A New Hybrid Stock Index Trend Price Prediction Model

Bashar Alhnaity[a,*], Maysam Abbod[a]

[a] *Department of Electronic and Computer Engineering, Brunel University London, UK*

Abstract

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 [1]. 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.

Keywords: Wavelet Transform, Naive-Bayes, Stock Market, Neural Networks, Support Vector Machine, Genetic Algorithm

---

1. Introduction

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

*Corresponding author
Email address: `Alhnaity.Bashar@gmail.com` (Bashar Alhnaity)

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. Wavelet analysis is considered to be a relatively new field in signal processing [2]. 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 [3]. 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 [4]. 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 timescales [5]. 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) [6] [7]. Therefore, this paper 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 [8].

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 [9] [10] [11]. 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 [12] [13] [14] [15] [16] [17]. 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 [18]. SVM and LS-SVM prediction results outperformed the other proposed approaches. SVM was used in Kim's [19] 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 [20] 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 [19] 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 [21] 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) [22] 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. 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 [23].

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) [24] 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 [25]. 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 [26], 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 [27]. 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.[28] 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 [29]. Kazem's (2013) [30] 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 [31] 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 [32] 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 [1]. 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 [19] [33] [20]. 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 [34] 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 up-dating, 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 [35]. 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) [36]. In this study a model neural network was used and it achieved a 60.81% prediction rate for the ISE100 index.

Charts, technical indicators, the Dow theory, Gann lines and Elliot waves are the technical analysis tools used to exploit recurring patterns [37]. 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 [38]. 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 [39] [40] [41] [42] [43]. 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 [44] [45] [46]. 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 [47]. 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 [45].

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 [48]. SVM was used in Kim's (2003)[19] 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 [20] 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 [19] 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 [21] 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.

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 [49]. 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 [50] 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 [51] using an SVM ensemble. The SVM ensemble outperformed the individual SVM classifier. 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 [52]. In their study, the experiment result indicated that the proposed method outperformed the other used method.

This paper 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 paper is organised as follows: Section 4.2 describes the prediction procedure of the hybrid model. Subsection 3.4 introduces the WT method. The hybrid combination steps are illustrated in Subsection 4.3. Section 4.4 presents the results based on real data sets. Finally Section 5 concludes this paper.

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 [53]. 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 [50]. 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 [54]. 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 [55]. Murphy summarised the three premises of technical analysis as follows [40]:

- 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 [11]. 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 [46]:

- 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 1: Technical indicators used as input variables [56]

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 [57].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 [54] [58] [59] [28] [36] [60] [33] [20] [61] [62] [63]. The utilised technical indicators are formed and illustrated in Table 1 and Table 2. Thus, in order to obtain these technical indicators, the daily closing, high and low price of three datasets are utilised. Therefore, after implementing the formulas in Table 2, the technical indicators are generated. The first data set is the daily closing price of the FTSE 100 index 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, and it is also used to demonstrate and validate the predictability of the proposed 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 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 [64].

A statistical summary of the eight technical indicators was calculated and is presented in Table 3

Table 2: Technical indicators formulas [63]

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)	$100 - 100/1 + \sum_{i=0}^{n-1} UP_{t-i} \|n\| \sum_{i=0}^{n-1} Dw_{t-i}/n$
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.

Eight technical indicators are considered as input variables to predict the future daily trend in the stock price index. 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. In accordance to traditional time series regression setting, the t-th output value y_t is presented in Equation 1. However the inputs are the eight attribute which are the technical indicators and the target is Δy.

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

Table 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

3. Prediction models

3.1. 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 [66] [67]. 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:

$g\left(x\right) = w\emptyset\left(x\right) + b.\emptyset\left(x\right)$ 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.

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

$$L\left(d,y\right)=\begin{cases}|d-y|-\varepsilon & |d-y|\geq\varepsilon\\ 0 & other,\end{cases} \tag{3}$$

C and ε in Equation 2 and 3 are prescribed parameters. C is called the regularisation constant while ε is referred to as the regularisation constant. $L\left(d,y\right)$ Is the intensive loss function and the term $C\frac{1}{N}\sum_{i=1}^{N}L\left(d_i,y_i\right)$ 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 transformed to the following:

$$R\left(w,\zeta,\zeta^*\right)=\frac{1}{2}ww^T+C\times\left(\sum_{i=1}^{N}\left(\zeta,\zeta^*\right)\right) \tag{4}$$

Subject to:

$$w\phi\left(x_i\right)+b_i-d_i\leq\varepsilon+\zeta_{i^*} \tag{5}$$

$$d_i-w\phi\left(x_i\right)-b_i\leq\varepsilon+\zeta_i \tag{6}$$

$$\zeta_i,\zeta_{i^*}\geq 0 \tag{7}$$

The decision equation (kernel function) comes up finally after the Lagrange multipliers are introduced and optimality constraints exploited. Equation 8 is the form of kernel Equation:

$$f\left(x\right)=\sum_{i}^{l}\left(\alpha_i-{\alpha_i}'\right)K\left(x_i,x_j\right)+b \tag{8}$$

Where $\alpha_i{}'$ and α_i are called Lagrange multipliers in equation 8. 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 9.

$$K\left(x_i, x_j\right) = exp\left(-\gamma||x_i - x_j||^2\right) \tag{9}$$

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 [68].

3.2. 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: $P\left(C|X\right)$, from $P\left(C\right)$, $P\left(C|X\right)$ and $P\left(X\right)$. It is stated by Byes' theorem that

$$P\left(C|X\right) = \frac{P\left(X|C\right)P\left(C\right)}{P\left(X\right)} \tag{10}$$

The posterior probability $P\left(C|X\right)$ 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 $P\left(X|C\right)$ 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 10 illustrates how the data is classified.

$$P\left(C_i|X\right) = \frac{P\left(X|C_i\right)P\left(C_i\right)}{P\left(X\right)} \tag{11}$$

Having many attributes in the given data sets $A_1, A_2, ...A_n$, can reflect on the computational time to compute $P\left(X|C_i\right)$. Therefore, the solution to reducing computation when evaluating $P\left(X|C_i\right)$ 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.

$$lP\left(X|C_i\right) = \prod_{k=1}^{n} P\left(x_k|C_i\right) = P\left(x_1|C_i\right) \times P\left(x_2|C_i\right) \times .... \times P\left(x_n|C_i\right) \tag{12}$$

x_k in Equation 12 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 13:

$$f\left(x, \mu, \sigma\right) = \frac{1}{\sqrt[\sigma]{2\pi}} e^{-(x-\mu)^{2\backslash 2\sigma^2}} \tag{13}$$

so,

$$P\left(x_k|C_i\right) = f\left(x_k, \mu_{c_i}, \sigma_{c_i}\right) \tag{14}$$

μ_{c_i} and σ_{c_i}in Equation 14 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 13, 14 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 15 happened, then the class label of observation X can be predicted as class C_i.

$$P\left(X|C_i\right)P\left(C_i\right) > P\left(X|C_i\right)P\left(C_j\right) for 1 \leq j \leq m; j \neq i. \tag{15}$$

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 [63].

3.3. Recurrent Neural Network

Recurrent Neural Network (RNN) is considered as an enhanced ANN architecture, and thus a variant of Elman's network [69]. 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 [70], 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. Figure1 demonstrates the RNN architecture, where $x_t^{[n]}\left(n = 1, 2, ...k + 1\right), u_t^{[1]}, u_t^{[2]}$ are the inputs for the RNN model at time t including bias node. The output of RNN model is $\check{y}_t$ and $d_t^{[f]}\left(f = 1, 2\right)$, $w_t^{[n]}\left(n = 1, 2, ...k + 1\right)$ 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 :$K\left(x\right) = \frac{1}{1+e^{-x}}$ and S in the Figure above presents the linear function: $J\left(x\right) = \sum_i x_i$.Function 16

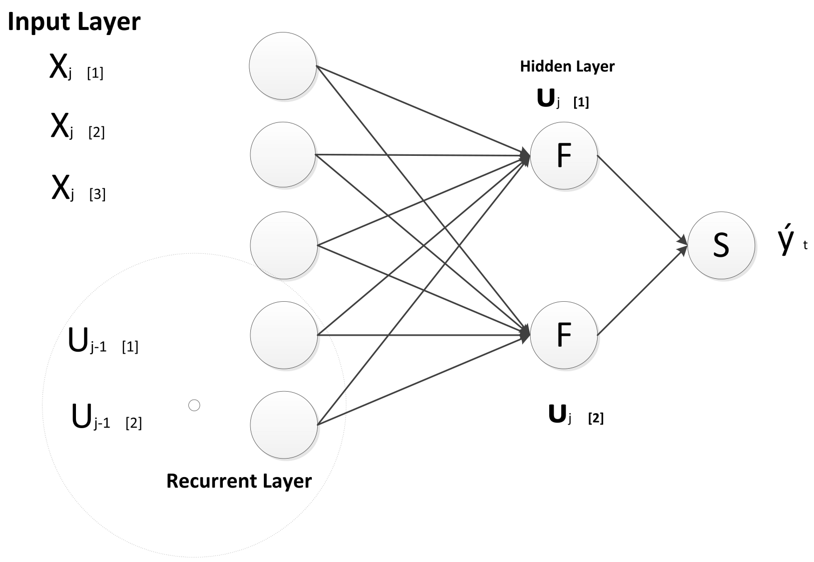

Figure 1: Architecture of Recurrent Neural Network.

illustrates the way in which the error is minimised.

$$E\left(d_t, w_t\right) = \frac{1}{T}\sum_{t=1}^{T}\left(y_t - \hat{y}_t\left(d_t, w_t\right)\right)^2 \tag{16}$$

3.4. 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.[71], 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.[72] 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 [6], 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 [73], [74], [75], [76], [77] and [78]. 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 [79]. The wavelet theory is based on Fourier analysis, where any function is represented as the sum of the sine and cosine functions. Equation 17 illustrates the wavelet admissibility condition, where the wavelet $\psi(t)$ is simply a function of time t that obeys a basic rule [80].

$$C_{\psi} = \int_{0}^{\infty} \frac{\psi\left(f\right)}{f} df < \infty \tag{17}$$

In Equation 17 the Fourier transform and frequency f function of $\psi\left(t\right)$ 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 18 and the mother wavelet also is represented in Equation 19, in both equations $j = 1, ..J$ where J is the wavelet decomposition level [81].

$$\phi_{j,k} = 2^{-j/2}\phi\left(t - 2^{j}k/2^{j}\right) \tag{18}$$

$$\psi_{j,k} = 2^{-j/2}\psi\left(t - 2^{j}k/2^{j}\right) \tag{19}$$

In Equations 18 and 19 which represent the two wavelet types (father and mother), the maximum scale sustainable by the number of data is denoted by J and satisfies:

$$\int \phi\left(t\right) dt = 1 and \int \psi\left(t\right) dt = 0 \tag{20}$$

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 20 illustrates. The projections which can give expansion coefficients are illustrated in Equation 21:

$$\begin{gathered} s_{J,k} = \int \phi_{J,k} f\left(t\right) dt \\ d_{J,k} = \int \psi_{J,k}\left(t\right) dt \left(j = 1, 2, ..J\right) \end{gathered} \tag{21}$$

Equation 22 defines the orthogonal wavelet series approximation to $f(t)$:

$$\begin{gathered} f\left(t\right) = \sum_k s_{J,k}\left(t\right) + \sum_k d_{J,k}\psi_{J,k}\left(t\right) + \sum_k d_{J-1,k}\psi_{J-1,k}\left(t\right) \\ +\ldots\ldots + \sum_k d_{1,k}\psi_{1,k}\left(t\right) \end{gathered} \tag{22}$$

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

$$\begin{gathered} f\left(t\right) = s_J\left(t\right) + D_J\left(t\right) + D_{J-1}\left(t\right) + \ldots + D_1\left(t\right) \\ S_J\left(t\right) = \sum_k s_{J,k}\phi_{J,k}\left(t\right) \\ D_J\left(t\right) = \sum_k s_{J,k}\psi_{J,k}\left(t\right) \end{gathered} \tag{23}$$

Equation 22 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 [82].

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 24 illustrates this case where:

$$n = n/2 + n/4 + \ldots + n/2^{J-1} + n/2^J \tag{24}$$

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 [83]. Thus, WT is adopted for preprocessing the data in this chapter three times.

3.5. Benchmark Prediction Model

In this paper 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. Simple Auto-regressive Model

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

$$\begin{aligned} \Delta y_t = a_1 y^{Sk}(t) + a_2 y^{SD}(t) + a_3 y^{MO}(t) + a_4 y^{RSI}(t) \\ + a_5 y^{MA}(t) + a_6 y^{EMA}(t) + a_7 y^{R}(t) + a_8 y^{MACD}(t) \end{aligned} \tag{25}$$

To predict next day direction closing price of stock index, equation 25 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.

3.5.2. 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 [24], [84]. Timmermann [85] stated in his study that using the simple average may work as well as more sophisticated approaches. In this paper, the simple average is used a benchmark combination model. Equation 26 illustrates the calculation of the combination prediction method at time t [86].

$$f_{SM}^{t} = \left(f_{M1}^{t} + f_{M2}^{t} + f_{M3}^{t}\right) \div 3 \tag{26}$$

4. Experimental results

4.1. Single approached model

Single prediction approach is illustrated in Figure 2. 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 2, while $(n+t)^{th}$ is the output days direction of the closing price. SVM, RNN and Naive Bayes are the employed prediction models.

The below tables illustrate the results of the single approached 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. Moreover, models

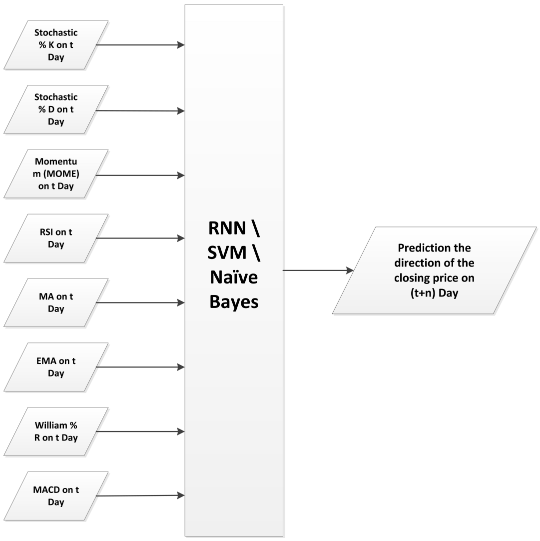

Figure 2: Single stage approach for predicting n day a head of time.

parameters were also determined and obtained by the analytic approach mentioned in subsection ?? for the BNN model, and subsection ?? for the SVM model. The testing results of the utilised models are summarised in Table 4 for the training data sets and Table 5 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 4.

From Table 4, 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 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 4 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

Table 4: 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	——

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 5. It can be observed from Table 5 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.

4.2. Hybrid Modelling and Prediction Procedures

The proposed approach in this paper 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,

Table 5: 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	——

and the weights of these three models are determined by the GA. Moreover, the purpose of the proposed method in this paper is to predict stock index price movement. Figure 3 shows the main procedures of the proposed model approach. The following points illustrate the implementation steps of the new hybrid model:

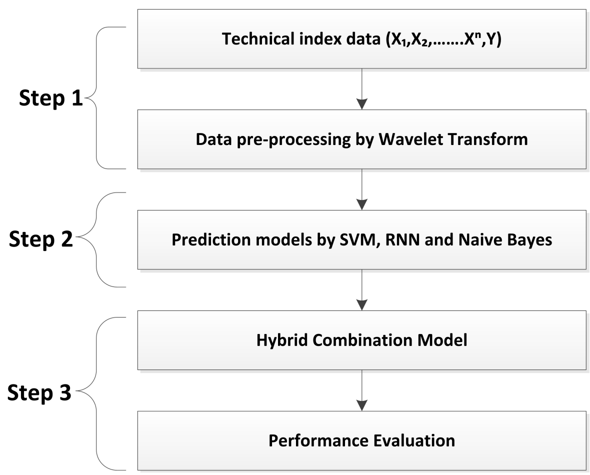

Figure 3: The overall process of WT based SVM, Naive Bayes and RNN hybrid methodology.

- Step one: selecting technical indicators as feature subsets. Subsection 2 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 3.4 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.
- 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 4.3.

4.3. 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 [24], [84]. 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 [85]. 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 [87] [88]:

$$\hat{y}_t = \sum_{i=1}^{m} w_i f_{it}\,(t = 1, 2, ...n) \tag{27}$$

$$\sum_{i=1}^{m} w_i = 1 \tag{28}$$

Equation 27 can be expressed in another from, such as:

$$\begin{aligned} &\hat{y} = FW \\ &\quad where \quad \hat{y} = [\hat{y_1}, \hat{y_2}, ...\hat{y_n}]^T, F = [f_{it}]_{n\times m} \end{aligned} \tag{29}$$

The error for the prediction model can be formed as Equation 30 illustrated.

$$\begin{gathered} e_t = y_t - \hat{y}_t = \sum_{i=1}^{m} w_i y_t - \sum_{i=1}^{m} w_i f_{it} = \\ \sum_{i=1}^{m} w_1 w_i (y_t - f_{it}) = \sum_{i=1}^{m} w_i e_{it} \end{gathered} \tag{30}$$

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

$$\hat{Y}_{combined_t} = \frac{w_1 \hat{Y}_{WT-SVM} + w_2 \hat{Y}_{WT-NaiveBayes} + w_3 \hat{Y}_{WT-RNN}}{(w_1 + w_2 + w_3)} \tag{31}$$

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, WT-SVM, WT-Naive Bayes and WT-RNN models, where the assigned weights are w_1, w_2, w_3 respectively, with $\sum_{i=1}^{3} w_i = 1.0 \leq w_i \leq 1$.

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 31 is the simplest combination method for the three prediction models. Nevertheless, in many cases equal weights cannot achieve the best prediction result. Therefore, this paper adopts a hybrid approach utilising GA as an optimiser to determine the optimal weight for each prediction model. Figure 4 illustrates the architecture of the WT-GA-WA hybrid model.

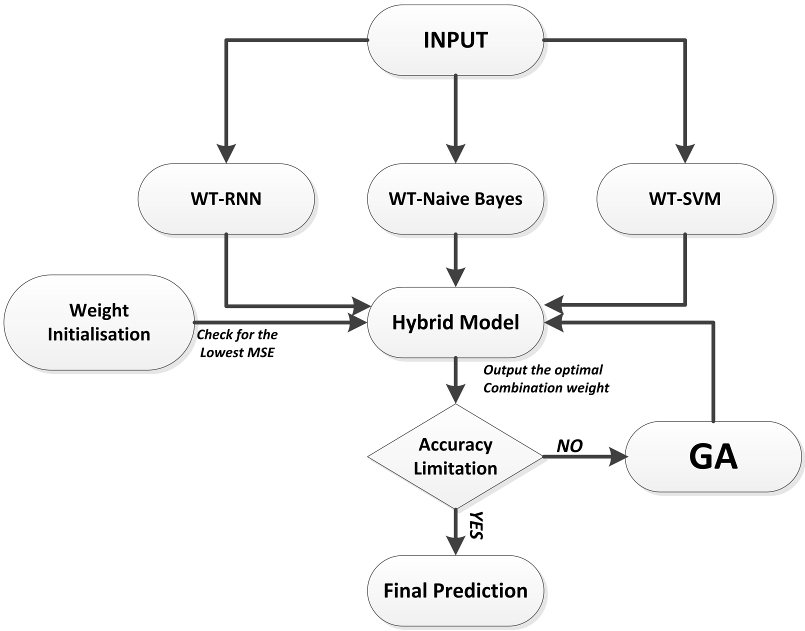

Figure 4: The flow chart of the WT-GA-WA hybrid model.

4.3.1. Genetic Algorithm

The GA is a well-known tool in computational methods modelled on a Darwinian selection mechanism. GA principles were proposed by Holland [89], and developed by Goldbreg [90] and Koza [91]. 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 [92].

- 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.

4.4. Discussions

The new hybrid model in this paper 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 paper 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 Subsection 4.1 with the results of the new hybrid model. As explained earlier in Section 4.2 the proposed approach consists of three steps. Thus, the selection criteria of the input variable is explained in Section 2. 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 4.3.

Table 6 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 6 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 Table 4, 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 6: 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**
FTSE100	WT-SVM	86.42	9.29	6.09	0.80	1.36
	WT-RNN	44.11	6.64	4.59	0.90	1.25
	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
S&P500	WT-SVM	8.16	1.51	2.85	0.78	1.94
	WT-RNN	5.42	2.32	1.22	0.85	1.32
	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
Nikkei225	WT-SVM	1574.11	39.67	28.40	0.87	1.74
	WT-RNN	725.20	26.92	20.55	0.94	1.18
	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 6. From the Table 6, 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 Table 4, the WT-GA-WA results have the lowest error, which implies that the proposed model in this chapter outperformed the single approach. Table 6 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 6. 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 Table 4, 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 7: 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 7 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 7 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 with the obtained results of single approaches in Table 5, 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.

5. Conclusion

This paper 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 paper. 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.

References

[1] Y. S. Abu-Mostafa, A. F. Atiya, Introduction to financial forecasting, Applied Intelligence 6 (3) (1996) 205–213.

[2] A. Cohen, I. Daubechies, P. Vial, Wavelets on the interval and fast wavelet transforms, Applied and Computational Harmonic Analysis 1 (1) (1993) 54–81.

[3] J. B. Ramsey, Z. Zhang, The analysis of foreign exchange data using waveform dictionaries, Journal of Empirical Finance 4 (4) (1997) 341–372.

[4] A. Popoola, K. Ahmad, Testing the suitability of wavelet preprocessing for tsk fuzzy models, in: Fuzzy Systems, 2006 IEEE International Conference on, IEEE, 2006, pp. 1305–1309.

[5] R. Gençay, F. Selçuk, B. Whitcher, Differentiating intraday seasonalities through wavelet multi-scaling, Physica A: Statistical Mechanics and its Applications 289 (3) (2001) 543–556.

[6] J. 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) (1999) 2593–2606.

[7] K. Papagiannaki, N. Taft, Z.-L. Zhang, C. Diot, Long-term forecasting of internet backbone traffic, Neural Networks, IEEE Transactions on 16 (5) (2005) 1110–1124.

[8] Y. Kim, W. N. Street, An intelligent system for customer targeting: a data mining approach, Decision Support Systems 37 (2) (2004) 215–228.

[9] J. C. Francis, E. Kirzner, Investments: analysis and management, McGraw-Hill New York, 1991.

[10] W. Leigh, N. Modani, R. 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) (2004) 515–530.

[11] A. W. Lo, H. Mamaysky, J. Wang, Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation, Tech. rep., National bureau of economic research (2000).

[12] W. Brock, J. Lakonishok, B. LeBaron, Simple technical trading rules and the stochastic properties of stock returns, Journal of finance (1992) 1731–1764.

[13] M. P. Taylor, H. Allen, The use of technical analysis in the foreign exchange market, Journal of international Money and Finance 11 (3) (1992) 304–314.

[14] B. LeBaron, Technical trading rules and regime shifts in foreign exchange, Advanced trading rules (1998) 5–40.

[15] A. W. Lo, J. Hasanhodzic, The evolution of technical analysis: financial prediction from Babylonian tablets to Bloomberg terminals, Vol. 139, John Wiley & Sons, 2011.

[16] D. Keller, Breakthroughs in Technical Analysis: New Thinking From the World's Top Minds, Vol. 61, John Wiley & Sons, 2010.

[17] C.-H. Park, S. H. Irwin, What do we know about the profitability of technical analysis?, Journal of Economic Surveys 21 (4) (2007) 786–826.

[18] P. Ou, H. Wang, Prediction of stock market index movement by ten data mining techniques, Modern Applied Science 3 (12) (2009) p28.

[19] K.-j. Kim, Financial time series forecasting using support vector machines, Neurocomputing 55 (1) (2003) 307–319.

[20] M. Kumar, 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.

[21] M. Kumar, 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.

[22] S.-H. Hsu, J. P.-A. Hsieh, T.-C. Chih, K.-C. Hsu, A two-stage architecture for stock price forecasting by integrating self-organizing map and support vector regression, Expert Systems with Applications 36 (4) (2009) 7947–7951.

[23] S. Makridakis, A. Andersen, R. Carbone, R. Fildes, M. Hibon, R. Lewandowski, J. Newton, E. Parzen, R. Winkler, The accuracy of extrapolation (time series) methods: Results of a forecasting competition, Journal of forecasting 1 (2) (1982) 111–153.

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

[25] S. Asadi, E. Hadavandi, F. Mehmanpazir, M. M. Nakhostin, Hybridization of evolutionary levenberg–marquardt neural networks and data preprocessing for stock market prediction, Knowledge-Based Systems 35 (2012) 245–258.

[26] E. Hadavandi, A. Ghanbari, S. Abbasian-Naghneh, Developing an evolutionary neural network model for stock index forecasting, in: Advanced Intelligent Computing Theories and Applications, Springer, 2010, pp. 407–415.

[27] T. Hill, M. O'Connor, W. Remus, Neural network models for time series forecasts, Management science 42 (7) (1996) 1082–1092.

[28] G. Armano, M. Marchesi, A. Murru, A hybrid genetic-neural architecture for stock indexes forecasting, Information Sciences 170 (1) (2005) 3–33.

[29] W. Shen, Y. Zhang, X. Ma, Stock return forecast with ls-svm and particle swarm optimization, in: Business Intelligence and Financial Engineering, 2009. BIFE'09. International Conference on, IEEE, 2009, pp. 143–147.

[30] A. Kazem, E. Sharifi, F. K. Hussain, M. Saberi, O. K. Hussain, Support vector regression with chaos-based firefly algorithm for stock market price forecasting, Applied Soft Computing 13 (2) (2013) 947–958.

[31] P.-F. Pai, K.-P. Lin, C.-S. Lin, P.-T. Chang, Time series forecasting by a seasonal support vector regression model, Expert Systems with Applications 37 (6) (2010) 4261–4265.

[32] C. Cheng, W. Xu, J. Wang, A comparison of ensemble methods in financial market prediction, in: Computational Sciences and Optimization (CSO), 2012 Fifth International Joint Conference on, IEEE, 2012, pp. 755–759.

[33] K.-j. Kim, I. Han, Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index, Expert systems with Applications 19 (2) (2000) 125–132.

[34] M. T. Leung, H. Daouk, A.-S. Chen, Forecasting stock indices: a comparison of classification and level estimation models, International Journal of Forecasting 16 (2) (2000) 173–190.

[35] A.-S. Chen, M. T. Leung, H. Daouk, Application of neural networks to an emerging financial market: forecasting and trading the taiwan stock index, Computers & Operations Research 30 (6) (2003) 901–923.

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

[37] B. Qian, K. Rasheed, Stock market prediction with multiple classifiers, Applied Intelligence 26 (1) (2007) 25–33.

[38] T. HellstrSom, K. HolmstrSom, Predicting the stock market.

[39] R. D. Edwards, J. Magee, Technical analysis of stock trends (new york institute of finance, business information and publishing, new york), EdwardsTechnical Analysis of Stock Trends1992.

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

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

[42] S. Martinez-Jaramillo, E. P. Tsang, An heterogeneous, endogenous and co-evolutionary gp-based financial market, Evolutionary Computation, IEEE Transactions on 13 (1) (2009) 33–55.

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

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

[45] S.-C. Chi, W.-L. Peng, P.-T. Wu, M.-W. 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, IEEE, 2003, pp. 291–296.

[46] C. N. Tan, A hybrid financial trading system incorporating chaos theory, statistical and artificial intelligence/soft computing methods, in: at the Queensland Finance Conference, 1999.

[47] J.-P. Huang, Prediction: Pure technical analysis might not work satisfactorily, in: Experimental Econophysics, Springer, 2015, pp. 167–180.

[48] M. M. Aldin, H. D. Dehnavi, S. 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) (2012) p25.

[49] A. Abraham, B. Nath, P. K. Mahanti, Hybrid intelligent systems for stock market analysis, in: Computational science-ICCS 2001, Springer, 2001, pp. 337–345.

[50] C.-F. Tsai, Y.-C. Lin, D. C. Yen, Y.-M. Chen, Predicting stock returns by classifier ensembles, Applied Soft Computing 11 (2) (2011) 2452–2459.

[51] J. Sun, H. Li, Financial distress prediction using support vector machines: Ensemble vs. individual, Applied Soft Computing 12 (8) (2012) 2254–2265.

[52] T.-J. Hsieh, H.-F. Hsiao, W.-C. 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) (2011) 2510–2525.

[53] C. D. Kirkpatrick II, J. Dahlquist, Technical analysis: the complete resource for financial market technicians, FT press, 2010.

[54] P.-C. Chang, T. W. Liao, J.-J. Lin, C.-Y. Fan, A dynamic threshold decision system for stock trading signal detection, Applied Soft Computing 11 (5) (2011) 3998–4010.

[55] G. S. Atsalakis, K. P. Valavanis, Surveying stock market forecasting techniques–part ii: Soft computing methods, Expert Systems with Applications 36 (3) (2009) 5932–5941.

[56] R. Bisoi, P. K. Dash, A hybrid evolutionary dynamic neural network for stock market trend analysis and prediction using unscented kalman filter, Applied Soft Computing 19 (2014) 41–56.

[57] A. Gorgulho, R. Neves, N. Horta, Applying a ga kernel on optimizing technical analysis rules for stock picking and portfolio composition, Expert systems with Applications 38 (11) (2011) 14072–14085.

[58] W. Shen, X. Guo, C. Wu, D. Wu, Forecasting stock indices using radial basis function neural networks optimized by artificial fish swarm algorithm, Knowledge-Based Systems 24 (3) (2011) 378–385.

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

[60] C.-L. Huang, C.-Y. Tsai, A hybrid sofm-svr with a filter-based feature selection for stock market forecasting, Expert Systems with Applications 36 (2) (2009) 1529–1539.

[61] J. Yao, C. L. Tan, H.-L. Poh, Neural networks for technical analysis: a study on klci, International journal of theoretical and applied finance 2 (02) (1999) 221–241.

[62] Y. Kara, M. A. Boyacioglu, Ö. K. 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) (2011) 5311–5319.

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

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

[65] C. Cortes, V. Vapnik, Support-vector networks, Machine learning 20 (3) (1995) 273–297.

[66] M. Pontil, A. Verri, Properties of support vector machines, Neural Computation 10 (4) (1998) 955–974.

[67] E. Osuna, R. Freund, F. Girosi, Support vector machines: Training and applications.

[68] I. Steinwart, A. Christmann, Support vector machines, Springer Science & Business Media, 2008.

[69] J. L. Elman, Finding structure in time, Cognitive science 14 (2) (1990) 179–211.

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

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

[72] J. Li, Z. Shi, X. Li, Genetic programming with wavelet-based indicators for financial forecasting, Transactions of the Institute of Measurement and Control 28 (3) (2006) 285–297.

[73] M. A. Ario, et al., Forecasting time series via discrete wavelet transform, Tech. rep., Society for Computational Economics (1998).

[74] A. Aussem, J. Campbell, F. Murtagh, Wavelet-based feature extraction and decomposition strategies for financial forecasting, Journal of Computational Intelligence in Finance 6 (2) (1998) 5–12.

[75] M. Gherman, R. Terebes, M. Borda, Time series analysis using wavelets and gjr-garch models, in: Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European, IEEE, 2012, pp. 2138–2142.

[76] H. Dongdong, Z. Wenhong, The forecast of price index based on wavelet neural network, in: Business Intelligence and Financial Engineering (BIFE), 2011 Fourth International Conference on, IEEE, 2011, pp. 32–36.

[77] J.-Z. Wang, J.-J. Wang, Z.-G. Zhang, S.-P. Guo, Forecasting stock indices with back propagation neural network, Expert Systems with Applications 38 (11) (2011) 14346–14355.

[78] L.-J. Kao, C.-C. Chiu, C.-J. Lu, C.-H. Chang, A hybrid approach by integrating wavelet-based feature extraction with mars and svr for stock index forecasting, Decision Support Systems 54 (3) (2013) 1228–1244.

[79] F. Abramovich, P. Besbeas, T. Sapatinas, Empirical bayes approach to block wavelet function estimation, Computational Statistics & Data Analysis 39 (4) (2002) 435–451.

[80] R. Gençay, F. Selçuk, B. J. Whitcher, An introduction to wavelets and other filtering methods in finance and economics, Academic press, 2001.

[81] J. B. Ramsey, C. Lampart, The decomposition of economic relationships by time scale using wavelets: expenditure and income, Studies in Nonlinear Dynamics & Econometrics 3 (1).

[82] A. Sharkasi, M. Crane, H. J. Ruskin, J. 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) (2006) 511–521.

[83] D. W. Patterson, Artificial neural networks: theory and applications, Prentice Hall PTR, 1998.

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

[85] A. Timmermann, Forecast combinations, Handbook of economic forecasting 1 (2006) 135–196.

[86] C. Xiang, W. Fu, Predicting the stock market using multiple models, in: Control, Automation, Robotics and Vision, 2006. ICARCV'06. 9th International Conference on, IEEE, 2006, pp. 1–6.

[87] S. Gupta, P. C. Wilton, Combination of forecasts: An extension, Management Science 33 (3) (1987) 356–372.

[88] C. Christodoulos, C. Michalakelis, D. Varoutas, Forecasting with limited data: Combining arima and diffusion models, Technological forecasting and social change 77 (4) (2010) 558–565.

[89] J. H. Holland, Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence, MIT press, 1992.

[90] D. E. Golberg, Genetic algorithms in search, optimization, and machine learning, Addion wesley 1989.

[91] J. R. Koza, Genetic programming: on the programming of computers by means of natural selection, Vol. 1, MIT press, 1992.

[92] B. Wu, C.-L. Chang, Using genetic algorithms to parameters (d, r) estimation for threshold autoregressive models, Computational Statistics & Data Analysis 38 (3) (2002) 315–330.