Ilkogretim Online - Elementary Education Online, 2020; Vol 19 (Issue 4): pp. 7981-8002
http://ilkogretim-online.org
doi: 10.17051/ilkonline.2020.04.765193

Comparative Analysis Of Alternative Time Series Models In Forecasting Karachi Stock Exchange

Muhammad Aamir[1,*], Hazrat Ali[1], Umair Khalil[1] and Muhammad Bilal[1,2]

[1]Department of Statistics, Abdul Wali Khan University, Mardan Pakistan
[2]Department of Mathematical Sciences, BUITEMS Quetta, Pakistan

Email: aamirkhan@awkum.edu.pk; hazrat.ali.gar@gmail.com; umairkhalil@awkum.edu.pk , muhammad.bilal3@buitms.edu.pk
[*]Corresponding author aamirkhan@awkum.edu.pk

ABSTRACT

The stability of prices is an important indicator of overall economic performance and is one of the main objectives of monetary policy. Pakistan's economy has a long history of unstable macroeconomic performance, especially the persistence of high inflation rates, which lasted for almost three decades. During this long period, many stability programs mostly backed by the IMF could not be implemented thoroughly and failed to achieve the desired outcomes and price stability. Researchers have identified factors including firm size, past stock performance, value, and growth as some of the factors affecting the stock exchange. While the current study uses only the Karachi Stock Exchange 100 (KSE100) index as a proxy and performs a time-series analysis to identify the best forecasting model which can help investors and government agencies to make up-to-date decisions. Recently, forecasting future observations based on time series data has received great attention in many fields of research. Several techniques have been developed to address this issue to predict the future behaviour of a particular phenomenon. In this study, two methodologies for forecasting the KSE100 index are used. The first is the linear time series modelling consisting of NAÏVE and Box-Jenkins methodology, while the second is the non-linear methods consisting of Generalized Autoregressive Conditional Heteroskedasticity (GARCH), Artificial Neural Network (ANN), and Support Vector Machine (SVM). These two approaches are used to obtain the static and dynamic forecasts of the KSE100 index daily data, and the accuracies are compared by using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Directional Statistics (DS) and Diebold-Marino (DM) Test. The results indicated that the ANN is the most effective machine learning approach for improving the forecasting accuracy of the KSE100 index. Thus, from this study, the recommended model for forecasting the KSE100 index data is ANN which can handle the non-linearity and non-stationarity. Hence, the ANN model is recommended and could be used for forecasting the KSE100 index data.

Key Words: Econometrics, Forecasting, KSE100 Index, Machine Learning, Time Series

1. INTRODUCTION

1.1 Background of the study

The stability of prices, as an important indicator of overall economic performance, is one of the main objectives of monetary policy. Pakistan's economy has a long history of unstable macroeconomic performance, especially the persistence of high inflation rates, which lasted for almost three decades. During this long period, many stability programs, mostly backed by the IMF could not be implemented thoroughly and failed to achieve the desired outcome, economic and price stability. Researchers have identified factors including firm size, past stock performance, value, and growth as some of the factors affecting the stock exchange. Fundamental analysis examines the company's stocks prices movements by their historical financial and accounting data. It includes an analysis of the company's earnings, expenses, profit, management experience, assets and liabilities, and industry dynamics. Such analysis helps investors to make investment strategies for getting surplus returns [1-4]. Five readily fundamental signals are considered more efficient by the financial analyst in predicting stock returns i.e., current ratio, leverage ratio, returns on assets, earnings per share, and price-earnings ratio. While the current study uses only the Karachi stock exchange (KSE) index and performs a time series analysis to identify the best forecasting model which can help investors and government agencies to make up-to-date decisions. On the other hand, due to the complex nature of the stock index, this is an open research area and for forecasting the stock index, most of the researcher has used different techniques [5-7]. Recently, forecasting future observations based on time series data has received great attention in many fields of research. Several techniques have been developed to address this issue to predict the future behavior of a particular phenomenon. The traditional approach based on Box and Jenkins' Autoregressive Integrated Moving Averages (ARIMA) and Generalized Autoregressive Conditional Heteroscedastic (GARCH) models is commonly used because the resulting models are easy to understand and interpretable [8, 9].

It has been widely accepted that predicting stock returns is not a simple task, since many market factors are involved, and their structural relationships are not perfectly linear. Recently, a promising data mining technique in machine learning has been proposed to uncover the predictive relationships of numerous financial and economic variables. Inspired by the fact that the determinant between these variables and their interrelationships over overstock returns changes over time, we explore this issue further by using data mining to uncover the recent relevant variables with the greatest predictive ability. In the near past, some advanced techniques are incorporated consisting of Support Vector Machines (SVM) and Artificial Neural Networks (ANN) that can handle the nonlinearity in time series and also overcome the problem of non-stationarity. The use of SVM and ANN is increasing rapidly because of their ability to form complex nonlinear systems for forecasting based on in-hand sample data [5]. Thus, because of the above reasons forecasting KSE is not a simple task. In the literature review, many models are found on KSE forecasting including the most popular Box-Jenkins methods, for example [10, 11], GARCH type models are included as in [12-14]. To the best of our knowledge, the support vector machine has not been yet applied for volatility or risk forecasting of the KSE100 index. This motivates us to fill this gap and contribute to the literature and also do a comparison among the traditional approaches and machine learning approaches.

In recent times, the hopes of financial researchers have been ignited because of the successful use of non-linear methods in other areas of research. The best advice one can hope to get from such a prediction is that it is best not to be in a particular market at all. Consequently, an effort has been made in this study to better understand the use of non-linear models KSE100 index, which is the biggest and most important stock exchange in Pakistan. In fact, in the year 2002, the KSE100 index was declared the world's best stock market by Bloomberg, and it was repeatedly ranked first by "Business Weekly" for more than six years in a row as the best-performing market globally.

In this study, two methodologies for forecasting the KSE100 index are used. The first is the linear time series modelling consisting of NAÏVE and Box-Jenkins methodology while the second is the non-linear methods consisting of GARCH, ANN, and SVM. These two approaches will be used to forecast the KSE100 index of daily data. The main purpose of this study is to evaluate the performance of ANN and SVM with alternative ARIMA, NAIVE, and GARCH univariate time series models to forecast the daily KSE100 index. Finally, a comparison of the forecasting accuracy of these models will be computed by using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), Directional statistics (DS), and Diebold-Marino (DM) Test.

1.2 An Overview of KSE

KSE is the oldest and largest stock exchange in Pakistan located in Karachi. It was founded soon after the creation of Pakistan dated 18 September 1947. KSE is Pakistan's largest tax-payer and contributed more than Rs. 4 billion during the fiscal year spanning 01 July 2006 to 30 June 2007 towards the national exchequer for more details visit (www.psx.com.pk). KSE was also listed among the 10 best stock markets in the world in 2015. According to Bloomberg, the Pakistani benchmark stock market index is the third-best performer in the world since 2009. In June 2015, Khaleej Times reported that since 2009, Pakistani equities delivered 26 percent a year for US dollar investors, making Karachi the best-performing stock exchange in the world. As of 10 July 2015, total market capitalization reached Rs. 7.33 trillion (US$72 billion approximately). Initially, the KSE began with a share index of 50 as the market grew a representative index is needed. Thus, the KSE100 index was introduced on 01 November 1991 with base points of 1000 and to date the most generally accepted measure of the exchange. The KSE100 index is composed of the 100 largest capitalization companies and there are different sectors but in the present study, only the KSE100 index is considered.

The Pakistan stock exchange (PSX) was established on 11 January 2016 after the merger of the Karachi, Lahore, and Islamabad stock exchanges. PSX's origins were laid with the establishment of the KSE in 1947, the Lahore Stock Exchange (LSE) in 1970, and the Islamabad Stock Exchange (ISE) in 1992. As of 26 July 2020, there were about 540 companies listed in PSX and the total market capitalization was Rs. 7.07031 trillion (US$43 billion). Thus, in this study, the KSE100 index is considered because it contributes more than 85% of shares in total [15].

1.3 KSE100 Index Data Used by Different Studies

KSE100 index data sets were used for forecasting the KSE100 index in the past. All time series were comprised of daily data. The KSE100 index data set is well-known and used as a benchmark for PSX with the reasons for its availability and easily accessible. In the previous studies presented above, different researchers used the same KSE100 index data with different classifications. Table 1 illustrated KSE100 index data sets with the time used in the previous studies.

Table 1: Description of data sets used by different researchers

Author	Period	Horizon	Data	Methods Employed
Rashid, et al. [16]	2000 to 2019	Daily	KSE100 index	GARCH
Muhammad and Ali [1]	2007 to 2017	Daily	KSE100 index	ARCH, GARCH and ARIMA
Ahmad [17]	Dec 2013 to Sep 2017	Daily	KSE100 index	GARCH and ARIMA
Zafar and Hassan [18]	2000 to 2014	Daily	KSE100 index	GARCH
Iqbal [19]	3 Jan 2005 to 30 Dec 2011	Daily	KSE100 index	ARIMA and GARCH
Khalid, et al. [20]	Jan 2009 to Dec 2012	Daily	KSE100 index	EMD with ARIMA
Haque and Naeem [21]	1998 to 2008	Daily	KSE100 index	GARCH
Iqbal, et al. [22]	January 1992 to June 2008	Daily	KSE100 index	ARIMA and GARCH
Fatima and Hussain [23]	1 Jan 2000 to 18 Oct 2002	Daily	KSE100 index	ARIMA and GARCH

The remainder of the study is organized as section 2 comprised the methodologies used in the study. Section 3 is composed of the analysis of the work on KSE100 index data which confirmed its worth. In the last section-4 concluded the work performed in this study.

2. METHODS USED IN THIS STUDY

2.1 ARIMA Model

The models which require that the time series should be stationary are the ARMA models. Stationarity implies that the mean, variance, and autocorrelation structure of the time series remains constant over time since the data is smooth having no seasonality and trend factors. The successive difference has been taken whenever the given time-series are not stationary preferably at most two consecutive differences. The Interdependencies in time series (Z_t) is measured by the AR terms while the dependence on preceding error terms is measured by MA terms [24, 25]. An ARMA model of order (p, q) for a univariate time series has the following form:

$$Z_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i Z_{t-i} + \sum_{j=1}^{q} \gamma_j\, \varepsilon_{t-j} + \varepsilon_t$$

Where p is the number of AR terms, q is the number of lagged error terms, $\alpha_i, i = 1,2, \dots, p$ is the coefficient of AR terms, and α_0 representing the constant term while $\gamma_j, j = 1,2, \dots, q$ is the coefficient of MA terms. $\varepsilon_t \sim N(0, \delta^2)$ is a white noise process having mean zero, variance δ^2, no autocorrelation, independently and identically distributed random variables. Using the Box-Jenkins technique, first of all, it can be assured that the given time series is stationary. Stationarity can be achieved by taking the successive differencing of the time series (usually at most two times). The ACF and PACF plots are used for choosing the proper order of the polynomials and contrasted with the theoretical bounds [26]. Some theoretical approaches are also used to obtain the appropriate ARIMA models like AIC [27] and BIC [28].

2.2 NAIVE Model

In the NAÏVE model, the current value at time t is used as a forecast for the next lead time t+1, which is as follows:

$\hat{\mathbf{y}}_{\mathbf{t+1}} = \mathbf{y}_{\mathbf{t}}.$

2.3 GARCH Model

The GARCH model is the variant of the ARCH model introduced by Bollerslev [29]. The GARCH model is used to capture the high-frequency effect on the previous or lagged squared error terms and is more important to detect time-series volatility. The effect of the GARCH is common in most cases where a lower order needs to capture the dynamics. The GARCH (p, q) model can be written as:

$$\gamma_t = \sigma_t \varepsilon_t$$
$$\sigma_t^2 = \phi + \sum_{i=1}^{q} \lambda_j \varepsilon_{t-j}^2 + \sum_{j=1}^{p} \varphi_j \sigma_{t-j}^2$$

Where p and q are the orders of the GARCH and ARCH terms and ε_t is a series of independently and identically distributed (iid) random variables with a zero mean and unit variance, $\phi > 0, \lambda_j \geq 0, \varphi_j \geq 0$, and $\sum_{i=1}^{max\,(q,p)} (\lambda_j + \varphi_j) < 1$. It is also noted that if $p = 0$, then the GARCH model is equivalent to the ARCH model. The AIC information criterion is used to determine the orders of p and q.

2.4 Support Vector Machine

SVM is an important machine learning (ML) method first introduced by [30], which can be applied to solve classification and regression problems. Initially, the SVM intended to solve classification problems, such as face identification, text classification, optical character recognition, etc. But recently the SVM is commonly used in regression and forecasting time series problems, see for more details [11, 13]. The fundamental principle of SVMs is to discover a hyperplane that separates a set of labelled data points into two distinct classes. SVM has been used for various time series models. On the other hand, they are the data points in which rounding errors are greater than are equal to the so-called SVM tube size. If the training set's sample observations are not completely separated, then alternatively a non-linear margin or boundary has to be made. To attain the nonlinear margin, the original input space is plotted into a higher dimensional space called the feature space. The features space is then searched for a hyperplane that can isolate the occurrences in the same feature space. If the classes are not separated linearly then we use the kernel function. For example, authors in [31] used SVM classifiers for analyzing the time series dengue data and genetic algorithm (GA) to determine the time-lags and subclass of climatic factors as key factors influencing the transmission of dengue. Figure 1 shows the linearly separable between the two classes [32]. Different types of SVM are used for different problems.

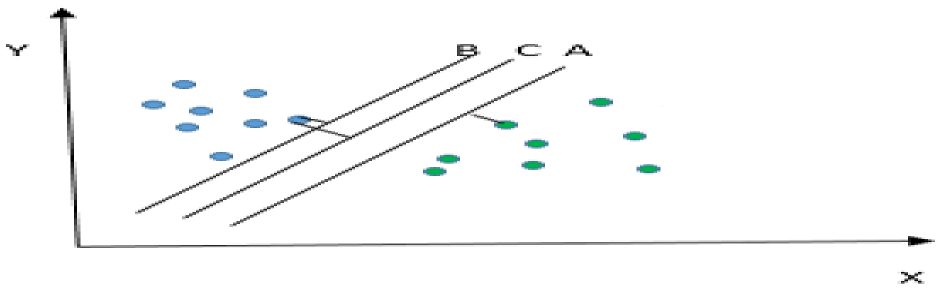

Figure 1: Linearly separable hyperplane between two classes

In Figure 1, a linearly separable hyperplane is a line that divides the data point into two classes (blue & green), written as

$$y = \alpha * x + \beta$$

$$\alpha * x + \beta - y = 0$$

The above equation in vector form hyperplane can be written as, let $X = (x, y)$ and $W = (\alpha, -1)$.

$$W.X + \beta = 0$$

Because of good and reliable performance in real-world classification problems, the SVM is extended to regression and time-series forecasting [33]. In SVM literature, when it is used for classification problems called Support Vector Classification and when used for regression problems called Support Vector Regression (SVR).

2.5 Artificial Neural Network

ANN has many distinguishing characteristics over the ARIMA models and is a successful alternative to linear models to handle the nonlinearity factors when forecasting the time series. One of the best characteristics of the ANN is the universal approximation of any non-linear function up to the best degree of accuracy as desired [34, 35]. ANN can have single or multiple layers. The proposed ANN can handle the nonlinearity in KSE100 index data and enhances forecasting accuracy. The basic type of neural network is Multi-layer Perceptron (MLP). MLP is a feed-forward neural network (FFNN) with one or more hidden layers between the input layer and the output layer. Feedforward means that the data is produced in one direction. The most commonly used neural network is the FFNN with a single hidden layer having one output node for forecasting the time series applications [36, 37].

In this study, the FFNN is considered to contain two layers in addition to the input layer. The first is the hidden layer and the second is the output layer. In the input layer, there are K inputs. These inputs are oftentimes weighted randomly. For each hidden layer, there are also N neurons. Practically, the best number of neurons in the hidden is equalled to $K \times 2 + 1$ [38, 39]. Each input variable X is separately weighted by a random weight ϕ automatically. The weighted inputs for K variables and N neurons are summed with the biased value E to formulate the input of the transfer function. The input variable (IPV) of the transfer function (f) can be formulated as follows:

$$IPV = \sum_{i=1}^{N} \sum_{j=1}^{K} \omega_{i,j} X_{j,t} + E$$

where, $j = 1,2, \ldots, K$ and $i = 1,2, \ldots, N$. The most commonly used types of transfer functions for hidden and output layers are commonly tan-sigmoid which generates outputs between -1 to +1, log-sigmoid which generate outputs between 0 and +1, and linear transfer functions which generate outputs between -1 to +1. Selecting a transfer function for the hidden and output layers is important for obtaining good results and depends on the nature of the data. Figure 2 displays the different types of transfer functions that can be specified using the ANN toolbox in MATLAB.

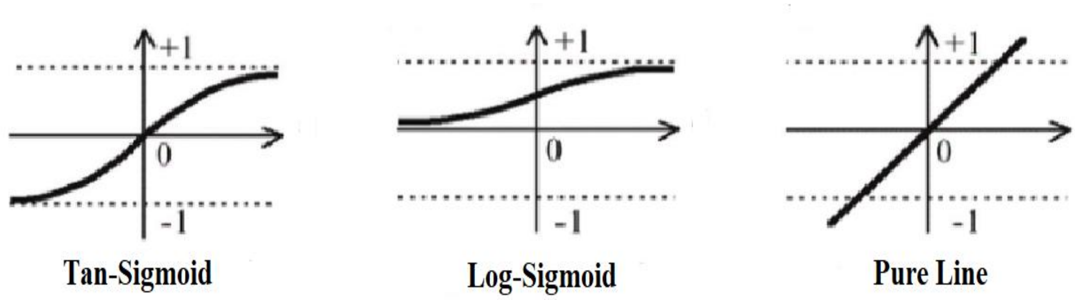

Figure 2: Transfer function types of ANN

where $V = X$ is an input vector for the hidden layer, $T = X$ is an input vector for an output layer, $T = \hat{X} = f(N)$ as an output vector for the first hidden layer, $T = \hat{X} = f(A)$ as an output vector for an output layer, $M = IPV$ as a summation function for the first hidden layer, $A = IPV$ as a summation function for an output layer, H represents the number of neurons in the first hidden layer and the number of inputs for an output layer, and O represents the number of the neurons in an output layer and equals to 1 oftentimes. Using the ANN toolbox, the FFNN that includes one hidden layer and one output layer with a single output can be represented in Figure 3 [40, 41].

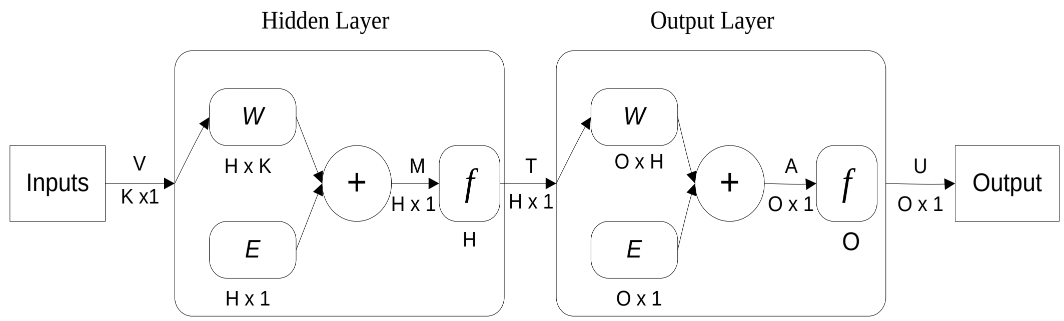

Figure 3: The general structure of ANN

ANN training process is employed to adjust the random weights and bias values in each layer to enhance the forecasting accuracy and obtain the desired output. The best training functions for backpropagation algorithms are Levenberg-Marquardt (LM) and Bayesian regularization (BR) [42, 43]. LM training function is oftentimes the fastest training algorithm and is recommended to minimize the MSE of a neural network [44-48]. BR training function determines the weights and bias values according to LM optimization. It minimizes the forecasting error by combining the classical error's sum of squares with two further terms called regularization [49-53].

2.6 Forecasting Accuracy Measures

To check the proposed model's competency with the other models used in this study would be measured by using five different evaluation criteria. The first is the MAE, the second is the RMSE and the third is MAPE typically these methods are defined as:

$$MAE = \frac{1}{f}\sum_{t=1}^{f}\left|\hat{X}_t - X_t\right|$$

$$RMSE = \sqrt{\frac{1}{f}\sum_{t=1}^{f}\left(\hat{X}_t - X_t\right)^2}$$

$$MAPE = \frac{1}{f}\sum_{t=1}^{f}\left|\frac{\hat{X}_t - X_t}{X_t}\right| \times 100$$

Where f signifies the total number of forecasts, $\hat{X}_t$ represents the forecasted value and X_t represents the original value at time t. The above-stated forecasting accuracy methods are common and measure the level of forecasts. The directional forecasts are suitable for Government agencies and investors because they are interested in the market trend. Directional statistic (DS) used for this purpose provides the directional forecasting accuracy of the competing models [54, 55], and is defined as follows:

$$DS = \frac{1}{f}\sum_{t=1}^{f} F_t \times 100$$

Where $F_t = 1$, if $(X_{t+1} - X_t) \times \left(\hat{X}_{t+1} - X_t\right) \geq 0$, otherwise $F_t = 0$ and f represent the total number of forecasts. The minimum values of MAE, RMSE, and MAPE, the greater the forecasting accuracy is. While the greater values of DS produce more accurate directional forecasts. In this study, the Diebold–Mariano (DM) test is also used which compares the forecasting errors of the two competing models.

$DM = \frac{\bar{E}}{\sqrt{Var(\bar{E})}}$,

where $\bar{E} = \frac{1}{f}\sum_{t=1}^{f} E_t$, $E_t = (X_t - f_{1,t})^2 - (X_t - f_{2,t})^2$, $Var(\bar{E}) = \frac{1}{f}(r_0 + 2\sum_{k=1}^{f} r_k)$ and $r_k = Cov(E_t, E_{t-k})$. $f_{1,t}$ signifies the forecasted values taken from the first model, while $f_{2,t}$ taken from the second model. The assumption required for the DM test is that the total number of forecasts from both models would be equal.

3 Data Analysis and Discussions

3.1 Data Used in the Study

The data used in this study was obtained from the publicly accessible Pakistan Stock Exchange (PSX) website (https://www.psx.com.pk/). The opening KSE100 index historical data is obtained from 1st January 2016 to 30th July 2020. The data is available in the online repository (https://dps.psx.com.pk/historical). This repository is for the KSE100 index visual dashboard operated by PSX limited. The data have a total number of 1137 intraday values. The descriptive statistics of the data are presented in Table 2.

Table 2: Descriptive statistics of KSE100 index data

Total No. of Observations	1137	Mean	39834
Minimum value	27247	Median	40238
Maximum value	52876	Standard Deviation	5297.2
1st Quartile	35691	Skewness	0.0962
3rd Quartile	42942	Kurtosis	2.4381

From Table 2, it can be seen that the mean of the series is 39434 with a standard deviation of 5297.2. The skewness and kurtosis of the series are 0.0962 and 2.4381, respectively which is apart from normality. The data is positively skewed (skewness > 0) and platy-kurtic (kurtosis < 3). To test the data normality Jarque-Bera test is performed and presented in Table 3.

Table 3: Normality test

Test	**Statistic**	**P-value**	**Decision**
Jarque-Bera	16.7107	0.0002	Data is not normally distributed

From Table 3, it is observed that the p-value of the Jarque-Bera test is less than 0.01, which showed that the null hypothesis of normality is not accepted. To check the correctness, robustness, and generalizability of the suggested procedure, the KSE100 index data is divided into training and testing sets of observations. The first 80% of the total observations of the time series were used as a training set whereas the rest 20% was used as the testing set to evaluate the model's performance [56, 57]. The KSE100 index time series contains a total of 1137 observations spanning (1 January 2016, to 30 July 2020), the first 910 observations from (1 January 2016 to 1 September 2019) belong to the training series, and the rest 227 from (3 September 2019 to 30 July 2020) part of the testing series. The original training and testing observations are shown in Figure 4 and the testing set of observations is more focused in the second plot of Figure 4.

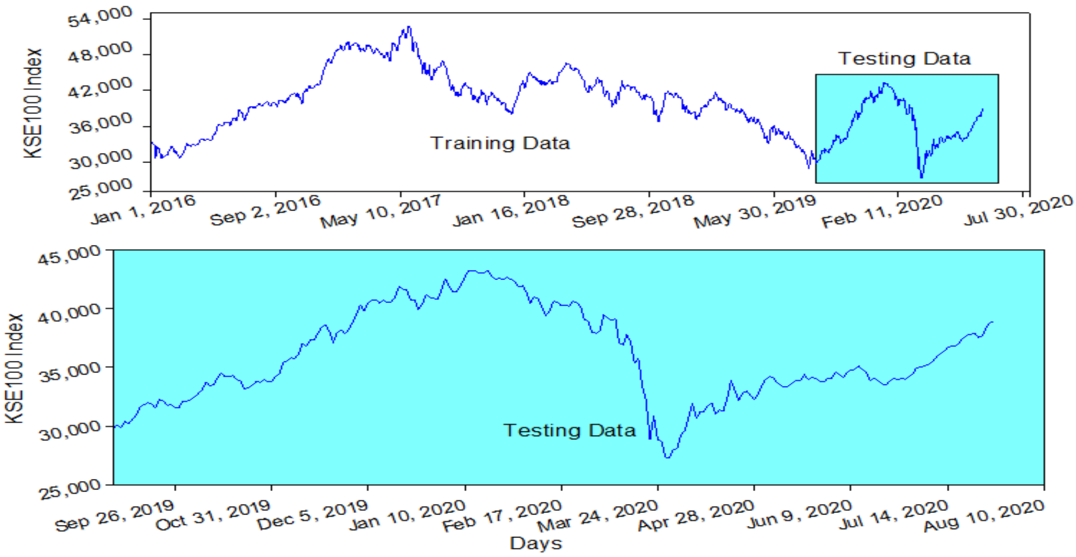

Figure 4: Plot of the KSE100 index original training and testing data

3.2 ARIMA Modelling Approach

In this section, the ARIMA model will be fitted for KSE100 index data for the training period and forecast for the testing period. One of the basic assumptions for the ARIMA model is stationarity which implies that the data should be stationary before applying any statistical test and estimation of parameters. To overcome the stationarity problem, the unit root ADF test is employed. The results obtained are presented in Table 4.

Table 4: ADF test and their P-values

Series	F-statistic	P-value	Decision
Original	-1.5906	0.7516	Not stationary
1st Difference	-9.1110	<0.0001	Stationary

Alternative hypothesis: The data is stationary

From Table 4, it is observed that the series is stationary, after taking the first difference of the series. After obtaining the stationarity, the best orders of AR term p and MA term q will be determined. In this study, the ACF and PACF plots are used to determine the order of the AR and MA terms. The ACF and PACF plots of the first difference series of KSE100 index data are presented in Figure 5.

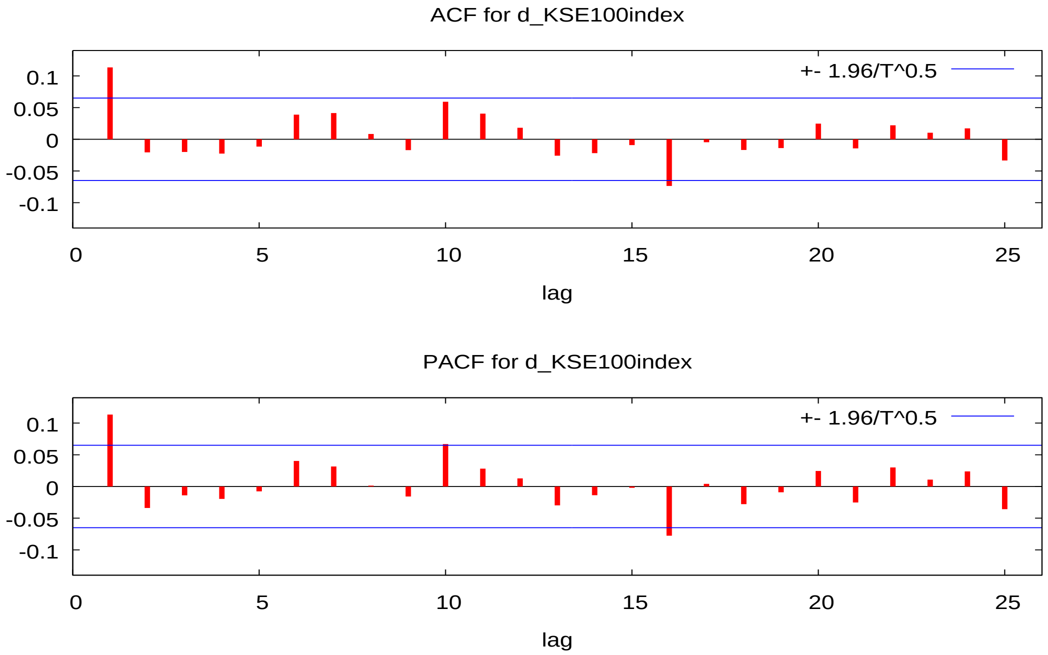

Figure 5: Plots of the ACF and PACF of the first difference of KSE100 index training data

By looking at the ACF plot of Figure 5, it is observed that there is only one significant spike at lag-1 and a second significant spike at lag-16. Therefore, the MA order to choose is 1. From the PACF plot, significant spikes are observed at lag 1, 10, and 16. Thus, a first guess would be the order of AR is 1. Therefore, from Figure 5 the first guess of the order of the ARMA model is (1,1) whereas the series was stationary after taking the first difference, so the recommended order of the ARIMA model is (1,1,1). After selecting the best ARIMA model, the next step is the estimation of parameters. The estimated ARIMA(1,1,1) model parameters are presented in Table 5.

Table 5: The estimated ARIMA model parameters and standard errors

Coefficient	**Estimate**	**Standard error**	**t-statistic**	**P-value**
AR1	0.1102	0.0332	3.320	0.0009
MA1	-0.9952	0.0036	-276.6	<0.0001

From Table 5, it is observed that the p-values of AR1 and MA1 are less than 0.01 which indicated that both coefficients are statistically significant. The next step is the diagnostic checking that the selected model is adequate and could be used for forecasting. The LB test checks the serial correlation among the fitted model residuals. The test stated that if the p-value is greater than 0.05 then the hypothesis of no serial correlation is not rejected and provided that the fitted model is adequate and can be used for forecasting [58]. The fitted ARIMA model diagnostic checking is carried out and the results are shown in Figure 6.

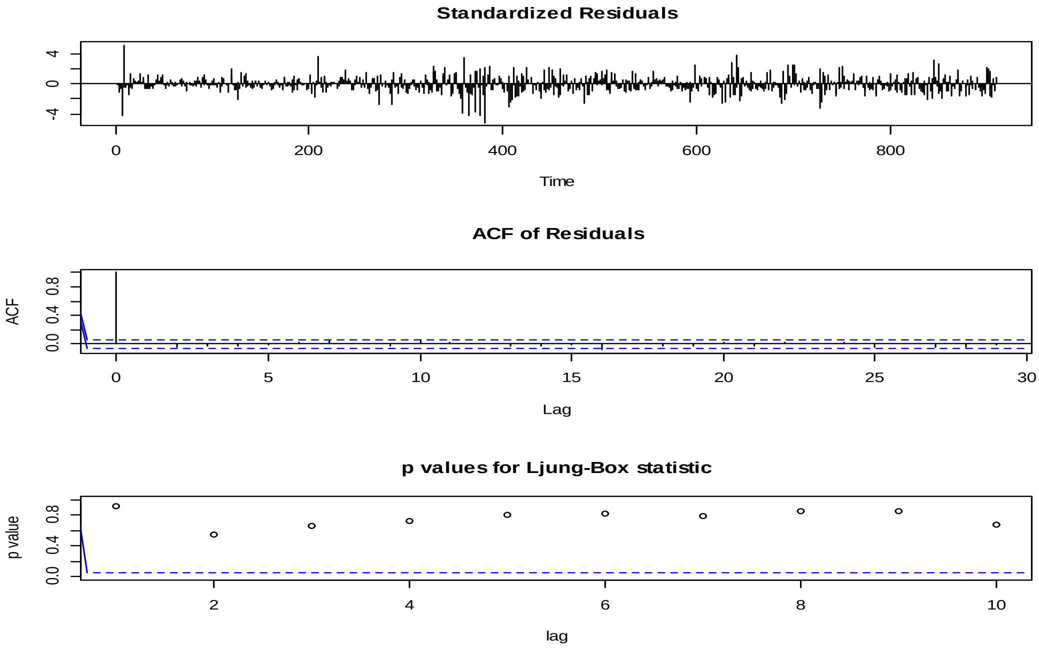

Figure 6: Diagnostic test results of the fitted ARIMA(1,1,1) model

The p-values of the LB test for all lagged values are greater than 0.05 and hence the hypothesis of no serial correlation among the fitted model residuals is statistically insignificant. This reveals that the fitted ARIMA(1,1,1) model is adequate and can be used for forecasting KSE100 index data. The last step of the Box-Jenkins methodology is forecasting. The forecast package of R-software is used for selected ARIMA model forecasting [59].

3.3 GARCH Modelling

Based on AIC, BIC, and AICc criterions the best-chosen model is GARCH(1,1). Before applying the GARCH model the data was converted to the log-returns using the relationship $Y_t = ln\,(Y_t/Y_{t-1})$ and for the final output converted again and transferred to the original units. The estimated parameters of the GARCH(1,1) model, are presented in Table 6.

Table 6: The estimated GARCH model parameters and standard errors

Coefficient	Estimate	Standard error	t-statistic	P-value
Constant	0.0005	0.0003	1.488	0.1366
$\boldsymbol{a_1}$	0.0000067	0.0000027	2.444	0.0145
$\boldsymbol{a_2}$	0.1278	0.0324	3.948	0.0001
$\boldsymbol{b_1}$	0.8215	0.0449	18.270	<0.0001

3.4 NAIVE Model

In this section, the NAÏVE model is fitted for KSE100 index data. In the NAÏVE model, the current value at time t is used as a forecast for the next observation t+1. i.e. $\hat{\mathbf{y}}_{\mathbf{t+1}} = \mathbf{y_t}$.

3.5 Artificial Neural Network Approach

The input structure of ANN is subject to the number of AR parameters of the ARIMA model. The ANN input for the KSE100 index data is 1st lag from the original series. In this study, for ANN the MATLAB Neural Network Time Series Tools (ntstool) was used. MATLAB has improved some of its toolboxes for Neural networks. Neural network toolboxes in MATLAB use many types of training algorithms, training functions, and transfer functions. The structure of ANN has been constructed by determining the following requirements.

(i). Feedforward backpropagation is used as a network type.
(ii). The input for KSE100 index data is Y_{t-1}. This input was automatically created by the MATLAB ntstool once the training input and target data were imported into the ANN structure.
(iii). Only Levenberg-Marquardt training algorithms were used as a training function in this study.
(iv). The weights were determined randomly depending on the strategies of the ANN toolboxes.
(v). The maximum number of neurons in the hidden layer is equal to $(2 \times k + 1)$ where k is the number of inputs [38, 60]. The number of hidden neurons for KSE100 index data is 3.
(vi). The nonlinearity of the KSE100 index was obligated to choose a nonlinear transfer function such as tan-sigmoid or log-sigmoid for the hidden layer to filter the nonlinearity. In this study, the Nonlinear Autoregressive (NAR) problem is selected from the ntstool.

The complete structure of ANN(1,3,1) is presented in Figure 7 for KSE100 index data. Training and testing processes only need to import the input and target variables for training and testing periods, respectively.

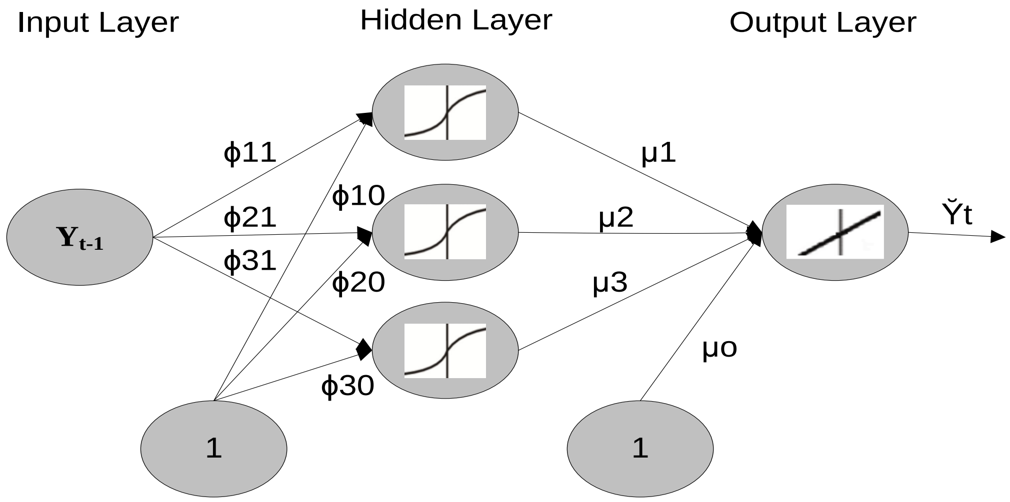

Figure 7: ANN structure for KSE100 index data

3.6 Support Vector Machine Approach

The first step is to normalize the KSE100 index data which increases the model accuracy. The following procedure is used to normalize the KSE100 index data in the (0 to 1) scalable range.

$$y_{norm} = {(y_t - y_{min})}\Big/{(y_{max} - y_{min})}$$

Where y_{norm}, represents the normalized values, y_t representing the original values, y_{min} used for the smallest and y_{max} used for the largest values of the original KSE100 index series. In this study, the input structure of SVM is subject to the number of AR parameters of the ARIMA model. To determine the input structure for the SVM the order of the AR term is used i.e., [61, 62]. For KSE100 index data, the AR(1) model is fitted. In other words, the SVM input for the KSE100 index data is 1st lag from the original series. The structure of SVM has been constructed by determining the following requirements.

(i). The KSE100 index time-series data are used spanning 01 January 2016 to 30 July 2020. To fit the SVM model the data are divided into P cases.

(ii). The KSE100 index time-series data using one input, so we take P =1 i.e., Y_{t-1}. This input was automatically created by the r environment once the training input and target data were imported into the SVM structure.

(iii). The next step after the input selection is the SVM model parameter selection. An SVM model is based on four parameters which are SVM type, SVM kernel, cost C, and gamma ε. The accuracy of the SVM model is highly dependent on these parameters. The cross-validation of a 10-fold procedure was applied for obtaining the best values of C and ε parameters. In step 1, the value of C ranging from 1 to 1000 and the value of ε ranging from 0.0001 to 0.05 was taken.

(iv). The best SVM model for KSE100 index data was selected with SVM type "nu-regression", SVM kernel "radial", cost C "719" and gamma ε "0.001". With these specifications, the SVM model produced the lowest RMSE.

In this study, one-step-ahead forecasting was determined. The training data is used to develop the SVM model.

3.7 ANALYSIS AND DISCUSSIONS

The original data was divided into 80% training and 20% testing parts [63-65]. Then, we evaluate models including NAIVE, ARIMA, GARCH, SVM, and ANN, to compare their performance using various accuracy metrics including MAE, RMSE, MAPE, DS, and DM test. These metrics provide different perspectives to assess predicting models. The first two are the absolute performance measures while the third and fourth are relative performance measures. Table 7 summarizes the NAÏVE, ARIMA, GARCH, SVM, and ANN forecasting accuracy measures for the 20% testing set of KSE100 index daily data.

Table 7: Forecasting accuracy measures of testing data for all models

Model	**MAE**	**RMSE**	**MAPE**	**DS**
NAÏVE	426.67	614.64	1.22	---
ARIMA	422.27	617.84	1.21	57.96
GARCH	424.63	613.97	1.21	56.63
SVM	307.12	393.27	0.94	84.95
ANN	**116.68**	**188.83**	**0.35**	**93.55**

In Table 7, the values of MAE for NAÏVE, ARIMA, GARCH, SVM, and ANN models are 426.67, 422.27, 424.63, 307.12, and 116.68, respectively. This reveals that ANN shows the lowest value (the best) among the other methods. Similarly, the RMSE values for NAÏVE, ARIMA, GARCH, SVM, and ANN are 614.64, 617.84, 613.97, 393.27, and 188.83, respectively, and show that ANN achieved better performance compared to the other methods. We highlighted the results for the ANN model indicating the smallest values of MAE and RMSE among all models. While the values of MAPE for NAÏVE, ARIMA, GARCH, SVM, and ANN models are 1.22, 1.21, 1.21, 0.94, and 0.35, respectively. Thus, the values of the MAPE for NAÏVE, ARIMA, and GARCH models are between 1 to 10, which reveals that these models fall in the range of good models whereas SVM and ANN models values are less than 1, which indicates that SVM and ANN models fall in the range of perfect models [25]. Since the value of MAPE for the ANN model is lower, therefore, the best-suited model for KSE100 index data is the ANN. Moreover, next is the DS for ARIMA, GARCH, SVM, and ANN models are 57.96, 56.63, 84.95, and 93.55, respectively. Directional forecasts are more important because investors and policymakers always look for the trend in which the market prices are going up or down. Regarding the directional forecasting, the DS values of the models ARIMA and GARCH shown in Table 7 were 57.96 and 56.63 for KSE100 index data which were very close to the random guess. Thus, for ARIMA and GARCH models, it was hard to attain directional forecasting at a satisfactory level, and the anticipated cause is the complex nature of the KSE100 index data (i.e., non-stationarity and non-linearity). Next is the forecasting performance of the SVM and ANN models, the DS values of these models were 84.95 and 93.55, respectively. These values are higher than a random guess and showed that SVM and ANN models had a higher capability of directional forecasting compared with ARIMA and GARCH models. The ANN model attained the highest percentage of directional forecast which is 93.55% for KSE100 index data and outperformed other models such as ARIMA, GARCH, and SVM. Thus, the ANN and SVM models achieved the highest percentage in directional forecasting while the ARIMA and GARCH models achieved the lowest values for the directional forecast. The ANN method shows significant performance compares to the rest of the methods based on 20% testing parts. Figure 8 shows the plot of all forecasting accuracy measures for all models.

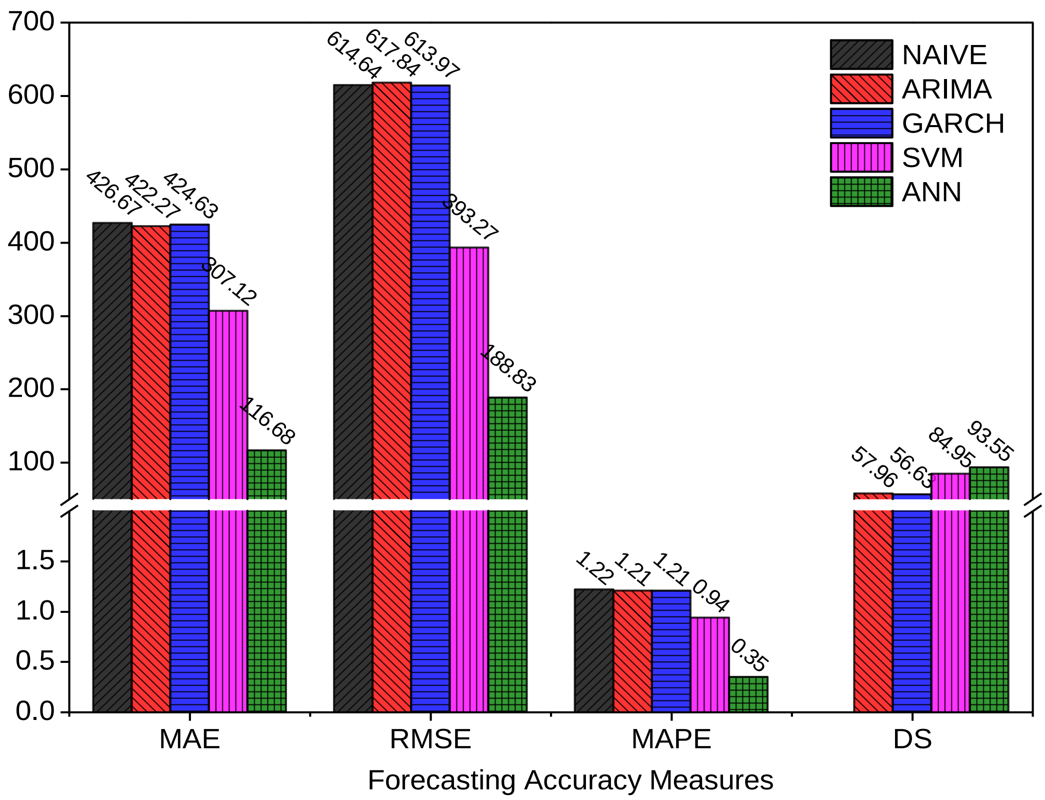

Figure 8: Plot of the forecasting accuracy measures for testing the data set of all models

To authenticate the dominance of the ANN model, the study also applied the DM test. The DM test results statistically confirmed the above conclusion. For testing data, the DM test results are presented in Table 8. The DM test results statistically confirmed the above conclusion. Firstly, the ANN model chooses as a benchmark that statistically outperformed SVM, GARCH, ARIMA, and NAÏVE models and their respective p-values were far below 0.001 for the KSE100 index testing data. Secondly, the SVM model selects as a benchmark model, and the corresponding p-values of the GARCH, ARIMA, and NAÏVE models were not statistically significant for KSE100 index testing data. The SVM model performed well but their performance was not statistically different from the GARCH, ARIMA, and NAÏVE models. Thirdly, the GARCH model selects as a benchmark model, and the corresponding p-values of the ARIMA and NAÏVE were not statistically significant which indicates that the performance of these three models is the same at a 1% level of significance. The last is the ARIMA model compared with the NAÏVE model, the performance of both models was statistically insignificant, which indicated that both models have the same forecasting accuracy.

Table 8: The DM test results of the KSE100 index testing data set

Model	SVM	GARCH	ARIMA	NAÏVE
ANN	- 3.2971 (0.001)*	-5.5482 (<0.001)*	-5.5646 (<0.001)*	-5.5843 (<0.001)*
SVM		0.0781 (0.938)	0.0349 (0.972)	0.1019 (0.919)
GARCH			-0.3920 (0.695)	-0.3520 (0.725)
ARIMA				-0.2920 (0.770)

" * " significant at 1%, Alternative hypothesis: Model forecasts are different

For a clearer picture, the original testing data and all fitted model data are plotted in Figure 9. The fitted models include NAÏVE, GARCH, ARIMA, SVM, and ANN. In Figure 9(a) the testing data set original values and one-step ahead forecasted values of all models are plotted. While in Figure 9(b) the last thirteen days' values are more focused to highlight the pattern of all models with the original testing values. Among all models, the testing forecasted values of the ANN model are very much close to the original testing set values which confirms the superiority of the ANN model. Thus, the recommended model for forecasting the KSE100 index data is the ANN. The ANN model can handle the complex data of the KSE100 index and provide more accurate forecasts.

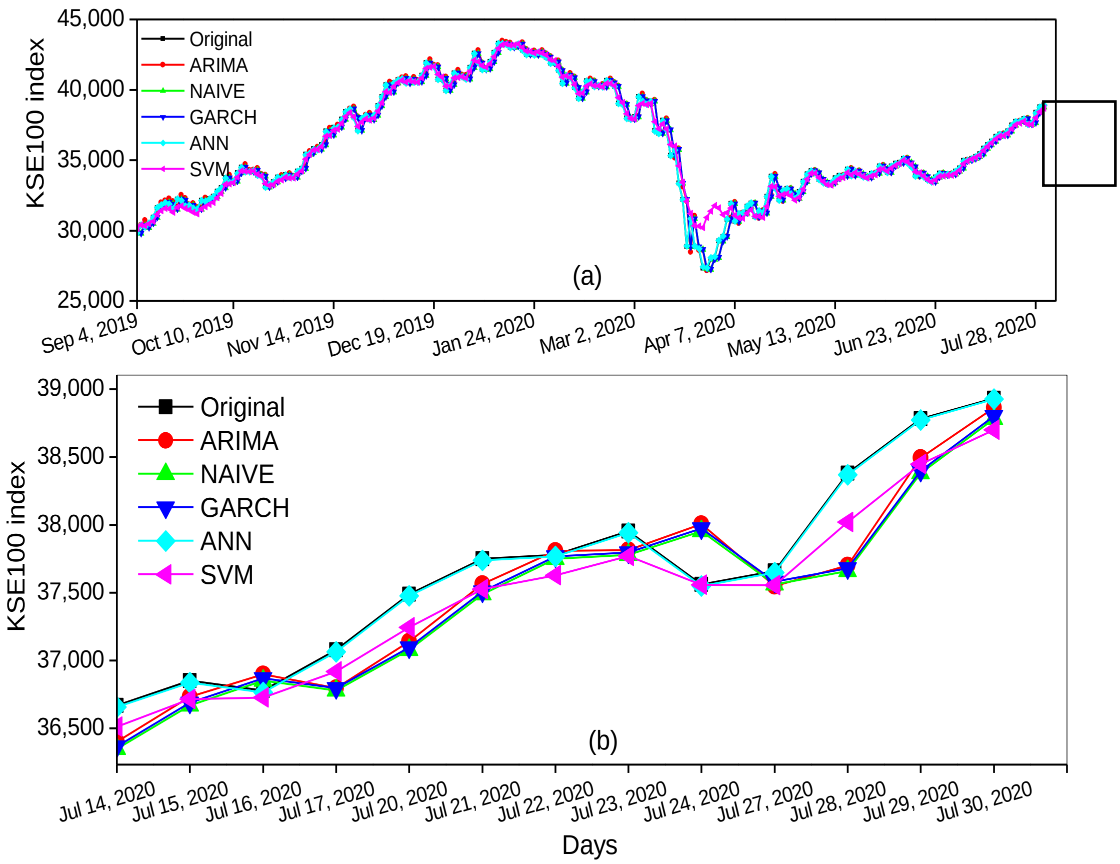

Figure 9: Plot of the original and forecasted values of all models (a) testing set (b) focused values of the last 13 days of the KSE100 index testing data set

3.8 Future Forecasts

The actual trend of detecting the KSE100 index in Pakistan has been increasing with exponential growth and is based on public behaviours and the Government's interference. Using the 1137 days of data from 01 January 2020 to 30 July 2020, ARIMA and ANN model, we forecasted the KSE100 index up to 31 August 2020. As we are trading with the nonstationary and nonlinear time series the ANN model is the best choice and provides a more accurate forecast for the future. The best model for KSE100 index data is used and daily forecasts are computed. The forecasting results of the model ANN demonstrated that there is an increase in the KSE100 index data with an exponential growth rate. While the ARIMA model forecasts show an increase in future values at a constant rate. Figure 10(a) shows the original and forecasted values of all models for testing data sets and also one-month ahead forecasts of ARIMA and ANN models. Figure 10(b) the one-month ahead forecasts of the ANN and ARIMA models for the KSE100 index are more focused to see a clearer picture of the data.

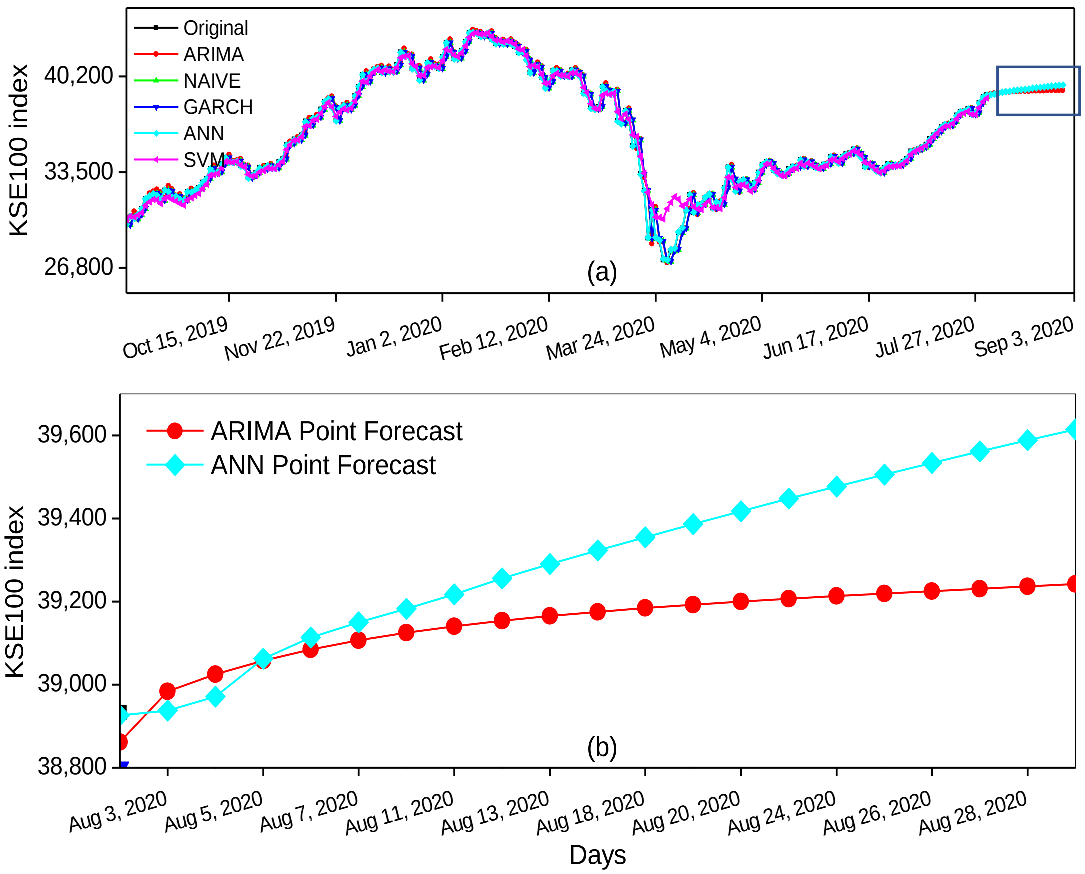

Figure 10: Plot of the original and forecasted values of all models (a) testing set (b) Plot of the one-month ahead forecasts of the ANN and ARIMA models for the KSE100 index

4 Conclusion

The stability of the stock exchange, is an important indicator of the overall economic performance of a country and, is one of the main objectives of monetary policy. However, maintaining stock exchange stability over a period requires forward-looking approaches due to the lags and dynamic structure of stock exchange index movements. Therefore, predicting and forecasting the stock exchange index has emerged as a crucial factor to be placed in monetary policy decisions for almost every monetary decision-making process. In this study, the KSE100 index forecasting accuracy of alternative time series models is evaluated. For this purpose, static and dynamic forecasts are produced, and the accuracies are compared by MAE, RMSE, MAPE, and DS measures as well as by DM test. The results indicated that the ANN was the most effective machine learning approach for improving the forecasting accuracy of the KSE100 index. Thus, the recommended model for forecasting the KSE100 index data is ANN which can handle the non-linearity pattern in the data. Hence, the ANN model could be used for forecasting the KSE100 index data. Therefore, the proposed model for forecasting the KSE100 index data is ANN for daily data. The dynamic forecast results are also computed one month ahead and presented in Figure 10.

REFERENCES

[1] S. Muhammad and G. Ali, "The relationship between fundamental analysis and stock returns based on the panel data analysis; evidence from karachi stock exchange (kse)," Research Journal of Finance and Accounting, vol. 9, pp. 84-96, 2018.

[2] P. M. Dechow, A. P. Hutton, L. Meulbroek, and R. G. Sloan, "Short-sellers, fundamental analysis, and stock returns," Journal of financial Economics, vol. 61, pp. 77-106, 2001.

[3] S. Moshiri and N. Cameron, "Neural networks versus econometric models in inflation forecasting," Journal of Forecasting, vol. 19, pp. 201-17, 2000.

[4] E. Nakamura, "Inflation forecasting using a neural network," Economics Letters, vol. 86, pp. 373-378, 2005.

[5] T. Koutroumanidis, K. Ioannou, and E. Zafeiriou, "Forecasting bank stock market prices with a hybrid method: the case of Alpha bank," Journal of Business Economics and Management, vol. 12, pp. 144-163, 2011.

[6] M. K. Okasha, "Using support vector machines in financial time series forecasting," International Journal of Statistics and Applications, vol. 4, pp. 28-39, 2014.

[7] P. Agnolucci, "Volatility in crude oil futures: a comparison of the predictive ability of GARCH and implied volatility models," Energy Economics, vol. 31, pp. 316-321, 2009.

[8] M. Aamir and A. Shabri, "Modelling and forecasting monthly crude oil price of Pakistan: A comparative study of ARIMA, GARCH and ARIMA Kalman model," in AIP Conference Proceedings, 2016, p. 060015.

[9] M. Aamir and A. B. Shabri, "Modelling and forecasting monthly crude oil prices of Pakistan: A comparative study of ARIMA, GARCH and ARIMA-GARCH models," Science International vol. 27, pp. 2365-2371, 2015 2015.

[10] P. Sadorsky, "Modeling and forecasting petroleum futures volatility," Energy Economics, vol. 28, pp. 467-488, 2006.

[11] W. Huang, Y. Nakamori, and S.-Y. Wang, "Forecasting stock market movement direction with support vector machine," Computers & operations research, vol. 32, pp. 2513-2522, 2005.

[12] K.-j. Kim, "Financial time series forecasting using support vector machines," Neurocomputing, vol. 55, pp. 307-319, 2003.

[13] F. E. Tay and L. Cao, "Application of support vector machines in financial time series forecasting," omega, vol. 29, pp. 309-317, 2001.

[14] R. A. Ahmed and A. B. Shabri, "Daily crude oil price forecasting model using arima, generalized autoregressive conditional heteroscedastic and support vector machines," American Journal of Applied Sciences, vol. 11, pp. 425-432, 2014.

[15] M. Tariq, Z. U. Haq, S. Jan, M. Jehangir, and M. Aamir, "Real Exchange Rate and Foreign Exchange Reserves: A Mercantilist View," Life Science Journal, vol. 10, 2013.

[16] T. Rashid, A. Ismail, and K. Rashid, "Time Series Analysis of Stock Market Volatility in Pakistan," Asian Journal of Probability and Statistics, pp. 12-23, 2020.

[17] M. Ahmad, Hasan, M., Rafiq, MY., Naeem, MA., & Naseem, MA. , "Time Series Modeling of KSE-100 Index " Smart Journal of Business Management Studies vol. 14, 2018 2018.

[18] N. Zafar and A. Hassan, "An Empirical Investigation of Herding: Case of Karachi Stock Exchange," Pakistan Journal of Life & Social Sciences, vol. 14, 2016.

[19] F. Iqbal, "Risk forecasting of Karachi Stock Exchange: A comparison of classical and Bayesian GARCH models," Pakistan Journal of Statistics and Operation Research, pp. 453-465, 2016.

[20] M. Khalid, M. Sultana, and F. Zaidi, "Stock market returns and direction prediction: an empirical study on karachi stock exchange," Mathematical Theory and Modeling, vol. 5, pp. 106-111, 2015.

[21] A. Haque and K. Naeem, "Forecasting volatility and Value-at-Risk of the Karachi Stock Exchange 100 index: Comparing distribution-type and asymmetry-type models," European Online Journal of Natural and Social Sciences, vol. 3, pp. pp. 208-219, 2014.

[22] J. Iqbal, S. Azher, and A. Ijaz, "Predictive ability of Value-at-Risk methods: evidence from the Karachi Stock Exchange-100 Index," EERI Research Paper Series2010.

[23] S. Fatima and G. Hussain, "Statistical models of KSE100 index using hybrid financial systems," in 2006 IEEE International Conference on Engineering of Intelligent Systems, 2006, pp. 1-6.

[24] D. C. Montgomery, C. L. Jennings, and M. Kulahci, Introduction to time series analysis and forecasting: John Wiley & Sons, 2015.

[25] W. Gao, M. Aamir, A. B. Shabri, R. Dewan, and A. Aslam, "Forecasting Crude Oil Price Using Kalman Filter Based on the Reconstruction of Modes of Decomposition Ensemble Model," IEEE Access, vol. 7, pp. 149908-149925, 2019.

[26] G. E. Box, G. M. Jenkins, and G. C. Reinsel, Time series analysis: forecasting and control vol. 734: John Wiley & Sons, 2011.

[27] R. Shibata, "Selection of the order of an autoregressive model by Akaike's information criterion," Biometrika, pp. 117-126, 1976.

[28] G. Schwarz, "Estimating the dimension of a model," The annals of statistics, vol. 6, pp. 461-464, 1978.

[29] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," Journal of econometrics, vol. 31, pp. 307-327, 1986.

[30] C. Cortes and V. Vapnik, "Support-vector networks," Machine learning, vol. 20, pp. 273-297, 1995.

[31] X. Fu, C. Liew, H. Soh, G. Lee, T. Hung, and L.-C. Ng, "Time-series infectious disease data analysis using SVM and genetic algorithm," in 2007 IEEE Congress on Evolutionary Computation, 2007, pp. 1276-1280.

[32] Z. Khan, M. Naeem, U. Khalil, D. M. Khan, S. Aldahmani, and M. Hamraz, "Feature selection for binary classification within functional genomics experiments via interquartile range and clustering," IEEE Access, vol. 7, pp. 78159-78169, 2019.

[33] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," Statistics and computing, vol. 14, pp. 199-222, 2004.

[34] G. Zhang, B. E. Patuwo, and M. Y. Hu, "Forecasting with artificial neural networks:: The state of the art," International journal of forecasting, vol. 14, pp. 35-62, 1998.

[35] C. Hamzaçebi, "Improving artificial neural networks' performance in seasonal time series forecasting," Information Sciences, vol. 178, pp. 4550-4559, 2008.

[36] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," Neurocomputing, vol. 50, pp. 159-175, 2003.

[37] G. P. Zhang and M. Qi, "Neural network forecasting for seasonal and trend time series," European journal of operational research, vol. 160, pp. 501-514, 2005.

[38] K. G. Sheela and S. N. Deepa, "Review on methods to fix number of hidden neurons in neural networks," Mathematical problems in engineering, vol. 2013, 2013.

[39] P. Xu, M. Aamir, A. Shabri, M. Ishaq, A. Aslam, and L. Li, "A New Approach for Reconstruction of IMFs of Decomposition and Ensemble Model for Forecasting Crude Oil Prices," Mathematical Problems in Engineering, vol. 2020, 2020.

[40] M. H. Beale, M. T. Hagan, and H. B. Demuth, "Neural network toolbox™ user's guide," in R2012a, The MathWorks, Inc., 3 Apple Hill Drive Natick, MA 01760-2098,, www. mathworks. com, 2012.

[41] N. Singh, A. Singh, and M. Tripathy, "Selection of hidden layer neurons and best training method for ffnn in application of long term load forecasting," Journal of Electrical Engineering, vol. 63, pp. 153-161, 2012.

[42] D. T. Bui, B. Pradhan, O. Lofman, I. Revhaug, and O. B. Dick, "Landslide susceptibility assessment in the Hoa Binh province of Vietnam: a comparison of the Levenberg–Marquardt and Bayesian regularized neural networks," Geomorphology, vol. 171, pp. 12-29, 2012.

[43] H. Yonaba, F. Anctil, and V. Fortin, "Comparing sigmoid transfer functions for neural network multistep ahead streamflow forecasting," Journal of Hydrologic Engineering, vol. 15, pp. 275-283, 2010.

[44] B. G. Kermani, S. S. Schiffman, and H. T. Nagle, "Performance of the Levenberg–Marquardt neural network training method in electronic nose applications," Sensors and Actuators B: Chemical, vol. 110, pp. 13-22, 2005.

[45] C. He and C. Ma, "A smoothing self-adaptive Levenberg–Marquardt algorithm for solving system of nonlinear inequalities," Applied Mathematics and Computation, vol. 216, pp. 3056-3063, 2010.

[46] A. P. Piotrowski and J. J. Napiorkowski, "Optimizing neural networks for river flow forecasting–Evolutionary Computation methods versus the Levenberg–Marquardt approach," Journal of hydrology, vol. 407, pp. 12-27, 2011.

[47] I. Mukherjee and S. Routroy, "Comparing the performance of neural networks developed by using Levenberg–Marquardt and Quasi-Newton with the gradient descent algorithm for modelling a multiple response grinding process," Expert Systems with Applications, vol. 39, pp. 2397-2407, 2012.

[48] N. Khan, D. Gaurav, and T. Kandl, "Performance evaluation of Levenberg-Marquardt technique in error reduction for diabetes condition classification," Procedia Computer Science, vol. 18, pp. 2629-2637, 2013.

[49] P. Kumar, S. Merchant, and U. B. Desai, "Improving performance in pulse radar detection using Bayesian regularization for neural network training," Digital signal processing, vol. 14, pp. 438-448, 2004.

[50] K. Hirschen and M. Schäfer, "Bayesian regularization neural networks for optimizing fluid flow processes," Computer methods in applied mechanics and engineering, vol. 195, pp. 481-500, 2006.

[51] J. Caballero and M. Fernández, "Artificial neural networks from MATLAB® in medicinal chemistry. Bayesian-regularized genetic neural networks (BRGNN): Application to the prediction of the antagonistic activity against human platelet thrombin receptor (PAR-1)," Current topics in medicinal chemistry, vol. 8, pp. 1580-1605, 2008.

[52] W. Jiang, L. Zhang, and P. Wang, "Nonlinear time series forecasting of time-delay neural network embedded with Bayesian regularization," Applied Mathematics and Computation, vol. 205, pp. 123-132, 2008.

[53] J. L. Ticknor, "A Bayesian regularized artificial neural network for stock market forecasting," Expert Systems with Applications, vol. 40, pp. 5501-5506, 2013.

[54] L. Yu, S. Wang, and K. K. Lai, "Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm," Energy Economics, vol. 30, pp. 2623-2635, 2008.

[55] L. Yu, S. Wang, and K. K. Lai, Foreign-Exchange-Rate Forecasting with artificial neural networks vol. 107: Springer Science & Business Media, 2010.

[56] L. Yu, W. Dai, and L. Tang, "A novel decomposition ensemble model with extended extreme learning machine for crude oil price forecasting," Engineering Applications of Artificial Intelligence, vol. 47, pp. 110-121, 2016.
[57] B. Zhu, X. Shi, J. Chevallier, P. Wang, and Y. M. Wei, "An Adaptive Multiscale Ensemble Learning Paradigm for Nonstationary and Nonlinear Energy Price Time Series Forecasting," Journal of Forecasting, 2016.
[58] S. Y. Wang, L.A. Yu and K.K. Lai,, "Crude oil price forecasting with TEI@ I methodology," Journal of Systematic Science. Complexity, vol. 18, pp. 145-166, 2005.
[59] R. J. Hyndman and Y. Khandakar, Automatic time series for forecasting: the forecast package for R: Monash University, Department of Econometrics and Business Statistics …, 2007.
[60] M. AAMIR, "CRUDE OIL PRICE FORECASTING BASED ON THE RECONSTRUCTION OF IMFS OF DECOMPOSITION ENSEMBLE MODEL WITH ARIMA AND FFNN MODELS," 2018.
[61] M. Aqil, I. Kita, A. Yano, and S. Nishiyama, "A Takagi-Sugeno Fuzzy System for the Prediction of River Stage Dynamic," Japan Agricultural Research Quarterly: JARQ, vol. 40, pp. 369-378, 2006.
[62] Z. He, X. Wen, H. Liu, and J. Du, "A comparative study of artificial neural network, adaptive neuro fuzzy inference system and support vector machine for forecasting river flow in the semiarid mountain region," Journal of Hydrology, vol. 509, pp. 379-386, 2014.
[63] M. Aamir and A. Shabri, "Improving Crude Oil Price Forecasting Accuracy via Decomposition and Ensemble Model by Reconstructing the Stochastic and Deterministic Influences," Advanced Science Letters, vol. 24, pp. 4337-4342, 2018.
[64] M. Aamir, A. Shabri, and M. Ishaq, "Improving forecasting accuracy of crude oil price using decomposition ensemble model with reconstruction of IMFs based on ARIMA model," Malaysian Journal of Fundamental and Applied Sciences, vol. 14, pp. 471-483, 2018.
[65] M. Aamir, A. Shabri, and M. Ishaq, "CRUDE OIL PRICE FORECASTING BY CEEMDAN BASED HYBRID MODEL OF ARIMA AND KALMAN FILTER," Jurnal Teknologi, vol. 80, 2018.