Financial Engineering Modelling Using Computational Intelligent Techniques: Financial Time Series Prediction

A thesis submitted for the degree of Doctor of Philosophy

by
Bashar Al-hnaity

Electronic and Computer Engineering Department
College of Engineering, Design and Physical Sciences
Brunel University London

November 2015

Abstract

Prediction of financial time series is described as one of the most challenging tasks of time series prediction, due to its characteristics and dynamic nature. In any investment activity, having an accurate prediction system will significantly benefit investors by guiding decision making, especially in trading, asset management and risk management. Thus, the attempts to build such systems have attracted the attention of practitioners in the market and also researchers for many decades.

Furthermore, the purpose of this thesis is to investigate and develop a new approach to predicting financial time series with consideration given to their dynamic nature. In this thesis, the prediction procedures will be carried out in three phases. The first phase proposes a new hybrid dynamic model based on Ensemble Empirical Mode Decomposition (EEMD), Back Propagation Neural Network (BPNN), Recurrent Neural Network (RNN), Support Vector Regression (SVR) and EEMD-Genetic Algorithm (GA)-Weighted Average (WA) to predict stock index closing price. EEMD in this phase is introduced as a preprocessing step to historical observation for the first time in the literature. The experimental results show that the EEMDD-GA-WA model performance is a notch above the other methods utilised in this phase. The second phase proposes a new hybrid static model based on Wavelet Transform (WT), RNN, Support Vector Machine (SVM), Nave Bayes and WT-GA-WA to predict the exact change of the stock index closing price. In this phase, the experimental results showed that the proposed WT-GA-WA model outperformed the rest of the models utilised in this phase. Moreover, the input data that are fed into the hybrid model in this phase are technical indicators.

The third phase in this research introduces a new Hybrid Heuristic-Rules-based System (HHRS) for stock price prediction. This phase intends to combine the output of the hybrid models in phase one and two in order to enhance the final prediction results.

Thus, to the best of our knowledge, this study is the only one to have carried out and tested this approach with a real data set.

The results show that the HHRS outperformed all suggested models over all the data sets. Thus, this indicates that combining different techniques with diverse types of information could enhance prediction accuracy.

Declaration

I certify that the effort in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree. I also certify that the work in this thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself have been duly acknowledged and referenced.

Signature of Student

Acknowledgements

This thesis is dedicated to my late sister soul Ayat, and my loving and caring father, as without his support I could never have achieved this, and my beloved mother Um Bashar for her continued prayers, moral support and patience. I also owe special thanks to my sisters and brothers for their continuous encouragement during my research. Thank you Alaa, Aseal, Farah, Ammar, Aseam and Mohamad. The person to whom I owe the most sincere appreciation is my supervisor, Dr. Maysam Abbod, whose influence runs through each page of this thesis. This thesis could not have been written without his guidance and knowledge.

Nomenclature

ABC	**A**rtificial **B**ee **C**olony
AI	**A**rtificial **I**ntelligent
ANFIS	**A**daptive **N**etwork Based **F**uzzy **I**nference **S**ystem
ANN	**A**rtificial **N**eural **N**etwork
APE	**A**verage **P**ercent **E**rror
AR	**A**uto **R**egressive
ARCH	**A**uto **R**egressive **H**eterocedasticity
ARIMA	**A**uto **R**egressive **I**ntegrated **M**oving **A**verage
ARMA	**A**uto **R**egressive **M**oving **A**verage
BH	**B**uy and **H**old
BPNN	**B**ack **P**ropagation **N**eural **N**etwork
CBR	**C**ase **B**ased **R**easoniong
CGA	**C**haotic **G**enetic **A**lgorithm
DA	**D**irectional **A**ccuracy
DT	**D**ecision **T**ree
EBNN	**E**lman **B**ackpropagation **N**eural **N**etworks
EEMD	**E**nsemble **E**mpirical **M**ode **D**ecomposition
EMA	**E**xponential **M**oving **A**verage
EMD	**E**mpirical **M**ode **D**ecomposition
EMH	**E**fficient **M**arket **H**ypothesis
ESN	**E**cho **S**tate **N**etwork
FD	**F**eature **D**iscretization
FNN	**F**uzzy **N**eural **N**etworks
GA	**G**enetic **A**lgorithm
GARCH	**A**uto **R**egressive **C**onditional **H**eterocedasticity
GMM	**G**eneralised **M**ethods of **M**oments
HHRS	**H**ybrid **H**euristic **R**ules **S**ystem
IBCO	**I**mproved **B**acterial **C**hemotaxis **O**ptimisation
IMF	**I**ntrinsic **M**ode **F**unctions
LASSO	**L**east **A**bsolute **S**hrinkage and **S**election **O**perator

LDA	**L**inear **D**iscriminant **A**nalysis
LM	**L**evenberg **M**arquardt
LT	**L**inear **T**ransformation
MA	**M**oving **A**verage
MACD	**M**oving **A**verage **C**onvergence and **D**ivergence
MAD	**M**ean **A**bsolute **D**ifference
MAE	**M**ean **A**bsolute **E**rror
MAPE	**M**ean **A**bsolute **P**ercentage Error
MAPE	**M**ean **A**bsolute **P**ercentage **E**rror
MARS	**M**ultivariate **A**daptive **R**egression **S**plines
MKSVR	**M**ultiple **K**ernel **S**upport **V**ector **R**egression
MLR	**M**ultiple **L**inear **R**egression
MOME	**M**omentum
MSE	**M**ean **S**quare **E**rror
NLICA	**N**onlinear **I**ndependent **C**omponent **A**nalysis
OLS	**O**rdinary **L**east **S**quares **M**ethods
PCA	**P**rincipal **C**omponent **A**nalysis
PNN	**P**robabilistic **N**eural **N**etwork
PSO	**P**article **S**warm **O**ptimization
QDA	**Q**uadratic **D**iscriminant **A**nalysis
R	**C**ross **C**orrelation **C**oefficient
RBFN	**R**adial **B**asis **F**unction **N**eural **N**etwork
RMSE	**R**oot **M**ean **S**quare **E**rror
RNN	**R**ecurrent **N**eural **N**etwork
ROI	**R**eturn **O**f an **I**nvestment
RSI	**R**elative **S**trength **I**ndex
RW	**R**andom **W**alk
SA	**S**imple **A**verage
SD	**S**tandard **D**eviation
SKSVR	**S**ingle **K**ernel **S**upport **V**ector **R**egression
SOM	**S**elf **O**rganizing **M**ap
SRCS	**S**tepwise **R**egression **C**orrelation **S**election
SSL	**S**emi **S**upervised **L**earning
SSVR	**S**easonal **S**uport **V**ector **R**egression
SV	**S**tochastic **V**olatility
SVM	**S**upport **V**ector **M**achine
SVR	**S**upport **V**ector **R**egression
TS	**T**abu **S**earch
WA	**W**eighted **A**verage
WT	**W**avelet **T**ransform

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Introduction

Taken at successive times in a sequence of observations is the definition of a time series. Extrapolating the discovered patterns from the historical data series into the future is what time series prediction methods attempt to achieve. The prediction domain of financial time series involves the projection of time series such as stock prices, interest rates, inflation, exchange rates and many other financial instruments into the future based on their historical values. In financial predicting, asset prices are considered as one of the most discussed topics. This is because it is important for management, asset pricing, portfolio optimisation and selection, option pricing and risk management to fully understand and predict the asset prices. To investors, whether they are individual or financial institutions, the accuracy of the prediction process is of great importance as the prediction output is utilised for investment decision making. However, there are number of factors that influence assets price such as economical, political and psychological factors [3] [129], and thus these factors introduce a high randomness into the financial time series assets price, which reflects on the prediction process and makes it a difficult task.

A financial time series is characterised by natural complexity. Firstly, the behaviour of the financial time series is nearly like a random walk process and also very much like white noise processes. Therefore, under these conditions, it is implied that theoretical prediction of a financial time series process is impossible [114]. Secondly, financial time series are characterised by features such as non-linearity and high non-stationary. Thirdly, they possess structure instability, which implies that external reasons or events influence the movement of the asset price. These events can be categorised as economic crises or abrupt changes in government policy [111] [222].

Due to these characteristics and the aforementioned complex features of financial time series, predicting them is considered as one of the most challenging tasks in time series prediction. Over the past two decades, many studies have investigated the predictability of financial time series, or more specifically stock prices. In the earlier studies on the predictability of stock prices, it was thought that there were no methods that could predict financial time series using historical data sets. Hence, at that time, only current price could be used as a best predictor of the future price. However, accordance to Malkiel and Fama [86] study on the efficient market hypothesis, prices of stock are informationally efficient and it is possible to use historical stock prices to predict future prices. Moreover, in the literature it has been shown that numerous studies employ different suitable methods that can lead to successful prediction values for financial time series.

Generally, in the literature, the utilised prediction techniques can be classified into two categories: 1) statistical models, 2) artificial intelligence models (AI). However, statistical models have failed to capture the non-linear patterns that exist in financial time series data as they are based on the assumption that a linear correlation exists in them. Therefore, and in order to tackle this limitation, AI techniques have been utilised to enhance the prediction process of financial time series. In addition, the Wolpert theorem [270] stated that no single AI technique could solve all problems. In this thesis, an investigation is carried out into the predictability of financial time series, and new prediction approaches are proposed in order to achieve better prediction results, compared to the statistical models and single-approach A1 techniques. The recent research on stock prediction has focused on two groups: 1) time series prediction and 2) trend prediction [288]. In this thesis, not only will the financial time series be predicted, but also the trend of asset prices will be studied in terms of predicting financial time series.

1.2 Motivation, Aim and Objective

Prediction of financial time series is considered as a challenging and important task in the field of finance. In any investment activities, having an accurate prediction system will significantly benefit investors by guiding decision making, especially in trading, asset management and risk management. Thus, the attempts to build such systems have attracted the attention of practitioners in the market and also researchers for many decades.

Nonetheless, given that the nature of financial time series is dynamic and complicated, perfect financial time series prediction and easy gains from trading and asset management should not be expected. In any attempts to predict financial time series, two main and fundamental questions should be answered:

- Based on historical observation, under what circumstances could a specific financial time series be predicted?
- How can static and dynamic models be combined to achieve high accuracy?

The aim of this research is to develop a hybrid system for accurate prediction of financial time series data. The above two questions illustrate the main two objectives of this research, which will be discussed in detail in this thesis, and are presented below:

- To assess the performance of specific prediction methods in terms of predicting financial time series, both stock price and the direction movement (trend). Different performance measurements are used in this thesis; if the predicting value of the out-of-sample (testing data set) in both models (predicting stock price or price direction) is better than the benchmark prediction values, then the proposed method has outperformed the benchmark models.

- To consider the dynamic and complex nature of financial time series when building and developing the new proposed methods is the main purpose of this thesis. Different techniques are introduced in the prediction process based on the nature of the data. Many points are considered when designing the new prediction models, such as preprocessing the original data set with new techniques from different areas such as signal processing.

1.3 Research Contributions

In this thesis the main contributions are as follows:

- Introducing new steps in distributing the in-sample data sets with respect to the time series consequences. It is concluded that using such steps can reduce the tolerance accuracy of the models and enhance the training quality of the utilised methods. Chapter 3
- Introducing a quantisation factor into the Support Vector Machine (SVM) and Support Vector Regression (SVR) models. To the best of the researchers knowledge, this step is proposed for the first time in SVR and SVM to model and predict financial time series. Chapter 3
- Designing a new Hybrid Model to predict financial time series. Ensemble Empirical Mode Decomposition was introduced for the first time into financial time series prediction. Chapter 4
- Designing a new hybrid model to predict the exact change in the stock price. To

the best of the researchers knowledge, this model is introduced for the first time for predicting stock market direction. Chapter 5

- Designing A New Hybrid Heuristic rules based System For Stock Price Prediction (HHRS). To the best of the researchers knowledge, this approach is introduced to predict financial time series by integrating two different approaches together for the first time. Chapter 6

1.4 Thesis Outline

The remaining chapters of this thesis are organised as follows:

- Chapter 2 provides a literature review on financial time series prediction. This includes some of the most popular efficient market hypotheses, the random walk theory, and the latest current time series prediction approaches and methodologies.
- In Chapter 3, the important conditions of the utilised data sets are represented. Steps for data preprocessing and data splitting are given. The predictability of single approach models is tested using statistical models and AI techniques. Different performance measurements are explained in this chapter.
- In Chapter 4, the design and architecture of the new proposed hybrid model are exhibited. The new data preprocessing EEMD technique is explained in detail. The performance of the new hybrid model is compared with other single approaches and methods in order to prove that the new proposed model has enhanced the prediction ability.
- Chapter 5 illustrates the ability of the new hybrid model to predict the exact change in the future movement of the stock price. The prediction preprocessing steps and the hybrid model architecture are exhibited in this chapter. The prediction results of the new hybrid model are compared in order to show that the new model has

outperformed the other utilised methods.

- Chapter 6 provides the concept of the new hybrid heuristic-rules-based system for stock price prediction. The architecture of the new hybrid model is explained. Furthermore, the rule-based system is exhibited for the first time in this chapter. Results and performance measurements are discussed in this chapter and the results are compared with other utilised methods.
- Finally, Chapter 7 provides a summary of the contributions of the thesis, and also presents a conclusion of the whole thesis. Moreover, some ideas for future research improvement and investigation are presented as future work for this thesis.

1.5 Publications

Portions of the work detailed in this thesis have been presented in national and international scholarly publications, as follows:

Al-hnaity, B., Abbod, M: "Ensemble SVR Model to Predict FTSE100 Index", In proceeding of ITISE2014, Granada, Spain, 2014.

Al-hnaity, B., Abbod, M: "A Novel Hybrid Ensemble Model to Predict FTSE100 Index by Combining Neural Network and EEMD", European Control Conference, ECC 2015, Jul 2015, Linz, Austria. Proceedings of the European Control Conference 2015.

Al-hnaity, B., Abbod, M: "Predicting FTSE100 Close price Using Hybrid Model" ,SAI Intelligent Systems Conference 2015 November 10-11,2015 London,UK.

Alaraj, M., Abbod, M., & Al-Hnaity, B. Evaluation of Consumer Credit in Jordanian Banks: A Credit Scoring Approach.UKsim 2015 March 2015 Cambridge.UK

Chapter 2

Literature Review

2.1 Introduction

For many years the predictability of the financial market has attracted the attention of market professionals, academics and investors [188]. However, the most recent upsurge in financial market predictability has come about through the adaptation of computational intelligent techniques to assist the rapidly growing volume of financial data. Moreover, financial time series modelling and predicting are widely acknowledged as arduous tasks [11]. This chapter will briefly provide a review of financial stock market prediction [243]. In the literature, the first attempts to develop temporal patterns in data were in 1662. Using bills of mortality, John Graunt at that time published several social and epidemiological comparisons [186] [149].

However, other theories suggest that the financial market cannot be predicted. The Efficient Market Hypothesis (EMH) is one of those theories. Section 2.2 will present a short discussion of this well-known hypothesis. Section 2.3 will present a review of previous prediction/forecasting attempts in the financial market, and then the most famous and widely used methods for financial prediction (fundamental analysis and technical analysis) will be presented in Sections 2.4 and 2.5. Section 2.6 presents the traditional financial time series prediction models. Furthermore, computational intelligence techniques and their first introduction into the financial domain will be discussed in Section 2.7 . The limitations of these techniques will be presented in the final section 2.8, and thus Chapter 3 will be briefly discussed.

2.2 Efficient Market Hypothesis

Financial market prediction is considered a difficult and highly complex domain. The efficient market hypothesis (EMH) states that financial market instruments such as stock prices follow a random walk pattern [85], which means that the probability of the stock going up or down is the same, concluding that any prediction model achieve more than 50% accuracy [245]. EMH was firstly proposed by Fama, who suggested that all the market information is already revealed and included in the stock price; therefore the natural and only possible way for stock prices to behave is by random walk [88]. In other words, all of the available information is always fully reflected in the stock prices, and as a result, no form of information can be used to enhance any prediction model to generate better accuracy. The stock price, according to EMH, will quickly and efficiently absorb any new information that arises in the market [178]. Fair game market is how Famas EMH described the investment activity in the financial market [85]. There are three kinds of Efficient Market levels identified by Fama:

- The weak form of EMH states that stock prices reflect all historical information, variations of price and volume of trade. Thus stock prices follow a random walk process, as future prices cannot be predicted using past prices.

- The semi-strong form of EMH states that stock prices reflect all public historical information, variations of price and volume of trade. Therefore, there is no advantage to using this information to predict the future price as everyone knows it.

- The strong form of EMH states that stock prices reflect all inside and public historical information, variations of price and volume of trade. Therefore, no one can gain an advantage from using such information to predict future prices since all the information is available to all market participants.

It was firstly suggested by [142] that stock price movement follows a random walk process, which is related to the weak form of EMH. The question of whether or not it is futile to attempt to predict future stock prices has given rise to deviations from EMH, despite the rich literature that supports EMH. Empirical and theoretical works have examined EMH and also argued whether stock market prices follow a random walk process or not.It has been argued by [177] that stock market prices can be predicted and do not follow a random walk process. The results of empirical and theoretical works are mixed and contradictory. Early studies were found supporting EMH [64] [85] [8]. It was claimed by [133] that among all the propositions in economics, EMH has the most solid supporting empirical evidence; however recent studies tend to show anomalies [93] [139] [169] [30]. EMH was stated as a false hypothesis in Famas second review [87]. In the next sections, several studies on predicting stock prices using different approaches will be presented.

2.3 Financial Time Series Prediction Attempts

John Graunt, as mentioned in the previous section, was one of the first scholars to make early prediction attempts. The definition of a time series, according to [28] [66], is a sequence of data points, measured typically at successive points in time and spaced at uniform time intervals. In other words, a time series is a collection of observations that measure some activities over time [44], whereas historical activities are recorded at equal consistent measurement space intervals, for example, day, week, month, etc. Time series are used in many fields, such as statistics, signal processing, pattern recognition, econometrics and mathematical finance. According to the literature, the first attempt to study financial time series behaviour was done by financial professionals and journalists, not academics. This became a long-standing tradition, as even up to the present day, many empirical studies have originated from the financial industry. These empirical studies have shown a potential gain where specialised data is the key to the success of any financial time series model. Charles Dow was the first publisher to publish the best example of a financial time series, in his editorials in the Wall Times between 1900 and 1902. The Dow editorial formed the basis of the Dow Theory and influenced what later became known as technical analysis [187].

Interest has grown among investors (individuals or institutions), stock fund managers and financial analysts in predicting financial time series, as a way of surviving in tough financial situations; the accurate prediction of any financial instrument has become an important issue in investment decision making. This huge interest and the ongoing financial global crisis have urgently increased the demand for accurately predicting stock price movement methods. Time series, both mathematical and statistical, are the most current models that financial prediction relies upon [9] [20] [23] [265] [25] [95] [181]. An extensive review was provided by [68] on time series prediction throughout 25 years. The review showed that during the past 25 years (until 2006), time series prediction grew with the advent of computational and statistical models, and a variety of mature prediction calculations and evaluation approaches. However, there are many unsolved problems remaining, and many others are arising, meaning that there is still a need for more work. Moreover, foreign exchange markets, bond markets and financial derivatives markets are the domains into which financial time series prediction can be extended [58] [250] [168].

For decades, financial time series prediction (especially stock price prediction) has been an attractive research topic due to the huge profit that can be gained if a highly accurate model is achieved. Therefore, researchers and financial analysts have strived to achieve higher accuracy with a variety of stock price prediction models. However, stock markets are characteristically dynamic, non-linear, complicated, non-parametric and chaotic in nature [234]. Furthermore, in the financial prediction domain many time series models have been suggested and implemented. Much recent work has been done on time series prediction, which reflects the importance of the topic [34] [33] [37] [55] [130] [138] [174] [179] [229] [276] [295]. In financial time series prediction models, auto-regressive conditional heteroscedasticity (ARCH) was proposed by Engle [83]; generalised auto-regressive conditional heterocedasticity (GARCH) was also proposed by Bollerslev [23]; and Box and Jenkins [25]; introduced the auto-regressive moving average (ARMA) and the auto-regressive integrated moving average (ARIMA) models. These are the best-known and most used time models. In addition, the auto-regressive (AR) model is the most popular and important model to predict financial time series [44]. However, all traditional statistical time series prediction models have a limited capability for modelling time series data; they do not yield an automated process, as they require additional historical data, and these data must be of normal distribution in order to obtain a good prediction performance [69]. As a result of these drawbacks, statistical time series models do not have the ability to track the complexity and non-stationary nature of the stock market [22]. Finally, in the literature, many attempts have been made in the stock price prediction domain utilising different methods, which can be categorised into four main categories: 1. Fundamental Analysis, 2. Technical Analysis, 3. Traditional financial time series prediction, and 4. Computational intelligence [70].A more detailed description of these prediction methods will be presented in the following sections.

2.4 Fundamental Analysis

This type of analysis is based on examining the companys financial statements and balance sheets to predict their shares future trends. Hence this type of analysis involves processing past records of assets, earnings, dividends, interest rates, sales, products, management and market information. Other commonly employed financial ratios are the current ratio, return on assets, and the liabilities ratio for prediction stock price. According to fundamental analysis, each stock has its intrinsic value, where the stock is undervalued if the intrinsic value is bigger than the stock price. So the investor should buy this stock in this case and not buy it in the opposite case [194]. Economic factors are also used in fundamental analysis. The static nature of macroeconomics data can influence the return of the individual stock and the stock index, and they also have a significant impact on the underlying growth of the company and its earning prospects. Moreover, stock market liquidity can be affected by economic variables. Inflation rates, employment figures and the price index are the main economic variables. In order for analysts to make a decision about whether or not to sell or buy stock, all of these factors should be taken into account [156].

2.5 Technical Analysis

Technical analysis is a study of the market itself where market action such as price (open, high, low, trading volume and close) can tell everything and is sufficient for prediction tasks. In technical analysis there are three main assumptions: 1. price and volume (market action) reflect everything; 2. prices move in trends; 3. history repeats itself. Technical analysis and EMH have a similar assumption, however each assumption has a totally different conclusion(s). New information, and whether it is fully and immediately reflected in the market. Technical analysts thus believe that in response to new information, stock prices move slowly. Hence, the driving forces hardly change in the market, and recurrent and predictable trends are shown in the prices. In other words, technical analysis, known as chartist, was based on studying the charts that depict historical market data, and from these a pattern will be derived to predict the market behaviour [91] [161] [171]. Finding any kind of pattern in the data is the main goal of technical analysis, so such findings can be used by analysts to make a prediction of stock or market movements. Despite its lack of popularity and the fact that it has not been accepted among analysts or academics in previous years, the use of technical analysis in recent years has increased [27] [238] [157] [170] [141] [203].

Charts, technical indicators, the Dow theory, Gann lines and Elliot waves are the technical analysis tools used to exploit recurring patterns [213]. Self-destruction is considered to be a big problem in technical techniques. Thus, the profitable opportunity will disappear quickly as the news of a better strategy becomes well-known and all the traders make the same decisions. As a result, a successful strategy should not be revealed. A successful trading strategy must be dynamic and self-adaptive, due to the regime shifting character of the market [115]. Generally, the ability to predict the direction of future prices has attracted the attention of researchers, and led to a focus on technical analysis in order to improve the return of investment [80] [190] [211] [182] [94]. Thus, price and volume are assumed to be the two most relevant factors in determining the direction and behaviour of a particular stock or market by technical analysis methods [17] [60] [233]. There are three main premises that technical analysis is based on. First, market action discounts everything. In another words, price changes are included in all the market price determining factors, such as macroeconomy, government interference, and investor psychology. Second, prices move in trends; this illustrates the main purpose of technical analysis which is to find the early stage of the trend and then the technicians will follow the trend until it reaches the end. Third, history might repeat itself. Evidently, by studying the past markets, data different chart patterns can be identified and categorised, which can lead to an early identification of a certain market trend. As these trends have occurred several times in the past, the assumption of the same trends happening again can be possible [127]. Technical indicators are the main technical analysis tool that has the ability to predict future price fluctuation and also to guide investors on whether to sell or buy a particular stock at the right time. Technical indicators come from mathematical formula, which are based on stock price and volume [60].

The purpose of providing the above information on technical analysis is not for building a theoretical or empirical justification. The purpose is to use technical indicators as a main tool of technical analysis as an input for the prediction model. More details on the suggested prediction model will be given in the following chapters.

2.6 Traditional Financial Time Series Prediction

In recent years, the ability to predict accurately has been crucial to many decision makers in different processes such as planning, organising, scheduling, selling and buying, strategy formulation, policy making and supply chain management. Hence the importance of highly accurate predictions has received much attention from researchers, and extensive effort has been invested into this area of research. Despite the huge interest and the number of studies that have been conducted on financial time series prediction, this field remains important and is an active topic of research at the present time and will continue to be in the future [288]. Due to security issues in the stock market, important financial bridges were developed between investors and listed companies, in which the dynamic changes in the stock market were directly reflected in the stock market index. Many benefits can be obtained from studying stock indices, such as providing a guiding reference for investors in order to make financial decisions, which additionally aids governments to develop the correct rules and reduce the negative effects of stock market volatility. They can also promote the stable development of stock markets, where stock indices closely develop the national economy [277].

In econometrics, forecasting, modelling and trading financial indices (financial time series) are considered to be among the most challenging and difficult topics. This is a result of the characteristics of financial time series: non-linear, non-stationary, noisy, with a high degree of uncertainty and hidden relationships. These characteristics are a reason for the information unavailability, affecting the behaviour of financial markets between past and future captured prices. Statistical tools have been used by many researchers in order to predict future stock prices, including the regression model, and the most popular methods, the GARCH model, the ARIMA model and the probabilistic model. In the regression model the output y depends on the input signals $x_1.....x_n$. Unfortunately, in the prediction of financial stock data, this type of model does not seem to do very well. Therefore, Busuttil and Kalnishkan [29] have developed two regression models to tackle such a weakness: the weight-controlled kernel aggregating algorithm for regression (WeCKAAR) and the kernel aggregating algorithm for regression with changing dependencies (KAARCh). The difference between these two new models is that WeCKAAR adds a decaying weight to the existing regression techniques, while KAARCh adds an aggregating algorithm to the trading signals.

Moreover, the GARCH model is used in traditional stock prediction and is also considered to be an important and popular model. The GARCH model assigns exponentially decreasing weights to the original regression model. It was used by Diebold and Mariano [74] to predict the exchange rate. However, the result happened to be the same as that of the traditional regression model. Linear and non-linear models were tested by Gencay using GARCH models in an attempt to achieve significant prediction signal data; however, the results were disappointing. The Box-Jenkins ARIMA is one of the most popular models in the prediction of financial time series data. Finding a similar pattern and combining the stock datas moving average value is how ARIMA works, which is slightly different to a traditional regression model. ARIMA was utilised to predict financial time series data by Timmermann and Granger [243]. However, ARIMA has shown limited system reliability [197]. Despite their success in solving time series prediction problems, traditional time series prediction models unfortunately have two major drawbacks: 1. in the stock market, stock index prices do not follow statistical assumptions, so as result of that some traditional times series models are not applicable to such datasets; 2. traditional time series methods require more historical data, and in order to describe daily observations, linguistic expressions are often utilised. Beside, stock data is well-known for its noises involutedly characteristics which would affect the prediction performance [56].

2.7 Computational Intelligence

The recent emergence of data mining and computational intelligence algorithms has tackled the computational demands on mathematical models to predict stock prices. The main focus of computational intelligence is designing and developing algorithms that have the capability to produce new knowledge and also to improve the performance of the algorithms [185]. Data minings main concern is to develop and implement algorithms for discovering the priori unknown relationships. Data mining was defined by Han and Kamber [112] as: The process of discovering interesting knowledge from large amounts of data stored in databases, data warehouses, or other information repositories. According to [21], discovering meaningful patterns, establishing effective models and rules, and the analysis of large quantities of data by automatic or semi-automatics means, are the main concepts of data mining. Recently, data mining has been widely used to discover patterns and to predict the future using historical data. Kovalerchuk and Vityaev (2005) stated in their study that data mining in general, and in finance specifically, is still more art than hard science [150].

Complexity and non-linearity are the main characteristics of the stock market price, thus it is considered to be a subtle and difficult system for humans to comprehend. This is why computational intelligence techniques have been introduced and extensively used in stock prediction. A number of studies have been carried out in this area. The recent studies on stock prediction can be categorised into roughly two types: 1. time series prediction [134] [32] [124] [126] [113]; and 2.trend prediction [245] [148] [125] [285] [57]. In time series prediction, the models are trained to fit the historical prices series of an individual stock index and are used to predict the future prices. However, trend prediction models are trained to obtain the relation between various technical variables and the (rise and decline) movement of stock prices. Generally, in modelling stock trend prediction, a wide range of data mining algorithms have been introduced including neural networks [245] [246] [156], support vector machines [159], logistic regressions [240] [84] [246], decision trees (DT) [246] [220], and naive Bayes [297]. In the next subsection, many examples of predictions of stock prices and stock trend directions are presented.

2.7.1 Different Single Prediction Models on Stock Market

In time series modelling there are a number of different approaches. Traditional statistical models are linear in their prediction of future value [23] [118] [217]. However, extensive research has introduced and utilised AI techniques such as Artificial Neural Networks (ANNs), Fuzzy Logic, Genetic Algorithms (GAs) and many other techniques; the results showed the prediction capability of such techniques [107] [160] [287]. Initially, theories of model building categorised stock index prediction models into two categories. The first category is based on statistical theories such as GARCH and stochastic volatility (SV) [78]. The second Category is based on artificial intelligence such as artificial neural networks (ANNs) [47], and the support vector machine (SVM) [129]. The most utilised approaches in many machine learning prediction algorithms for predicting stock price and stock index values are the Artificial Neural Network (ANN) and Support Vector Regression (SVR). However, learning the patterns is not the same in each algorithm. Artificial neural network (ANN) [72] [143] [208] [296], and other statistical prediction methods [48] [123] [132] attempt to predict stock prices under different circumstances such as market atmosphere or economic conditions. Artificial neural networks are one of the most popular prediction techniques. This technique simulates the same learning process as the human biological neural network, by developing models from extremely complex and non-linear formulas. It is employed to perform better prediction and analysis output and different parameters are used to train the neural network. In stock price prediction, the ANN model has achieved fruitful results. ANN was introduced for the first time by [266] to detect unknown regularities in stock price changes. San Diego's stock market was predicted by [100] where 63% accuracy was obtained for stocks rising and 74.4% for stocks falling. An efficient accuracy was generated by [230] with an ANN model for predicting the European stock market between 1991 and 1997. Shachmurove and Witkowska [223] point out, after comparing the results from the ordinary least squares method (OLS) and ANN, that ANN is a better prediction tool. The Indian stock market was predicted by an ANN model, and the root mean square error (RMSE) and mean absolute error (MSE) were used to evaluate the model; they showed a small margin of error [79]. Also, Zhu et al. [296] stated that the ANN prediction model performs better when predicting NASDAQ, DJIA and STI. According to [286] ANN in stock prediction is the most efficient technique. Several studies have indicated that ANN outperformed statistical regression models [218] and discriminant analysis [282]. There are two main methodologies for utilising Artificial Intelligent (AI) techniques for modelling stock prediction [292]. The first methodology uses past data to predict future stock prices by considering stock price variations as time series. Artificial Intelligent techniques are used in this approach as predictors [41] [46] [202]. However, these prediction models face a major drawback that due to the enormous noise and high dimensionality of stock price data. Therefore, among all existing prediction models, none can deliver a satisfactory performance, as the studies of [286] and [180] have observed. The second methodology considers technical indices and qualitative factors for predicting stock market and trend analysis. The purpose of merging technical indicators, for stock price prediction, is to permit the exploitation of tolerance for imprecision, uncertainty and partial truth in order to achieve tractability, robustness and low cost solutions. Many attempts have been made to predict financial markets, ranging from traditional time series approaches to Artificial Intelligent (AI) techniques [286], including ARCH-GARCH models [82], ANNs and other evolutionary computation methods [51] [284] [57] [50].

Technical analysts point out that the recent price reflects all the important information about the stock, therefore future prices can easily be predicted if the trends in the movement are observed. Moreover, there are many macro-economical factors that influence the movement of the stock market such as political events, firms policies, general economic conditions, the commodity price index, the bank rate, the bank exchange rate, investors expectations, institutional investors choices, the movement of other stock markets, the psychology of investors, etc.[184]. The calculation of the stock indices value is based on high stocks market capitalisation. In addition, statistical information can be gained from stock prices value by various technical parameters. Hence stock indices are derived from high market capitalisation and stocks prices, and an overall picture of the economy can be given by stock indices, taking into account the above-mentioned macro-economic factors. Ten data mining techniques were used in [196] to predict the Hang Seng index price movement.

In recent years, Support Vector Machine (SVM) has received huge attention and popularity from researchers, and it has also been considered a state-of-the-art technique in both regression and classification. The Support Vector Machine has great potential and superior performance in practical applications. Moreover, it can be regarded as a statistical learning theory algorithm with the principles of risk minimisation structure, which have been adapted by SVM. Therefore, this unique capability has drawn the attention of investors as well as researchers with the aim of utilising such a technique in financial time series prediction. As a result of the introduction of the insensitivity loss function SVM by Vapnik [249], the regression model SVM become known as the Support Vector Regression. Hence SVR solves many problems regarding non-linear estimation and financial time series prediction, and the attention paid to utilising such a model has also increased dramatically. There are many studies on using SVM and SVR to predict financial time series. Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), K-nearest neighbour classification, Nave Bayes based on kernel estimation, Logit model, Tree-based classification, neural networks, Bayesian classification with Gaussian process, Support Vector Machine (SVM) and Least Squares Support Vector Machine (LS-SVM) were included in the approaches in [196]. SVM and LS-SVM prediction results outperformed the other proposed approaches.

Technical indicators such as moving average, RSI, CCI, MACD, etc. were used to predict the movement direction of the Tehran exchange price index. In this study, the effectiveness of employing technical indicators for prediction of the movement of the index price was evaluated [7]. SVM was used in Kim's (2003)[144] model to predict the direction of the daily stock price change of the Korean composite stock price index (KOSPI). 12 technical indicators were selected as initial attributes; stochastic K%, stochastic D%, stochastic slow D%, momentum, ROC, Williams%R, A/D oscillator, disparity5, disparity10, OSCP, CCI and RSI. The movement directions of S&P CNX NIFTY indices were predicted by [153] using SVM and random forest, and they were compared with the result of traditional and logit models and ANN. The same technical indicators that were used by [144] have been used in this study. The SVM prediction result outperformed random forest, neural network and traditional models. NIKKEI225 index predictability was investigated with SVM. Thus, linear discriminant analysis, quadratic discriminant analysis and Elman backpropagation neural networks were compared with SVM. However, the experiment result showed that SVM outperformed the other classification methods. The usefulness of ARIMA, ANN, SVM and random forest regression models were investigated by [154] to predict the S&P CNX NIFTY index. Statistical and financial measurements were used via a trading experiment to assess the performance of the three non-linear models and the linear model. In their study, the SVM model outperformed the other model as the empirical results suggested. Hus et al.(2009) [122] developed a stock price prediction model by integrating two stage architectures, a self-organisation map and a support vector regression to examine seven major stock indices. The empirical result of the proposed two-stage architectural model indicated an alternative promising stock price prediction model.

2.7.2 The Development of Prediction Models: Multi or Hybrid Approaches?

There has been a failure to utilise single artificial techniques to capture the non-stationary property and accurately describe the moving tendencies that exist in financial time series. This is due to its fluctuation and the dynamic changes in the relationship between independent and dependent variables. Such fluctuation and structural changes, which are often caused by political events, economic conditions, traders expectations and other environmental factors, are important characteristics of financial time series. Therefore, utilising a hybrid model or combining several models has become a common practice in order to improve the prediction accuracy. According to M-competition, combining more than one model often leads to enhancing the prediction performance [176].

Various models and theories have been implemented by researchers in order to improve the prediction performance. Different techniques have been incorporated into single machine learning algorithms. Zhang and Wu (2009) [293] integrated a back propagation neural network (BPNN) with an improved Bacterial Chemotaxis Optimization (IBCO). Another method was proposed combining data preprocessing methods, genetic algorithms and the Levenberg-Marquardt (LM) algorithm in learning BPNN by [14]. In their studies, data transformation and the selection of input variables was used under data preprocessing in order to improve the models prediction performance. Moreover, the obtained result has proved that the model is capable of dealing with data fluctuations as well as yielding a better prediction performance. In addition, a hybrid artificial intelligence model was proposed by [106], combining genetic algorithms with a feed forward neural network to predict the stock exchange index. The latest research indicates that, especially in the short term, prediction models based on artificial intelligence techniques outperform traditional statistical based models [116]. However, and also as a result of the dramatic move in stock indices in response to many complex factors, there is plenty of room for enhancing intelligent prediction models. Therefore, researchers have introduced and tried to combine and optimise different algorithms, and they have taken the initiative to build new hybrids models in order to increase prediction speed and accuracy.

There are many example of these attempts. Armano et al.[11] optimised GA with ANN to predict stock indices; SVM were also combined with PSO in order to carry out the prediction of the stock index by Shen and Zhang [227]. Kazem's (2013) [140] prediction model, chaotic mapping, firefly algorithm and support vector regression (SVR) was proposed to predict stock market prices. In their study, the SVR-CFA model was introduced for the first time and the results were compared with SVR-GA (Genetic Algorithm), SVR-CGA (Chaotic Genetic Algorithm), SVR-FA (Firefly Algorithm) and the ANN and ANFIS model, whereas the new adopted model (SVR-CFA) outperformed the other compared models. The Seasonal Support Vector Regression (SSVR) model was developed by [199] to predict seasonal time series data. In order to determine their parameters, a hybrid genetic algorithm and tabu search (GA,TS) algorithms were implemented. And also on the same data sets, Sessional Auto-regressive Integrated Moving Average (SARIMA) and SVR were used for prediction, however the empirical results based on prediction accuracy indicated that the proposed model SSVR outperformed both SVR and SARIMA. A novel hybrid model to predict future evolutions of various stock indices was developed by integrating a genetic algorithm based on optimal time scale feature extractions with support vector machines. Neural networks, pure SVMs and traditional GARCH models were used as a benchmark and prediction performances were compared. The proposed hybrid prediction performance models were the best. Root mean squared error (RMSE) is one of the main utilised prediction models for performance measurement, however the reduction in this standard statistical measurement was significantly high.

According to [54] using ensemble learning algorithms to predict financial time series can powerfully improve the prediction performance of the base learner.

Generally, the financial time series prediction process, and precisely stock market prediction, is considered to be a challenging task since the nature of the stock market is essentially dynamic, non-linear, complicated, non-parametric and chaotic [3]. In financial time series prediction modelling, SVM and ANN have been used successfully. However, many studies have indicated that such tools had some limitations in learning the patterns as a result of stock market datas tremendous characteristics, such as noise, dimensional complexity and being non-stationary. Consequently, on such data characteristics, ANN often exhibited an inconsistent and unpredictable performance. Therefore, it is quite difficult to predict stock price movements directions [144] [145] [153]. Moreover, this study used a back propagation neural network and case-based reasoning (CBR) to exam the feasibility of the proposed model for financial prediction, whereas the experimental result indicated that SVM showed a promising alternative for prediction and outperformed BPNN and CBR.

Various prediction models to predict index direction movement were examined by [163] based on multivariate classification techniques, where parametric and non-parametric models were used and compared. In their study, classification models (discriminant analysis, logit, probit and probabilistic neural networks) outperformed the level estimation models (adaptive exponential smoothing, vector auto-regression with Kalman filter updating, multivariate transfer function and multi-layered feed forward neural network) as suggested by empirical experimentation for predicting the stock market movement direction and maximising investment return. The Taiwan stock exchange index was predicted by [46]. To predict the index direction, a probabilistic neural network (PNN) was used. Generalised methods of moments (GMM) with a Kalman filter and random walk were compared with PNNs statistical performance. The PNN empirical result indicated a stronger prediction power than the rest of the compared methods. Another input for prediction of the stock ISE100 index price movement used technical indicators such as MA, momentum, RSI, stochastics K%, and moving average convergence-divergence (MACD) [75]. In this study a model neural network was used and it achieved a 60.81% prediction rate for the ISE100 index.

Through time, several techniques have been employed and developed to predict stock trends. Classical regression methods were used initially. However, as a result, stock data has been categorised as a non-stationary time series data, and other non-linear machine learning techniques have been introduced. The hybridisation of soft computing techniques to predict stock market and trend analysis was introduced by [1]. In their study, the NASDAQ-100 index of the NASDAQ stock market was predicted for one day a head by a neural network and a neuro-fuzzy system. The trend prediction results of the proposed hybrid model were promising. The training ensemble model, therefore, represents a single hypothesis. It is not necessary in this approach to be contained within the approaches space of the models from which it is built. Such an approach has shown a flexibility in the functions it can represent. Therefore, the flexibility of the theory enabled them to over fit the training data more than a single model. However, in practice, other ensemble techniques, especially bagging, tend to reduce over-fitting problems of the training data. Prediction performance was investigated by [246] utilising ensemble methods to analyse stock. Two methods were considered in their study: majority voting and bagging. Additionally, the performances were compared of two types of classifier ensemble and a single classifier (neural network, decision trees, and logistic regression). The experiment result, based on prediction accuracy, indicated that the multiple classifier outperformed the single classifier. A new financial distress prediction (EDP) method was proposed by [232] using an SVM ensemble. The SVM ensemble outperformed the individual SVM classifier.

In data mining, preprocessing steps and feature selection have the ability to filter redundant and/or irrelevant features [245]. The results of feature selection in simpler models are easier interpretation, and faster induction and structural knowledge [52]. However, identifying more representative features and improving stock prediction are challenging issues, and many studies have claimed that in stock prediction modelling, verifying the feature selection is the success key to this process [245]. In stock prediction models the most common and adopted feature selection algorithms include: stepwise regression analysis (SRS), principle component analysis (PCA), decision tree (DT), and information gain [245] [240] [84]. All of these feature selection algorithms have only one ability which is revealing the underlying correlations/associations; they do not have the ability to determine the direct influence of stock features on stock price. Moreover, other analysis methods have been introduced by researchers to enhance the prediction accuracy. Different methods from the signal processing area were combined with AI techniques and the results were promising. Wavelet Transformation (WT) was combined with a back propagation neural network by [119]. In their study, the experiment result indicated that the proposed method outperformed the other used method. A novel hybrid model to predict Shanghai securities index by [257] integrated Empirical Mode Decomposition (EMD) with Support Vector Regression. The result from the proposed method indicated that combining EMD with SVR enhanced the prediction performance.

2.7.3 Surveying Stock Market Prediction Techniques

Generally, stock market prediction focus on building approaches in order to achieve best results using the available input data. However the successful key in stock market prediction is developing the least complex model. As mentioned in the previous section and subsection, modelling stock price prediction has seen many changes and different methods have also been introduced and integrated. In the literature, it has been verified that no single method or model works well in all situations. Wolpert [268] stated that there is no single computational view that solves all problems.

This subsection review the related scientific articles that focused on artificial intelligence and data mining techniques derived and applied to predict stock market. More than 60 related scientific articles applied to stock market prediction have been reviewed. The results of this survey are presented in two summary tables. The first table lists input variables to the stock market model. The second table lists a summary of the prediction models, data preprocessing, sample size and the models performance measures for each paper. Building a cohesive presentation of the recent implemented techniques on stock prediction modelling is the purpose of this survey, which will help to draw a conclusion for this chapter.

2.7.3.1 Input Variables

The utilised number of variables in each model differs. Generally the average number of input variables is between four and ten; however, there are some cases where only the daily closing price are used as an input variables [42] [281] [49] [213] [167]. However, [70] [137] use 15 and 10 input variables, respectively. Tables 2.1 and 2.2 summarise the surveyed articles, and thus focus on the input variable choices for each article. In the literature there are specific techniques that are widely utilised in choosing the most important input variables for the forecasting process among a large number of candidates, based on how each input affects the results. Thus, the most commonly used inputs are the stock index opening or closing price, as well as the daily highest and lowest values, supporting the statement that soft computing methods use quite simple input data to provide predictions.

About 40% of the surveyed articles use data stock or index prices as input; that is, the daily closing price or some indicator depending on it.

Table 2.1: Input variable choices

Article	Input Variables
[70]	Fifteen Technical indicators and eleven fundamental variables to predict PETR4
[225]	Twelve technical indicators to predict Shanghai Composite Indices
[137]	Ten technical indicators to predict Istanbul Stock Exchange (ISE) National 100 index
[119]	Fourteen Technical indicators to predict (DJIA, FTSE100, NIKKEI225,TAIEX)
[291]	Fundamental and momentum variables to predict Shanghai Stock Exchanges
[42]	Daily closing price to predict Ten different stocks price
[120]	Sixteen Technical Indicators to predict TAIEX
[281]	Daily closing price to predict TAIEX
[204]	Sixteen financial economical indexes to predict 200 company stock price listed in KOSPI200
[283]	Daily closing price to predict Shanghai Composite Index
[49]	Daily closing price to predict TAIEX, DJIA, S and P500 and IBOVESPA
[129]	Eight Macroeconomic input variables to predict NIKKEI225 index
[198]	Daily closing price to predict Ten stocks
[213]	Dow Jones daily return time series to predict a return sign (up or down)
[167]	Daily closing price of all Stander and poor 500 stocks
[279]	Five technical indicators to predict daily stock movement data from NZX
[257]	Daily closing price to predict Shanghai securities index
[277]	Daily closing price to predict Shanghai and Shenzhen CSI300 index
[4]	Technical and fundamental variable to predict Dell and Nokia daily stock price
[241]	Different number of input to predict the return of FTSE100
[228]	Daily closing price to predict Petro China
[159]	Twenty Nine Technical indices to predict NASDAQ direction movement
[6]	Predict 870 companies daily closing price from kuala lumpur stock exchange
[71]	Macroeconomic and Technical indecatoer to predict PETR4
[53]	Thirteen Technical indicators to predict TAIEX
[258]	Internal factor(daily closing price) external factor(USDHKD, USKRW and stander and poor500) to predict the direction of KOSPI200 and HSI index
[125]	Twenty three Technical indicators to predict the trend of Taiwan and Korea stock markets
[67]	Three future index contracts of NIKKEI225 (SGX-DT, OSE and CME) to predict future price and four technical indicators to predict Shanghai B-Share stock index closing price
[77]	Daily closing price to predict A300 and NASDAQ
[172]	Daily closing price to predict Shanghai B-Share stock index
[136]	NIKKEI225 future contracts were used to predict NIKKEI225 closing cash index and four technical indicators to predict SSEC index closing price
[145]	Thirteen technical indicators to predict the direction of KOSPI
[144]	Twelve Technical indicators to predict the direction of KOSPI
[173]	NIKKEI225 index futures prices are used to forecast NIKKEI opining cash index TAIEX futures prices and seven technical indicators are used to forecast TAIEX closing cash index
[263]	Ten technical indicators to predict TAIEX index
[245]	The fundamental and macroeconomic indexes are used to predict the listed electronic corporations stocks which are published by TSE(Taiwan stock exchange)

Table 2.2: (continued for Table 2.1)Input variable choices

Article	Input Variables
[212]	Daily closing price to predict the Mexican Stock Exchange
[275]	S&P500 return to predict the direction of return
[22]	Technical indicators are used to predict the direction of four main Indian indices; BSE, IBM, RIL and Oracle
[253]	Daily closing price to predict SSE, HS300, DJIA ans standard and poor500
[108]	Fundamental and technical indicators to predict DAX stock price
[76]	Monthly closing price to predict Down Jones index
	Daily closing price to predict IBM stock price
[105]	Daily stock exchange rate to predict NASDAQ
[206]	Ten technical indicators to predict the value of CNX Nifty and Standard and poor BSE
[205]	Ten technical indicators to predict the trend of CNX Nifty, standard and poor BSE Sensex indices and two stocks Reliance industries and Infosys Ltd
[293]	Six technical indicators to predict the value of Standard and poor500
[38]	Technical indicators to predict stocks of Citigroup and Motors Liquidation company
[135]	Daily closing price to predict Composite index of China (SSE), Bovespa index of Brazil, Dow Jones index US (DJ) and Nikkei225 index of Japan
[294]	Daily closing price to predict Shanghai market index and Doe Jones index
[155]	The daily stock price to predict thirty six companies in NYSE and NASDAQ stocks
[124]	Daily closing price to predict the next day of Taiwan index futures (FITX) price index
[264]	Daily closing price to predict TAIEX index
[254]	The monthly closing price to predict the trend of SZIL from China and monthly opening price to predict the trend of DJIAI from the US
[107]	Open, high, low and close price to preict next day closing price of IBM and Dell corporation IT sector and British and Ryanair airlines from airline sector
[56]	Daily closing price to predict Taiwan stock exchange capitalisation weighted stock index TAIEX
[251]	Monthly closing price to predict SCI index price. the data collected from Shanghai stock exchange
[22]	Daily closing price and two technical indicators to predict next day trend of BSE,IBM stock market, Reliance stock market(RIL) and Oracle stock market
[226]	Twelve technical indicators to predict the trend of Shanghai Composite Indices
[252]	Daily closing price to predict Shanghai Composite Index(SCI)

The daily closing price are used by [42], [281], [283], [49], [198] [167], [257], [277], [228], [172], [212], [253], [105], [135], [294], [155], [124], [56], [22] and [252]. About 45% of the surveyed articles use as inputs technical analysis factors that are sometimes combined with daily or previous stock index prices, as in [70], [225], [137], [119], [120], [279], [159], [53], [125], [145], [144], [173], [263], [22], [206], [205], [293], [38] and [226]. The technical analysis factors range from 2 to 30, with most articles using mostly a combination of all previously described variables, and also fundamental analysis indicators and statistical data.

2.7.3.2 Prediction Methodology

Each surveyed paper is classified with respect to data preprocessing, sample size, type of implemented prediction model and performance measurements.

Table 2.3: Comparisons of the related work

Article	Data preprocessing	Sample size	Prediction model model	Performance measures
[70]	[0.2,0.8]	M:11years	ANN	MAPE,RMSE THEIL
[225]	-	D:3MONTHS	ANFSA+RBFNN ARIMA BP,SVM RBF,GA+RBF,PSO+RBF	Forecastratio Errorratio
[137]	[-1,0,1]	2733	SVM,BP,OLS	t-Test prediction performance%
[119]	WT/SRCS	D:6years	BPNN,ABCRNN	RMSE
[291]	[0,1] PCA,DT,LASSO	8347	Baselin models	Precision% Accuracy%
[42]	-	5229	Hybrid C&RT-ANN	Accuracy%
[120]	[0,1]	3540	Hybrid SOM-GA	MSE,RMSE MAPE,MAE,APE
[281]	-	3years	SKSVR,MKSVR, ARIMA,FNN	RMSE
[204]	[-1,1]	403	SSL,SVN,ANN,BH	ROI
[283]	EMD	2years	EMD-SVM	Absolute error, MSE
[49]	-	-	FUZZY TIME SERIES-KSVM,GARCH, GIR-GARCH, FUZZY-GARCH	RMSE
[129]	[-1,1]	676	RW,LDA,QDA,EBNN,SVM	Hit ratio%
[198]	-	2Months	ARIMA,SVM, Hybrid-ARIMA-SVM	MAE,MAPE, MSE,RMSE
[213]	Log	1024	ANN.KNN,DT,Ensembles	ERROR RATE%
[167]	Log-PCA	1100	ESN,BPNN,Elman,RBF, ESN+PCA	APE
[279]	[-1,0,1]	1665	Rough set	-
[257]	EMD	220	SVR, EMD-SVR	RMSE,AME,AMRE

Table 2.4: (continued for Table 2.3)Comparisons of the related work

Article	Data preprocessing	Sample size	Prediction model model	Performance measures
[172]	-	1000	MARS,SVR, BPNN,MLR	RMSE,MAD MAPE,DA
[136]	NLICA	1000	SVR,PC-SVR NLICA-SVR LICA-SVR	RMSE,MAD, MAPE,RMSPE DS
[145]	[0,1]	2928	BPLT,GALT, GAFD	HIT RATIO% P-value
[144]	[0,1]	2928	SVM,BPNN	HIT RATIO% P-value
[173]	ICA	1144	SVR,ICA-SVR RANDOM WALK	RMSE,NMSE,DS, MAD,CD,CP
[263]	-	1year	ANFIS	RMSE
[245]	[0,1]	7years	PCA-MLPNN, GA-MLPNN, MLPNN, CART-MLPNN	CONFUSION MATRIX T-test
[212]	-	800	PSO-Ensemble-NN	MSE,T,P-value
[275]	Return rate	13827	AR,NN	MSE
[22]	[0,1]	905	MLP,RBFNN, ANFIS,FLANN, FLANN,DNN LLRBFNN,LLWNN	RMSE, AMAPE, MAPE
[253]	[0,1]	SSE1513, S&P 1539 DJIA 1532	PCA-STNN	MAE, RMSE MAPE
[108]	[0,1]	40years	GA-NN,GR-NN, RBE,BNNMAS	MAPE
[76]	-	157	ANN,ADANN	SMAPE
[105]	[-1,1]	146	MLP,HybridANN, DAN2	MSE,MAD MAD%
[206]	-	10years	ANN,SVR, RANDOM FOREST, SVR-RANDOM FOREST	MAPE,MAE RMSE MSE
[205]	[-1,+1]	10years	ANN,SVM, Naive-bayes, Random Forest	Accuracy% F-measure
[293]	-	2350	IBCO,BPNN, IBCO-BPNN	 MSE
[38]	[0,1]	378	EPCNN,BPN,TSK	MAPE
[135]	Log/WT/MARS	1000	WT-SVR,SVR,ANFIS WT-MARS,ARIMA WT-MARS-SVR	RMSE, MAD,MAPE, RMSPE
[294]	PCA	1200	SVM,PSOSVM, LPP,NN,RBFN	 Accuracy Rate%

Table 2.5: (continued for Table 2.3)Comparisons of the related work

Article	Data preprocessing	Sample size	Prediction model model	Performance measures
[254]	[0,1]	SIZ216	ESM,BPNN,GA	MAE
		DJIAI240	GA-ESM-BPNN	RMSE,DA
			ARIMA	MAPE,ME
[107]	SRA	IBM491	ARIMA,ANN	
		DELL491	GA	
		BA594		MAPE
		RA471	CGF	
[56]	EMD	6years	EMD,SVR,	
			EMD-SVR,AR	RMSE
[251]	EMD	220	SVR,EMD,	AME,AMRE
			EMD-SVR	RMSE
[22]	[0,1]	905	DE,UKF	AMAPE
			DEUKF,RTRL	RMSE,MAPE
[226]	K-means	30groups	AFAS,K-means,	Eror ratio%
			ARIMA,SVM,GA,	
			PSO,RBFNN	Forecast ratio%
[252]	[0,1]	204	WT,WD,	MAE
			WD-BPNN,	RMSE
			BPNN	MAPE
[277]	EMD	2years	SVM,EMD-SVM	MAPE,NRMSE
[4]	[0,1]	-	ANN	Accuracy%
[241]	[0,1]	1260	GA-MLP,GA-SVM	Accuracy%
[228]	-	320	ARMA,BPNN,	
			ARM-BPNN	ERROR RAIT%
[159]	F-score,SSFS	1065	SVM,BPNN	Accuracy%
				T-test,P-value
[6]	3phase clustering	870	SSE	QGR,RATIO
[71]	[0,1]	2384	MLPANN	RMSE,MRPE
[53]	-	9years	SVR,Fuzzy time series-SVR	RMSE
[258]	PCA	10years	SVM,ANN,RW	HIT RATIO%
[125]	Wrapper	-	SVM,KNN,BP,DT,LR	Accuracy%
[67]	ICA	1000	BPNN,NLICA-BPNN	RMSE,MAD
				DS,MAPE
				RMSPE
[77]	WT	560	WT-BP,BP	APE,MAPE
			WT-ARIMA,ARIMA	RMSE
[155]	-	14years	Ensemblr,ANN,GA	Accuracy%
[124]	[0,1]	1540	SVR,	MSE
			SOFM-SVR,	MAE
			SOFM	MAPE
[264]	-	1YEAR	ANFIS,AR	RMSE

The prediction performance can be affected by many factors; input data preprocessing and proper sampling are two of the main factors. Selecting and preprocessing the input data is considered to be a main step in eliminating redundant inputs. Input data in many cases has a large range of value reducing the effectiveness of the training procedures. Therefore, a normalisation step is suggested in order to tackle this issue. Data normalisation and scaling techniques 21 studies. In addition, other techniques were used and introduced to pre-process the input data, such as (PCA) Principal Component Analysis, (WT) Wavelet Transform, (ICA) Independent Component Analysis, and (EMD) Empirical Mode Decomposition. However, it is stated that not all articles provide details about data pre- processing, or whether any pre-processing occurs. Hence, all articles referring to data pre-processing are a crucial and necessary step. The chosen sample size by most authors is daily data. Tables 2.4, 2.3 and 2.5 summarised the chosen sample size. Moreover, the tables also exhibited the specific models and techniques that have been used in the surveyed articles. The most commonly implemented techniques are SVM, BPNN, SVR, GA, AR, and ARIMA.

Tables Tables 2.4, 2.3 and 2.5 present a list of performance measures to evaluate each authors approach. Performance measures can be classified as statistical measures and non-statistical measures.

- Statistical measures included the Root Mean Square Error (RMSE), the Mean Absolute Error (MAE) and the Mean Squared Error (MSE), and statistical indicators such as autocorrelation, the correlation coefficient and standard deviation.

- Non-statistical performance measures include measures that are related to the economical side of prediction. Hit Rate is considered to be the most utilised measure; it measures the percentage of correct predictions of the model. Additionally, the other two measures that deal with the profitability of the model are the annual rate of return and the average annual profit of the model.

2.8 Summary

This chapter provided a review on financial time series prediction, including the popular efficient market hypothesis, and the early attempts at financial time series prediction. It presented two main constituents, fundamental and technical analysis. The literature review on current time series prediction methodologies and model evaluations is also given. Different examples from the literature of different computational intelligence techniques for financial time series prediction were presented, and also the development of modelling and predicting financial time series data was given. Moreover, this chapter has surveyed articles that have applied artificial intelligence techniques to predict financial time series (stock price). The survey has focused on input data, prediction methodology, and performance measures. The observation is that AI techniques are suitable for stock market prediction. Experiments demonstrate that soft computing techniques outperform conventional models in most cases. They deliver better results than a prediction system with higher predicting accuracy. However, difficulties arise when defining the structure of the models. The proposed methodology in this thesis will try to overcome problems such as model structures, data preprocessing, inability to catch dynamics and non-linearity of time series, and lack of understanding of data, while taking advantage of the explicit model specification and variable information of stock prices.

Chapter 3

Intelligent Financial Data Modelling Techniques

3.1 Introduction

In modern financial time series prediction, predicting stock prices has been regarded as one of the most challenging applications. Thus, since the beginning of the stock market, various models have been proposed to support investors with more precise predictions. However, stock prices are influenced by a different number of factors and non-linear relationships between factors existed in different periods, such that prediction of the value of stock prices or trends is considered to be an extremely difficult task for investors.

Additionally, researchers have proposed numerous conventional numerical prediction models. However, traditional statistical models such ARCH, GARCH, ARMA, ARIMA and AR have failed to capture the complexity and the behaviour of stock prices. George Box stated in his work that essentially, all models are wrong, but some are useful [24]. Therefore, researchers have introduced more advanced non-linear techniques including Support Vector Machine (SVM), Support Vector Regression (SVR), Neural Network (NN) and

Naive Bayes (NB). Furthermore, these techniques are employed in this thesis to predict the value of the stock index closing price and the direction movement. In this chapter, a single approach is demonstrated and implemented to predict financial data (stock index). Thus, two approaches are adopted in this chapter. First, dynamic models are designed to predict financial time series closing price by the above-mentioned techniques. Furthermore, the architecture and parameters for each technique are described in the next section and subsections. Finally, technical indicators will be used as an input to predict the direction of movement of the index price.

3.2 Financial Data

Financial data are characteristically non-linear, non-stationary, noisy, with a high degree of uncertainty and hidden relationships. These characteristics are reasons for information unavailability, affecting the behaviour of financial markets between past and future captured prices. In financial markets, predicting the stock index is considered to be an important tool for its participants. Thus, in order to guard against risk, investors rely on the stock index and government institutions to monitor the market fluctuations. Moreover, researchers refer to the stock index in their financial studies regarding issues such as pricing financial derivatives and portfolio selection. Therefore, researchers, in order to carry out accurate prediction, have tried different models and algorithms, and have achieved magnificent results.

The simple meaning of prediction is to understand which variables lead to predict other variables. This entails comprehension of the timing of lead-lag relations between many variables, understanding the statistical significance of these lead-lag relations, and learning which variables are the more important ones to watch as signals for predicting market movements. Thus, with increasing financial market volatility and internationalised capital flow, better prediction is a key element of successful financial decision making. In portfolio management by commerce and investors, accurate prediction methods are crucial. All over the world and every day, the stock market is a place where vast amount of capital are invested and traded. As a result of the emergence of online brokerages, trading stocks is becoming more accessible to the general public, and such changes mean that the influence of the stock market goes beyond macro-economic effects. More precisely, stock markets now influence the well-being of the general public.

The creation of a robust, intelligent system that can accurately predict the stock index has always been a subject of great interest for many investors and financial analysts. However, the stock market is dynamic and unpredictable [5]. There are many factors that stock market movement reacts to, such as political, economic and social [191] [109], so predicting the stock market is complex.

In recent years, the stock market has become a popular investment channel due to its high return of investment compared with other investment instruments. The prediction of the stock index has attracted not just investors but also private institutions. On the other hand, due to the inherently noisy and non-stationary nature of stock index prices, accurate predictions are considered a challenging task [3] [165]. There are many factors that affect stock index prices such as firms policies, political events, general economic conditions, commodity price indexes, interest and exchange rates and investors expectations and psychological factors. Investors consider that stock index prices are always the most important information. However, the nature of stock index prices is essentially dynamic, non-linear, non-parametric and chaotic. Therefore, investors must handle time series that are non-stationary and noisy with frequent structural breaks [193] [259]. As a result of its high potential profit, stock market financial time series prediction has received a huge focus from investors, speculators and researchers. The revolution and the advance in both analytical and computational methods have led to a number of interesting new approaches to financial time series mining, based on non-linear and non-stationary models [11].

Predicting the future and having the ability to achieve high accuracy is crucial to many decision processes in planning, organising, scheduling, purchasing, strategy formulation, policy making, supply-chain management and so on. As a result of this importance, the area of prediction has received much attention and magnificent efforts have been made in the past by researchers. However, it is still an active and an important research domain of human activity at the present time and will continue to be so in the future [288]. In financial prediction, stock index prices have always been a subject of interest for most investors and financial analysts. Nevertheless, making decisions about whether to buy or sell has remained a very difficult task for investors as there are many different factors that affect stock prices [40] [260]. In business and finance, stock price prediction has remained one of the main important topics. Moreover, stock market environmental characteristics are very complicated, dynamic, stochastic and thus difficult to predict [261] [278] [248] [221]. Resulting from a huge number of factors such economic, political or psychological, financial time series present complex behaviour. Therefore, this complex, dynamic and noisy nature of financial time series is the reason why modelling and predicting financial indices is a very interesting but challenging topic.

In this thesis works are divided into three phases. The FTSE 100 (UK), S&P 500 (USA) and NIKKEI 225 (Japan) are the real financial data sets used to demonstrate the presented models. Firstly, in this chapter a single approach and its results for two phases are presented. In the follow subsections and sections, more details about the data sets, the single approaches and prediction modelling procedures are explained.

3.2.1 Financial Times Series Data

According to Han et al. [112], a time series data set consists of sequences of numeric values obtained over repeated measurements of time. The values are typically measured at equal time intervals (e.g. every minute, hour or day). The main assumption that is considered to be a fundamental point of predicting time series is that a pattern existing in historical observations will continue to exist in the future. In other words, this implies that in order to predict the time series, there must be no structural breaks or regime switching in the time series. In this thesis, the financial time series that are used to validate the single approach proposed models in this chapter are the closing daily prices of three main world indices. The first data set is the FTSE 100 daily closing price from the London stock market. This data set covers the period from 03/01/1984 to 30/10/2014. Figure3.1 shows the time series plot. The second set is the Nikkei 225 daily index of

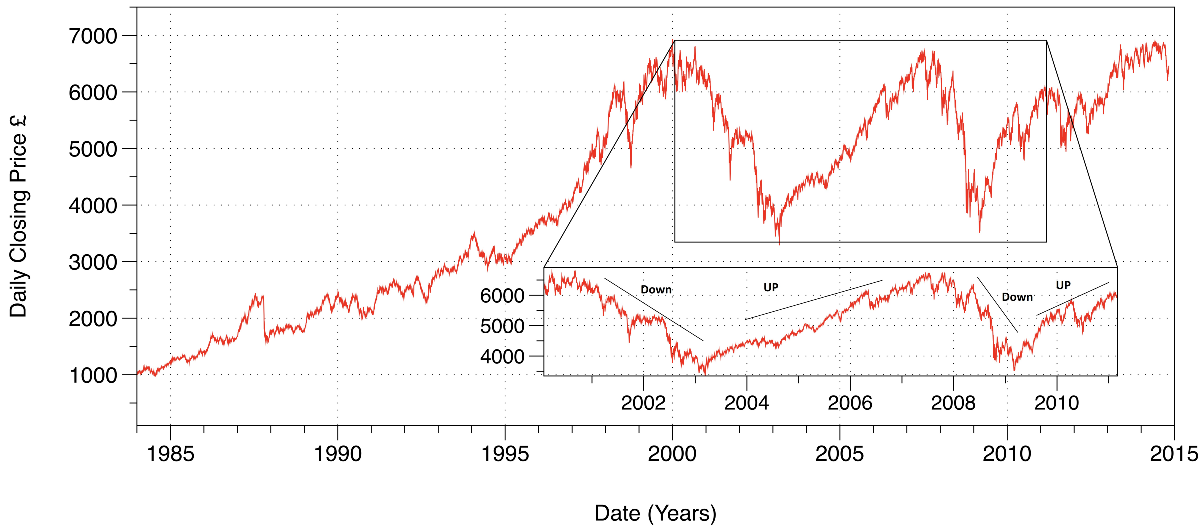

Figure 3.1: Daily Closing Price for the FTSE100 Index.

the closing price from the Japanese stock market. The data set covers the period from 04/01/1984 to 30/06/2015. Figure3.2 shows the time series plot.

The S&P500 (SPX) index of daily closing price from the USA stock market is the third data set. The data set covers the period from 03/01/1950 to 30/06/2015. Figure 3.3 shows the time series plot.

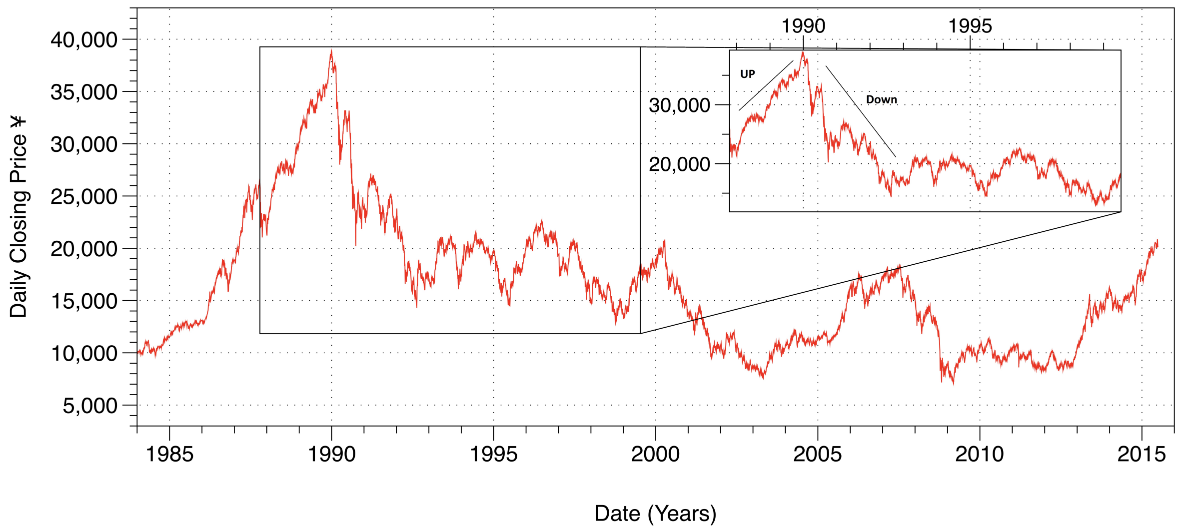

Figure 3.2: Daily Closing Price for the Nikkie 225 Index.

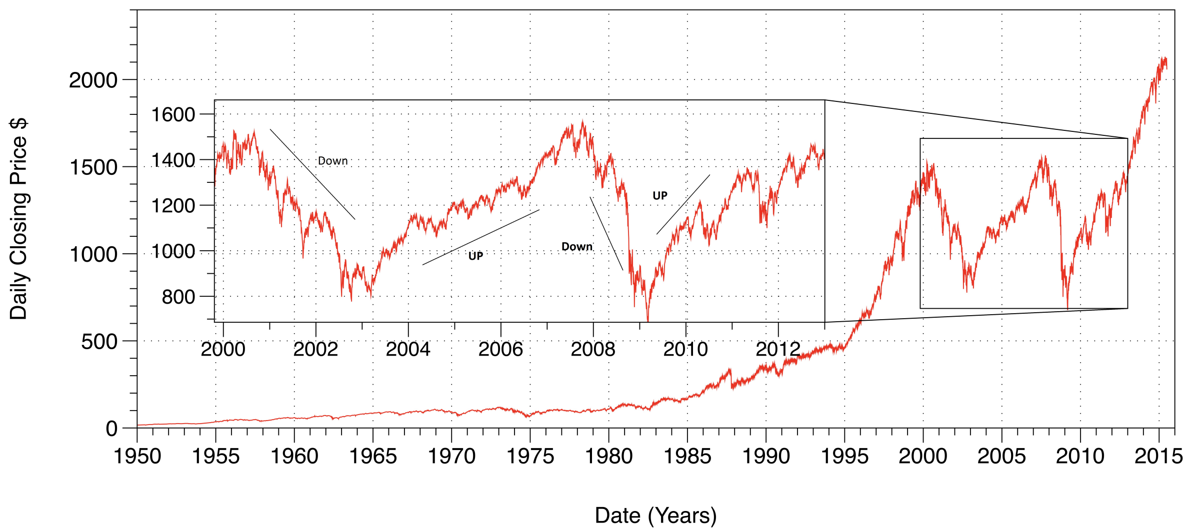

Figure 3.3: Daily Closing Price for the SPX500 Index.

3.2.2 Technical Indicators

Technical indicators are derived from technical analysis. In this chapter, the technical indicators are presented in order to be used as an input to test the proposed models of this thesis. Technical analysis in stock markets falls under the efficient market hypothesis (EMH), which asserts that free market information is efficient and states that stock prices are ultimately made freely. This can lead to the construction of an ambidextrous instrument to be used to predict the stock trends by studying the past trading data which is concerned with price and volume, rather than other external drivers such as financial statements, news and economic factors [147]. Thus, hidden relevant information on a relationship between past prices and trading volume can be reflected directly by technical analysis, and also a complicated price-based mechanism tends to repeat itself as a result of the assumption that investors collectively tend to engage in a patterned and recognisable behaviour. Therefore, the focus of the technicians is to identify patterns and trends in associated past prices and trading volume information, and thus analyse statistics that are generated from various marketing activities through various methods, in order to evaluate the performance of the stock indices [246]. Moreover, exploring the markets internal information and assuming that all of the necessary factors are hidden in the stock exchange information, are the main findings when utilising technical analysis [39]. In the light of these considerations, academics and practitioners have used technical analysis to predict the direction of stock prices based on historical stock index data.

Technical analysis, compared with fundamental analysis, is one of the most popular methods of future stock index price prediction [15]. Murphy summarised the three premises of technical analysis as follows [190]:

- Everything is discounted by market actions: the effect of supply and demand, which is the basis for all economic and fundamental analysis, can be reflected in the price and also every change in the market is ultimately reflected in the market price itself. The main concern of technical analysis in not studying the reasons for price action, it is focusing on the study of the price action instead.

- In the market, prices move in trends: this point is considered as a foundation of almost all technical systems attempts to analysis trends and trading in the direction of the trend. The trend in motion is more likely to continue than to reverse the underlying premise.

- The probability that history repeats itself: such assumptions have derived from the study of human psychology, which does not tend to change over time. However, such a view could lead to the identification of chart patterns, which are observed to recur over time, revealing traits of a bullish or bearish market psychology.

Technical analysis is intended to identify regularities in the form of time series of price and volume under the above principles by extracting non-linear patterns from trading noisy data [171]. Moreover, many attempts have been made by technicians to identify price patterns and market trends through by a number of techniques and tools that are based on price and volume transformations in the stock market, and to study those patterns, in which are included specifically technical indicators. Technical analysts believe in the assumption of collective repetition, whereby investors repeat the behaviour of the investors who preceded them.

Subjective judgements are used by some technical analysts in order to determine the optimal pattern, particularly that instrument which reflects a given time and what the interpretation of the pattern should be; Restated, or other which are employed strictly mechanical or systematic approaches such through technical in order to pattern identification and interpretation. Therefore, in technical analysis, the selection of technical indicators has become an interesting and important issue. Technical indicators have been widely used in prediction stock index price direction, however in order to select such indicators, different criteria should be followed, upon which prediction systems are highly dependent. These criteria can be summarised as the following points [233]:

- Availability: data should be easily obtained.

- The historical data base must be sufficient: in order to process the data, there must be enough samples for the testing system and machine learning.

- Indicators must be correlated to the price: in other words, the data should be somehow related and relevant to the price of the security (whether it is lagging, leading coincidental, or noise).

- Data must be in a periodic order: the availability of the data in a predictable frequency (daily, weekly, monthly, and annually) is a must in the data.
- Data must be reliable: as a result of globalisation, the fast changing pace of financial environments and the dramatic increase in financial market volatility, obtaining the right and reliable data is a very difficult process. Furthermore, enhanced and developed technical tools have been introduced with emphasis on computer-assisted techniques and specially designed computer software.

Table 3.1: Technical indicators used as input variables [22]

Technical indicators	**Explanation**
Stochastic oscillator (%K,%D)	Stochastic oscillator is a momentum indicator that uses support and resistance level. The term stochastic refers to the location of current price in relation to its price range over a period of time.
Momentum (MOME)	Momentum is the empirically observed tendency for rising asset prices to rise further, and falling prices to keep falling prices to keep falling.
Relative strength index (RSI)	It is intend to chart the current and historical strength or weakness of a stock or market based on the closing price of a recent trading period.
Williams %R (%R)	Williams %R is usually plotted using negative values. For the purpose of analysis and discussion, simply ignore the negative symbols. It is best to wait for the securitys price to change direction before placing your trades.
Moving average convergence and divergence (MACD)	MACD shows the difference between a fast and slow exponential moving average (EMA) of closing price. Fast means a short period average, and slow means a long period one.
Moving average (MA)	Moving averages are used to emphasize the direction of trend and smooth out price, volume fluctuations that can confuse interpretation.
Exponential Moving average (EMA)	EMA gives more weight to recent price and the longer the period of EMA the less total weight.The ability of picking up the changes on price is the. main advantage of EMA.

The above given criteria selections are the reasons that computer-assisted methods are selected and used by technical indicators, which are normally formulated to be applied to stock price prediction [102].In this chapter, technical indicators are used to predict the direction. Furthermore, the selection of the eight technical indicators as feature subsets was based on the related review of domain experts and prior researchers [39] [225] [235] [11] [75] [124] [145] [153] [280] [137] [206]. The utilised technical indicators are formed and illustrated in Table 3.1 and Table 3.2. Thus, in order to obtain these technical indicators, the daily closing, high and low price of the S&P 500 are utilised as Figures 3.3 and 3.4 illustrate. Therefore, after implementing the formulas in Table 3.2, the technical indicators for S&P 500 index are formed and illustrated in Figure 3.5. For the FTSE 100 also daily closing, high and low prices in Figures 3.1 and 3.6 are used to obtain the technical indicators which are formed and illustrated in Figure 3.7. Furthermore, the daily closing, high and low prices in Figures 3.2 for the Nikkei 225 are used to obtain the technical indicators for this index as Figure 3.9 illustrates.

Table 3.2: Technical indicators formulas [206]

Technical indicators	**Formulas**
Stochastic %K	$C_t - L_n/H_n - L_n \times 100$
Stochastic%D	$\sum_{i=0}^{n-1} \%K_{t-i}/n$
Momentum (MOME)	$C_t - C_{t-4}$
Relative strength index (RSI)	$100 - 100/1 + \sum_{i=0}^{n-1} UP_{t-i} \lvert n \rvert \sum_{i=0}^{n-1} Dw_{t-i}/n$
Moving average (MA5)	$\mathrm{MA}\,(m)_t = \left(\sum_{i=n}^{n} x_i \right) /m, m = 5$
Exponential moving average (EMA(5))	$EMA(m)_i = (1/m) \times \sum_{i=t}^{n} DI_i, m = 5$
William %R	$H_n - C_t/H_n - L_n \times 100$
Moving average convergence and divergence (MACD)	$MACD\,(5)_i = \frac{1}{5} \sum_{t=i-5}^{i} (EMA(5)_t - EMA(5)_t),$

C is the closing price, L the low price, H the high price, Up the upward price change, Dw the downward price change.

A statistical summary of the eight technical indicators was calculated and is presented in Table 3.3

Table 3.3: Summary statistics for the selected indicators

Index name	Indicators name	Max	Min	Mean	Standard deviation
FTSE100	Stochastic %K	100	0.02	56.97	31.04
	Stochastic %D	100	0.057	56.66	28.93
	Momentum (MOME)	603.20	-1264.90	7.34	147.79
	RSI	100	6.90	53.54	16.72
	Moving average (MA5)	6886.60	993.48	4168.82	1754.73
	EMA(5)	6892.23	996.60	4168.82	1754.51
	William %R	-0.032	-100	-42.88	30.79
	MACD	144.52	-318.28	4.76	47.78
S&P500	Stochastic %K	100	0.009	58.62	30.69
	Stochastic %D	100	0.087	57.60	29.27
	Momentum (MOME)	160.65	-309.96	1.39	22.56
	RSI	100	2.32	54.40	18.05
	Moving average (MA5)	2127.95	16.77	471.96	541.59
	EMA(5)	2125.65	16.79	471.96	541.57
	William %R	-0.01	-100	-42.87	30.54
	MACD	34.79	-77.20	0.88	7.46
Nikkei225	Stochastic %K	100	0.0056	55.51	32.25
	Stochastic %D	100	0.18	55.00	30.22
	Momentum (MOME)	3735	-4853	14.93	749.16
	RSI	100	4.73	-45.23	32.08
	Moving average (MA5)	38797.80	7177.69	16447.47	6315.76
	EMA(5)	38743.92	7171.75	16447.47	6314.15
	William %R	100	4.73	-45.23	32.08
	MACD	795.75	-1697.12	9.17	269.65

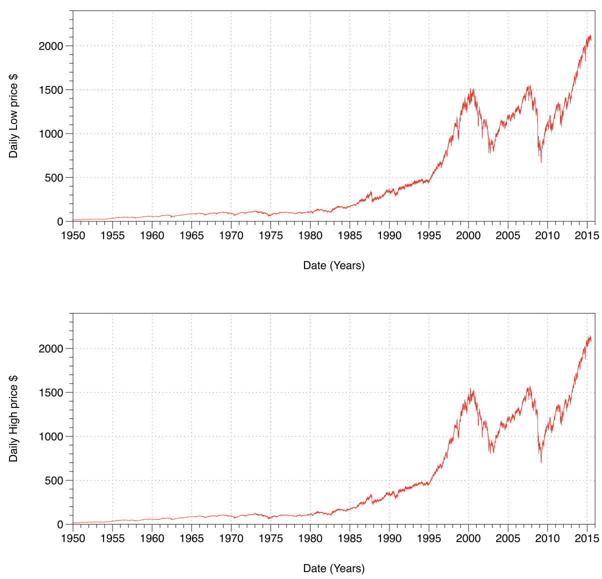

Figure 3.4: Daily High and Low Price of S&P 500.

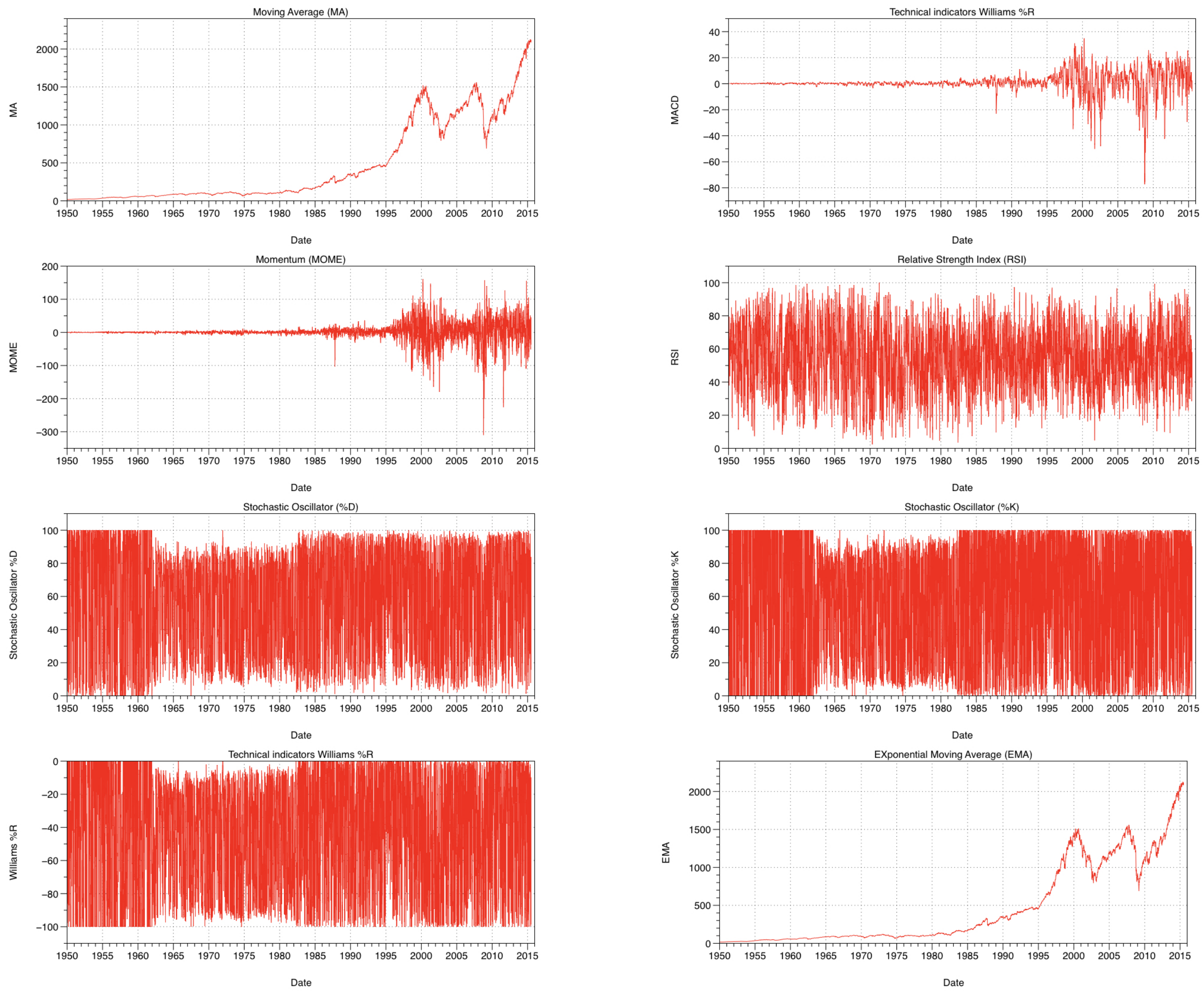

Figure 3.5: S&P 500 Technical indicators

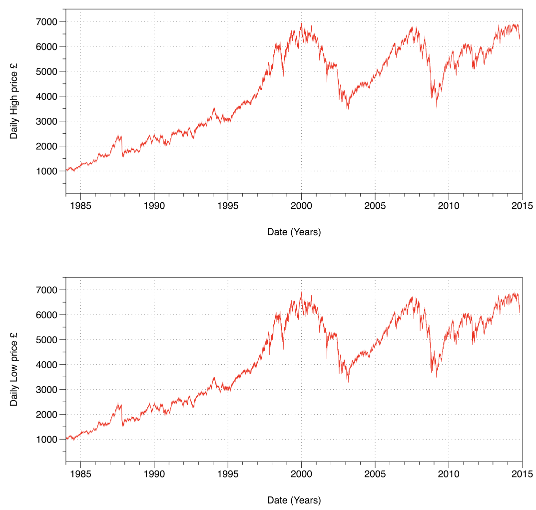

Figure 3.6: Daily High and Low Price of FTSE 100.

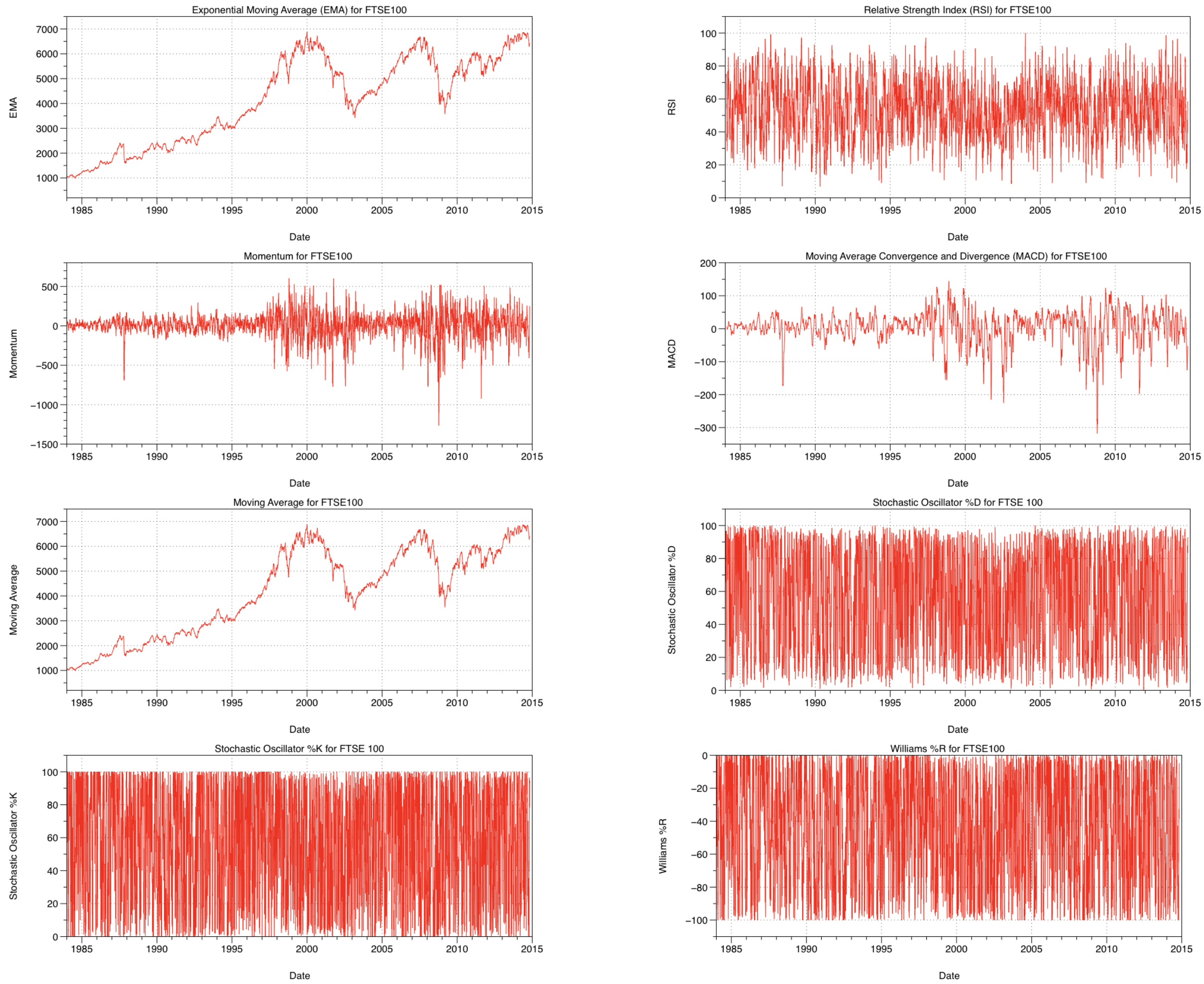

Figure 3.7: FTSE 100 Technical indicators

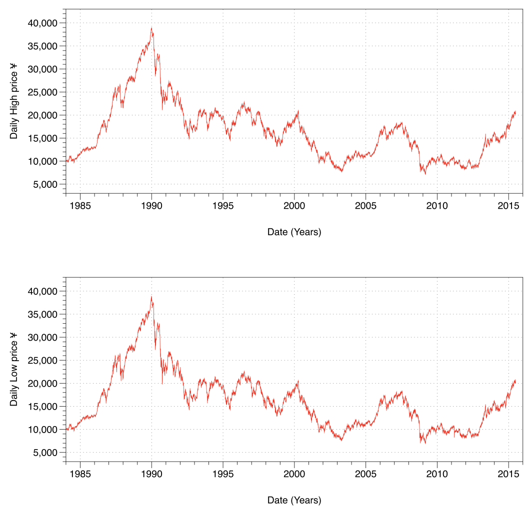

Figure 3.8: Daily High and Low Price of Nikkei 225.

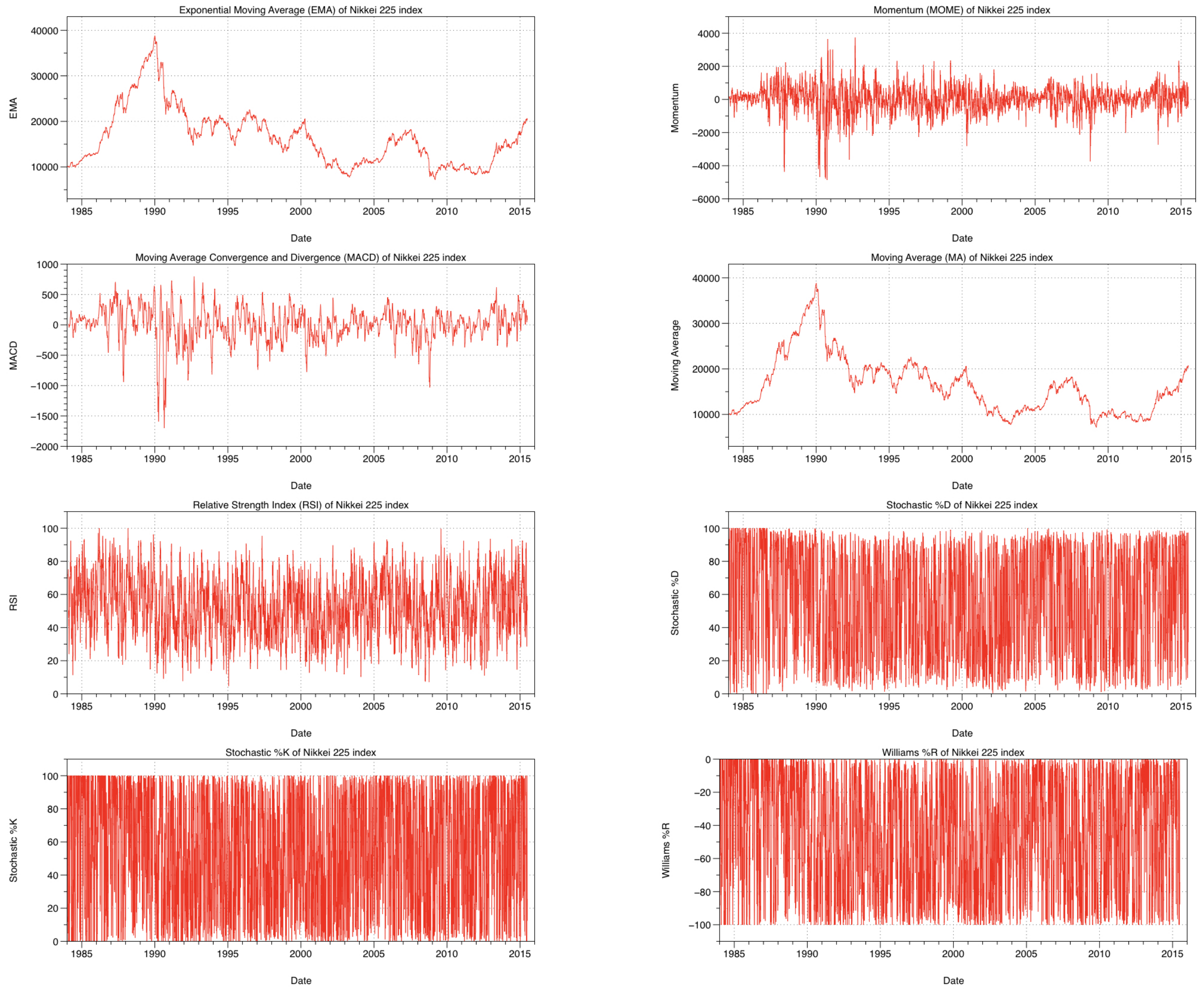

Figure 3.9: Nikkei 225 Technical indicators

3.3 Prediction Evaluation

3.3.1 Standard Statistical Measure

The accuracy of a prediction is referred to as goodness of fit, which means how well the prediction models fit a set of observations. Moreover, prediction accuracy can also be regarded as an optimists term for prediction errors, as is stated in [13]. Hence, prediction error represents the difference between the predicted value and the actual value. Therefore, the prediction error can be defied as: if x_n is the actual observations at time n and $\hat{x_n}$ is the predicted value for the same period, then the prediction error takes the following form:

$$E_n = x_n - \tilde{x_n} \tag{3.1}$$

Generally, in equation 3.1 $\hat{x_n}$ is one step ahead prediction so E_n is the one step ahead prediction error. The accuracy measurement can be classified under five categorises, according to [131]: scale-dependent measures, measures based on percentage errors, measures based on relative errors, relative measures and scaled error. However, in this research the most popular and commonly available statistical accuracy measures used are presented in Table 3.4:

Table 3.4: Statistical accuracy measures [131]

Name	**Category**
Mean Square Error (MSE)	Scale-dependent measures.
Root Mean Square Error (RMSE)	Scale-dependent measures.
Mean Absolute Error (MAE)	Scale-dependent measures.
Cross correlation coefficient (R)	Method of computing the degree of relationship between variable.
Standard Deviation (SD)	An statistical test used to observe variations.

MSE, RMSE and MAE are scale-dependent measures which are commonly utilised to compare different methods on the same data set. In addition to standard statistical measures, standard deviation is used to observe variations. Moreover, a well-known cross correlation method of computing the degree of relationship existing between the variables is also used in this study [253]. The Table 3.5 illustrates the utilised prediction accuracy measure formulas.

Table 3.5: Error metrics equation to measure prediction accuracy [253] [67]

Abbrev.	**Formulas**
MSE	$= \frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2$
RMSE	$= \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}$
MAE	$= \frac{1}{N}\sum_{i=t}^{N}\frac{\lvert x_i - \hat{x}_i\rvert}{\lvert x_i\rvert} =$
R	$= \frac{\sum_{i=1}^{N}(x_i - \bar{x_i})\left(\hat{x}_i - \hat{\bar{x}}_i\right)}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x_i})^2 \sum_{i=1}^{N}\left(\hat{x}_i - \hat{\bar{x}}_i\right)^2}}$
SD	$= \sqrt{\frac{\sum(x_i - \hat{x}_i)^2}{N-1}}$

These statistical analyses, after they are computed, will provide the required information regarding the prediction accuracy and will strengthen the conclusions. This is so for all five (MSE, RMSE, MAE, R and SD) for the retained statistical error, whereas the lower output is the better prediction accuracy of the model concerned. However, after considering the pros and cons of the above-mentioned statistical accuracy measures, MSE and RMSE are chosen for performance comparison on the same data set, and for comparison across data sets, MAE is utilised in this research [32].

3.3.2 In-sample out-of-sample Test

In the given prediction method, in-sample testing is used, utilising all of the available historical data in order to fit a model of interest, which is likely to lead to understating the prediction error, thus overestimating its predictability. The selection and estimation processes with in-sample testing are designed to calibrate the prediction procedure to the historical data, and thus the pattern of these data may not persist into the future.

In addition, the selected models by in-sample fit might not be the best to predict out-of-sample data. Therefore, and in order to assess the goodness-of-fit of the proposed method, the predictors used out-of-sample tests. The evaluation of out-of-sample prediction accuracy begins with the division of the historical data series into a training period (in-sample data) and a test period (out-of- sample data sets). The training period is utilised to identify and estimate a model while the test period is reserved to assess the models prediction accuracy [236].

3.4 Prediction Procedures of Financial Data

Generally this research includes three phases, and this chapter presents the first stage of phases 1 and 2. The holistic framework of the proposed prediction models is shown in Figure 3.10. The phase 1 model predicts the next day stock index closing price. Therefore, this chapter presents the first stage of this phase, where a single approach model is introduced as an initial prediction stage. In this stage, the parameters for each proposed method will be identified in order to have an optimal prediction result and then in order to enhance the result, different ensemble techniques will be introduced. For phase 2, a prediction model for stock index direction is built in this chapter. Thus, the first stage will be explained in this chapter, where the parameters for single approaches are determined in order to obtain the optimal results.

As a result of utilising a dynamic time series data set in phase 1, this phase is called a dynamic time series prediction model. However, statistical methods were utilised to generate the technical indicators which will be fed into the prediction model in phase 2, so this model is called a static prediction model to predict the direction of stock index.

To summarise, the prediction procedures for the first stages of phases one and two are listed as follows: Firstly the dynamic prediction model of financial time series, phase1

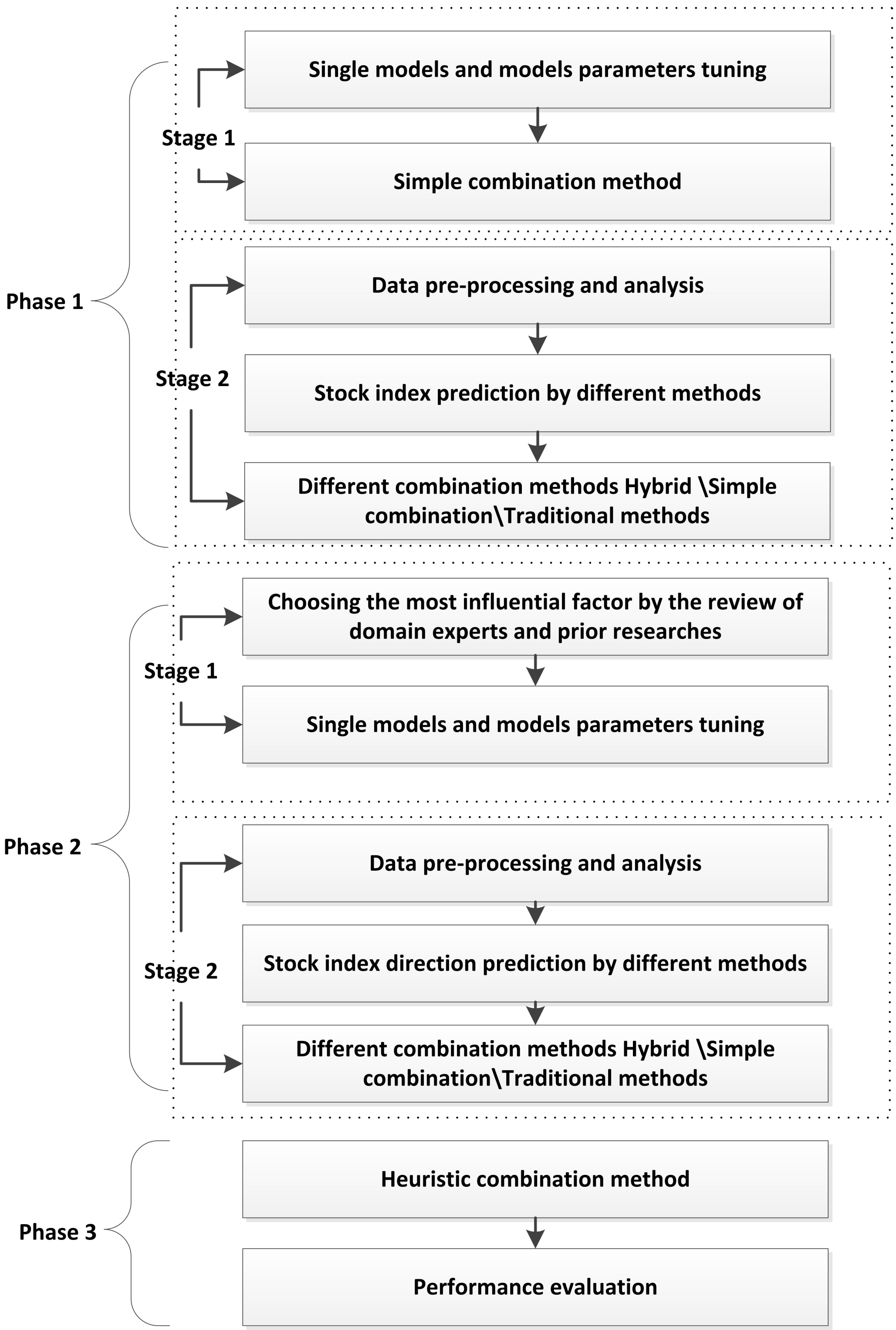

Figure 3.10: Holistic framework of prediction models.

stage1:

- To optimise the closing price of the index to derive a target price time series;
- To estimate the algorithm parameters using the in-sample data set, applying the models to the out-of-sample data set of the target price and evaluating the prediction results with the proposed statistical performance measurement;
- To introduce simple combination techniques in order to enhance the final prediction result of the single prediction approach;

Secondly, the static prediction models of the index future direction, phase 2 stage 1:

- To select the technical indicators as feature subsets through the review of domain experts and prior research;
- To optimise the direction of the closing price to derive a target direction movement of the closing price;
- To estimate the utilised algorithm parameters using the in-sample data set and applying the models to the out-of-sample data set of the target price and to evaluate the prediction results by the proposed statistical performance measurement;

3.4.1 Research Data

As mentioned in section 3.2 there are two different financial data sets utilised in this research; the financial time series data (daily closing price) which is presented in subsection 3.2.1, and technical indicators data as presented in subsection 3.2.2. Moreover, each data set will be fed into a different prediction model. In prediction procedures section 3.4 the closing price (financial time series data) of the chosen data set will be fed into the phase 1 model. The first data set is the daily closing price of the FTSE 100 index as shown in Subsection 3.2.1. It is used to demonstrate the predictability of the single approach model. The first 7788 observations (03/01/1984 - 29/10/2013) are used as the in-sample data set (training set). The last 250 observations (30/10/2013- 30/10/2014) are used as the out-sample set (testing set).

The second data set is the daily closing price of the S&P 500 index, as explained in Subsection 3.2.1, and it is also used to demonstrate and validate the predictability of the proposed single approached method. The first 16230 observations (03/01/1950 - 02/07/2014) are used as the in-sample data set (training set). The last 250 observations (03/07/2014 - 30/06/2015) are used as the out-sample set (testing set).

The third data set is the daily closing price of Nikkie 225 index as Subsection 3.2.1 illustrated, and it is also utilised to demonstrate the predictability of the proposed single approached methods. The first 7508 observations (04/01/1984 - 04/07/2014) are used as the in-sample data set (training set). The last 250 observations (07/07/2014 - 30/06/2015) are used as the out-sample set (testing set). Therefore, it is considered a necessary step to divide each data set into 2 subsets as explained previously, where the in-sample training set is the largest set and it is used by the proposed methods to learn the pattern presented in the data. Thus, in order to discover a robust model it is highly recommended to have a long training duration and large sample [289].

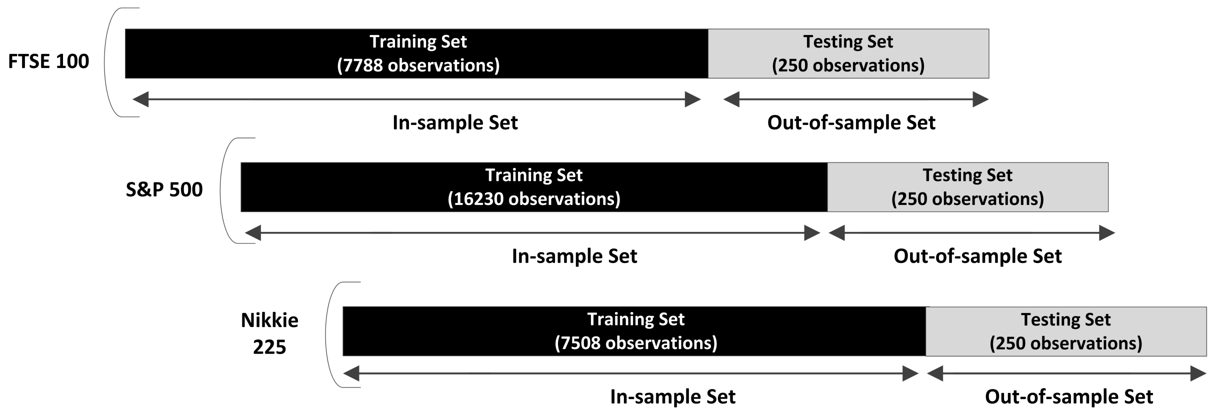

Figure 3.11: In-sample and out-of-sample data sets.

The out-sample testing set, whose length considerably less then the training set, is used to test the performance of estimated model from the in-sample (training set). Figure 3.11 illustrates the data division into the in-sample out-of-sample sets. Moreover, in phase 2 the same division of the data occurs, however different types of data sets are fed in, in order to demonstrate the predictability of the proposed models. Subsection 3.2.2 illustrates the data characteristics that are utilised in phase 2 stage 1.

3.4.2 Data Preprocessing

Selecting and preprocessing the data are crucial steps in any modelling effort, particularly for generalising the new predictive model. Data sets are divided into two sets: training and testing, which are explained in subsection 3.4.1. Thus, as mentioned in section 3.4 this chapter presents the first stage of phases one and two. In the first stage of phase 1 the preprocessing steps for financial time series data are as follows:

- Data Clearing:

  In the real world, data tends to be incomplete and inconsistent. Therefore, the routine attempt at data Clearing is to fill in missing values and to correct inconsistencies in the data. The daily closing price of FTSE 100, S&P 500 and Nikkie 225 are used as an input, which, as a result of the public holidays in the stock market, caused missing values for each data set. Treating the missing data can be done by different methods: ignore the tuple, fill in the missing value manually, use a global constant to fill in the missing value, use the attribute mean to fill in the missing value, use the attribute mean for all samples belonging to the same class as the given tuple and use the most probable value to fill in the missing value. However, this research adopted the ignore the tuple method. In other words, in order to treat the missing information in the data set, we exclude the tuples from the data set [112].

- Attribute creation:

It must be ensured that the data presented in section can be sufficiently utilised in the proposed prediction models. Therefore, and according to Fuller [92], a real value time series is considered to be a set of random variables indexed in time or it can be a set of observations ordered in time, such as performance degradation data. In accordance with the traditional setting of time series regression, the broad concept of the target value is represented as an unknown function for the input vector x_t of the p-lagged last value of y itself, as Equation 3.2 illustrates, where p is the number of back time steps. Thus the input array and the output array are respectively as follow in Equations 3.3 and 3.4, where t is the number of training data.

$$y_t = f\left(x_t\right) = f\left(y_{t-1}, y_{t-2}, \ldots.., y_{t-p}\right) \tag{3.2}$$

$$X = \begin{bmatrix} y_1 & y_2 & \cdots & y_p \\ y_2 & y_3 & \cdots & y_{p+1} \\ y_{t-p} & y_{t-p+1} & \cdots & y_{t-1} \end{bmatrix} = \begin{vmatrix} x_{p+1} \\ x_{p+2} \\ \vdots \\ x_t \end{vmatrix} \tag{3.3}$$

$$Y = \begin{vmatrix} y_{p+1} \\ y_{p+2} \\ \vdots \\ y_t \end{vmatrix} \tag{3.4}$$

Moreover, these steps are for training the in-sample data set and testing the out-of-sample data set. However, this research adopts a new approach by forming the training in-sample data set and testing the out-of-sample data set in order to achieve the best prediction results. The following forms illustrate the steps of this formation: For training data: $Tr = [x_t......x_{t+5}]\,[x_{t+6}]$, where m is a random number permutation $1 < m < p$, p is the data size. For testing data: $Ts = [x_t....x_{t+5}]\,[x_{t+6}]$, where t is $1 : p$, P is the testing data set size. Figure 3.12 illustrates the architecture of the data preprocessing.

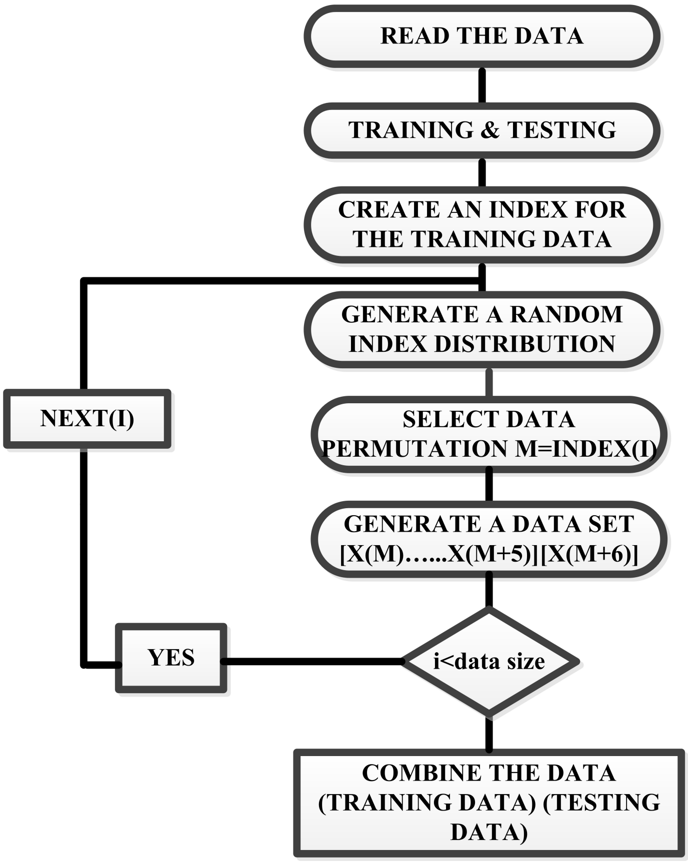

Figure 3.12: The data preprocessing frame.

However, phase 2 stage 1 preprocessing steps are as follow:

- Attribute creation:

  Eight technical indicators are considered as input variables to predict the future daily trend in the stock price index, more informations about these features are available in section 3.2.2. Thus, these variables are a results of implementing a different statistical methods. Thus, the matlab codes for obtaining the inputs are as follows: macd, willpctr, kperiods, dperiods, rsindex, tsmom and tsmovavg. Section 3.2.2 explains the steps of calculating these features. In accordance to traditional time series regression setting, the t-th output value y_t is presented in Equation 3.3. However the inputs are the eight attribute which are the presented technical indicators in section 3.2.2 and the target is Δy.

- Data Clearing:

  As a result of creating the attribute by different statistical techniques there might be some missing values. Therefore, data clearing techniques are used to check the completeness of the data. As mentioned earlier, the missing data can be treated by different methods. This research uses the attribute mean to fill in the missing value [112].

3.5 Methodologies and Empirical Results

There are many methods for predicting the stock price, such as fundamental methods, technical methods and traditional time series methods. Prediction techniques can be classified into statistical models and artificial intelligence models. In Chapter 2 a cutting edge literature review was provided on this domain. As explained in the holistic framework of prediction models in Figure 3.10, this chapter presents the first stage of phase 1 and 2. Thus, each phase investigates the predictability of the proposed models by using different input data sets. Phase 1 stage 1 predicts the closing stock index price by using the previous historical data, however, in stage 1 phase 2 eight technical indicators were chosen to predict the direction of the closing price of the stock index.

The single prediction approach is described in the following sections and subsections, while the focus of these prediction approaches are divided into two directions. First the single approach is illustrated, along with the prediction models of the stock index closing price using the historical daily closing price. Figure 3.13 illustrates the basic stages of the single prediction approach, where n^{th} day ahead is the prediction task. The historical closing prices of each data set are the inputs to the prediction models. Three data sets are used individually in this approach to be predicted by SVR, PBNN and RNN techniques,

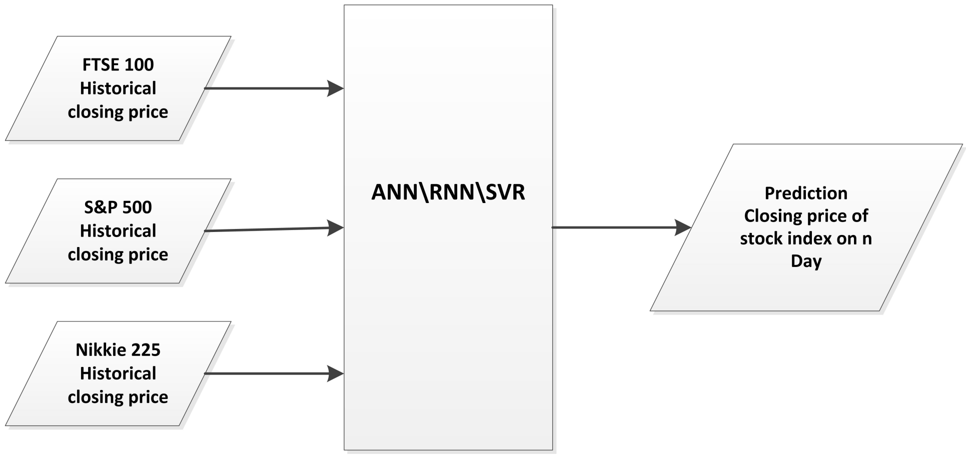

Figure 3.13: Single stage approach for predicting n day a head of time.

which are described in the following subsections.

The second single prediction approach is illustrated in Figure 3.14. Here, n^{th} is the prediction task of day a head of time t. In this approach, and as explained earlier, eight technical indicators are used as an inputs describing t^{th} day, which are summarised in section 3.2.2, while $(n+t)^{th}$ is the output days direction of the closing price. SVM, RNN and Naive Bayes are the employed prediction models, which are described in the following subsections.

3.5.1 Benchmark Prediction Model

In this thesis a traditional prediction model, the Simple Auto-regressive model (AR) is used, in order to benchmark the performance efficiency of the utilised models. Moreover, the simple average in this thesis is used as a benchmark combination method.

3.5.1.1 Simple Auto-regressive Model

The Auto-regressive (AR) model in this study is used as a benchmark model to evaluate the prediction power between the utilised models based on the relative improvements i

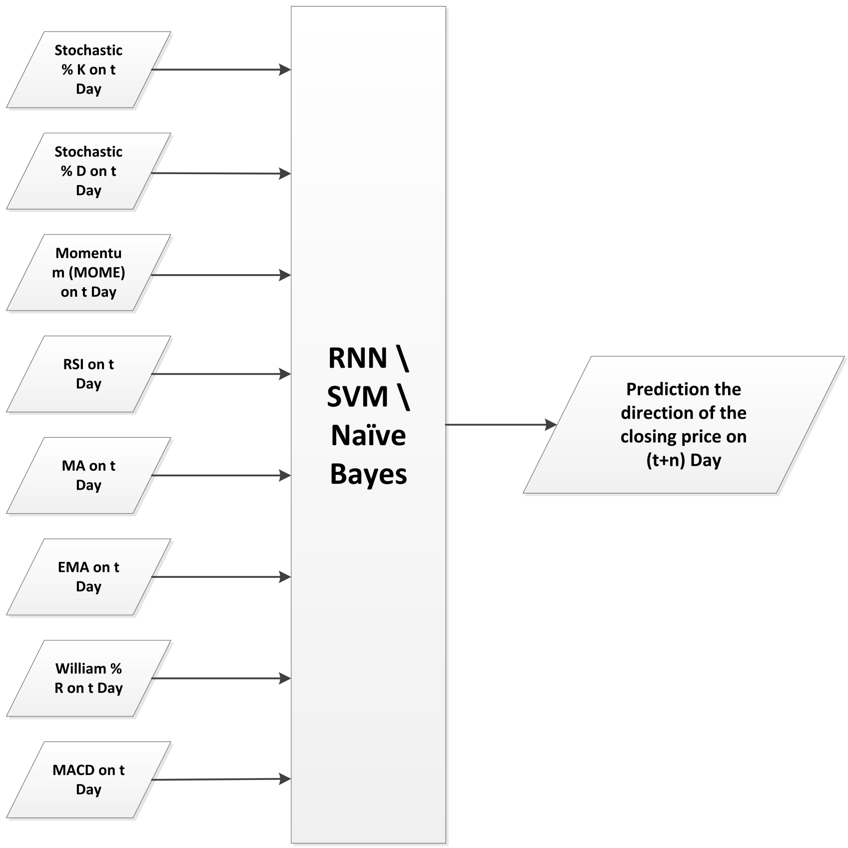

Figure 3.14: Single stage stock direction approach for predicting n day a head of time .

root mean square error. Equation 3.5 illustrates AR model used.

$$y_t = a_1 y\,(t-1) + a_2 y(t-2) + ..., a_t y\,(t-p) \tag{3.5}$$

$$\begin{aligned}\Delta y_t = a_1 y^{Sk}\,(t) + a_2 y^{SD}(t) + a_3 y^{MO}(t) + a_4 y^{RSI}(t) \\ + a_5 y^{MA}(t) + a_6 y^{EMA}(t) + a_7 y^{R}(t) + a_8 y^{MACD}\,(t)\end{aligned} \tag{3.6}$$

In Equation 3.5 $y\,(t)$ is the predicted stock price based on the past close daily price, $y\,(t) - y\,(t-n)$ and the coefficients of AR model are $a_1 - a_n$. 5 lagged daily price is the order of which is used in the AR model was varied and found to give better prediction result. Moreover, to predict next day direction closing price of stock index, equation 3.6 is employed. Δy is the target direction price at the time t and the inputs are y^{Sk}= Stochastic %K, y^{SD}= Stochastic %D, y^{MO}= Momentum, y^{RSI}= RSI, y^{MA}= Moving Average, y^{EMA}= Exponential Moving Average,y^{R}= William %R and y^{MACD}= Moving

Average Convergence and Divergence . The model coefficients were determined by using the Implemented regress function in the MATLAB.

3.5.1.2 Prediction Combination Techniques

Combining different prediction techniques has been investigated widely in the literature. In short-term predictions, combining the various techniques is more useful according to [293], [12]. Timmermann [242] stated in his study that using the simple average may work as well as more sophisticated approaches. In this research, the simple average is used a benchmark combination model. In this chapter, two stages of the first two phases are presented, and in each stage three prediction techniques are used for the prediction model. Equation 3.7 illustrates the calculation of the combination prediction method at time t [275].

$$f_{SM}^{t} = \left(f_{M1}^{t} + f_{M2}^{t} + f_{M3}^{t}\right) \div 3 \tag{3.7}$$

3.5.2 Artificial Neural Network

In recent years, predicting financial time series utilising Artificial Neural Networks (ANNs) has increased dramatically. The idea of ANN can be seen before reaching the output, where the filtration of the inputs through one or more hidden layers, each of which consists of hidden units, or nodes, is considered to be the main idea. Thus, final output is related to the intermediate output [68].

The ability of learning from data through adaptive changing structure based on external or internal information that flows through the network during the learning phase and generates output variables based on learning is one of the most important advantages of ANN. Furthermore, the non-linear nature of ANN is also a valuable quality. ANN is classified as a non-linear data modelling tool, thus, one of the main purposes of utilising such a model is to find the patterns in data or to model a complex relationships between inputs and outputs. Hence, an explicit model-based approach fails, but ANNs adapts to irregularities and unusual features in a time series.

The application of ANNs has been popularly utilised in financial time series prediction modelling. A comprehensive review of ANNs and their application in various financial domains is given in Chapter 2. However, the same as with any other technique, ANNs have some disadvantages such as not allowing much understanding of the data, which might be caused by it not being an explicit model. Therefore, providing a black box for the prediction process is considered as a disadvantage. The danger of over-fitting the in-sample training data is also a major ANN method drawback [158]. In term of the goodness of fit, the performance on in-sample data sets, which ANNs are trained on, is reasonably good. However, in out-of-sample sets its performance is conditional on not breaking the structure in the data sets. According to Balestrassi et al.[18], excessive training time and a large number of parameters that must be experimentally selected in order to generate good predictions are considered to be the other drawbacks facing ANN applications. In this thesis, two ANN architectures, BPNN and RNN, are used to predict financial data, and subsections 3.5.2.1 and 3.5.2.2 demonstrate these models.

3.5.2.1 Back Propagation Neural Network

In modelling time series with non-linear structures, the most commonly used structure is three layers feed forward back propagation [73]. The weights are determined in the back propagation process by building connections among the nodes based on data training, producing a least-mean-square error measure of the actual or desired and the estimated values from the output of the neural network. The initial values are assigned for the connection weights. In order to update the weights, the error between the predicted and actual output values is back propagated via the network. Minimising the errors in the desired and predicted output attempts takes place after the procedure of supervised learning [152]. The architecture of this network contains a hidden layer of neurons with a non-linear transfer function and an output layer of neurons with a non-linear transfer function and an output layer of neurons with linear transfer functions. Figure 3.15 illustrates the architecture of a back propagation network, where $x_j\,(j = 1, 2, ..., n)$ represent the input variables; $z_i\,(j = 1, 2, ..., m)$ represent the outputs of neurons in the hidden layer; and $y_t\,(t = 1, 2, ..., l)$ represent the outputs of the neural network [254]. In theory,

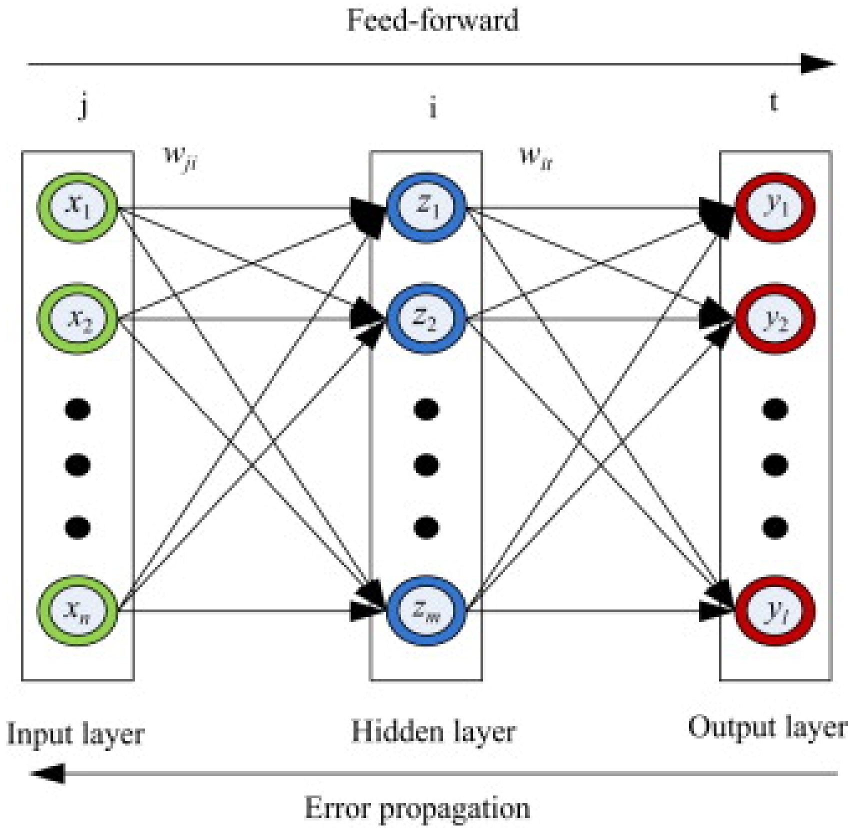

Figure 3.15: Architecture of feed forward back propagation neural network [254].

the neural network has the ability to simulate any kind of data pattern given a sufficient training. Training the neural network will determine the perfect weight to achieve the correct outputs. The following steps illustrate the training process of updating the weights values [89]. The first stage is hidden layers; the bellow equation explains how the outputs of all the neurons in the hidden layer are calculated:

$$net_i = \sum_{j=0}^{n} w_{ji} x_j v i = 1, 2, ...m \tag{3.8}$$

$$z_i = f_H \left(net_i \right) i = 1, 2, ...m \tag{3.9}$$

Where net_i is the activation value of the ith node, z_i is the output of hidden layer, and f_H is called the activation Equation of a node, in this research a sigmoid function is utilised. Equation 3.10 explains the utilised sigmoid activation equation.

$$f_H \left(x \right) = \frac{1}{1 + exp(-x)} \tag{3.10}$$

Second stage the output: The outputs of all the neurons in the output layer are given as equation 3.11 illustrates:

$$y_t = f_t \left(\sum_{i=0}^{m} w_{it} z_i \right) t = 1, 2, ...l \tag{3.11}$$

The activation equation is $f_t \left(t = 1, 2, ..., l \right)$, Which is usually a line equation. The weights are assigned with random values initially, and are modified by the delta rule according to the learning samples traditionally. The topology in this study is determined by the trial and error method, which was conducted to choose the best number of neuron experiments different ranges of 20 to 5 neurons in a hidden layer two layer feed forward back propagation network were tried. The model stopped training after reaching the pre-determined number of epochs. The ideal topology was selected based on the lowest Mean Square Error.

3.5.2.2 Recurrent Neural Network

Recurrent Neural Network (RNN) is considered as an enhanced ANN architecture, and thus a variant of Elman's network [81]. The ability to form more complex computations than the static feed forward network is the reason behind adopting the RNN algorithm. Furthermore, the capability of learning temporal pattern sequences which are context- or time-dependent is also an advantage of utilising such a method. Embodying a short-term memory by activating a feedback network is also one of the main features of a simple Recurrent Neural Network (RNN). According to Tenti [239], requiring more substantial memory and connections in simulations, in comparison with a back propagation network, is one of the main disadvantages of RNN, and therefore this increase leads to high computational timing. However, utilising RNN can yield better results. Figure3.16

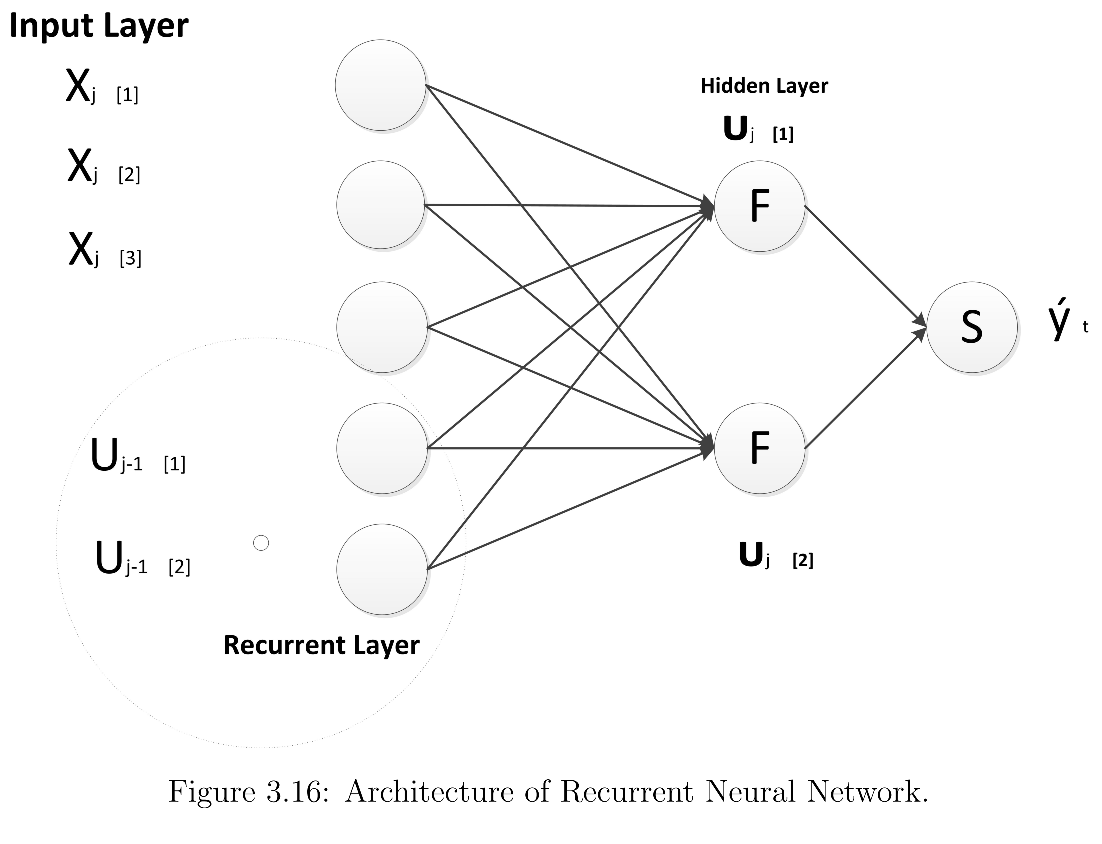

Figure 3.16: Architecture of Recurrent Neural Network.

demonstrates the RNN architecture, where $x_t^{[n]}\,(n = 1, 2, ...k + 1)\,, u_t^{[1]}, u_t^{[2]}$ are the inputs for the RNN model at time t including bias node. The output of RNN model is $\check{y}_t$

and $d_t^{[f]}$ $(f = 1, 2)$, $w_t^{[n]}$ $(n = 1, 2, ...k + 1)$ are presenting the weights of the network. The output of the hidden nodes is $U_t^{[f]}, f = (1, 2)$ at time t. The activation function in this model is sigmoid $F : K(x) = \frac{1}{1+e^{-x}}$ and S in the Figure above presents the linear function: $J(x) = \sum_i x_i$.Function 3.12 illustrates the way in which the error is minimised.

$$E(d_t, w_t) = \frac{1}{T} \sum_{t=1}^{T} (y_t - \hat{y}_t(d_t, w_t))^2 \tag{3.12}$$

3.5.2.3 BPNN and RNN Parameters Determination

A side from the above settings described, for RNN and BPNN algorithms, there were still a number of parameters to be decided. The number of hidden layer, number of neurons (n), value of learning rate (lr), and momentum constant (mc) are PBNN, RNN models parameters that must be efficiently determined.

Table 3.6: Model selection results of the single BPNN prediction model FTSE100, S&p500 and Nikkei225

Number of nodes for two hidden layer	**Learning rate** lr	**Momentum constant** mc	**FTSE100 Training** MSE	**S&P500 Training** MSE	**Nikkei225 Training** MSE
5-5	0.3	2	2425	58	55715
	0.4	4	2420	57	55255
	0.5	6	2401	59	53829
	0.6	8	2422	58	54680
	0.7	10	2421	59	56651
10-5	0.3	2	2409	57.7	52941
	0.4	4	2409	58.17	53362
	0.5	6	2987	58.07	53763
	0.6	8	2404	93	53707
	0.7	10	2382	58.16	45089
15-5	0.3	2	2428	57.9	53027
	0.4	4	2410	56.13	51945
	0.5	6	2402	58.01	53840
	0.6	8	2375	57.6	56725
	0.7	10	2399	58.32	54588
20-5	0.3	2	2339	56.63	52561
	0.4	4	2390	58.17	52743
	0.5	6	2411	57.98	51675
	0.6	8	2401	58.16	54839
	0.7	10	2413	56.09	71976

Since there are no general rules for designing the model topology, a trial and error set was conducted to choose the best number of parameters for both the RNN and PBNN models during the training process. The model stopped training after reaching the predetermined number of epochs. The ideal topology was selected based on the lowest Mean Square Error. According to Chauvin and Rumelhart [45], one hidden layer is sufficient to model a complex system with the desired performance. However, in this chapter, two hidden layers are used. The nodes of two hidden layers is set to be $[5-5]$, $[10-5]$, $[15-5]$ and $[20-5]$. The learning rate is considered to be an important parameter in training the model, and thus during the training process a set of learning rates of 0.3, 0.4, 0.5, 0.6, and 0.7 are tested. In addition to the above-mentioned parameters, a set of momentum constants, $2, 4, 6, 8$ and 10 are also tested during the training process.

Table 3.7: Model selection results of the single RNN prediction model FTSE100, S&p500 and Nikkei225

Number of nodes for two hidden layer	**Learning rate** lr	**Momentum constant** mc	**FTSE100 Training** MSE	**S&P500 Training** MSE	**Nikkei225 Training** MSE
	0.3	2	2397	58.24	55739
	0.4	4	2413	58.08	55901
5-5	0.5	6	2410	59.01	56119
	0.6	8	2417	58.07	57009
	0.7	10	2403	57.78	57011
	0.3	2	2409	59.01	56612
	0.4	4	2396	58.40	57214
10-5	0.5	6	2409	58.23	56654
	0.6	8	2408	59.16	57851
	0.7	10	2396	60.81	56631
	0.3	2	2407	58.9	54420
	0.4	4	2419	59.17	56440
15-5	0.5	6	2579	58.13	58620
	0.6	8	2377	58.87	57781
	0.7	10	2422	57.12	54391
	0.3	2	2389	56.85	54196
	0.4	4	2423	55.61	54663
20-5	0.5	6	2399	57.91	55339
	0.6	8	2406	58.61	54293
	0.7	10	2400	56.96	68976

In this chapter and as mentioned earlier in section 3.5 two kinds of modelling are carried out to Predicting the next day direction price of the FTSE100, S&P500 and Nikkei225 stock indices, where RNN is one of the utilised algorithms. RNN parameters are determined using the proposed method in this section, where the testing results with a combination of different hidden layers, learning rates and momentum constants are summarised in Table 3.8. It is observed from Table 3.8 that the parameters which give the minimum MSE and hence the best topology for the RNN model for the three data sets is the FTSE100 data set. The model topology is [20-5] (20, 5 nodes for the two hidden layers with learning rate of 0.3 and momentum constant of 2). The S&P500 model topology is [10-5] (10, 5 nodes for the two hidden layers) with learning rate of 0.3 and momentum constant of 2. For the Nikkei225 it is [10-5] (10, 5 nodes for the two hidden layers) with learning rate of 0.7 and momentum constant of 10.

Table 3.8: Model selection results of the single RNN Trend prediction model FTSE100, S&p500 and Nikkei225

Number of nodes for two hidden layer	**Learning rate** lr	**Momentum constant** mc	**FTSE100 Training** MSE	**S&P500 Training** MSE	**Nikkei225 Training** MSE
	0.3	2	2436	0.303	55348
	0.4	4	2474	0.305	55758
5-5	0.5	6	2444	0.3031	56029
	0.6	8	2438	0.303	55174
	0.7	10	2460	0.3048	54867
	0.3	2	2442	0.302	55843
	0.4	4	2493	0.305	56004
10-5	0.5	6	2437	0.3054	55019
	0.6	8	2438	0.302	55541
	0.7	10	2439	0.305	54952
	0.3	2	2444	0.302	55902
	0.4	4	2435	0.304	55554
15-5	0.5	6	2437	0.306	55206
	0.6	8	2434	0.305	55429
	0.7	10	2439	0.3046	56073
	0.3	2	2426	0.304	55499
	0.4	4	2441	0.306	56073
20-5	0.5	6	2438	0.305	56012
	0.6	8	2439	0.308	55482
	0.7	10	2437	0.302	56049

Furthermore, the second prediction model for the next day closing price of the same data sets using RNN is also introduced in this research. The parameters of this model are illustrated in Table 3.7. The best observed parameters to set up an RNN model giving the minimum MSR are shown in Table 3.7. Firs, for the FTSE100 data set model the topology is, [20-5] (20, 5 nodes for the two hidden layers) with learning rate of 0.6 and momentum constant of 8. Second, the S&P500 model topology is [20-5] (20, 5 nodes for the two hidden layers) with learning rate of 0.4 and momentum constant of 4. Third, for the Nikkei225 it is [20-5] (20, 5 nodes for the two hidden layers) with learning rate of 0.3 and momentum constant of 2.

In addition to those models, BPNN is also used for modelling and predicting the same data sets, with Table 3.6 explaining the combination of parameter results. The minimum MSE combination of parameters to set up the BPNN prediction model for the three data sets, as illustrated in Table 3.6, are as follow: Firs, for the FTSE100 data set the model topology is, [20-5] (20, 5 nodes for the two hidden layers) with learning rate of 0.3 and momentum constant of 2. Second, the S&P500 model topology is [20-5] (20, 5 nodes for the two hidden layers) with learning rate of 0.7 and momentum constant of 10. Third, for Nikkei225 it is [10-5] (10, 5 nodes for the two hidden layers) with learning rate of 0.7 and momentum constant of 10.

3.5.3 Support Vector Machine

The SVM theory was developed by Vladimir Vapnik in 1995. It is considered as one of the most important breakthroughs in machine learning field and can be applied in classification and regression [65]. In modelling the SVM, the main goal is to select the optimal hyperplane in high dimensional space, ensuring that the upper bound of the generalisation error is minimal. SVM can only directly deal with linear samples but mapping the original space into a higher dimensional space can make the analysis of a non-linear sample possible [209] [195]. For example, if the data point (x_i, y_i) was given randomly and independently generated from an unknown function, the approximate function form by SVM is as follow:

$g(x) = w\emptyset(x) + b.\emptyset(x)$ Is the feature and non-linear mapped from the input space x. w and b are both coefficients and can be estimated by minimising the regularised risk equation.

$$R(C) = C\frac{1}{N}\sum_{i=1}^{N} L(d_i, y_i)\frac{1}{2} + ||w||^2 \tag{3.13}$$

$$L(d, y) = \begin{cases} |d-y| - \varepsilon & |d-y| \geq \varepsilon \\ 0 & other, \end{cases} \tag{3.14}$$

C and ε in Equation 3.13 and 3.14 are prescribed parameters. C is called the regularisation constant while ε is referred to as the regularisation constant. $L(d, y)$ Is the intensive loss function and the term $C\frac{1}{N}\sum_{i=1}^{N} L(d_i, y_i)$ is the empirical error while the $\frac{1}{2} + ||w||^2$ indicates the flatness of the function. The trade-off between the empirical risk and flatness of the model is measured by C. Since introducing positive slack variables ζ and ζ^* equation 3.14 transformed to the following:

$$R(w, \zeta, \zeta^*) = \frac{1}{2}ww^T + C \times \left(\sum_{i=1}^{N}(\zeta, \zeta^*)\right) \tag{3.15}$$

Subject to:

$$w\phi(x_i) + b_i - d_i \leq \varepsilon + \zeta_{i^*} \tag{3.16}$$

$$d_i - w\phi(x_i) - b_i \leq \varepsilon + \zeta_i \tag{3.17}$$

$$\zeta_i, \zeta_{i^*} \geq 0 \tag{3.18}$$

The decision equation (kernel function) comes up finally after the Lagrange multipliers are introduced and optimality constraints exploited. Equation 3.19 is the form of kernel Equation:

$$f(x) = \sum_{i}^{l}(\alpha_i - \alpha_i{}')K(x_i, x_j) + b \tag{3.19}$$

Where ${\alpha_i}'$ and α_i are called Lagrange multipliers in equation 3.19. They satisfy the equalities, $\alpha_i \times {\alpha_i}' = 0, \alpha_i \geq 0, {\alpha_i}' \geq 0$. The kernel value is the same with the inner product of two vectors x_i and x_j in the feature space $\phi\left(x_i\right)$ and $\phi\left(x_j\right)$. The most popular kernel function is Radial Basis Function (RBF) it is form in Equation 3.20.

$$K\left(x_i, x_j\right) = exp\left(-\gamma||x_i - x_j||^2\right) \tag{3.20}$$

Theoretical background, geometric interpretation, unique solutions and mathematical tractability are the main advantages which have made SVM attractive to researchers and investors alike, and it can be applied to many applications in different fields such as predicting financial time series [231].

3.5.4 Support Vector Regression

As explained in subsection 3.5.3, the idea of SVM is to construct a hyperplane or set of hyperplanes in a high or infinite dimensional space, which can be used for classification. In the regression problem the same margin concept used in SVM is used. The goal of solving regression problems is to construct a hyperplane that is close to as many of the data points as possible. Choosing a hyperplane with a small norm is considered the main objective, while simultaneously minimising the sum of the distances from the data points to the hyperplane [274].

In the case of solving a regression problem using SVM, SVM became known as the support vector regression (SVR) where the aim is to find a function f with parameters w and b by minimizing the following regression risk:

$$R\left(f\right) = \frac{1}{2}\left(x, w\right) + C\sum_{i=1}^{N} l\left(f\left(x_i\right), y_i\right) \tag{3.21}$$

C in Equation 3.21 is a trade-off term, the margin in SVM is the first term which is used in measuring VC-dimension [89].

$$f(x, w, b) = (w, \phi(x)) + b, \tag{3.22}$$

In the Equation 3.22 $\phi(x) : x \rightarrow \Omega$ is kernel function, mapping x into in the high dimensional space. SVR and as proposed by [274]. The $-\varepsilon$ insensitive loss function is used as follows:

$$l(y, f(x)) = \begin{cases} 0, if|y - f(x) < \varepsilon \\ |y - f(x)| - \varepsilon, Otherwise \end{cases} \tag{3.23}$$

Equation 3.24 constrained minimisation problem is equivalent to previous minimisation Equation 3.21.

Min

$$y\left(w, b, \zeta^* = \frac{1}{2}(w, w) + C\sum_{i=1}^{N}(\zeta_i + \zeta_{i^*})\right) \tag{3.24}$$

Subject to:

$$y_i - ((w, \phi(x_i) + b)) \leq \varepsilon + \zeta_i, \tag{3.25}$$

$$((w, \phi(x_i)) + b) - y_i \leq \varepsilon + \zeta_{i^*}, \tag{3.26}$$

$$\zeta_i^* \geq 0 \tag{3.27}$$

In sample (x_i, y_i) the ζ_i and ζ_{i^*} measure the up error and down error. Maximizing the dual function or in other words construct the dual problem of this optimization problem (primal problem) by large method is a standard method to solve the above minimization problem. There are four common kernel functions, among all theses four kernel function, this study will be utilising radial basis function (RBF). In accordance to [59], [237] and [31] RBF kernel function is the most widely applied in SVR. Function 3.28 is defining the kernel RBF, where the width of the RBF is denoted by σ. Furthermore, Cherkassky and Ma [59] suggested that the value of σ must be between 0.1 and 0.5 in order for SVR model to achieve the best performance. In this thesis σ value is determined as 0.1.

$$K\left(x_i, x_i\right) = exp\left(\frac{-||x_i - x_j||^2}{2\sigma^2}\right) \tag{3.28}$$

3.5.4.1 Determining the Parameters in SVM, SVR Models

This research used SVM to predict the direction movement of the chosen data sets and SVR to predict the daily closing price of the chosen data sets. If the parameter of each model is set properly, that could improve the prediction output for each model. In both models, the RBF kernel function has been used, as indicated in Functions 3.16, 3.17 and 3.16 for SVM and in Function 3.25, 3.26 and 3.27. Two parameters should determine C and γ in both models SVM and RBF.However, in the literature there are no general rules for choosing those parameters.Thus, this research adopting the most common approach to search the best C and γ values, which is grid search approach [293]. The grid search approach was proposed by Lin et al [121] using a validation process in order to produce good generalisations with which to decide parameters. The criteria for choosing the optimal C and γ parameters is by trying pairs of C and γ, and the best combination of these parameters that can generate the minimum mean square error MSE is chosen to set up the prediction model.

On the grid and after identifying a better region, on this region a more specific grid search can be conducted [121]. Finally, after determining these lowest cross validation prediction error parameters C and γ, choosing these parameters to be used in creation the prediction models. Figure 3.17 illustrates the process of the grid search algorithm when building the SVM, SVR prediction model.

A set of exponentially growing sequences of C and γ are used and the best parame-

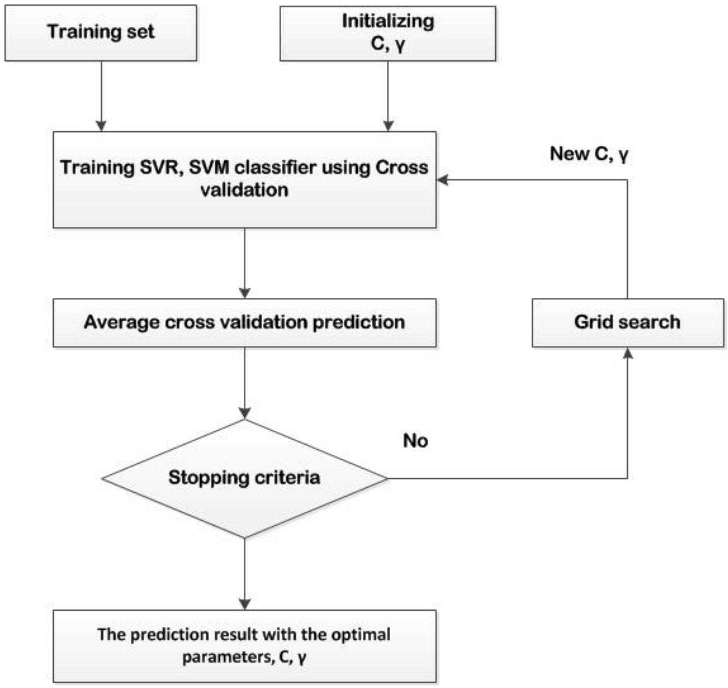

Figure 3.17: Grid search algorithm for parameters selection

ter combination results are illustrated in Table 3.9 for SVR next day closing price prediction model. Different combination of parameters sets ($C = 60, 70, 80, 90, 100$) and ($\gamma = 0.0001, 0.0002, 0.0003, 0.0004, 0.0005$) are shown in Table 3.9 and Table 3.10.

Table 3.9 shows that the parameters for the FTSE100 prediction next day closing price is ($C = 100, \gamma = 0.0001$), which gives the minimum mean square error MSE in the training data set and is the best one for setting up the prediction model. For S&P500 the next day closing price prediction combination parameters are illustrated in Table 3.9, thus ($C = 100, \gamma = 0.0001$) are the best combination parameters which give the minimum MSE in training data set. Furthermore, the best combination results to set up the SVR prediction next day closing price model of the Nikkei225 parameters are ($C = 100.\gamma = 0.0003$) which give the minimum MSE in the training data set as Table 3.9 shows.

Moreover, Figure 3.18 illustrates the grid search result for choosing the best combination parameters for the FTSE100 data set. Figure 3.19 shows the parameter selection for

Table 3.9: Model selection results of the single SVR prediction model FTSE100, S&P500 and Nikkei225

C	γ	**FTSE100 Training** MSE	**S&P500 Training** MSE	**Nikkei225 Training** MSE
	0.0001	2925.9	2449.8	41986
	0.0002	2926.8	2453.8	41986
60	0.0003	2927.6	2457.6	41986
	0.0004	2928.3	2461.1	41895
	0.0005	2928.9	2464.4	41994
	0.0001	2556.5	2169.8	41949
	0.0002	2565.1	2184.6	41934
70	0.0003	2572.7	2199.7	41984
	0.0004	2579.5	2214.6	41988
	0.0005	2585.5	2228.7	41949
	0.0001	2400.6	2073.8	41896
	0.0002	2411.6	2092.4	41869
80	0.0003	2421.6	2110.0	41897
	0.0004	2430.7	2126.5	41879
	0.0005	2438.7	2141.8	41867
	0.0001	2194.8	1947.8	41821
	0.0002	2208.1	1971.9	41823
90	0.0003	2220.2	1994.6	41825
	0.0004	2231.2	2015.4	41832
	0.0005	2241.2	2034.5	41852
	0.0001	1929.4	1786.6	41718
	0.0002	1944.2	1813.4	41715
100	0.0003	1957.9	1841.0	41703
	0.0004	1970.7	1868.1	41709
	0.0005	1982.4	1893.4	41712

the Nikkei225 data set. The grid search selection for the S&P500 SVR prediction model parameters is explained in Figure 3.20.

Table 3.10 shows the results of the parameters combination to set up a prediction model using SVM for the FTSE100, S&P500 and Nikkei225 data sets. As Table 3.10 presents, MSE is used to determine which combination of parameters is the best, where the lowest MSE of the parameters combination is chosen to set up the prediction models. Firstly, for FTSE100 data set the best combination of parameters are $C = 100, \gamma = 0.0003$. Secondly, for S&P500 the best combination of parameters are $C = 100, \gamma = 0.0003$. Thirdly, for Nikkei225 the best combination of parameters are $C = 90, \gamma = 0.0002$.

Additionally, Figure 3.21 represents the lowest and highest MSE of the grid search results

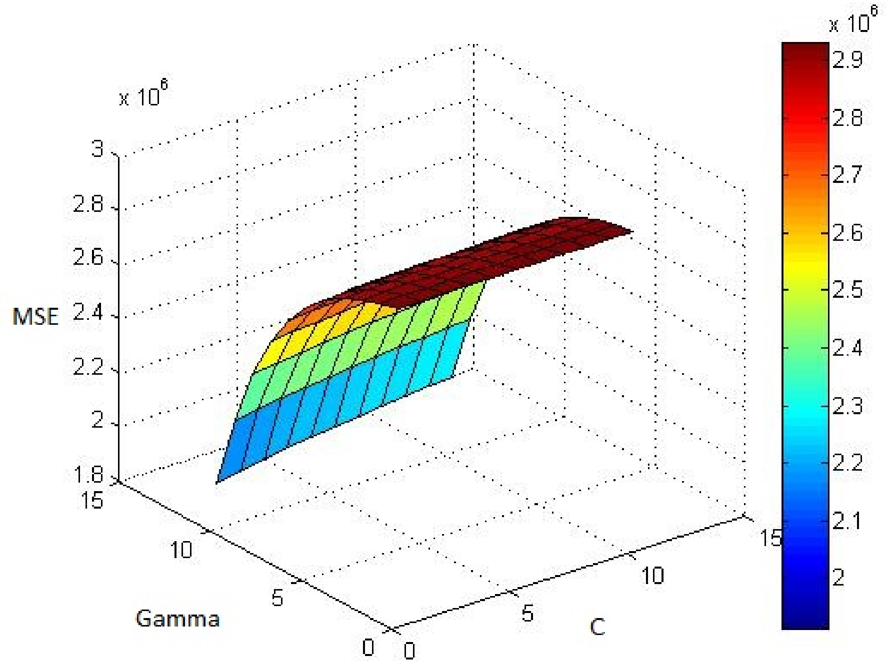

Figure 3.18: C and γ parameters for SVR FTSE100 prediction model

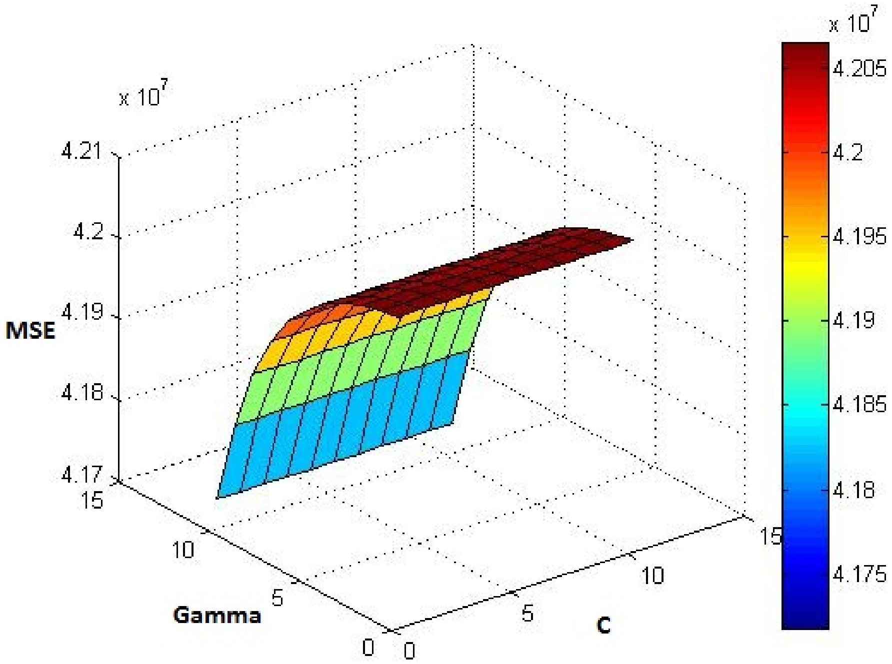

Figure 3.19: C and γ parameters for SVR Nikkei225 prediction model

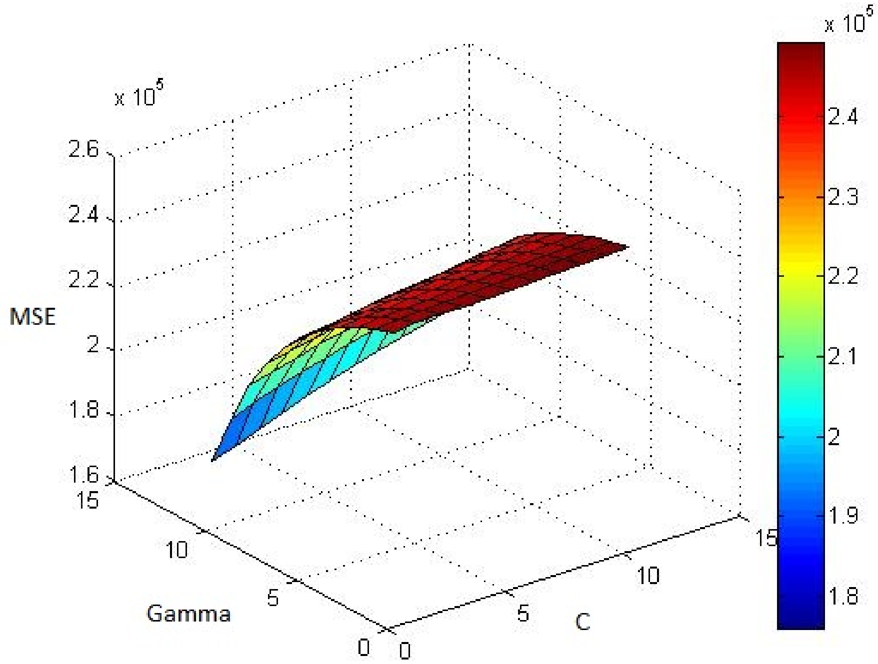

Figure 3.20: C and γ parameters for SVR S&P500 prediction model

for the Nikkei225 stock index. Figure 3.22 also illustrates the MSE of the parameters combination by grid search for the FTSE100 stock index. Finally, Figure 3.23 shows the output of the gird search process for the C and γ, SVM model parameters for S&P500 stock index.

Moreover, to build both SVM and SVR prediction models, the $LIBSVM-SVR$ package proposed by Chang and Lin [35] is adapted in this research.

3.5.4.2 Quantization Factor

This research adopts a new approach to enhance the prediction output of SVM and SVR models. The quantisation factor is for the first time introduced to SVM and SVR. In this paper, this factor has been added to the SVM and SVR model. As explained above in the methodology section, in SVM and SVR the model input is (x_i, y_i). After adding the optimal factors which were determined by trial and error, the optimal factor was chosen

Table 3.10: Model selection results of the single SVM Trend prediction model FTSE100, S&P500 and Nikkei225

C	γ	**FTSE100 Training** MSE	**S&P500 Training** MSE	**Nikkei225 Training** MSE
	0.0001	1881.2	19.0549	53131
	0.0002	1882.0	19.0551	53031
60	0.0003	1881.9	19.0603	53303
	0.0004	1881.5	19.0600	53230
	0.0005	1882.3	19.0559	53122
	0.0001	1681.4	12.4246	51930
	0.0002	1681.8	12.4265	51983
70	0.0003	1682.3	12.4305	51945
	0.0004	1683.4	12.4248	51936
	0.0005	1683.6	12.4336	51953
	0.0001	7023.8	2.0738	50270
	0.0002	1441.0	7.01208	50366
80	0.0003	1431.5	7.0301	50293
	0.0004	1432.5	7.02451	50309
	0.0005	1431.1	7.02602	50320
	0.0001	3342.8	1.9478	47995
	0.0002	1136.5	3.3226	47891
90	0.0003	1140.1	3.3522	48001
	0.0004	1140.6	3.4016	48021
	0.0005	1140.3	3.40095	47999
	0.0001	1273.7	1.7866	45116
	0.0002	829.9	1.2841	45129
100	0.0003	826.4	1.2195	44910
	0.0004	830.7	1.3091	44919
	0.0005	829.4	1.3005	44940

from a range of factors between 10 and 100. The optimal factor was selected based on the lowest mean square error. The steps below illustrate the change in the input after introducing the quantisation factor:

$$Xprim = X_i \div factor \tag{3.29}$$

$$Yprim = Y_i \div factor \tag{3.30}$$

The model inputs become $(Xprim_i, Yprim_i)$. After applying the above models (SVM and SVR) and to obtain the final prediction result, the chosen factor multiplied with the

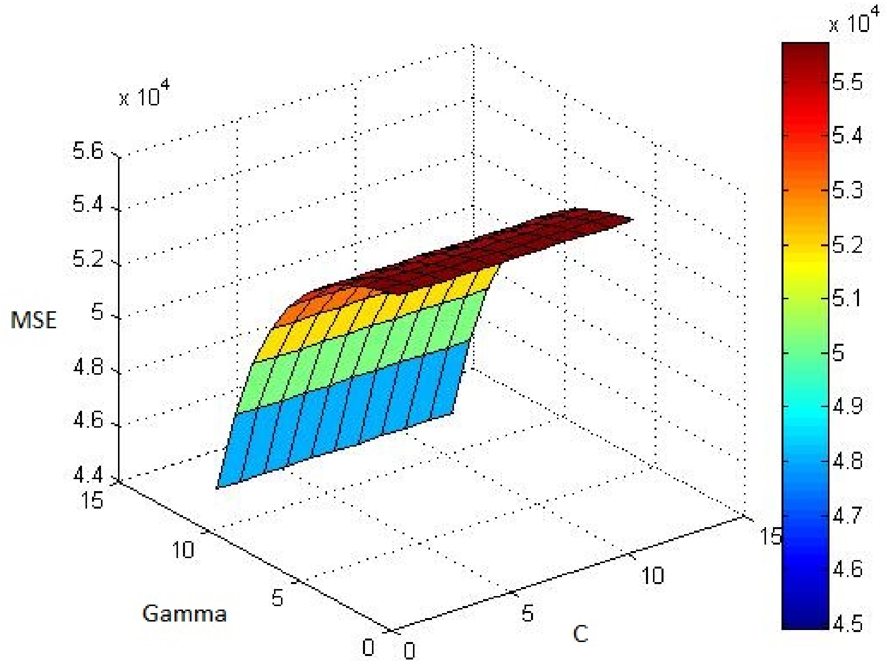

Figure 3.21: C and γ parameters for SVM Nikkei225 Trend prediction model

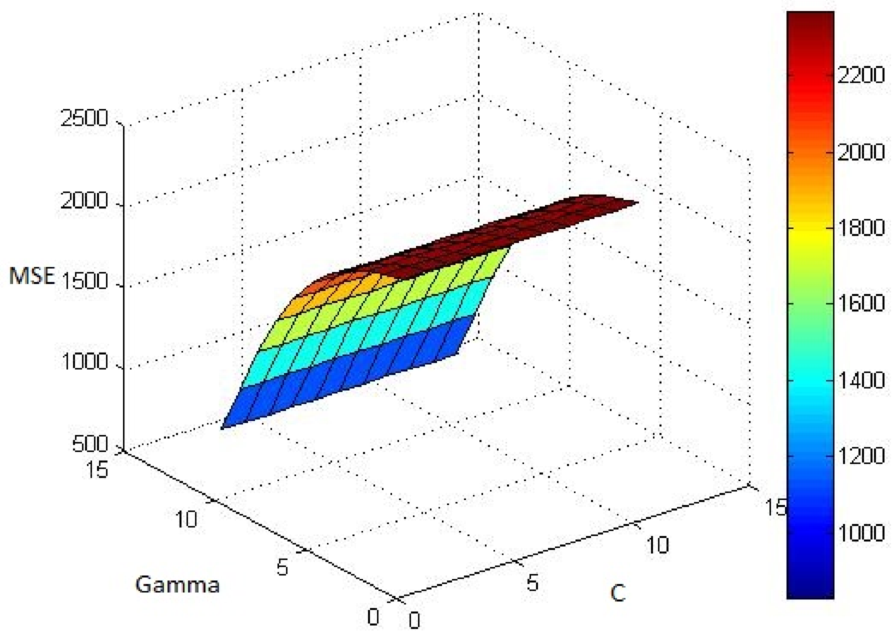

Figure 3.22: C and γ parameters for SVM FTSE100 Trend prediction model

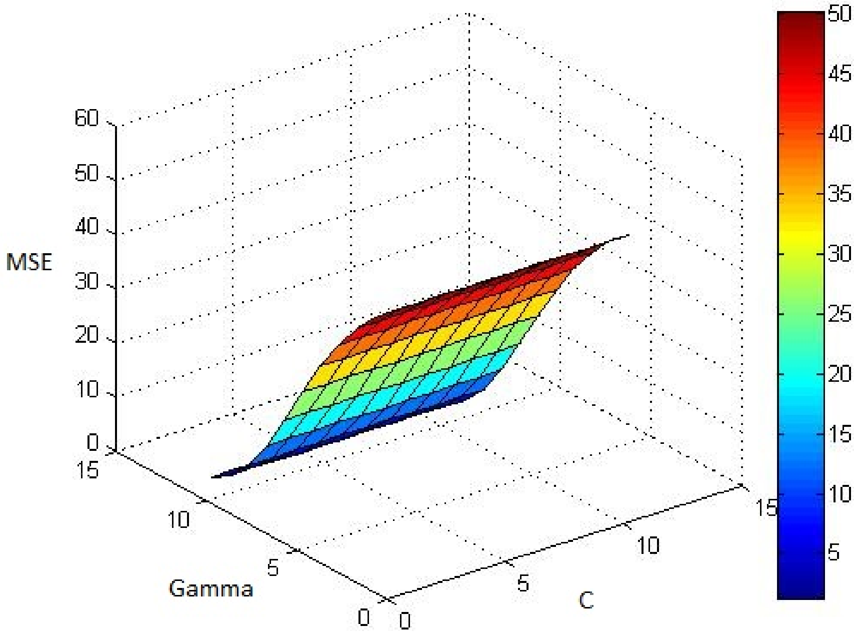

Figure 3.23: C and γ parameters for SVM S&P500 Trend prediction model

output of each model as illustrated in Equation 3.31 and 3.32:

$$Xprim_{pred} = Xprim \times Factor \tag{3.31}$$

$$Yprim_{pred} = Yprim \times Factor \tag{3.32}$$

This method has been proposed to enhance the performance and prediction ability of the SVM and SVR techniques. To the best of our knowledge, this is the first time that this approach has been introduced and utilised in financial time series predictions (stocks indices predictions). The testing results of adding the quantisation factor in SVR and SVM prediction models are summarised in Table 3.11 and Table 3.12 respectively. A number of quantisation factor, between 10-100, are tested. Table 3.11 and 3.12 show the results of the prediction models SVR and SVM, and thus the factors that are generating the lowest MSE will be chosen to set up the prediction models. Furthermore, the tests were carried out on the training set of the three stock indices (FTSE100, S&P500 and

Table 3.11: Model Quantization factor selection results of the single SVR prediction model FTSE100, S&P500 and Nikkei225

QFactor	**FTSE100 Training** *MSE*	**S&P500 Training** *MSE*	**Nikkei225 Training** *MSE*
–	629937	171	41229512
10	2366	58	3233567
20	2384	58.3	267896
30	2403	58.4	81254
40	2410	58.5	57782
50	2414	58.56	53524
60	2418	58.58	52964
70	2419	58.61	52738
80	2417	58.64	52869
90	2422	58.65	53024
100	1912	55	34720

Nikkei225). First, for the FTSE100, the best $Qfactor$ in the SVR model is 100, thus the generated MSE is 1912 which is the lowest achievable score between the rest of the factors. Second, for S&P500 the best $Qfactor$ is 100. Third, for Nikkei225 the best $Qfactor$ is 100. Moreover, in the SVM prediction model the best $Qfactor$ is 100 for the FTSE100 data set, 10 is the best $Qfactor$ for S&P500 and 10 is the best $Qfactor$ for the Nikkei225 data set.

Table 3.12: Model Quantization factor selection results of the single SVM Trend prediction model FTSE100, S&P500 and Nikkei225

QFactor	**FTSE100 Training** *MSE*	**S&P500 Training** *MSE*	**Nikkei225 Training** *MSE*
–	2385	137	54899
10	2276	57	36032
20	2385	58.1	43248
30	2410	58.28	46637
40	2413	58.34	49073
50	2419	58.39	50401
60	2423	58.41	51446
70	2427	58.42	52299
80	2429	58.43	53011
90	2430	58.45	53419
100	833	58.46	53706

3.5.5 Naive Bayes

In accordance with the Nave Bayes classifier, classes are conditionally independent. Predicting the probability of data belonging to a particular class is the main process of the Bayesian classifier. Therefore, the concept of the Bayes theorem is used for probability prediction. Thus, the Bayes theorem is useful in that process, as it provides a way of calculating the posterior probability: $P(C|X)$, from $P(C)$, $P(C|X)$ and $P(X)$. It is stated by Byes' theorem that

$$P(C|X) = \frac{P(X|C)\,P(C)}{P(X)} \tag{3.33}$$

The posterior probability $P(C|X)$ gives the probability of hypothesis C being true given that event X has occurred. C is the hypothesis in this research, which is the probability of belonging to class price direction movement $\Delta Price_{t+1}$ and X is the test data sets. The occurrence condition probability $P(X|C)$ of the event X given hypothesis C is true. However, it can be estimated from the training data. A summary of how naive Bayesian classifier, or simple Bayesian classifier working is as follows:

In assumption of m classes $C_1, C_2, ...C_m$ the occurrence event of test data X is given. In this the test data will be classified into highest probability by Bayesian classifies. The Bayes theorem Equation 3.33 illustrates how the data is classified.

$$P(C_i|X) = \frac{P(X|C_i)\,P(C_i)}{P(X)} \tag{3.34}$$

Having many attributes in the given data sets $A_1, A_2, ...A_n$, can reflect on the computational time to compute $P(X|C_i)$. Therefore, the solution to reducing computation when evaluating $P(X|C_i)$ can be through making a class conditional independence using the nave assumption. In other words, presuming that the attributes values are conditionally independent of one another, given the class label of the tuple which means that there are no dependent relationships between attributes.

$$lP\left(X|C_i\right) = \prod_{k=1}^{n} P\left(x_k|C_i\right) = P\left(x_1|C_i\right) \times P\left(x_2|C_i\right) \times .... \times P\left(x_n|C_i\right) \tag{3.35}$$

x_k in Equation 3.35 presents the value of attribute A_k. Thus, whether it is categorical or continuous, the computation of $P\left(x_k|C_i\right)$ depends on. $P\left(x_k|C_i\right)$ is the number of observations of class C_i in the training data set, if A_k happened to be categorical, where, x_k value of A_k is divided by the number of observations of class C_i in the training data set. When a Gaussian distribution is fitted to the data, if A_k is continuous, the $P\left(x_k|C_i\right)$ value will be calculated as illustrated in Equation 3.36:

$$f\left(x,\mu,\sigma\right) = \frac{1}{\sqrt[\sigma]{2\pi}} e^{-(x-\mu)^{2\backslash 2\sigma^2}} \tag{3.36}$$

so,

$$P\left(x_k|C_i\right) = f\left(x_k, \mu_{c_i}, \sigma_{c_i}\right) \tag{3.37}$$

μ_{c_i} and σ_{c_i}in Equation 3.37 respectively are the mean and standard deviation of the A_k attribute value for training tuples of class C_i. In order to estimate $P\left(x_k|C_i\right)$, μ_{c_i} and σ_{c_i} value should be plugged in Equation 3.36, 3.37 together with x_k. To predict the label class X, $P\left(X|C_i\right) P\left(C_i\right)$ must be evaluated for each class C_i. However, if and only if Equation 3.38 happened, then the class label of observation X can be predicted as class C_i.

$$P\left(X|C_i\right) P\left(C_i\right) > P\left(X|C_i\right) P\left(C_j\right) for 1 \leq j \leq m; j \neq i. \tag{3.38}$$

There are many other proposals that utilise Bayesian classifiers, such as a theoretical justification for other classifiers, especially the ones that do not explicitly use Bayes theorem. As an example, for some algorithms, such as neural networks, curve fitting and nave Bayesian, under specific assumptions it can be demonstrated that these algorithms output the maximum posteriori hypothesis [206].

3.6 Results and Discussion

To further explain the presented artificial intelligence techniques (SVM, SVR, BPNN, RNN and Nave Bayes) and their capability for understanding the pattern in historical financial time series data, SVM,SVR, BPNN, RNN and Nave Bayes are applied to demonstrate their predictability of real-world financial time series. In addition, this chapter presents two direction prediction models.First, a dynamic model is built to predict the next day closing price using SVR, BNN and RNN. The input for this model is the historical closing price of FTSE100, S&P500 and Nikkei225. Second, a static model is built to predict the next day price direction of FTSE100, S&P500 and Nikkei225, whereas the input data for these models are the technical indicators for each data set. More details about these two models can be found in section 3.5. Moreover, the implementation platform was carried out via Matlab. The result of this experiment will be shown in the following section.

3.6.1 Dynamic Model Prediction Results

For evaluating the performance of the dynamic proposed prediction models using BNN, RNN and SVR, the daily closing prices of FTSE100, S&P and Nikkei225 are used in this thesis. Figure 3.13 in section 3.5 illustrates the single model approach.For predicting the FTSE 100, Nikkei 225 and S&P 500 next day closing prices, the FTSE 100, Nikkei 225 and S&P 500 historical closing prices are used as a prediction variables. Furthermore, subsections 3.2.1 and 3.4.2 explain the input features and the preprocessing stages of the data sets. The prediction results of the proposed models are compared to the traditional statistics model, called Auto-regressive (AR), and to the simple average combination model. The models parameters were determined in subsections 3.5.4.2, 3.5.4.1 and 3.5.2.3.

The FTSE 100, S&P 500 and Nikkei 225 closing price index prediction results using AR, SVR, RNN, BPNN and Simple average (SA) models are computed and listed in Table 3.13. Moreover, each model was run twenty times and the average of the results was calculated, along with the Standard Deviation (SD).

Table 3.13: The prediction result of training data sets for FTSE 100, S&P 500 and Nikkei 225 using SVR,RNN,AR and SA.

Index name	Models	MSE	RMSE	MAE	R	SD
FTSE100	AR	2429.91	49.29	31.93	0.99	——
	SVR	2424.77	49.24	31.89	0.99	2.36
	RNN	2416.33	49.15	32.11	0.99	2.51
	BPNN	2418	49.18	37.65	0.99	2.42
	SA	2417	49.17	37.49	0.92	——
S&P500	AR	58.31	7.63	3.43	0.99	——
	SVR	58.62	7.65	3.45	0.99	2.032
	RNN	57.87	7.60	3.45	0.99	2.021
	BPNN	57.57	7.58	3.46	0.99	2.050
	SA	57.81	7.60	3.45	0.95	——
Nikkei225	AR	55943.34	236.52	160.06	0.99	——
	SVR	53147.7	230.53	157.58	0.99	1.61
	RNN	53676.96	231.67	158.98	0.99	2.32
	BPNN	96473.96	283.66	175.35	0.99	3.64
	SA	67756.20	260	162.92	0.92	——

From Table 3.13 it can be found that the MSE, RMSE, MAE, R and SD values of the AR, SVR, RNN, BNN and SA models for the training FTSE 100, Nikkei 225 and S&P 500 data sets are relatively similar with fractional differences. This indicates that there is a small deviation between the prediction values of the utilised models. Thus it can be concluded that none of these models, in comparison with one another, have achieved the best results in terms of prediction error and accuracy.

The prediction results of the FTSE 100, Nikkei 225 and S&P 500 testing data sets are summarised in table 3.14. It can be observed from Table 3.14 that AR model outperformed the rest of proposed methods by achieving the smallest MSE, RMSE and MAE values for the FTSE 100 and S&P 500 testing data sets. However, the results of AR, SVR, RNN, BPNN and SA, in predicting Nikkei 225 testing data set, are very similar

Table 3.14: The prediction result of testing data sets for FTSE 100, S&P 500 and Nikkei 225 using SVR,RNN,AR and SA.

Index name	Models	MSE	RMSE	MAE	R	SD
FTSE100	AR	1815.46	42.60	31.95	0.95	——
	SVR	1912.04	43.72	34.43	0.95	1.48
	RNN	2248.38	47.23	38.10	0.95	2.34
	BPNN	2188	46.75	37.65	0.95	1.44
	SA	2033.77	45.09	35.99	0.95	——
S&P500	AR	240.56	15.51	11.68	0.96	——
	SVR	244.99	15.65	11.83	0.97	1.21
	RNN	336.28	18.33	14.42	0.96	2.05
	BPNN	602.97	22.85	18.12	0.94	3.45
	SA	267.22	16.34	12.21	0.96	——
Nikkei225	AR	34806.38	186.56	134.88	0.99	——
	SVR	34720.67	186.33	135.36	0.99	2.92
	RNN	34822.62	186.60	135.39	0.99	2.70
	BPNN	35547.91	188.50	137.028	0.99	4.00
	SA	34761.99	186.44	135.32	0.99	——

with a fractional differences. The cross correlation coefficient R of each data sets is presented in Table 3.14 and 3.13. It can be observed that the results of R for all the methods are drawing near to 1, which implies that the predicted values and the actual values do not deviate too much. Moreover, each method was run twenty times and the standard deviation was calculated. It can be observed that the results of SD for all models are relatively small, which implies that the models are not running randomly.

Figure 3.24 depicts the actual testing data set of the Nikkie 225 daily closing price and the predicted values from the AR, RNN, BPNN, SVR and SA. As can be observed from Figure 3.24 all of the utilised methods have generated good prediction results. The predicted values are also very close to the actual values and to one another.

Figure 3.25 presents the actual FTSE 100 closing price of the testing data set and the predicted values from AR, SVR, BPNN, RNN and SA. It also can be observed that the predicted obtained value from all the utilised models are very close to the actual values and to one another. The predicted values of AR, SVR, RNN, BNN and SA for S&P 500 closing price testing data set are illustrated in Figure 3.26. It can be observed from

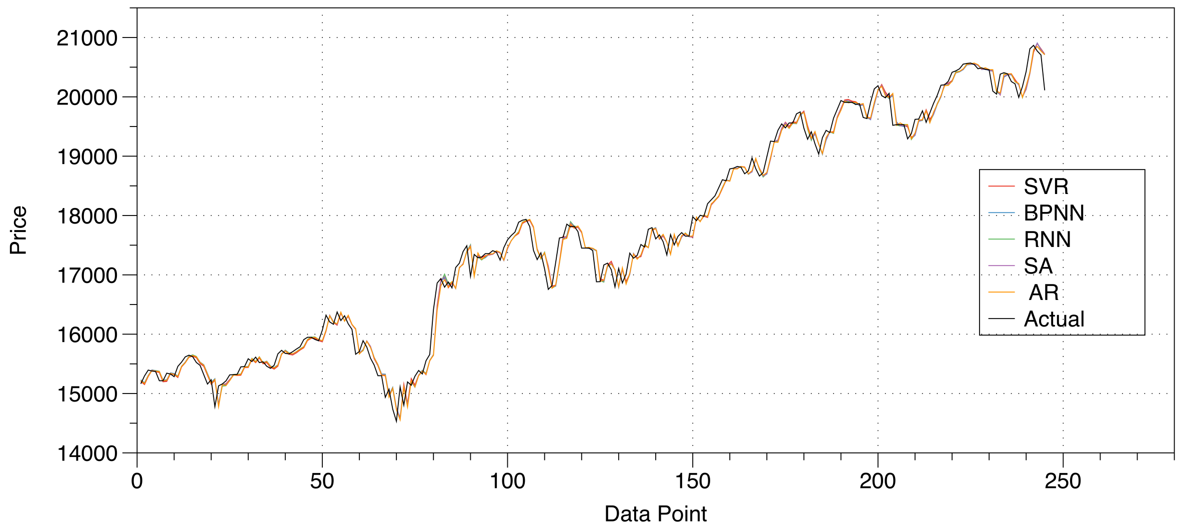

Figure 3.24: The actual Nikkei 225 closing price Index and its predicted values from AR, SVR, RNN, BPNN and SA.

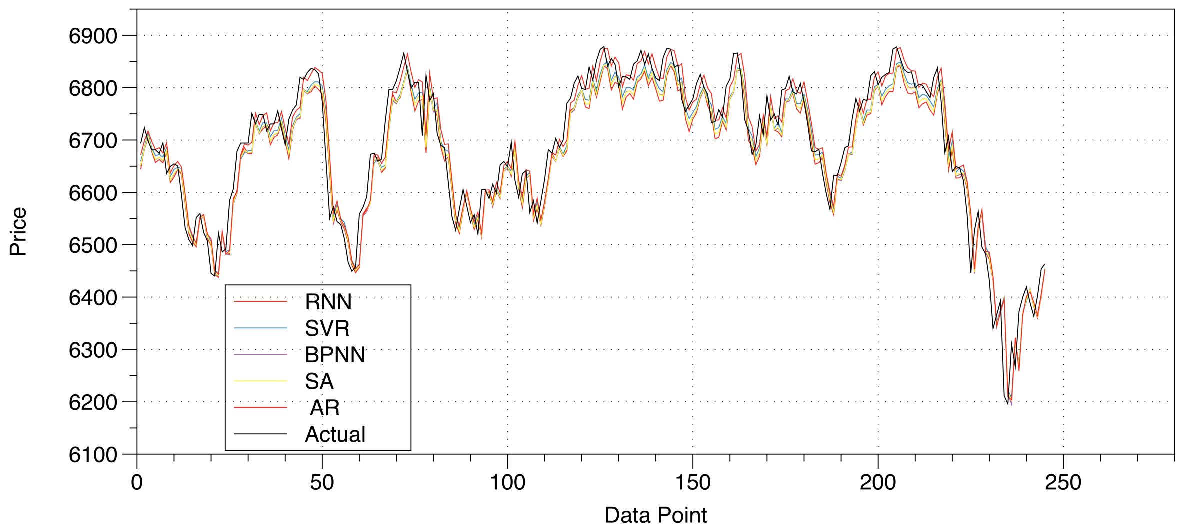

Figure 3.25: The actual FTSE 100 closing price Index and its predicted values from AR, SVR, RNN, BPNN and SA models.

Figure 3.26 that the AR, SVR, RNN and SA predicted values are close to the actual values than with the BPNN model.

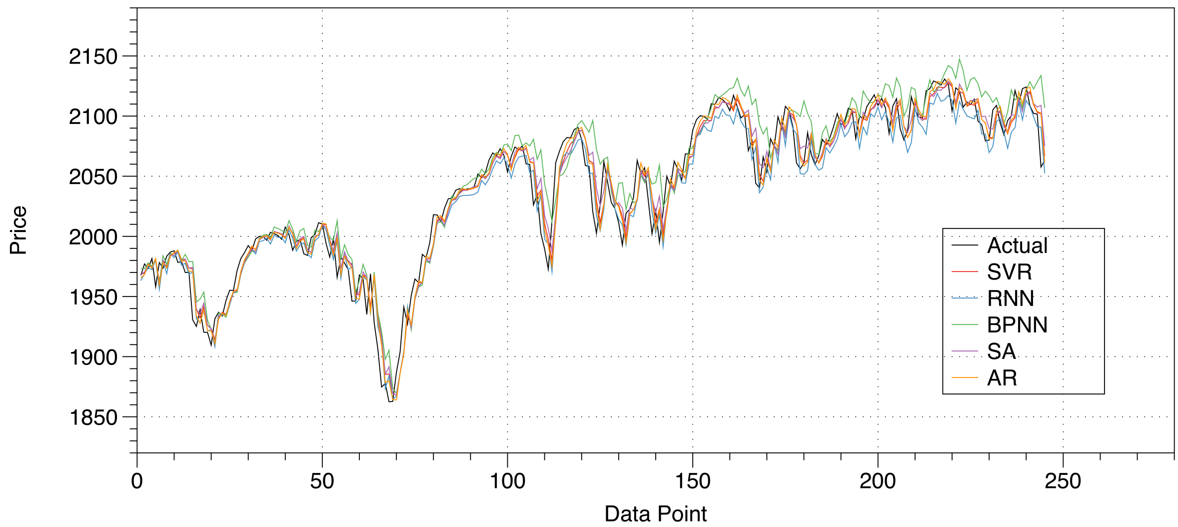

Figure 3.26: The actual S&P 500 closing price Index and its predicted values from AR, SVR, RNN, BPNN and SA models.

3.6.2 Static Model Prediction Results

This section presents the results of the static model. Thus, the next day closing prices of the FTSE 100, S&P 500 and Nikkei 255 are predicted using the RNN, SVM, AR, Naive Bayes and SA models. Eight technical indicators were used as an input in the prediction models; more details on the models features are exhibited in subsection 3.2.2. Moreover, models parameters were also determined and obtained by the analytic approach mentioned in subsection 3.5.2.3 for the BNN model, and subsection 3.5.4.1 for the SVM model. The testing results of the utilised models are summarised in Table 3.15 for the training data sets and Table 3.16 for the testing data sets.

The FTSE 100, Nikkei 225 and S&P 500 direction closing price predicted results using AR, SVM, RNN, Naive Bayes and SA are computed and listed in Table 3.15.

From Table 3.15, it can be found that MSE, RMSE and MAE of the Nive Bayes model are respectively 1409.31, 37.54 and 21.46. It can be observed that these values are smaller than the values of the rest of the models for FTSE 100 direction closing price. In addition, it indicates that there is a smaller deviation between the models prediction error. For

Table 3.15: The Trend prediction result of training data sets for FTSE 100, S&P 500 and Nikkei 225 using SVM,RNN,AR, Naive Bayes and SA.

Index name	Models	MSE	RMSE	MAE	R	SD
FTSE100	AR	2434.81	49.34	31.87	2.06	——
	SVM	2431.22	49.30	31.82	0.18	1.85
	RNN	2416.34	49.15	31.85	0.02	2.23
	Naive Bayes	1409.31	37.54	21.46	0.73	2.92
	SA	2086.33	45.6	27.26	0.63	——
S&P500	AR	58.35	7.63	3.44	0.01	——
	SVM	58.46	7.64	3.42	0.06	2.082
	RNN	56.22	7.49	3.45	0.16	2.17
	Naive Bayes	9.38	3.06	0.97	0.92	2.57
	SA	42.02	6.48	3.28	0.56	——
Nikkei225	AR	55945.01	236.52	159.93	0.05	——
	SVM	53706.66	231.74	157.43	0.20	1.73
	RNN	53024.89	230.21	158.41	0.20	5.40
	Naive Bayes	36068.98	189.91	87.93	0.72	2.74
	SA	36779.62	191.78	140.69	0.64	——

the prediction of the S&P 500 closing price direction, the prediction error results indicate that the Nave Bayes model has outperformed the rest of the models. It can also be observed from Table 3.15 that when predicting the Nikkei stock index direction price, Nave Bayes obtained the smallest error. Thus, it can be concluded that the Nave Bayes model provides better prediction results than the AR, SVM, RNN and SA in terms of prediction error.

The prediction results for the testing data sets of FTSE 100, S&P 500 and Nikkei 225 are illustrated in Table 3.16. It can be observed from Table 3.16 that none of the utilised models provide good prediction results. The predicted values of the testing data sets of FTSE 100, S&P 500 and Nikkei 225 are very poor and there is a huge deviation between the actual and predicted values, as the results of R of each model show.

Thus, it can be concluded that the single model approach has failed to performed sufficiently and a more complex approach is needed in order to enhance the prediction results and improve the quality of the prediction process.

Figure 3.27 depicts the Actual S&P 500 price direction values and predicted values from

Table 3.16: The Trend prediction result of testing data sets for FTSE 100, S&P 500 and Nikkei 225 using SVM,RNN,AR, Naive Bayes and SA.

Index name	Models	MSE	RMSE	MAE	R	SD
FTSE100	AR	1789.25	42.29	31.68	0.01	——
	SVM	1768.34	42.05	31.95	0.12	2.26
	RNN	1794.85	42.36	31.97	0.10	1.58
	Naive Bayes	5074	71.23	56.64	0.09	2.17
	SA	2043.15	45.20	35.14	0.09	——
S&P500	AR	237	15.40	11.66	0.02	——
	SVM	245.91	15.65	11.84	0.09	1.72
	RNN	230.56	15.15	11.36	0.01	1.92
	Naive Bayes	328.72	18.13	13.80	0.10	1.05
	SA	257.63	16.05	121.15	0.03	——
Nikkei225	AR	34292.98	185.18	133.48	0.06	——
	SVM	34054.72	184.53	132.97	0.02	2.05
	RNN	34796.2	186.52	135.69	0.08	1.92
	Naive Bayes	82789	287.7319	230.47	0.06	3.74
	SA	37179.01	192.81	143.23	0.08	——

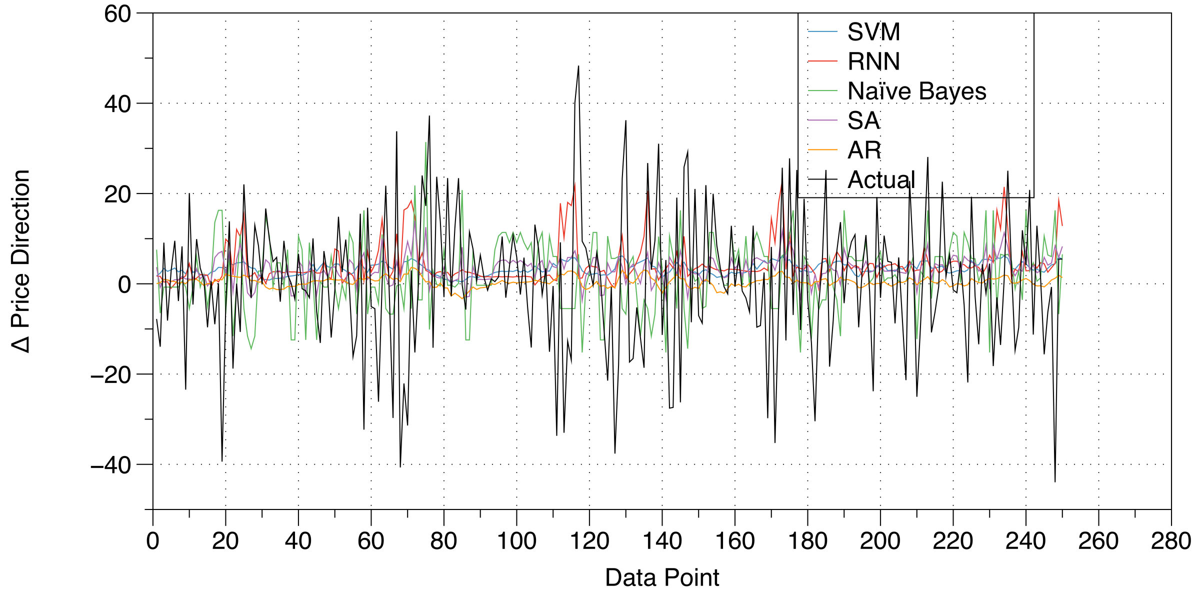

Figure 3.27: The actual S&P 500 closing direction price Index and its predicted values from AR, SVM, RNN, Naive Bayes and SA models.

the SVM, RNN, Naive Bayes, AR and SA models. The predicted values of the Naive Bayes Model are closer to the actual values in comparison to the values of the other models. Figures 3.28 and 3.29 illustrate the Nikkei 225, and FTSE 100 predicted values of the utilised models and the actual values. It can be also observed that in both data

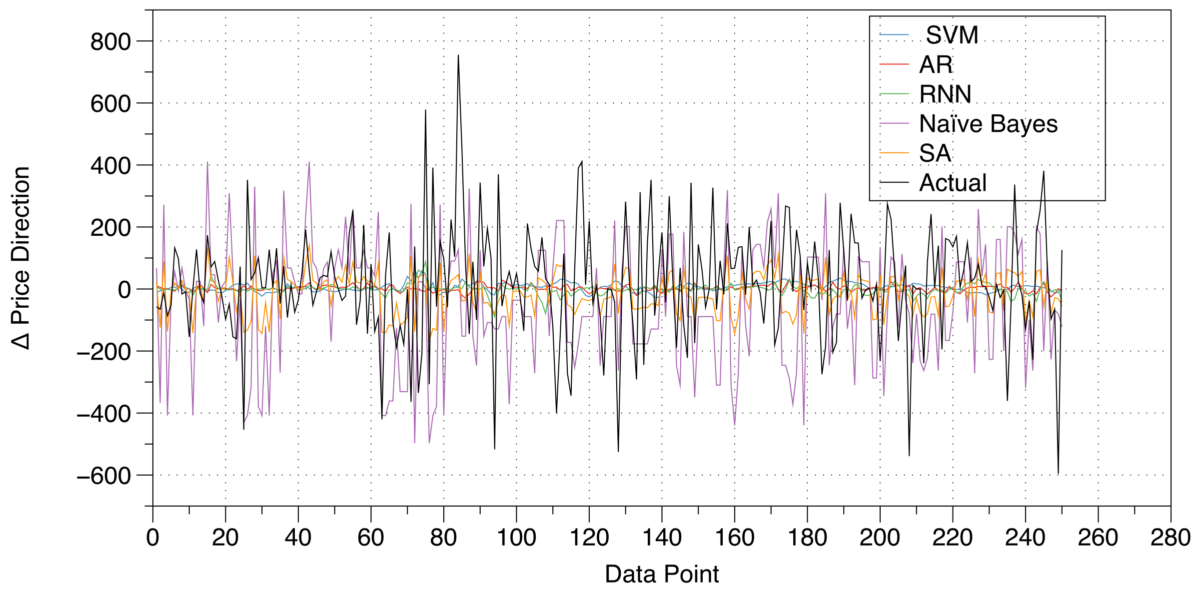

Figure 3.28: The actual Nikkei 225 closing direction price Index and its predicted values from AR, SVM, RNN, Naive Bayes and SA models.

sets, the Naive Bayes predicted values are closer to the actual value than those of the SVM, AR, RNN and SA models.

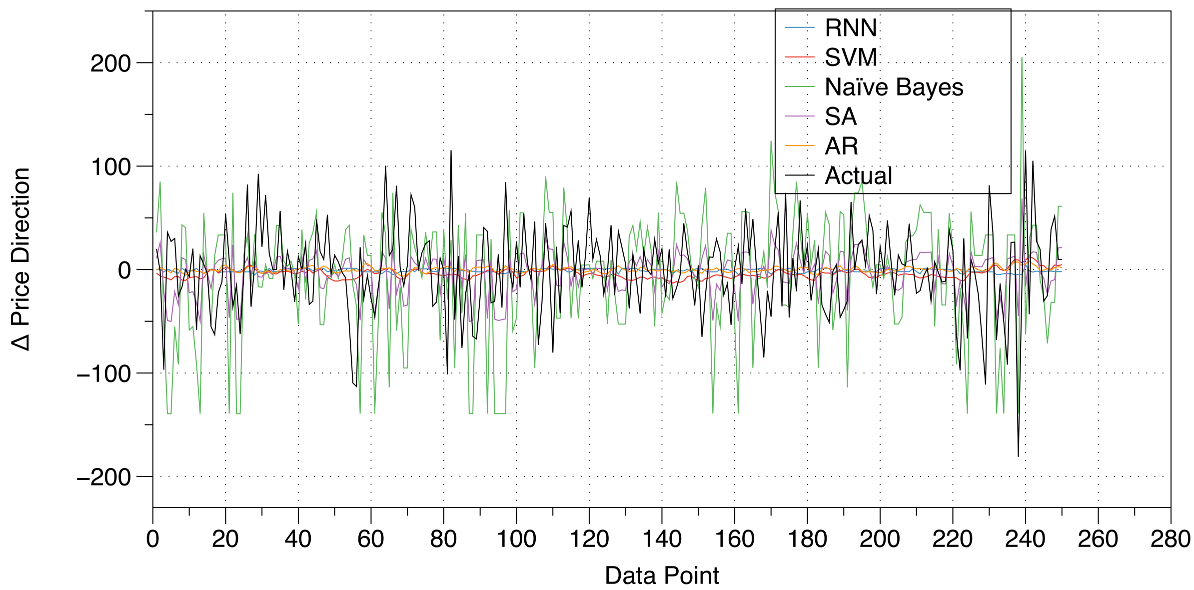

Figure 3.29: The actual FTSE 100 closing direction price Index and its predicted values from AR, SVM, RNN, Naive Bayes and SA models.

3.7 Summary

In this chapter, the main focus was on building single approach models to predict real world financial time series. Furthermore, the performance of single models, Support Vector Regression (SVR), Recurrent Neural Network (RNN), Back Propagation Neural Network (BPNN), Auto-regressive (AR), Simple Average (SA) and Support Vector Machine (SVM) were investigated. Thus, this chapter carried out two prediction directions. Firstly, a dynamic model was established to predict the FTSE 100, Nikkei 225 and S&P 500 next day closing prices, using the historical time series data and previous closing prices as inputs. SVR, BPNN and RNN were applied as prediction models and a simple combination method, a simple average, is used. The results of these models were compared with the Autoregressive benchmark model. Secondly, a static model was built to predict stock index price direction. Eight technical indicators were used as model inputs. RNN, SVM and Nave Bayes are the techniques used to predict the FTSE 100, Nikkei 225 and S&P 500 stock index next day price direction. The results of these models are compared with the benchmark model AR.

One of the contributions of this work is that the training data sets in the dynamic models are distributed randomly with respect to the time series, which is found to be helpful in providing more predictability than with a normal distribution of the data. This step of distributing the training data sets can be explained by the nature of financial time series; that is, the fluctuation of closing prices over time is very high and this step can help the model to learn the pattern in the training data set and perform very well in the testing data set.

Quantisation factors were proposed for the SVR and SVM prediction models in order to improve their prediction performance. In addition, procedures for extracting the target data sets in the static models were also proposed in this chapter.

The empirical results were obtained by applying the dynamic and static models to predict the next day closing price and next day closing price direction of the FTSE 100, Nikkei 225 and S&P 500. SVR, RNN,BP and the combination simple average were compared with the benchmark model AR. However, it was observed from the results that the predicted values of the models were very similar, which indicated that there are small deviations between the models prediction values. Hence, the outputs of the adopted models were not satisfactory, and there is a need for more accurate methods. Furthermore, and in order to tackle this problem, the next chapter will investigate and implement a new approach to minimise the errors and provide better predictability models.

Chapter 4

A New Hybrid Financial Time Series Prediction Model Based on EEMD-BPNN-SVR-RNN-GA

4.1 Introduction

Due to financial time series' inherent characteristics such as being non-linear, non-stationary, and noisy, with a high degree of uncertainty and hidden relationships, single artificial intelligence and other conventional techniques have failed to capture its non-stationary property and accurately describe its moving tendency. Therefore, research and market participants have paid a great deal of attention to tackling such problems. Thus, various models' architecture and new algorithms have been introduced and developed in the literature to alleviate the influence of noise [262]. However, the noise characteristic can refer to the unavailability of information, which affects the behaviour of financial markets and past and future captured prices. The information which is not included in the prediction models are considered as noise, while the non-stationery characteristics imply that the financial time series distributions fluctuate over time. Therefore, the prediction of stock prices and detecting their noise is considered a very challenging and difficult financial topic.

In the previous chapter, single approach models were used to predict financial time series; however, the results were not satisfactory. Therefore, this chapter adopts a novel three steps hybrid intelligent prediction model based on Neural Network Back Propagation (BPNN), Recurrent Neural Network (RNN), Support Vector Machine (RNN), Genetic Algorithm (GA) and ensemble empirical mode decomposition (EEMD). This chapter is structured as follows: in Section 4.2 the overall process of the hybrid model, prediction steps and prediction techniques (EEMD, RNN, SVR, BPNN and GA) will be presented. Section 4.3 presents the experimental results of the proposed hybrid model. Moreover, this chapter is representing stage two of phase one as the holistic framework of the prediction models illustrated in Figure 3.10 in Chapter 3.

4.2 Hybrid modelling and prediction procedures

According to Li and Ma [166], traditional linear methods and the majority of sophisticated non-linear machine learning models have failed to capture the complexity and the non-linearities that exist in financial time series, particularly during periods of uncertainty such the credit crisis in 2008 [241]. Financial time series characteristics imply that the statistical distributions of the time series can change over time. The cause of these changes may be caused by economic fluctuations, or political and environmental events [110] [3]. As a result, it has been verified in the literature that no single method or model works well to capture financial time series characteristics properly and accurately describe its moving tendency, which leads to different and inaccurate financial time series prediction results [290] [43]. In order to address these issues, a novel three steps hybrid model is proposed to alleviate the influence of noise, utilising complete ensemble empirical mode decomposition with adaptive noise, Back Propagation neural network BPNN, Recurrent Neural Network RNN, Support Vector Regression SVR and Genetic Algorithm GA.

In the literature there are many methods for combining artificial intelligence techniques with EMD used in the area of stock index prediction [283] [257] [56] [103]. However, to the researchers knowledge there are no applications of artificial intelligence techniques combined with EEMD modelling and predicting stock indices. This thesis introduces EEMD into SVR, BPNN and RNN in order to enhance the modelling capability, for the first time in the literature. Figure 4.1 illustrates the overall process of the proposed approach for stock index prediction.

The following points illustrate the implementation steps of the new hybrid model:

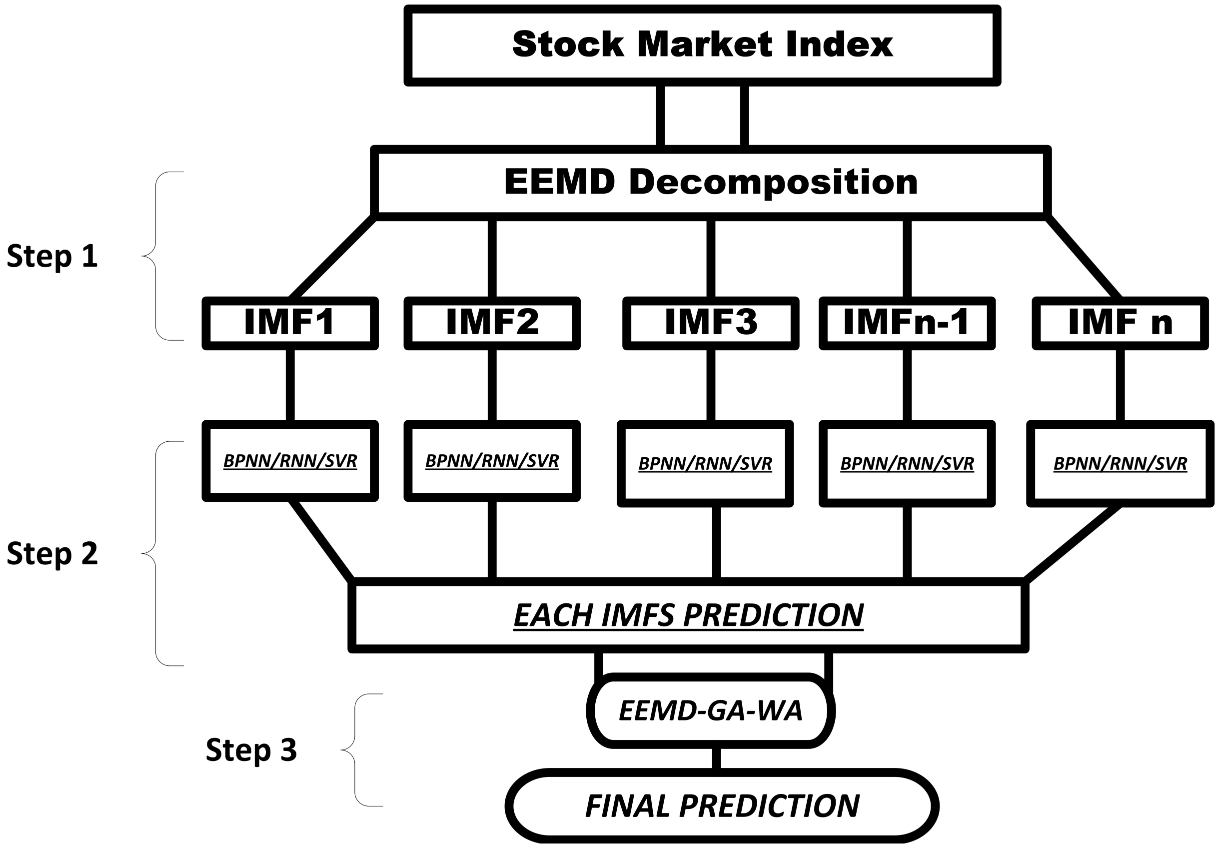

Figure 4.1: The overall process of EEMD based SVR, BPNN and RNN hybrid methodology.

- In step one, the proposed EEMD has been used to decompose the original stock market index to several intrinsic mode functions.Subsection 4.2.1 illustrates the data decomposition process by the proposed EEMD technique.

- In the second step BPNN, RNN and SVR have been used to predict IMFs. In addition, the RNN, SVR and RNN methods and model topologies are illustrated (see Chapter3).
- In step three, the final prediction value can be obtained through the sum of the previous steps final predicted result. The weighted average combination function was utilised to combine the different EEMD-RNN, EEMD-BPNN and EEMD-SVR methods. An optimiser genetic algorithm has been used to determine the weight of the combiner; more details on this step are available in Subsection 4.2.2.

Applying the proposed method can more fully capture the local fluctuations of the original data. Therefore, the proposed method decomposed the original data into different IMFs where each IMF has simpler frequency components and stronger regularity compared with the original data. Furthermore, EEMD reduced the complexity and improved the efficiency and accuracy of the prediction model.

4.2.1 Ensemble Empirical Mode Decomposition

4.2.1.1 Empirical Mode Decomposition

Empirical mode decomposition (EMD) is a method that is characterised by the ability to analyse non-linear and non-stationary signals. EEMD performed the EMD over an ensemble of the signal plus Gaussian white noise in order to prevent any mixing problems by populating the whole time-frequency space and to benefit from the dyadic filter bank behaviour of the EMD [256] [128] [273]. The EMD transformation mechanism of the signal $x(t)$ is defined by the decomposition of the signal into a small number of intrinsic mode functions (IMFs) or modes. There are two conditions the signal must satisfy in order to be considered as an IMF: (i) the number of extreme values and the number of zero crossings in the whole data set must either be equal or differ at most by one; (ii) everywhere the mean value of the upper and lower envelope must be zero. As the above point defines the IMFs, the decomposition processes of the signal are explained by the following steps [128]:

- First step: for any signal $x(t)$ the local maxima and minima will be identified. All maxima and maxima will be connected to produce an upper and lower envelop by a cubic spline curve. m is defined as the mean value of upper and lower envelopes and the difference between $x(t)$ and m defined as h:

$$h = x(t) - m \tag{4.1}$$

- Second step: h Will be taken as a new original signal $x(t)$ and then the operation will be repeated in step $(a)\,k$ times until h is an IMF. To judge whether h is an IMF or not the below function is considered as a termination criterion:

$$D_k = \frac{\sum_{t=0}^{T} \left|h_{(K-1)} h_K(t)\right|^2}{\sum_{t=0}^{T} \left|h_{K-1}(t)\right|^2} \tag{4.2}$$

  In Equation(4.2), when D_K is smaller than a predetermined value, h_K can be viewed as an IMF. The first IMF will designate as $c_1 = h_K$.

- Third step: When c_1 is determined, the residue r_1 can be obtained by separating c_1 from the rest of the data as follows:

$$r_1 = x(t) - c_1 \tag{4.3}$$

After obtaining r_1, the operation will be repeated in steps (1) and (2) as r_1 is the new signal $x(t)$, until obtaining the second $IMFc_2$. Time j will be considered until r_j is smaller than a predetermined value or r_j becomes a monotone function in order to obtain all the IMFs. A series of IMFs and a residue r in the end will be obtained after applying the above steps.

EMD has many drawbacks; mode mixing is one of the main drawbacks which usually cause intermittency in the analysing signal. This drawback can imply either a dingle IMF consisting of signals of dramatically disparate scales or a signal of the same scale appearing in different IMF components. To prevent such problems, a new noise-assisted data analysis method EEMD was proposed. In EEMD, the ensemble of the trails' mean is the true IMF component. A decomposition result of the signal plus a white noise of finite amplitude consisting of each trail [273].

Several white noise studies have shown that the EMD method is an effective self-adaptive dyadic filter bank when applied to white noise, demonstrating the benefits of the EEMD method [90] [272]. In EEMD, IMF's components are defined as the mean of the corresponding IMFs obtained via EMD over an ensemble of trials, which are generated by adding different realization of white noise of finite variance to the original signal $x\,[n]$. The below point describes the EEMD algorithm [244]:

- Add white noise series to the original signal. In other words generate $x^i\,[n] = x\,[n] + w^i\,[n]$, where $w^n\,[n]\,(i = 1, ..., I)$ are different realization of Gaussian noise.
- Decompose the signal with added white noise into IMFs by EMD. Each $x^i\,[n]\,(i = 1, ...I)$, is fully decomposed by EMD getting their modes $IMF_i^K\,[n]$, where $= 1....K$ indicates the modes.
- Repeating the previous two steps for a certain number of times with different white noise each time and obtain the corresponding IMF components of the decomposition.
- Calculate the mean of all the corresponding IMF components and take the mean as the final result for each IMF. In another words assign IMF_K as the $K - th$

mode of $x[n]$, obtained as the average of the corresponding $IMF_K^i : IMF_K[n] = \frac{1}{I}\sum_{i=1}^{I} IMF_i^K[n]$.

4.2.1.2 Ensemble Empirical Mode Decomposition with Adaptive Noise

The above point explained how, when observing operations in the EMD, each is $x^i[n]$ decomposed independently from other realizations and so for each one a residue $r_k^i[n] = r_{k-1}^i[n] - IMF_k^i[n]$ is obtained. In this proposed method, the decomposition mode will be noted as $\widetilde{IMF_K}$ and the unique first residue propose as:

$$r_1[n] = x[n] - \widetilde{IMF_1}[n] \tag{4.4}$$

In Equation (4.4) $\widetilde{IMF_1}[n]$ is obtained in the same way as in EEMD. After that, the first EMD is used over an ensemble of r_n plus different realization of a given noise to obtain $\widetilde{IMF_2}$ by averaging. Then, the next residue is defined as: $r_2 = r_1[n] - \widetilde{IMF_2}[n]$. Until the stopping criterion is reached, the previous steps with the rest of modes will be continued. This method can be described as following algorithm if $x[n]$ is the target data.

- To obtain the first modes and compute, decompose by $EMD\ I$ realization $x[n] + \varepsilon_0 w^i[n]$.

  $\widetilde{IMF_1}[n] = \frac{1}{i}\sum_{i=1}^{I} IMF_1^i[n] = \overline{IMF_1}[n]$.

- In the first stage ($k = 1$)calculate the first residue as in Equation (4.4): $r_1[n] = x[n] - \widetilde{IMF_1}[n]$.

- Decompose realisations $r_1[n] + \varepsilon_1 E_1(w^i[n]), i = 1...I$ until their first EMD mode and define the second mode:

  $\widetilde{IMF_2}[n] = \frac{1}{I}\sum_{i=1}^{I} E_1(r_1[n] + \varepsilon_1 E_1(w^i[n]))$

- For $k = 2...K$ calculate the $k - th$ residue:

$$r_k[n] = r_{(k-1)}[n] - \widetilde{IMF_k}[n] \tag{4.5}$$

- Decompose realizations $r_k[n] + \varepsilon_k E_k(w^i[n]), i = 1...I$, Until their first EMD mode and define the $(k+1) - th$ mode as:

$$\widetilde{IMF_{(k+1)}}[n] = \frac{1}{I}\sum_{i=1}^{I} E_1\left(r_k[n] + \varepsilon_k E_k\left(w^i[n]\right)\right) \tag{4.6}$$

- Then for next k go to step 4. Steps 4 to 6 are performed until it is no longer feasible to be decomposed the obtained residue. The final residue satisfies:

$$R[n] = x[n] - \sum_{k=1}^{K} \widetilde{IMF_k}, \tag{4.7}$$

When K the total number of modes is complete. Therefore, the given signal $x(n)$ can be expressed as:

$$x[n] = \sum_{k=1}^{K} \widetilde{IMF_k} + R[n]. \tag{4.8}$$

Equation (4.8) makes the proposed decomposition complete and provides an exact reconstruction of the original data. At each stage on observable selection of SNR is allowed by ε_i coefficients. It has been suggested by Wu and Hung [273] that regarding the amplitude of the added noise, small amplitude values for data dominated should be used by high-frequency signals and vice versa.

4.2.2 Hybrid Combination Model EEMD-GA-WA

Combining different prediction techniques has been investigated widely in the literature. In short-range predictions, combining the various techniques is more useful according to [293], [12], while according to the Timmermann study, using a simple average may work as well as more sophisticated approaches. However, using one model can produce more accurate predictions than any other methods. Therefore, simple averages would not be sufficient in such cases [242]. Compared with different prediction models, the hybrid prediction method is based on a certain linear combination. The assumption for the actual value in period t by model i is $f_{it}\,(i = 1, 2, ...m)$, and the corresponding prediction error will be $e_{it} = y_t - f_{it}$. The weight vector will be $W = [w_1, w_2, ...w_m]^T$ .. Then in the hybrid model the predicted value is computed as follows [104] [61]:

$$\hat{y}_t = \sum_{i=1}^{m} w_i f_{it}\,(t = 1, 2, ...n) \tag{4.9}$$

$$\sum_{i=1}^{m} w_i = 1 \tag{4.10}$$

Equation 4.9 can be expressed in another from:

$$\begin{aligned} &\hat{y} = FW \\ &where \quad \hat{y} = [\hat{y_1}, \hat{y_2}, ...\hat{y_n}]^T, F = [f_{it}]_{n\times m} \end{aligned} \tag{4.11}$$

The error for the prediction model can be formed as in Equation 4.12.

$$\begin{gathered} e_t = y_t - \hat{y}_t = \textstyle\sum_{i=1}^{m} w_i y_t - \sum_{i=1}^{m} w_i f_{it} = \\ \textstyle\sum_{i=1}^{m} w_1 w_i\,(y_t - f_{it}) = \sum_{i=1}^{m} w_i e_{it} \end{gathered} \tag{4.12}$$

This research proposed a hybrid model that combines EEMD-SVR, EEMD-BPNN and EEMD-RNN.

$$\hat{Y}_{combined_t} = \frac{w_1\hat{Y}_{EEMD-SVR} + w_2\hat{Y}_{EEMD-BPNN} + w_3\hat{Y}_{EEMD-RNN}}{(w_1 + w_2 + w_3)} \tag{4.13}$$

The prediction values in period t are $\hat{Y}_{combined_t}$, $\hat{Y}_{EEMD-SVR}$, $\hat{Y}_{EEMD-BPNN}$ and $\hat{Y}_{EEMD-RNN}$ for the hybrid, EEMD-RNN, EEMD-BPNN and EMMD-SVR models, where the assigned weights are w_1, w_2, w_3 respectively, with $\sum_{i=1}^{3} w_i = 1.0 \leq w_i \leq 1$.

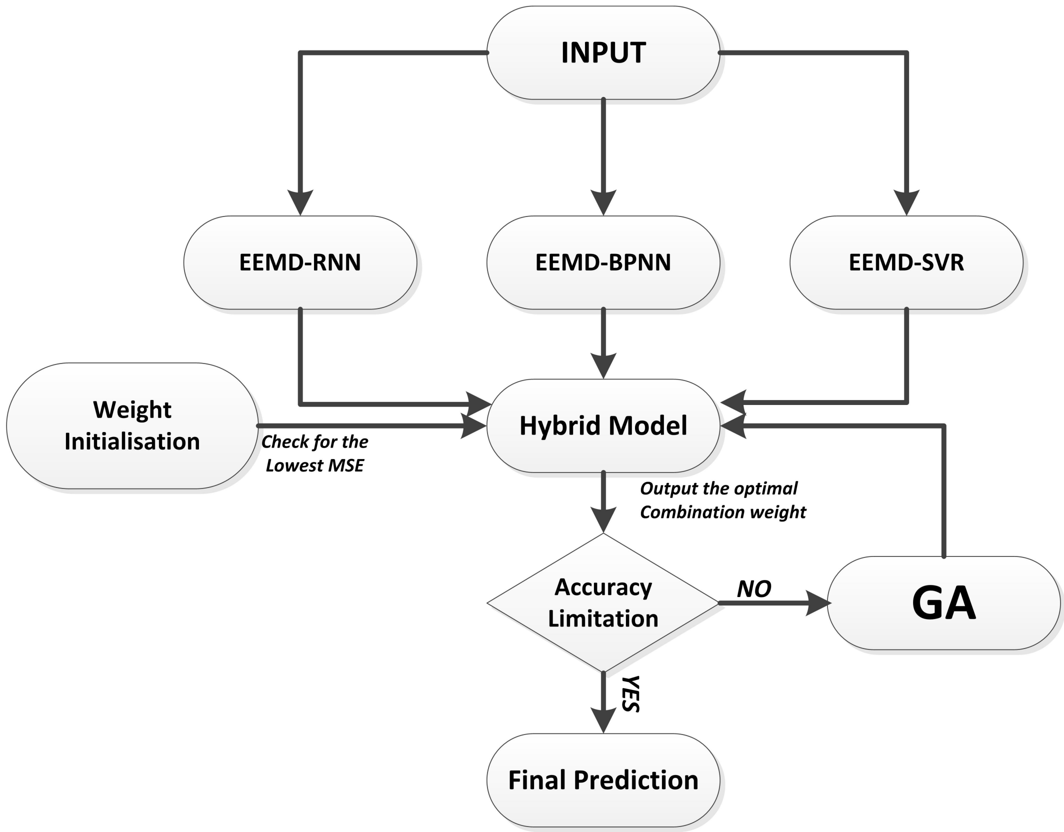

Figure 4.2: The flow chart of the EEMD-GA-WA hybrid model.

The most important step in developing a hybrid prediction model is to determine the perfect weight for each individual model. Setting $w_1 = w_2 = w_3 = 1/3$ in Equation 4.13 is the simplest combination method for the three prediction models. Nevertheless, in many cases equal weights cannot achieve the best prediction result. Therefore, this chapter adopts a hybrid approach utilising the Genetic Algorithm GA as an optimiser to determine the optimal weight for each prediction model. Figure 4.2 illustrates the architecture of EEMD-GA-WA hybrid model.

4.2.2.1 Genetic Algorithm

The GA is a well-known tool in computational methods modelled on a Darwinian selection mechanism. GA principles were proposed by Holland [117], and developed by Goldbreg [101] and Koza [151]. Thus, the main purpose of using such an algorithm is to solve optimisation problems such as determining the optimal weights for the proposed hybrid model in this research. In comparison with other conventional optimisation methods, GA has many differences, which contribute in making GA more efficient in searching for the optimal solution. The following points exhibit those differences [271].

- Computing the strings in GA algorithms is done by encoding and decoding discrete points than using the original parameter values. Thus, GAs tackle problems associated with discontinuity or non-differentiability functions, where traditional calculus methods have failed to work. Therefore, due to the adaptation of the binary strings, such characteristics allow GAs to better compute logic operations.

- The prior information is not important, and thus there is no need for such information as the primary population is randomly generated. GA uses a fitness function in order to evaluate the suggested solution.

- Initialisation, selection and reproduction, whether crossover or mutation, are what GA depends on in the searching process, which involves random factors. As a result, the searching process in GAs for every single execution will be stand-alone, even the ones under identical parameter settings, which perhaps may affect the results.

4.3 Experimental Results

In this chapter a new hybrid model based on EEMD, SVR, BPNN, RNN and EEMD-GA-WA methods is constructed to predict the FTSE 100, S&P 500 and Nikkei 225 next day closing prices. To demonstrate the validity of the proposed methods, this section compares the results of the single approach models from Chapter 3 with the results of the hybrid models from this chapter. In the suggested model there are four steps involved in stock index prediction models. Firstly, the original stock index data sets are decomposed by EEMD. Secondly SVR, RNN and BPNN are used to predict each IMF and the residue. Thirdly, in order to obtain the prediction results for the original data sets the prediction results of each IMF and the residue must be combined. Fourthly, the predicting results of the three models by the proposed method EEMD-GA-WA must be combined with the predicting results of the hybrid model. Section 4.2 illustrates each step in detail.

The FTSE 100 data set is decomposed by EEMD into 13 different IMFs and IMF 14 is the residue, as shown in Figure 4.3.

In this study the standard deviation of the noise added is 0.2 and the ensemble size is 500. Figure 4.3 illustrates the different IMFs with different frequencies. The frequencies from IMF 1 to IMF 6 are much higher and they mainly reflect the randomness of the information of the original FTSE 100 data set. However, the periodic trends of IMF 7 to IMF 11 are obvious and they are called the periodic components of the original data set. IMF 12, IMF 13 and IMF 14 are called the trend components. Figure 4.4 presents the S&P 500 data set after being decomposed by EEMD.

As Figure 4.4 illustrates, different IMFs have different frequencies. From IMF 1 to IMF 9 the frequencies of these IMFs are much higher and they reflect the randomness of the information of the original S&P 500 data set. IMF 10 to IMF 11 periodic trends are obvious and these are called periodic components of the original S&P 500 data set. Moreover, IMF 12, 13 and IMF 14 are called the trend components.

The decomposed Nikkei 225 by EEMD is shown in Figure 4.5.

As shown in Figure 4.5 for the decomposed Nikkei 225, different IMFs have different frequencies. Thus, IMF 1 to IMF 8 frequencies are much higher and they mainly reflect

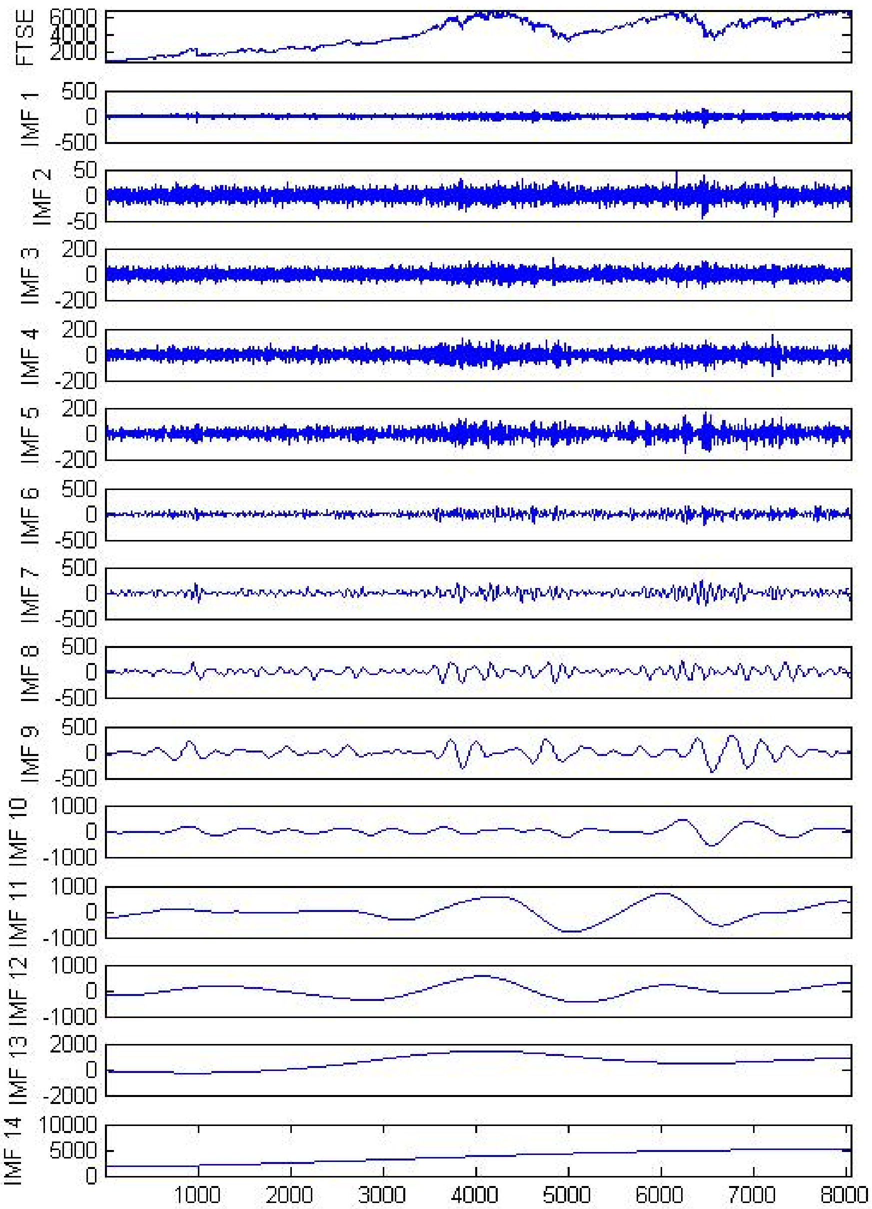

Figure 4.3: The decomposition of FTSE 100 index closing price

the randomness of the information of the original Nikkei 225 data set. IMF 9 and IMF 10 obviously show the periodic trend so they are called the periodic components. The rest (IMF 11 to IMF 14) are called the trend components.

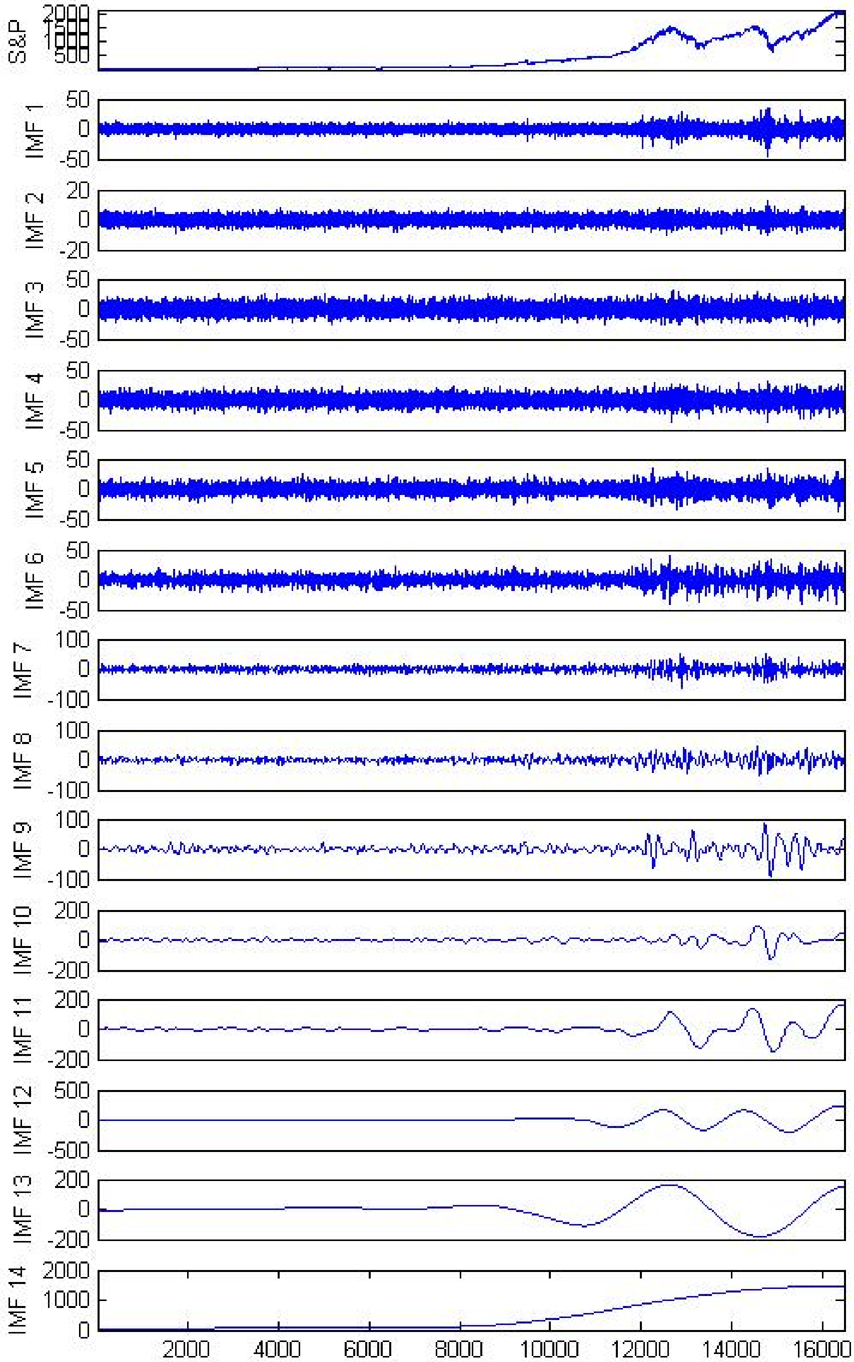

Figure 4.4: The decomposition of S&P 500 index closing price

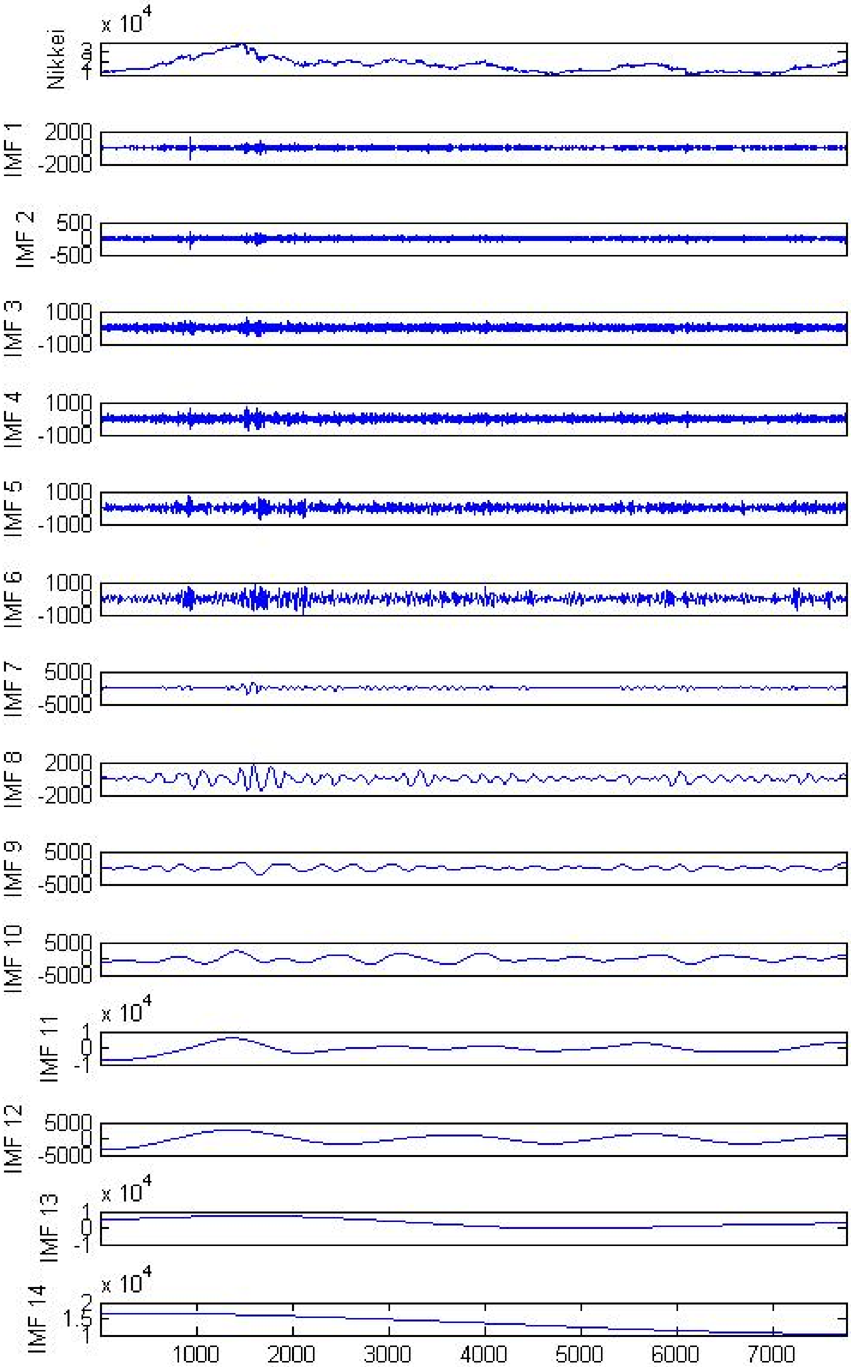

Figure 4.5: The decomposition of Nikkei 225 index closing price

After FTSE 100, Nikkei 225 and S&P 500 are decomposed by EEMD, the prediction problem is changed into predicting each IMF and the residue for the data sets. BPNN, RNN and SVR are used for the predicting of each IMF and the residue. The FTSE 100 is decomposed into 14 IMFs and in each IMF there are 8038 observations. The first 7788 observations are selected as a training data set and the remaining (250) are selected as the testing data set. For S&P 500 there are 14 IMFs and in each IMF there are 16480 observations. The first 16230 observations are set as the training data set and the remaining (250) observations are selected as the testing data set. And finally for the Nikkei 225 there are 14 IMFs and in each IMF there are 7758 observations. The first 7508 are selected as the training data set and the remaining 250 observations are selected as the testing data set. Moreover, the same approach which was followed in Chapter 3 Section 3.4.1 in preparing the in-sample out-of-sample data sets is adopted in this section. Prediction for the FTSE 100, S&P 500 and Nikkei 225 are obtained by simply adding up the prediction results of each IMF and the residue. The simple average (SA) is also used in this chapter as a benchmark combination method; more information about SA is available in chapter 3 section 3.5.1. In addition, the proposed model EEMD-GA-WA prediction results were obtained by combining the final results of the three models as explained in Section 4.2.2.

For the FTSE 100, the predicted results of the methods based on EEMD-RNN, EEMD-BPNN, EEMD-SA and EEMD-GA-WA are shown in Figure 4.6. It can be observed from Figure 4.6 that the predicted values obtained from the proposed models EEMD-RNN, EEMD-BPNN, EEMD-SA and EEMD-GA-WA are closer to the actual price than the predicted values of the single models Figure 3.25. Moreover, the EEMD-GA-WA predicted values are closer to the actual values than with the other models.

The actual S&P 500 closing price values and predicted values of the EEMD-SVR, EEMD-RNN, EEMD-BPNN, EEMD-SA and EEMD-GA-WA models are presented in Figure 4.7.

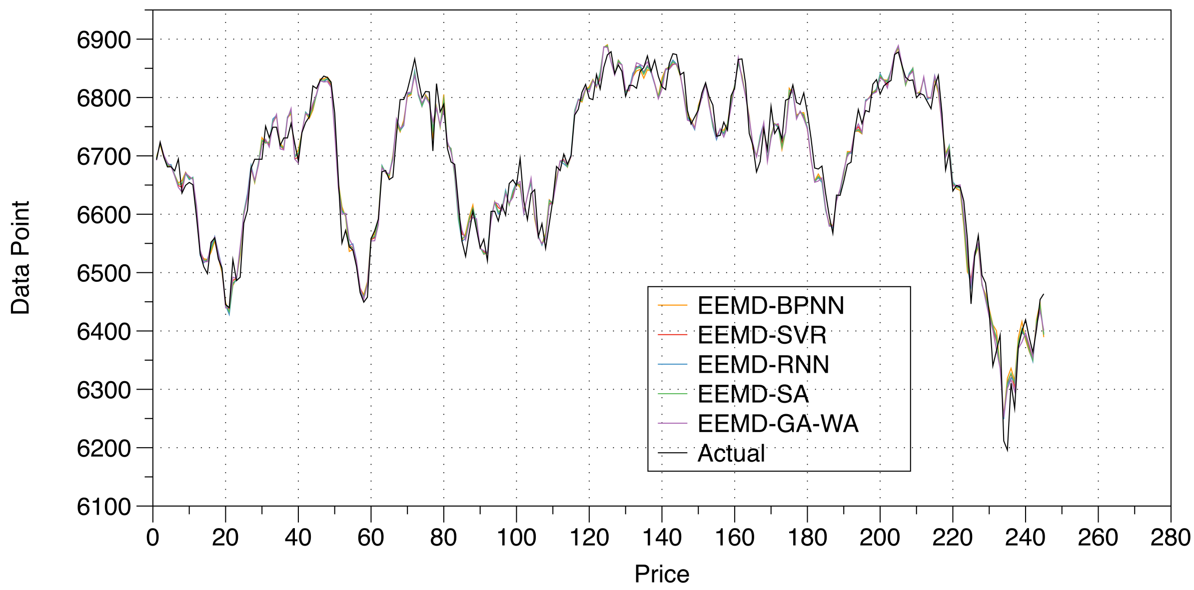

Figure 4.6: The actual FTSE100 closing price Index and its predicted values from EEMD-SVR, EEMD-RNN, EEMD-BPNN, EEMD-SA and EEMD-GA-WA models.

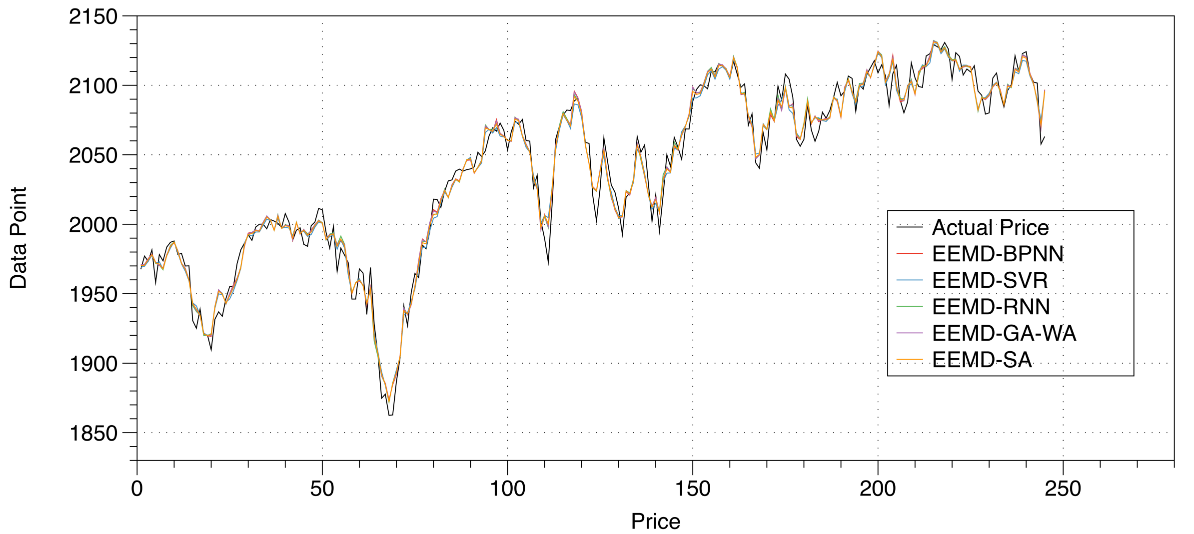

Figure 4.7: The actual S&P500 closing price Index and its predicted values from EEMD-SVR, EEMD-RNN, EEMD-BPNN, EEMD-SA and EEMD-GA-WA models.

It can be noticed from Figure 4.7 that the predicted values from the EEMD-SVR, EEMD-RNN, EEMD-BPNN, EEMD-SA and EEMD-GA-WA models are closer to the actual values of S&P 500 than those from the single approach models from Figure 3.26. Moreover, the predicted values from the EEMD-SVR, EEMD-RNN, EEMD-BPNN, EEMD-SA and EEMD-GA-WA models have a smaller deviation between their results and the actual values of the S&P 500 closing prices, as Figure 4.7 indicates. Figure 4.8 depicts the actual Nikkei 225 closing price values and predicted values from the EEMD-SVR, EEMD-RNN, EEMD-BPNN, EEMD-SA and EEMD-GA-WA models. From the figure it can be observed that the predicted values of all the models have smaller deviations with the actual values. Thus, in comparison with the single-approach- predicted values from Figure 3.24, the results in Figure 4.8 indicate that the EMD-SVR, EEMD-RNN, EEMD-BPNN, EEMD-SA and EEMD-GA-WA models predicted values are closer to the actual Nikkei 225 closing prices. Thus, it can be concluded that the predicted values of the proposed models in this chapter are closer to the actual values than those of the single approach models in Chapter 3.

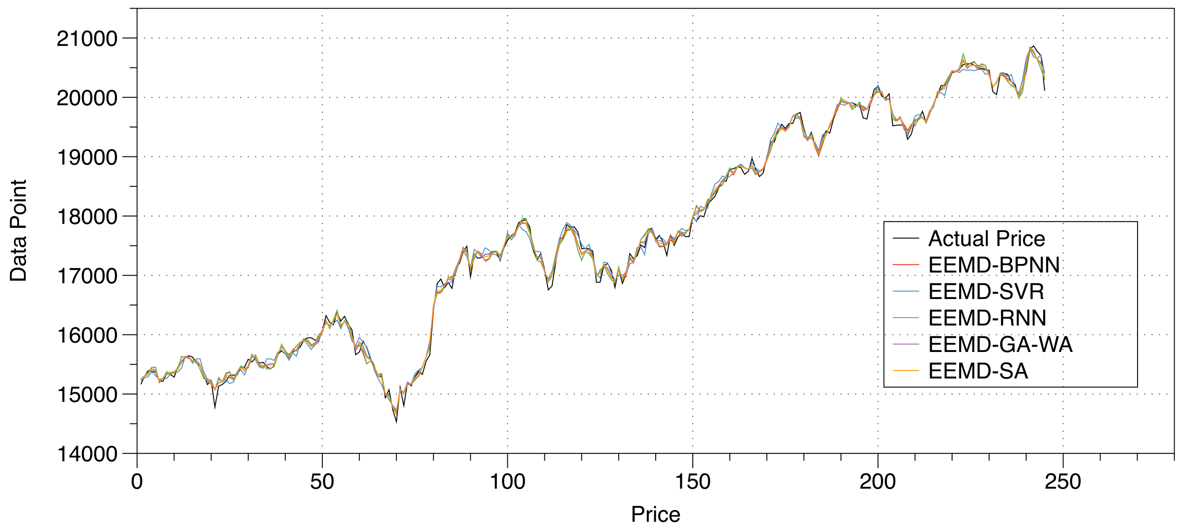

Figure 4.8: The actual Nikkei 225 closing price Index and its predicted values from EEMD-SVR, EEMD-RNN, EEMD-BPNN, EEMD-SA and EEMD-GA-WA models.

The FTSE 100 closing price index prediction results of the training data set using EMD-SVR, EEMD-RNN, EEMD-BPNN, EEMD-SA and EEMD-GA-WA models are computed and listed in Table 4.1. It can be found from Table 4.1 that the MSE, RMSE and MAE of the EEMD-GA-WA model for FTSE 100 predicted values are, respectively, 704.05, 26.53 and 20.15. Thus, it can be observed that these results are the smallest among the rest of the proposed models. This indicates that there is smaller deviation between the actual and predicted values utilising the suggested model EEMD-GA-WA. Moreover, compared to the predicted results obtained from the single approach in chapter 3 Table 3.13, the EEMD-GA-WA model has the lowest MSE, RMSE and MAE. The cross-correlation coefficient R results for all prediction models of FTSE 100 indicate that the prediction values and the actual values do not deviate too much. In addition, each method was run twenty times and the standard deviation was calculated. Thus, it can be observed that the results of SD for all models are relatively small, which implies that the models are not running randomly.

Table 4.1: The prediction result of training data sets for FTSE 100, S&P 500 and Nikkei 225 using EEMDSVR,EEMDRNN, EEMDBPNN, GA-WA and SA.

Index name	**Models**	**MSE**	**RMSE**	**MAE**	**R**	**SD**
FTSE100	EEMD-SA	714.86	26.73	20.45	0.99	——
	EEMD-SVR	749.70	27.38	20.71	0.99	1.61
	EEMD-RNN	723.54	26.89	20.63	0.99	1.86
	EEMD-BPNN	707.44	26.59	20.41	0.99	1.74
	EEMD-GA-WA	704.05	26.53	20.15	0.99	1.94
S&P500	EEMD-SA	39.11	6.25	3.88	0.99	——
	EEMD-SVR	40.27	6.34	2.58	0.99	1.071
	EEMD-RNN	39.73	6.30	2.61	0.99	1.067
	EEMD-BPNN	39.44	6.28	2.24	0.99	1.19
	EEMD-GA-WA	39.40	6.27	2.18	0.99	1.48
Nikkei225	EEMD-SA	9741.54	98.69	74.47	0.99	——
	EEMD-SVR	12751.14	112.92	85.38	0.99	1.68
	EEMD-RNN	11882.82	109.00	84.14	0.99	1.64
	EEMD-BPNN	12131.96	110.14	84.38	0.99	1.58
	EEMD-GA-WA	9716.02	98.56	74.02	0.99	1.79

The predicted values for the training data set of the S&P 500 closing price index are illustrated in Table 4.1. From Table 4.1, it can be found that the MSE, RMSE and MAE are respectively, 39.40, 6.27 and 2.12 of EEMD-GA-WA. Thus, it can be observed that these results have the smallest error between the all utilised models. Therefore, this indicates that there is less deviation between the predicted and actual values utilising the proposed model EEMD-GA-WA. In comparison with the obtained results from the single approach in chapter 3 Table 3.13, the EEMD-GA-WA results have the lowest error, which implies that the model proposed in this chapter outperformed the single approach.

Table 4.1 exhibits the cross-correlation coefficient R results for the S&P 500, and it can be observed that the prediction values and the actual values do not deviate too much. Furthermore, the SD results are relatively small, which implies that the models are not running randomly.

Table 4.1 presents the predicted values of the Nikkei 225 closing price for the training data set. As the Table shows, the EEMD-GA-WA model results have the lowest error among all the utilised models, which indicates that there is a smaller deviation between the actual and predicted values using EEMD-GA-WA. In addition, compared to the single approach results in chapter 3 Table 3.13, EEMD-GA-WA performed much better and achieved the lowest error. Moreover, the results of the cross-correlation coefficient R results for the Nikkei 225, indicate that the deviation between the actual and the predicted values is not that much. Thus, the SD results are relatively small, which implies that the models are not running randomly.

The testing data set for the FTSE 100, S&P500 and Nikkei 225 closing price indices prediction results utilising the proposed methods are computed and listed in Table 4.2. From Table 4.2, it can be found that the MSE, RMSE and MAE of the EEMD-GA-WA model outperformed the rest of the models and have the lowest error, which indicates that there is a smaller deviation between the actual and predicated values utilising EEMD-GA-WA. Moreover, compared to the results of the single approaches in chapter 3 Table 3.14, the EEMD-GA-WA, EEMD-SVR, EEMD-RNN, EEMD-BPNN and EEMD-SA models have the lowest MSE,RMSE and MAE.

Thus, it can be concluded that introducing EEMD to predict the FTSE 100, Nikkei 225 and S&P 500 has enhanced the obtained results and provided a better predict result than the single approach models in terms of prediction error and accuracy.

Table 4.2: The prediction result of testing data sets for FTSE 100, S&P 500 and Nikkei 225 using EEMDSVR,EEMDRNN, EEMDBPNN, EEMD-GA-WA and EEMD-SA.

Index name	Models	MSE	RMSE	MAE	R	SD
FTSE100	EEMD-SA	556.31	23.58	17.95	0.98	——
	EEMD-SVR	534.47	23.11	17.58	0.98	1.45
	EEMD-RNN	556.29	23.58	18.05	0.99	1.58
	EEMD-BPNN	598.46	24.46	18.60	0.98	1.67
	EEMD-GA-WA	518.90	22.77	17.53	0.99	1.06
S&P500	EEMD-SA	168.24	12.97	9.03	0.99	——
	EEMD-SVR	170.25	13.04	9.22	0.98	1.38
	EEMD-RNN	175.19	13.23	9.0	0.99	1.54
	EEMD-BPNN	177.63	13.32	9.01	0.99	1.34
	EEMD-GA-WA	118.46	10.88	7.92	0.99	1.86
Nikkei225	EEMD-SA	7895.01	88.85	69.88	0.99	——
	EEMD-SVR	9757.61	98.78	78.63	0.99	1.82
	EEMD-RNN	8632.35	92.91	72.82	0.99	1.53
	EEMD-BPNN	900.5	94	74.91	0.99	1.49
	EEMD-GA-WA	7254.00	85.95	68.47	0.99	1.92

4.4 Summary

This chapter proposed a new three-step prediction model by integrating EEMD and RNN, BPNN, SVR, GA-WA for financial time series. The proposed models (EEMD-RNN, EEMD-SVR, EEMD-PBNN, EEMD-SA and EEMD-GA-WA) first use EEMD to decompose the FTSE 100, Nikkei 225 and S&P 500 into different IMFs. Since the financial time series data is inherently noisy, EEMD is utilised to reconstruct the criterion of the original unsteady and non-linear data into certain components which have fixed frequency and periodicity. Furthermore, SVR, PBNN and RNN are applied to the predictions of each IMF. After each IMF is predicted in the built RNN, PBNN and SVR models, the final prediction results are obtained by adding the predicted results of each IMF for each model.

One of the contributions of this thesis is introducing EEMD as a preprocessor to add white noise and decompose the raw data into a finite set of IMFs, which have high correlations and a simpler frequency. Moreover, this chapter compared the obtained results of the proposed model with the single approach models in Chapter 3. Thus, the empirical results showed that using EEMD as a preprocessor advances the simplification of RNN, SVR and BPNN modelling and also obtains a much more precise model than the single approach models. Therefore, the new proposed model proved that it is better at predicting non-linear and strong noise data than any other single model employed in this thesis.

In addition, the weighted average combination method using GA to optimise the weight (the EEMD-GA-WA model) was introduced to combine the final prediction results of EEMD-SVR, EEMD-PBNN and EEMD-RNN in order to enhance the prediction accuracy and minimise the error. According to the experiments, the proposed hybrid model EEMD-GA-WA produced lower prediction errors than the EEMD-SVR, EEMD-PBNN and EEMD-RNN models.

Therefore, it can be concluded that the proposed hybrid model EEMD-GA-WA outperformed the EEMD-SVR, EEMD-PBNN and EEMD-RNN models.

Chapter 5

A New Hybrid Model for Stock Index Trend Price Prediction Based on Wavelet transforms-RNN-SVM and Naive Bayes

5.1 Introduction

Predicting stock price tendencies is considered to be one of the most challenging tasks in stock market mining. A number of factors influence stock market performance, such as political events, general economic conditions and trader expectations. Market traders rely heavily on different types of intelligent systems to help them to make decisions and to deal with stocks and futures [3]. However, up to date these available systems have limited ability and capability. However, to date these available systems have had limited ability and capability. Moreover, financial experts also find it difficult to make accurate predictions, due to the uncertainties involved and the characteristics of market trends.

Generally, among market experts there is no consensus on the effectiveness of predicting a financial time series.

In the literature, many prediction techniques have been developed and introduced to predict stock trends. Traditional regression models were one of the earliest models used to predict stock trends. However, these models do not perform satisfactorily, as stock data are categorised as non-stationary, non-linear time series data. Therefore, new models from different domains, which have the ability to understand and learn from non-linear data sets, have been introduced to predict stock direction movements. ANN, SVM, and Nave Bayes are three machine learning algorithms, and other techniques from the signal processing domain, such as Wavelet Transformation (WT), are most widely used for predicting stock and stock price movements. More information about SVM, ANN and Nave Bayes is available in Chapter 2 and 3. Wavelet analysis is considered to be a relatively new field in signal processing [63]. It can be defined as a mathematical function that decomposes data into different frequency components, where studying each component is considered in order to match it to its resolution scale, where the scale denotes a time horizon [216]. Generally, wavelet filtering is relatively close to the volatile and time-varying characteristics of real-world time series and therefore it is not limited by any assumptions such as being stationary [210]. The process of decomposing the data set into different scales by wavelet transformation makes the data after decomposition more useful for distinguishing seasonality, revealing structural breaks and volatility clusters, and identifying local and global dynamic properties of a process at specific time-scales [97]. Experimental results have shown that wavelets algorithms are useful, particularly for analysing, modelling and predicting the behaviour of any diverse data set such as financial instruments (stocks, exchange rates) [214] [201]. Therefore, this chapter adopts WT using the Haar wavelet to decompose the time series, in order to tackle any problems associated with data redundancy and being non-stationery. Thus, numerous practical problems are related to the quite large number of input variables, which are quite large. In addition, these input variables are characteristically redundant. As a results of such characteristic, eliminating the redundant variables could lead to enhancing the performance of prediction models. Moreover, by reducing the data dimensionality and eliminating the redundant variables, the interpret-ability of predictive model can be enhanced [146].

In Chapter 2 a single approach model is presented to predict the direction of stock prices. However, the results were not satisfactory. Therefore, this chapter proposed a new stock movement prediction approach which integrates Wavelet Transform, SVM, RNN and Naive Bayes to enhance the prediction capability. Moreover, a hybrid model is proposed to combine the output of the WT-SVM, WT-RNN and WT-Naive Bayes as the weights of these proposed methods are determined by the GA. The chapter is organised as follows: Section 5.2 describes the prediction procedure of the hybrid model. Subsection 5.2.1 introduces the WT method. The hybrid combination steps are illustrated in Subsection 5.2.2. Section 5.3 presents the results based on real data sets. Finally Section 5.4 concludes this chapter.

5.2 Hybrid Modelling and Prediction Procedures

The proposed approach in this chapter consists of three steps. The first step utilises wavelet transformation in order to decompose the predictor variable under different basic functions and decomposition stages to generate the sub-series. In the second step, the obtained sub-series from the first step is applied in SVM, RNN and Naive Bayes as a new input variable to build a prediction model. Finally, in the third step, the SVM, RNN and Naive Bayes are combined, and the weights of these three models are determined by the GA. Moreover, the purpose of the proposed method in this chapter is to predict stock index price movement. Thus, this chapter carries out the stage two of phase two as explained earlier in Chapter 3 Section 3.4 in Figure 3.10.

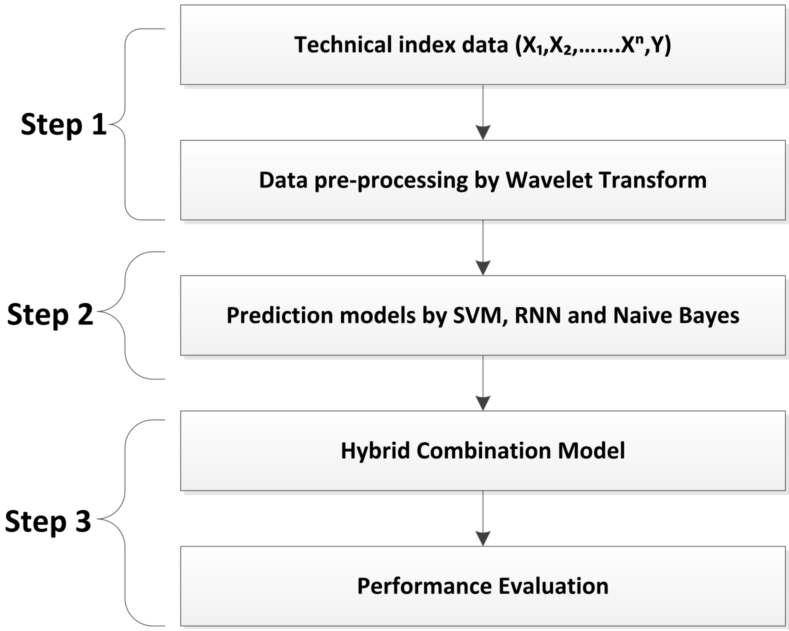

Figure 5.1: The overall process of WT based SVM, Naive Bayes and RNN hybrid methodology.

To evaluate the proposed method, this chapter uses the eight technical indicators of the three stock market indices, FTSE 100, Nikkei 225 and S&P500; more details about these indicators are available in Chapter 3 Section 3.2.2. Figure 5.1 shows the main procedures of the proposed model approach. The following points illustrate the implementation steps of the new hybrid model:

- Step one: selecting technical indicators as feature subsets. Subsection 3.2.2 in Chapter 3 explained and summarised the selected technical indicators and how they were selected. The proposed WT is applied for preprocessing the input data. As mentioned earlier, three levels of wavelet preprocessing are implemented and RMSE is used to measure the performance of WT. Subsection 5.2.1 illustrates the WT process. The purpose of this step is to remove the noise in the original data.

- In the second step, after the data has been decomposed by WT, the Naive Bayes, RNN and SVM models are used to predict the new input variables. In addition, the Nave Bayes, RNN and SVM methods and model topologies are exhibited in Chapter 3.

- Step three; the weighted average combination function was utilised to combine the different WT-RNN, WT-BPNN and WT-SVR methods. An optimiser genetic algorithm is used to determine the weight of the combiner; more details on this step are available in Subsection 5.2.2.

5.2.1 Wavelet Transform (WT)

The capability of revealing certain aspects of data is one of the main advantages of wavelet transformation which is missing in other signal analysis techniques. Such aspects can be trend, breakdown points, discontinuities in higher derivatives and self-similarity. Therefore, such an advantage makes wavelet transformation suitable for the analysis of non-linear and non-stationary financial time series. In addition, the ability to decompose a time series into multiple resolution constituent time series by wavelet transformation is also a main reason for using it in time series data analysis. According to Gencay et al.[96], insights into the dynamics of financial time series are provided by wavelet methods, in a way that is beyond what other standard time series methodologies could provide. Moreover, Jin et al.[164] argue that, in practice, local characteristics of a non-stationary time series can be indicated by the obtained wavelet coefficients at a time scale space in order to identify the state of the system, which often extracts features.

According to Ramsey [214], analysing economic and financial data using the wavelet approach has many benefits such as flexibility in handling very irregular data series, time scale decomposition of data and non-parametric representation of each individual time series, as well as determining whether a time series can be predicted at the corresponding prediction horizon. Therefore, as a result of such benefits, researchers have introduced this prediction approach using wavelet transformation. Early studies in the use of wavelet transformation are discussed in [10], [16], [99], [77], [251] and [135]. Thus, decomposing the signal into time scale components by wavelet transformation is the main reason for utilising such an approach, and then treating each approximation at each time scale as a separate series. Although the wavelet transformation approach has shown significant results in predicting financial time series, decomposition with wavelet transformation has a major disadvantage, which is that the cumulative error from predicting each resolution level could be greater than those from predicting the original time series directly. Therefore, in order to tackle such drawbacks, an investigation is necessary to find out under which situations a decomposition method with wavelet transformation outperforms single resolution approaches.

In this chapter, the Haar wavelet is applied as a main wavelet transformation tool. Thus, the wavelets role is not just decomposing the data into terms of times and frequency, but it also has another important role, which is reducing the processing time significantly. In this chapter n denotes the size of time series, and then the used wavelet decomposition can be determined in $O(n)$ time [2]. The wavelet theory is based on Fourier analysis, where any function is represented as the sum of the sine and cosine functions. Equation 5.1 illustrates the wavelet admissibility condition, where the wavelet $\psi(t)$ is simply a function of time t that obeys a basic rule [98].

$$C_\psi = \int_0^\infty \frac{\psi(f)}{f} df < \infty \tag{5.1}$$

In Equation 5.1 the Fourier transform and frequency f function of $\psi(t)$ is C_ψ. Wavelet transform (WT) can be defined as that mathematical tool which can be implemented to solve numerous applications problems such as signal processing and image analysis. Thus, the first implementation of WT was to solve the problem associated with the Fourier transformations as they occur. As a result of the data characteristics, such as being non-stationary or the fact that the data are localised in time, space, or frequency, the occurrence could take place. Within the given function /family there are two types of wavelets, depending on the rule of normalisation. The first type is the father wavelet where the smooth and low-frequencies are described. The second type of wavelet is the mother wavelet where the high-frequency components details are described. The father wavelet is represented in Equation 5.2 and the mother wavelet also is represented in Equation 5.3, in both equations $j = 1, ..J$ where J is the wavelet decomposition level [215].

$$\phi_{j,k} = 2^{-j/2}\phi\left(t - 2^{j}k/2^{j}\right) \tag{5.2}$$

$$\psi_{j,k} = 2^{-j/2}\psi\left(t - 2^{j}k/2^{j}\right) \tag{5.3}$$

In Equations 5.2 and 5.3 which represent the two wavelet types (father and mother), the maximum scale sustainable by the number of data is denoted by J and satisfies:

$$\int \phi\left(t\right) dt = 1 and \int \psi\left(t\right) dt = 0 \tag{5.4}$$

For example a time series data function $f(t)$ considered as an input by wavelet analysis, where a sequence of projections can be built up onto father and mother wavelets which is indexed by both $\{k\}, k = \{0, 1, 2, ...\}$ and by $\{s\} = 2^{j}\{j = 1, 2.J\}$. Thus in order to analysis real discrete sample data, a lattice for making calculations is required. Therefore, from a mathematical perspective it is convenient to use a dyadic expansion as Equation 5.4 illustrates. The projections which can give expansion coefficients are illustrated in Equation 5.5:

$$\begin{gathered} s_{J,k} = \int \phi_{J,k} f\left(t\right) dt \\ d_{J,k} = \int \psi_{J,k}\left(t\right) dt \left(j = 1, 2, ..J\right) \end{gathered} \tag{5.5}$$

Equation 5.6 defines the orthogonal wavelet series approximation to $f(t)$:

$$\begin{gathered} f\left(t\right) = \sum_k s_{J,k}\left(t\right) + \sum_k d_{J,k}\psi_{J,k}\left(t\right) + \sum_k d_{J-1,k}\psi_{J-1,k}\left(t\right) \\ +...... + \sum_k d_{1,k}\psi_{1,k}\left(t\right) \end{gathered} \tag{5.6}$$

Briefly it can be also represented in another form as Equation shows:

$$\begin{aligned} f\left(t\right) = s_J\left(t\right) + D_J\left(t\right) + D_{J-1}\left(t\right) + ... + D_1\left(t\right) \\ S_J\left(t\right) = \textstyle\sum_k s_{J,k}\phi_{J,k}\left(t\right) \\ D_J\left(t\right) = \textstyle\sum_k s_{J,k}\psi_{J,k}\left(t\right) \end{aligned} \tag{5.7}$$

Equation 5.6 illustrates how WT is used to calculate the coefficient of the wavelet series with finite extent for a discrete signal $f_1, f_2, ...f_n$. Mapping the vectors $f = f_1, f_2, ...f_n$ to a vector of n wavelet coefficient $w = w_1, w_2, ...w_n$ by WT, where both the smoothing coefficient $S_{J,k}$ and detail coefficient $d_{J,k}, j = 1, 2, ..J$ are contained. The $S_{J,k}$ denotes the underlying smooth behaviour of signal at scale coarse 2^J, whereas, $d_{j,k}$ can be described as a coarse scale deviation from the smooth behaviour, and thus $d_{j-1,k}, ..d_{1,k}$ provides the progressively finer scale deviations from the smooth behaviour [224].

In the case of n is divisible by 2^J, $d_{1,k}$, which contains $n/2$ observations at the finest scale $2^1 = 2$, and $n/4$ observations in $d_{2,k}$ at the second finest scale $2^1 = 2$. Therefore, $d_{j,k}$ and $s_{j,k}$ each contain $n/2^J$ observations. Equation 5.8 illustrates this case where:

$$n = n/2 + n/4 + ... + n/2^{J-1} + n/2^J \tag{5.8}$$

Thus $f(t)$ represents the original data, S_1 denotes the approximation signal, where D_1 is a detailed signal. In this Chapter the multi-resolution decomposition is defined for the signal by specifying that the coarsest scale is S_J and $S_{J-1} = S_J + D_J$.In addition, $S_{j-1} = S_j + D_j$ which$\{S_J, S_{J-1}, ..S_1\}$ and thus with ever increasing refinement, the sequence of multi-resolution approximations of the function $f(t)$. $\{S_J, D_J, D_{J-1}, ...D_j, ..., D_1\}$ is given the corresponding multi-resolution decomposition of $f(t)$. A set of orthogonal signal components are represented in the sequence of terms $S_J, D_J, D_{J-1}, ...D_j, ..., D_1$ which represents the signal at resolutions 1 to J. Thus each D_{J-k} is providing the orthogonal increment to the representation of function $f(t)$ at the scale 2^{J-k}. Therefore, in the case of the data pattern being very rough, the wavelet process is repeatedly applied. Generally the main aim of preprocessing the data is to minimise the error between the signal before and after the transformation. Thus, in this process the noise in the data can be removed. One of the main reasons for preprocessing the data is to reduce the adaptive noise in the training pattern which will tackle any risk problem related to over-fitting in the training phase [207]. Thus, WT is adopted for preprocessing the data in this chapter three times.

5.2.2 Hybrid Combination Model WT-GA-WA

In the literature, different combination methods of prediction techniques have been widely investigated. In short-range predictions, combining various techniques is more useful according to [293], [12]. According to the Timmermann study, using the simple average may work as well as more sophisticated approaches. However, using one model can produce more accurate predictions than any other methods. Therefore, simple averages would not be sufficient in such cases [242]. Compared with different prediction models, the hybrid prediction method is based on a certain linear combination. The assumption for the actual value in period t by model i is $f_{it}\,(i = 1, 2, ...m)$, the corresponding prediction error will be $e_{it} = y_t - f_{it}$. And also the weight vector will be $W = [w_1, w_2, ...w_m]^T$ .. Then in the hybrid model the predicted value is computed as follow [104] [61]:

$$\hat{y}_t = \sum_{i=1}^{m} w_i f_{it}\,(t = 1, 2, ...n) \tag{5.9}$$

$$\sum_{i=1}^{m} w_i = 1 \tag{5.10}$$

Equation 5.9 can be expressed in another from, such as:

$$\begin{aligned} &\hat{y} = FW \\ &\quad where \quad \hat{y} = [\hat{y_1}, \hat{y_2}, ...\hat{y_n}]^T, F = [f_{it}]_{n\times m} \end{aligned} \tag{5.11}$$

The error for the prediction model can be formed as Equation 5.12 illustrated.

$$e_t = y_t - \hat{y}_t = \sum_{i=1}^{m} w_i y_t - \sum_{i=1}^{m} w_i f_{it} = \\ \sum_{i=1}^{m} w_1 w_i \left(y_t - f_{it}\right) = \sum_{i=1}^{m} w_i e_{it} \tag{5.12}$$

This research proposes a hybrid model combining WT-SVM, WT-Naive Bayes and WT-RNN.

$$\hat{Y}_{combined_t} = \frac{w_1 \hat{Y}_{WT-SVM} + w_2 \hat{Y}_{WT-NaiveBayes} + w_3 \hat{Y}_{WT-RNN}}{(w_1 + w_2 + w_3)} \tag{5.13}$$

The prediction values in period t are $\hat{Y}_{combined_t}, \hat{Y}_{WT-SVM}, \hat{Y}_{WT-NaiveBayes}$ and $\hat{Y}_{WT-RNN}$ for the hybrid, EWT-SVM, WT-Naive Bayes and WT-RNN models, where the assigned weights are w_1, w_2, w_3 respectively, with $\sum_{i=1}^{3} w_i = 1.0 \leq w_i \leq 1$.

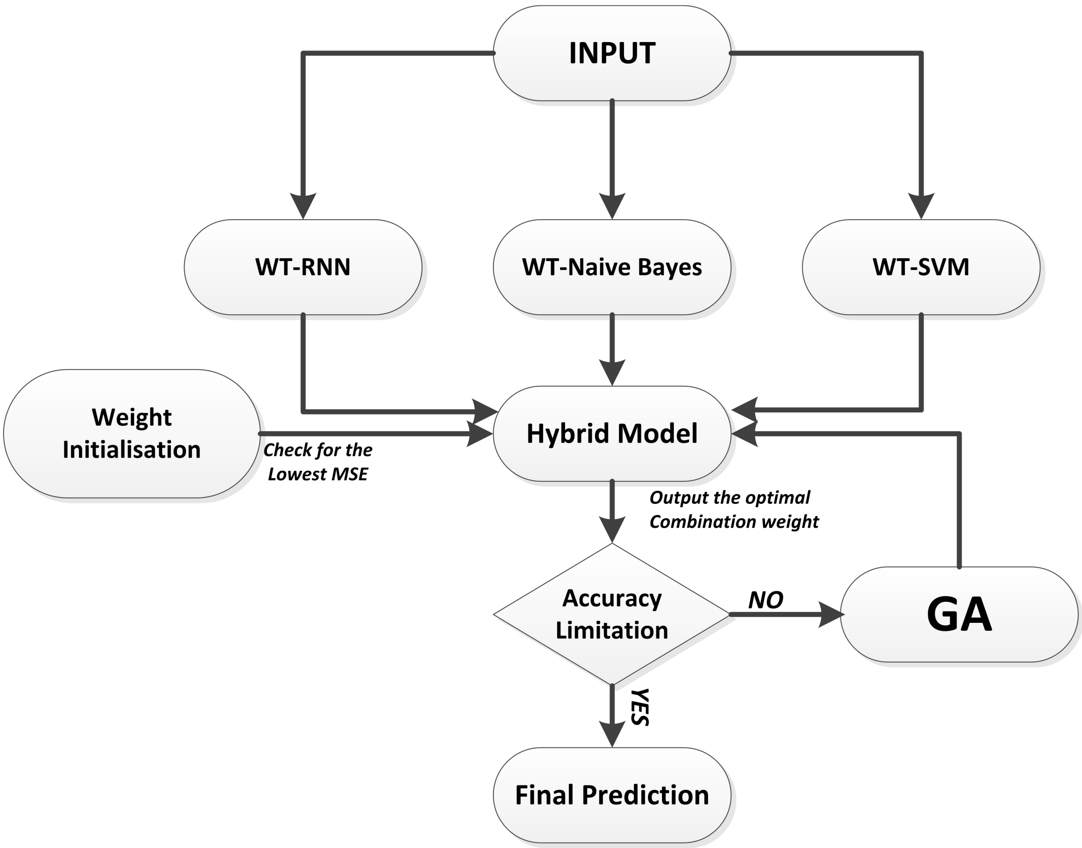

Figure 5.2: The flow chart of the WT-GA-WA hybrid model.

The most important step in developing a hybrid prediction model is to determine the perfect weight for each individual model. Setting $w_1 = w_2 = w_3 = 1/3$ in Equation 5.13 is the simplest combination method for the three prediction models. Nevertheless, in many cases equal weights cannot achieve the best prediction result. Therefore, this chapter adopts a hybrid approach utilising GA as an optimiser to determine the optimal weight for each prediction model. Figure 5.2 illustrates the architecture of the WT-GA-WA hybrid model. Moreover, the GA algorithm is explained in Chapter 4 Subsection 4.2.2.1.

5.3 Results and Discussions

The new hybrid model in this chapter is based on the WT, SVM, RNN, Nave Bayes and WT-GA-WA hybrid models. These are constructed to predict the direction of the closing prices of the FTSE 100, S&P 500 and Nikkei 225 indices. To the best of the researchers knowledge this chapter adopts a new prediction method to predict the exact changes in the closing price of stock indices. To demonstrate the validity of the new adopted methods, this section compares the results of the single approach from Chapter 3 with the results of the new hybrid model. As explained earlier in Section 5.2 the proposed approach consists of three steps. Thus, the selection criteria of the input variable is explained in Section 3.2.2 Chapter 3. The first step is to decompose the input variable by WT. The second step is to predict those decomposed variables by SVM, RNN and Naive Bayes. The third step is to combine the predicted results by the proposed hybrid model as explained in Section 5.2.2.

Table 5.1 shows the predicted results of the direction movement of the closing price for the FTSE 100 training data set using the WT-SVM, WT-RNN, WT-Naive Bayes, WT-SA and WT-GA-WA models. It can be found from Table 5.1 that the MSE, RMSE and MAE of the WT-GA-WA model for the FTSE 100 predicted values are, respectively, 23.82, 4.88 and 3.37. Thus, it can be observed that these results are the smallest among the rest of the proposed models. Therefore, it indicates that there is smaller deviation between the actual and predicted values utilising the suggested model WT-GA-WA. Moreover, compared to the obtained predicted results form single approach in chapter 3 Table 3.15, the WT-GA-WA model has the lowest MSE, RMSE and MAE. The cross correlation coefficient R results for all prediction models of FTSE 100 indicates that the prediction values and the actual values are not deviating too much. In addition, each methods were run twenty times and the standard deviation was calculated. Thus, it can be observed that the results of SD for all models are relatively small, which implies that the models are not running randomly.

Table 5.1: The Trend prediction result of training data sets for FTSE 100, S&P 500 and Nikkei 225 using WT-SVM, WT-RNN, WT-SA, WT-Naive Bayes and WT-GA-WA hybrid model.

Index name	**Models**	**MSE**	**RMSE**	**MAE**	**R**	**SD**
	WT-SVM	86.42	9.29	6.09	0.80	1.36
	WT-RNN	44.11	6.64	4.59	0.90	1.25
FTSE100	WT-Naive Bayes	40.28	6.34	4.02	0.99	1.07
	WT-SA	39.56	6.28	3.94	0.95	——
	WT-GA-WA	23.82	4.88	3.37	0.99	1.49
	WT-SVM	8.16	1.51	2.85	0.78	1.94
	WT-RNN	5.42	2.32	1.22	0.85	1.32
S&P500	WT-Naive Bayes	5.5	2.34	1.11	0.96	0.87
	WT-SA	6.36	2.52	1.35	0.93	——
	WT-GA-WA	0.005	0.07	0.009	0.99	1.71
	WT-SVM	1574.11	39.67	28.40	0.87	1.74
	WT-RNN	725.20	26.92	20.55	0.94	1.18
Nikkei225	WT-Naive Bayes	0.41	0.64	0.02	0.99	2.40
	WT-SA	409.93	20.24	15.18	0.91	——
	WT-GA-WA	0.40	0.63	0.02	0.96	1.95

The S&P 500 predicted values of direction movement of closing price for the training data set are illustrated in Table 5.1. From the Table 5.1, it can be found that the MSE, RMSE and MAE are respectively, 0.005, 0.07 and 0.009 of WT-GA-WA. Thus, it can be observed that these results are the smallest error between the all utilised models. This indicates that there is smaller deviation between the predicted and actual values utilising the proposed model WT-GA-WA. In comparison with the obtained results from the single approach in chapter 3 Table 3.15, the WT-GA-WA results have the lowest error, which implies that the proposed model in this chapter outperformed the single approach. Table 5.1 exhibits the cross-correlation coefficient R results for the S&P 500. It can be observed that the prediction values and the actual values do not deviate too much. However, SD results are relatively small, which implies that the models are not running randomly.

The computed predicted results for the Nikkei 225 direction movement of the closing price for the training data sets are presented in 5.1. The table shows that the WT-GA-WA model results have the lowest error among all the utilised models, which indicates that there is a smaller deviation between the actual and the predicted values using WT-GA-WA. In addition, compared to the single approach results in chapter 3 Table 3.15, WT-GA-WA performed much better and achieved the lowest error. Moreover, the results of the cross-correlation coefficient R for the Nikkei 225 indicate that the deviation between the actual and the predicted values are not that much. Thus, SD results are relatively small, which implies that the models are not running randomly.

Table 5.2: The Trend prediction result of testing data sets for FTSE 100, S&P 500 and Nikkei 225 using WT-SVM, WT-RNN, WT-SA, WT-Naive Bayes and WT-GA-WA hybrid model.

Index name	**Models**	**MSE**	**RMSE**	**MAE**	**R**	**SD**
FTSE100	WT-SVM	47.50	6.89	5.63	0.93	1.52
	WT-RNN	106.75	10.33	8.00	0.83	1.42
	WT-Naive Bayes	199.97	14.14	11.63	0.48	1.29
	WT-SA	74.53	8.63	6.55	0.84	——
	WT-GA-WA	35.95	5.95	4.34	0.94	1.17
S&P500	WT-SVM	46.20	6.79	2.633	0.88	1.03
	WT-RNN	55.57	7.45	2.70	0.55	0.92
	WT-Naive Bayes	54.58	7.38	3.05	0.89	0.94
	WT-SA	51.98	7.20	5.14	0.83	——
	WT-GA-WA	42.43	6.51	2.56	0.89	1.50
Nikkei225	WT-SVM	1344.27	36.66	26.16	0.81	2.34
	WT-RNN	2544.99	50.44	33.56	0.76	1.91
	WT-Naive Bayes	6010.71	77.52	58.31	0.54	2.34
	WT-SA	2261.00	47.55	34.90	0.78	——
	WT-GA-WA	1318.03	36.30	25.18	0.88	1.04

Table 5.2 presents the predicted results of the direction movement of the FTSE 100,

S&P500 and Nikkei 225 closing prices for the testing data set using the proposed models. It can be found from Table 5.2 that the MSE, RMSE and MAE values of the WT-GA-WA model are significantly smaller than the rest of the models, which indicates that there is a smaller deviation between the actual and predicated values utilising WT-GA-WA. In addition, comparing the obtained results of the proposed models in this chapter with the obtained results of single approaches in chapter 3 Table 3.16, the WT-GA-WA, WT-SVM, WT-RNN, WT-Nave Bayes and WT-SA models have the lowest MSE, RMSE and MAE. Thus, it can be concluded that integrating WT into the utilised AI techniques to predict the trends of the FTSE 100, S&P 500 and Nikkei 225 indices has enhanced the prediction results and minimised the prediction errors.

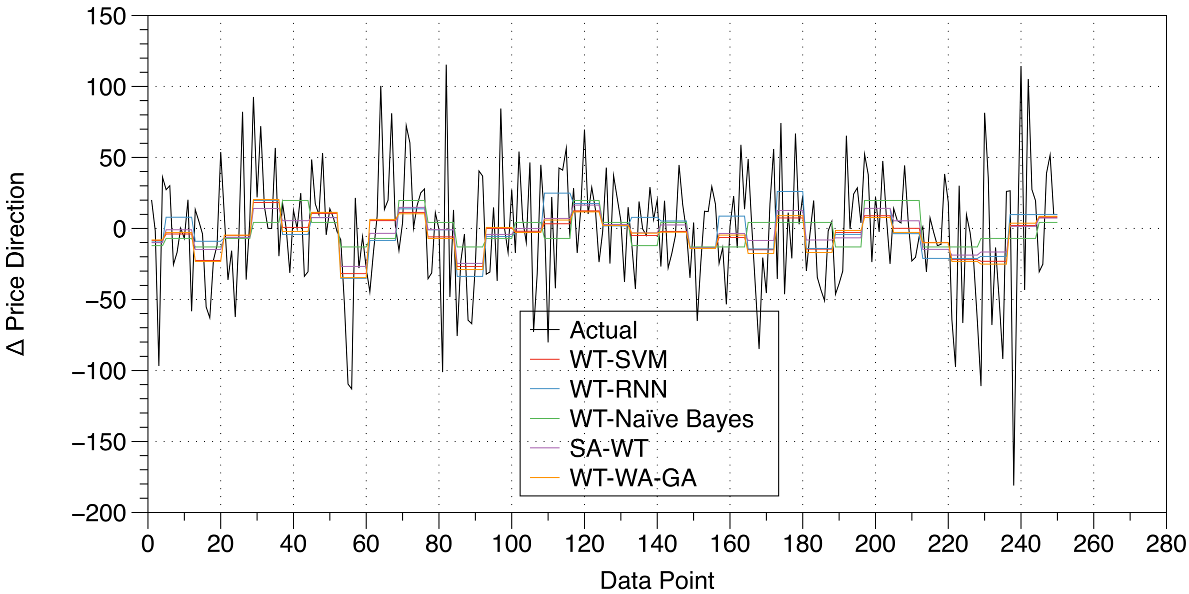

Figure 5.3: The actual FTSE 100 closing direction price Index and its predicted values from WT-SVM, WT-RNN, WT-Naive Bayes, WT-SA and WT-GA-WA hybrid models.

The prediction values of the FTSE 100 closing price direction movement versus the actual closing price direction movement of the proposed models are illustrated in Figure 5.3. It can be observed from Figure 5.3 that the predicted values obtained from WT-GA-WA, WT-SVM, WT-RNN, WT-Naive Bayes and WT-SA models are closer to the actual price than the predicted values of the single approach models in Figure 3.29. In addition, the WT-GA-WA predicted values are closer to the actual values than the other utilised models.

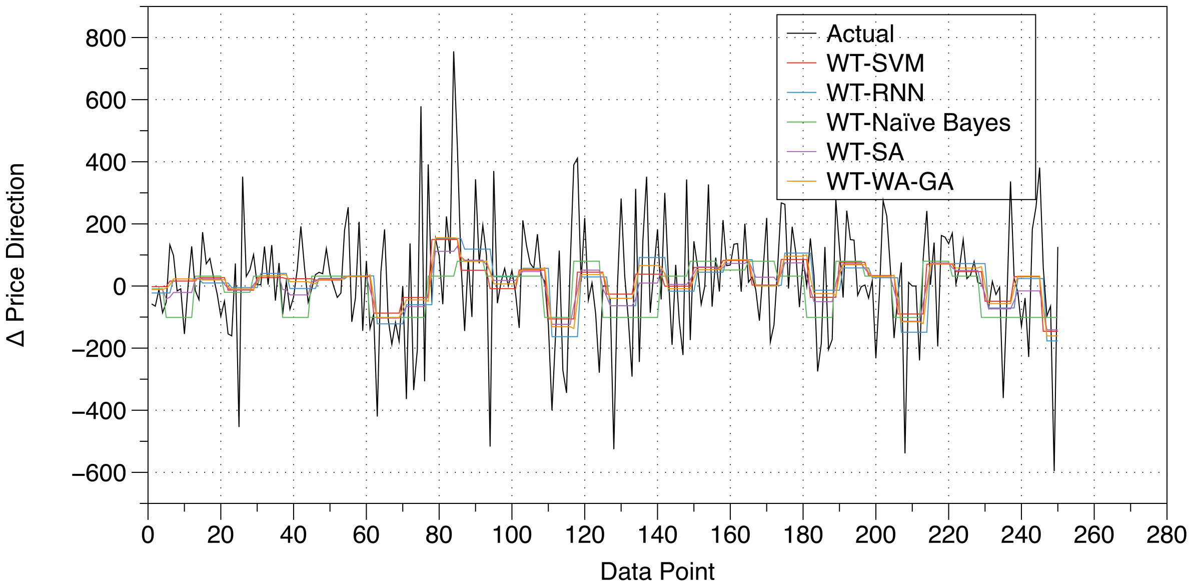

Figure 5.4: The actual Nikkei 225 closing direction price Index and its predicted values from WT-SVM, WT-RNN, WT-Naive Bayes, WT-SA and WT-GA-WA hybrid models.

The predicted direction movement of the Nikkei 225 closing price values by the WT-GA-WA, WT-SVM, WT-RNN, WT-Naive Bayes and WT-SA models and actual values are illustrated in Figure 5.4. Thus, it can be noticed from Figure 5.4 that the predicted values from the WT-GA-WA, WT-SVM, WT-RNN, WT-Naive Bayes and WT-SA models are closer to the actual values of the Nikkei 225 than those from the single approach models in Figure 3.28. In addition, the predicted values from the WT-GA-WA, WT-SVM, WT-RNN, WT-Naive Bayes and WT-SA models have a smaller deviation between their results and the actual values of the Nikkei 225 closing prices direction movement, as Figure 5.4 indicates.

Figure 5.5 depicts the actual S&P 500 closing price direction movement values and predicted values from the WT-GA-WA, WT-SVM, WT-RNN, WT-Naive Bayes and WT-SA models. From the figure it can be observed that the predicted values of all the models have a smaller deviation with the actual values. Thus, in comparison with the single-

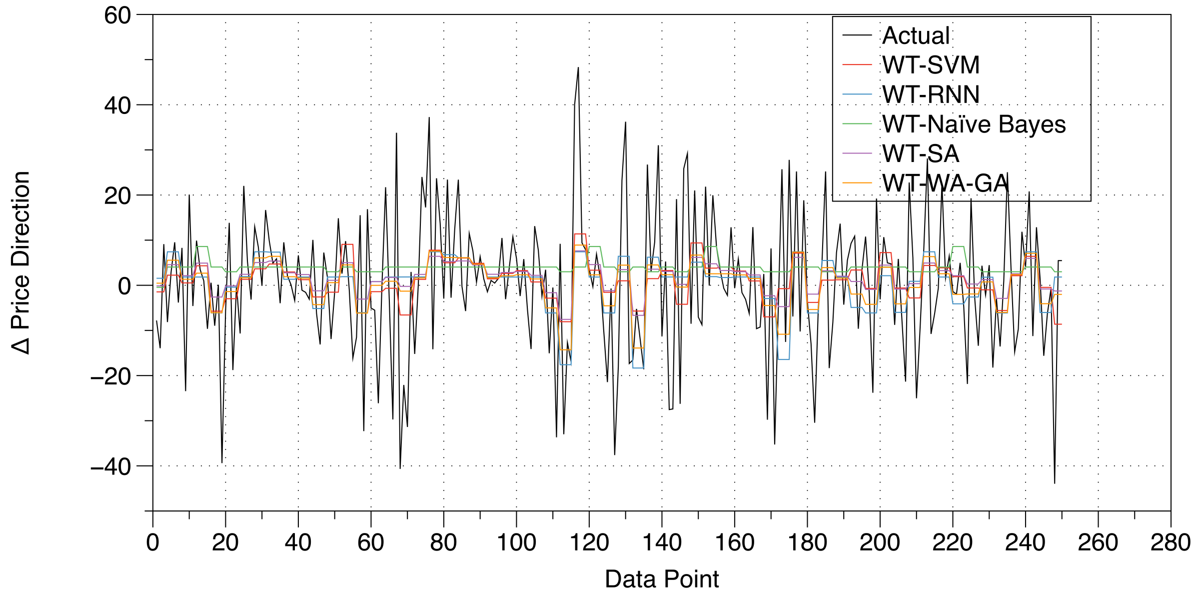

Figure 5.5: The actual S&P 500 closing direction price Index and its predicted values from WT-SVM, WT-RNN, WT-Naive Bayes, WT-SA and WT-GA-WA hybrid models.

approach predicted values from figure 3.27, the results in figure 5.5 indicate that the WT-GA-WA, WT-SVM, WT-RNN, WT-Naive Bayes and WT-SA predicted values are closer to the actual S&P 500 direction movement of the closing prices. Thus, it can be concluded that the predicted values of the proposed models in this chapter are closer to the actual values than those of the single approach models in chapter 3.

5.4 Summary

This chapter presented a new hybrid prediction model to predict the exact change in the closing prices of the FTSE 100, Nikkei 225 and S&P500 indices. To the best of the researchers knowledge the proposed model is introduced for the first time in this chapter.As explained earlier, this chapter exhibited stage two of phase two of the proposed prediction model for this study. Thus, the aim of the proposed method is to build an integrated hybrid system, WT-GA-WA, to predict the stock closing price direction movement. The hybrid system involves three steps: 1) preprocessing the data using

Wavelet Transform (WT) this step implements WT to decompose the data in order to eliminate the noise; 2) application of SVM, RNN and Naive Bayes is used to predict the decomposed data; 3) the use of the WT-GA-WA hybrid model to combine the predicted data.

The proposed approach is compared with the single approach models and the benchmark model. Thus, the results from implementing the new models greatly outperform the other single approach models. Moreover, in order to create more comparative benchmark problems and also to provide sufficient evidence that the proposed models are robust, other data sets were utilised .

The results show that the proposed models are applicable. Therefore, it can be concluded that the proposed model WT-GA-WA greatly outperforms the WT-SVM, WT-RNN, WT-Naive Bayes and WT-SA models.

Chapter 6

A New Hybrid Heuristic Rules Based System for Stock Price Prediction

6.1 Introduction

Stock price prediction is an important topic in the financial domain, and it has received in recent years considerable attention from market participants and researchers. However, this topic is described by many researchers as one of the most challenging tasks due to its high volatility, complexity, dynamics and turbulence. A number of methodologies have been used in an attempt to predict the stock price. These various methodologies can be broadly categorised into fundamental analysis, technical analysis, traditional time series techniques and computational intelligence techniques, which are described in detail in previous chapters. Moreover, and after much investigation into the increasing movements of market prices, the researcher concluded that the stock prices do not follow the random walk assumption. Thus, they are behaving in a highly non-linear, dynamic manner. In addition, the random walk theory of stock price future movements may merely be a veil of randomness that shrouds a noisy non-linear process [247]. As a result, and in order to remove this veil and to deliver reliable prediction results, the application of AI was introduced. However, it has been commonly reported that there is no single computational method that solves all the problems. Therefore, utilising and introducing a hybrid model, or combining different models, has become a common practice in order to improve prediction performance.

The early attempts in this field have expanded dramatically since the early research of [219] [19]. A comprehensive review and annotated bibliography are provided by Clemen [62] in this area. Moreover, Abraham et al [1] were one of the earliest scholars to have introduced hybridised soft computing techniques to build automated stock market prediction and trend analysis. Generally, the main purpose of combining different prediction models is to use each unique feature of the combined models in order to capture the variety of different patterns in the data. Theoretical and empirical findings have indicated that combining different models can be an effective and efficient way to improve prediction performance [175] [192] [200] [267]. In the literature there are many different utilised combination methods that have been introduced by researchers in order to improve the prediction performance. However, recent studies have shown that the direction of research on building stock prediction approaches can be roughly grouped into two types: time series prediction, and trend prediction. To the best of the researchers knowledge, this chapter introduces for the first time a new hybrid heuristic-rules-based combination system using a static (trend prediction) model and a dynamic (time series prediction) model in order to improve financial time series prediction performance. Moreover, this chapter represents phase three as explained earlier in Figure 3.10.

The remainder of this chapter is organised as follows: Section 6.2 introduces the proposed heuristic combination rules. Section 6.3 presents the results and discussions. Finally the summary of this Chapter is drawn in Section 6.4.

6.2 The Hybrid Heuristic Rules Based System Procedures

Hybridisation and combination approaches have been a common practise in modelling and predicting stock prices after the literature verified that the single approach did not solve all the problems. Thus, it was reported that in the prediction process, the hybrid approach system achieved a better performance level in comparison with traditional approach systems [26] [36] [143] [162] [189]. Thus, from the above discussion and that in chapter 2, it can be concluded that each algorithm can tackle this problem. However, it is also noted that each algorithm has its own limitations. Therefore, and in order to improve the prediction accuracy, this chapter proposes a new hybrid combination model. In accordance with Wolperts theorem [269], Hybrid Intelligent Systems (HISs) are described as a free set of computational intelligence techniques combined together in order to solve a given problem, which covers all the computational phases from data preprocessing up to the final output. Figure 6.1 illustrates the procedures of the new hybrid prediction method.

It includes three phases, as shown in Figure 6.1. Phase 1 is a hybrid dynamic approach to predict the next day closing price of the FTSE 100 Nikkei 225 and S&P 500 indices, using the historical closing price as inputs. This approach is explained in detail in Chapter 4. Phase 2 is a hybrid static approach to predict the exact direction movement of the FTSE 100 Nikkei 225 and S&P 500 next day closing prices by using the technical indicators as inputs. The prediction procedures of this approach are explained in Chapter 5. Phase 3 exhibits the procedures of the heuristic combination rules, using dynamic and static hybrid approaches. As the first two phases are explained in the previous chapters, this section will focus on explaining Phase 3.

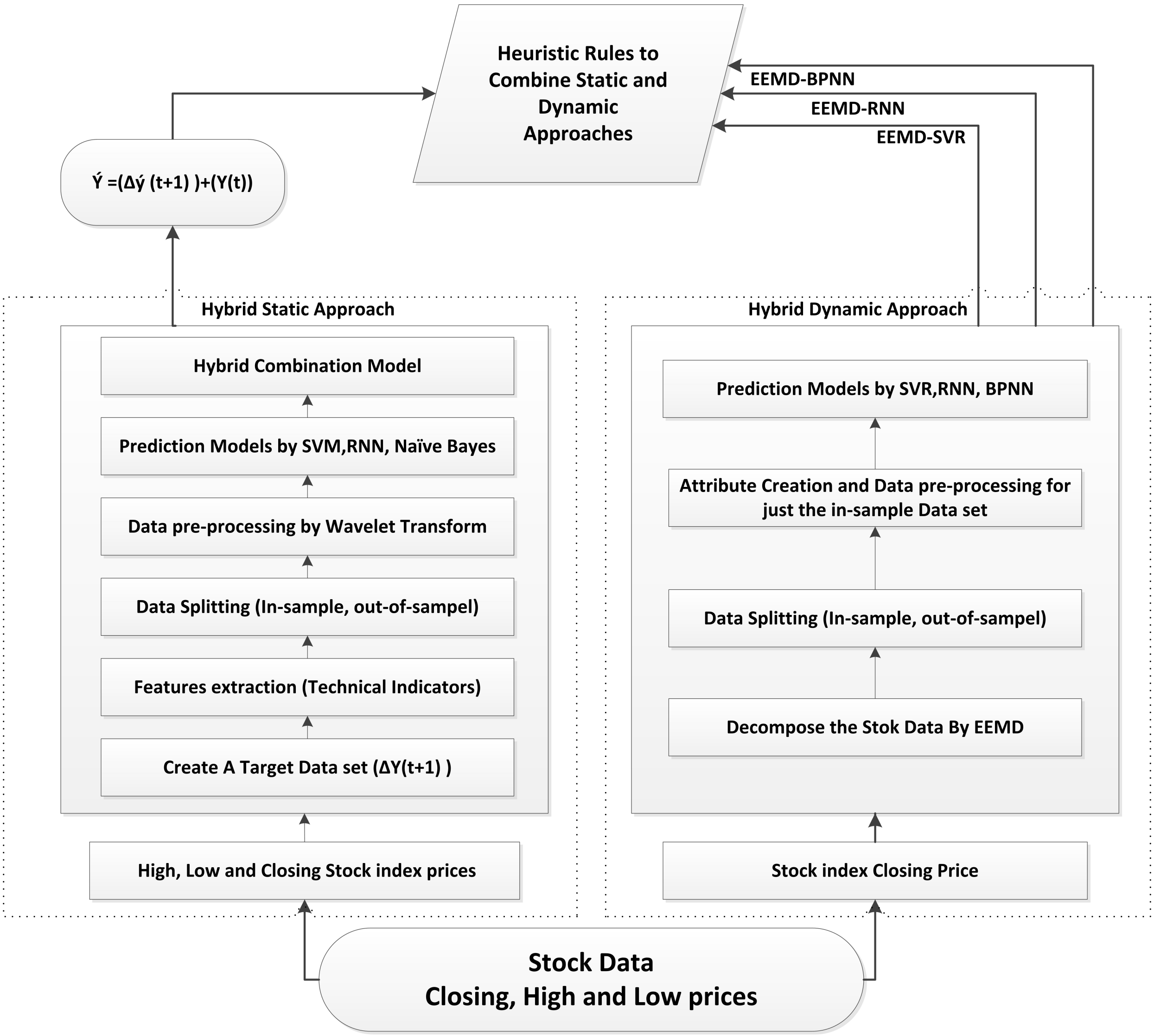

Figure 6.1: The Procedure of The Hybrid Heuristic Combination System.

6.2.1 Heuristic Rules

The heuristic approach is defined as that set of realistic and practical approaches that are easier to put into practice [255]. The general heuristic rule in this study is to combine the prediction results of the dynamic model where the prediction results of the static model are used as a factor in the fitness function alongside the threshold factor. After explaining the general purpose of the heuristic rules, these rules are transformed into IF-THEN rules. In this chapter the basic idea is to set rules for the system theory to convert heuristic rules into IF-THEN rules to build a combination model in order to enhance the prediction accuracy. The transformation into IF-THEN rules can be achieved by the following steps:

- Before the simulation is started, a few steps are required to prepare the input data sets. As Figure 6.1 illustrates, there are four inputs to the heuristic combination method; one is from the static model and the rest are from the dynamic model. The first preprocessing step is to transform the output of the static model $\Delta\hat{y}_{t+1}$ into $\hat{Y}_t$ as equation 6.1 illustrates.

$$\hat{Y}_t = (\Delta\hat{y}_{t+1} + Y_t) \tag{6.1}$$

  Where $\Delta\hat{y}_{t+1}$ is the trend predicted value from the static model, Y_t is the previous closing day price of the stock index and then $\hat{Y}_t$ is the predicted closing price of the day t.

- The second step is making sure that the same data size of the in-sample out-of-sample data sets remain the same as the one in the dynamic and static models.

- The third step is setting up the IF-THEN rules, after transforming the input data set and splitting the input data sets into in-sample and out-of-sample data sets. The IF-THEN rules base can be set. In the presented experimental protocol the aim is to combine the three dynamic models where the static model is used as a reference to determine the absolute error in the fitness function.

After implementing these steps, the next step is to generate the new IF-THEN rules. Equations 6.2 and 6.3 illustrate the input and the output which will be defined as an ultimate rule. Moreover, the threshold factor is determined by trial and error techniques based on the RMSE of the in-sample (training) data set.

$$f^t_{HHRS} = \left(f^t_{EEMD-SVR} + f^t_{EEMD-PBNN} + f^t_{EEMD-RNN}\right) \div 3 \tag{6.2}$$

$$Abs - E = \hat{Y}_t - f^t_{HHRS} \tag{6.3}$$

Where in Equation 6.2 f^t_{HHRS} is the predicted value of Hybrid heuristic Rules System (HHRS), $f^t_{EEMD-SVR}$ is the predicted value of EEMD-SVR model, $f^t_{EEMD-PBNN}$ is the predicted value of EEMD-PBNN and $f^t_{EEMD-RNN}$ is the predicted value of EEMD-RNN. $Abs - E$ is the absolute error of $\hat{Y}_t$ the static model obtained predicted value of the next day closing price of the day t - the f^t_{HHRS}.

Rule i:

IF $Abs - E_i$ is $> Threshold_i$, Then $f^t_{HHRS} = \left(f^t_{EEMD-SVR} + f^t_{EEMD-PBNN}\right) \div 2$

IF $Abs - E_i$ is $> Threshold_i$, Then $f^t_{HHRS} = \left(f^t_{EEMD-SVR} + f^t_{EEMD-RNN}\right) \div 2$

IF $Abs - E_i$ is $> Threshold_i$, Then $f^t_{HHRS} = \left(f^t_{EEMD-PBNN} + f^t_{EEMD-RNN}\right) \div 2$

$Abs - E_i$ and $Threshold_i$ are the term of the IF-THEN rules. Thus in the process of in-sample data, the main task is to find the suitable threshold factor for the fitness equation 6.3. Figure 6.2 represents the operation process of the IF-THEN rules. Thus the steps are explained step-by-step in the following points:

- Set up the IF-THEN rules parameters and initialise a set of rules.
- Generate a current rule and start the combination procedures.
- Search for the solution rules that obtain the minimum error.
- Remove the least accurate. Then repeat the previous point.
- After the best IF-THEN rule is determined, the predicted value is generated and the model performance is measured.

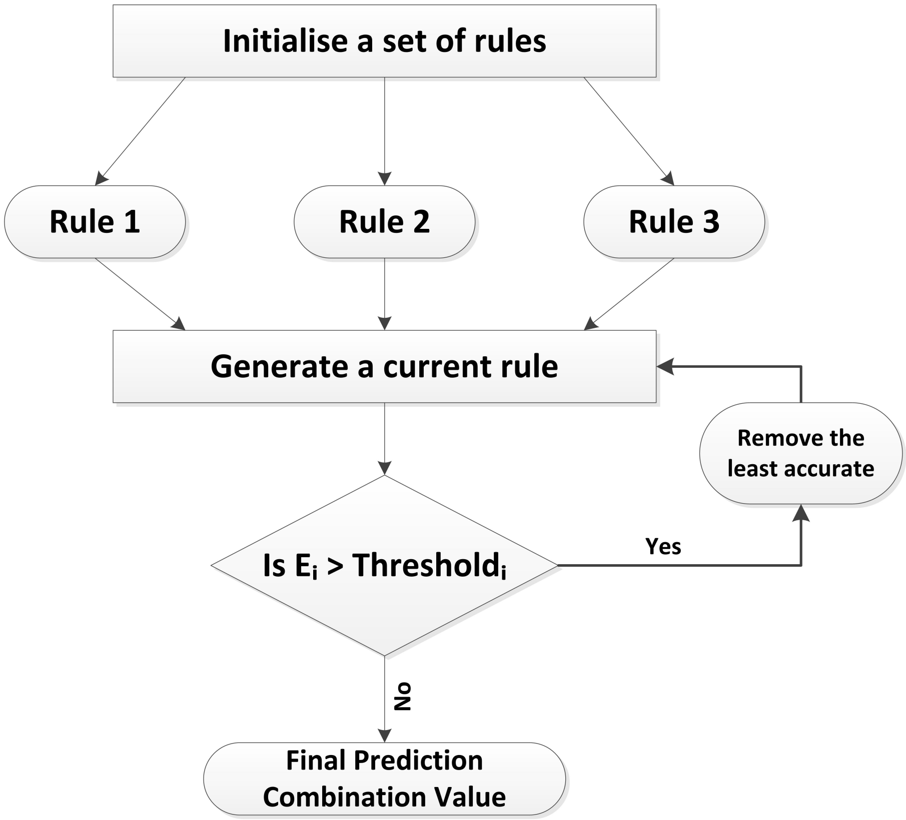

Figure 6.2: The Operation Process of IF-THEN Rules.

6.3 Results and Discussions

The proposed system HHRS is implemented in this section using the predicted values from earlier explained hybrid models, dynamic model and static model. Thus, the proposed HHRS system performance is compared with the benchmark model AR from Chapter 3. The SA model is used to combine the single approach models in Chapter 3 and the EEMD-GA-WA model which is presented in Chapter 4, in order to demonstrate the validity of the proposed system for predicting the FTSE 100, S&P 500 and Nikkei 225 next day closing prices.

Table 6.1 represents the FTSE 100 closing price index prediction results of the testing data set using the SA, AR, EEMD-GA-WA and HHRS models. SA, AR and EEMD-GA-WA are listed in Table 6.1 in order to compare them with the HHRS proposed system. Moreover, the threshold was determined by a set of trial and error conducted to choose the best value of the threshold; different ranges of values were conducted to choose the optimal threshold value from a range of values. The optimal value was selected based on the lowest RMSE on the training data set. Hence, the determined threshold for modelling FTSE 100, Nikkei 225 and S&P 500 are respectively, 30, 550 and 10. It can be found from Table 6.1 that the MSE, RMSE and MAE of the HHRS system for the FTSE 100 predicted values are, respectively, 35.79, 5.98 and 3.52. Thus, it can be observed that these results are the smallest in comparison with the other listed models. This indicates that there is smaller deviation between the actual and predicted values utilising the proposed system HHRS. Moreover, compared to the obtained predicted results from the single approach in Table 3.14 and the hybrid proposed models in Table 4.2, the HHRS system obtained the lowest MSE, RMSE and MAE. The cross-correlation coefficient R results for all of the prediction models for the FTSE 100 indicate that the prediction values and the actual values do not deviate too much. Hence, the HHRS system has outperformed all of the proposed models in this thesis significantly.

In addition, the predicted values for the testing data set of the S&P 500 closing price index are illustrated in Table 6.1. Table 6.1 shows the MSE, RMSE and MAE values which are respectively, 77.76, 8.81 and 7.00 for the HHRS system. Thus, it can be observed that these results contain the smallest error between all of the utilised models. This indicates that there is smaller deviation between the predicted and actual values utilising the proposed model HHRS. In comparison with the obtained results from the single approach in Chapter 3 Table 3.14 and the hybrid proposed models in Table 4.2, the HHRS results have the lowest error, which implies that the proposed model in this chapter outperformed the single approach. Table 6.1 presents the cross correlation coefficient R results for the S&P 500, and it can be observed that the prediction values and the actual values do not deviate too much.

Table 6.1 presents the predicted value of the Nikkei 225 closing price for the testing data set. As the table shows, the results of HHRS have the lowest error in comparison with all the utilised models, which indicates that there is a smaller deviation between the actual and the predicted values using HHRS. In addition, compared to the single approach results in Table 3.14 and the hybrid proposed models in Table 4.2, HHRS outperformed all of the single approach models. Moreover, the results of the cross- correlation coefficient R results for the Nikkei 225 indicate that the deviation between the actual and the predicted values is not that much.

Thus, it can be concluded from the results illustrated in Table 6.1 that the proposed model HHRS has outperformed all of the utilised approaches in this study in all of the data sets.

Table 6.1: The prediction result of testing data sets for FTSE 100, S&P 500 and Nikkei 225 using HHRS, EEMD-GA-WA, EEMD-SA, AR and SA.

Index name	**Models**	**MSE**	**RMSE**	**MAE**	**R**
FTSE100	SA	2033.77	45.09	35.99	0.95
	AR	1815.46	42.60	31.95	0.95
	EEMD-GA-WA	518.90	22.77	17.53	0.99
	HHRS	35.79	5.98	3.52	0.99
S&P500	SA	267.22	16.34	12.21	0.96
	AR	240.56	15.51	11.68	0.96
	EEMD-GA-WA	118.46	10.88	7.92	0.99
	HHRS	77.76	8.81	7.00	0.99
Nikkei225	SA	34761.99	186.44	135.32	0.99
	AR	34806.38	186.56	134.88	0.99
	EEMD-GA-WA	7254.00	85.95	68.47	0.99
	HHRS	6883.62	82.96	64.67	0.99

The predicted values of the FTSE 100 by the SA, AR, EEMD-GA-WA and HHRS models are shown in Figure 6.3. It can be observed from Figure 6.3 that the predicted values obtained from the proposed model HHRS are closer to the actual price than the other predicted values of the other models.

The actual S&P 500 closing price values and predicted values of the SA, AR, EEMD-GA-WA and HHRS models are presented in Figure 6.4. It can be observed from Figure 6.4 that the predicted values from the HHRS model are closer to the actual values of the S&P 500 than the rest of the models from Figure 3.26 and from Figure 4.7.

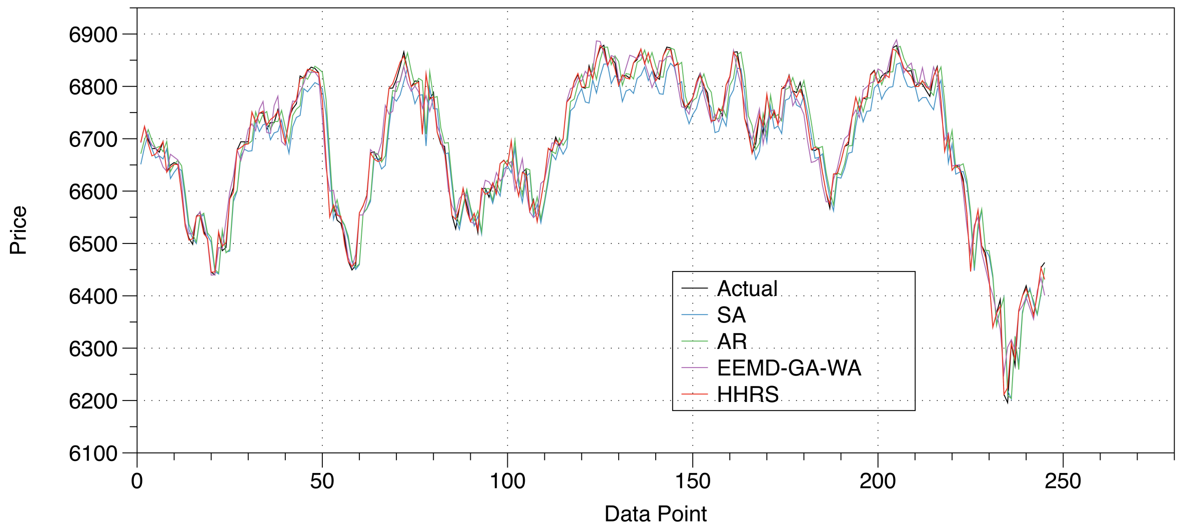

Figure 6.3: The actual FTSE 100 closing price Index and its predicted values from SA, AR, EEMD-BPNN, EEMD-GA-WA and HHRS models.

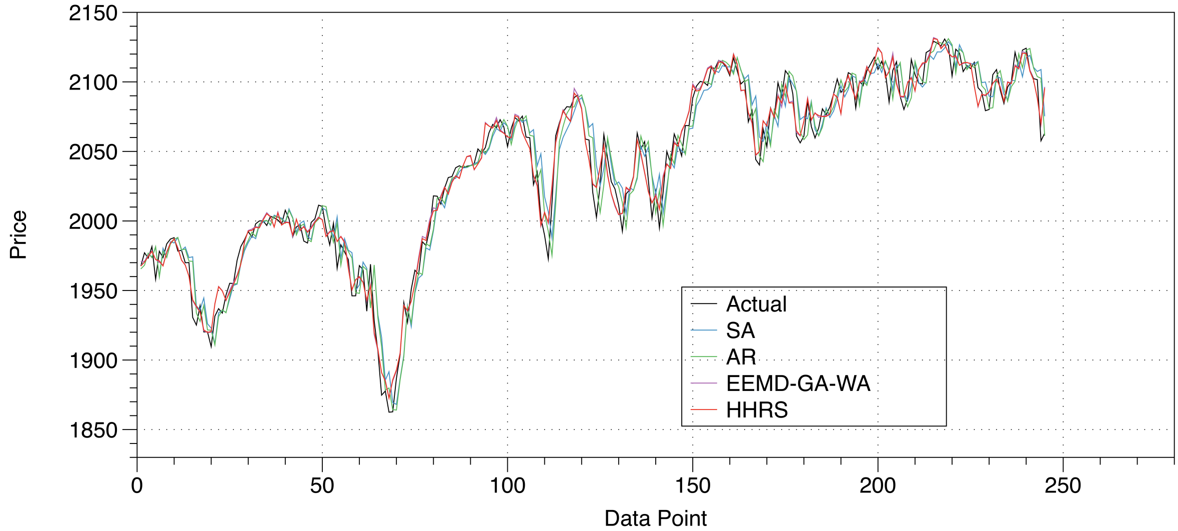

Figure 6.4: The actual S&P 500 closing price Index and its predicted values from SA, AR, EEMD-BPNN, EEMD-GA-WA and HHRS models.

Figure 6.5 depicts the actual Nikkei 225 closing price values and predicted values from the SA, AR, EEMD-GA-WA and HHRS models. From the Figure 6.5 it can be observed that the predicted values of all the models have smaller deviations with the actual values. Thus, in comparison with single approach predicted values from Figure 3.24 and the hybrid model from Figure 4.8, the results in Figure 6.5 indicate that the HHRS model predicted value is closer to the actual Nikkei 225 closing prices. Thus, it can be concluded that the predicted values of the proposed models in this chapter are closer to the actual values than those of the single approach models in chapter 3 and the hybrid approach models in Chapter 4.

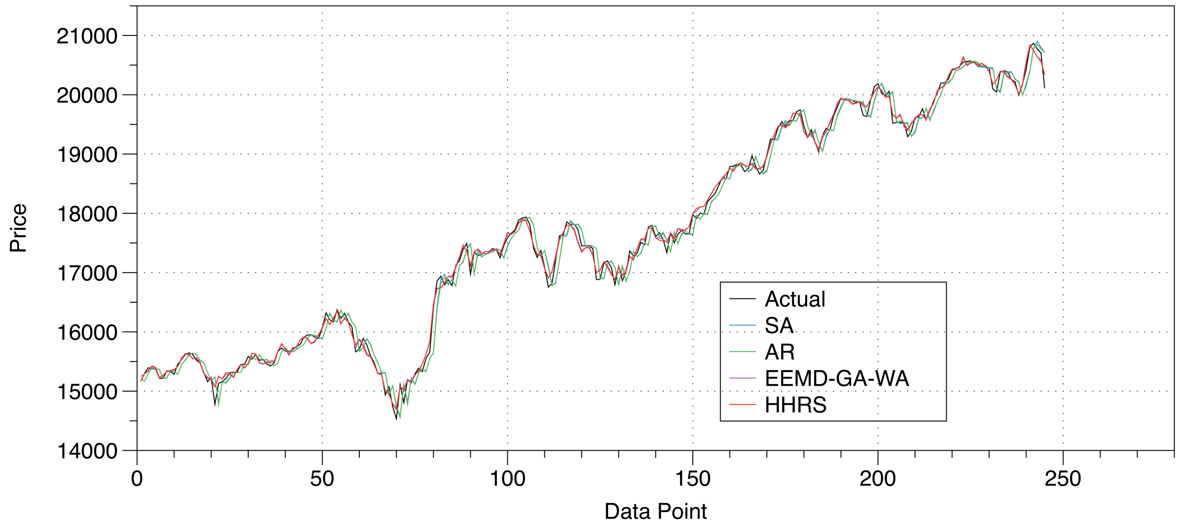

Figure 6.5: The actual Nikkei 225 closing price Index and its predicted values from SA, AR, EEMD-BPNN, EEMD-GA-WA and HHRS models.

6.4 Summary

This chapter presents a new approach based on the IF-ThEN rules system for building a stock price predicting expert system HHRS, and thus improving the prediction accuracy is the main aim. In the first phase a hybrid dynamic approach is proposed to predict the next day closing price for the chosen data set. Therefore, this phase is divided into three stages. Stage one: decompose the stock data by EEMD into a different number of IMFs. The second stage is predicting each IMF by RNN, PBNN and SVR. The third stage is calculating the sum of the IMF in order to obtain the final prediction results. Thus, this phase is presented in detail in Chapter 4. In the second phase, a static hybrid model was built where technical indicators were used as model input variables. This phase is also divided into three stages. The first stage is to create target data sets; this approach intends to predict the exact change of stock price, and then use the statistical tool to generate the technical indicators. The second stage is transforming the data sets by WT and then predicting it by using the SVM, RNN and Naive Bayes AI techniques. The third stage is combining the predicted results of the three AI techniques by WA, where GA is used as an optimiser to determine the weights. Moreover this phase is explained step-by-step in Chapter 5. In this chapter, the static model is used as a support model to enhance the prediction accuracy, where it is used as a reference in the fitness function in the proposed model. According to a survey of 692 fund managers in five countries by Menkhoff [183], the majority of fund managers rely on technical analysis. Thus, this study highlights the importance of technical indicators in supporting the decision makers in the market to make a decision with regard to selling or buying in the market. Phase Three is presented and explained in the earlier sections, where the output of Phases 1 and 2 are fed into the IF-then rules approach in order to combine the predicted value and enhance the final prediction results. The HHRS approach has the following new features:

- Transforming the predicted direction next day price into a full day next day price.
- For accuracy purposes, creating the IF-THEN rules in order to build a hybrid prediction expert system.
- In order to assure that the created rules are working with optimum solutions, the transformed predicted values which are generated from technical indicators (static approach) are used in the fitness function.

To evaluate the performance of the HHRS system, comparisons of different models such as SA, AR and EEME-GA-WA are listed in Table 6.1 and other previous tables from Chapter 4. As can be observed from the comparisons with different utilised methods, the HHRS system prediction results are much better than those from PBNN, RNN, SVR, AR, SA, EEMD-SVR, EEMD-RNN, EEMD-PBNN, SA-EEMD and EEMD-GA-WA, which confirms that the HHRS system is the best.

Chapter 7

Conclusions and Further Work

7.1 Introduction

This chapter presents the conclusion that was reached based on the accumulated findings of this thesis. Firstly, the undertaken work in this thesis is summarised and the key contributions are highlighted. Furthermore, the limitations of the research are discussed in order to generate future work suggestions. The remainder of this chapter is organised as follows: section 7.2 provides the overall conclusion of the thesis and highlights the main contributions. Finally, Section 7.3 demonstrates the suggested further works.

7.2 Conclusions

The main purpose of this research is to build a model that is able to predict financial time series considering their dynamic and complex nature. As mentioned earlier in Chapter 3, the prediction process will be carried out into three phases. In Chapter 1 two important question have been answered in this thesis in terms of predicting financial times series with specific methods and benefiting from the prediction of financial time series. The modelling of financial time series at early stages of prediction focuses on building a single approach model to predict the real-world financial time series. Thus at this stage the parameters of the single approach are determined as in 3 stage 1 of phases one and two. Some new factors have been introduced in order to enhance the prediction ability of the single approach model and the results show that after introducing the factor the prediction error reduced. However, as the aim of this research is to reach optimal and satisfactory results, terminating the research here was not an option.

A new approach was introduced in Chapter 4 using EEMD for the first time as an preprocessing tool. Three new prediction step approaches were introduced to predict the next day closing price. It can be concluded from the empirical results that the proposed hybrid method EEMD-GA-WA has outperformed the single approach model and the two-stage prediction models.

In addition, stock price direction was predicted using a three-step hybrid model. WT was used in order to remove the noise from the historical observations. To the best of the researchers knowledge, Chapter 5 provided the literature with a new prediction approach where the exact change in the closing price was successfully predicted with satisfactory results. Furthermore, the aim of the proposed method is to build an integrated hybrid system, WT-GA-WA, to predict the exact change of the next day closing price. To demonstrate the validity of the model, real data sets were used: FTSE 100, Nikkei 225 and S&P 500. It can be concluded from the empirical results that WT-GA-WA achieved the best results among all the utilised methods.

The second fundamental question raised in Chapter 1 has been answered in Chapter 6 in terms of combining dynamic models with static ones; the advantages of this integration can be noticed in the enhanced prediction values. The empirical results were obtained by implementing the proposed models HHRS to predict the FTSE 100, Nikkei 225 and S&P 500, following the proposed prediction procedures. The suggested model has achieved the best prediction performance among all the methods utilised in the thesis.

7.3 Future Works

The work of the above research has by no mean answered all existing suggestions nor solved all problems surrounding such a system. Therefore, many extensions may take place.

- Generalisation:

  Three data sets were used in this study. Therefore, and in order to generalise and validate the proposed model, different types of data set should be taken into consideration for future research.

- Complexity of the model (computational time):

  In order to reduce the computational time, further research should focus on alternative solutions to tackle such issues. For example, multi-core system could be one of the solutions, or the use of a Field-Programmable Gate Array (FPGA).

- Introduce new variables:

  As there are many factors influencing the stock price, future research should consider more input variables in order to achieve better accuracy.

References

[1] Ajith Abraham, Baikunth Nath, and Prabhat Kumar Mahanti. Hybrid intelligent systems for stock market analysis. In *Computational science-ICCS 2001*, pages 337–345. Springer, 2001.

[2] Felix Abramovich, Panagiotis Besbeas, and Theofanis Sapatinas. Empirical bayes approach to block wavelet function estimation. *Computational Statistics & Data Analysis*, 39(4):435–451, 2002.

[3] Yaser S Abu-Mostafa and Amir F Atiya. Introduction to financial forecasting. *Applied Intelligence*, 6(3):205–213, 1996.

[4] AA Adebiyi, CK Ayo, M Adebiyi, and SO Otokiti. An improved stock price prediction using hybrid market indicators. *African Journal of Computing & ICT*, 5(5):124–135, 2012.

[5] AA Adebiyi, CK Ayo, Marion O Adebiyi, and SO Otokiti. Stock price prediction using neural network with hybridized market indicators. *Journal of Emerging Trends in Computing and Information Sciences*, 3(1):1–9, 2012.

[6] Saeed Aghabozorgi and Ying Wah Teh. Stock market co-movement assessment using a three-phase clustering method. *Expert Systems with Applications*, 41(4):1301–1314, 2014.

[7] Mahmood Moein Aldin, Hasan Dehghan Dehnavi, and Somayye Entezari. Evaluating the employment of technical indicators in predicting stock price index variations using artificial neural networks (case study: Tehran stock exchange). *International Journal of Business and Management*, 7(15):p25, 2012.

[8] Sidney S Alexander. Price movements in speculative markets: Trends or random walks. *Industrial Management Review (pre-1986)*, 2(2):7, 1961.

[9] Edward I Altman. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of finance*, 23(4):589–609, 1968.

[10] Miguel A Ario et al. Forecasting time series via discrete wavelet transform. Technical report, Society for Computational Economics, 1998.

[11] Giuliano Armano, Michele Marchesi, and Andrea Murru. A hybrid genetic-neural architecture for stock indexes forecasting. *Information Sciences*, 170(1):3–33, 2005.

[12] J Scott Armstrong. Combining forecasts: The end of the beginning or the beginning of the end? *International Journal of Forecasting*, 5(4):585–588, 1989.

[13] Jon Scott Armstrong. *Principles of forecasting: a handbook for researchers and practitioners*, volume 30. Springer Science & Business Media, 2001.

[14] Shahrokh Asadi, Esmaeil Hadavandi, Farhad Mehmanpazir, and Mohammad Masoud Nakhostin. Hybridization of evolutionary levenberg–marquardt neural networks and data pre-processing for stock market prediction. *Knowledge-Based Systems*, 35:245–258, 2012.

[15] George S Atsalakis and Kimon P Valavanis. Surveying stock market forecasting techniques–part ii: Soft computing methods. *Expert Systems with Applications*, 36(3):5932–5941, 2009.

[16] Alexandre Aussem, Jonathan Campbell, and Fionn Murtagh. Wavelet-based feature extraction and decomposition strategies for financial forecasting. *Journal of Computational Intelligence in Finance*, 6(2):5–12, 1998.

[17] E Michael Azoff. *Neural network time series forecasting of financial markets*. John Wiley & Sons, Inc., 1994.

[18] Pedro Paulo Balestrassi, E Popova, AP de Paiva, and JW Marangon Lima. Design of experiments on neural network's training for nonlinear time series forecasting. *Neurocomputing*, 72(4):1160–1178, 2009.

[19] John M Bates and Clive WJ Granger. The combination of forecasts. *Or*, pages 451–468, 1969.

[20] William H Beaver. Financial ratios as predictors of failure. *Journal of accounting research*, pages 71–111, 1966.

[21] Michael J Berry and Gordon Linoff. *Data mining techniques: for marketing, sales, and customer support*. John Wiley & Sons, Inc., 1997.

[22] Ranjeeta Bisoi and Pradipta K Dash. A hybrid evolutionary dynamic neural network for stock market trend analysis and prediction using unscented kalman filter. *Applied Soft Computing*, 19:41–56, 2014.

[23] Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327, 1986.

[24] George EP Box and Norman Richard Draper. *Empirical model-building and response surfaces*, volume 424. Wiley New York, 1987.

[25] George EP Box and Gwilym M Jenkins. *Time series analysis: forecasting and control, revised ed*. Holden-Day, 1976.

[26] Anthony Brabazon and Peter B Keenan. A hybrid genetic model for the prediction of corporate failure. *Computational Management Science*, 1(3-4):293–310, 2004.

[27] William Brock, Josef Lakonishok, and Blake LeBaron. Simple technical trading rules and the stochastic properties of stock returns. *Journal of finance*, pages 1731–1764, 1992.

[28] Peter J Brockwell and Richard A Davis. *Introduction to time series and forecasting*. Springer Science & Business Media, 2006.

[29] Steven Busuttil and Yuri Kalnishkan. Weighted kernel regression for predicting changing dependencies. In *Machine Learning: ECML 2007*, pages 535–542. Springer, 2007.

[30] Kirt C Butler and SJ Malaikah. Efficiency and inefficiency in thinly traded stock markets: Kuwait and saudi arabia. *Journal of Banking & Finance*, 16(1):197–210, 1992.

[31] Li-Juan Cao and Francis EH Tay. Support vector machine with adaptive parameters in financial time series forecasting. *Neural Networks, IEEE Transactions on*, 14(6):1506–1518, 2003.

[32] Qing Cao, Karyl B Leggio, and Marc J Schniederjans. A comparison between fama and french's model and artificial neural networks in predicting the chinese stock market. *Computers & Operations Research*, 32(10):2499–2512, 2005.

[33] Oscar Castillo and Patricia Melin. Simulation and forecasting complex economic time series using neural networks and fuzzy logic. In *Neural Networks, 2001. Proceedings. IJCNN'01. International Joint Conference on*, volume 3, pages 1805–1810. IEEE, 2001.

[34] Oscar Castillo and Patricia Melin. Hybrid intelligent systems for time series prediction using neural networks, fuzzy logic, and fractal theory. *Neural Networks, IEEE Transactions on*, 13(6):1395–1408, 2002.

[35] Chih-Chung Chang and Chih-Jen Lin. Libsvm: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3):27, 2011.

[36] Fi-John Chang and Ya-Ting Chang. Adaptive neuro-fuzzy inference system for prediction of water level in reservoir. *Advances in Water Resources*, 29(1):1–10, 2006.

[37] Jing-Rong Chang, Liang-Ying Wei, and Ching-Hsue Cheng. A hybrid anfis model based on ar and volatility for taiex forecasting. *Applied Soft Computing*, 11(1):1388–1395, 2011.

[38] Pei-Chann Chang et al. A novel model by evolving partially connected neural network for stock price trend forecasting. *Expert Systems with Applications*, 39(1):611–620, 2012.

[39] Pei-Chann Chang, T Warren Liao, Jyun-Jie Lin, and Chin-Yuan Fan. A dynamic threshold decision system for stock trading signal detection. *Applied Soft Computing*, 11(5):3998–4010, 2011.

[40] Pei-Chann Chang and Chen-Hao Liu. A tsk type fuzzy rule based system for stock price prediction. *Expert Systems with applications*, 34(1):135–144, 2008.

[41] Pei-Chann Chang, Yen-Wen Wang, and Wen-Ning Yang. An investigation of the hybrid forecasting models for stock price variation in taiwan. *Journal of the Chinese Institute of Industrial Engineers*, 21(4):358–368, 2004.

[42] Tsung-Sheng Chang. A comparative study of artificial neural networks, and decision trees for digital game content stocks price prediction. *Expert Systems with Applications*, 38(12):14846–14851, 2011.

[43] C Chatfield. What is the bestmethod of forecasting? *Journal of Applied Statistics*, 15(1):19–38, 1988.

[44] Chris Chatfield. *The analysis of time series: an introduction*. CRC press, 2013.

[45] Yves Chauvin and David E Rumelhart. *Backpropagation: theory, architectures, and applications*. Psychology Press, 1995.

[46] An-Sing Chen, Mark T Leung, and Hazem Daouk. Application of neural networks to an emerging financial market: forecasting and trading the taiwan stock index. *Computers & Operations Research*, 30(6):901–923, 2003.

[47] CH Chen. Neural networks for financial market prediction. In *Neural Networks, 1994. IEEE World Congress on Computational Intelligence., 1994 IEEE International Conference on*, volume 2, pages 1199–1202. IEEE, 1994.

[48] Fang Chen, Chengling Gou, Xiaoqian Guo, and Jieping Gao. Prediction of stock markets by the evolutionary mix-game model. *Physica A: Statistical Mechanics and its Applications*, 387(14):3594–3604, 2008.

[49] Mu-Yen Chen and Bo-Tsuen Chen. A hybrid fuzzy time series model based on granular computing for stock price forecasting. *Information Sciences*, 294:227–241, 2015.

[50] Shu-Heng Chen. Genetic algorithms and genetic programming in computational finance: An overview of the book. In *Genetic Algorithms and Genetic Programming in Computational Finance*, pages 1–26. Springer, 2002.

[51] Shyi-Ming Chen. Forecasting enrollments based on fuzzy time series. *Fuzzy sets and systems*, 81(3):311–319, 1996.

[52] You-Shyang Chen and Ching-Hsue Cheng. Evaluating industry performance using extracted rgr rules based on feature selection and rough sets classifier. *Expert Systems with Applications*, 36(5):9448–9456, 2009.

[53] You-Shyang Chen, Ching-Hsue Cheng, and Wei-Lun Tsai. Modeling fitting-function-based fuzzy time series patterns for evolving stock index forecasting. *Applied intelligence*, 41(2):327–347, 2014.

[54] Cheng Cheng, Wei Xu, and Jiajia Wang. A comparison of ensemble methods in financial market prediction. In *Computational Sciences and Optimization (CSO), 2012 Fifth International Joint Conference on*, pages 755–759. IEEE, 2012.

[55] Ching-Hsue Cheng, Guang-Wei Cheng, and Jia-Wen Wang. Multi-attribute fuzzy time series method based on fuzzy clustering. *Expert Systems with Applications*, 34(2):1235–1242, 2008.

[56] Ching-Hsue Cheng and Liang-Ying Wei. A novel time-series model based on empirical mode decomposition for forecasting taiex. *Economic Modelling*, 36:136–141, 2014.

[57] Ching-Hsue Cheng, Liang-Ying Wei, and You-Shyang Chen. Fusion anfis models based on multi-stock volatility causality for taiex forecasting. *Neurocomputing*, 72(16):3462–3468, 2009.

[58] Chongcheul Cheong, Young-Jae Kim, and Seong-Min Yoon. Can we predict exchange rate movements at short horizons? *Journal of Forecasting*, 31(7):565–579, 2012.

[59] Vladimir Cherkassky and Yunqian Ma. Practical selection of svm parameters and noise estimation for svm regression. *Neural networks*, 17(1):113–126, 2004.

[60] Sheng-Chai Chi, Wei-Ling Peng, Pei-Tsang Wu, and Ming-Wei Yu. The study on the relationship among technical indicators and the development of stock index prediction system. In *Fuzzy Information Processing Society, 2003. NAFIPS 2003. 22nd International Conference of the North American*, pages 291–296. IEEE, 2003.

[61] Charisios Christodoulos, Christos Michalakelis, and Dimitris Varoutas. Forecasting with limited data: Combining arima and diffusion models. *Technological forecasting and social change*, 77(4):558–565, 2010.

[62] RT Clemen. Combining forecasts: A review and annotated bibliography', international journal of forecasting, 5 (4), 559-83. *INTERNATIONAL LIBRARY OF CRITICAL WRITINGS IN ECONOMICS*, 108:132–156, 1999.

[63] Albert Cohen, Ingrid Daubechies, and Pierre Vial. Wavelets on the interval and fast wavelet transforms. *Applied and Computational Harmonic Analysis*, 1(1):54–81, 1993.

[64] Paul H Cootner. The random character of stock market prices. 1964.

[65] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine learning*, 20(3):273–297, 1995.

[66] Paul SP Cowpertwait and Andrew V Metcalfe. *Introductory time series with R*. Springer Science & Business Media, 2009.

[67] Wensheng Dai, Jui-Yu Wu, and Chi-Jie Lu. Combining nonlinear independent component analysis and neural network for the prediction of asian stock market indexes. *Expert systems with applications*, 39(4):4444–4452, 2012.

[68] Jan G De Gooijer and Rob J Hyndman. 25 years of time series forecasting. *International journal of forecasting*, 22(3):443–473, 2006.

[69] Jan G De Gooijer and Kuldeep Kumar. Some recent developments in non-linear time series modelling, testing, and forecasting. *International Journal of Forecasting*, 8(2):135–156, 1992.

[70] Fagner A de Oliveira, Cristiane N Nobre, and Luis E Zárate. Applying artificial neural networks to prediction of stock price and improvement of the directional prediction index–case study of petr4, petrobras, brazil. *Expert Systems with Applications*, 40(18):7596–7606, 2013.

[71] Fagner Andrade De Oliveira, Luis Enrique Zárate, Marcos de Azevedo Reis, and Cristiane Neri Nobre. The use of artificial neural networks in the analysis and prediction of stock prices. In *Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on*, pages 2151–2155. IEEE, 2011.

[72] Vijay S Desai and Rakesh Bharati. The efficacy of neural networks in predicting returns on stock and bond indices*. *Decision Sciences*, 29(2):405–423, 1998.

[73] Luis A Diaz-Robles, Juan C Ortega, Joshua S Fu, Gregory D Reed, Judith C Chow, John G Watson, and Juan A Moncada-Herrera. A hybrid arima and artificial neural networks model to forecast particulate matter in urban areas: the case of temuco, chile. *Atmospheric Environment*, 42(35):8331–8340, 2008.

[74] Francis X Diebold and Robert S Mariano. Comparing predictive accuracy. *Journal of Business & economic statistics*, 2012.

[75] AI Diler. Predicting direction of ise national-100 index with back propagation trained neural network. *Journal of Istanbul Stock Exchange*, 7(25-26):65–81, 2003.

[76] Juan Peralta Donate, Paulo Cortez, Germán Gutiérrez Sánchez, and Araceli Sanchis De Miguel. Time series forecasting using a weighted cross-validation evolutionary artificial neural network ensemble. *Neurocomputing*, 109:27–32, 2013.

[77] Huang Dongdong and Zeng Wenhong. The forecast of price index based on wavelet neural network. In *Business Intelligence and Financial Engineering (BIFE), 2011 Fourth International Conference on*, pages 32–36. IEEE, 2011.

[78] Garland B Durham. Sv mixture models with application to s&p 500 index returns. *Journal of Financial Economics*, 85(3):822–856, 2007.

[79] Goutam Dutta, Pankaj Jha, Arnab Kumar Laha, and Neeraj Mohan. Artificial neural network models for forecasting stock price index in the bombay stock exchange. *Journal of Emerging Market Finance*, 5(3):283–295, 2006.

[80] Robert D Edwards and John Magee. Technical analysis of stock trends (new york institute of finance, business information and publishing, new york). *EdwardsTechnical Analysis of Stock Trends1992*, 1992.

[81] Jeffrey L Elman. Finding structure in time. *Cognitive science*, 14(2):179–211, 1990.

[82] Robert Engle. Garch 101: The use of arch/garch models in applied econometrics. *Journal of economic perspectives*, pages 157–168, 2001.

[83] Robert F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007, 1982.

[84] David Enke and Suraphan Thawornwong. The use of data mining and neural networks for forecasting stock market returns. *Expert Systems with applications*, 29(4):927–940, 2005.

[85] Eugene F Fama. The behavior of stock-market prices. *Journal of business*, pages 34–105, 1965.

[86] Eugene F Fama. Efficient capital markets: A review of theory and empirical work*. *The journal of Finance*, 25(2):383–417, 1970.

[87] Eugene F Fama. Efficient capital markets: Ii. *The journal of finance*, 46(5):1575–1617, 1991.

[88] Eugene F Fama. Random walks in stock market prices. *Financial analysts journal*, 51(1):75–80, 1995.

[89] Durdu Ömer Faruk. A hybrid neural network and arima model for water quality time series prediction. *Engineering Applications of Artificial Intelligence*, 23(4):586–594, 2010.

[90] Patrick Flandrin, Gabriel Rilling, and Paulo Goncalves. Empirical mode decomposition as a filter bank. *Signal Processing Letters, IEEE*, 11(2):112–114, 2004.

[91] Jack Clark Francis and Eric Kirzner. *Investments: analysis and management.* McGraw-Hill New York, 1991.

[92] Wayne A Fuller. *Introduction to statistical time series*, volume 428. John Wiley & Sons, 2009.

[93] Liam A Gallagher and Mark P Taylor. Permanent and temporary components of stock prices: Evidence from assessing macroeconomic shocks. *Southern Economic Journal*, pages 345–362, 2002.

[94] A García-Almanza. *New classification methods for gathering patterns in the context of genetic programming.* PhD thesis, PhD Thesis, Department of Computing and Electronic Systems, University of Essex, 2008.

[95] Fernando Gaxiola, Patricia Melin, Fevrier Valdez, and Oscar Castillo. Interval type-2 fuzzy weight adjustment for backpropagation neural networks with application in time series prediction. *Information Sciences*, 260:1–14, 2014.

[96] R Gençay, F Selçuk, and B Whitcher. An introduction to wavelets and other filtering methods in finance and economics. *Waves in Random Media*, 12(3):399–399, 2002.

[97] Ramazan Gençay, Faruk Selçuk, and Brandon Whitcher. Differentiating intraday seasonalities through wavelet multi-scaling. *Physica A: Statistical Mechanics and its Applications*, 289(3):543–556, 2001.

[98] Ramazan Gençay, Faruk Selçuk, and Brandon J Whitcher. *An introduction to wavelets and other filtering methods in finance and economics.* Academic press, 2001.

[99] Mircea Gherman, Romulus Terebes, and Monica Borda. Time series analysis using wavelets and gjr-garch models. In *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*, pages 2138–2142. IEEE, 2012.

[100] Antonio Glaria Bengoechea. Stock market indices in santiago de chile: forecasting using neural networks. In *Neural Networks, 1996., IEEE International Conference on*, volume 4, pages 2172–2175. IEEE, 1996.

[101] David E Golberg. Genetic algorithms in search, optimization, and machine learning. *Addion wesley*, 1989, 1989.

[102] António Gorgulho, Rui Neves, and Nuno Horta. Applying a ga kernel on optimizing technical analysis rules for stock picking and portfolio composition. *Expert systems with Applications*, 38(11):14072–14085, 2011.

[103] Kousik Guhathakurta, Indranil Mukherjee, and A Roy Chowdhury. Empirical mode decomposition analysis of two different financial time series and their comparison. *Chaos, Solitons & Fractals*, 37(4):1214–1227, 2008.

[104] Sunil Gupta and Peter C Wilton. Combination of forecasts: An extension. *Management Science*, 33(3):356–372, 1987.

[105] Erkam Guresen, Gulgun Kayakutlu, and Tugrul U Daim. Using artificial neural network models in stock market index prediction. *Expert Systems with Applications*, 38(8):10389–10397, 2011.

[106] Esmaeil Hadavandi, Arash Ghanbari, and Salman Abbasian-Naghneh. Developing an evolutionary neural network model for stock index forecasting. In *Advanced Intelligent Computing Theories and Applications*, pages 407–415. Springer, 2010.

[107] Esmaeil Hadavandi, Hassan Shavandi, and Arash Ghanbari. Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting. *Knowledge-Based Systems*, 23(8):800–808, 2010.

[108] Reza Hafezi, Jamal Shahrabi, and Esmaeil Hadavandi. A bat-neural network multi-agent system (bnnmas) for stock price prediction: Case study of dax stock price. *Applied Soft Computing*, 29:196–210, 2015.

[109] H Haleh, B Akbari Moghaddam, and S Ebrahimijam. A new approach to forecasting stock price with ekf data fusion. *International Journal of Trade, Economics and Finance*, 2(2):49–58, 2011.

[110] James W Hall. Adaptive selection of us stocks with neural nets. *Trading on the edge: neural, genetic, and fuzzy systems for chaotic financial markets. New York: Wiley*, pages 45–65, 1994.

[111] James D Hamilton. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the Econometric Society*, pages 357–384, 1989.

[112] Jiawei Han, Micheline Kamber, and Jian Pei. *Data mining: concepts and techniques: concepts and techniques*. Elsevier, 2011.

[113] Md Rafiul Hassan, Kotagiri Ramamohanarao, Joarder Kamruzzaman, Mustafizur Rahman, and M Maruf Hossain. A hmm-based adaptive fuzzy inference system for stock market forecasting. *Neurocomputing*, 104:10–25, 2013.

[114] Thomas Hellström and Kenneth Holmström. Predictable patterns in stock returns. In *University, Department of*. Citeseer, 1998.

[115] Thomas HellstrSom and Kenneth HolmstrSom. Predicting the stock market. 1998.

[116] Tim Hill, Marcus O'Connor, and William Remus. Neural network models for time series forecasts. *Management science*, 42(7):1082–1092, 1996.

[117] John Henry Holland. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. MIT press, 1992.

[118] David A Hsieh. Chaos and nonlinear dynamics: application to financial markets. *The journal of finance*, 46(5):1839–1877, 1991.

[119] Tsung-Jung Hsieh, Hsiao-Fen Hsiao, and Wei-Chang Yeh. Forecasting stock markets using wavelet transforms and recurrent neural networks: An integrated system based on artificial bee colony algorithm. *Applied soft computing*, 11(2):2510–2525, 2011.

[120] Chih-Ming Hsu. A hybrid procedure for stock price prediction by integrating self-organizing map and genetic programming. *Expert Systems with Applications*, 38(11):14026–14036, 2011.

[121] Chih-Wei Hsu, Chih-Chung Chang, Chih-Jen Lin, et al. A practical guide to support vector classification, 2003.

[122] Sheng-Hsun Hsu, JJ Po-An Hsieh, Ting-Chih Chih, and Kuei-Chu Hsu. 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, 2009.

[123] Chenyi Hu and Ling T He. An application of interval methods to stock market forecasting. *Reliable Computing*, 13(5):423–434, 2007.

[124] Cheng-Lung Huang and Cheng-Yi Tsai. A hybrid sofm-svr with a filter-based feature selection for stock market forecasting. *Expert Systems with Applications*, 36(2):1529–1539, 2009.

[125] Chenn-Jung Huang, Dian-Xiu Yang, and Yi-Ta Chuang. Application of wrapper approach and composite classifier to the stock trend prediction. *Expert Systems with Applications*, 34(4):2870–2878, 2008.

[126] Chien-Feng Huang. A hybrid stock selection model using genetic algorithms and support vector regression. *Applied Soft Computing*, 12(2):807–818, 2012.

[127] Ji-Ping Huang. Prediction: Pure technical analysis might not work satisfactorily. In *Experimental Econophysics*, pages 167–180. Springer, 2015.

[128] Norden E Huang, Zheng Shen, Steven R Long, Manli C Wu, Hsing H Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung, and Henry H Liu. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 454, pages 903–995. The Royal Society, 1998.

[129] Wei Huang, Yoshiteru Nakamori, and Shou-Yang Wang. Forecasting stock market movement direction with support vector machine. *Computers & Operations Research*, 32(10):2513–2522, 2005.

[130] Kunhuang Huarng. Effective lengths of intervals to improve forecasting in fuzzy time series. *Fuzzy sets and systems*, 123(3):387–394, 2001.

[131] Rob J Hyndman and Anne B Koehler. Another look at measures of forecast accuracy. *International journal of forecasting*, 22(4):679–688, 2006.

[132] Huseyin Ince and Theodore B Trafalis. Short term forecasting with support vector machines and application to stock price prediction. *International Journal of General Systems*, 37(6):677–687, 2008.

[133] Michael C Jensen. Some anomalous evidence regarding market efficiency. *Journal of financial economics*, 6(2):95–101, 1978.

[134] Angelos Kanas. Non-linear forecasts of stock returns. *Journal of Forecasting*, 22(4):299–315, 2003.

[135] Ling-Jing Kao, Chih-Chou Chiu, Chi-Jie Lu, and Chih-Hsiang Chang. A hybrid approach by integrating wavelet-based feature extraction with mars and svr for stock index forecasting. *Decision Support Systems*, 54(3):1228–1244, 2013.

[136] Ling-Jing Kao, Chih-Chou Chiu, Chi-Jie Lu, and Jung-Li Yang. Integration of nonlinear independent component analysis and support vector regression for stock price forecasting. *Neurocomputing*, 99:534–542, 2013.

[137] Yakup Kara, Melek Acar Boyacioglu, and Ömer Kaan Baykan. Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the istanbul stock exchange. *Expert systems with Applications*, 38(5):5311–5319, 2011.

[138] Nilesh N Karnik and Jerry M Mendel. Applications of type-2 fuzzy logic systems to forecasting of time-series. *Information Sciences*, 120(1):89–111, 1999.

[139] Manolis G Kavussanos and Everton Dockery. A multivariate test for stock market efficiency: the case of ase. *Applied Financial Economics*, 11(5):573–579, 2001.

[140] Ahmad Kazem, Ebrahim Sharifi, Farookh Khadeer Hussain, Morteza Saberi, and Omar Khadeer Hussain. Support vector regression with chaos-based firefly algorithm for stock market price forecasting. *Applied Soft Computing*, 13(2):947–958, 2013.

[141] David Keller. *Breakthroughs in Technical Analysis: New Thinking From the World's Top Minds*, volume 61. John Wiley & Sons, 2010.

[142] Maurice George Kendall and A Bradford Hill. The analysis of economic time-series-part i: Prices. *Journal of the Royal Statistical Society. Series A (General)*, 116(1):11–34, 1953.

[143] Hyun-jung Kim and Kyung-shik Shin. A hybrid approach based on neural networks and genetic algorithms for detecting temporal patterns in stock markets. *Applied Soft Computing*, 7(2):569–576, 2007.

[144] Kyoung-jae Kim. Financial time series forecasting using support vector machines. *Neurocomputing*, 55(1):307–319, 2003.

[145] Kyoung-jae Kim and Ingoo Han. Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. *Expert systems with Applications*, 19(2):125–132, 2000.

[146] YongSeog Kim and W Nick Street. An intelligent system for customer targeting: a data mining approach. *Decision Support Systems*, 37(2):215–228, 2004.

[147] Charles D Kirkpatrick II and Julie Dahlquist. *Technical analysis: the complete resource for financial market technicians*. FT press, 2010.

[148] Yoshinori Kishikawa and Shozo Tokinaga. Prediction of stock trends by using the wavelet transform and the multi-stage fuzzy inference system optimized by the ga. *IEiCE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, 83(2):357–366, 2000.

[149] Judy L Klein. Statistical visions in time. *Cambridge: Cambridge University*, 1997.

[150] Boris Kovalerchuk and Evgenii Vityaev. Data mining for financial applications. In *Data Mining and Knowledge Discovery Handbook*, pages 1203–1224. Springer, 2005.

[151] John R Koza. *Genetic programming: on the programming of computers by means of natural selection*, volume 1. MIT press, 1992.

[152] Miroslav Kubat. Neural networks: a comprehensive foundation by simon haykin, macmillan, 1994, isbn 0-02-352781-7., 1999.

[153] Manish Kumar and M Thenmozhi. Forecasting stock index movement: A comparison of support vector machines and random forest. In *Indian Institute of Capital Markets 9th Capital Markets Conference Paper*, 2006.

[154] Manish Kumar and M Thenmozhi. Support vector machines approach to predict the s&p cnx nifty index returns. In *10th Capital Markets Conference, Indian Institute of Capital Markets Paper*, 2007.

[155] Yung-Keun Kwon and Byung-Ro Moon. Evolutionary ensemble for stock prediction. In *Genetic and Evolutionary Computation–GECCO 2004*, pages 1102–1113. Springer, 2004.

[156] Monica Lam. Neural network techniques for financial performance prediction: integrating fundamental and technical analysis. *Decision Support Systems*, 37(4):567–581, 2004.

[157] Blake LeBaron. Technical trading rules and regime shifts in foreign exchange. *Advanced trading rules*, pages 5–40, 1998.

[158] Blake LeBaron and Andreas S Weigend. A bootstrap evaluation of the effect of data splitting on financial time series. *Neural Networks, IEEE Transactions on*, 9(1):213–220, 1998.

[159] Ming-Chi Lee. Using support vector machine with a hybrid feature selection method to the stock trend prediction. *Expert Systems with Applications*, 36(8):10896–10904, 2009.

[160] Yi-Shian Lee and Lee-Ing Tong. Forecasting time series using a methodology based on autoregressive integrated moving average and genetic programming. *Knowledge-Based Systems*, 24(1):66–72, 2011.

[161] William Leigh, Naval Modani, and Ross Hightower. A computational implementation of stock charting: abrupt volume increase as signal for movement in new york stock exchange composite index. *Decision Support Systems*, 37(4):515–530, 2004.

[162] George Lekakos and George M Giaglis. A hybrid approach for improving predictive accuracy of collaborative filtering algorithms. *User Modeling and User-Adapted Interaction*, 17(1-2):5–40, 2007.

[163] Mark T Leung, Hazem Daouk, and An-Sing Chen. Forecasting stock indices: a comparison of classification and level estimation models. *International Journal of Forecasting*, 16(2):173–190, 2000.

[164] Jin Li, Zhu Shi, and Xiaoli Li. Genetic programming with wavelet-based indicators for financial forecasting. *Transactions of the Institute of Measurement and Control*, 28(3):285–297, 2006.

[165] Tao Li, Qi Li, Shenghuo Zhu, and Mitsunori Ogihara. A survey on wavelet applications in data mining. *ACM SIGKDD Explorations Newsletter*, 4(2):49–68, 2002.

[166] Yuhong Li and Weihua Ma. Applications of artificial neural networks in financial economics: a survey. In *Computational Intelligence and Design (ISCID), 2010 International Symposium on*, volume 1, pages 211–214. IEEE, 2010.

[167] Xiaowei Lin, Zehong Yang, and Yixu Song. Short-term stock price prediction based on echo state networks. *Expert systems with applications*, 36(3):7313–7317, 2009.

[168] Shi-Miin Liu and Chih-Hsien Chou. Parities and spread trading in gold and silver markets: A fractional cointegration analysis. *Applied Financial Economics*, 13(12):899–911, 2003.

[169] Andrew W Lo. The adaptive markets hypothesis: Market efficiency from an evolutionary perspective. *Journal of Portfolio Management, Forthcoming*, 2004.

[170] Andrew W Lo and Jasmina Hasanhodzic. *The evolution of technical analysis: financial prediction from Babylonian tablets to Bloomberg terminals*, volume 139. John Wiley & Sons, 2011.

[171] Andrew W Lo, Harry Mamaysky, and Jiang Wang. Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation. Technical report, National bureau of economic research, 2000.

[172] Chi-Jie Lu, Chih-Hsiang Chang, Chien-Yu Chen, Chih-Chou Chiu, and Tian-Shyug Lee. Stock index prediction: A comparison of mars, bpn and svr in an emerging market. In *Industrial Engineering and Engineering Management, 2009. IEEM 2009. IEEE International Conference on*, pages 2343–2347. IEEE, 2009.

[173] Chi-Jie Lu, Tian-Shyug Lee, and Chih-Chou Chiu. Financial time series forecasting using independent component analysis and support vector regression. *Decision Support Systems*, 47(2):115–125, 2009.

[174] Richard Maclin, Jude W Shavlik, et al. Combining the predictions of multiple classifiers: Using competitive learning to initialize neural networks. In *IJCAI*, pages 524–531. Citeseer, 1995.

[175] Spyros Makridakis. Why combining works? *International Journal of Forecasting*, 5(4):601–603, 1989.

[176] Spyros Makridakis, A Andersen, Robert Carbone, Robert Fildes, Michele Hibon, Rudolf Lewandowski, Joseph Newton, Emanuel Parzen, and Robert Winkler. The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *Journal of forecasting*, 1(2):111–153, 1982.

[177] Burton G Malkiel. The efficient market hypothesis and its critics. *Journal of economic perspectives*, pages 59–82, 2003.

[178] Burton Gordon Malkiel. *A random walk down Wall Street: including a life-cycle guide to personal investing.* WW Norton & Company, 1999.

[179] Jianchang Mao. A case study on bagging, boosting and basic ensembles of neural networks for ocr. In *Neural Networks Proceedings, 1998. IEEE World Congress on Computational Intelligence. The 1998 IEEE International Joint Conference on*, volume 3, pages 1828–1833. IEEE, 1998.

[180] Vadim Marmer. Nonlinearity, nonstationarity, and spurious forecasts. *Journal of Econometrics*, 142(1):1–27, 2008.

[181] A Marszałek and T Burczyński. Modeling and forecasting financial time series with ordered fuzzy candlesticks. *Information Sciences*, 273:144–155, 2014.

[182] Serafin Martinez-Jaramillo and Edward PK Tsang. An heterogeneous, endogenous and coevolutionary gp-based financial market. *Evolutionary Computation, IEEE Transactions on*, 13(1):33–55, 2009.

[183] Lukas Menkhoff. The use of technical analysis by fund managers: International evidence. *Journal of Banking & Finance*, 34(11):2573–2586, 2010.

[184] Kai Miao, Fang Chen, and ZG Zhao. Stock price forecast based on bacterial colony rbf neural network [j]. *Journal of Qingdao University (Natural Science Edition)*, 20(2):50–54, 2007.

[185] Ryszard S Michalski, Jaime G Carbonell, and Tom M Mitchell. *Machine learning: An artificial intelligence approach*. Springer Science & Business Media, 2013.

[186] Terence C Mills. *Forecasting Financial Markets (Volume 2)*. Edward Elgar, 2002.

[187] Terence C Mills and Raphael N Markellos. *The econometric modelling of financial time series*. Cambridge University Press, 2008.

[188] Terence C Mills and TC Mills. *Predicting the Unpredictable?: Science and Guesswork in Financial Market Forecasting*, volume 87. Institute of Economic Affairs London, 1992.

[189] Sung-Hwan Min, Jumin Lee, and Ingoo Han. Hybrid genetic algorithms and support vector machines for bankruptcy prediction. *Expert systems with applications*, 31(3):652–660, 2006.

[190] John J Murphy. *Technical analysis of the financial markets: A comprehensive guide to trading methods and applications*. Penguin, 1999.

[191] Arman Khadjeh Nassirtoussi, Teh Ying Wah, and David Ngo Chek Ling. A novel forex prediction methodology based on fundamental data. *African Journal of Business Management*, 5(20):8322–8330, 2011.

[192] Paul Newbold and Clive WJ Granger. Experience with forecasting univariate time series and the combination of forecasts. *Journal of the Royal Statistical Society. Series A (General)*, pages 131–165, 1974.

[193] Kyong Joo Oh and Kyoung-jae Kim. Analyzing stock market tick data using piecewise nonlinear model. *Expert Systems with Applications*, 22(3):249–255, 2002.

[194] Dennis Olson and Charles Mossman. Neural network forecasts of canadian stock returns using accounting ratios. *International Journal of Forecasting*, 19(3):453–465, 2003.

[195] Edgar Osuna, Robert Freund, and Federico Girosi. Support vector machines: Training and applications. 1997.

[196] Phichhang Ou and Hengshan Wang. Prediction of stock market index movement by ten data mining techniques. *Modern Applied Science*, 3(12):p28, 2009.

[197] Ping-Feng Pai. System reliability forecasting by support vector machines with genetic algorithms. *Mathematical and Computer Modelling*, 43(3):262–274, 2006.

[198] Ping-Feng Pai and Chih-Sheng Lin. A hybrid arima and support vector machines model in stock price forecasting. *Omega*, 33(6):497–505, 2005.

[199] Ping-Feng Pai, Kuo-Ping Lin, Chi-Shen Lin, and Ping-Teng Chang. Time series forecasting by a seasonal support vector regression model. *Expert Systems with Applications*, 37(6):4261–4265, 2010.

[200] Franz C Palm and Arnold Zellner. To combine or not to combine? issues of combining forecasts. *Journal of Forecasting*, 11(8):687–701, 1992.

[201] Konstantina Papagiannaki, Nina Taft, Zhi-Li Zhang, and Christophe Diot. Long-term forecasting of internet backbone traffic. *Neural Networks, IEEE Transactions on*, 16(5):1110–1124, 2005.

[202] Kamban Parasuraman and Amin Elshorbagy. Wavelet networks: an alternative to classical neural networks. In *Neural Networks, 2005. IJCNN'05. Proceedings. 2005 IEEE International Joint Conference on*, volume 5, pages 2674–2679. IEEE, 2005.

[203] Cheol-Ho Park and Scott H Irwin. What do we know about the profitability of technical analysis? *Journal of Economic Surveys*, 21(4):786–826, 2007.

[204] Kanghee Park and Hyunjung Shin. Stock price prediction based on a complex interrelation network of economic factors. *Engineering Applications of Artificial Intelligence*, 26(5):1550–1561, 2013.

[205] Jigar Patel, Sahil Shah, Priyank Thakkar, and K Kotecha. Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. *Expert Systems with Applications*, 42(1):259–268, 2015.

[206] Jigar Patel, Sahil Shah, Priyank Thakkar, and K Kotecha. Predicting stock market index using fusion of machine learning techniques. *Expert Systems with Applications*, 42(4):2162–2172, 2015.

[207] Dan W Patterson. *Artificial neural networks: theory and applications*. Prentice Hall PTR, 1998.

[208] Raúl Pino, José Parreno, Alberto Gomez, and Paolo Priore. Forecasting next-day price of electricity in the spanish energy market using artificial neural networks. *Engineering Applications of Artificial Intelligence*, 21(1):53–62, 2008.

[209] Massimiliano Pontil and Alessandro Verri. Properties of support vector machines. *Neural Computation*, 10(4):955–974, 1998.

[210] Ademola Popoola and Khurshid Ahmad. Testing the suitability of wavelet preprocessing for tsk fuzzy models. In *Fuzzy Systems, 2006 IEEE International Conference on*, pages 1305–1309. IEEE, 2006.

[211] Martin J Pring. *Technical analysis explained: The successful investor's guide to spotting investment trends and turning points*. McGraw-Hill Professional, 2002.

[212] Martha Pulido, Patricia Melin, and Oscar Castillo. Particle swarm optimization of ensemble neural networks with fuzzy aggregation for time series prediction of the mexican stock exchange. *Information Sciences*, 280:188–204, 2014.

[213] Bo Qian and Khaled Rasheed. Stock market prediction with multiple classifiers. *Applied Intelligence*, 26(1):25–33, 2007.

[214] James B Ramsey. The contribution of wavelets to the analysis of economic and financial data. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 357(1760):2593–2606, 1999.

[215] James B Ramsey and Camille Lampart. The decomposition of economic relationships by time scale using wavelets: expenditure and income. *Studies in Nonlinear Dynamics & Econometrics*, 3(1), 1998.

[216] James B Ramsey and Zhifeng Zhang. The analysis of foreign exchange data using waveform dictionaries. *Journal of Empirical Finance*, 4(4):341–372, 1997.

[217] T Subba Rao and MM Gabr. *An introduction to bispectral analysis and bilinear time series models*, volume 24. Springer Science & Business Media, 2012.

[218] Apostolos Nicholas Refenes, Achileas Zapranis, and Gavin Francis. Stock performance modeling using neural networks: a comparative study with regression models. *Neural networks*, 7(2):375–388, 1994.

[219] David J Reid. Combining three estimates of gross domestic product. *Economica*, pages 431–444, 1968.

[220] Na Ren, M Zargham, and Sanaz Rahimi. A decision tree-based classification approach to rule extraction for security analysis. *International Journal of Information Technology & Decision Making*, 5(01):227–240, 2006.

[221] Tae Hyup Roh. Forecasting the volatility of stock price index. *Expert Systems with Applications*, 33(4):916–922, 2007.

[222] Huntley Schaller and Simon Van Norden. Regime switching in stock market returns. *Applied Financial Economics*, 7(2):177–191, 1997.

[223] Yochanan Shachmurove, Dorota Witkowska, et al. *Utilizing artificial neural network model to predict stock markets*. University of Pennsylvania, Center for Analytic Research in Economics and the Social Sciences, 2000.

[224] Adel Sharkasi, Martin Crane, Heather J Ruskin, and Jose A Matos. The reaction of stock markets to crashes and events: A comparison study between emerging and mature markets using wavelet transforms. *Physica A: Statistical Mechanics and its Applications*, 368(2):511–521, 2006.

[225] Wei Shen, Xiaopen Guo, Chao Wu, and Desheng Wu. Forecasting stock indices using radial basis function neural networks optimized by artificial fish swarm algorithm. *Knowledge-Based Systems*, 24(3):378–385, 2011.

[226] Wei Shen, Xiaopen Guo, Chao Wu, and Desheng Wu. Forecasting stock indices using radial basis function neural networks optimized by artificial fish swarm algorithm. *Knowledge-Based Systems*, 24(3):378–385, 2011.

[227] Wei Shen, Yunyun Zhang, and Xiaoyong Ma. Stock return forecast with ls-svm and particle swarm optimization. In *Business Intelligence and Financial Engineering, 2009. BIFE'09. International Conference on*, pages 143–147. IEEE, 2009.

[228] Shuzhen Shi, Wenlong Liu, and Minglu Jin. Stock price forecasting using a hybrid arma and bp neural network and markov model. In *Communication Technology (ICCT), 2012 IEEE 14th International Conference on*, pages 981–985. IEEE, 2012.

[229] SR Singh. A simple method of forecasting based on fuzzy time series. *Applied mathematics and computation*, 186(1):330–339, 2007.

[230] Manfred Steiner and Hans-Georg Wittkemper. Portfolio optimization with a neural network implementation of the coherent market hypothesis. *European journal of operational research*, 100(1):27–40, 1997.

[231] Ingo Steinwart and Andreas Christmann. *Support vector machines*. Springer Science & Business Media, 2008.

[232] Jie Sun and Hui Li. Financial distress prediction using support vector machines: Ensemble vs. individual. *Applied Soft Computing*, 12(8):2254–2265, 2012.

[233] Clarence NW Tan. A hybrid financial trading system incorporating chaos theory, statistical and artificial intelligence/soft computing methods. In *at the Queensland Finance Conference*, 1999.

[234] Tuan Zea Tan, Chai Quek, and Geok See Ng. Brain-inspired genetic complementary learning for stock market prediction. In *Evolutionary Computation, 2005. The 2005 IEEE Congress on*, volume 3, pages 2653–2660. IEEE, 2005.

[235] Mieko Tanaka-Yamawaki and Seiji Tokuoka. Adaptive use of technical indicators for the prediction of intra-day stock prices. *Physica A: Statistical Mechanics and its Applications*, 383(1):125–133, 2007.

[236] Leonard J Tashman. Out-of-sample tests of forecasting accuracy: an analysis and review. *International Journal of Forecasting*, 16(4):437–450, 2000.

[237] Francis EH Tay and Lijuan Cao. Application of support vector machines in financial time series forecasting. *Omega*, 29(4):309–317, 2001.

[238] Mark P Taylor and Helen Allen. The use of technical analysis in the foreign exchange market. *Journal of international Money and Finance*, 11(3):304–314, 1992.

[239] Paolo Tenti. Forecasting foreign exchange rates using recurrent neural networks. *Applied Artificial Intelligence*, 10(6):567–582, 1996.

[240] Suraphan Thawornwong and David Enke. The adaptive selection of financial and economic variables for use with artificial neural networks. *Neurocomputing*, 56:205–232, 2004.

[241] Konstantinos Theofilatos, Andreas Karathanasopoulos, Peter Middleton, Efstratios Georgopoulos, and Spiros Likothanassis. Modeling and trading ftse100 index using a novel sliding window approach which combines adaptive differential evolution and support vector regression. In *Artificial Intelligence Applications and Innovations*, pages 486–496. Springer, 2013.

[242] Allan Timmermann. Forecast combinations. *Handbook of economic forecasting*, 1:135–196, 2006.

[243] Allan Timmermann and Clive WJ Granger. Efficient market hypothesis and forecasting. *International Journal of Forecasting*, 20(1):15–27, 2004.

[244] Maria E Torres, Marcelo Colominas, Gastón Schlotthauer, Patrick Flandrin, et al. A complete ensemble empirical mode decomposition with adaptive noise. In *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on*, pages 4144–4147. IEEE, 2011.

[245] Chih-Fong Tsai and Yu-Chieh Hsiao. Combining multiple feature selection methods for stock prediction: Union, intersection, and multi-intersection approaches. *Decision Support Systems*, 50(1):258–269, 2010.

[246] Chih-Fong Tsai, Yuah-Chiao Lin, David C Yen, and Yan-Min Chen. Predicting stock returns by classifier ensembles. *Applied Soft Computing*, 11(2):2452–2459, 2011.

[247] Ray Tsaih, Yenshan Hsu, and Charles C Lai. Forecasting s&p 500 stock index futures with a hybrid ai system. *Decision Support Systems*, 23(2):161–174, 1998.

[248] Philip M Tsang, Paul Kwok, Steven O Choy, Reggie Kwan, SC Ng, Jacky Mak, Jonathan Tsang, Kai Koong, and Tak-Lam Wong. Design and implementation of nn5 for hong kong stock price forecasting. *Engineering Applications of Artificial Intelligence*, 20(4):453–461, 2007.

[249] Vladimir N Vapnik. An overview of statistical learning theory. *Neural Networks, IEEE Transactions on*, 10(5):988–999, 1999.

[250] Luis M Viceira. Bond risk, bond return volatility, and the term structure of interest rates. 2007.

[251] Jian-Zhou Wang, Ju-Jie Wang, Zhe-George Zhang, and Shu-Po Guo. Forecasting stock indices with back propagation neural network. *Expert Systems with Applications*, 38(11):14346–14355, 2011.

[252] Jian-Zhou Wang, Ju-Jie Wang, Zhe-George Zhang, and Shu-Po Guo. Forecasting stock indices with back propagation neural network. *Expert Systems with Applications*, 38(11):14346–14355, 2011.

[253] Jie Wang and Jun Wang. Forecasting stock market indexes using principle component analysis and stochastic time effective neural networks. *Neurocomputing*, 156:68–78, 2015.

[254] Ju-Jie Wang, Jian-Zhou Wang, Zhe-George Zhang, and Shu-Po Guo. Stock index forecasting based on a hybrid model. *Omega*, 40(6):758–766, 2012.

[255] Jue Wang, Abdel-Rahman Hedar, Shouyang Wang, and Jian Ma. Rough set and scatter search metaheuristic based feature selection for credit scoring. *Expert Systems with Applications*, 39(6):6123–6128, 2012.

[256] Tong Wang, Mingcai Zhang, Qihao Yu, and Huyuan Zhang. Comparing the applications of emd and eemd on time–frequency analysis of seismic signal. *Journal of Applied Geophysics*, 83:29–34, 2012.

[257] Wei Wang, Hong Zhao, Qiang Li, and Zhixiong Liu. A novel hybrid intelligent model for financial time series forecasting and its application. In *Business Intelligence and Financial Engineering, 2009. BIFE'09. International Conference on*, pages 279–282. IEEE, 2009.

[258] Yanshan Wang and In-Chan Choi. Market index and stock price direction prediction using machine learning techniques: An empirical study on the kospi and hsi. *arXiv preprint arXiv:1309.7119*, 2013.

[259] Yi-Fan Wang. Mining stock price using fuzzy rough set system. *Expert Systems with Applications*, 24(1):13–23, 2003.

[260] Gary R Weckman, Sriram Lakshminarayanan, Jon H Marvel, and Andy Snow. An integrated stock market forecasting model using neural networks. *International Journal of Business Forecasting and Marketing Intelligence*, 1(1):30–49, 2008.

[261] CH Wei. Hybrid learning fuzzy neural models in stock forecasting. *Journal of Information and Optimization Sciences*, 26(3):495–508, 2005.

[262] Liang-Ying Wei. A hybrid anfis model based on empirical mode decomposition for stock time series forecasting. *Applied Soft Computing*, 42:368–376, 2016.

[263] Liang-Ying Wei, Tai-Liang Chen, and Tien-Hwa Ho. A hybrid model based on adaptive-network-based fuzzy inference system to forecast taiwan stock market. *Expert Systems with Applications*, 38(11):13625–13631, 2011.

[264] Liang-Ying Wei, Ching-Hsue Cheng, and Hsin-Hung Wu. A hybrid anfis based on n-period moving average model to forecast taiex stock. *Applied Soft Computing*, 19:86–92, 2014.

[265] Robert Craig West. A factor-analytic approach to bank condition. *Journal of Banking & Finance*, 9(2):253–266, 1985.

[266] Halbert White. Economic prediction using neural networks: The case of ibm daily stock returns. In *Neural Networks, 1988., IEEE International Conference on*, pages 451–458. IEEE, 1988.

[267] Robert L Winkler. Combining forecasts: A philosophical basis and some current issues. *International Journal of Forecasting*, 5(4):605–609, 1989.

[268] David H Wolpert. The supervised learning no-free-lunch theorems. In *Soft Computing and Industry*, pages 25–42. Springer, 2002.

[269] David H Wolpert. The supervised learning no-free-lunch theorems. In *Soft Computing and Industry*, pages 25–42. Springer, 2002.

[270] David H Wolpert and William G Macready. No free lunch theorems for optimization. *Evolutionary Computation, IEEE Transactions on*, 1(1):67–82, 1997.

[271] Berlin Wu and Chih-Li Chang. Using genetic algorithms to parameters (d, r) estimation for threshold autoregressive models. *Computational Statistics & Data Analysis*, 38(3):315–330, 2002.

[272] Zhaohua Wu and Norden E Huang. A study of the characteristics of white noise using the empirical mode decomposition method. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 460, pages 1597–1611. The Royal Society, 2004.

[273] Zhaohua Wu and Norden E Huang. Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in adaptive data analysis*, 1(01):1–41, 2009.

[274] Yaqing Xia, Yulong Liu, and Zhiqian Chen. Support vector regression for prediction of stock trend. In *Information Management, Innovation Management and Industrial Engineering (ICIII), 2013 6th International Conference on*, volume 2, pages 123–126. IEEE, 2013.

[275] C Xiang and WM Fu. Predicting the stock market using multiple models. In *Control, Automation, Robotics and Vision, 2006. ICARCV'06. 9th International Conference on*, pages 1–6. IEEE, 2006.

[276] RN Yadav, Prem Kumar Kalra, and Joseph John. Time series prediction with single multiplicative neuron model. *Applied soft computing*, 7(4):1157–1163, 2007.

[277] Jian-Hui Yang and Xiao-zhen Dai. Prediction of shanghai and shenzhen 300 index based on emd-svm model. In *Systems and Informatics (ICSAI), 2012 International Conference on*, pages 1209–1212. IEEE, 2012.

[278] Qiang Yang and Xindong Wu. 10 challenging problems in data mining research. *International Journal of Information Technology & Decision Making*, 5(04):597–604, 2006.

[279] JingTao Yao and Joseph P Herbert. Financial time-series analysis with rough sets. *Applied Soft Computing*, 9(3):1000–1007, 2009.

[280] Jingtao Yao, Chew Lim Tan, and Hean-Lee Poh. Neural networks for technical analysis: a study on klci. *International journal of theoretical and applied finance*, 2(02):221–241, 1999.

[281] Chi-Yuan Yeh, Chi-Wei Huang, and Shie-Jue Lee. A multiple-kernel support vector regression approach for stock market price forecasting. *Expert Systems with Applications*, 38(3):2177–2186, 2011.

[282] Youngohc Yoon, George Swales Jr, and Thomas M Margavio. A comparison of discriminant analysis versus artificial neural networks. *Journal of the Operational Research Society*, pages 51–60, 1993.

[283] Honghai Yu and Haifei Liu. Improved stock market prediction by combining support vector machine and empirical mode decomposition. In *Computational Intelligence and Design (ISCID), 2012 Fifth International Symposium on*, volume 1, pages 531–534. IEEE, 2012.

[284] Hui-Kuang Yu. Weighted fuzzy time series models for taiex forecasting. *Physica A: Statistical Mechanics and its Applications*, 349(3):609–624, 2005.

[285] Lean Yu, Huanhuan Chen, Shouyang Wang, and Kin Keung Lai. Evolving least squares support vector machines for stock market trend mining. *Evolutionary Computation, IEEE Transactions on*, 13(1):87–102, 2009.

[286] Lotfi A Zadeh. The role of fuzzy logic in modeling, identification and control. *Modeling Identification and Control*, 15(3):191, 1994.

[287] MH Zarandi, Esmaeil Hadavandi, and IB Turksen. A hybrid fuzzy intelligent agent-based system for stock price prediction. *International Journal of Intelligent Systems*, 27(11):947–969, 2012.

[288] Defu Zhang, Qingshan Jiang, and Xin Li. Application of neural networks in financial data mining. In *International Conference on Computational Intelligence*, pages 392–395. Citeseer, 2004.

[289] Dongsong Zhang and Lina Zhou. Discovering golden nuggets: data mining in financial application. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 34(4):513–522, 2004.

[290] G Peter Zhang. Time series forecasting using a hybrid arima and neural network model. *Neurocomputing*, 50:159–175, 2003.

[291] Xiangzhou Zhang, Yong Hu, Kang Xie, Shouyang Wang, EWT Ngai, and Mei Liu. A causal feature selection algorithm for stock prediction modeling. *Neurocomputing*, 142:48–59, 2014.

[292] Y-Q Zhang, Somasheker Akkaladevi, G Vachtsevanos, and Tsau Young Lin. Granular neural web agents for stock prediction. *Soft Computing*, 6(5):406–413, 2002.

[293] Yudong Zhang and Lenan Wu. Stock market prediction of s&p 500 via combination of improved bco approach and bp neural network. *Expert systems with applications*, 36(5):8849–8854, 2009.

[294] Guo Zhiqiang, Wang Huaiqing, and Liu Quan. Financial time series forecasting using lpp and svm optimized by pso. *Soft Computing*, 17(5):805–818, 2013.

[295] Zhi-Hua Zhou, Yuan Jiang, Yu-Bin Yang, and Shi-Fu Chen. Lung cancer cell identification based on artificial neural network ensembles. *Artificial Intelligence in Medicine*, 24(1):25–36, 2002.

[296] Xiaotian Zhu, Hong Wang, Li Xu, and Huaizu Li. Predicting stock index increments by neural networks: The role of trading volume under different horizons. *Expert Systems with Applications*, 34(4):3043–3054, 2008.

[297] Yi Zuo and Eisuke Kita. Stock price forecast using bayesian network. *Expert Systems with Applications*, 39(8):6729–6737, 2012.