Forecasting stock market by Using Artificial Neural Network and Support Vector Machine

Pakistan Institute of Development Economics

Thesis Submitted

By

SAHEEM AHMED

Supervisor: Dr. Faheem Aslam

Co-Supervisor: Dr. Saud Ahmed Khan

M. Phil Economics and Finance

2019

PAKISTAN INSTITUTE OF DEVELOPMENT ECONOMICS, ISLAMABAD

CERTIFICATE

This is to certify that this thesis entitled **"Forecasting Stock Market by Using Artificial Neural Network and Support Vector Machine."** submitted by **Mr. Saheem Ahmad** is accepted in its present form by the Department of Economics and Finance, Pakistan Institute of Development Economics (PIDE) Islamabad as satisfying the requirements for partial fulfillment of the Degree of Master of Philosophy in Economics and Finance.

Supervisor:

Dr. Faheem Aslam
Assistant Professor,
COMSATS University,
Islamabad

Supervisor:

Dr. Saud Ahmad Khan
Assistant Professor,
PIDE, Islamabad

Internal Examiner:

Dr. Ahmad Fraz
Assistant Professor,
PIDE, Islamabad.

External Examiner:

Dr. Waseem Shahid Malik
Professor,
SBP, Memorial Chair,
University of Peshawar,
Peshawer

Head, Department of Economics and Finance:

Dr. Ahsan Ul Haq
Assistant Professor,
PIDE, Islamabad.

October 25, 2019

Dedication

I am dedicating this thesis to four most adorable people. Their existence in my life is like a heavenly precious gift. First and foremost dedicate it to my Parents because they were always there for me when I need them and thought the value of hard work. They always support me with my decisions related to my future. Thank you so much for always having faith in me and for showing selfless support.

Next I want to dedicate to my brother Fasih Ahmed, my sister Masira and my fiancé Arhama. They motivated me every time when get exhausted. Thank you for always being there for me and helping me with my work.

Last but not the least I also want to mention my teacher and supervisor Dr. Faheem Aslam who consistently helped me through his advice and brilliant ideas which are always fruitful for my work. Thank you sir to make me able to complete my thesis.

Acknowledgement

First and foremost, I want to thanks Allah Rabbul Izzat for letting me see this thesis thought and provided all that I needed to complete my thesis. Allah almighty always took care of everything that would have stopped me in my tracks and strengthened me even through my most difficult time.

My sincere appreciation goes to my supervisor Dr. Fahim Aslam whose contribution and constructive criticism has pushed me to make more effort and to work in precise manner. His valuable suggestions make me able to complete this research work. My appreciation is also for my co-supervisor Dr. Saud Ahmed for his support and suggestions helped me to complete my thesis.

I also want to thanks to my friends Mohsin Rasheed, Nasir Rao, Liaquat Ali, Tanzeel-ur-Rahman and Umar mehmood for your moral support and encouragement lead me to complete my thesis.

Finally, I pay my deep gratitude to my beloved parents; Mr. Abdul Karim and Mrs. Shabana Naz, Brother Fasih Ahmed, Sister Masira and my fiancé Arhama for their love, care and prayers have made me able to complete this journey.

List of Tables

List of Figures

Table of Contents

Abstract

In the business sector, it is always a difficult task to predict the daily prices of stock market due to its nonlinear behavior in stock market prices. Hence, there are numerous study has been conducted in respect of the prediction of the direction of stock market movement. Several studies provide solid corroboration that models using traditional regression techniques face irresistible problems due to model ambivalence. So here there is a need of discriminating model used to minimize the high risk and maximize returns. Computing techniques are rational method for forecasting stock prices. This study presents a computational approach for predicting the stock market indexes of six different countries. It is a review of artificial neural network (ANN) and support vector machine (SVM) pertain and achieve to predict the stock market prices. Daily stock price data from 2000 to 2018 has been used for prediction. Six stock market indexes of Asian region have been selected to predict the daily stock prices by using 30 variables. The proposed methods consists of three steps in first step it uses technical analysis based on historical data to calculate important indicators. In second step it identified the most important indicators which actually represents the data set and then finally use ANN and SVM to predict the stock market prices by using important variables. For better performance, the study have performed comparative analysis to find out better model in each country. The results concluded that SVM and ANN has ability to predict the stock market and the main finding of the study is that by using the most important feature extraction the prediction accuracy has been increased and as a result investors can earn huge profit by using these techniques.

CHAPTER 1

INTRODUCTION:

From the last few years, forecasting the stock prices is becoming an attractive topic for researchers in time series domain. Stock market research field was developed non-linear, non-parametric and complicated in nature. Many researchers and investors are trying to improve their forecasting technique because little change in improvement will lead towards the huge returns.

Fama (1970) proposed the **efficient market hypothesis** states that asset prices fully reflect all available information. A direct implication is that it is impossible to "beat the market" consistently and market prices will only react to new information. Fama also argued that stocks always trade on their fair value, it make impossible for investors to either purchase undervalued stocks or sell stocks for high prices. As such it should be impossible to outperform the overall market through market timing or expert stock selection. The only way an investor can possibly obtain abnormal returns is by chance or by purchasing riskier investments.

There are many researchers who opposed EMH theory and suggested that there are many ways for forecasting. Sridharan (2013) and Wafi and Hassan (2015) disagree with it and says that there are methods and technologies available which make them to predict the future price of the stocks. They are Technical analysis and Fundamental Analysis, which are traditional methods of forecasting stock market price. But these methods are not efficient as Wafi and Hassan (2015) documented that there is lack of information in these methods so it is a big question mark on the efficiency on these models. There are some other methods for forecasting stock market are ARMA model and GARCH model but the problem with these model is of restriction of stationary data.

Many researchers now focus on machine learning methods which are better than fundamental and technical analysis methods. These methods are also effective in non-stationary dataset and helpful in volatility situations. There are number of machine learning methods but two methods have been discussed artificial neural network and support vector machine.

Artificial Neural Networks (ANN), are becoming very popular in financial applications. ANN is nonparametric and nonlinear model. ANN allow you to completely utilize the data and make the data to determine the parameters and structure of a model without any restriction in the indicators of model. Neuron network is an approach that helps in regression. It has three layers: hidden layer, input layer and output layer.

Another method used in this study is support vector machine. Vapnik and Chervonenkis (1963) invented support vector machine (SVM). Support Vector Machine is one of the good binary classifiers. It takes the decision of classifying the data in different dimensions most points in one category fall on one side while most points in the other category fall on the other side of the boundary. Then elements in one category have sum should be greater than 0, while elements in the other category will have the sum should be less than 0. In our classification, classify it like $y \in \{-1, 1\}$.

1.1 Plan of the Study:

According to the plan of the study first I have identified the problem statement and objective of the study. Afterword find out the research gap by focused on literature review that how past studies have worked by using ANN and SVM. After that the theoretical framework in which I have discussed all the variables. Later methodology and results have been analyzed and concludes with comments on possible future work in the same area and some conclusions.

1.2 Problem Statement:

Forecasting the stock market is an important problem in finance. There are a many models for forecasting but the efficiency of those models is a question mark. Investors want to earn abnormal profit, it is only possible when you forecast the stock price accurately close to the actual one because inaccuracy lead towards the huge risk. Suhas (2013) and Wafi and Hassan (2015) proposed that there are some efficient methods and technologies available which make them to predict the stock prices. Furthermore, Kara, Boyacioglu and Baykan (2011) stated that one can forecast the stock prices accurately by using machine learning methods. Henrique, Sobreiro and Kimura (2018) documented that SVM has the prediction power, especially when you update the model periodically and also stated that accuracy of prediction increased in low volatility.

1.2.1 Questions:

Following questions are answering in this research:

- What is the prediction of stock market indexes by using ANN and SVM?
- What are the reflective indicators identified by the models in each index?
- Comparison of market prediction accuracy between neural network and support vector machine.

1.2.2 Objectives:

This study pursues the following specific objectives:

- The main objective is to predict the stock market indexes.
- To identify the most important indicators in each country.
- To find out the better model for each stock market index by comparing ANN and SVM.

1.3 Research Gap:

Forecasting through machine learning is becoming familiar to the researchers, but there is only limited work available as compare to statistical models. Following research gap I found from the literature:

In previous studies, many researchers using ARCH and GARCH model for prediction purpose like Grachev (2017) forecasted stock market direction using ARCH and GARCH. Similarly, Vasudevan and Vetrivel (2016) forecast stock market volatility by using GARCH model. Li, Yang and Li (2017) predicted the shanghai stock market direction using ARIMA model.

In this present era researchers are focusing on machine learning methods but most of the studies used only single technique or with single model like (Kara et al., 2011) forecasted the direction of Istanbul stock exchange by using SVM and ANN. Similarly, Qu and Zhang (2016) predicted the Chinese stock index by using SVM. In this thesis, 6 Asian countries index have taken for analyzing the direction of stock market by using ANN and SVM

Secondly, up till now comparison of these technique is very rare and the comparison in selected stock market indexes is not available. I have compared these two models to check the significance level in each stock market index, because results may vary in each stock market index due to difference in economies.

1.4 Significance of the Study:

This study is very important for the investors, the person who wants to invest in stock market has a purpose of high return and want to become wealthy. Any model which helps stock market investors to pick the right stock at the right time, it will make investors very wealthy. That's why many investors, professionals and researchers are looking for a system or model for prediction of stock prices, which leads them towards the huge returns. . For those people there is a need of a new model that help them to forecast the stock market easily and accurately. Usage of these methods will be beneficial for the investors because it produces very good results with high accuracy and little improvement in accuracy will provide abnormal profit to stock market investors.

CHAPTER 2

Literature Review:

Researchers uses artificial neural network in their policy reports and suggested that it is one of the best method for forecasting the stock market prices. Devadoss & Ligori (2013) documented that for forecasting stock prices of Bombay stock market, they used Multilayer networks. They forecasted stock prices and also compared two architecture methods NN1 (3-6-1) and NN2 (3-16-1). Neural network forecasting of stock market is showing that the neural networks are able to forecast the stock prices, even if there is volatility in the market. These methods of neural networks are tested on the data of Bombay stock exchange and the results are predicted. For input data they used the historical prices and then train it with some variables. The Indian stock market which is highly volatile, if the forecasting is possible in such market with high accuracy, so it is very useful for the investors to get benefit from these methods of forecasting. It is highly recommended for the investors to forecast the stock prices by neural networks and earn huge profit. He also suggested that not only the historical prices used for forecasting, but for improvement one can use international stock market dataset and can use macro-economic variables as well. Fundamental and technical analysis indicators can also be used for the stock market prediction purpose and can check the improvement of neural networks.

(Qiu et al., 2016), their study uses the daily stock price data of Japanese stock market and predicting the daily stock market movements. The author is using artificial neural network for forecasting the Japanese stock market index. For improving the efficiency of predicting of stock market, the author uses Artificial Neural Network model with genetic algorithm (GA) and also used hybrid GA-ANN model and then compare the accuracy of both models. Then he assign weights to variables and used both the model GA and Hybrid model and compare the results with the actual data. The study suggested that Type 2 model which is hybrid GA-ANN model provides better result than type 1 model GA model. The type 2 model has the accuracy of prediction of stock market 81.27 % that is good result. The author also compare this model with other relevant models and documented that there method is more efficient and providing more effective results as compare to any other model. The author also suggested that the accuracy can be improved more by use of other variables that may affect the forecasting performance. He also suggested that the use of other machine learning methods may be provide better results than this method.

Addai (2016) predicted the stock market index by using machine learning techniques. The researcher is using the daily price data of the standard and poor 500 index (S&P 500). There are five techniques which are used by author. The techniques are artificial neural network, Quadratic discriminant Analysis (QDA), and Logit model, K-Nearest Neighbour-hood classification (KNN) and Linear discriminant analysis (LDA). From the empirical data the results are concluded. It was documented by the author that artificial neural network performs better than any other techniques which are used by him. He further said that any technique could not perform 100% but artificial neural network performed well, it performed almost 61%. The author also documented that the investor can use ANN model for his portfolio with some level of confidence interval. He also suggested that for predicting the price, one should use high and low prices of a day and then predict the next day prices. He also suggest to use economic variables for better results.

Maciel and Ballini (2010) documented by the author that the main purpose is to analyze the financial time series forecasting by using neural networks and their ability to forecast the trend of different index such as European, North American and Brazilian stock market index. The authors are also comparing the neural networks with traditional forecasting method which is generalized autoregressive conditional heteroskedasticity (GARCH) model. The author proposed that artificial neural network has the capability to predict the direction of stock market index, also proposed that if it is properly trained with the related variables than robustness can be improved more. The researchers also documented that, in terms of statistics artificial neural network performed well as compare to GARCH model. ANN performed well even the data contains noisy information and high volatility. Moghaddam, Moghaddam and Esfandyari (2016) used ANN for forecasting the daily stock price of the index NASDAQ. They took the data from Jan 2015 to June 2015 only 99 days data. First 70 days they used for training the dataset and while other remaining 29 days used for the testing purpose and a result of ANN the result forecasted is validated, is almost same as the actual result.

Lawrence (1997) documented that this study is a survey on the application of neural networks in forecasting stock market prices. With their ability to discover patterns in nonlinear and chaotic systems, neural networks offer the ability to predict market directions more accurately than current techniques. Common market analysis techniques such as technical analysis, fundamental analysis, and regression are discussed and compared with neural network performance. Also, the Efficient Market Hypothesis (EMH) is presented and contrasted with chaos theory and neural networks.

This studied surveyed the application of neural networks to financial systems. It demonstrated how neural networks have been used to test the Efficient Market Hypothesis and how they outperform statistical and regression techniques in forecasting share prices. Although neural networks are not perfect in their prediction, they outperform all other methods and provide hope that one day we would more fully understand dynamic, chaotic systems such as the stock market.

Selvamuthu, Kumar and Mishra (2019) in this article they used neural networks based on three different learning algorithms, i.e. Scaled Conjugate Gradient, Levenberg Marquardt and Bayesian Regularization for stock market prediction based on trick data as 15 min data of an Indian company and compared their results. All three algorithm provide an accuracy of 99%

using tick data. the accuracy over 15 min dataset drops to 97.0%, 96.2% and 98.9% for SCG, LM and Bayesian Regularization respectively which is significantly poor in comparison with that of results obtained using tick dataset.

Kim and Kim (2019) propose a model, called the feature fusion long short-term memory-convolutional neural network (LSTM-CNN) model, that combines features learned from different representations of the same data, namely, stock time series and stock chart images, to predict stock prices. The proposed model is composed of LSTM and a CNN, which are utilized for extracting temporal features and image features. They measure the performance of the proposed model relative to those of single models (CNN and LSTM) using SPDR S&P 500 ETF data. Our feature fusion LSTM-CNN model outperforms the single models in predicting stock prices. In addition, discovered that a candlestick chart is the most appropriate stock chart image to use to forecast stock prices. Thus, this study shows that prediction error can be efficiently reduced by using a combination of temporal and image features from the same data rather than using these features separately.

Song, Zhou and Han (2018) surveyed and compared the predictive power of five neural network models, namely, back propagation (BP) neural network, radial basis function (RBF) neural network, general regression neural network (GRNN), support vector machine regression (SVMR), least squares support vector machine regression (LS-SVMR). They apply the five models to make price prediction of three individual stocks, namely, Bank of China, Vanke A and Kweichou Moutai. Adopting mean square error and average absolute percentage error as criteria. They find BP neural network consistently and robustly outperforms the other four models.

Some of the researchers uses Support vector machine and documented that it is a good method for forecasting and also stated that it produces very significant result. Hakob (2017) used statistical analysis tools and prediction model machine learning method is presented to predict the stock market prices. He made three portion, in first portion he used useful indicators from the historical data by using the technical analysis. In second portion he selected most important variable by using the method of two variable model. And in the end he forecasted the stock price data by using the support vector machine (SVM). He compared the two models performance SVM with the proposed method to check the effectiveness of the selected variables. The study conclude that the SVM perform well if you use the related variable which represent the actual data. Hakob (2015) documented that he used a combined model for forecasting purpose based on Principle component analysis (PCA) and Artificial neural network (ANN) for predicting the financial time series data. For selecting the important variables he used technical analysis. The PCA approach and ANN model used data for training purpose and forecast the data of NASDAQ Baltic stock market. To evaluate the performance Mean square error was used and this lead to conclude that forecasting data the alternative proposed method can be used.

(Henrique et al., 2018) documented that the efficient market hypothesis suggest that one cannot forecast the market movement, so there are some machine leaning algorithms which helps in forecasting the stock price movements and make investors able to earn abnormal profit of their investments. One method the author used is Support vector machine (SVM) and author compare it some traditional method Random walk, which is derived by efficient market hypothesis. The author used the daily data of stock market for forecasting the day to day prices. The results proposed that Support vector machine has ability to predict the stock prices and providing tremendous results. They also suggested that the SVM gives better result in a case of less volatility in the market. One important contribution is the comparison between the

Support vector machine and random walk model. According to Random walk model markets movement are unpredictable in long run. In this context the SVM provides better result than random walk methods, it gives better results with the use of linear Kernel. It also shows that SVM also work in volatility, even in high volatility it forecast the stock prices. But in low volatility it works very well and gives results with high accuracy. The study also suggests that the use of more indicators can be useful for improving the results. The one can also use technical and as well as fundamental indicators as variables and can improve the efficiency of machine learning methods.

Qu and Zhang (2016) used an assumption that every high frequency causes momentum. With this assumption the author used the daily stock market data of Chinese stock exchange CSI 300 index. The author used Support vector regression with optimal parameters new Kernel and then compare the results with sigmoid kernel. The author concluded that the new kernel method with SVR perform well as compare to other method, and also said that the investor can earn huge capital gain with this simple strategy.

Patel, Shah, Thakkar and Kotecha (2015) predicted the stock market by using different machine learning methods. The author used the daily stock market data of Bombay stock exchange market and forecasted the 1, 10, 15 and 30 days stock market prices. The author used 2 machine learning methods support vector machine and artificial neural network in the first stage and in 2nd stage used fusion approach by using Support Vector Regression (SVR), Artificial Neural Network (ANN) and random forest (RF) forecasting models. The author uses 10 indicators as input variables forecast the stock prices and then compare the results of the model. The empirical results suggested that the machine learning methods have ability to forecast the stock market. Machine learning methods have high accuracy for forecasting even in high volatility situation. The author further suggest that the efficiency can be more improve by using the authentic variables. He also said that more indicators like technical and fundamental indictors may be more beneficial for forecasting the stock market. The machine learning method for prediction is very useful for investors to reap more gain.

Nayak and Mishra (2015) proposed that for prediction of Indian stock market author used support vector machine. The main objective is to find out the predicted price, momentum and volatility of Bombay stock exchange and CNX. The author used technical analysis tools for prediction of stock market. The writer used support vector machine with kernel function for forecasting the stock prices, up and down movement and momentum of Bombay stock exchange index. The author forecast the 1 day, 1 week and 1 month prices of stock market by using SVM and KNN forecasting methods. The author suggested that the SVM is very useful for forecasting the prices, movement and momentum of stock market. The empirical results proposed that SVM has ability to forecast the stock market index with high accuracy and is very important for investors to use this methods to earn abnormal profit.

Huang, Nakamori and Wang (2005) investigated the predictability of 1nancial movement direction with SVM by forecasting the weekly movement direction of NIKKEI 225 index. To evaluate the forecasting ability of SVM, compare its performance with those of Linear Discriminant Analysis, Quadratic Discriminant Analysis and Elman Back propagation Neural Networks. The experiment results show that SVM outperforms the other classi1cation methods. Further, proposed a combining model by integrating SVM with the other classi1cation methods. The combining model performs best among all the forecasting methods

Cao, Liang and Ni (2012) further explored the relationship between the Internet financial information and financial markets, including the relations between the Internet financial information content, Internet financial information sentimental value and stock price. They collect the news of three stocks in the Chinese stock market in the GEM in a few large portals, use the text sentiment analysis algorithm to calculate the sentimental value of the corresponding

Internet financial information, combined with the stock price data, implant Support Vector Machines to analyzes and forecasts on the stock price, the accuracy of the prediction of stock prices has been improved.

Rustam and Kintandani (2019) forecasted the stock market of Jakarta composite index (JKSE). Several studies have focused on the prediction of stock prices using machine learning, while one uses support vector regression (SVR). Therefore, this study examines the application of SVR and particle swarm optimization (PSO) in predicting stock prices using stock historical data and several technical indicators, which are selected using PSO. Subsequently, a support vector machine (SVM) was applied to predict stock prices with the technical indicator selected by PSO as the predictor. The study found that stock price prediction using SVR and PSO shows good performances for all data, and many features and training data used by the study have relatively low error probabilities. Thereby, an accurate model was obtained to predict stock prices in Indonesia.

Madge and Bhatt (2015) stated that Support Vector Machine is a machine learning technique used in recent studies to forecast stock prices. This study uses daily closing prices for 34 technology stocks to calculate price volatility and momentum for individual stocks and for the overall sector. These are used as parameters to the SVM model. The model attempts to predict whether a stock price sometime in the future will be higher or lower than it is on a given day. Find little predictive ability in the short-run but definite predictive ability in the long-run.

Nayak, Naik and Behera (2015) surveyed the role of SVM in various data mining tasks like classification, clustering, prediction, forecasting and others applications. In broader point of view, we have reviewed the number of research publications that have been contributed in various internationally reputed journals for the data mining applications and also suggested a possible no. of issues of SVM. The main aim of this paper is to extrapolate the various areas of SVM with a basis of understanding the technique and a comprehensive survey, while offering researchers a modernized picture of the depth and breadth in both the theory and applications. Although, this paper is by no means a meticulous review of the literature in the application of support vector machine to application of data mining. Also the fundamentals of the support vector machines (SVMs) have been discussed along with the different formulations of the optimization problem resulting from the training of such machines.

Meesad and Rasel (2013) combined different kinds of windowing function with a support vector machine. This is a new way to apply different kinds of windowing function as data preprocess step to feed the input into the machine learning algorithm for pattern recognition. From the result analysis, one can say that, SVR models that are built by using rectangular window and flatten window operator are good to predict stock price for 1 day a-head, 5 days a-head and 22 days a-head. Because Error rate (MAPE) between actual and predicted price values in those models is quite acceptable because of low margin difference. In this study, support vector regression (SVR) analysis is used as a machine learning technique in order to predict the stock market price as well as to predict stock market trend. Moreover, different types of windowing operators are used as data preprocess or input selection technique for SVR models. This is a new approach which uses different types of windowing functions as data preprocess for predicting time series data. Support vector regression is a useful and powerful machine learning technique to recognize pattern of time series dataset. It can produce good prediction result if the value of important parameters can be determined properly. Different kinds of Windowing operators are used in this experiment in order to feed more reliable inputs into regression models. This study is done on a well-known company of Dhaka stock exchange (DSE), named ACI group of company Limited. Four year's historical time series dataset are collected from the DSE from 2009 to 2012, as daily basis for experimentations. Finally, predicted results from Win SVR models are compared with actual price values of DSE to evaluate the model prediction performance.

From the result analysis, one can say that, SVR models that are built by using rectangular window and flatten window operator are good to predict stock price for 1 day a-head, 5 days a-head and 22 days a-head. Because Error rate (MAPE) between actual and predicted price values in those models is quite acceptable because of low margin difference.

To increase the predictive performance, some of the authors have used hybrid machine learning algorithms. (Wang et al., 2014) find out the direction of stock prices using SVM. He suggested that the result is near to the actual one and it is more effective model. The indicator is using for input variables in TA indicator. (Nayak et al., 2015) used a hybrid KNN and SVM system for forecasting stock values. It showed better result as compare to that traditional artificial neural networks. Furthermore, (Patel et al, 2015) used RF, SVM and hybrid neural network to predict the daily stock prices. Comparison between the results with the same algorithm, concluded by the author algorithm with hybrid uses have better result than same isolated algorithms. Dash and Dash (2016) used neural network with input variable TA and compared it to the same machine learning algorithm in stock investment decision and find out that neural network perform better as compare to any other machine learning models.

Araújo, Oliveira and Meira (2015) stated that there is a lot of space for the development of this model. They proposed a mathematical model for forecasting the stock prices of Brazilian stock exchange, which is measured in seconds. Manahov, Hudson and Gebka (2014) used the genetic learning algorithm to investigate the profitability of the FOREX market, which is measured in minutes and applied changes in price. Qu and Zhang (2016) used SVR method in Chinese market used minute prices.

There are many other studies applied forecasting techniques to the data of higher frequencies data and suggested that the SVR is providing much better results as compare to any other model. The results are almost accurate as the actual one. Gallo and Brownlees (2006) proposed HFT model on immediate change in the price variations that is on a ticks price vary according to sale or purchase transaction performed. They said that for investors it is a good thing by which he can forecast the stock price and can earn huge profit because machine learning methods are very useful for forecasting and almost provide accurate results.

Jabbarzadeh, Shavvalpour, Khanjarpanah and Dourvash (2016) forecasted the stock market direction by using nonlinear probability models and different technical indicators. The author used methods like probit, logit and extreme value model with proposed indicators. The proposed indicators are price channel index, relative strength index, momentum, moving average, random walk and moving average convergence divergence. The researcher is predicting the direction of standard and poor 500 index. The empirical result of the study is showing that extreme value model is performing well as compared to other models. In addition the models are showing good performance in term of error criteria. The forecasting power of extreme value model is 72.72% and in reality criteria it is 60.73%. In term of error measurement criterion, extreme value model is the best model among all the models which are used by the author.

There are some literature available about the comparison between artificial neural network and support vector machine. (Kara et al., 2011) forecasted the stock prices by using two different methods of machine learning techniques and then compared their performance. The researchers have used daily stock market data of Istanbul stock exchange (ISE). The two techniques which have used are artificial neural network (ANN) and Support Vector Machine (SVM). There are ten technical indicators has been used by these models to forecast the prices of stock market. From the results, it is showing that both the model performed well but according to comparison artificial neural network performed better than Support vector machine in Istanbul stock market index. The Artificial neural network performance is 75.74%, while performance of support vector machine is 71.52%. The author further suggested that the accuracy can be improved by using other variables like technical and fundamental indicators.

Paik and Kumari (2017) documented that they used three models Support vector machine (SVM), Artificial Neural Network (ANN) and Extreme learning machine (ELM) for forecasting the stock prices. ANN is nonlinear model to forecast the stock price better because it has no restrictions. Prediction is better by SVM because it uses marginal values rather than the average values. For forecasting the stock prices with the help of ELM uses fast training mechanism. These three technique are very useful for studying and as well as for prediction the direction of stock market prices. In this statement of belief authors concluded that, these three models perform well and has the ability to predict the direction of stock market. They also wrote that, if you talk about the comparison than SVM performed better as compare to artificial neural network.

Khan (2009) strives on the relative importance of dividends, retained earnings, and other determinants in the explanation of stock prices in Bangladesh with particular stock price of the companies associated with Dhaka Stock Exchange (henceforth DSE), an emerging capital market of Bangladesh. The prime objective of this study is to study determinants of market share price and to examine their functional relationships with the market price of common stocks trades in DSE. Used different models to explain the dynamic relationships of market price of common stocks with the determinants of market share price like dividends, retained earnings, lagged price earnings ratio and market price of previous year. The results of the empirical analysis evidences that dividends, retained earnings and other determinants have dynamic relationship with market share price. Findings also suggest that the overall impact of dividend on stock prices is comparatively better that that of retained earnings and expected dividends play an important role in the determination of stock prices whatever determinants, like lagged price earnings ratio or lagged price, are considered.

Széliga, Verdes, Granitto and Ceccatto (2003) refined and complemented a previously-proposed artificial neural network method for learning hidden signals forcing nonstationary behavior in time series. The method adds an extra input unit to the network and feeds it with the proposed profile for the unknown perturbing signal. The correct time evolution of this new input parameter is learned simultaneously with the intrinsic stationary dynamics underlying the series, which is accomplished by minimizing a suitably-defined error function for the training process. They incorporate here the use of validation data, held out from the training set, to accurately determine the optimal value of a hyper parameter required by the method. Furthermore, evaluate this algorithm in a controlled situation and show that it outperforms other existing methods in the literature. Finally, discuss a preliminary application to the real-world sunspot time series and link the obtained hidden perturbing signal to the secular evolution of the solar magnetic field.

The literature review concluded that there are lot of studies which states that one can predict the stock market by using different machine learning methods. Some studies are there which conclude that machine learning methods are better than that of traditional methods like ARCH, GARCH models. But there are only few work available which focus on more than one country. In this study our main focus is to predict the stock market of ASIAN region by using machine learning methods which is very rare in past studies.

CHAPTER 3

3.1 Theoretical Framework:

Fama (1970) proposed efficient market hypothesis (EMH) it suggests that market are efficient, no system can beat the market. There is an ongoing debate on the validity of efficient market hypothesis. But some researchers like Devadoss and Ligori (2013) and Qiu, Song and Akagi (2016) suggested that one can forecast the stock prices by machine learning methods, include Artificial Neural Network, Support Vector Machine and many others. By selection of most important variables in machine learning methods, one can easily predict the prices of the stock and as a result investors of stock market can earn abnormal profit.

There are many variables which have been used in previous studies. (Patel et al., 2015) and (Kara et al., 2011) documented that the use of more important variables lead towards high accuracy of the models. For SVM number of variables has been used. But (Henrique et al., 2018) suggested the four variables: simple moving, average (SMA), weighted moving average (WMA), the accumulation/distribution oscillator (ADO), average true range (ATR). Addai (2016) suggested that the use of economic variables may improve the accuracy of the models. Khan (2009) suggested the following variables: earnings per share, market price of previous year, divided per share, net asset, ratio of stock price per share, volume of stock transaction and error coefficient. In this study, the forecasted or the predicted stock price is the dependent variable with 15 independent variables which have been identified by the models. From these variables SVM identified 10 important indicators while ANN identified 5 important indicators, which are defined as follows:

- **Exponential Moving Average (EMA):**

It is a type of moving average (MA) that places a greater weight and significance on the most recent data points. It is also known as the exponentially weighted moving average.

- **Rate of Change (ROC):**

The Rate of Change (ROC) indicator calculates the percentage change in between the current price and the n number of previous prices. It is also referred as momentum. ROC fluctuate from positive to negative and it shows the overbought and oversold condition.

- **Commodity Channel Index (CCI):**

CCI is used to detect the cycling trend, the name commodity is using here because it was first used to describe the trend for commodity. It is oscillate above and below from the mean line. It is like as to momentum because there is no limit of upward or downward values. The prices it takes is of 14 days period.

- **Average Directional Movement Index (ADX):**

ADX is used to detect the trend strength by this it will help in reducing the risk and an increase in profit. But it will never determine that whether the trend is bullish or bearish. The time period usually used is 14 periods but you can enhance it.

- **Average True Range (ATR):**

ATR usually tells the volatility of the stock. It describe somehow that what will be the expected price of a security on a given day, it is referred as moving average which usually used the data of 14 days to describe the volatility.

- **Chaikin Money Flow (CMF):**

CMF is a technical indicator used to describe the money flow volume over a given period of time. It is a metric which defines the selling and buying condition of a security on a single period. When CMF is close to 1 than it is buying pressure and when it is close to zero then it is selling condition.

- **Relative Strength Index (RSI):**

The Relative Strength Index (RSI), is an indicator for momentum that measures the speed and change in prices. The RSI value occurs between zero and 100. Here consider RSI as overbought condition when it crosses 70 and consider as oversold condition when it crosses 30 towards zero. It is also used to find out the current trend of the stock.

- **Close Location Value (CLV):**

CLV used in technical analysis to determine that the value of a security is close to day high and day low price. It is ranges between 1 to -1, if it is close to 1 then it is close to high or ot is close to -1 then it is equal to day low.

- **AroonUp and AroonDn:**

There are two indicators are found in Aroon that AroonUp and AroonDn. It basically used to describe the trend changes in the prices of a security. It is also determine trend strength. Strong uptrends will regularly see new upward trends while strong downward trends will see downward trend. The AroonUp determines the upward trend and AroonDn determine the downward trend.

- **Percentage Bandwidth (%BB):**

Percentage BB or bandwidth is an oscillator which is derived from standard Bollinger BB indicator. It is used to check the volatility of any security. This indicator is consists of three lines. The middle line shows the moving average of 20 days, while 1^{st} and 3^{rd} line shows the 2 standard deviation which is above and below the middle line. It is useful to determine the trading signals and it also defines the trend.

- **True High and True Low:**

True High is determined as the higher of yesterday's high and the close 2 days ago and true low is described as the lower of yesterday's low and the close 2 days ago.

- **Money Flow Index (MFI):**

The MFI is an indicator used for technical analysis for defining buying and selling pressure. By analyzing both price and volume it will define the buying and selling points. It makes a oscillator that moves between a range 0 and 100, when the MFI increases it is overbought condition and when the MFI goes down then it is oversold condition.

- **True Range (TR):**

There are three parts which define the true range.

The distance between the close price of yesterday to low price of today.

The distance between the close price of yesterday to today's high price.

The distance between today high to today low.

14 period prices are used in raw true range to give average true range. The true range can different types of moving averages like simple, weighted, exponential moving average and etc.

- **William %R (WPR):**

Williams's %R, it is also known as Williams Percent Range. It is an indicator that remain between zero and -100. It is used to measure the overbought and oversold levels. It also help to find out the entry and exit points in the market. In practice, it takes the stock's price to the high-low range over a period of 14 days.

- **Chande Momentum Oscillator (CMO):**

The CMO is one of the technical indicator used for analysis. It is the difference of the sum of higher close prices and the sum of all lower close prices and then divide it to the sum of all the prices have taken at that time period. The range is between -100 to +100 and the time period usually used is 20 days. It defines the overbought and oversold condition when CMO is +50 then it is overbought condition while when CMO reaches at -50 then it is oversold condition. CMO is almost similar to RSI.

- **Oscillator / Summation:**

The McClellan Oscillator and Summation Indexes were developed by Sherman and Marian McClellan in 1969. They were developed to gain an advantage in selecting entry and exit times in the stock market. There are two important interpretations by this oscillator. One is when oscillator is positive it means that it shows money is coming in the market and when it became negative it means money is going out from the market. Secondly it is very helpful to tell the overbought and oversold condition in the market.

Table No: 3.1

Technical Indicators	**Calculations**
Exponential Moving Average	(CLOSE (i) * P) + EMA (i-1) * (1-P) close (i): closing price for the current period P: percentage of price use.
Rate of Change	(Price- price n periods ago)/ (Price n periods ago)
Commodity Channel Index	(Close t + High t + Low t) / 3
Average Directional Movement Index	MA [(DI) – (-DI)] / [(DI) + (-DI)] DI: plus directional indicator -DI: Minus directional indicator
Average True Range	ATR = (High t – Low t), (High t – Close t-1), (Low t – Close t-1)
Chaikin Money Flow	CMF= Sum of money flow volume of 21 days / Sum of volume of 21 days

	Money flow volume = Money flow multiplayer X volume for that period Money flow multiplayer = {(close – high) – (High – Close)} / (High – Low)
Relative Strength Index	100 – [100 / (1 + (Average of Upward change in price / Average of Downward Change in price)]
Close Location Value	[(Close – low) – (high – Close)] / (High – Low)
AroonUp and AroonDn	AroonDn = 100*{(n – Low t) / n}
	AroonUp = 100*{(n – High t) / n}
Percentage Bandwidth	(Current price – Lower Bands) / (Upper Band – Lower Band)
Money Flow Index	MFI = 100 – 100 (1 + Money flow ratio) Money flow ratio = (Positive money flow / Negative money flow) Raw Money flow = Volume * Typical prices. Typical Price = (Close + High + Low) / 3
True Range	ABS (High – Low), ABS (High – close.1). ABS (Close.1 – Low) ABS: shows the absolute value Close 1: Yesterday close price
William %R	(highest high – stock price) / (highest high – lowest low) x -100
Chande Momentum Oscillator	CMO = (sH – sL) / (sH + sL) * 100

Oscillator / Summation:	{MA (first period) – MA (second period) Summation}

3.2 Data and Methodology:

Artificial neural network is one of the widely used method for forecasting stock market. It has ability to predict the stock market even the data is non-stationary. Zhang, Patuwo and Hu (1998) and (Széliga et al., 2003) suggest that ANNs provides useful properties that some of the traditional linear and non-linear regression models cannot. Such as ANN manage non stationary data as well. Support Vector machine is also a better method for predicting stock market as compare to any other traditional methods. Qu and Zhang (2016) and (Henrique et al., 2018) stated that SVM has more prediction accuracy compare to other methods. They also documented that SVM also work in volatility, even in high volatility it forecast the stock prices.

These methods are better than that of traditional methods which are usually use for forecasting. But some researchers says that Artificial Neural network is better than Support vector machine and some says that Support vector machine is better. (Kara et al., 2011) forecasted Istanbul stock market direction and stated that ANN prediction performance is 75% and SVM prediction performance is 71%. Similarly, Paik and Kumari (2017) predicted stock market and documented that SVM prediction accuracy is better than ANN and as well as Extreme learning Machine method.

This study aims to forecast the direction of stock market indexes. For this purpose, 6 indexes have selected from Asian region. The selected countries are India, Pakistan, Malaysia, China, Japan and Hong Kong. These are the main countries of ASIA which almost represents South Asia, East Asia and South East Asia. These countries also have economic relationship as there are trade relations between these countries. For forecasting daily stock price data has been taken and the data period is different for each index but it is between January, 2000 to December,

2018. From this data I have forecasted the direction of these stock market indexes by using two machine learning methods. After prediction compared both the models to find out the better model in each stock index.

First, model have trained on 80% of dataset and test on 20% dataset with the help of 30 indicators. Later the techniques identified the important features and with important variables prediction accuracy has been analyzed. The comparison between the models have also been analyzed in each stock index, because results may vary in each country due to different economies. Following table is showing the country name, index, code and time period.

Table No: 3.2

COUNTRY	INDEX	Yahoo finance CODE	START DATE	END DATE
China	SSE	^SSEC	January 2003	December 2018
Hong Kong	HIS	^HIS	January 2001	December 2018
Japan	NIKKEI 225	^N225	January 2002	December 2018
Malaysia	KLCI	^KLSE	January 2000	December 2018
India	BSE	^BSESN	January 2003	December 2018
Pakistan	KSE	^KSE	January 2010	December 2018

3.3 Empirical Model:

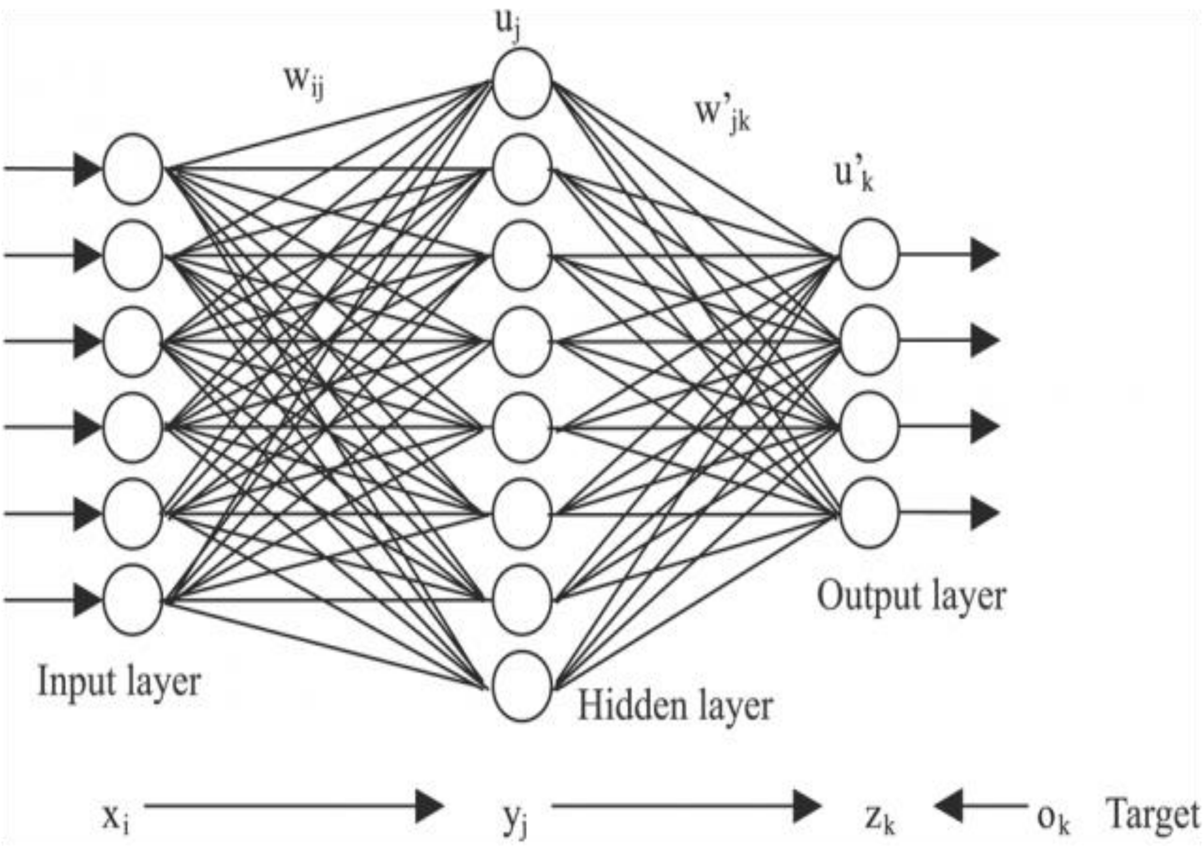
Two models Artificial Neural Network (ANN) and Support Vector Machine (SVM) has been used for prediction and to make comparison between them. There are 30 indicators to forecast the direction of stock market.

3.3.1 Artificial Neural Network:

There are three layers of artificial neural network which are input layer, hidden layer and output layer. In input layer you regress data with the help of indicators, in hidden layer it process the data with help of indicators and in third layer it gives out the result.

$$y = \alpha + \alpha 1x1 + \alpha 2x2 + \cdots + anxn + \varepsilon$$

Figure No: 3.1



Source of figure: extremetech.com

Where,

y: Forecasted Value or Output layer

α: Coefficient which are showing weights,

x: Variables as shown in Table: 01

xn: Input layer

3.3.2 Support Vector Machine:

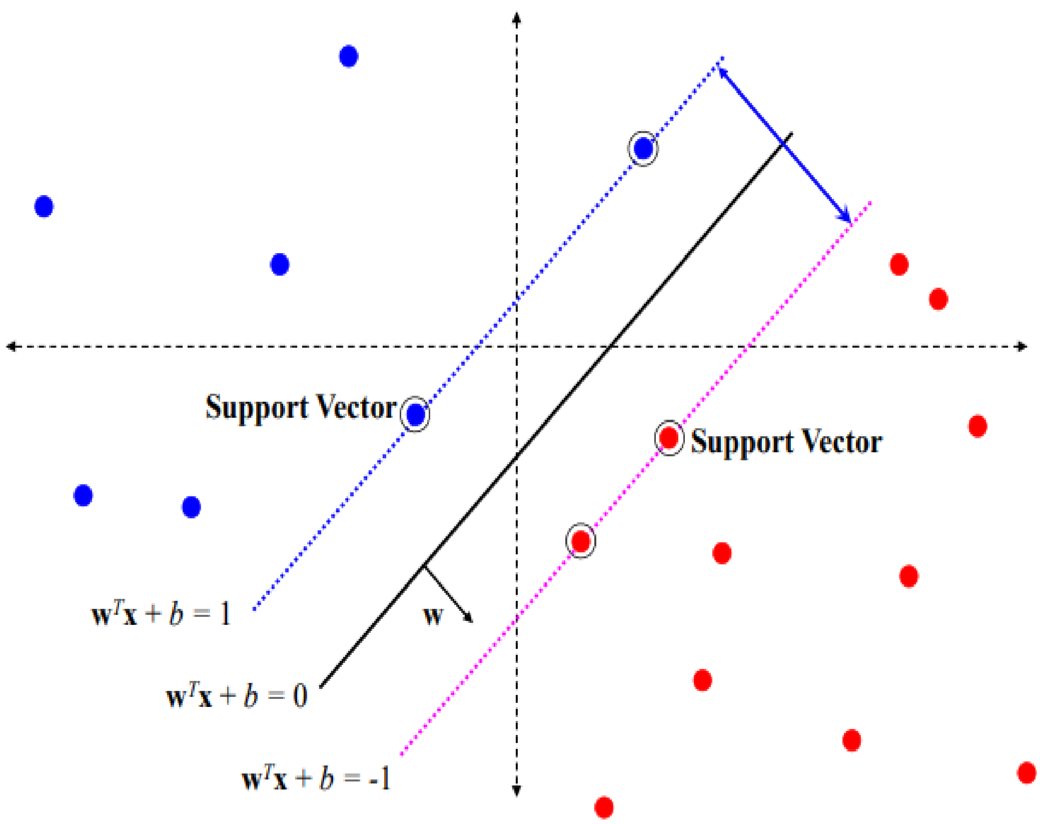
SVM is basically use for classification purpose. With the help of independent variables, it classifies observations between groups. SVM first train the dataset with the help of variables and then separate the dataset in number of dimensions. This separation line is known as hyper plane which define the direction of observations. Following equation defines the hyper plane.

$WXi + b = 0$…………….. (i)

$WXi + b > 1$ para $Yk = 1$ …………….. (ii)

$WXi + b < 1$ para $Yk = -1$…………… (iii)

Figure No: 3.2



Source of figure: datasciencecentral.com

Where,

Y: sample classification

W: Weights

X: Variables

b: Constant

The results which have been extracted from the ANN and SVM are shown by confusion matrix. So I am first describing confusion matrix and indicators of balanced accuracy.

3.4 Confusion Matrix:

It is formed from the four outcomes, which is produced as a result of binary classification. The four outcomes are True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN).

Where,

TP shows accurate positive prediction

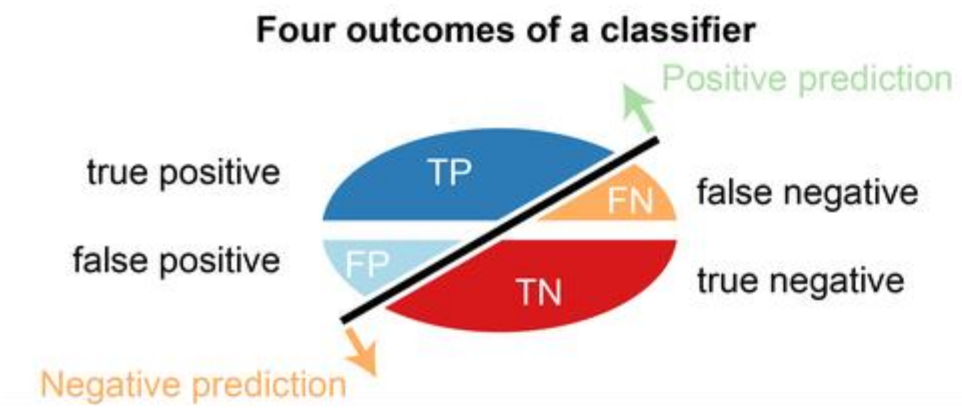
TN shows accurate negative prediction

FP shows inaccurate positive prediction

FN shows inaccurate negative prediction

Figure No: 3.3

Confusion Matrix



Source of figure: classeval.wordpress.com

Confusion Matrix shows the following results, which defines the predictive performance.

3.5 Accuracy:

Accuracy is calculated as the no. of all correct predictions divided by the total no. of the dataset. The best accuracy is 1.0, whereas the worst is 0.0. It can also be calculated by 1 – ERR.

3.6 Sensitivity (Recall or True Positive Rate):

Sensitivity is calculated as the number of correct positive predictions divided by the total number of positives. It is also called recall or true positive rate. The best sensitivity is 1.0 and the worst is 0.0.

3.7 Specificity (True negative rate):

Specificity is calculated as the number of correct negative predictions divided by the total number of negatives. It is also called true negative rate. The best specificity is 1.0 and the worst is 0.0.

3.8 Precision (Positive predictive value):

Precision is calculated as the number of correct positive predictions divided by the total number of positive predictions. It is also called positive predictive value. The best precision is 1.0 and the worst is 0.0.

3.9 False positive rate:

False positive rate is calculated as the number of incorrect positive predictions divided by the total number of negatives. The best false positive rate is 0.0 and the worst is 1.0. It can also calculated as 1 – specificity.

3.10 Cohen's kappa:

Kappa statistic is a very good measure that can handle very well both multi-class and imbalanced class problems. According to their scheme a value < 0 is indicating no agreement, 0–0.20 as slight agreement, 0.21–0.40 as fair agreement, 0.41–0.60 as moderate agreement, 0.61–0.80 as substantial agreement, and 0.81–1 as perfect agreement.

Table No: 3.3

Balanced Accuracy indicators	**Calculations**
Accuracy	(TP + TN) / (P + N)
Kappa Value	(Total Accuracy – Random Accuracy) / (1-Random Accuracy)
Sensitivity	TP / P
Specificity	TN / N
Positive Predictive Value	TP / (TP + FP)
Negative Predictive Value	FP / (TN + FP)

CHAPTER 4

Results and Interpretations:

There are two model approach for each technique. Before use any approach first the study train the models on 80% dataset and then test the model on 20% dataset. For 1st approach model the technique test the dataset by using all the selected variables and for 2nd approach model, the study used only important variables which have been find out by the technique itself. There are two methods of machine learning are using in which one is support vector machine. First model extracted results by using all the selected variables. The input frame of the 1st model is as follows.

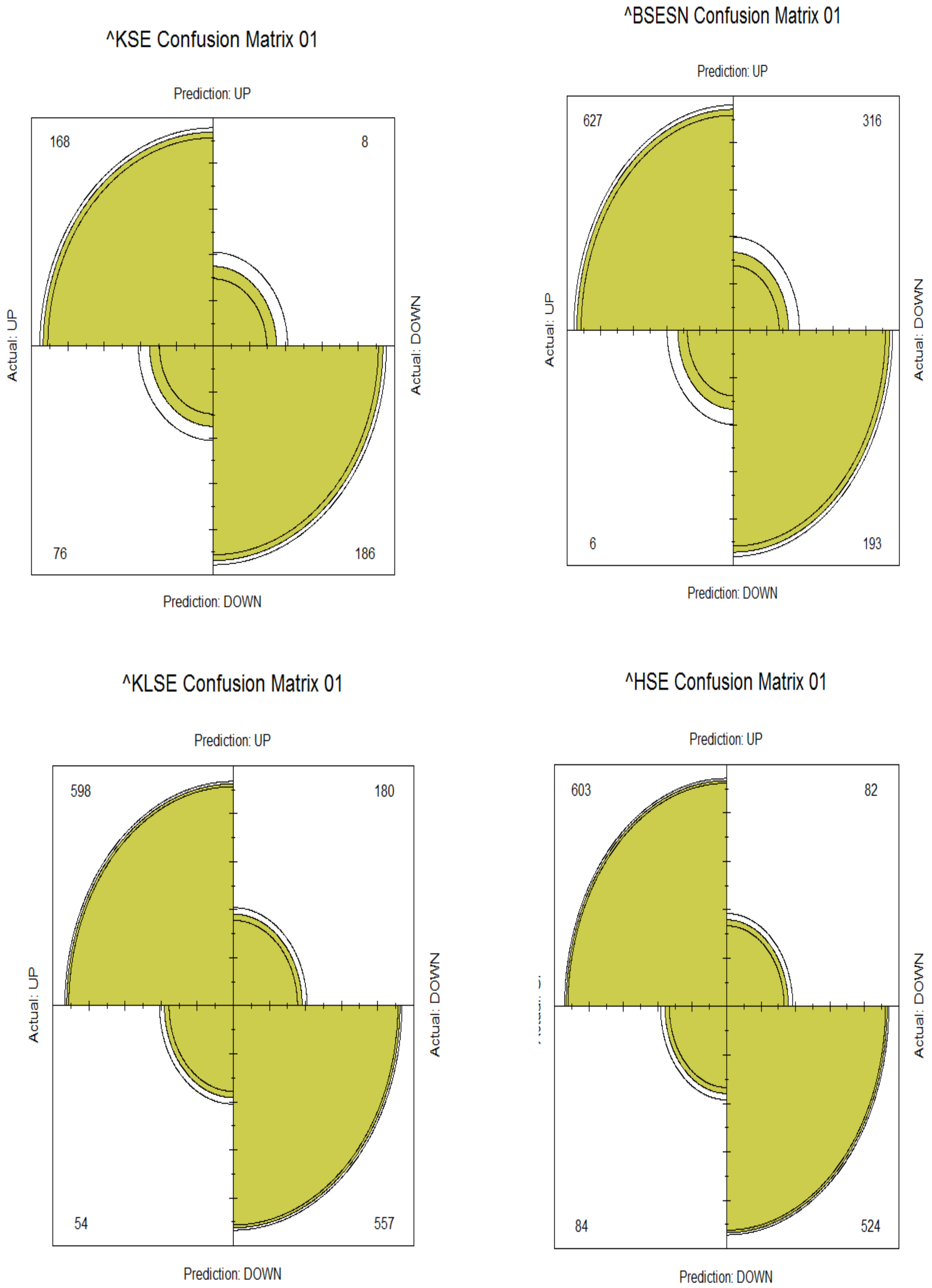
"Class ~

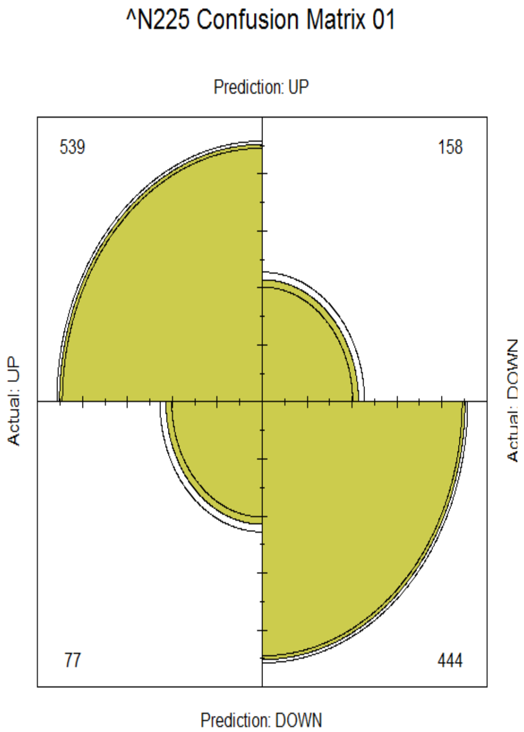
RSI, EMAcross, MACD, SMI, WPR, ADX, CCI, CMO, ROC, aroonUp, aroonDn, oscillator, TR, ATR, TrueHigh, TrueLow, LowerBB, MiddleMA, UpperBB, PercentageB, ChaikinAD, CLV, CMF, Donchain_high, Donchian_mid, Donchain_low, DPO, MFI, OBV, SAR"

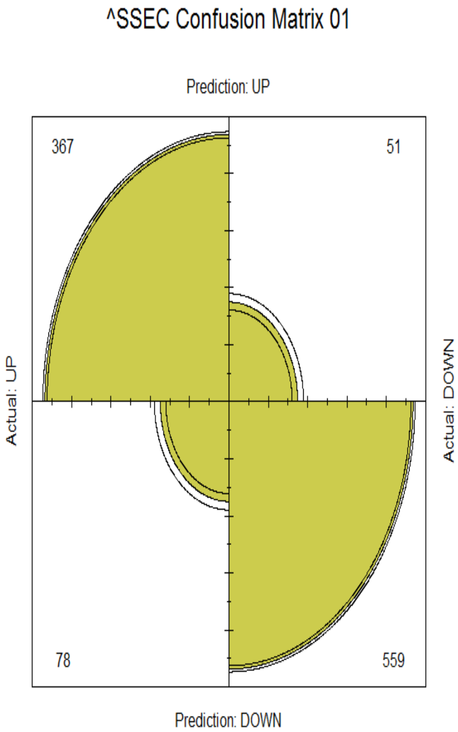
Class shows the dependent variable while other are the selected independent variables. SVM results extracted by using first model approach are as follows:

Figure No: 4.1

Confusion Matrix







The confusion matrices shows the correct and wrong prediction of prices for each index. True positive (TP) and True Negative (TN) are correct prediction while False Positive (FP) and False Negative (FN) are the wrong predicted prices.

Table No: 4.1

Balanced Accuracy indicators	Indexes					
	KSE	**BSESN**	**KLSE**	**HSE**	**N225**	**SSEC**
Accuracy	0.8082	0.718	0.8315	0.8716	0.8071	0.8777
95% CI	(0.7682, 0.844)	(0.691, 0.744)	(0.8108, 0.8509)	(0.8521, 0.8894)	(0.7838, 0.8289)	(0.8564, 0.8969)
No Information Rate	0.5571	0.5543	0.5306	0.5313	0.5057	0.5782
P-Value [Acc > NIR]	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2e-16	< 2.2e-16	< 2e-16

Kappa	0.6248	0.3931	0.6655	0.7423	0.6135	0.7472
McNamara’s Test P-Value	2.67E-13	< 2.2e-16	3.05E-16	0.9381	1.80E-07	2.21E-02
Sensitivity	0.6885	0.9905	0.9172	0.8777	0.875	0.8247
Specificity	0.9588	0.3792	0.7558	0.8647	0.7375	0.9164
Pos Pred Value	0.9545	0.6649	0.7686	0.8803	0.7733	0.878
Neg Pred Value	0.7099	0.9698	0.9116	0.8618	0.8522	0.8776
Balanced Accuracy	**0.8236**	**0.6848**	**0.8365**	**0.8712**	**0.8063**	**0.8706**

The results concluded that the balanced prediction accuracy is more than 80% for all the indexes except in case of BSESN from support vector machine. For KSE accuracy is high near Kappa value shows the substantial agreement, rejected Ho because P value is less than 0.05. Specificity is very high. Positive predicted value shows the correct up price prediction it is almost accurate here. For BSESN prediction accuracy is not that much high as of KSE. Kappa value states the slightly agreement between observed and expected values, while P value shows the significant result. Sensitivity is very high and negative predicted value shows that prediction of down values is almost 100% correct. Now for KLSE prediction accuracy is is high Kappa value shows substantial agreement. Sensitivity and negative predicted value is high here which shows the high prediction accuracy in down prices.

For HSE the prediction accuracy is also high. Kappa value shows the substantial agreement between observed and expected values. P value is less than 0.05 so reject Ho. Sensitivity and specificity both are good here. Positive predicted value and negative predicted value are high which shows high accuracy in prediction for both up and down prices of the index. Next country is Japan, its symbol is N225, its prediction accuracy is high but Kappa states the substantial agreement. Sensitivity is high. Positive predicted value and negative predicted value both are good here. For SSEC the accuracy of prediction is very high as compare to other indexes. Kappa value shows the substantial agreement. Sensitivity and specificity both have high values. Positive predicted value and negative predicted value are also high shows the high prediction accuracy.

For second model approach the study have to find out the best variables, which actually defines the dataset of each country which I have selected. The most important variables find out by the model are as follows.

Table No: 4.2

Importance Variables						
	KSE	**BSESN**	**KLSE**	**HSE**	**N225**	**SSEC**
CLV	0.9292	0.9516	0.9152	0.9255	0.9516	0.955
ROC	0.8215	0.7904	0.7985	0.7164	0.7904	0.7651
WPR	0.738	0.7237	0.7278	0.6476	0.7237	0.7098
CCI	0.7198	0.715	0.7184	0.6437	0.715	0.7022
Percentage BB	0.6359	0.6245	0.6396	0.5762	0.6245	0.6184
Aroon Up	0.6178	0.5899	0.5746	0.5329	0.5899	0.574
CMF	0.6102	0.6101	0.5752	0.5846	0.6101	0.613
CMO	0.5955	0.6004	0.6064	0.5631	0.6004	0.614
MFI	0.5716	0.5674	0.5798	0.5402	0.5674	0.5783
Oscillator	0.5568	0.5497	0.5564	0.5219	0.5497	0.5733

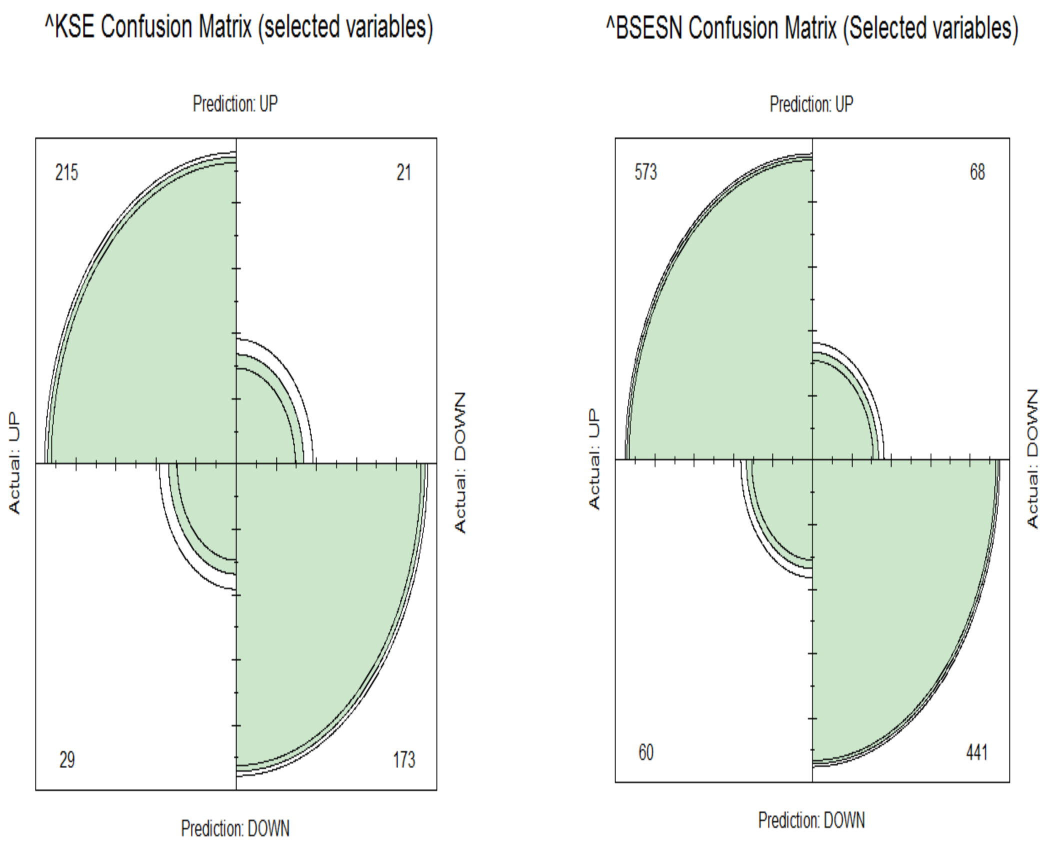
CLV percentage is most important variable find out by SVM as compare to any other variable. While ROC is 2nd most important variable and WPR is 3rd most important variable. Similarly, all these variables are most important variables which has been used and it actually defines the dataset for all the countries which I have selected. Now the study converse about the 2nd approach of the model which states about the important variable. The input frame of the 2nd model is as follows.

"Class ~

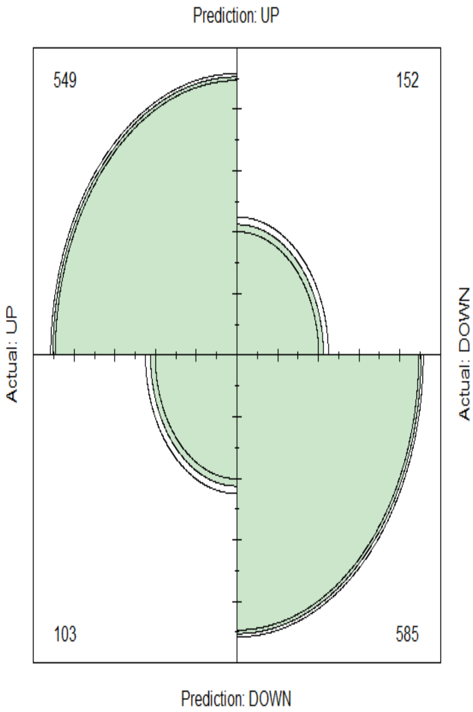
CLV+ROC+WPR+CCI+PercentageBB+AroonUp+CMF+CMO+MFI+Oscillator

Class shows the dependent variable while other 10 indicators are the most important variable. The confusion matrices extracted by using important variables are as follows.

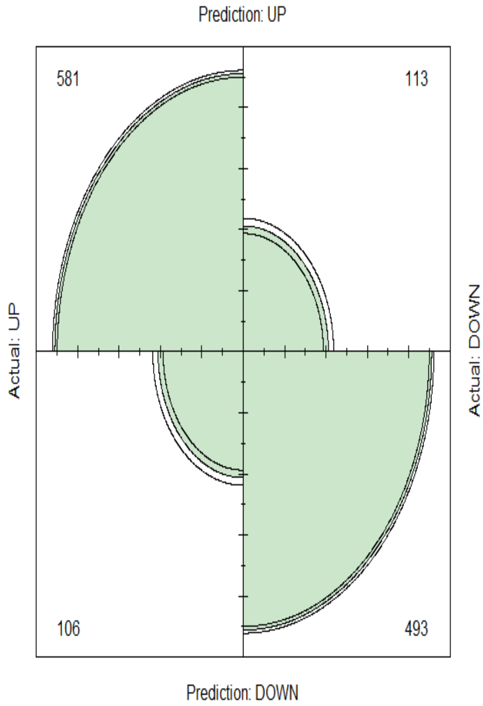
Figure No: 4.2



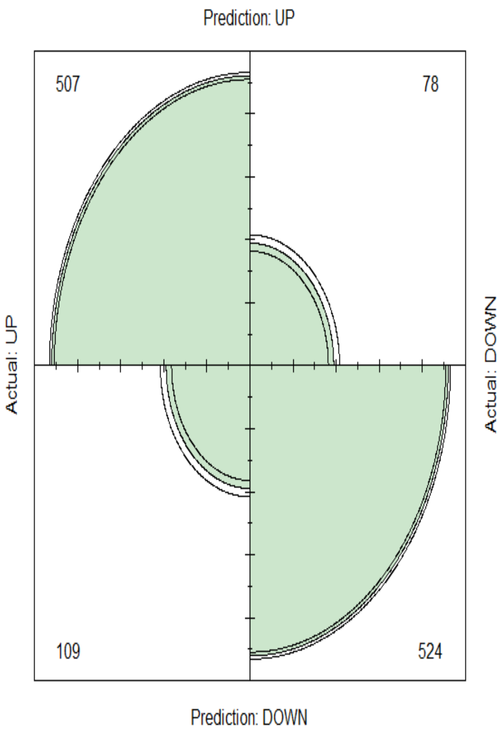
^KLSE Confusion Matrix (Selected variables)



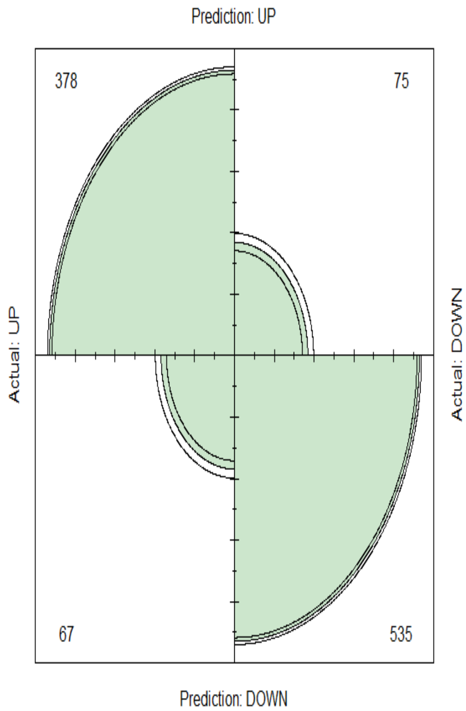
^HSE Confusion Matrix (selected variables)



^N225 Confusion Matrix (Selected Variables)



^SSEC Confusion Matrix (Selected Variables)



The confusion matrices shows the correct and wrong prediction price for each selected index. True positive indicate that up prices have been predicted accurately while True negative shows that down prices have been predicted accurately. Following table examine the balanced accuracy for each index extracted by using important variables

Table No: 4.3

Balanced Accuracy indicators	**Indexes**					
	KSE	**BSESN**	**KLSE**	**HSE**	**N225**	**SSEC**
Accuracy	0.8881	0.8879	0.8164	0.8306	0.8465	0.8654
95% CI	(0.8548, 0.9161)	(0.8682, 0.9056)	(0.795, 0.8364)	(0.809, 0.8507)	(0.825, 0.8663)	(0.8433, 0.8854)
No Information Rate	0.5571	0.5543	0.5306	0.5313	0.5057	0.5782
P-Value [Acc > NIR]	< 2e-16	< 2e-16	< 2.2e-16	< 2e-16	< 2e-16	< 2e-16
Kappa	0.7732	0.7728	0.633	0.6597	0.6931	0.7247
McNamara's Test P-Value	1.00E+00	0.5361	2.65E-03	0.6852	0.02825	0.5569
Sensitivity	0.9016	0.9052	0.842	0.8457	0.8231	0.8494
Specificity	0.8711	0.8664	0.7938	0.8135	0.8704	0.877
Pos Pred Value	0.898	0.8939	0.7832	0.8372	0.8667	0.8344
Neg Pred Value	0.8756	0.8802	0.8503	0.823	0.8278	0.8887
Balanced Accuracy	**0.8864**	**0.8858**	**0.8179**	**0.8296**	**0.8467**	**0.8476**

The results extracted that there is significant improvement in the accuracy of the model. Prediction accuracy has been increased by using important variables for 3 countries including KSE, BSESN and N225 but there is a minor decrease in accuracy of other 3 countries. For BSESN, there is a distinguish change in accuracy. In 1st model the accuracy was 68% but by using important variable there is almost 20% increase in the prediction accuracy. Kappa value has also increased and shows the substantial agreement. Sensitivity, specificity, Positive predicted value and negative predicted value are high here and states that the prediction accuracy for up and down prices is high. Similarly, for KLSE there is a minor decrease in accuracy, it was 83.14% but now it become 81.64%. As a result the Kappa value has also declined but shows the substantial agreement. P value is less than 0.05. Negative predicted value has value as compare to positive predicted value so it states that the prediction accuracy in down prices is higher than that of up prices.

For HSE there is a minor decline in prediction accuracy but Kappa value indicate the substantial agreement. Sensitivity and specificity both has been decreased and but still shows the high values. Next selected country is Japan, its symbol is N225. The prediction accuracy for N225 has been increased, it was 80.71% now it has become 84.65% so the value of Kappa has also been increased and indicate substantial agreement. P value is also significant. Sensitivity has been decreased buy there is an increase in specificityPositive predicted value has increased here and negative price value has been decreased but shows the high accuracy in predictions. For SSEC, there is a minor decrease in accuracy it has been reached from 87.77% to 86.54%. So the Kappa has also decreased. Sensitivity, specificity, Positive predicted value and negative predicted value is high shows the high accuracy for both up and down prices.

The study again run the model to find out the best parameters, which helps me to generate the optimize results. The actual parameter which is using Cost: 1000 and Gamma: 0.0001 now the study find out best parameter for each country index as shown in table below.

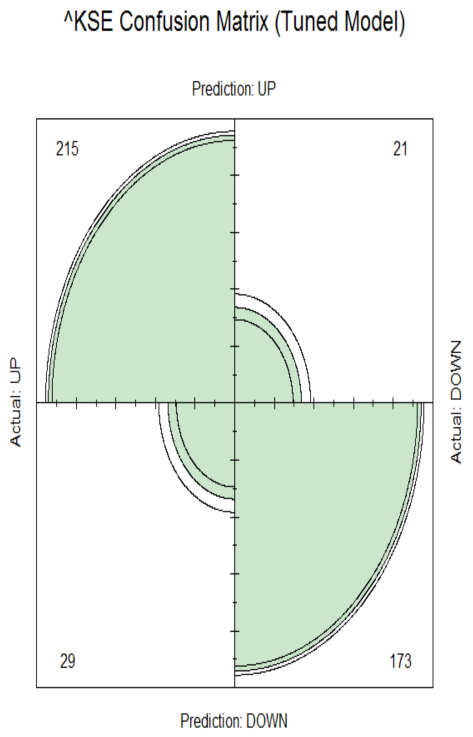
Table No: 4.4

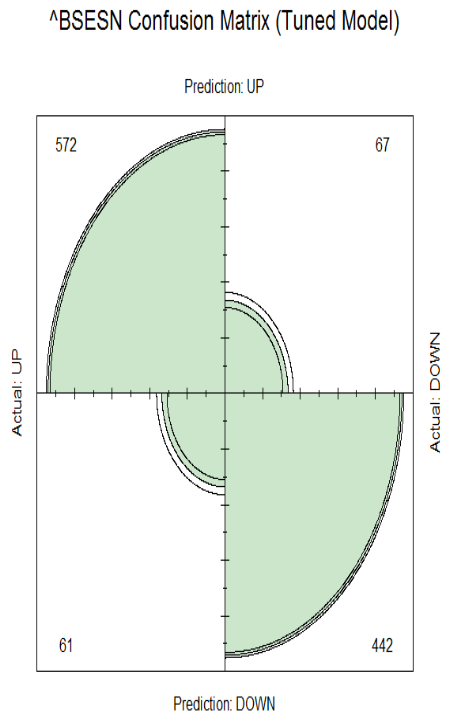
BEST PARAMETERS						
	KSE	BSESN	KLSE	HSE	N225	SSEC
COST	100	10	100	10	10	10
GAMMA	0.001	0.001	0.1	0.01	0.0001	0.0001

Cost is 100 for KSE, BSESN and for CSE. But for HSE, N225 and SSEC it is 10. Gamma is 0.001 for KSE, BSESN and CSE. For HSE it is 0.01. For N225 and for SSEC the Gamma is 0.0001. The result concluded by using the best parameters are as follows:

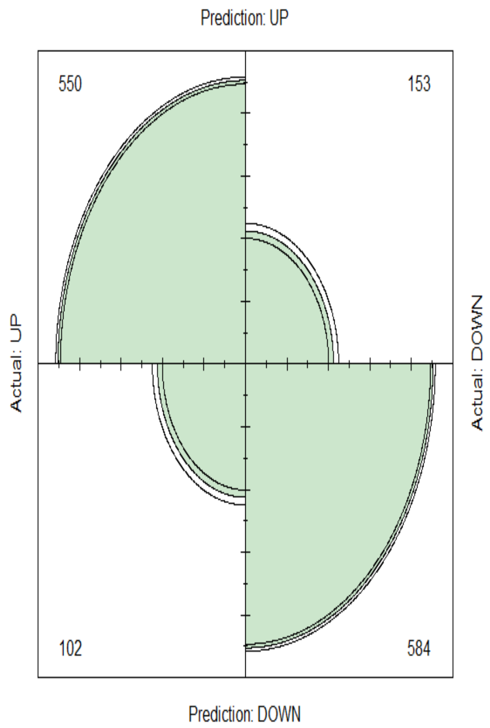
Figure No: 4.3

Now the study used the best parameter to get the better results for each index. The study used the important variable approach but use the optimize parameters for each index. Following matrices shows the results explored by using optimized parameters.

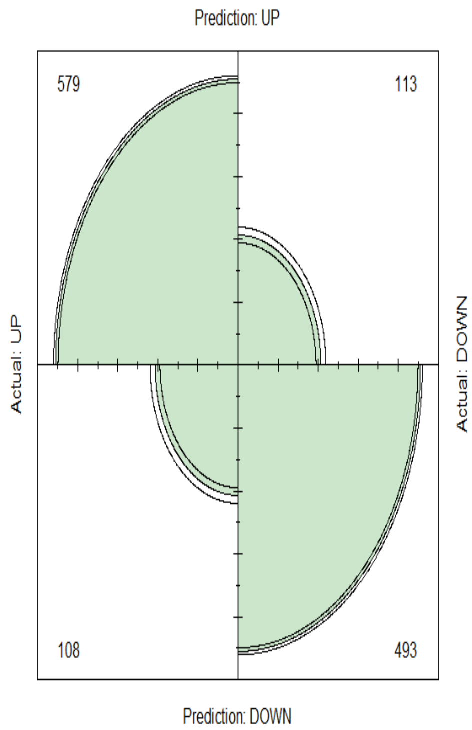




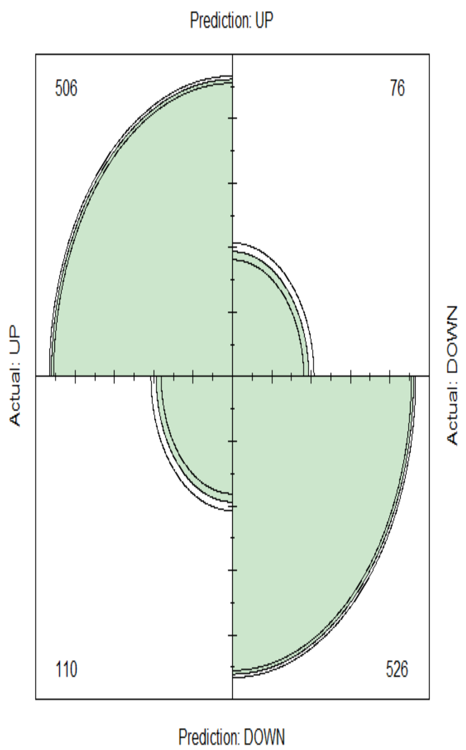
^KLSE Confusion Matrix (Tuned Model)



^HSE Confusion Matrix (Tuned Model)



^N225 Confusion Matrix (Tuned Model)



^SSEC Confusion Matrix (Tuned model)

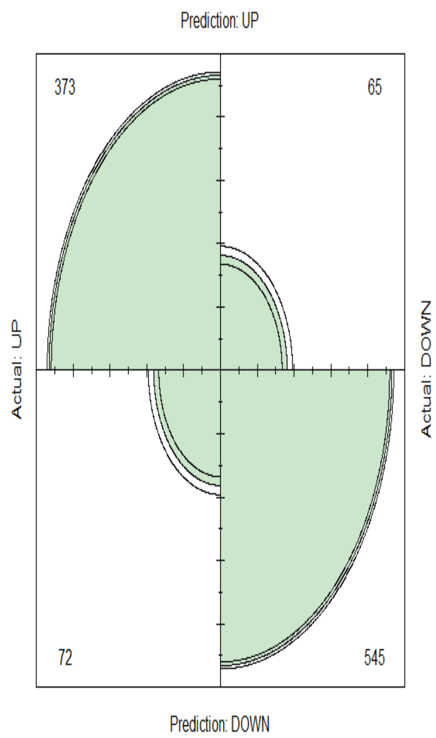


Table No: 4.5

Balanced Accuracy Indicators	Indexes					
	KSE	BSESN	KLSE	HSE	N225	SSEC
Accuracy	0.8858	0.8879	0.8461	0.8164	0.8473	0.8701
95% CI	(0.8523, 0.9141)	(0.8682, 0.9056)	(0.8194, 0.8702)	(0.795, 0.8364)	(0.8258, 0.867)	(0.8483, 0.8898)
No Information Rate	0.5571	0.5543	0.5197	0.5306	0.5057	0.5782
P-Value [Acc > NIR]	< 2e-16	< 2.2e-16	< 2e-16	< 2.2e-16	< 2e-16	< 2e-16
Kappa	0.7696	0.7729	0.6924	0.6331	0.6947	0.7332
McNamara's Test P-Value	0.3222	< 2.2e-16	2.00E-02	1.74E-03	0.01553	0.6082
Sensitivity	0.8811	0.9036	0.8199	0.8436	0.8214	0.8382
Specificity	0.8918	0.8684	0.8744	0.7924	0.8738	0.8934
Pos Pred Value	0.911	0.8951	0.8759	0.7824	0.8694	0.8516
Neg Pred Value	0.8654	0.8787	0.8177	0.8513	0.827	0.8833
Balanced Accuracy	**0.8865**	**0.886**	**0.8471**	**0.818**	**0.8476**	**0.8476**

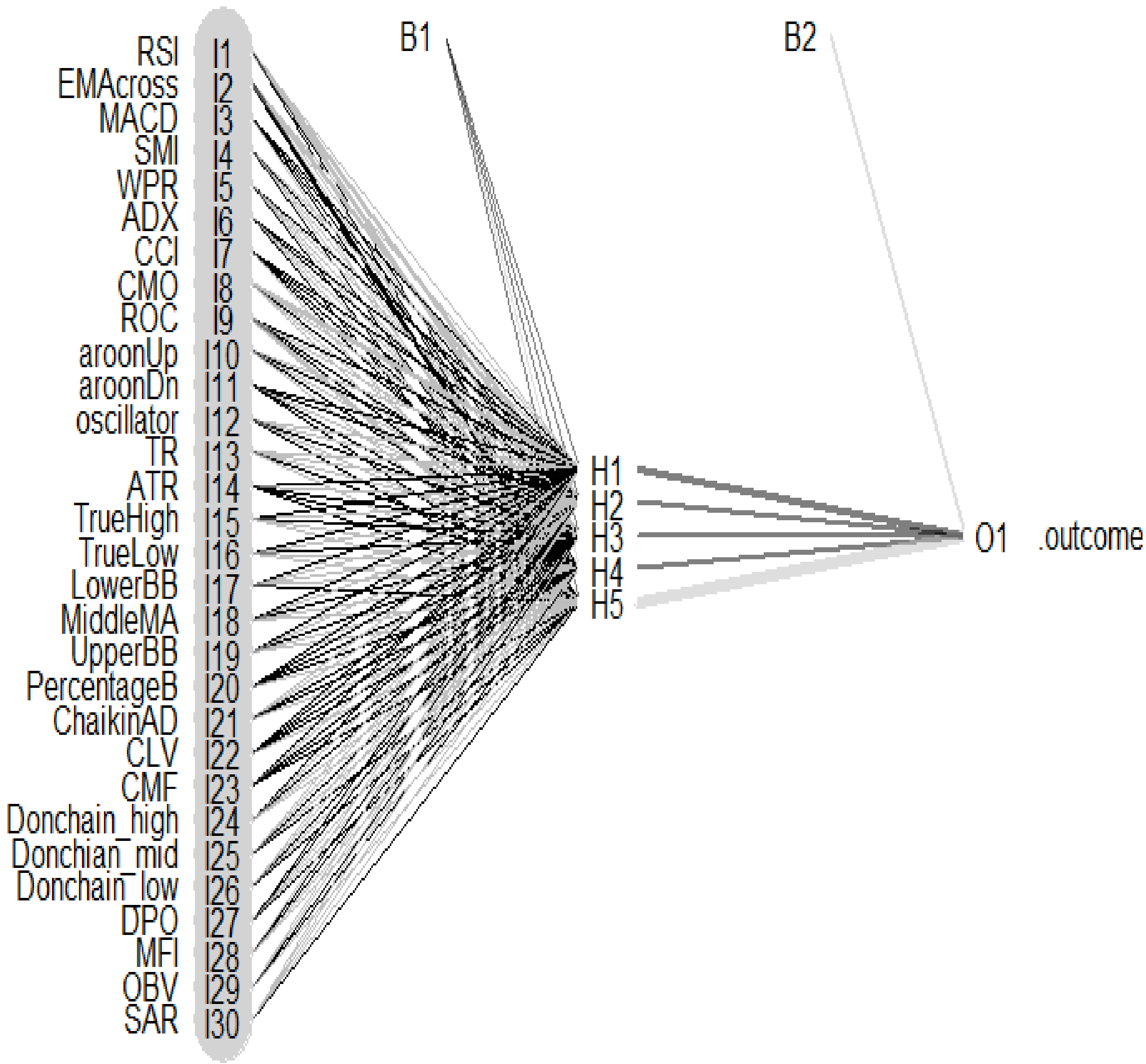
The balanced accuracy has been enhanced little for all the countries by using optimized parameters except for HSE. For KSE the balanced accuracy by using selected variables was 88.64% but now it has become 88.65%. For BSESN, the balanced accuracy was 88.58% and now it has been increased and reached to 88.6%. For KLSE there is a minor increased in balanced accuracy that it become 81.8% as it was 81.79%.

In case of HSE, there is a minor decrease in accuracy, balanced accuracy was 82.96% but now it has become 82.82%. In case of N225 balance accuracy was 84.67%, now it has been increased and reached to 84.76%. Minor increase in specificity and positive predicted value make its accuracy higher. For china, its symbol is SSEC the balanced accuracy is same here but there is an increase in accuracy of the model, which makes it better than selected variables model.

Another method which is using to check the prediction accuracy is artificial neural network. There are three layers of neural network 1st is input layer in which information is given, 2nd is hidden layer by which it can process the dataset and 3rd layer provide the result. Following graph shows that how neural network work with given 30 indicators.

Figure No: 4.4

ANN Process by using 30 indicators



The first layer is showing the 30 indicators, while B1 is showing the hidden layer and B2 is the outcome of the process of input and hidden layers. There are also two model approach first we examine the 1st model approach.

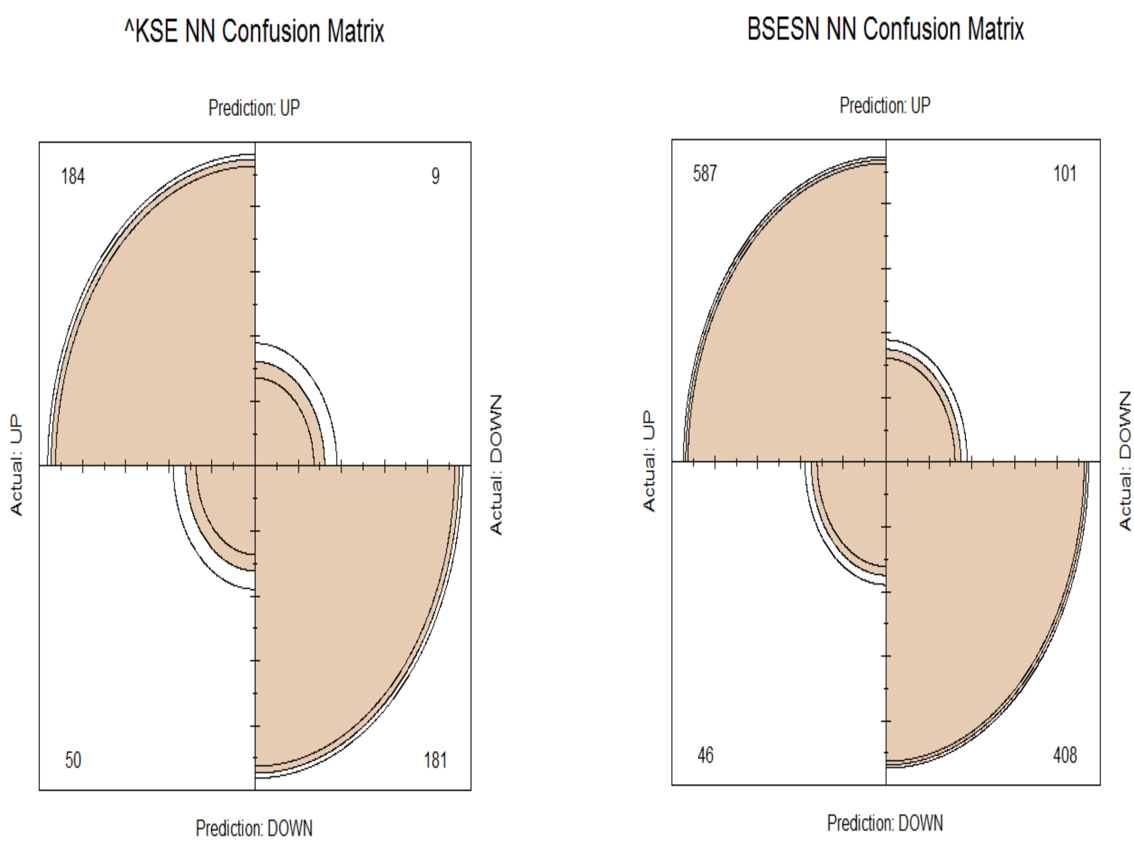
For 1st model approach the input frame is

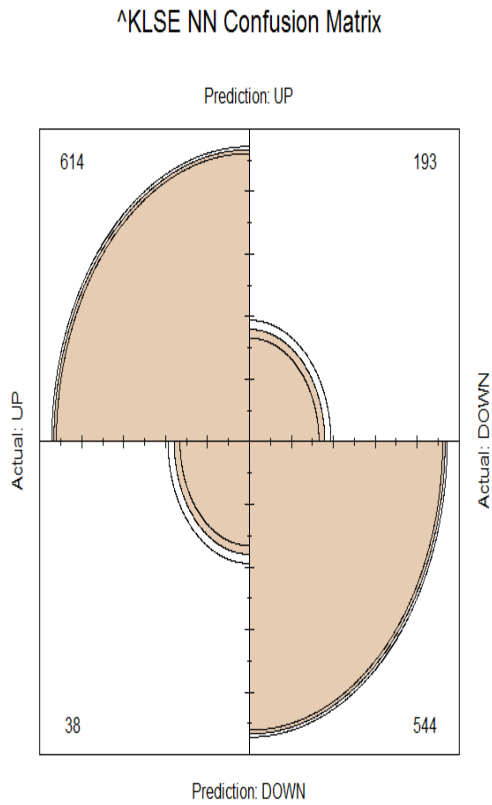
"Class ~

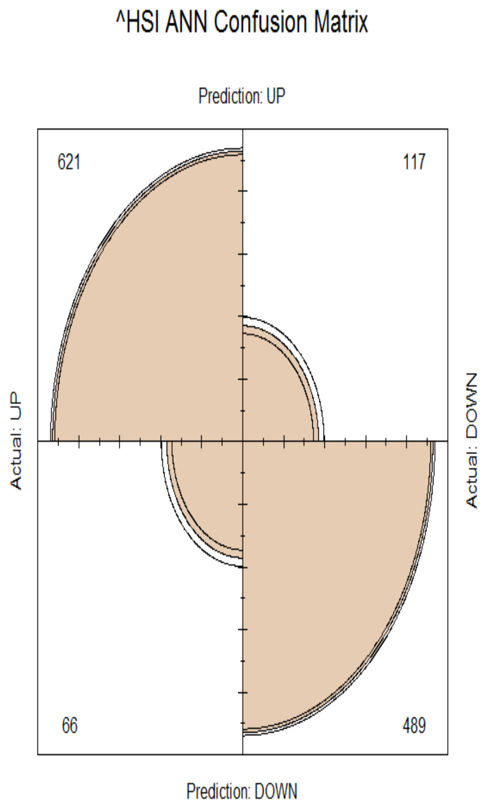
RSI, EMAcross, MACD, SMI, WPR, ADX, CCI, CMO, ROC, aroonUp, aroonDn, oscillator, TR, ATR, TrueHigh, TrueLow, LowerBB, MiddleMA, UpperBB, PercentageB, ChaikinAD, CLV, CMF, Donchain_high, Donchian_mid, Donchain_low, DPO, MFI, OBV, SAR”

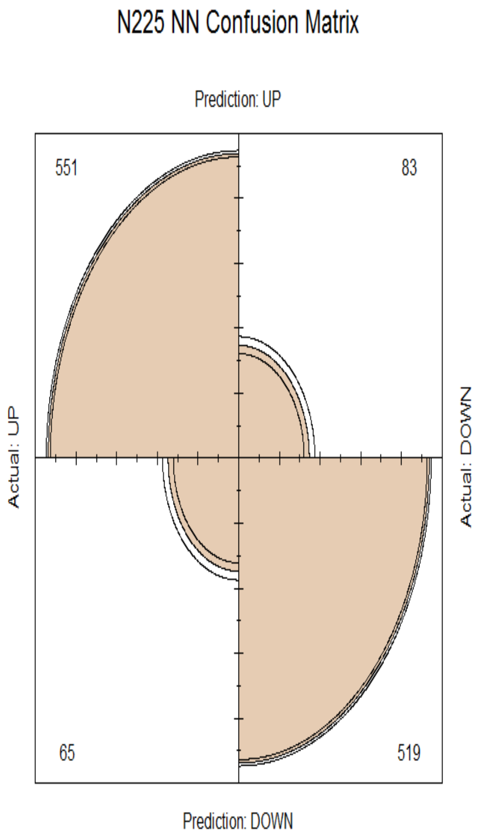
Class indicating the dependent variable while other are the independent variable. The results extracted from the neural networks by using all the variables is showing in the following graph of confusion matrix.

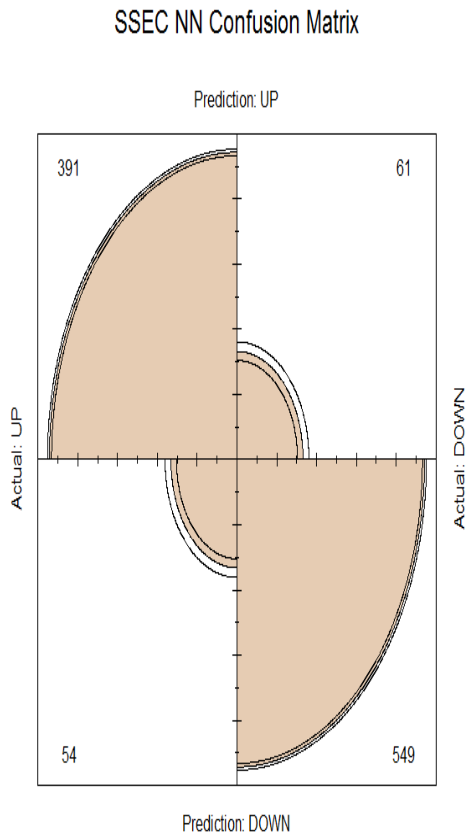
Figure No: 4.5




^KLSE NN Confusion Matrix
Prediction: UP
614
193
38
544
Prediction: DOWN
Actual: UP
Actual: DOWN


^HSI ANN Confusion Matrix
Prediction: UP
621
117
66
489
Prediction: DOWN
Actual: UP
Actual: DOWN


N225 NN Confusion Matrix
Prediction: UP
551
83
65
519
Prediction: DOWN
Actual: UP
Actual: DOWN


SSEC NN Confusion Matrix
Prediction: UP
391
61
54
549
Prediction: DOWN
Actual: UP
Actual: DOWN

The confusion matrices shows the correct and false prediction of all the prices. True positive and true negative shows the correct prediction while false positive and false negative indicate towards the inaccurate predictions of the prices. The following table shows the balanced accuracy and the indicators of balanced accuracy by using all the variables.

Table No: 4.6

Balanced Accuracy Indicators	**Indexes**					
	KSE	**BSESN**	**KLSE**	**HSE**	**N225**	**SSEC**
Accuracy	0.8608	0.8713	0.8337	0.8585	0.8785	0.891
95% CI	(0.8242, 0.8924)	(0.8505, 0.8902)	(0.8131, 0.8529)	(0.8383, 0.877)	(0.8588, 0.8963)	(0.8706, 0.9092)
No Information Rate	0.5519	0.5543	0.5306	0.5313	0.5057	0.5782
P-Value [Acc > NIR]	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2.2e-16	<2e-16	<2e-16
Kappa	0.7243	0.7367	0.6707	0.7144	0.7569	0.777
McNamara's Test P-Value	1.91E-07	8.44E-06	< 2.2e-16	2.19E-04	0.1623	0.5758
Sensitivity	0.7863	0.9273	0.9417	0.9039	0.8945	0.8787
Specificity	0.9526	0.8016	0.7381	0.8069	0.8621	0.9
Pos Pred Value	0.9534	0.8532	0.7608	0.8415	0.8691	0.865
Neg Pred Value	0.7835	0.8987	0.9347	0.8811	0.8887	0.9104
Balanced Accuracy	**0.8695**	**0.8645**	**0.8399**	**0.8554**	**0.8783**	**0.8893**

The result concluded that the accuracy prediction by using all the 30 indicators is also high from artificial neural networks. The prediction accuracy is more than 85% for all the indexes but in case of KLSE the prediction accuracy is near to 85%. And Kappa value shows the substantial agreement between expected and observed values for all the indexes. For KSE. P value is significant here. Specificity is and Positive predicted value is very high which shows high accuracy in positive prediction. For BSESN, Ho rejected because P value is less than 0.05. Sensitivity is very high. Positive predicted value and negative predicted value is also high which shows high accuracy in positive and negative predictions. For KLSE Sensitivity and negative predicted value is very high which shoes high accuracy in negative prediction prices. For HSE, sensitivity, positive predicted value and negative predicted value is high shows accuracy in predictions.

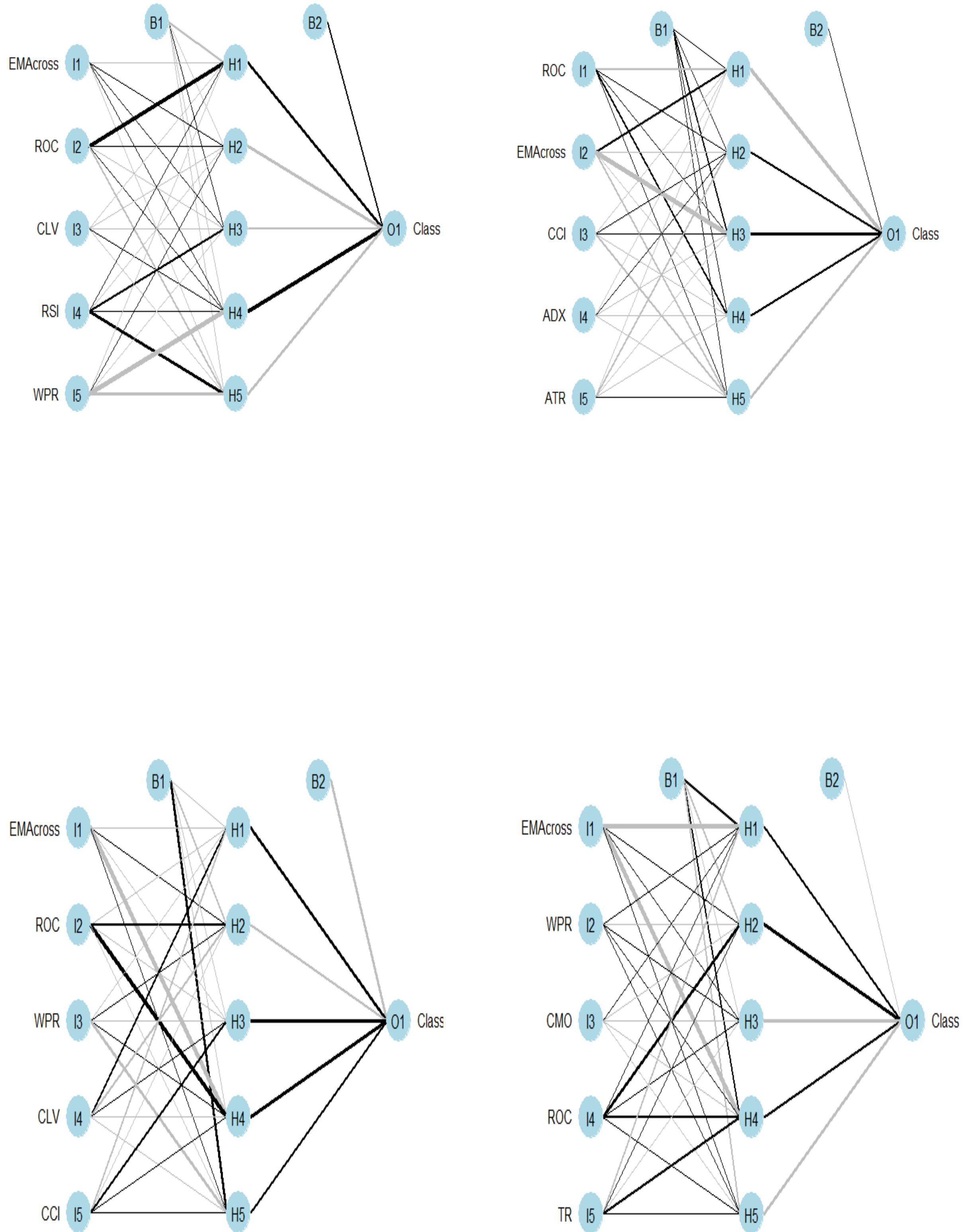
For N225 sensitivity, specificity, positive predicted value and as well as negative predicted value is shows high accuracy in prediction of both positive and negative values. The last country is China and its symbol is SSEC. P value is significant here. Sensitivity, specificity, positive predicted value and as well as negative predicted value is very high and as a result the prediction accuracy is very high. Now the study find out the best indicators by using ANN, which actually defines the dataset. Following table shows the most important variables selected by the model itself.

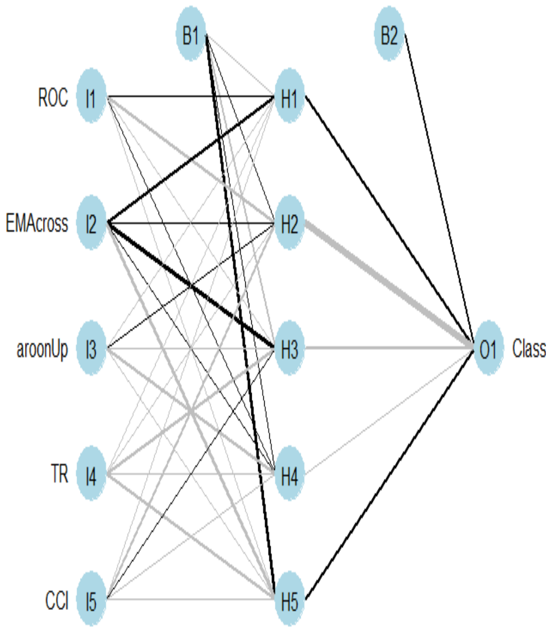
Table No: 4.7

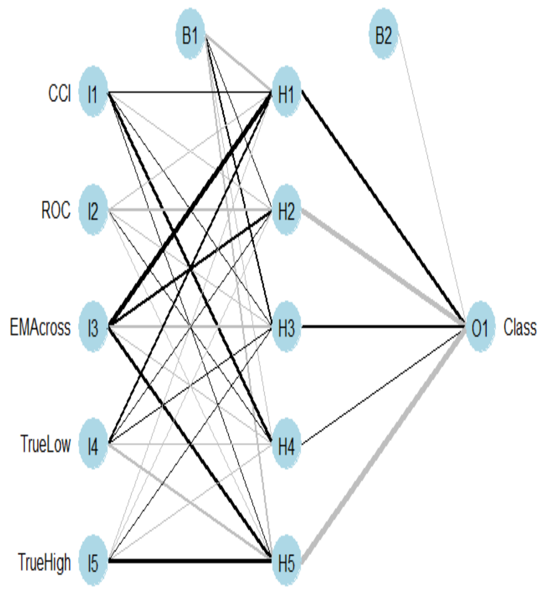
Important Variables					
KSE	**BSESN**	**KLSE**	**HSE**	**N225**	**SSEC**
EMAcross	EMAcross	EMAcross	EMAcross	ROC	CCI
ROC	ROC	ROC	ROC	EMAcross	ROC
CLV	CCI	WPR	AroonDn	AroonUp	EMAcross
RSI	ADX	CLV	TR	TR	TrueLow
WPR	ATR	CCI	CCI	CCI	TrueHigh

Figure No: 4.6

Following figures shows the ANN process for each index by using the important variables.







These above graphs showing that how neural network is producing results with the help of input and hidden layer by using important variables. For each country index there are 5 important variables which are extracted by the neural networks. The input frame of the 2nd model is as follows.

KSE: "Class ~ EMAcross+ROC+CLV+RSI+WPR"

BSESN: "Class ~ EMAcross+ROC +CCI+ADX+ATR"

KLSE: "Class ~ EMAcross+ROC +WPR+CLV+CCI"

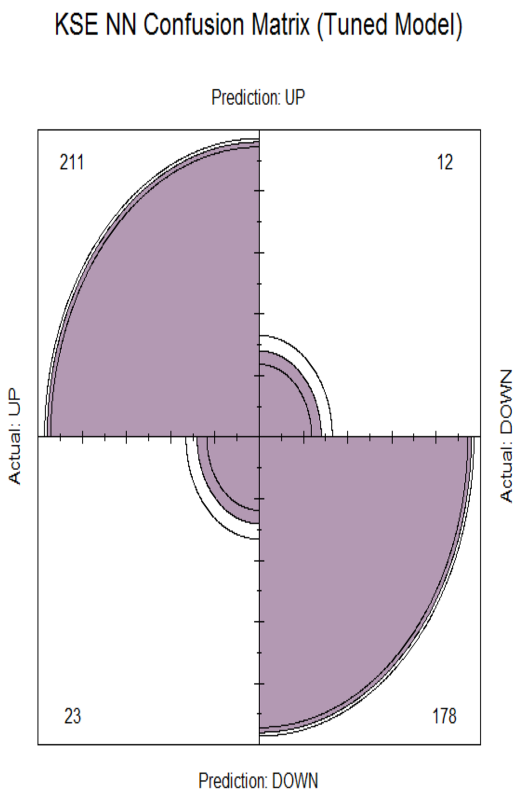
HSE: "Class ~ ROC+ EMAcross+AroonUp+TR+CCI"

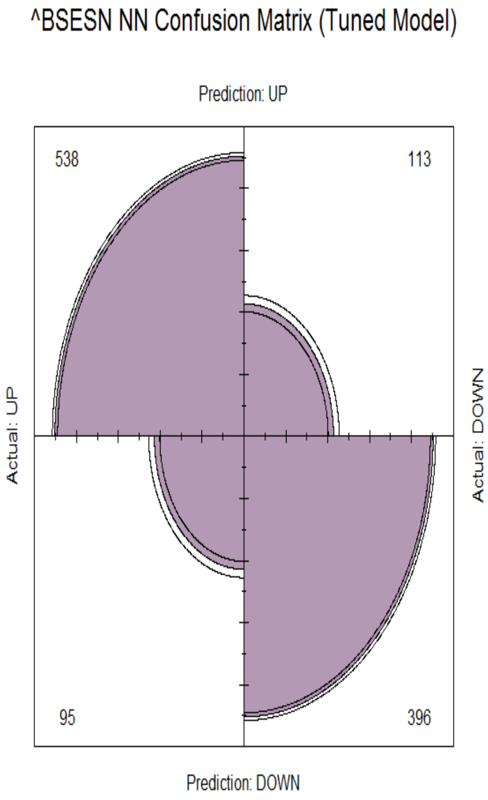
N225: “Class ~ ROC+ EMAcross+AroonUp++TR+CCI”

SSEC: “Class ~ CCI+ROC+EMAcross +TrueLow+TrueHigh”

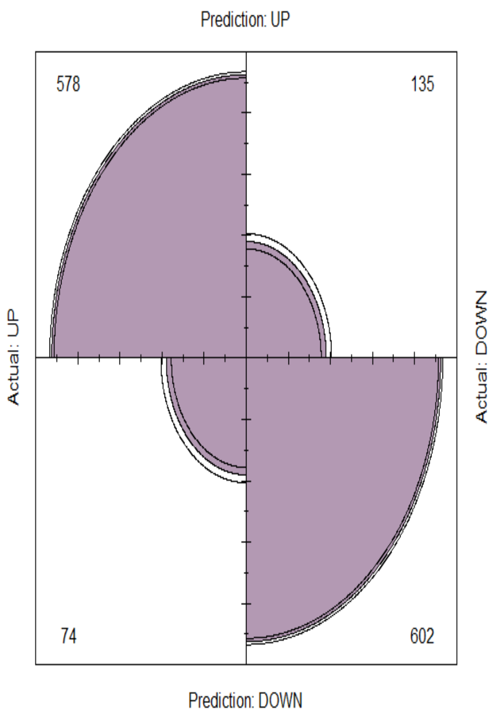
Class shows the dependent variable while other 5 indicators are the most important independent variables for each index. Results extracted from the above important variables are shown by confusion matrix:

Figure No: 4.7

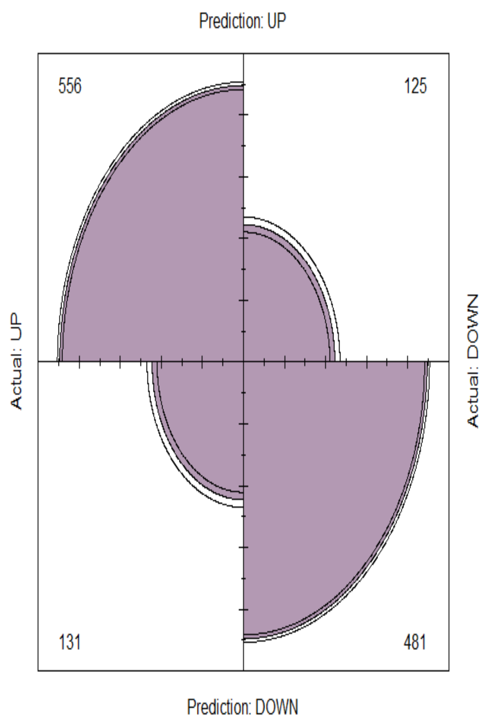




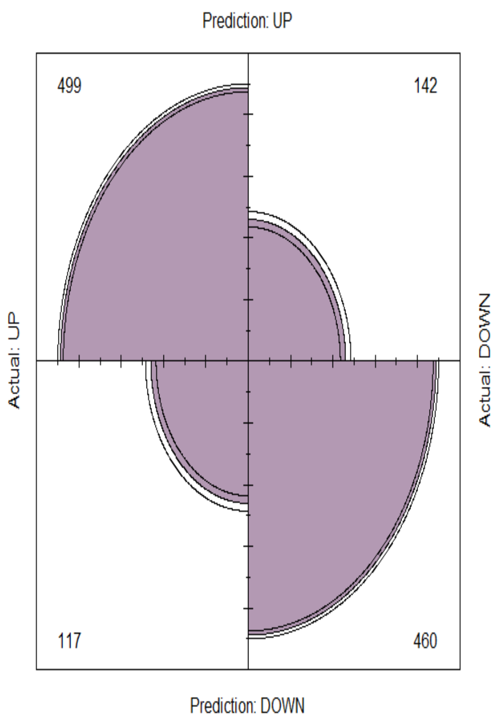
^KLSE NN Confusion Matrix (Tuned Model)



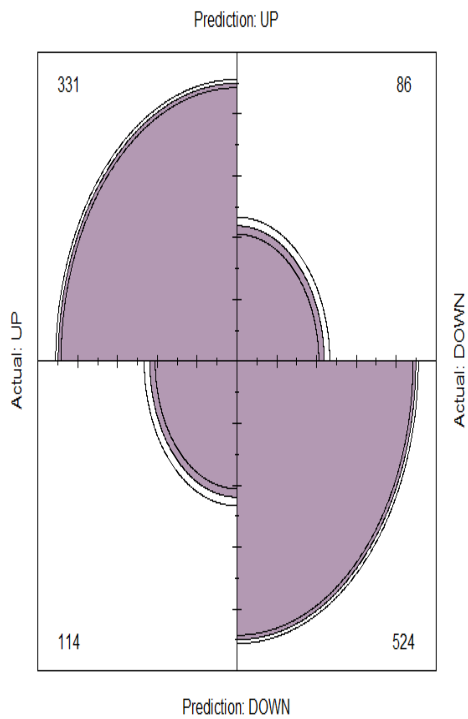
^HSI ANN Confusion Matrix (Tuned Model)



N225 NN Confusion Matrix (Tuned Model)



SSEC NN Confusion Matrix (Tuned Model)



The correct and false prediction of prices is shown by confusion matrix. The above figures shows the accurate and inaccurate prediction of prices for each index.

Table No: 4.8

Balanced Accuracy Indicators	**Indexes**					
	KSE	**BSESN**	**KLSE**	**HSE**	**N225**	**SSEC**
Accuracy	0.9175	0.8179	0.8495	0.8291	0.7874	0.8104
95% CI	(0.8871, 0.9418)	(0.7942, 0.8398)	(0.8296, 0.8679)	(0.8074, 0.8492)	(0.7633, 0.81)	(0.7854, 0.8337)
No Information Rate	0.5519	0.5543	0.5306	0.5313	0.5057	0.5782
P-Value [Acc > NIR]	<2e-16	<2e-16	< 2.2e-16	<2e-16	<2e-16	<2e-16
Kappa	0.834	0.6301	0.6996	0.6566	0.5745	0.608
McNamara's Test P-Value	0.09097	0.2385	3.32E-05	7.88E-01	0.1359	0.05624
Sensitivity	0.9017	0.8499	0.8865	0.8428	0.8101	0.7438
Specificity	0.9368	0.778	0.8168	0.8135	0.7641	0.859
Pos Pred Value	0.9462	0.8264	0.8107	0.8367	0.7785	0.7938
Neg Pred Value	0.8856	0.8065	0.8905	0.8203	0.7972	0.8214
Balanced Accuracy	**0.9193**	**0.8140**	**0.8517**	**0.8282**	**0.7871**	**0.8014**

By using the 5 variables, it is concluded that these variables are very important more than 80% accuracy in the results is due to the important variables. For KSE the accuracy has been enhanced and reached to 91.75%. Ho rejected here because P value is less than 0.05. Sensitivity, specificity, Positive predicted value and negative predicted value is very high and states that the prediction accuracy is higher in both cases of up and down values. Similarly, for BSESN the accuracy is high and kappa shows the substantial agreement. Positive predicted value and negative predicted value is more than 80% which indicate more than 80% accuracy in prediction. For KLSE the accuracy is also high and Kappa value is indicating substantial agreement. Sensitivity and negative predicted value is high shows high accuracy in down predicted values.

For HSE the accuracy has declined and reached to 80.2%. Kappa value has also decreased but still shows the substantial agreement between expected and observed values. Reject Ho because P value is less than 0.05. Next country is Japan, its symbol is N225 and its prediction accuracy has declined and reached to 78.74%. Kappa value indicate the moderate agreement because the value is less than 60%. For SSEC the accuracy is good and kappa value shows substantial agreement. P value is less than 0.05 shows significant results. Negative predicted value is more than 80% shows the high accuracy in negative predicted value. Now the study examine the comparative analysis between these models for all indexes to find out the better in each index.

4.1 Comparative Analysis between ANN and SVM:

Now I am going to compare both models to find out the better model which is giving more accuracy after using important variables. The following table shows the comparison between ANN and SVM and determine that which the better model in each country is.

Table No: 4.9

Balanced Accuracy						
	KSE	BSESN	KLSE	HSE	N225	SSEC
SVM	0.8865	0.886	0.8471	0.818	0.8476	0.8476
ANN	0.9193	0.814	0.8517	0.8282	0.7871	0.8014

The result concluded that both the models have ability to predict the stock market with high accuracy. By comparing both the model it is extracted that for some countries ANN is better than SVM and similarly for some countries SVM is better than ANN. For KSE the balanced accuracy is 88.65% by using SVM and 91.93% by using ANN, so the result concluded that for KSE ANN is a better model than that of SVM. For BSESN balanced accuracy is 88.6% by suing SVM and 81.4% by using ANN, so the results stated that SVM better than ANN in case of BSESN. Similarly, the balanced accuracy of KLSE by using SVM is 84.71% and 85.17% by suing ANN, so result concluded that ANN is better in case of KLSE.

For HSE balanced accuracy is 81.8% by using SVM and by using ANN it is 82.82%, in this case ANN is better than that of SVM. For N225 the balanced accuracy by using SVM is 84.76% and by using ANN the balanced accuracy is 78.71%, so result concluded that SVM is better than ANN as SVM has high accuracy than ANN. For China, it's symbol is SSEC the balanced accuracy is 84.76% by using SVM and by using ANN the balanced accuracy is 80.14%, so the results states that for SSEC SVM is better than ANN. The results extracted by both models have high accuracy but in case of KSE, KLSE and HIS, ANN has better prediction accuracy than that of SVM. But in other three countries BSESN, N225 and SSEC, SVM has ability to predict with more accuracy than ANN.

Chapter No: 05

Conclusion and Recommendations:

Stock market daily price prediction is very difficult for investors and any model which help investor to predict the stock prices accurately will leads him towards huge profit. In this study two methods of machine learning have used to predict the direction of stock market indexes of six Asian countries. First I have trained the models on 80% of the dataset and identified the important features for each index. Later both techniques tested the dataset by using important variables and analyzed the prediction accuracy. The results extracted that both machine learning methods ANN and SVM have ability to predict the stock market. By using the important features, accuracy has been enhanced for all indexes and investor can earn huge profit by using these techniques.

Machine learning has been around for a number of decades, but as its use continues to accelerate, policymakers must confront the broad challenges it poses. Machine learning is now demonstrably relevant to the realm not just of technology or science, but also of wider public policy. Algorithms are there that can read our medical records, decide if we qualify for a loan, identify our political beliefs, and recommend a prison sentence. The one can make decisions with the help of machine learning methods on the basis of historical data that how one can get loan, to decide what is the best place to transfer the patient after surgery and also recommend a prison sentence. The machine learning technology community has already started taking many of the concerns outlined in this section seriously but much more remains to be done.

However, for future studies the prediction accuracy can further increased by two ways. The first is to use more variables because may be I did not use some variables which are more important than selected variables, it will affect the prediction accuracy. Second, use other machine learning methods like Decision tree, genetic algorithm it may enhance the prediction accuracy.

References:

Addai, S. (2016). *Financial forecasting using machine learning* (Doctoral dissertation, University of Cape Town).

Adebiyi, A. A., Adewumi, A. O., & Ayo, C. K. (2014). Comparison of ARIMA and artificial neural networks models for stock price prediction. *Journal of Applied Mathematics*, *2014*.

Adebiyi, A. A., Ayo, C. K., Adebiyi, M. O., & Otokiti, S. O. (2012). An improved stock price prediction using hybrid market indicators. *African Journal of Computing & ICT*, *5*(5), 124-135.

Araújo, R. D. A., Oliveira, A. L., & Meira, S. (2015). A hybrid model for high-frequency stock market forecasting. *Expert Systems with Applications*, *42*(8), 4081-4096.

Brownlees, C. T., & Gallo, G. M. (2006). Financial econometric analysis at ultra-high frequency: Data handling concerns. *Computational Statistics & Data Analysis*, *51*(4), 2232-2245.

Cao, R., Liang, X., & Ni, Z. (2012, December). Stock price forecasting with support vector machines based on web financial information sentiment analysis. In *International Conference on Advanced Data Mining and Applications* (pp. 527-538). Springer, Berlin, Heidelberg.

Dash, S., & Patra, B. (2016). Genetic Diagnosis of Cancer by Evolutionary Fuzzy-Rough based Neural-Network Ensemble. *International Journal of Knowledge Discovery in Bioinformatics (IJKDB)*, *6*(1), 1-16.

Devadoss, A. V., & Ligori, T. A. A. (2013). Forecasting of stock prices using multi layer perceptron. *International journal of computing algorithm*, *2*, 440-449.

Grachev, O. Y. (2017). Application of Time Series Models (ARIMA, GARCH, and ARMA-GARCH) for Stock Market Forecasting.

Grigoryan, H. (2015). Stock market prediction using artificial neural networks. Case Study of TAL1T, Nasdaq OMX Baltic Stock. *Database Systems Journal*, *6*(2), 14-23.

Grigoryan, H. (2017, May). Stock Market Trend Prediction Using Support Vector Machines and Variable Selection Methods. In *2017 International Conference on Applied Mathematics, Modelling and Statistics Application (AMMSA 2017)*. Atlantis Press.

Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2018). Stock price prediction using support vector regression on daily and up to the minute prices. *The Journal of Finance and Data Science*, *4*(3), 183-201.

Huang, W., Nakamori, Y., & Wang, S. Y. (2005). Forecasting stock market movement direction with support vector machine. *Computers & operations research*, *32*(10), 2513-2522.

Jabbarzadeh, A., Shavvalpour, S., Khanjarpanah, H., & Dourvash, D. (2016). A multiple-criteria approach for forecasting stock price direction: nonlinear probability models with application in S&P 500 Index. *International Journal of Applied Engineering Research*, *11*(6), 3870-3878.

Kara, Y., Boyacioglu, M.A., & Baykan, Ö.K. (2011). Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange. Expert Syst. Appl., 38, 5311-5319.

Khan, S. H. (2009). Determinants of share price movements in Bangladesh: Dividends and retained earnings.

Kim, T., & Kim, H. Y. (2019). Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data. *PloS one*, *14*(2), e0212320.

Lawrence, R. (1997). Using neural networks to forecast stock market prices. *University of Manitoba*, *333*.

Li, C., Yang, B., & Li, M. (2017). Forecasting Analysis of Shanghai Stock Index Based on ARIMA Model. In *MATEC Web of Conferences* (Vol. 100, p. 02029). EDP Sciences.

Li, C., Yang, B., & Li, M. (2017). Forecasting Analysis of Shanghai Stock Index Based on ARIMA Model. In *MATEC Web of Conferences* (Vol. 100, p. 02029). EDP Sciences.

Maciel, L. S., & Ballini, R. (2010). Neural networks applied to stock market forecasting: an empirical analysis. *Journal of the Brazilian Neural Network Society*, *8*(1), 3-22.

Madge, S., & Bhatt, S. (2015). Predicting stock price direction using support vector machines. *Independent work report spring*.

Manahov, V., Hudson, R., & Gebka, B. (2014). Does high frequency trading affect technical analysis and market efficiency? And if so, how?. *Journal of International Financial Markets, Institutions and Money*, *28*, 131-157.

Meesad, P., & Rasel, R. I. (2013, May). Predicting stock market price using support vector regression. In *2013 International Conference on Informatics, Electronics and Vision (ICIEV)* (pp. 1-6). IEEE.

Moghaddam, A. H., Moghaddam, M. H., & Esfandyari, M. (2016). Stock market index prediction using artificial neural network. *Journal of Economics, Finance and Administrative Science*, *21*(41), 89-93.

Nayak, J., Naik, B., & Behera, H. (2015). A comprehensive survey on support vector machine in data mining tasks: applications & challenges. *International Journal of Database Theory and Application*, *8*(1), 169-186.

Nayak, J., Naik, B., & Behera, H. (2015). A comprehensive survey on support vector machine in data mining tasks: applications & challenges. *International Journal of Database Theory and Application*, *8*(1), 169-186.

Nayak, R. K., Mishra, D., & Rath, A. K. (2015). A Naïve SVM-KNN based stock market trend reversal analysis for Indian benchmark indices. *Applied Soft Computing*, *35*, 670-680.

Paik, P., & Kumari, M. B. (2017). Stock Market Prediction Using ANN, SVM, ELM: A Review. *Internat. J. Emerging Trends Technol. Comput. Sci.*, *6*(3), 88-94.

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

Qiu, M., & Song, Y. (2016). Predicting the direction of stock market index movement using an optimized artificial neural network model. *PloS one*, *11*(5), e0155133

Qiu, M., Song, Y., & Akagi, F. (2016). Application of artificial neural network for the prediction of stock market returns: The case of the Japanese stock market. *Chaos, Solitons & Fractals*, *85*, 1-7.

Qu, H., & Zhang, Y. (2016). A New Kernel of Support Vector Regression for Forecasting High-Frequency Stock Returns. *Mathematical Problems in Engineering*, *2016*.

Rustam, Z., & Kintandani, P. (2019). Application of Support Vector Regression in Indonesian Stock Price Prediction with Feature Selection Using Particle Swarm Optimisation. *Modelling and Simulation in Engineering*, *2019*.

Selvamuthu, D., Kumar, V., & Mishra, A. (2019). Indian stock market prediction using artificial neural networks on tick data. *Financial Innovation*, *5*(1), 16.

Song, Y. G., Zhou, Y. L., & Han, R. J. (2018). Neural networks for stock price prediction. *arXiv preprint arXiv:1805.11317*.

Sridharan, S. A. (2013). *Fundamental Analysis and Equity Volatility* (Doctoral dissertation, Stanford University).

Széliga, M. I., Verdes, P. F., Granitto, P. M., & Ceccatto, H. A. (2003). Artificial neural network learning of nonstationary behavior in time series. *International journal of neural systems*, *13*(02), 103-109.

Vasudevan, R. D., & Vetrivel, S. C. (2016). Forecasting stock market volatility using GARCH models: evidence from the Indian stock market. *Asian Journal of Research in Social Sciences and Humanities*, *6*(8), 1565-1574.

Wafi, A. S., Hassan, H., & Mabrouk, A. (2015). Fundamental Analysis Vs Technical Analysis in The Egyptian Stock Exchange–Empirical Study. *International Journal of Business and Management Study–IJBMS*, *2*(2).

Wang, K., Liang, C., Liu, J., Xiao, H., Huang, S., Xu, J., & Li, F. (2014). Prediction of piRNAs using transposon interaction and a support vector machine. *BMC bioinformatics*, *15*(1), 419.

Xiao, Y., Xiao, J., Lu, F., & Wang, S. (2014). Ensemble ANNs-PSO-GA approach for day-ahead stock E-exchange prices forecasting. *International Journal of Computational Intelligence Systems*, *7*(2), 272-290.

Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks:: The state of the art. *International journal of forecasting*, *14*(1), 35-62.