamental indicators in terms of RMSE averages both overall and per sector. For this part of the analysis, the models implementing the state layer or the feature selection method were not included.

| Sector | # of Companies | ANN | SVR | DT | LR | Avg. RMSE | BestPerformer_RMSE | Random Walk Model | Best vs RW statistical Outperformance |
|---|---|---|---|---|---|---|---|---|---|
| Consumer Discretionar | 20 | 0.2803 | 0.1767 | 0.2381 | 0.3969 | 0.2730 | 0.1552 | 0.4032 | 100% |
| Consumer Staples | 16 | 0.1187 | 0.0770 | 0.0992 | 0.1704 | 0.1163 | 0.0739 | 0.2060 | 100% |
| Energy | 13 | 0.2870 | 0.2096 | 0.2511 | 0.3475 | 0.2738 | 0.1711 | 0.4008 | 100% |
| Financials | 14 | 0.2169 | 0.1288 | 0.2187 | 0.3304 | 0.2237 | 0.1230 | 0.2920 | 100% |
| Health Care | 18 | 0.2042 | 0.1175 | 0.1437 | 0.3065 | 0.1930 | 0.1066 | 0.3672 | 100% |
| Industrials | 31 | 0.1895 | 0.1210 | 0.1275 | 0.2363 | 0.1686 | 0.1110 | 0.2826 | 100% |
| Information Technology | 13 | 0.2634 | 0.1875 | 0.2845 | 0.3530 | 0.2721 | 0.1697 | 0.3451 | 92% |
| Materials | 11 | 0.1820 | 0.1084 | 0.1013 | 0.2204 | 0.1530 | 0.0970 | 0.3034 | 100% |
| Utilities | 11 | 0.1401 | 0.0806 | 0.0699 | 0.1816 | 0.1180 | 0.0679 | 0.2128 | 100% |
| **Overall** | **147** | **0.2095** | **0.1339** | **0.1687** | **0.2834** | **0.1989** | **0.1198** | **0.3142** | **99%** |

### Table 4-6 Average RMSE for models using fundamental indicators

Table 4-6 includes the average performances from the machine learning methods (ANN, SVR, DT, LR), as well as the average for the best performer (achieved lowest RMSE) of the 4 machine learning methods. Furthermore, the performance of the random walk method is included, in order to assess how the models using the fundamental indicators performed against the random walk method. For each company in the study, the forecasting errors of the models using the fundamental indicators were compared to that of the random walk method and the percentage of cases with **statistically significant (p=0.05)** better performance was noted. Similarly to the observations from the technical indicator analysis, with the exception of the Information Technology sector, all the models significantly outperformed the random walk method. The only company where the random walk model was not outperformed was Linear Technology Corporation (LLTC).

When the fundamental indicators are used, analysing the performance of the models of the various machine learning methods used indicates that SVR was the best performing machine learning method in 52% of the cases and Decision Trees outperformed the rest in 48% of the case.

### 4.1.3 Combined indicators

The combined indicators were created by merging the 10 fundamental indicators with the 10 technical indicators. Running the 20 indicators through the Boruta algorithm (see Section 4.1) indicated that all the inputs were relevant to the forecasting of the output. Similarly to what was done with the technical and fundamental input sets, for each company, the combined set of indicators were run through a CRF (see Section 4.1) and ranked (1 being more influential, 20 being least influential). Figure 4-3 shows the ranking for each variable averaged across all the companies.

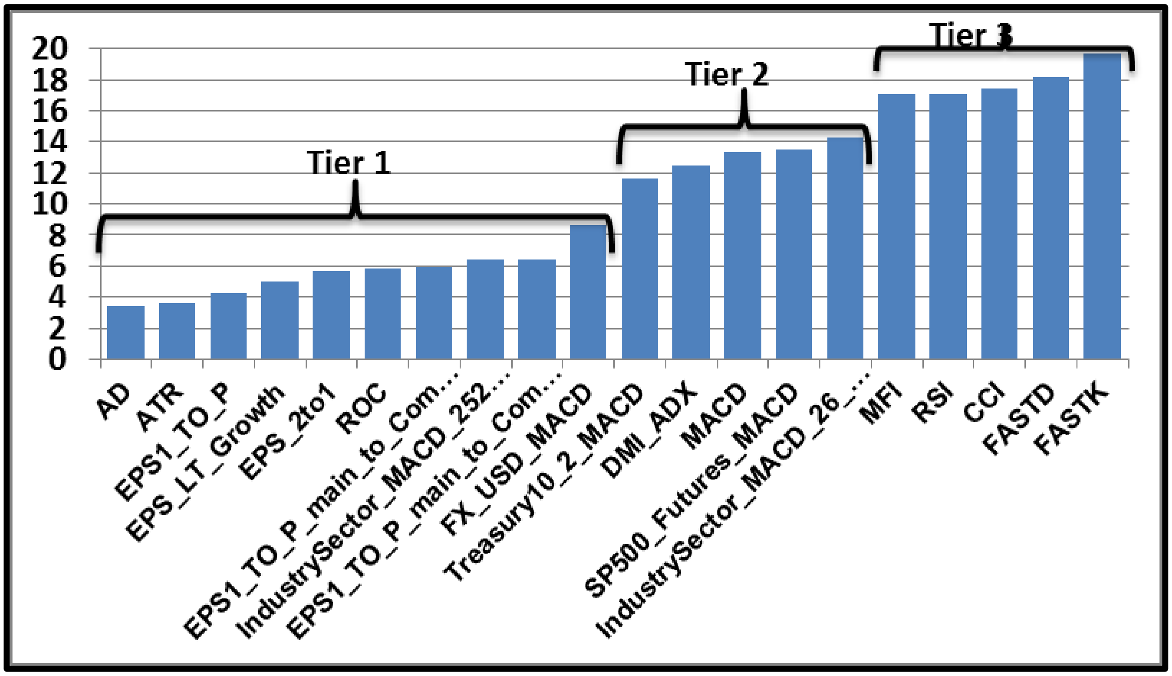


Figure 4-3 Average ranking of combined indicators (1 = influential, 20 least influential)

As shown in Figure 4-3, the relatively more influential indicators on average were a combination of the most influential indicators (Tier1) from the technical indicators and fundamental indicators. From the technical indicators, AD, ATR, and ROC were part of this group. From the Fundamentals, the company-related information, competitor related data, and long-term portion of the industry-related data were part of Tier 1. FX_USD_MACD was in between the Tier 1 and Tier2 indicators but since it was closer to Tier 1, it is shown as part of it in Figure 4-3. Thus, the more influential group of indicators from the combined Input set is more heavily comprised of the fundamental indicators, even though the two technical indicators (AD, ATR) ranked most influential. The middle tier for the combined indicators is formed of a combination of the technical indicators (Tier 2 from the technical indicators shown in Figure 4-1) and fundamental indicators (lowest-performing relative to other fundamentals, Tier 3 in Figure 4-2). The least influential set of indicators within the combined inputs, Tier 3, is composed of the technical indicators that ranked consistently less influential relative to other indicators across all companies.

Table 4-7 summarizes the performance of the models using the combined indicators in terms of RMSE averages in the overall and per sector. For this part of the analysis, models implementing the state layer or the feature selection method were not included.

| Sector | # of Companies | ANN | SVR | DT | LR | Avg. RMSE | BestPerformer_RMSE | Random Walk Model | Best vs RW statistical Outperformance |
|---|---|---|---|---|---|---|---|---|---|
| Consumer Discretionar | 20 | 0.1778 | 0.0778 | 0.1784 | 0.3318 | 0.1915 | 0.0778 | 0.4032 | 100% |
| Consumer Staples | 16 | 0.0847 | 0.0329 | 0.0647 | 0.1453 | 0.0819 | 0.0329 | 0.2060 | 100% |
| Energy | 13 | 0.1844 | 0.0763 | 0.1591 | 0.2753 | 0.1738 | 0.0763 | 0.4008 | 100% |
| Financials | 14 | 0.1604 | 0.0718 | 0.1403 | 0.2608 | 0.1583 | 0.0715 | 0.2920 | 100% |
| Health Care | 18 | 0.1619 | 0.0565 | 0.1420 | 0.2889 | 0.1623 | 0.0565 | 0.3672 | 100% |
| Industrials | 31 | 0.1388 | 0.0585 | 0.1008 | 0.2155 | 0.1284 | 0.0585 | 0.2826 | 100% |
| Information Technology | 13 | 0.1956 | 0.0771 | 0.1795 | 0.3347 | 0.1967 | 0.0771 | 0.3451 | 100% |
| Materials | 11 | 0.1356 | 0.0552 | 0.0974 | 0.2072 | 0.1238 | 0.0552 | 0.3034 | 100% |
| Utilities | 11 | 0.0963 | 0.0333 | 0.0729 | 0.1443 | 0.0867 | 0.0333 | 0.2128 | 100% |
| **Overall** | **147** | **0.1487** | **0.0604** | **0.1260** | **0.2469** | **0.1455** | **0.0604** | **0.3142** | **100%** |

**Table 4-7 Average RMSE for models using combined indicators**

The best of the models using the combined input set was able to **statistically significantly (p=0.05)** outperform the Random walk method in all cases. Also, the overall average of all the models was 0.2126 when using the technical indicators, and 0.1989 when using the fundamental indicators, which was further, reduced to 0.1455 when using the combined inputs.

Analysing the performance of the models along the lines of the various machine learning methods used indicates that SVR was the best performing machine learning method in 99% of the cases and that Decision Trees outperformed the rest in 1% of the cases. This is in stark contrast to the case observed when using the technical indicators only, which was dominated by Decision Trees. Thus, if one was using only technical indicators in conducting the experiments, Decision trees would have come out as the best machine learning method, but with the addition of the fundamental indicators into the input space, the decision with regards to the best machine learning method shifts to SVR. Therefore, if one was to use only technical indicators during the experiments, this could inadvertently introduce a bias into the conclusions (such as the best performing machine learning method) that might be drawn by the researcher. Table 4-8 shows the average of RMSE across the companies for various machine learning methods when using different inputs.

| | ANN | SVR | DT | LR |
|---|---|---|---|---|
| Technical | 0.2378 | 0.1865 | 0.1360 | 0.2899 |
| Fundamental | 0.2095 | 0.1339 | 0.1687 | 0.2834 |
| Combined | 0.1487 | 0.0604 | 0.1260 | 0.2469 |

**Table 4-8 Average RMSE per Machine Learning method and Input type**

Table 4-8 indicates that the level of improvement achieved when using the combined indicators is much more pronounced for SVR and to a lesser degree for ANN, than it is for DT or LR. Table 4-8 also indicates that the choice of the optimum machine learning method to use could be dependent on the input set as well.

### 4.1.4 Summary

Based on the findings from running the input sets through the Boruta algorithm (see Section 4.1), it was observed that all the features in the input sets were relevant to the forecasting problem. In order to assess the relative importance of the individual features in the input sets, CRF (see Section 4.1) were run with various input sets and the mean decrease in accuracy per feature was captured. In the case of technical indicators, regardless of the industry or sector of the company, features ATR, AD, and ROC were consistently ranked as relatively highly influential, which were followed by another grouping of features (DMI_ADX and MACD) which were more in the middle, and MFI, RSI, CCI, FASTD, and FAST K ranked relatively less influential compared to the rest. In the case of the fundamental indicators, the company related indicators, the competitor related indicators and the long term portion of the industry-related indicators (EPS_1_TO_P, EPS_LT_Growth, and EPS_2to1, EPS1_TO_P_main_to_Comp1, EPS1_TO_P_main_to_Comp2, IndustrySector_MACD_252_12) ranked as being influential. Although this general tendency for the features to be influential holds on a sector basis, the levels of relative influence exerted by these differ on a company basis. IndustrySector_MACD_26_12, SP500_Futures_MACD and the Treasury_10_2 ranked relatively less influential compared to the rest, regardless of industry or sector. The combined input followed along the lines of the trends that emerged from the technical and fundamental input sets. The relatively more influential indicators for the combined set came from a combination of the most influential indicators from the technical and fundamental sets but were more heavily populated by the fundamental indicators. The least influential indicators for the combined input set was the least influential indicators from the technical input set. It was observed that the relatively more influential indicators (Tier 1, Figure 4-3) in the combined input set was formed of both the technical and fundamental indicators, which indicate that there is added value from using the technical and fundamental indicators together. It was also observed that when using only technical indicators DT-based models outperformed the other machine learning methods. However, with the introduction of the fundamental data, SVR was able to improve its performance and especially in the case of models with combined indicators outperform the rest. Therefore, choosing a machine learning method which performs well based on the relevant set of inputs adds value in terms of improving forecasting accuracy. Furthermore, using only a specific subset of data (e.g. technical indicators) could result in making partial observations about the performance of the machine learning-based forecasting models.

## 4.2 Relative performance review of models using different input sets

Having investigated the features making up the technical, fundamental and combined input sets, the performance of models using these input sets were compared against each other. The tendency by machine learning researchers to use technical data rather than fundamental data has been stated in Chapter 2. One of the questions investigated through the simulations was whether models using fundamental data outperformed models using technical data. Are machine learning researchers ignoring an influential set of inputs in their research (i.e. Fundamentals) or is this not an important factor? Another hypothesis posed was whether using a combination of technical and fundamental indicators yields better forecasting performance than using either one in isolation. In an effort to investigate these queries, the forecasting performance (RMSE) of models using inputs from the technical, fundamental, and combined data sets were evaluated on the test data and compared for each company and machine learning method (ANN, SVR, DT, LR) in the study. In order to isolate the impact of using different input types, the feature selection layer and the state layer of the framework was excluded from this part of the analysis. The 147 companies in the study and the 4 machine learning methods utilized provided 588 cases to compare and analyse. In order to ensure the robustness of the results, paired t-tests using the significance level of 0.05 were used in these comparisons.

### 4.2.1 Technical versus fundamental

Table 4-9 shows the average RMSE for forecasting models using the different machine learning methods and the technical and fundamental indicators as inputs. For each machine learning method and business sector, the set of indicators yielding a lower RMSE is shown in bold. Across all the machine learning methods and companies, the Fundamentals achieve a lower RMSE (0.1989 versus 0.2126).

| | | ANN | | SVR | | DT | | LR | | Overall | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Sector** | **# of Companies** | **Technical** | **Fundamental** | **Technical** | **Fundamental** | **Technical** | **Fundamental** | **Technical** | **Fundamental** | **Technical** | **Fundamental** |
| Consumer Discretionary | 20 | 0.3178 | **0.2803** | 0.2381 | **0.1767** | **0.1905** | 0.2381 | **0.3762** | 0.3969 | 0.2807 | **0.2730** |
| Consumer Staples | 16 | 0.1379 | **0.1187** | 0.1157 | **0.0770** | **0.0897** | 0.0992 | **0.1654** | 0.1704 | 0.1272 | **0.1163** |
| Energy | 13 | **0.2686** | 0.2870 | 0.2187 | **0.2096** | **0.1544** | 0.2511 | **0.3201** | 0.3475 | **0.2405** | 0.2738 |
| Financials | 14 | **0.2123** | 0.2169 | 0.1749 | **0.1288** | **0.1283** | 0.2187 | **0.2804** | 0.3304 | **0.1990** | 0.2237 |
| Health Care | 18 | 0.2875 | **0.2042** | 0.2306 | **0.1175** | 0.1582 | **0.1437** | 0.3667 | **0.3065** | 0.2608 | **0.1930** |
| Industrials | 31 | 0.2108 | **0.1895** | 0.1672 | **0.1210** | **0.1165** | 0.1275 | 0.2536 | **0.2363** | 0.1870 | **0.1686** |
| Information Technology | 13 | 0.3180 | **0.2634** | 0.2411 | **0.1875** | **0.1766** | 0.2845 | 0.3833 | **0.3530** | 0.2798 | **0.2721** |
| Materials | 11 | 0.2074 | **0.1820** | 0.1635 | **0.1084** | 0.1151 | **0.1013** | 0.2532 | **0.2204** | 0.1848 | **0.1530** |
| Utilities | 11 | 0.1642 | **0.1401** | 0.1125 | **0.0806** | 0.0840 | **0.0699** | 0.1935 | **0.1816** | 0.1386 | **0.1180** |
| **Overall** | **147** | 0.2378 | **0.2095** | 0.1865 | **0.1339** | **0.1360** | 0.1687 | 0.2899 | **0.2834** | 0.2126 | **0.1989** |

**Table 4-9 Average RMSE performance of models with technical versus fundamental indicators**

Reviewing the results displayed in Table 4-9 shows that, with the exception of the Financials and Health care sector, models using the fundamental indicators performed better on average than the models using the technical indicators. Looking at the level of the machine learning methods show that for the SVR models, regardless of the business sector, the models with fundamental indicators outperformed their counterparts. The models using ANN and SVR as the machine learning methods performed better when using the fundamental indicators relative to when using the technical indicators as inputs. The models using DT and LR had a more mixed set of results depending on the business sector. Table 4-

10 displays the number of cases where models with technical (shown as T) or fundamental (shown as F) outperformed its counterpart and this was **statistically significant (p=0.05)**.

| Industry | # of companies | ANN | | SVR | | DT | | LR | | Total | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | F | T | F | T | F | T | F | T | F | T |
| Consumer Discretionary | 20 | 14 | 2 | 15 | 1 | 10 | 6 | 10 | 6 | 49 | 15 |
| Consumer Staples | 16 | 13 | 1 | 12 | 2 | 11 | 3 | 11 | 3 | 47 | 9 |
| Energy | 13 | 7 | 6 | 10 | 3 | 2 | 11 | 6 | 7 | 25 | 27 |
| Financials | 14 | 9 | 3 | 10 | 2 | 4 | 8 | 6 | 6 | 29 | 19 |
| Health Care | 18 | 15 | 1 | 16 | 0 | 12 | 4 | 11 | 5 | 54 | 10 |
| Industrials | 31 | 19 | 11 | 29 | 1 | 17 | 13 | 22 | 8 | 87 | 33 |
| Information Technology | 13 | 10 | 2 | 11 | 1 | 5 | 7 | 5 | 7 | 31 | 17 |
| Materials | 11 | 8 | 2 | 10 | 0 | 6 | 4 | 8 | 2 | 32 | 8 |
| Utilities | 11 | 8 | 2 | 10 | 0 | 9 | 1 | 9 | 1 | 36 | 4 |
| **Overall** | **147** | **103** | **30** | **123** | **10** | **76** | **57** | **88** | **45** | **390** | **142** |

**Table 4-10 Number of Cases of statistically significant (p=0.05) outperformance by technical (T) and fundamental (F)**

In 90% of the cases considered (532/588), there was a **statistically significant (p=0.05)** outperformance of the models using one input set over the other. In 390 (73.7%) of those cases, the models using the fundamental indicators outperformed models that were using the technical indicators. This outperformance was more pronounced in the cases where the machine learning methods used were ANN and SVR, to be able to pick up on information from the fundamental indicators more than the models using DT and LR as machine learning methods. Furthermore, with the exception of the Energy and Financials sectors, the models using fundamental indicators outperformed their counterparts on a comparatively more frequent basis. Based on the 10 fundamental and 10 technical indicators included in the study and for the forecasting horizon of 252 days, the results of the simulations indicate that in majority of the cases, using the fundamental indicators over the technical indicators in financial time series forecasting can improve the performance of the model. Thus, fundamental analysis appears to perform better when the task at hand involves 252 day forecasting out.

### 4.2.2 Combined versus technical

Table 4-11 shows the average RMSE for forecasting models using the different machine learning methods and technical and combined indicators as inputs. For each machine learning method and business sector, the set of indicators yielding a lower RMSE is shown in bold. In the overall, across all the machine learning methods and companies, the models with the combined indicators achieve a lower RMSE (0.1455 versus 0.2126) than the models using the technical indicators.

| | | ANN | | SVR | | DT | | LR | | Overall | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Sector** | **# of Companies** | **Technical** | **Combined** | **Technical** | **Combined** | **Technical** | **Combined** | **Technical** | **Combined** | **Technical** | **Combined** |
| Consumer Discretionary | 20 | 0.3178 | **0.1778** | 0.2381 | **0.0778** | 0.1905 | **0.1784** | 0.3762 | **0.3318** | 0.2807 | **0.1915** |
| Consumer Staples | 16 | 0.1379 | **0.0847** | 0.1157 | **0.0329** | 0.0897 | **0.0647** | 0.1654 | **0.1453** | 0.1272 | **0.0819** |
| Energy | 13 | 0.2686 | **0.1844** | 0.2187 | **0.0763** | **0.1544** | 0.1591 | 0.3201 | **0.2753** | 0.2405 | **0.1738** |
| Financials | 14 | 0.2123 | **0.1604** | 0.1749 | **0.0718** | **0.1283** | 0.1403 | 0.2804 | **0.2608** | 0.1990 | **0.1583** |
| Health Care | 18 | 0.2875 | **0.1619** | 0.2306 | **0.0565** | 0.1582 | **0.1420** | 0.3667 | **0.2889** | 0.2608 | **0.1623** |
| Industrials | 31 | 0.2108 | **0.1388** | 0.1672 | **0.0585** | 0.1165 | **0.1008** | 0.2536 | **0.2155** | 0.1870 | **0.1284** |
| Information Technology | 13 | 0.3180 | **0.1956** | 0.2411 | **0.0771** | **0.1766** | 0.1795 | 0.3833 | **0.3347** | 0.2798 | **0.1967** |
| Materials | 11 | 0.2074 | **0.1356** | 0.1635 | **0.0552** | 0.1151 | **0.0974** | 0.2532 | **0.2072** | 0.1848 | **0.1238** |
| Utilities | 11 | 0.1642 | **0.0963** | 0.1125 | **0.0333** | 0.0840 | **0.0729** | 0.1935 | **0.1443** | 0.1386 | **0.0867** |
| **Overall** | **147** | 0.2378 | **0.1487** | 0.1865 | **0.0604** | 0.1360 | **0.1260** | 0.2899 | **0.2469** | 0.2126 | **0.1455** |

**Table 4-11 Average RMSE performance of models with technical versus combined indicators**

Table 4-11 also shows that based on the average RMSE on a business sector basis, models with the combined indicators also outperform their counterparts. Reviewing across the various machine learning methods, only in the case of DT did the technical indicators on average achieve a slight outperformance for some companies in the Energy, Financials and Information Technology business sectors.

Furthermore, investigation as to the statistical significance of any difference in the performance of the models using the technical and combined indicators were carried out. Table 4-12 shows the number of cases where one was able to outperform the other in a **statistically significant (p=0.05)** way. In general, the models using the combined indicators outperform the models using the technical indicators in 84.52% of the cases, whereas technical only outperform in 1.53% of the cases, and the majority of those happen when the machine learning method is DT. The outperformance of the combined indicators generates its lowest relative performance in the Financial sector.

| Industry | # of compani | ANN | | SVR | | DT | | LR | | Total | | Total (%) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | T | C | T | C | T | C | T | C | T | C | T | C |
| Consumer Discretionary | 20 | 0 | 18 | 0 | 20 | 1 | 14 | 0 | 15 | 1 | 67 | 1.25% | 83.75% |
| Consumer Staples | 16 | 0 | 16 | 0 | 16 | 1 | 13 | 0 | 14 | 1 | 59 | 1.56% | 92.19% |
| Energy | 13 | 0 | 12 | 0 | 11 | 2 | 6 | 0 | 8 | 2 | 37 | 3.85% | 71.15% |
| Financials | 14 | 2 | 11 | 0 | 13 | 0 | 6 | 0 | 7 | 2 | 37 | 3.57% | 66.07% |
| Health Care | 18 | 0 | 17 | 0 | 18 | 0 | 14 | 0 | 17 | 0 | 66 | 0.00% | 91.67% |
| Industrials | 31 | 0 | 30 | 0 | 31 | 2 | 21 | 0 | 25 | 2 | 107 | 1.61% | 86.29% |
| Information Technology | 13 | 0 | 13 | 0 | 13 | 1 | 7 | 0 | 12 | 1 | 45 | 1.92% | 86.54% |
| Materials | 11 | 0 | 11 | 0 | 11 | 0 | 7 | 0 | 9 | 0 | 38 | 0.00% | 86.36% |
| Utilities | 11 | 0 | 11 | 0 | 11 | 0 | 8 | 0 | 11 | 0 | 41 | 0.00% | 93.18% |
| **Overall** | **147** | **2** | **139** | **0** | **144** | **7** | **96** | **0** | **118** | **9** | **497** | **1.53%** | **84.52%** |

**Table 4-12 (C)ombined versus (T)echnical statistically significant (p=0.05) outperformance distribution**

### 4.2.3 Combined versus fundamental

Table 4-13 shows the average RMSE for forecasting models using the different machine learning methods and the fundamental and combined indicators as inputs. For each machine learning method and business sector, the set of indicators yielding a lower RMSE is shown in bold. Overall, across all the machine learning methods and companies, the models with the combined indicators achieve a lower RMSE (0.1455 versus 0.1989) than models using the fundamental indicators.

| Sector | # of Companies | ANN Combined | ANN Fundamental | SVR Combined | SVR Fundamental | DT Combined | DT Fundamental | LR Combined | LR Fundamental | Overall Combined | Overall Fundamental |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Consumer Discretionary | 20 | **0.1778** | 0.2803 | **0.0778** | 0.1767 | **0.1784** | 0.2381 | **0.3318** | 0.3969 | **0.1915** | 0.2730 |
| Consumer Staples | 16 | **0.0847** | 0.1187 | **0.0329** | 0.0770 | **0.0647** | 0.0992 | **0.1453** | 0.1704 | **0.0819** | 0.1163 |
| Energy | 13 | **0.1844** | 0.2870 | **0.0763** | 0.2096 | **0.1591** | 0.2511 | **0.2753** | 0.3475 | **0.1738** | 0.2738 |
| Financials | 14 | **0.1604** | 0.2169 | **0.0718** | 0.1288 | **0.1403** | 0.2187 | **0.2608** | 0.3304 | **0.1583** | 0.2237 |
| Health Care | 18 | **0.1619** | 0.2042 | **0.0565** | 0.1175 | **0.1420** | 0.1437 | **0.2889** | 0.3065 | **0.1623** | 0.1930 |
| Industrials | 31 | **0.1388** | 0.1895 | **0.0585** | 0.1210 | **0.1008** | 0.1275 | **0.2155** | 0.2363 | **0.1284** | 0.1686 |
| Information Technology | 13 | **0.1956** | 0.2634 | **0.0771** | 0.1875 | **0.1795** | 0.2845 | **0.3347** | 0.3530 | **0.1967** | 0.2721 |
| Materials | 11 | **0.1356** | 0.1820 | **0.0552** | 0.1084 | **0.0974** | 0.1013 | **0.2072** | 0.2204 | **0.1238** | 0.1530 |
| Utilities | 11 | **0.0963** | 0.1401 | **0.0333** | 0.0806 | **0.0729** | 0.0699 | **0.1443** | 0.1816 | **0.0867** | 0.1180 |
| **Overall** | **147** | **0.1487** | 0.2095 | **0.0604** | 0.1339 | **0.1260** | 0.1687 | **0.2469** | 0.2834 | **0.1455** | 0.1989 |

**Table 4-13 Average RMSE performance of models with fundamental versus combined indicators**

Reviewing both on a business sector level and across the machine learning algorithms, the trend remains that the combined indicators consistently outperform their counterparts using the fundamental indicators.

Table 4-14 shows the number of cases where either the fundamental or combined indicators based models were able to outperform the other in a **statistically significant (p=0.05)** way. In general, the models using the combined indicators outperform the models using the fundamental indicators in 81.63% of the cases, whereas the models with the fundamental indicators only outperform in 1.36% of the cases. Similarly to what was observed with the technical indicators, the majority of the cases where the fundamental indicator-based models were able to outperform their counterparts using the combined indicators occur with DT as the machine learning method.

| Industry | # of compani | ANN | | SVR | | DT | | LR | | Total | | Total (%) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | F | C | F | C | F | C | F | C | F | C | F | C |
| Consumer Discretionary | 20 | 0 | 20 | 0 | 20 | 3 | 10 | 0 | 17 | 3 | 67 | 3.75% | 83.75% |
| Consumer Staples | 16 | 0 | 16 | 0 | 16 | 0 | 12 | 0 | 13 | 0 | 57 | 0.00% | 89.06% |
| Energy | 13 | 0 | 13 | 0 | 13 | 0 | 8 | 0 | 12 | 0 | 46 | 0.00% | 88.46% |
| Financials | 14 | 1 | 13 | 0 | 14 | 0 | 8 | 0 | 13 | 1 | 48 | 1.79% | 85.71% |
| Health Care | 18 | 0 | 18 | 0 | 18 | 1 | 7 | 0 | 14 | 1 | 57 | 1.39% | 79.17% |
| Industrials | 31 | 0 | 31 | 0 | 31 | 1 | 17 | 0 | 18 | 1 | 97 | 0.81% | 78.23% |
| Information Technology | 13 | 0 | 12 | 0 | 13 | 0 | 5 | 0 | 8 | 0 | 38 | 0.00% | 73.08% |
| Materials | 11 | 0 | 11 | 0 | 11 | 1 | 4 | 0 | 8 | 1 | 34 | 2.27% | 77.27% |
| Utilities | 11 | 0 | 11 | 0 | 11 | 1 | 4 | 0 | 10 | 1 | 36 | 2.27% | 81.82% |
| **Overall** | **147** | **1** | **145** | **0** | **147** | **7** | **75** | **0** | **113** | **8** | **480** | **1.36%** | **81.63%** |

Table 4-14 (C)ombined versus (F)undamental statistically significant (p=0.05) outperformance distribution

### 4.2.4 Combined versus technical & fundamental

A further hypothesis posed in RQ2 was whether using a combination of technical and fundamental indicators yields better forecasting performance than using either in isolation. To that end, the forecasting performance (RMSE) of forecasting models (ANN, SVR, DT, LR) using inputs from the technical, fundamental, and combined data sets were evaluated on the test data and compared for each company in the study. In order to isolate the impact of using different input types, the state layer of the framework was excluded from this part of the analysis. Paired t-tests (significance level of 0.05) were conducted comparing the forecasting errors of models using the combined indicators with that of models using technical indicators and also models using the combined indicators with that of the fundamental indicators. Table 4-15 summarizes the number of cases by the machine learning method where the combined inputs scenario outperformed (lower RMSE) in a **statistically significant (p=0.05)** way both of the models with the technical indicators and the fundamental indicators.

| Input | ANN | SVR | DT | LR | Total |
|---|---|---|---|---|---|
| Combined | 140 | 144 | 66 | 111 | 461 |
| % of Total | 95% | 98% | 45% | 76% | 78% |

**Table 4-15 Cases where the models using the combined indicators outperform models that using either technical or fundamental indicators**

In 78% of the 588 (4 machine learning methods x 147 companies) cases reviewed, the models using the combined indicators outperformed their counterparts, and this outperformance was more pronounced with ANN and SVR, than it was with DT and LR. Figure 4-4 displays the average RMSE per industry when using NN and SVR models with different inputs for 252 days forecasting horizon. Regardless of the industry, the Random Walk (RW) method is outperformed (lower RMSE) by models using technical indicators (T), which are in turn outperformed by models using fundamental indicators (F), and all were outperformed by models using the combined indicators (C).

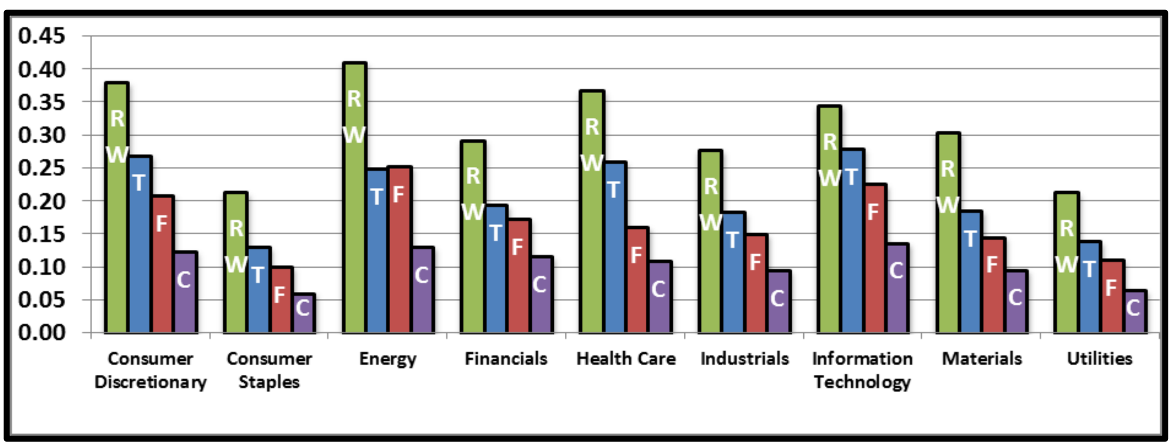


**Figure 4-4 Average RMSE of Random Walk (RW) method and models using technical (T), fundamental (F), and combined (C) inputs for 252 days forecasting**

In order to ascertain whether these observations were only valid for the 252 days forecasting horizon, further simulations were carried out for 126 days forecasting horizon. Figure 4-5 displays the average RMSE per industry when using ANN and SVR models with different inputs for

126 days forecasting horizon, where similar trends as in the 252 days forecasting can be observed. Again, regardless of the industry, the Random Walk (RW) method is outperformed (lower RMSE) by models using technical indicators (T), which are in turn outperformed by models using fundamental indicators (F), and all were outperformed by models using the combined indicators (C). Comparison of fundamental (F) and technical (T) indicator-based models show that on average fundamental analysis-based models (overall RMSE of 0.1464) outperform technical analysis-based ones (overall RMSE of 0.1693) ones regardless of the company's sector. The gap between the forecasting performances of models using technical and fundamental indicators is narrower for firms in sectors such as Financials and Energy, whilst the gap is wider for Health Care.

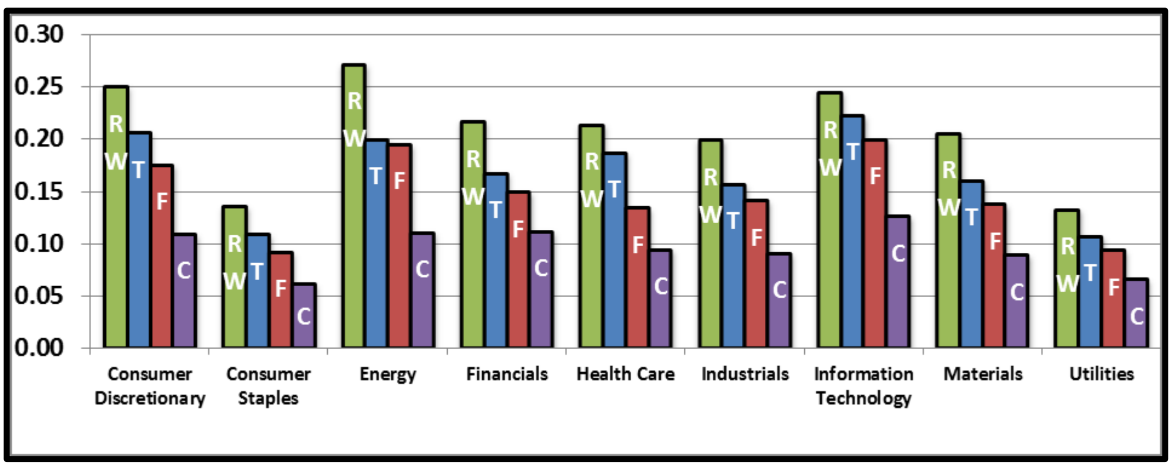


**Figure 4-5 Average RMSE of Random Walk (RW) method and models using technical (T), fundamental (F), and combined (C) inputs for 126 days forecasting**

### 4.2.5 Summary of relative performance review

In Section 4.2, the performance of the models using different inputs sets were compared. When comparing the models using the fundamental indicators to those using the technical indicators, the models with the fundamental indicators outperformed their counterparts in 66.3% of the cases (a relatively larger portion of this outperformance was with the models using ANN and SVR). On the other hand, the models using the technical indicators outperformed their counterparts in 23.8% of the cases (a relatively larger portion of this outperformance was with the models using DT and LR). Thus, with regards to the question of whether the fundamental or technical analysis is better at determining the future stock prices, from a machine learning point of view they are both relevant yet the fundamental analysis is relatively better performing than the technical analysis. Furthermore, the models using the combined indicators were able to **significantly (p=0.05)** outperform their counterparts using the technical or fundamental indicators in 84.52% and 81.63% of the cases respectively. Finally, in 78% of the cases models using the combined indicators were able to significantly (p=0.05) outperform both models using either technical or fundamental indicators in isolation. These results do suggest that synergy can be achieved by combining technical and fundamental indicators as opposed to in isolation. This further confirms and supports the views expressed by finance practitioners as covered in Chapter 2.

## 4.3 Summary of analysis of the input/feature sets

The aim of Chapter 4 was to investigate the impact of the various input sets on the predictive performance of the models. Section 4.1 investigated the features in the input sets and has shown that they are all relevant, yet some are more influential than others. Section

4.2 provided a relative comparison of the models using various input sets and has shown that fundamental indicator-based models can outperform technical indicator based models, but that using the combined indicator set outperforms both. Chapter 5 reviews the performance of the state layer.

# Chapter 5. Market State Sensitivity

Chapter 4 provided an analysis of the relevance and impact of various input sets in generating forecasts of stock prices. In addition to the inputs, another factor that can impact the forecasting performance of machine learning-based models is the challenge posed by the high volatility of the stock prices introduced by outside factors (such as overall market movements, or political shifts, etc.), as highlighted in Chapter 2. Chapter 5 reviews the results of the experiments in order to shed light on whether accounting for the state of the overall stock market could improve forecasting performance of machine learning models in predicting an individual stock's price movements in the future. In defining the approach to identify and capture the market states, one parameter to choose was the indicator (e.g. VIX vs. RSI of SP500) to represent the mood of the market and another parameter to choose was cluster size (level of granularity for the market's moods).

The rest of the chapter is organized as follows: Section 5.1 analyses which variable performed best in representing the states or the moods of the overall stock market, Section 5.2 compares the performance of the various levels of granularity at being able to capture the market moods, Section 5.3 compares the performance of models with state layer against models without the state layer, and Section 5.4 provides a summary of the chapter.

## 5.1 Market Mood Identification

As stated in Chapter 2, external factors such as the fluctuations of the overall stock market do have an impact on the stock prices of companies. To address the challenge of accounting for the states of the market in the forecasting process, the following approach has been proposed: explicitly identify the various states and “moods” of the market and then develop forecasting models for the stock price in question based on these moods. The first parameter to be decided upon for this approach was what type of market sentiment indicators can be used to represent the various moods of the market. Achelis (2000) described market indicators that can be used to track the movement of the market, and from the various indicators described, VIX, Relative Strength Indicator (RSI) of SP500, and Put-to-Call ratio were chosen as the variables to define and capture the potential moods exhibited by the stock market. Once the variables to represent the moods were selected, the potential market moods were further defined by the number of moods that the market can exhibit represented by the cluster sizes chosen (3, 5 and 7). In order to investigate which of these market mood indicators were most effective at being able to capture the states of the overall stock market, forecasting models using the various state layer definitions (3 indicators with 3 cluster sizes each) were tested on the test data for each company, machine learning method (ANN, SVR, DT, LR) and input set (technical, fundamental, combined) combination.

Table 5-1 shows the average RMSE values by each market mood indicator across various business sectors. The averages were calculated by using the performance of the model per company using a particular market mood indicator. Table 5-1 shows that overall VIX and Put-Call ratio achieved similar results (albeit VIX is slightly better with 0.0695 vs. 0.0711) and that RSI of SP500 was not performing on a par with the other two market mood indicators. Reviewing at a business sector level shows a similar trend where with the exception of the Information Technology sector, the average RMSE for models using the VIX market mood indicator are slightly better than those using Put-Call Ratio. The RMSE values displayed in bold in Table 5-1 show the best performing market mood indicator per business sector. Both Put-Call Ratio and VIX outperformed RSI of SP500 regardless of the business sector. In addition, Table 5-1 displays the percentage of cases where one market mood indicator was the best compared to the alternative definitions. VIX was the best performer in 61% of the cases, followed by Put-Call Ratio in 39% cases and RSI of SP500 in none. Certain business sectors such as Energy, Financials, Industrials and Materials, seem to “favour” VIX more heavily compared to other business sectors. One explanation as to the dismal performance of RSI of SP500 in relation to others is that VIX and Put-Call Ratio are forward-looking indicators, whereas the RSI of SP500 is a backwards-looking indicator.

| Sector | # of Companies | Average RMSE | | | % of Cases Overall Best | | |
|---|---|---|---|---|---|---|---|
| | | Best of VIX | Best of RSI of SP500 | Best of Put-Call Ratio | Best of VIX | Best of RSI of SP500 | Best of Put-Call Ratio |
| Consumer Discretionary | 20 | **0.0857** | 0.1279 | 0.0898 | 55% | 0% | 45% |
| Consumer Staples | 16 | **0.0406** | 0.0560 | 0.0400 | 56% | 0% | 44% |
| Energy | 13 | **0.0871** | 0.1255 | 0.0880 | 77% | 0% | 23% |
| Financials | 14 | **0.0777** | 0.1012 | 0.0811 | 64% | 0% | 36% |
| Health Care | 18 | **0.0704** | 0.1008 | 0.0719 | 50% | 0% | 50% |
| Industrials | 31 | **0.0642** | 0.0932 | 0.0673 | 65% | 0% | 35% |
| Information Technology | 13 | 0.0954 | 0.1367 | **0.0940** | 62% | 0% | 38% |
| Materials | 11 | **0.0621** | 0.0855 | 0.0624 | 64% | 0% | 36% |
| Utilities | 11 | **0.0407** | 0.0575 | 0.0408 | 55% | 0% | 45% |
| **Overall** | **147** | **0.0695** | **0.0990** | **0.0711** | **61%** | **0%** | **39%** |

**Table 5-1 Overall Comparison of VIX, RSI of SP500, and Put-Call Ratio**

Further analysis was conducted comparing the performance of the different market mood indicators from the input set type provided to the models. Table 5-2 shows the average RMSE of the best performing models per each market mood indicator by input type and across different business sectors (the best performing market mood indicator is shown in bold). Overall, VIX is the best market mood indicator for all three input types, followed by Put-Call Ratio and finally by RSI of SP500.

| Sector | # of Companies | Technical | | | Fundamental | | | Combined | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Best of VIX | Best of RSI of SP500 | Best of Put-Call Ratio | Best of VIX | Best of RSI of SP500 | Best of Put-Call Ratio | Best of VIX | Best of RSI of SP500 | Best of Put-Call Ratio |
| Consumer Discretionary | 20 | **0.1780** | 0.2286 | 0.1834 | 0.1491 | 0.1994 | **0.1445** | **0.0898** | 0.1284 | 0.0899 |
| Consumer Staples | 16 | 0.0906 | 0.1089 | **0.0905** | **0.0695** | 0.0985 | 0.0720 | **0.0406** | 0.0560 | 0.0411 |
| Energy | 13 | 0.1703 | 0.1916 | **0.1685** | **0.1656** | 0.2174 | 0.1737 | **0.0871** | 0.1255 | 0.0880 |
| Financials | 14 | 0.1341 | 0.1614 | **0.1334** | **0.1201** | 0.1681 | 0.1279 | **0.0777** | 0.1012 | 0.0817 |
| Health Care | 18 | **0.1586** | 0.2031 | 0.1733 | **0.1159** | 0.1602 | **0.1159** | **0.0704** | 0.1008 | 0.0719 |
| Industrials | 31 | **0.1215** | 0.1446 | 0.1222 | **0.1062** | 0.1431 | 0.1063 | **0.0642** | 0.0932 | 0.0673 |
| Information Technology | 13 | 0.1929 | 0.2350 | **0.1827** | **0.1593** | 0.2101 | 0.1645 | 0.0954 | 0.1367 | **0.0940** |
| Materials | 11 | 0.1179 | 0.1424 | **0.1123** | 0.1035 | 0.1397 | **0.0974** | **0.0621** | 0.0855 | 0.0624 |
| Utilities | 11 | **0.0792** | 0.0947 | 0.0845 | 0.0682 | 0.0915 | **0.0679** | **0.0407** | 0.0575 | 0.0408 |
| **Overall** | **147** | **0.1388** | **0.1692** | **0.1403** | **0.1175** | **0.1588** | **0.1186** | **0.0700** | **0.0991** | **0.0713** |

**Table 5-2 Comparison of VIX, RSI of SP500, and Put-Call Ratio by Input Set Type**

When using the technical indicators as inputs, for certain industries (e.g. Information Technology, Energy, Consumer Staples, Financials, and Materials) Put-Call Ratio performs slightly better than VIX. However, this trend is reversed in favour of VIX when using the fundamental indicators. Yet, the differences in the averages of RMSE are relatively small in magnitude. When using the combined indicators, with the exception of the IT sector, VIX outperforms Put-Call Ratio. Regardless of the input type, models using RSI of S&P500 in the definition of their state layer do not perform well against models using VIX and Put-Call Ratio.

A further review was undertaken to understand the performance of the models with the state layer from the perspective of the machine learning method used. Table 5-3 shows the average RMSE per each machine learning method and market mood indicator.

| ML Method | Best of VIX | Best of RSI of SP500 | Best of Put-Call Ratio |
|---|---|---|---|
| ANN | 0.1427 | 0.1887 | **0.1402** |
| SVR | **0.0709** | 0.1006 | 0.0720 |
| DT | 0.1259 | 0.1484 | **0.1220** |
| LR | 0.2287 | 0.2630 | **0.2228** |
| **Overall** | **0.1420** | **0.1752** | **0.1393** |

**Table 5-3 Average RMSE per market mood indicator by machine learning method**

The Put-Call ratio outperforms VIX for all machine learning methods except SVR. The SVR models are able to deliver a marked outperformance compared to the other machine learning methods. Given that the best-performing market mood indicator when using SVR is VIX (followed closely by Put-Call Ratio), this translates into VIX being the market mood with best performance (61% in Table 5-1).

In summary, our review of the three market mood indicators have shown RSI of SP500 to not perform as well as the others, and that in general VIX and Put-Call Ratio have a similar performance with VIX slightly outperforming Put-Call Ratio (61% versus 39%) overall. With respect to the input sets used, VIX is still the best performer for all three input sets, with Put-Call Ratio being a close second. For models using the technical or fundamental input sets, based on the business sector the best performing market mood indicator shifts between VIX and Put-Call Ratio.

## 5.2 Review of levels of granularity in Market Mood Identification

Cluster sizes of 3, 5 and 7 were chosen to capture the various moods of the stock market represented by VIX, RSI of SP500 and Put-Call Ratio. Table 5-4 gives the average RMSE performance of the models by various clusters when using VIX as the market mood indicator.

| Sector | # of Companies | Averge RMSE | | | % of Cases Best | | |
|---|---|---|---|---|---|---|---|
| | | 3 | 5 | 7 | 3 | 5 | 7 |
| Consumer Discretionary | 20 | **0.0907** | 0.0956 | 0.1056 | 85% | 15% | 0% |
| Consumer Staples | 16 | **0.0406** | 0.0446 | 0.0494 | 100% | 0% | 0% |
| Energy | 13 | **0.0879** | 0.0944 | 0.1040 | 85% | 15% | 0% |
| Financials | 14 | **0.0784** | 0.0856 | 0.0874 | 79% | 14% | 7% |
| Health Care | 18 | **0.0705** | 0.0774 | 0.0837 | 94% | 6% | 0% |
| Industrials | 31 | **0.0647** | 0.0711 | 0.0742 | 87% | 13% | 0% |
| Information Technology | 13 | **0.0956** | 0.1022 | 0.1172 | 85% | 15% | 0% |
| Materials | 11 | **0.0626** | 0.0684 | 0.0742 | 82% | 18% | 0% |
| Utilities | 11 | **0.0407** | 0.0461 | 0.0505 | 100% | 0% | 0% |
| **Overall** | **147** | **0.0705** | **0.0764** | **0.0829** | **88%** | **11%** | **1%** |

**Table 5-4 Average RMSE for VIX models by Cluster size and % of cases where each cluster size had the lowest RMSE**

Overall and across the business sectors, the performance of the VIX based models degrades as the cluster size increases (0.0705, 0.0764, and 0.0829 for 3, 5 and 7 respectively). As shown in Table 5-4, the models using clusters of 3 have outperformed their counterparties in 88% of the cases. As shown in Table 5-5, in the case of RSI of SP500, the models using clusters of 3 are even more successful (wherein 100% of the cases cluster size 3 is the best).

| Sector | # of Companies | Averge RMSE | | | % of Cases Best | | |
|---|---|---|---|---|---|---|---|
| | | 3 | 5 | 7 | 3 | 5 | 7 |
| Consumer Discretionary | 20 | **0.1279** | 0.1542 | 0.1726 | 100% | 0% | 0% |
| Consumer Staples | 16 | **0.0560** | 0.0689 | 0.0786 | 100% | 0% | 0% |
| Energy | 13 | **0.1255** | 0.1506 | 0.1711 | 100% | 0% | 0% |
| Financials | 14 | **0.1012** | 0.1262 | 0.1424 | 100% | 0% | 0% |
| Health Care | 18 | **0.1008** | 0.1253 | 0.1405 | 100% | 0% | 0% |
| Industrials | 31 | **0.0932** | 0.1139 | 0.1306 | 100% | 0% | 0% |
| Information Technology | 13 | **0.1367** | 0.1685 | 0.1921 | 100% | 0% | 0% |
| Materials | 11 | **0.0855** | 0.1053 | 0.1201 | 100% | 0% | 0% |
| Utilities | 11 | **0.0575** | 0.0711 | 0.0817 | 100% | 0% | 0% |
| **Overall** | **147** | **0.0990** | **0.1213** | **0.1376** | **100%** | **0%** | **0%** |

**Table 5-5 Average RMSE for RSI of SP500 models by cluster size**

Finally, as shown in Table 5-6, in the case of the models using the Put-Call Ratio as the market mood indicator, the cluster size of 3 still dominates (94%), but cluster sizes of 5 and 7 are selected as the best in 3% of the cases each.

| Sector | # of Companies | Averge RMSE | | | % of Cases Best | | |
|---|---|---|---|---|---|---|---|
| | | 3 | 5 | 7 | 3 | 5 | 7 |
| Consumer Discretionary | 20 | **0.0902** | 0.0988 | 0.1060 | 90% | 5% | 5% |
| Consumer Staples | 16 | **0.0412** | 0.0457 | 0.0500 | 94% | 0% | 6% |
| Energy | 13 | **0.0881** | 0.0971 | 0.1068 | 92% | 8% | 0% |
| Financials | 14 | **0.0811** | 0.0870 | 0.0963 | 100% | 0% | 0% |
| Health Care | 18 | **0.0728** | 0.0798 | 0.0852 | 89% | 6% | 6% |
| Industrials | 31 | **0.0673** | 0.0745 | 0.0810 | 97% | 3% | 0% |
| Information Technology | 13 | **0.0956** | 0.1026 | 0.1090 | 85% | 8% | 8% |
| Materials | 11 | **0.0624** | 0.0715 | 0.0779 | 100% | 0% | 0% |
| Utilities | 11 | **0.0408** | 0.0470 | 0.0496 | 100% | 0% | 0% |
| **Overall** | **147** | **0.0715** | **0.0787** | **0.0852** | **94%** | **3%** | **3%** |

**Table 5-6 Average RMSE for Put-Call models by cluster size**

Another analysis reviewed whether the input set used had any bearing on the issue of ideal cluster size. Table 5-7 shows the average RMSE of the models by input type for the various definitions of the market moods. The outperformance (best cluster size performances are highlighted in bold) of the cluster size 3 is on the average again evident regardless of the input type being used.

| Sector | VIX | | | RSI of SP500 | | | Put-Call Ratio | | |
|---|---|---|---|---|---|---|---|---|---|
| | 3 | 5 | 7 | 3 | 5 | 7 | 3 | 5 | 7 |
| Technical | **0.1656** | 0.1851 | 0.2032 | **0.1910** | 0.2308 | 0.2602 | **0.1641** | 0.1898 | 0.2036 |
| Fundamental | **0.1431** | 0.1572 | 0.1743 | **0.1791** | 0.2145 | 0.2416 | **0.1454** | 0.1571 | 0.1734 |
| Combined | **0.0907** | 0.1098 | 0.1297 | **0.1195** | 0.1550 | 0.1855 | **0.0918** | 0.1121 | 0.1321 |
| **Overall** | **0.1332** | **0.1507** | **0.1691** | **0.1632** | **0.2001** | **0.2291** | **0.1338** | **0.1530** | **0.1697** |

Table 5-7 Average RMSE of State Layer implementation by Input type

In summary, the smaller cluster size has performed better than the others. The better performance of cluster size 3 compared to the others can be attributed to the fact that as the level of granularity increases, the number of instances that are available to be part of the training set decreases, and as a result of this decrease the predictive performance of the model is negatively impacted.

## 5.3 State Layer versus no State Layer

The hypothesis put forth was that accounting for the states of the market and capturing the volatility caused in the stock’s price due to this market mood should improve performance of the machine learning-based financial time series forecasting. To evaluate the impact of introducing this state layer, for each company in the study, the best performing “definition” of the state layer implementation (market mood indicator, cluster size, input type, machine learning model combination) have been identified and compared against the models that are the best performing (input type, machine learning model combination which achieved lowest RMSE) without the state layer. With regards to the models with the state layer, 98% had SVR for the machine learning method, and 2% had DT. In addition, 98% had the combined indicators as the input set, and 2% had the fundamental indicators. Figure 5-1 shows the distribution with regards to the market mood indicators and cluster size for the state layer.

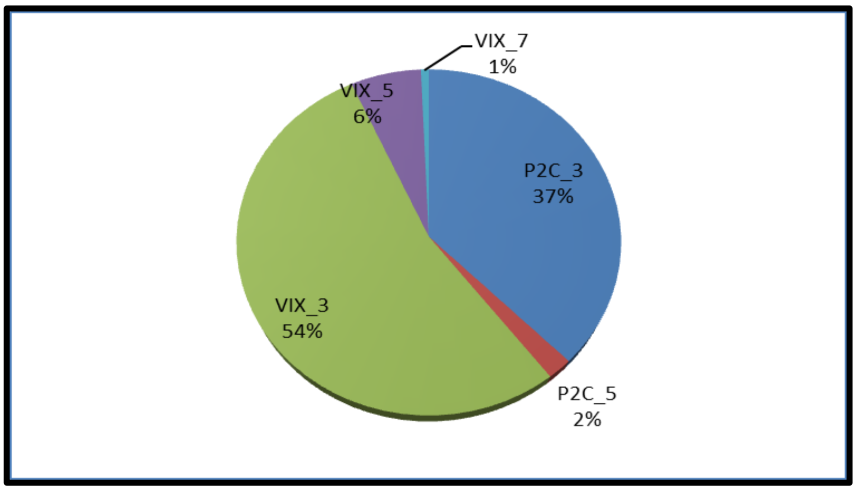


**Figure 5-1 The distribution of the cases where a particular market mood indicator achieved the lowest RMSE versus others**

Table 5-8 shows the average RMSE of the best models with and without the state layer, as well as the % of cases where each was better than their counterpart.

| Sector | # of Companies | Average RMSE: Best of Models with State Layer | Average RMSE: Best of Models without State Layer | % of Cases Overall Best: Models with State Layer | % of Cases Overall Best: Models without State Layer |
|---|---|---|---|---|---|
| Consumer Discretionary | 20 | 0.0836 | **0.0778** | 20% | 80% |
| Consumer Staples | 16 | 0.0385 | **0.0329** | 13% | 88% |
| Energy | 13 | 0.0845 | **0.0763** | 23% | 77% |
| Financials | 14 | 0.0766 | **0.0711** | 21% | 79% |
| Health Care | 18 | 0.0685 | **0.0565** | 17% | 83% |
| Industrials | 31 | 0.0636 | **0.0585** | 23% | 77% |
| Information Technology | 13 | 0.0901 | **0.0771** | 8% | 92% |
| Materials | 11 | 0.0603 | **0.0552** | 18% | 82% |
| Utilities | 11 | 0.0397 | **0.0333** | 0% | 100% |
| **Overall** | **147** | **0.0676** | **0.0604** | **17%** | **83%** |

**Table 5-8 Average RMSE of best models with and without the market state layer**

Overall, the models without the state layer tend to dominate their counterparts (83% vs. 17%, overall). In business sectors such as Utilities, the models with state layer implementation had no cases where they outperformed the models without the state layer. In business sectors such as Consumer Discretionary, Energy, Financials, and Industrials, in 1 out of every 4 or 5 cases models with the state layer outperformed their counterparts.

In order to ascertain if similar observations would be made under different forecasting horizons, further simulations were run on a randomly selected subset (85 companies) of the companies with a forecasting horizon of 126 days (i.e. 6 trading months). In the overall, when the forecasting horizon was 252 days, models that did not account for the market states explicitly outperformed those that did in 82% of the cases considered and achieved a lower RMSE on the average (0.0696 vs. 0.0616). However, when the forecasting horizon was 126 days, the models that accounted for the market states explicitly outperformed those that did in 47% of the cases considered and achieved a lower RMSE on average (0.0476 vs. 0.0550).

Therefore, the impact of the various states of the overall stock market are more pronounced when the forecasting horizon is 126 days and less so for longer forecasting horizons.

As a further analysis, Figure 5-2 displays the average RMSE per industry for a 126 days forecasting horizon. Companies in the Consumer Staples, Utilities, and materials industries are not sensitive to the states of the economy, and this appears reasonable as these are all industries where consumers cannot shrink their spending on regardless of the market state in the short term. On the other hand, companies in the Information Technology, Health Care, Financials, Energy, and Consumer Discretionary are sensitive to the states of the economy. For companies in the industries which appear not to be sensitive to the states of the economy (e.g. Consumer Staples, Utilities, and materials industries) regardless of the forecasting horizon (126 or 252), it might be sensible to investigate in the future whether substituting the state of the industry in the place of the stock market would result in any different observations.

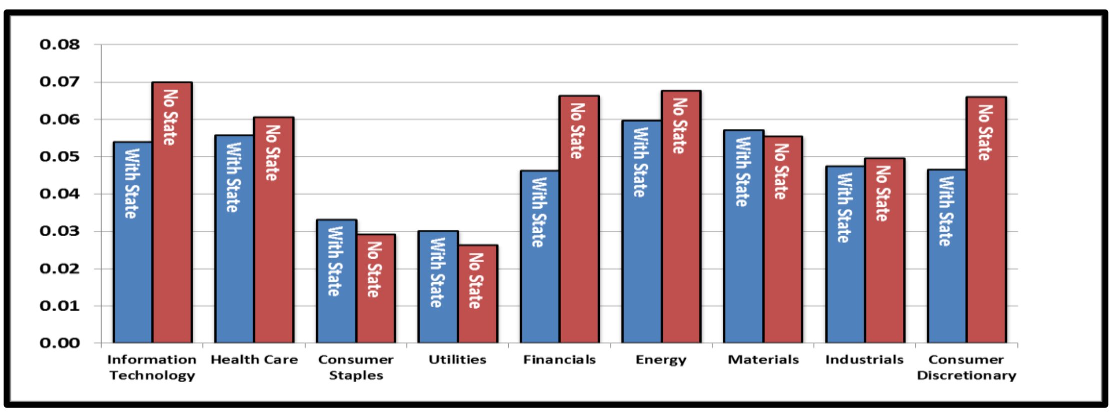


**Figure 5-2 Average RMSE by industry of models with state layer (With State) versus without the state layer (No State) when the forecasting horizon is 126 days**

Further review by using the input sets was conducted as shown in Table 5-9. Table 5-9 shows that when using fundamental or technical indicators alone, accounting for the states of the stock market does improve the forecasting model's performance in 66% and 49.7% of the cases. Thus the models using solely the technical or fundamental indicators could benefit (i.e. achieve lower RMSE) from explicitly accounting for the states of the market in the forecasting process.

| | Average RMSE | | % of Cases Overall Best | |
|---|---|---|---|---|
| **Sector** | **Best of Models with State Layer** | **Best of Models without State Layer** | **Models with State Layer** | **Models without State Layer** |
| Technical | 0.1333 | 0.1349 | 49.7% | 50.3% |
| Fundamental | 0.1115 | 0.1198 | 66.0% | 34.0% |
| Combined | 0.0680 | 0.0604 | 15.6% | 84.4% |
| **Overall** | **0.1043** | **0.1051** | **0.4376** | **0.5624** |

**Table 5-9 Models with state layer versus without the state layer comparison by input set (252 days)**

Table 5-9 only displays results for 252 days forecasting yet identical observations were made in the case of 126 days forecasting. This provides additional insight to our earlier analysis that accounting for the states of the market becomes less impactful with an increasing forecasting horizon, in that the model inputs used are also a factor. When generating a forecast for 252 days, the synergy achieved by using a combination of technical and fundamental indicators does surpass the benefits that are achieved by explicitly accounting for the states of the market.

Another analysis was conducted to investigate potential differences derived from the machine learning method used. Thus, the best performing models with and without the state layer implementation were compared to each other for each machine learning method. Table

5-10 displays the average RMSE as well as the percentage of cases where models with and without the state layer implementation outperformed the other.

| | Average RMSE | | % of Cases Overall Best | |
|---|---|---|---|---|
| Sector | Best of Models with State Layer | Best of Models without State Layer | Models with State Layer | Models without State Layer |
| ANN | **0.0720** | 0.1483 | 100.0% | 0.0% |
| SVR | 0.0695 | **0.0604** | 15.0% | 85.0% |
| DT | **0.0706** | 0.1075 | 99.3% | 0.7% |
| LR | **0.0733** | 0.2382 | 100.0% | 0.0% |
| **Overall** | **0.0713** | **0.1386** | **79%** | **21%** |

**Table 5-10 By ML method performance comparison of the models with state layer versus without the state layer**

Table 5-10 shows that, with the exception of the SVR machine learning method, the performance of the models using other machine learning methods improved as a result of the implementation of the state layer.

## 5.4 Summary

Chapter 5 provided a review of the market state layer of the framework which is used to account for the impact of the market moods on the stock price forecasting. Section 5.1 looked at which variables were effective in being able to represent the moods of the market and found that among the variables chosen, VIX and Put-call ratio were effective, while RSI of SP500 proved to not be relevant. Section 5.2 considered the level of granularity that is needed to capture the moods of the stock market and, based on the performance of the models, cluster size 3 performed better than the larger cluster sizes (e.g. 7). Section 5.3 compared the models implementing the market state layer to models without the market state layer and found the market state layer to be ineffective when using the combined indicators and SVR, but more effective when using the technical or fundamental indicators with other machine learning methods (e.g. ANN, DT). Also, the market state layer was more effective when the forecasting horizon is 126 versus 252 days. Companies belonging to certain industries (Consumer Staples, Utilities, and materials industries) were observed to be less sensitive to the market states regardless of the horizon. In Chapter 6, the performance of the proposed framework is compared against the random walk method and other base case (ANN model using the technical indicators); further, it reviews the contribution of the various layers of the framework.

# Chapter 6. Review of Framework performance

The proposed framework provides a structured approach to identifying the relevant inputs and accounting for the market moods in generating a forecast for the price of a stock. Chapter 6 provides a review of the performance of the proposed framework, followed by an analysis of the impact of the various layers on the framework's performance. The rest of the chapter is organized as follows: Section 6.1 provides an overall performance review for the framework versus base case, Section 6.2 investigates the impact of the various layers of the framework, Section 6.3 highlights potential benefits that can be achieved by using the framework in financial forecasting, and Section 6.4 summarizes the chapter.

## 6.1 Framework Performance vs. base cases

In Table 6-1 for each company in the study, the make-up of the model picked by the framework, its performance ("Framework RMSE") and the performance of the random walk method are shown. The make-up of the model picked by the framework represents the model that the framework considers to have the highest chance of making a successful forecast of the stock price of the specific company in question over the forecasting horizon (all the results displayed is for 252 days forecasting horizon). The make-up of the model is defined by the input set (T=technical, F=fundamental, C=combined), Feature Selection method applied (NO_FS = No Feature Selection, PCA_FS= PCA), State Layer implemented (NO_STATE= No state layer, otherwise: market mood indicator_Number of clusters (e.g.

VIX_3)), and Machine Learning method utilized (e.g. ANN, SVR, DT, LR).

| Ticker | Company Name | Input | FS | State | ML | Framework RMSE | Random Walk RMSE |
|---|---|---|---|---|---|---|---|
| AAPL | Apple Inc. | C | NO_FS | NO_STATE | SVR | 0.10549364 | 0.74512788 |
| ABT | Abbott Laboratories Common Stoc | C | NO_FS | NO_STATE | SVR | 0.02435051 | 0.19793112 |
| ADBE | Adobe Systems Incorporated | C | NO_FS | NO_STATE | SVR | 0.05118473 | 0.25732677 |
| ADM | Archer-Daniels-Midland Company | C | NO_FS | NO_STATE | SVR | 0.05035005 | 0.27956633 |
| AEP | American Electric Power Company | C | NO_FS | NO_STATE | SVR | 0.02603053 | 0.18727413 |
| AES | The AES Corporation Common Stoc | C | NO_FS | NO_STATE | SVR | 0.05911327 | 0.25161240 |
| AFL | AFLAC Incorporated Common Stock | C | NO_FS | NO_STATE | DT | 0.07867825 | 0.29648499 |
| ALL | Allstate Corporation (The) Comm | C | NO_FS | NO_STATE | SVR | 0.04240880 | 0.27667219 |
| AMAT | Applied Materials, Inc. | C | NO_FS | NO_STATE | SVR | 0.05419715 | 0.20248990 |
| APC | Anadarko Petroleum Corporation | C | NO_FS | NO_STATE | SVR | 0.06321735 | 0.31488622 |
| APD | Air Products and Chemicals, Inc | C | NO_FS | NO_STATE | SVR | 0.04504671 | 0.20727499 |
| BA | Boeing Company (The) Common Sto | C | NO_FS | NO_STATE | SVR | 0.04932477 | 0.27898208 |
| BAX | Baxter International Inc. Commo | C | NO_FS | NO_STATE | SVR | 0.02979397 | 0.21284918 |
| BIIB | Biogen Inc. | C | NO_FS | NO_STATE | SVR | 0.06132141 | 0.53622280 |
| BMY | Bristol-Myers Squibb Company Co | C | NO_FS | NO_STATE | SVR | 0.03264484 | 0.23099540 |
| BSX | Boston Scientific Corporation C | C | NO_FS | NO_STATE | SVR | 0.05528014 | 0.29959324 |
| CAT | Caterpillar, Inc. Common Stock | C | NO_FS | NO_STATE | SVR | 0.08806195 | 0.37927431 |
| CELG | Celgene Corporation | C | NO_FS | NO_STATE | SVR | 0.07804985 | 0.48612890 |
| CHD | Church & Dwight Company, Inc. C | C | NO_FS | NO_STATE | SVR | 0.02829043 | 0.25995595 |
| CHK | Chesapeake Energy Corporation C | C | NO_FS | NO_STATE | SVR | 0.08876782 | 0.40760766 |
| CINF | Cincinnati Financial Corporatio | C | NO_FS | NO_STATE | SVR | 0.04355244 | 0.23738912 |
| CL | Colgate-Palmolive Company Commo | C | NO_FS | NO_STATE | SVR | 0.02408827 | 0.17627511 |
| CMS | CMS Energy Corporation Common S | C | NO_FS | NO_STATE | SVR | 0.03953039 | 0.27650839 |
| COG | Cabot Oil & Gas Corporation Com | C | NO_FS | NO_STATE | SVR | 0.11653847 | 0.65890906 |
| COST | Costco Wholesale Corporation | C | NO_FS | NO_STATE | SVR | 0.02788882 | 0.24609869 |
| CPB | Campbell Soup Company Common St | C | NO_FS | NO_STATE | SVR | 0.02796742 | 0.13767353 |
| CSX | CSX Corporation | C | NO_FS | NO_STATE | SVR | 0.05136183 | 0.32979079 |
| CTAS | Cintas Corporation | C | NO_FS | NO_STATE | SVR | 0.03843585 | 0.24884571 |
| CVS | CVS Health Corporation Common S | C | NO_FS | NO_STATE | SVR | 0.04236352 | 0.25252772 |
| D | Dominion Resources, Inc. Common | C | NO_FS | NO_STATE | SVR | 0.02364292 | 0.18485535 |
| DD | E.I. du Pont de Nemours and Com | C | NO_FS | NO_STATE | SVR | 0.08604734 | 0.26057020 |
| DE | Deere & Company Common Stock | C | NO_FS | NO_STATE | SVR | 0.05654389 | 0.32496339 |
| DHI | D.R. Horton, Inc. Common Stock | C | NO_FS | NO_STATE | SVR | 0.08851957 | 0.43384447 |
| DHR | Danaher Corporation Common Stoc | C | NO_FS | NO_STATE | SVR | 0.03818075 | 0.20455229 |
| DOV | Dover Corporation Common Stock | C | NO_FS | NO_STATE | SVR | 0.05653336 | 0.27255998 |
| DOW | Dow Chemical Company (The) Comm | C | NO_FS | NO_STATE | SVR | 0.07817437 | 0.37234413 |
| DUK | Duke Energy Corporation (Holdin | C | NO_FS | NO_STATE | SVR | 0.03001128 | 0.19932775 |
| DVN | Devon Energy Corporation Common | C | NO_FS | NO_STATE | SVR | 0.06630390 | 0.28405833 |
| EIX | Edison International Common Sto | C | NO_FS | NO_STATE | SVR | 0.03430624 | 0.25402295 |
| EMN | Eastman Chemical Company Common | C | NO_FS | NO_STATE | SVR | 0.06464146 | 0.41654188 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| EMR | Emerson Electric Company Common | C | NO_FS | NO_STATE | SVR | 0.04987530 | 0.20009444 |
| EOG | EOG Resources, Inc. Common Stoc | C | NO_FS | NO_STATE | SVR | 0.09059972 | 0.39975913 |
| EQT | EQT Corporation Common Stock | C | NO_FS | NO_STATE | SVR | 0.04878073 | 0.34019680 |
| ESRX | Express Scripts Holding Company | C | NO_FS | NO_STATE | SVR | 0.06043036 | 0.41875718 |
| FLIR | FLIR Systems, Inc. | C | NO_FS | NO_STATE | SVR | 0.08208861 | 0.67997816 |
| FMC | FMC Corporation Common Stock | C | NO_FS | NO_STATE | SVR | 0.04614153 | 0.35904635 |
| GD | General Dynamics Corporation Co | C | NO_FS | NO_STATE | SVR | 0.04498400 | 0.33069753 |
| GE | General Electric Company Common | C | NO_FS | NO_STATE | SVR | 0.05919845 | 0.21194583 |
| GILD | Gilead Sciences, Inc. | C | NO_FS | NO_STATE | SVR | 0.06861406 | 0.21157830 |
| GIS | General Mills, Inc. Common Stoc | C | NO_FS | NO_STATE | SVR | 0.02033832 | 0.42400987 |
| GPS | Gap, Inc. (The) Common Stock | C | NO_FS | NO_STATE | SVR | 0.07193856 | 0.37204759 |
| GT | The Goodyear Tire & Rubber Comp | C | NO_FS | NO_STATE | SVR | 0.09659488 | 0.40297171 |
| GWW | W.W. Grainger, Inc. Common Stoc | C | NO_FS | NO_STATE | SVR | 0.04289988 | 0.54320853 |
| HAL | Halliburton Company Common Stoc | C | NO_FS | NO_STATE | SVR | 0.08178048 | 0.32344501 |
| HAS | Hasbro, Inc. | C | NO_FS | NO_STATE | SVR | 0.05618086 | 0.45067396 |
| HD | Home Depot, Inc. (The) Common S | C | NO_FS | NO_STATE | SVR | 0.06189368 | 0.49513809 |
| HOLX | Hologic, Inc. | C | NO_FS | NO_STATE | SVR | 0.08395443 | 0.27447480 |
| HON | Honeywell International Inc. Co | C | NO_FS | NO_STATE | SVR | 0.05974918 | 0.35854658 |
| HRB | H&R Block, Inc. Common Stock | C | NO_FS | NO_STATE | SVR | 0.05394880 | 0.35349303 |
| HRL | Hormel Foods Corporation | C | NO_FS | NO_STATE | SVR | 0.03391079 | 0.60181540 |
| HRS | Harris Corporation Common Stock | C | NO_FS | NO_STATE | SVR | 0.04355938 | 0.25318584 |
| IFF | Internationa Flavors & Fragranc | C | NO_FS | NO_STATE | SVR | 0.03894113 | 0.23714597 |
| INTC | Intel Corporation | C | NO_FS | NO_STATE | SVR | 0.04189523 | 0.23137873 |
| ITW | Illinois Tool Works Inc. Common | C | NO_FS | NO_STATE | SVR | 0.05166607 | 0.28633913 |
| JBHT | J.B. Hunt Transport Services, I | C | NO_FS | NO_STATE | SVR | 0.04315457 | 0.26383611 |
| JCI | Johnson Controls International | C | NO_FS | NO_STATE | SVR | 0.11301866 | 0.23642066 |
| JNJ | Johnson & Johnson Common Stock | C | NO_FS | NO_STATE | SVR | 0.01916471 | 0.34159712 |
| JWN | Nordstrom, Inc. Common Stock | C | NO_FS | NO_STATE | SVR | 0.08986920 | 0.21284951 |
| KEY | KeyCorp Common Stock | C | NO_FS | NO_STATE | SVR | 0.07719226 | 0.26833004 |
| KMB | Kimberly-Clark Corporation Comm | C | NO_FS | NO_STATE | SVR | 0.02528319 | 0.26465936 |
| KR | Kroger Company | C | NO_FS | NO_STATE | SVR | 0.03645721 | 0.13770759 |
| KSS | Kohl's Corporation Common Stock | C | NO_FS | NO_STATE | SVR | 0.05003608 | 0.40685840 |
| KSU | Kansas City Southern Common Sto | C | NO_FS | NO_STATE | SVR | 0.07382989 | 0.23862319 |
| L | Loews Corporation Common Stock | C | NO_FS | NO_STATE | SVR | 0.05350719 | 0.18473010 |
| LB | L Brands, Inc. | C | NO_FS | NO_STATE | SVR | 0.11238414 | 0.23882414 |
| LEN | Lennar Corporation Class A Comm | C | NO_FS | NO_STATE | SVR | 0.12804504 | 0.18788546 |
| LLTC | Linear Technology Corporation | C | NO_FS | NO_STATE | SVR | 0.04525669 | 0.43775316 |
| LLY | Eli Lilly and Company Common St | C | NO_FS | NO_STATE | SVR | 0.02808530 | 0.23253887 |
| LMT | Lockheed Martin Corporation Com | C | NO_FS | NO_STATE | SVR | 0.04316689 | 0.54073240 |
| LOW | Lowe's Companies, Inc. Common S | C | NO_FS | NO_STATE | SVR | 0.04448875 | 0.55891011 |
| LRCX | Lam Research Corporation | C | NO_FS | NO_STATE | SVR | 0.07860010 | 0.15441128 |
| MAS | Masco Corporation Common Stock | C | NO_FS | NO_STATE | SVR | 0.09039215 | 0.19932716 |
| MCHP | Microchip Technology Incorporat | C | NO_FS | NO_STATE | SVR | 0.04760829 | 0.24637107 |
| MKC | McCormick & Company, | C | NO_FS | NO_STATE | SVR | 0.02934290 | 0.25517601 |
| MRO | Marathon Oil Corporation Common | C | NO_FS | NO_STATE | SVR | 0.06968130 | 0.30718633 |
| MU | Micron Technology, Inc. | C | NO_FS | NO_STATE | SVR | 0.14094628 | 0.38302032 |
| MUR | Murphy Oil Corporation | C | NO_FS | NO_STATE | SVR | 0.04922495 | 0.21892651 |
| MYL | Mylan N.V. | C | NO_FS | NO_STATE | SVR | 0.04618584 | 0.21334538 |
| NBL | Noble Energy Inc. Common Stock | C | NO_FS | NO_STATE | SVR | 0.07613724 | 0.18751365 |
| NFX | Newfield Exploration Company Co | C | NO_FS | NO_STATE | SVR | 0.07352821 | 0.25815895 |
| NI | NiSource Inc Common Stock | C | NO_FS | NO_STATE | SVR | 0.04959846 | 0.38580132 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| NOC | Northrop Grumman Corporation Co | C | NO_FS | NO_STATE | SVR | 0.04351845 | 0.58355437 |
| NSC | Norfolk Southern Corporation Co | C | NO_FS | NO_STATE | SVR | 0.05452413 | 0.25021050 |
| PCG | Pacific Gas & Electric Co. Comm | C | NO_FS | NO_STATE | SVR | 0.02640022 | 0.28473527 |
| PDCO | Patterson Companies, Inc. | C | NO_FS | NO_STATE | SVR | 0.03898562 | 0.28710916 |
| PEG | Public Service Enterprise Group | C | NO_FS | NO_STATE | SVR | 0.03087401 | 0.42634769 |
| PFE | Pfizer, Inc. Common Stock | C | NO_FS | NO_STATE | SVR | 0.02888557 | 0.27915343 |
| PG | Procter & Gamble Company (The) | C | NO_FS | NO_STATE | SVR | 0.02171181 | 0.24420506 |
| PGR | Progressive Corporation (The) C | C | NO_FS | NO_STATE | SVR | 0.03270989 | 0.28194140 |
| PH | Parker-Hannifin Corporation Com | C | NO_FS | NO_STATE | SVR | 0.07101242 | 0.14958053 |
| PHM | PulteGroup, Inc. Common Stock | C | NO_FS | NO_STATE | SVR | 0.14745275 | 0.19261001 |
| PNR | Pentair plc. Ordinary Share | C | NO_FS | NO_STATE | SVR | 0.05051897 | 0.18918221 |
| PPG | PPG Industries, Inc. Common Sto | C | NO_FS | NO_STATE | SVR | 0.04282962 | 0.22789915 |
| PX | Praxair, Inc. Common Stock | C | NO_FS | NO_STATE | SVR | 0.03151686 | 0.10610067 |
| R | Ryder System, Inc. Common Stock | C | NO_FS | NO_STATE | SVR | 0.05667932 | 0.17171997 |
| REGN | Regeneron Pharmaceuticals, Inc. | C | NO_FS | NO_STATE | SVR | 0.16463100 | 0.27361889 |
| RHI | Robert Half International Inc. | C | NO_FS | NO_STATE | SVR | 0.04951237 | 0.84263784 |
| ROST | Ross Stores, Inc. | C | NO_FS | NO_STATE | SVR | 0.04139102 | 0.20856332 |
| RTN | Raytheon Company Common Stock | C | NO_FS | NO_STATE | SVR | 0.04015791 | 0.27159470 |
| SEE | Sealed Air Corporation Common S | C | NO_FS | NO_STATE | SVR | 0.07227067 | 0.32009453 |
| SHW | Sherwin-Williams Company (The) | C | NO_FS | NO_STATE | SVR | 0.04757886 | 0.19740063 |
| SLB | Schlumberger N.V. Common Stock | C | NO_FS | NO_STATE | SVR | 0.06310134 | 0.26860410 |
| SNA | Snap-On Incorporated Common Sto | C | NO_FS | NO_STATE | SVR | 0.04730214 | 0.50685374 |
| SO | Southern Company (The) Common S | C | NO_FS | NO_STATE | SVR | 0.02013905 | 1.10572086 |
| SPLS | Staples, Inc. | C | NO_FS | NO_STATE | SVR | 0.04098056 | 0.24500234 |
| SWK | Stanley Black & Decker, Inc. Co | C | NO_FS | NO_STATE | SVR | 0.04635371 | 0.37556067 |
| SWKS | Skyworks Solutions, Inc. | C | NO_FS | NO_STATE | SVR | 0.12515828 | 0.42163108 |
| SWN | Southwestern Energy Company Com | C | NO_FS | NO_STATE | SVR | 0.10398431 | 0.23808398 |
| SYY | Sysco Corporation Common Stock | C | NO_FS | NO_STATE | SVR | 0.02488130 | 0.40908761 |
| TJX | TJX Companies, Inc. (The) Commo | C | NO_FS | NO_STATE | SVR | 0.03994756 | 0.33049754 |
| TSN | Tyson Foods, Inc. Common Stock | C | NO_FS | NO_STATE | SVR | 0.05722380 | 0.30494997 |
| TXN | Texas Instruments Incorporated | C | NO_FS | NO_STATE | SVR | 0.08119236 | 0.16694800 |
| UNM | Unum Group Common Stock | C | NO_FS | NO_STATE | SVR | 0.04626373 | 0.21501873 |
| UNP | Union Pacific Corporation Commo | C | NO_FS | NO_STATE | SVR | 0.04680932 | 0.30992883 |
| URBN | Urban Outfitters, Inc. | C | NO_FS | NO_STATE | SVR | 0.06284046 | 0.25100582 |
| UTX | United Technologies Corporation | C | NO_FS | NO_STATE | SVR | 0.06188340 | 0.58513871 |
| VMC | Vulcan Materials Company (Holdi | C | NO_FS | NO_STATE | SVR | 0.05350575 | 0.67385020 |
| VRTX | Vertex Pharmaceuticals Incorpor | C | NO_FS | NO_STATE | SVR | 0.10966032 | 0.22231880 |
| WBA | Walgreens Boots Alliance, Inc. | C | NO_FS | NO_STATE | SVR | 0.05002597 | 0.12706327 |
| WDC | Western Digital Corporation | C | NO_FS | NO_STATE | SVR | 0.13064617 | 0.36767489 |
| WEC | WEC Energy Group, Inc. Common S | C | NO_FS | NO_STATE | SVR | 0.02623747 | 0.27759074 |
| WMT | Wal-Mart Stores, Inc. Common St | C | NO_FS | NO_STATE | SVR | 0.02629147 | 0.39456905 |
| WY | Weyerhaeuser Company Common Sto | C | NO_FS | NO_STATE | SVR | 0.05788378 | 0.23641545 |
| XLNX | Xilinx, Inc. | C | NO_FS | NO_STATE | SVR | 0.05606655 | 0.30387319 |
| XRAY | DENTSPLY SIRONA Inc. | C | NO_FS | NO_STATE | SVR | 0.03815877 | 0.43621179 |
| ZION | Zions Bancorporation | C | NO_FS | NO_STATE | SVR | 0.06856307 | 0.18636596 |
| FITB | Fifth Third Bancorp | T | NO_FS | VIX_7 | SVR | 0.20350206 | 0.28341339 |
| HAR | Harman International Industries | C | NO_FS | VIX_3 | SVR | 0.12329376 | 0.65014137 |
| HBAN | Huntington Bancshares Incorpora | T | NO_FS | VIX_3 | SVR | 0.20377585 | 0.25518736 |
| IR | Ingersoll-Rand plc (Ireland) | C | NO_FS | VIX_3 | SVR | 0.09128222 | 0.45139584 |
| MMM | 3M Company Common Stock | C | NO_FS | VIX_3 | SVR | 0.04269826 | 0.20255160 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| MRK | Merck & Company, Inc. Common St | C | NO_FS | VIX_3 | SVR | 0.05110468 | 0.54360647 |
| PNC | PNC Financial Services Group, I | C | PCA_FS | VIX_3 | SVR | 0.05490985 | 0.17901356 |
| RCL | Royal Caribbean Cruises Ltd. Co | C | NO_FS | VIX_3 | SVR | 0.11084075 | 0.31069549 |
| ROK | Rockwell Automation, Inc. Commo | C | NO_FS | VIX_3 | SVR | 0.09880902 | 0.21043556 |
| STI | SunTrust Banks, Inc. Common Sto | C | NO_FS | VIX_3 | SVR | 0.10627592 | 0.14475274 |
| TXT | Textron Inc. Common Stock | C | NO_FS | VIX_3 | SVR | 0.09529620 | 0.29315219 |

**Table 6-1 Model make-up per each company as picked by framework, and its performance (RMSE) versus Random Walk Model**

The framework **(statistical significance, p-value of 0.05)** outperformed the random walk model for each company included in the study. In 11 of the 150 cases, the framework picked models with the state layer, where VIX with a size 3 cluster was the predominant definition of the state layer for these models. With the exception of one case (DT for Aflac Incorporated), SVR was the machine learning method picked consistently by the framework. combined indicators dominated the input layer wherein all cases, with the exception of two (technical indicators for Fifth Third Bancorp and Huntington Bancshares), the models picked by the framework used the combined indicators. Table 6-2 displays the average RMSE by business sector for the models picked by the framework, and the random walk model.

| | | Average RMSE | | |
|---|---|---|---|---|
| Sector | # of Companies | Framework | Random Walk | ANN with Technical Indicators |
| Consumer Discretionary | 20 | 0.0787 | 0.3534 | 0.3178 |
| Consumer Staples | 16 | 0.0329 | 0.2822 | 0.1379 |
| Energy | 13 | 0.0763 | 0.3236 | 0.2686 |
| Financials | 14 | 0.0802 | 0.2428 | 0.2123 |
| Health Care | 18 | 0.0566 | 0.3166 | 0.2875 |
| Industrials | 31 | 0.0588 | 0.3356 | 0.2108 |
| Information Technology | 13 | 0.0771 | 0.3209 | 0.3180 |
| Materials | 11 | 0.0552 | 0.3071 | 0.2074 |
| Utilities | 11 | 0.0333 | 0.3485 | 0.1642 |
| **Overall** | **147** | **0.0614** | **0.3175** | **0.2378** |

**Table 6-2 Average RMSE by Business Sector**

As stated in Chapter 2, the overall tendency in machine learning research is towards using the technical indicators and ANNs are among the most widely implemented machine learning methods for financial time series forecasting. Thus, included in Table 6-2 for comparison against the framework is the performance of the models that use only technical indicators and ANN as the machine learning method. Overall the framework was able to deliver RMSE improvement **(statistically significant p=0.05)** versus the Random Walk method (0.0614 versus 0.3175) and against the comparison case of ANN with technical indicators (0.0614 versus 0.2378). This performance comparison was consistently observed across all business sectors.

## 6.2 Impact of the various framework layers

### 6.2.1 Impact of the input layer

In order to assess the impact of the input layer and the RMSE improvement attributable to this layer, an analysis was conducted to isolate the change in RMSE based on input sets used while keeping the rest of the model the same. The input type and machine learning method combination picked by the framework was used but without any implementation of the state layer. As a comparison case, the performance of the models with the same machine learning method but using only the technical indicators and without any implementation of the state layer was chosen. The average RMSE per business sector is shown in Table 6-3.

| Sector | # of Companies | Average RMSE: Models with Framework Input & ML | Average RMSE: Models with Technical Ind. And Framework ML | Average RMSE: Average Improvement in RMSE |
|---|---|---|---|---|
| Consumer Discretionary | 20 | 0.0778 | 0.2381 | 0.1603 |
| Consumer Staples | 16 | 0.0329 | 0.1157 | 0.0828 |
| Energy | 13 | 0.0763 | 0.2187 | 0.1425 |
| Financials | 14 | 0.0759 | 0.1724 | 0.0965 |
| Health Care | 18 | 0.0565 | 0.2306 | 0.1741 |
| Industrials | 31 | 0.0585 | 0.1672 | 0.1087 |
| Information Technology | 13 | 0.0771 | 0.2411 | 0.1640 |
| Materials | 11 | 0.0552 | 0.1635 | 0.1084 |
| Utilities | 11 | 0.0333 | 0.1125 | 0.0793 |
| **Overall** | **147** | **0.0608** | **0.1862** | **0.1254** |

**Table 6-3 Impact of the input layer of the framework**

An improvement in average RMSE of 0.1254 overall can be attributed to the input layer of the framework.  The level of improvement experienced is different for various business sectors.

### 6.2.2 Impact of the machine learning layer

In order to assess the impact of the machine learning layer and the RMSE improvement attributable to this layer, an analysis was conducted to isolate the change in RMSE based on using the framework-suggested machine learning method.  For reasons covered in Section 3.2.5.8 ANN was selected as comparison case for the chosen machine learning method. Thus, the performance of the models suggested by the framework (input type and machine learning combination) was compared against the same input type but using ANN for the machine learning methods. In both cases, the state layer was not implemented. The average RMSE by business sector for each case is shown in Table 6-4.

| Sector | # of Companies | Average RMSE | | |
|---|---|---|---|---|
| | | Models with Framework Input & ML | Models with ANN with Framework Input Type | Average Improvement in RMSE |
| Consumer Discretionary | 20 | 0.0778 | 0.1778 | 0.1001 |
| Consumer Staples | 16 | 0.0329 | 0.0847 | 0.0518 |
| Energy | 13 | 0.0763 | 0.1844 | 0.1082 |
| Financials | 14 | 0.0759 | 0.1577 | 0.0818 |
| Health Care | 18 | 0.0565 | 0.1619 | 0.1054 |
| Industrials | 31 | 0.0585 | 0.1388 | 0.0803 |
| Information Technology | 13 | 0.0771 | 0.1956 | 0.1185 |
| Materials | 11 | 0.0552 | 0.1356 | 0.0804 |
| Utilities | 11 | 0.0333 | 0.0963 | 0.0631 |
| **Overall** | **147** | **0.0608** | **0.1485** | **0.0877** |

**Table 6-4 Impact of the machine learning layer of the framework**

In the experiments conducted, the framework's ability to compare various options of machine learning methods and pick the best performing one have improved RMSE by 0.0877 on average.

### 6.2.3 Impact of the state layer

In order to assess the impact of the state layer of the framework, the analysis was conducted to isolate the change in RMSE based on using the framework-suggested state layer compared to not using the state layer at all. Table 6-5 compares the performance of the models suggested by the framework with their counterparts without state layer implementation. Given that in the majority of the cases the framework did not choose to implement the state layer as part of the optimum model, the comparison of the results of the two cases did not yield any difference when viewed across 147 cases. Thus, only the cases where the framework picked a state layer implementation were reviewed. Table 6-5 shows those cases and compares the performance against the same model make up but without the state layer implementation.

| Ticker | Company Name | Input Type | State Layer | ML | Framework | No State Layer | Difference |
|---|---|---|---|---|---|---|---|
| FITB | Fifth Third Bancorp | T | VIX_7 | SVR | 0.2035 | 0.2155 | **0.0120** |
| HAR | Harman International Industries | C | VIX_3 | SVR | 0.1233 | 0.1078 | (0.0155) |
| HBAN | Huntington Bancshares Incorpora | T | VIX_3 | SVR | 0.2038 | 0.1354 | (0.0684) |
| IR | Ingersoll-Rand plc (Ireland) | C | VIX_3 | SVR | 0.0913 | 0.0789 | (0.0124) |
| MMM | 3M Company Common Stock | C | VIX_3 | SVR | 0.0427 | 0.0477 | **0.0050** |
| MRK | Merck & Company, Inc. Common St | C | VIX_3 | SVR | 0.0511 | 0.0488 | (0.0023) |
| PNC | PNC Financial Services Group, I | C | VIX_3 | SVR | 0.0549 | 0.0476 | (0.0073) |
| RCL | Royal Caribbean Cruises Ltd. Co | C | VIX_3 | SVR | 0.1108 | 0.1087 | (0.0022) |
| ROK | Rockwell Automation, Inc. Commo | C | VIX_3 | SVR | 0.0988 | 0.1056 | **0.0068** |
| STI | SunTrust Banks, Inc. Common Sto | C | VIX_3 | SVR | 0.1063 | 0.1099 | **0.0036** |
| TXT | Textron Inc. Common Stock | C | VIX_3 | SVR | 0.0953 | 0.0887 | (0.0066) |
| **Overall** | | | | | **0.1074** | **0.0995** | **(0.0079)** |

**Table 6-5 Impact of the state layer of the framework**

Out of the 11 cases reviewed, the state layer was able to improve RMSE in 4 cases (with Difference column values highlighted in bold).

## 6.3 Use of Framework in Financial Forecasting

As part of the experiments, the framework was implemented to forecast percentage change in stock price in 252 days for a select group of companies (147). Sections 6.2.1 through 6.2.3 focused on the performance of the framework as delivered via the different layers of the framework. One of the benefits of the proposed framework is to provide the means to bring the researchers and practitioners closer to each other. One of the ways the proposed framework achieves this is through the forecasting model that is output per each company. The framework does take the view to generate a model definition after looking through the various input sets and machine learning methods made available to it, and come up with the model that will likely produce successful forecasts. Being able to generate a model definition that is customized per each company does interest the practitioners who can easily integrate the forecasts generated through the framework into their algorithmic trading strategies. Another way the framework adds value is through the comparative results that it generates along the way to selecting the forecasting model. This is, in turn, enabled through the layered approach taken in the framework's design. The fact that it can accommodate very easily use of various alternatives at each layer makes it a useful tool for researchers. For example, it is easy to change up the machine learning methods used, the input variable sets used, as well as using various market mood indicators. Therefore, not only does the framework accommodates generating forecasts but it also lends itself to generating results which can be comparatively studied (State vs. No state, various Machine Learning methods, various input types) along different dimensions. Thus, the layered approach taken in defining the framework lends itself to being able to generate comparative results which can provide insights and further the understanding that goes along with the results, which again would serve to bridge the gap between practitioners and researchers.

## 6.4 Summary of the Framework Performance chapter

Chapter 6 has provided a review of the performance of the proposed framework. In Section 6.1 the framework is compared to the benchmark cases and has been shown to outperform them significantly **(statistical significance, p=0.05)**. Section 6.2 provided a review of the contribution of the various layers (input, machine learning and state) of the framework towards improved forecasting accuracy. The analysis of the simulations conducted indicate that having relevant inputs and accounting for the states of the stock market can positively impact the forecasting performance of machine learning models. The framework proposed provides these benefits jointly by bringing it all under one common umbrella.

# Chapter 7. Conclusion and Future Work

Chapter 7 provides an overview of the major opportunities that were identified and motivated the work in this thesis, and summarizes the key conclusions from the investigations into the results of the experiments covered in chapters 4, 5 and 6; finally, future work is considered. Section 7.1 covers the conclusions from the work undertaken, where it provides a review of the main motivations behind the work, the research questions posed, and the results of the observations from the experiments conducted. Section 7.2 provides ideas for future work.

## 7.1 Conclusions

### 7.1.1 Motivation of Work

As covered in Section 2, two research needs have been identified in the area of stock price forecasting using machine learning methods: finding the relevant inputs, and accounting for the moods of the stock market. The first opportunity deals with a potential disconnect between finance practitioners and machine learning researchers with respect to the data used for investment decision making. Finance practitioners who carry out or advise with regards to making investment decisions typically belong to one of the two main schools of thought: the technical, or fundamental analysis. The technical analysis is based on past price and volume data only, whereas the fundamental analysis is based on the financial drivers of the company, industry, and economy. Historically, these two approaches have been seen as mutually exclusive, yet benefits of using them in conjunction have been articulated and appear to be gaining momentum in the Finance circles. Machine learning methods have been successfully applied to stock price forecasting, however, researchers tend to predominantly use the technical indicators and shy away from the fundamental indicators as these are relatively harder to obtain.

The second thread for the research was with regards to accounting for the moods of the overall stock market. The dynamic nature of the stock markets, and the influence of fluctuations of the overall stock market and other external factors (such as political, etc.) on individual stock prices have posed challenges for researchers when forecasting stock prices using machine learning methods (Cavalcante et al., 2016). Applying clustering methods to financial time-series data of the individual stock and developing localized models for the resulting clusters have been successfully used. This was mainly done to address that the relationship between predictors and predicted variable tends to change over time. Inspired by this method, the approach proposed was to create clusters based on the moods of the overall stock market first and to develop models using the timeframe and training data from the relevant cluster to forecast the price of the stock.

### 7.1.2 Evaluation of investigations

With respect to RQ1 and RQ2, an investigation was undertaken into whether there was a difference in using the technical versus fundamental indicators, and whether using them together was better than using them in isolation. Experiments were set up implementing models forecasting 252 day out stock price change for 147 companies of S&P 500 index using the technical indicators, fundamental indicators and a combined set of indicators. Models using ANN, SVR, DT, and LR machine learning methods have been trained on each of these 3 input sets and tested on out-of-sample test data. The forecasting accuracy of these models using the 3 input sets have been compared to generate observations and provide insights. Comparing the forecasting accuracy, it was shown that both the technical indicators and the fundamental indicators are relevant for stock price forecasting. Furthermore, it was shown that in 66% of the cases, the models with the fundamental indicators were able to **statistically significantly (p=0.05)** outperform models with the technical indicators. The reverse was only true in 24% of cases. In response to RQ1, whether the technical or fundamental analysis was better from machine learning forecasting point of view, with respect to 252 day out stock price forecasting, the models using indicators supplied by the fundamental analysis outperformed models using indicators supplied by the technical analysis. Furthermore, the models with the combined indicators **statistically significantly (p=0.05)** outperformed the models with the technical indicators in 84.5% of cases, and outperformed models using the fundamental indicators in 81% of cases, and in 78% of the cases outperform both. Thus, the results provide support for the view of finance practitioners that using the technical and fundamental analysis together is beneficial.

Why would it matter that machine learning researchers predominantly use the technical indicators and tend to ignore the fundamental indicators? When using solely the technical indicators, comparison of models with various machine learning methods implemented has shown that DT has outperformed the others in 99% of the cases, with SVR outperforming in the remainder of the cases. Thus, a machine learning researcher who might be using solely the technical indicators would probably conclude that DT is the better machine learning method for the problem. When using solely the fundamental indicators, DT and SVR emerge as the best performers in approximately half of the cases. However, when using the combined indicators, SVR dominated the other ML methods. Thus, focusing in on a subset of the relevant data (i.e. the technical indicators), can result in generating an incomplete picture and can affect the outcomes of analysis. Therefore, this suggests it is important to use both fundamental and technical indicators together to ensure that the models are exposed to a more representative spectrum of the influential factors available.

A review of the features making up the individual input sets have revealed the following

- In the case of fundamental analysis, company, competitor, and industry data provide relatively more influential features and that the overall macroeconomic data was relatively less influential.
- The more influential indicators in the combined input set was formed of both technical indicators and fundamental indicators which provide evidence of the synergistic effect from using them as complements and not substitutes

With regards to RQ3, experiments were carried out to investigate whether accounting for states of the overall stock market within the forecasting process in this way would result in any improvement over the traditional approach of training the models using the most recent available time-series data. Therefore, for the 147 companies the performance of forecasting models with the state layer identifying the states of the stock market by creating clusters of predefined sizes (3, 5, and 7) using selected market mood indicators (VIX, RSI of SP500, and Put-Call Ratio) were contrasted with models developed in the traditional manner. With respect to comparing the performance of the various market mood definitions used, it was observed that the smallest cluster size (i.e. 3) has performed better than larger cluster sizes (5 and 7) and that VIX and Put-Call Ratio had similar performance where RSI of SP500 was relatively ineffective in capturing the mood of the overall stock market.

The performance of the state layer is impacted by the input set selected. The percentage of cases where models with state layer outperform their counterparts is 49.7% when using the technical indicators. This increases to 66% when using the fundamental indicators, but drops to 15.6% when using the combined indicators. It was also found that the performance of the state layer is somewhat dependent on the machine learning method used. Performance of the models using ANN, DT, and LR as the machine learning method significantly **(p=0.05)** improved when the state layer was used, however in the case of SVR it has worsened slightly (0.0695 vs 0.0604). Therefore the combined nature of the indicators and SVRs ability to generalize effectively have nullified the contribution provided by the state layer.

Based on the investigation of the input and state dimensions, a novel framework was proposed which allows for identification of relevant indicator sets, accounts for the various states of the overall stock market, and allows the flexibility to identify effective machine learning methods. With respect to investigation towards the RQ4, for the 147 companies, the framework was implemented to generate forecasts for 252 days out stock price change. In all the cases, the framework was able to outperform the benchmark base case of the Random Walk method (RMSE of 0.0614 versus 0.3175). Given that ANNs have been implemented widely and that researchers have a tendency to use the technical indicators, this was identified as a comparison benchmark. The framework outperformed the ANN-based models with technical indicators (0.0614 vs 0.2378). The framework generates this improved performance through its various layers (input, machine learning, and state). In the experiments conducted the majority of the benefit was generated through the input and machine learning layers.

The results from the experiments indicate that the choice of which machine learning method is best should be based on the input sets available. The ability of the forecasting models to benefit from input sets also depends on the machine learning method used. It was shown that the models utilizing ANN and SVR were able to utilize the information from fundamental indicators more than DT and LR methods. For example, the RMSE for DT and LR based models exposed to technical, fundamental, and combined indicators were relatively close to each other with average rates of improvement achieved from using combined indicators of 16% and 23% respectively. However, for ANN, and to a greater degree for SVR-based models, the RMSE improvement obtained from using the combined indicators versus technical or fundamental were much more pronounced (33% ~ 60%). Thus, having a framework that can accommodate the selection of the best performing machine learning method based on the input sets would be valuable.

Naturally, the observations made from this body of work do have certain limitations. For example, only the forecasting horizons of 252 and 126 days were considered. Different forecasting horizons (60 days, 30 days, etc.) might result in different type of observations being made. Another limitation is with regards to the features that were picked for the fundamental and the technical indicator sets. Also, the data set for this study only contained companies from S&P500 that only includes very large US corporations. Thus, small to medium enterprises were not part of the data set, and the study was US-centric. Furthermore, the size of the data set (147 companies only, and covering a period of x years) should also be considered as a limitation of the work contained in this thesis.

## 7.2 Future Work

The future work can be broadly categorized into two main groups. Firstly, taking the existing body of work included in the thesis and implementing enhancements/changes to the proposed framework via changing the parameters (different inputs, different machine learning methods, and different forecasting horizons) and approaches taken (different approaches to implementation of the state layer). Section 7.2.1 describes some potential future work that can be undertaken with regards to this first category. Secondly, applying the framework on not just stock price forecasting but other domains where there is time series forecasting. Section 7.2.2 describes some potential future work that can be undertaken with regards to this second category.

### 7.2.1 Stock price forecasting related future work

As pointed out in Cavalcante et al. (2016) and Heaton et al. (2016), deep learning-based methods present a lot of opportunity for financial forecasting. Through the many layers that features are passed through Deep Learning does present various advantages over the "shallow" machine learning models (as implemented in the thesis) where more complex interactions between features can be detected, and the available input data is not necessarily limited to the ones provided/recommended by the experts (i.e. supervised learning). For example, in our experiments, the input features were based on the technical and fundamental analysis and were selected from indicators used by other researchers. Deep Learning does provide opportunities for being able to provide a larger set of input features at raw form and allow the deep learning layers to perform the feature extraction. This is mainly done through stacking many hidden layers, autoencoders, to enable the necessary feature extraction.

Another opportunity provided by deep learning approaches is to be able to keep track and have a memory of the hidden states of the environment as the learning takes place. Recurrent Neural Networks (RNN) have been implemented but are known to perform poorly when long term learning is need due to "vanishing and exploding gradients that can result from propagating through the many layers" (Heaton et al., 2016). A form of RNN that is immune to this weakness is Long Short-Term Memory (LSTM) model, where the LSTMs additionally include an "input gate, a forget gate, and input modulation gate, and a memory cell" (Heaton et al., 2016). Chen et al. (2015) provided 30-day historical data sequences to an LSTM model which in return generated 3 days out forecast for Shanghai and Shenzen stock markets. The model by Chen et al. (2015) was able to outperform the Random Walk method in terms of accuracy (14.3% vs. 27.2%). Bao et al. (2017) proposed a model for forecasting financial stock markets which was formed of three components: Wavelet transform (WT), Stacked autoencoders (SAE), and LSTM. The Wavelet transform is applied to remove noise from the financial time series, this is then fed to the SAE component, which through its many layers ensures that the feature extraction is carried out and finally LSTM component tracks the learning through time steps and enhance to forecasting capability. In order to assess the effectiveness of the SAE component, Bao et al. (2017) compared their proposed model versus a conventional RNN, plain LSTM, and combination of only the Wavelet Transformation and LSTM components. These models were applied in the prediction of six stock markets, which were representative of various levels of stock market maturity: Developing (CSI 300, Nifty 50), Relatively Developed (Hang Seng Index, Nikkei 225), and Developed (S&P500 and DJIA index). The proposed model outperforms the rest based on average MAPE, the RNN performs the poorest, followed by LSTM, WLSTM. This is observed regardless of the maturity level of the stock market. Implementation of the model proposed by Bao et al. (2017) using the data and cases from this thesis can provide further insight into the power of the combined indicators.

### 7.2.2 Application of the framework in other domains where there is time series forecasting

There are a large number of domains where the framework developed can be put to use. One area where the framework developed can be applied is in health-related time-series forecasts. This encompasses a very large area of research. For example, Zhang et al. (2014) compared the performance of ARIMA, SVR and two decomposition methods (regression and exponential smoothing) at being able to forecast the likelihood of future epidemic diseases (e.g. brucellosis, hepatitis A, hepatitis B, etc.) in China. Zhang et al. (2014) noted that the outbreaks of these are influenced by factors such as "temperature, rainfall and sunshine, etc. "and also that "the extent of the seasonality is not quite similar among them." The simulations run by Zhang et al. (2014) compared the forecasting performance (MAPE, MAE, MSE) of the methods chosen and found that no one method was able to emerge as the dominant forecasting method but "that support vector machine generally outperforms the conventional ARIMA model and decomposition methods." The authors explain the relatively poor performance of the ARIMA by the fact that there was a "level shift" in the data in particular years, which threw off the models. This type of a study can also be replicated through the framework where in place of the stock market moods, other external influential factors (such as weather-related information) can be used and individual models can be developed to forecast the occurrence of future epidemic diseases.

Another area where the framework developed can be applied is in energy-related forecasting. Electricity load forecasting is one such area where there is non-linear relationships and high volatility in the data. Yang et al. (2019) applied Least Squares Support Vector Machines (LSSVM) to predict the half-hour electricity load of the following week, which is an example of a short-term load forecasting problem. Mohan et al. (2018) noted that electricity load data has “non-linear and non-stationary characteristics” and has “dependency to various exogenous factors including time, day, weather, seasonal economic aspects, and social activities”, which “make the load forecasting a difficult task.” There are a lot of parallels between the data characteristics for electricity load forecasting and financial time series forecasting and this seems to make it a natural candidate for the framework to be applied to.

# References


Aghabozorgi, S., Shirkhorshidi, S. A., Teh Ying Wah, Y. T. (2015). Time-series clustering – A decade review. Information Systems. 53, 16-38.

Amini, A., Rahnama, G., Alinezhad, A. (2015). Ranking and Managing Stock in the Stock Market Using Fundamental and Technical Analyses. Journal of Modern Processes in Manufacturing and Production, 4(3), 45-57.

Atsalakis, G. S., and Valavanis, K. P. (2009). Surveying stock market forecasting techniques - Part II: Machine learning methods. Expert Systems with Applications. 36, 5932-5941.

Bettman, J. L., Sault, S. S., and Schultz, E. L. (2009). Fundamental and technical analysis: substitutes or complements? Accounting and Finance , 21-36.

Bao W, Yue J, Rao Y (2017) A deep learning framework for financial time series using stacked autoencoders and long-short term memory. PLoS ONE 12(7): e0180944.

Bellman, R. (1961). Adaptive Control Processes: A Guided Tour. Princeton: Princeton University Press.

Campanella, F., Mustilli, M., and D'Angelo, E. (2016). Efficient Market Hypothesis and Fundamental Analysis: An Empirical Test in the European Securities Market. Review of Economics & Finance. 27-42.

Cavalcante, C. R., Brasileiro, C. R., Souza, L. F. V., Nobrega, P. J., Oliveira A. L. (2016). Computational Intelligence and Financial Markets: A Survey and Future Directions. Expert Systems with Applications. 55, 194-211.

Cavalcante, R. C. , and Oliveira, A. L. (2015). An approach to handle concept drift in financial time series based on Extreme Learning Machines and explicit Drift Detection. International joint conference on neural networks(IJCNN). 1–8).

Cavalcante, R. C. , and Oliveira, A. L. (2016). FEED: Feature Extraction for Explicit Concept Drift Detection in Time Series. International joint conference on neural networks (IJCNN).

Chandwani D., Saluja S. M. (2014). Stock Direction Forecasting Techniques: An Empirical Study Combining Machine Learning System with Market Indicators in the Indian Context. International Journal of Computer Applications. 92(11), 8-17.

Chen, H., Cheng-Few, L., Wei-Kang, S. (2016). Technical, Fundamental, and Combined Information for Separating Winners from Losers. Pacific-Basin Finance Journal. 39, 224-242.

Chen, K., Zhou, Y., Dai, F. (2015). A LSTM-based method for stock returns prediction: A case study of China stock market. IEEE International Conference on Big Data (Big Data), Proceedings, 2823-2824.

Cherif, A., Cardot, H., Bone R. (2011). SOM time series clustering and prediction with recurrent neural networks. Neurocomputing. 74, 1936 - 1944.

D'Urso, P., Cappelli, C., Lallo, D. D., Massari, R. (2013). Clustering of financial time series. Physica A: Statistical Mechanics and its Applications. 392, 2114-2129.

Earnings Yield Ratio. In Investopedia. (2016). Available at: http://www.investopedia.com/terms/e/earningsyield.asp [Accessed September 2016].

Esfahanipour, A., and Aghamiri, W. (2010). Adapted Neuro-Fuzzy Inference system on indirect approachTSKfuzzy rule base for stock market analysis. Expert Systems with Applications , 37 (7), 4742-4748.

FRED. (2016). Available: https://fred.stlouisfed.org/series/DTWEXM [September 2016].

Fu, T. (2011). A review on time series data mining.Engineering Applications of Artificial Intelligence. 24, 164-181.

Hadavandi, E., Ghanbari, A., and Abbasian-Naghneh, S. (2010a). Developing an Evolutionary Neural Network Model for Stock Index Forecasting.Advanced Intelligent Computing. Theories and Applications - Communications in Computer and Information Science. 93, 407-415.

Hadavandi, E., Shavandi, H., and Ghanbari, A. (2010b). Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting. Knowledge-Based Systems , 23, 800-808.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. (2009). The WEKA Data Mining Software: An Update. SIGKDD Explorations , Volume 11, Issue 1.

Hassan, M. R. (2009). A combination of hidden Markov model and fuzzy model for stock market forecasting. Neurocomputing , 3439-3446.

Hassan, M. R., and Nath, B. (2005). Stock market forecasting using hidden markov model: a new approach. 5th International Conference on Intelligent System Design and Application, (s. 192-196). Poland.

Hassan, M. R., Nath, B., and Kirley, M. (2007). A fusion model of HMM, ANN and GA for stock market forecasting. Expert Systems with Applications , 33 (1), 171-180.

Heaton, J.B., Polson, N.G., Witte, J.H. (2016). Deep Learning In Finance. arXiv preprint arXiv:1602.06561

Hong, K., Wu, E. (2016). The roles of past returns and firm fundamentals in driving US stock price movements. International Review of Financial Analysis. 43, 62-75.

Hothorn, T., Buehlmann, P., Dudoit, S., Molinaro, A., Laan V. D. M. (2006). Survival Ensembles. Biostatistics, 7(3), 355--373.

Hsu, S. H. (2009). A two-stage architecture for stock price forecasting by integrating self-organizing map and support vector regression. Expert Systems with Applications , 36 (4), 7947-7951.

Hsu, M., Lessman, S., Sung, M., Ma T. (2016). Bridging the divide in financial market forecasting: machine learners vs. financial economists. Expert Systems with Applications , 215 - 234.

Hull, J. C. (2009). Options, Futures, and Other Derivatives. New Jersey: Pearson Education International.

Hyndman, R. J. and G. Athanasopoulos (2014). Forecasting: principles and practice. OTexts

I/B/E/S. (2016). Available: I/B/E/S International [September 2016]. Retrieved from Wharton Research Data Service.

Ibidapo, I., Adebiyi, A., Okesola, O. (2017). Soft Computing Techniques for Stock Market Prediction: A Literature Survey. Covenant Journal of Informatics and Communication Technology. 5-2.

Jiang Y., Fang. X. (2015). Bull, bear or any other states in the us stock market? Economic Modelling. 44, 54–58.

Khoa, N., Sakakibara, K., and Nishikawa, I. (2006). Stock Price Forecasting using Back Propagation Neural Networks with Time and Profit Based Adjusted Weight Factors. SICE-ICASE International Joint Conference, (5484-5488).

Krantz, M. (2016). Fundamental Analysis For Dummies. Indianapolis, Indiana, USA: Wiley Publishing.

Krollner, B., Vanstone, B., and Finnie, G. (2010). Financial time series forecasting with machine learning techniques: A survey. Paper presented at the European symposium on artificial neural networks: Computational and machine learning. Bruges, Belgium.

Kuhn, M. . Contributions from Jed Wing, Steve Weston, Andre Williams, Chris Keefer, Allan Engelhardt, Tony Cooper, Zachary Mayer, Brenton Kenkel, the R Core Team and Michael Benesty. (2014). caret:
Classification and Regression Training. R package version 6.0-37. http://CRAN.Rproject.org/package=caret

Kuhn, M., Johnson, K. (2013). Applied Predictive Modeling.Springer.

Kursa, B. M., Rudnicki R. W. (2010). Feature Selection with the Boruta Package. Journal of Statistical Software. 36(11), 1-13.

Laboissiere, A. L., Fernandes, A.S. R., Lage, G. G. (2015). Maximum and minimum stock price forecasting of Brazilian power distribution companies based on artificial neural networks. Applied Soft Computing.35, 66-74.

Lo, A. W. (2004). The adaptive markets hypothesis. Journal of Portfolio Management. 30, 15-29.

Lo, A. W. (2017). Adaptive Markets Financial Evolution at the Speed of Thought. Press Princeton.

Lorenzo. D. R. (2013). Basic Technical Analysis of Financial Markets: A Modern Approach (Perspectives in Business Culture). Springer-Verlag Mailand.

Lui, Y. M. (1998). The use of fundamental and technical analyses by foreign exchange dealers: Hong Kong evidence. Journal of International Money and Finance , 535-545.

Malkiel, B. G. (2003). The EfŽ ficient Market Hypothesis and Its Critics. Journal of Economic Perspectives. 17, 59 – 82.

Malkiel, B. G. (2016). A Random Walk Down Wall Street: The Time-Tested Strategy for Successful Investing, 11th revised edition. W. W. Norton & Company.

Manahov, V., Hudson, R. (2014). A note on the relationship between market efficiency and adaptability – New evidence from artificial stock markets. Expert Systems with Applications. 41, 7436-7454.

Mohan N., Soman K. P., Kumar S. S. (2018) A data-driven strategy for short-term electric load forecasting using dynamic mode decomposition model. Applied Energy 232, 229–244.

Munnix, M., Takashi, S., Rudi S., Leyvraz, F., Seligman, T, Guhr, T., Stanley, E. H. (2012). Identifying States of a Financial Market. Scientific reports. 2. 644.

MSCI USA IMI SECTOR INDEXES. (2016). Available at: https://www.msci.com/msci-usa-imi-sector-indexes [Accessed September 2016].

Nazário, T. F. R., e Silva, L. J., Sobreiro, A. V., Herbert, K. (2017). A literature review of technical analysis on stock markets. The Quarterly Review of Economics and Finance, Elsevier. 66, 115-126.

Nassirtoussi, K. A., Aghabozorgi, S., Wah, Y. T., Ngo, C. L. D. (2014). Text mining for market prediction: A systematic review. Expert Systems with Applications. 41, 7653-7670.

Oberlechner, T. (2001). Importance of Technical and Fundamental Analysis in the European Foreign Exchange Market. International Journal of Finance and Economics , 81-93.

Palit, A. K. , Popovic, D. (2006). Computational intelligence in time series forecast- ing: Theory and engineering applications (Advances in industrial control) . Springer- Verlag .

Patel, J. , Shah, S. , Thakkar, P. , and Kotecha, K. (2015). Predicting stock market index using fusion of machine learning techniques. Expert Systems with Applications, 42 (4), 2162–2172.

Quandl.com. (2016). WIKI Prices. Available at: https://www.quandl.com/product/WIKIP/WIKI/PRICES-Quandl-End-Of-Day-Stocks-Info [Accessed December 2016].

R Core Team. (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rockefeller, B. (2011). Technical Analysis for Dummies. Indianapolis, Indiana: Wiley Publishing.

Romero, J. P., Balch T. (2014). What Hedge Funds Really Do: An Introduction to Portfolio Management.Business Expert Press.

Russell, S., and Norvig, P. (2010). Artificial Intelligence: A modern Approach. New Jersey: Pearson.

Ryan A. J., Ulrich M. J. (2014). xts: eXtensible Time Series. R package version 0.9-7. http://CRAN.R-project.org/package=xts

Schwager, J. D., and Turner, S. C. (1995). Futures: Fundamental Analysis. Wiley.

Shonkwiler, R. W. (2013). Finance with Monte Carlo. New York: Springer.

Strobl, C., Boulesteix, A., Zeileis, A., Hothorn, T.(2007). Bias in Random Forest Variable Importance Measures: Illustrations, Sources and a Solution. BMC Bioinformatics, 8(25)

Strobl, C., Boulesteix, A., Kneib, T., Augustin, T., Zeileis, A.(2008). Conditional Variable Importance for Random Forests. BMC Bioinformatics, 9(307)

Torgo, L. (2017). Data mining with R : learning with case studies, 2nd Edition. Chapman & Hall/CRC, Florida.

Thomsett, M. C. (2006). Getting Started in Fundamental Analysis. Hoboken, New Jersey: Wiley Publishing.

Thomsett, M. C. (2015). Getting Started Getting Started in Stock Analysis, Illustrated Edition. Hoboken, Singapore: John Wiley Publishing.

Tkáč, M., Verner, R. (2016). Artificial neural networks in business: Two decades of research. Applied Soft Computing. 38, 788-804.

Tsinaslanidis, P. E. , and Kugiumtzis, D. (2014). A prediction scheme using perceptually important points and dynamic time warping. Expert Systems with Applications. 41 , 6848–6860.

Tsinaslanidis, E. P., Zapranis, D. A. (2016). Technical Analysis for Algorithmic Pattern Recognition. Springer.

Ulrich, J. (2015). TTR: Technical Trading Rules. R package version 0.23-0. http://CRAN.R-project.org/package=TTR

Urquhart, A., Hudson, R. (2013). Efficient or adaptive markets? Evidence from major stock markets using very long run historic data. International Review of Financial Analysis. 28, 130-142.

Vanstone, B., Finnie, G.(2009). An empirical methodology for developing stockmarket trading systems using artificial neural networks". Expert Systems with Applications. 36(3), 6668-6680.

Vanstone, B., Tan, C. (2003). A Survey of the Application of Machine learning to Investment and Financial Trading. Proceedings of the Australian and New Zealand Intelligent Information Systems Conference.Australian and New Zealand Intelligent Information Systems., (s. 211-216). Sydney.

Wafi, S. A., Hassan, H., Mabrouk, A. (2015a). Fundamental Analysis Models in Financial Markets – Review Study. Procedia Economics and Finance. 30, 939-947.

Wafi, A. S., Hassan, H. & Mabrouk, A., (2015b). Fundamental Analysis vs. Technical Analysis in the Egyptian Stock Exchange – Empirical Study. Proc. of the Third Int. Conf. on Advances in Economics, Management and Social Study.

Weng, B., Ahmed, A. M., Megahed, M. F. (2017). Stock market one-day ahead movement prediction using disparate data sources. Expert Systems with Applications. 79, 153-163.

Wu, S., Lee, S. (2015). Employing local modeling in machine learning based methods for time-series prediction, Expert Systems with Applications. 42, 341-354

Xu, W., Zuo, M., Zhang, M., and He, R. (2010). Constraint bagging for stock price prediction using neural networks. Proceedings of the 2010 International Conference on Modelling, Identification and Control. Okayama, Japan.

Yahoo Finance. (2016). Available at: https://finance.yahoo.com/ [Accessed September 2016].

Yang A., Li W., Yang X. (2019) Short-term electricity load forecasting based on feature selection and Least Squares Support Vector Machines. Knowledge-Based Systems, 163, 159–173.

Zwart G., M. T. (2009). The economic value of fundamental and technical information in emerging currency markets. Journal of International Money and Finance , 28, 581-604.

Zhang X., Zhang T., Young A. A., Li X. (2014) Applications and Comparisons of Four Time Series Models in Epidemiological Surveillance Data. PLoS ONE 9(2): e88075.

Zhong, X., and Enke, D. (2017). Forecasting daily stock market return using dimensionality reduction. Expert Systems with Applications. 67, 126-139.

# APPENDIX

## Appendix A

### List of Companies included in the study

| Ticker | Company Name | Sector | Industry |
|---|---|---|---|
| GPS | Gap, Inc. | Consumer Discretionary | Apparel Stores |
| JWN | Nordstrom, Inc. | Consumer Discretionary | Apparel Stores |
| LB | L Brands, Inc. | Consumer Discretionary | Apparel Stores |
| ROST | Ross Stores, Inc. | Consumer Discretionary | Apparel Stores |
| URBN | Urban Outfitters, Inc. | Consumer Discretionary | Apparel Stores |
| JCI | Johnson Controls International | Consumer Discretionary | Auto Parts |
| KSS | Kohl's Corporation | Consumer Discretionary | Department Stores |
| TJX | TJX Companies, Inc. | Consumer Discretionary | Department Stores |
| HAR | Harman International Industries | Consumer Discretionary | Electronic Equipment |
| HD | Home Depot, Inc. | Consumer Discretionary | Home Improvement Stores |
| LOW | Lowe's Companies, Inc. | Consumer Discretionary | Home Improvement Stores |
| SWK | Stanley Black & Decker, Inc. Co | Consumer Discretionary | Machine Tools & Accessories |
| DHI | D.R. Horton, Inc. | Consumer Discretionary | Residential Construction |
| LEN | Lennar Corporation Class A Comm | Consumer Discretionary | Residential Construction |
| PHM | PulteGroup, Inc. | Consumer Discretionary | Residential Construction |
| RCL | Royal Caribbean Cruises Ltd. Co | Consumer Discretionary | Resorts & Casinos |
| GT | The Goodyear Tire & Rubber Comp | Consumer Discretionary | Rubber & Plastics |
| SNA | Snap-On Incorporated to | Consumer Discretionary | Small Tools & Accessories |
| SPLS | Staples, Inc. | Consumer | Specialty Retail, Other |

| | | | |
|---|---|---|---|
| | | Discretionary | |
| HAS | Hasbro, Inc. | Consumer Discretionary | Toys & Games |
| CHD | Church & Dwight Company, Inc. C | Consumer Staples | Cleaning Products |
| COST | Costco Wholesale Corporation | Consumer Staples | Discount, Variety Stores |
| WMT | Wal-Mart Stores, Inc. t | Consumer Staples | Discount, Variety Stores |
| WBA | Walgreens Boots Alliance, Inc. | Consumer Staples | Drug Stores |
| ADM | Archer-Daniels-Midland Company | Consumer Staples | Farm Products |
| SYY | Sysco Corporation | Consumer Staples | Food Wholesale |
| KR | Kroger Company to | Consumer Staples | Grocery Stores |
| CVS | CVS Health Corporation | Consumer Staples | Health Care Plans |
| HRL | Hormel Foods Corporation n | Consumer Staples | Meat Products |
| TSN | Tyson Foods, Inc. | Consumer Staples | Meat Products |
| CL | Colgate-Palmolive Company | Consumer Staples | Personal Products |
| KMB | Kimberly-Clark Corporation Comm | Consumer Staples | Personal Products |
| PG | Procter & Gamble Company | Consumer Staples | Personal Products |
| CPB | Campbell Soup Company t | Consumer Staples | Processed & Packaged Goods |
| GIS | General Mills, Inc. toc | Consumer Staples | Processed & Packaged Goods |
| MKC | McCormick & Company, Incorporat | Consumer Staples | Processed & Packaged Goods |
| APC | Anadarko Petroleum Corporation | Energy | Independent Oil & Gas |
| CHK | Chesapeake Energy Corporation C | Energy | Independent Oil & Gas |
| COG | Cabot Oil & Gas Corporation Com | Energy | Independent Oil & Gas |
| DVN | Devon Energy Corporation n | Energy | Independent Oil & Gas |
| EOG | EOG Resources, Inc. toc | Energy | Independent Oil & Gas |
| EQT | EQT Corporation | Energy | Independent Oil & Gas |
| MRO | Marathon Oil Corporation n | Energy | Independent Oil & Gas |
| MUR | Murphy Oil Corporation | Energy | Independent Oil & Gas |
| NBL | Noble Energy Inc. | Energy | Independent Oil & Gas |
| NFX | Newfield Exploration Company Co | Energy | Independent Oil & Gas |
| SWN | Southwestern Energy Company Com | Energy | Independent Oil & Gas |
| HAL | Halliburton Company toc | Energy | Oil & Gas Equipment & Services |

| | | | |
|---|---|---|---|
| SLB | Schlumberger N.V. | Energy | Oil & Gas Equipment & Services |
| AFL | AFLAC Incorporated | Financials | Accident & Health Insurance |
| UNM | Unum Group | Financials | Accident & Health Insurance |
| WY | Weyerhaeuser Company to | Financials | Lumber, Wood Production |
| PNC | PNC Financial Services Group, I | Financials | Money Center Banks |
| STI | SunTrust Banks, Inc. to | Financials | Money Center Banks |
| HRB | H&R Block, Inc. | Financials | Personal Services |
| ALL | Allstate Corporation  Comm | Financials | Property & Casualty Insurance |
| CINF | Cincinnati Financial Corporatio | Financials | Property & Casualty Insurance |
| L | Loews Corporation | Financials | Property & Casualty Insurance |
| PGR | Progressive Corporation  C | Financials | Property & Casualty Insurance |
| FITB | Fifth Third Bancorp | Financials | Regional - Midwest Banks |
| HBAN | Huntington Bancshares Incorpora | Financials | Regional - Midwest Banks |
| KEY | KeyCorp | Financials | Regional - Midwest Banks |
| ZION | Zions Bancorporation | Financials | Regional - Pacific Banks |
| BIIB | Biogen Inc. | Health Care | Biotechnology |
| CELG | Celgene Corporation | Health Care | Biotechnology |
| GILD | Gilead Sciences, Inc. | Health Care | Biotechnology |
| REGN | Regeneron Pharmaceuticals, Inc. | Health Care | Biotechnology |
| VRTX | Vertex Pharmaceuticals Incorpor | Health Care | Biotechnology |
| BMY | Bristol-Myers Squibb Company Co | Health Care | Drug Manufacturers - Major |
| JNJ | Johnson & Johnson | Health Care | Drug Manufacturers - Major |
| LLY | Eli Lilly and Company t | Health Care | Drug Manufacturers - Major |
| MRK | Merck & Company, Inc. t | Health Care | Drug Manufacturers - Major |
| PFE | Pfizer, Inc. | Health Care | Drug Manufacturers - Major |
| MYL | Mylan N.V. | Health Care | Drugs - Generic |
| ESRX | Express Scripts Holding Company | Health Care | Health Care Plans |
| ABT | Abbott Laboratories toc | Health Care | Medical Appliances & Equipment |
| BSX | Boston Scientific Corporation C | Health Care | Medical Appliances & Equipment |
| HOLX | Hologic, Inc. | Health Care | Medical Appliances & Equipment |

| | | | |
|---|---|---|---|
| PDCO | Patterson Companies, Inc. | Health Care | Medical Equipment Wholesale |
| BAX | Baxter International Inc. | Health Care | Medical Instruments & Supplies |
| XRAY | DENTSPLY SIRONA Inc. | Health Care | Medical Instruments & Supplies |
| NOC | Northrop Grumman Corporation Co | Industrials | Aerospace/Defense - Major Diversified |
| TXT | Textron Inc. | Industrials | Aerospace/Defense - Major Diversified |
| BA | Boeing Company to | Industrials | Aerospace/Defense Products & Services |
| GD | General Dynamics Corporation Co | Industrials | Aerospace/Defense Products & Services |
| LMT | Lockheed Martin Corporation Com | Industrials | Aerospace/Defense Products & Services |
| RTN | Raytheon Company | Industrials | Aerospace/Defense Products & Services |
| UTX | United Technologies Corporation | Industrials | Aerospace/Defense Products & Services |
| CTAS | Cintas Corporation | Industrials | Business Services |
| DHR | Danaher Corporation toc | Industrials | Diversified Machinery |
| DOV | Dover Corporation | Industrials | Diversified Machinery |
| GE | General Electric Company n | Industrials | Diversified Machinery |
| HON | Honeywell International Inc. Co | Industrials | Diversified Machinery |
| IR | Ingersoll-Rand plc (Ireland) | Industrials | Diversified Machinery |
| ITW | Illinois Tool Works Inc. n | Industrials | Diversified Machinery |
| MMM | 3M Company | Industrials | Diversified Machinery |
| ROK | Rockwell Automation, Inc. | Industrials | Diversified Machinery |
| CAT | Caterpillar, Inc. | Industrials | Farm & Construction Machinery |
| DE | Deere & Company | Industrials | Farm & Construction Machinery |
| MAS | Masco Corporation | Industrials | General Building Materials |
| EMR | Emerson Electric Company n | Industrials | Industrial Electrical Equipment |
| PH | Parker-Hannifin Corporation Com | Industrials | Industrial Equipment & Components |
| PNR | Pentair plc. Ordinary Share | Industrials | Industrial Equipment & Components |
| GWW | W.W. Grainger, Inc. toc | Industrials | Industrial Equipment Wholesale |
| CSX | CSX Corporation | Industrials | Railroads |
| KSU | Kansas City Southern to | Industrials | Railroads |

| | | | |
|---|---|---|---|
| NSC | Norfolk Southern Corporation Co | Industrials | Railroads |
| UNP | Union Pacific Corporation | Industrials | Railroads |
| R | Ryder System, Inc. | Industrials | Rental & Leasing Services |
| FLIR | FLIR Systems, Inc. | Industrials | Scientific & Technical Instruments |
| RHI | Robert Half International Inc. | Industrials | Staffing & Outsourcing Services |
| JBHT | J.B. Hunt Transport Services, I | Industrials | Trucking |
| ADBE | Adobe Systems Incorporated | Information Technology | Application Software |
| HRS | Harris Corporation | Information Technology | Communication Equipment |
| WDC | Western Digital Corporation | Information Technology | Data Storage Devices |
| AAPL | Apple Inc. | Information Technology | Electronic Equipment |
| INTC | Intel Corporation | Information Technology | Semiconductor - Broad Line |
| TXN | Texas Instruments Incorporated | Information Technology | Semiconductor - Broad Line |
| SWKS | Skyworks Solutions, Inc. | Information Technology | Semiconductor - Integrated Circuits |
| XLNX | Xilinx, Inc. | Information Technology | Semiconductor - Integrated Circuits |
| LLTC | Linear Technology Corporation | Information Technology | Semiconductor - Specialized |
| MCHP | Microchip Technology Incorporat | Information Technology | Semiconductor - Specialized |
| AMAT | Applied Materials, Inc. | Information Technology | Semiconductor Equipment & Materials |
| LRCX | Lam Research Corporation | Information Technology | Semiconductor Equipment & Materials |
| MU | Micron Technology, Inc. | Information Technology | Semiconductor- Memory Chips |
| DD | E.I. du Pont de Nemours and Com | Materials | Agricultural Chemicals |
| APD | Air Products and Chemicals, Inc | Materials | Chemicals - Major Diversified |
| DOW | Dow Chemical Company Comm | Materials | Chemicals - Major Diversified |
| EMN | Eastman Chemical Company n | Materials | Chemicals - Major Diversified |
| FMC | FMC Corporation | Materials | Chemicals - Major Diversified |
| PX | Praxair, Inc. | Materials | Chemicals - Major Diversified |
| VMC | Vulcan Materials Company (Holdi | Materials | General Building Materials |

| | | | |
|---|---|---|---|
| SEE | Sealed Air Corporation | Materials | Packaging & Containers |
| IFF | Internationa Flavors & Fragranc | Materials | Specialty Chemicals |
| PPG | PPG Industries, Inc. to | Materials | Specialty Chemicals |
| SH W | Sherwin-Williams Company | Materials | Specialty Chemicals |
| NI | NiSource Inc | Utilities | Diversified Utilities |
| PEG | Public Service Enterprise Group | Utilities | Diversified Utilities |
| AEP | American Electric Power Company | Utilities | Electric Utilities |
| AES | The AES Corporation toc | Utilities | Electric Utilities |
| CMS | CMS Energy Corporation | Utilities | Electric Utilities |
| D | Dominion Resources, Inc. n | Utilities | Electric Utilities |
| DUK | Duke Energy Corporation (Holdin | Utilities | Electric Utilities |
| EIX | Edison International to | Utilities | Electric Utilities |
| PCG | Pacific Gas & Electric Co. Comm | Utilities | Electric Utilities |
| SO | Southern Company | Utilities | Electric Utilities |
| WEC | WEC Energy Group, Inc. | Utilities | Electric Utilities |

**Table A-1 Companies included in the study**