Hybrid Modeling of ARIMA, ANN and SVM for Macro Variables Forecasting in Pakistan

By

Rizwan Ali

Registration No: PIDE2016MPHILETS08

Supervised By

Dr. Hafsa Hina

Co-Supervised By

Dr. Amena Urooj

Department of Econometrics and Statistics

Pakistan Institute of Development Economics

Islamabad, Pakistan

2019

Pakistan Institute of Development Economics

CERTIFICATE

This is to certify that this thesis entitled: **"Hybrid Modeling of ARIMA, ANN and SVM for Macro Variables Forecasting in Pakistan"** submitted by Mr. Rizwan Ali is accepted in its present form by the Department of Econometrics and Statistics, Pakistan Institute of Development Economics (PIDE), Islamabad as satisfying the requirements for partial fulfillment of the degree in **Master of Philosophy in Econometrics**.

Supervisor:

Dr. Hafsa Hina
Assistant Professor
PIDE, Islamabad

Co-supervisor:

Dr. Amena Urooj
Assistant Professor
PIDE, Islamabad

External Examiner:

Dr. Mudassar Rashid Khawaja
Assistant Professor
Department of Economics
COMSATS, Islamabad

Head,
Department of Econometrics and Statistics:

Dr. Amena Urooj

DEDICATION

This thesis is dedicated to my family for their love, support and prayers. I am nothing without my family who desires to see me succeed in life.

ACKNOWLEDGEMENT

In the name of Allah the most Merciful and Beneficent

All the acclamations and appreciations are for **ALLAH**, the Almighty, the greatest of all, on whom ultimately we depend for sustenance and guidance. I would like to thank Almighty Allah for giving me opportunity, determination and strength to do my research. His continuous grace and mercy was with me throughout my life and ever more during the tenure of my research. All respects to His **Holy Prophet** (PBUH) who has guided his ummah to seek knowledge from cradle to grave and enable us to win honor of life and whose teachings has served as a beacon of light for humanity and the hours of despair and darkness.

I offer my humble and affable gratitude to my research Supervisor **Dr. Hafsa Hina** and co-supervisor **Dr. Amna Urooj** for their very affectionate and kind attitude guidance encouraging discussion and keen interest throughout my research work that enabled me to carry out my research work more confidently. I would like to express my sincere gratitude to **Dr. Hafsa Hina** for her continuous support in my research like a mother, for her patience, motivation, enthusiasm, and immense knowledge. Her guidance helped me in all the time of research and writing of this thesis. I could not have a better advisor and mentor for my MPhil study. Thank you Ma'am!

I am also grateful to **Madam Nadia Hassan, Dr. Saud Ahmad Khan, Madam Uzma Zia, Sarabjot Singh** for providing their help and guidance in my studies and research. I believe that I could not come to this level without their help.

I sincerely thank to my **Mother** and **Father** for their love, trust, prayers and financial support during my whole career , special thanks to my elder brother **Imran Ali** who treats me as a prince. Thank you bhai for being my guardian and believing in me . I also

appreciates my cute younger brother **Irfan Ali** for his love and care. No word can express my thanks to my **loving sisters** whom affections, prayers, care and support help me not only during my studies but throughout my life.

I express my special thanks to my friends (**Zeeshan Fazal, Hanzala Zulfiqar, Faridullah Kakar, Hamid Ali, Imad Ali, Tahir Mehmood, Asif Butt, Omair Ahmad, Muneer Khichi, Maqsood Ahmad (Sooda Mahi) and my whole colleagues of 2016 badge**) for their encouragement, moral support, and care . At the end, I would like to say thank you **shanay, hanzla, zubairay and shah sahib**, I enjoyed spending time with you. Thanks for giving me such a joyful time.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Time series forecasting remains a challenging task owing to its nonlinear, complex and chaotic behavior. Autoregressive integrated moving average (ARIMA) models are most frequently used since long time in forecasting. Artificial neural networks (ANN) is considered a good alternative to traditional ARIMA model in time series forecasting and often regard superior than ARIMA in forecasting performance. In recent literature Support vector machines (SVM) is becoming famous for solving nonlinear regression problems and time series forecasting. In this study, a hybrid methodology is used which combines the linear ARIMA with nonlinear models of ANN and SVM in order to improve the forecasting performance of Pakistan's macroeconomic variables such as inflation, exchange rate and stock return. The forecasting performance of all models i.e., ARIMA, ANN, SVM, ARIMA-ANN and ARIMA-SVM are compared on the basis of RMSE and MAE. The results indicate that the best forecasting model to achieve high forecast accuracy is the hybrid ARIMA-SVM.

Key words: Time series forecasting, ARIMA, ANN, SVM, Hybrid models

CHAPTER 1

INTRODUCTION

1.1 Introduction

Time series data are not very simple. They are subject to many complexities, instabilities, and nonlinearities which emerge due to dynamic behavior of economy. Due to complexity, instability and nonlinearity in most of time series data, it is not easy to make perfect and accurate predictions for it by using different conventional and advanced methods. Therefore, we never succeed in making accurate and perfect future forecasting by using different conventional and advanced methods. But it does not mean to stop making predictions as predictions keep the things moving in the hope of better future. It also helps policy makers in taking corrective measures and planning for the development of better future.

Autoregressive Integrated Moving Average (ARIMA) model has been used since long for time series forecasting and giving reliable results. It is well accepted owing to its statistical properties as well as popular Box-Jenkins (1976) methodology but it has limitations in nonlinear data set handling (Zhu and Wei, 2013). Now a days, novel neural network techniques have been widely used for nonlinear data sets because of their feature of fast learning and pattern recognizing for complex data sets. Artificial neural networks (ANN) is famous and well accepted technique due to its flexibility to nonlinear data sets and gives reliable results for financial time series, but it also has disadvantages of over fitting and not giving much understanding of data. Support vector machines (SVMs) is considered as a major breakthrough in machine learning and is widely applied to classification and regression analysis. It is high performing algorithm with little tuning. The long term and accurate predictions play a crucial role in any of business and investment strategy. Most of the studies have shown that a hybrid model of different linear and nonlinear models have some powerful tools to advance the forecast performance of each individual model (Khashei and Bijari, 2011).

Linear models such as ordinary least squares (OLS), maximum likelihood estimation (MLE), generalized linear model (GLM), Cholesky Decomposition, analysis of variance (ANOVA) Methods and ARIMA are used to analyze the linear dependency among the dependent and independent variables. There are many limitations for these models in case of nonlinearity and complex data set. Often these models are used for nonlinear data set by making different prior assumptions and taking data transformation. But in real life data are not normal as assumed, so it may lead to selection of the model which can be inappropriate and does not replicate the factual shape of the data set. However, linear models are preferred due to their simplicity, requirement of small memory space and speed of convergence in comparisons with nonlinear models which usually have low convergence (Petruseva *et al.*, 2017). When the influence of the nonlinearity to the overall specification of the model is very small, also one cannot assume a nonlinear model to perform fine than a linear model. But linear models cannot perform well for nonlinear functions as their counter parts can.

Nonlinear models like SVM, ANN, K-Nearest Neighbors, Decision Trees like CART, Random Forest and Naive Bayes take much attention for their importance of tackling the nonlinear and complex data set issues. So, it is difficult to decide either linear or nonlinear models are best that's why we choose both for our forecasting and will make their hybrid for sake of fruitful results.

In this study, forecasting performance of linear ARIMA, nonlinear models ANN, SVM and hybrid ARIMA-ANN and ARIMA-SVM would be compared. An appropriate combination of linear and nonlinear models yields a more precise predictions than any separate linear and nonlinear models for forecasting time series data (Babu *et al.*, 2014).

1.2 Objective of the Study

The objective of the study is to propose hybrid model of both linear and nonlinear models to achieve accuracy in forecasting for empirical illustrations. Macro variables like Inflation, Exchange rate and Stock return are used to compare forecasting ability of proposed hybrid models.

1.3 Significance of the Study

The significance of the research is the affirmed need in the financial and Marco variable time series data to be able to forecast precisely up to maximum level. The investors in the financial world are regularly involved in decision making that will pursue their aims on risks, returns and policies. Even though, obtaining precise market forecasts is difficult because financial time series forecasts are one of the most demanding task of time series forecasting (Cocianu and Grigoryan, 2015; Cao and Tay, 2001 and Al-hnaity and Abbod, 2016). Because in general time series data are often highly instable, noisy and deterministically chaotic in nature. Various univariate and multivariate measures have been taken for time series forecasting originating from linear and nonlinear techniques. In this study selected linear and nonlinear models and their hybrids models are applied with the objective of enhancing forecast performance of the actual measurements. For that reason the importance of this study is to discover the use of hybrid model for time series forecasts in Pakistan by performing a comprehensive experimental study comprises of linear, nonlinear and hybrid models. The following actions are taken to evaluate the time series forecasts performance of different models:

- Forecasts are made from ARIMA model.

- Forecasts are drawn from ANN and SVM.
- Hybrids ARIMA-ANN and ARIMA-SVM are devised.
- Forecasts are made by hybrid models of ARIMA-ANN and ARIMA-SVM.
- Finally, mutually comparing the results of all five models run in this study.

These three types of different models such as linear, nonlinear and hybrids are not used simultaneously for evaluating the performance of time series forecasting in case of Pakistan. So, it increase the value of this study in the sense of exclusive attributes of above mentioned models which are used to explore the forecasts performance of time series data in Pakistan.

1.4 Organization of the study

The organization of the thesis is as. Chapter 1 consists of introduction. Chapter 2 includes a detailed reviewed literature of methods used for time series forecasting. The data description and methodological framework followed for this thesis is presented in Chapter 3. While Chapter 4 provides the experimental result and discussion and comparison of obtained forecasts from all techniques. Finally, Chapter 5 concludes the thesis and the limitation of the study have some recommendations for extending work.

CHAPTER 2

LITERATURE REVIEW

2.1. Introduction

Several studies have been conducted for the sake of forecasting time series data arising from various fields. In literature different models are suggested based on their forecasting performance that are consisted of linear, nonlinear and hybrid models. In this chapter existing literature on ARIMA, ANN, SVM and hybrid modeling and their applications in different fields has been reviewed. The use of these models has been also be discussed with their weaknesses and strengths based on their forecast performance on different kind of data sets. The chapter proceeds from linear to nonlinear and then to their hybrid models. The usage of these models such as ARIMA, ANN and SVM in literature of Pakistan has been also discussed. ARIMA modeling was used almost in every field of forecasting in Pakistan whereas limited applications of ANN has been found in case of Pakistan. The applications of SVM modeling exist in pattern recognition and classification in literature of Pakistan but SVM has no evidence of its use in forecasting. Also there has been no effort is made to develop hybrid modeling for forecasting the Pakistan's economy.

2.1.1 Literature Review of Autoregressive Integrated Moving Average (ARIMA)

Meyler *et al.* (1998) have forecasted Irish inflation using ARIMA model. They have stated that ARIMA models are theoretically more validated and can be unexpectedly robust as compared to alternative modeling approaches. They have emphasized on forecast performance and suggested there should be more focus on minimizing out of sample forecasts errors rather than minimizing in sample 'goodness of fit'. Saeed and Zakria (2000) have used an empirical study of the Box Jenkins ARIMA methodology to forecast the wheat production in Pakistan. The diagnostic checking showed that

ARIMA with order (2,2,1) is suitable. The fifteen year ahead forecasting was done which showed a good fit with 95% confidence interval. Sultana *et al.* (2014) have forecasted inflation and economic growth for Pakistan by using ARIMA and decomposition methodology on monthly series. They have compared out of sample forecasts of both time series methods based on mean absolute deviation and sum of square of errors in which ARIMA give better forecasts performance.

Farooqi (2014) has presented ARIMA model to forecast the future annual values of imports and exports of Pakistan. Standard statistical techniques such as AIC and diagnostic check were used to determine the validity of the fitted model. The author has founded that ARIMA model with order (2,2,2) and (1,2,2) are suitable for imports and exports respectively. Jafri *et al.* (2012) have examined the rate of dust fall by using ARIMA model and Stochastic models a case study to Quetta, Pakistan. The aim of the study was to control the pollutants specially heavy and toxic metals present in the particular matters. Study showed that ARIMA model give comparatively better forecasts for rate of dust fall.

2.1.2 Literature Review of Artificial Neural Networks (ANN)

Zhang *et al*. (1999) have bonded the gap between theoretical development and the real world applications of ANN and presented the general framework for understanding the role of neural network. They have predicted the bankruptcy and check the robustness of the model through cross validation. Tkacz and Hu (1999) have concluded that neural networks are best as compared to their traditional counterparts. These model can capture more essential non linearities among financial variables and real output growth at longer horizons and performed poor for 1-quarter forecasts. They have said that neural networks are very robust to exploit the nonlinear relationships between variables to provide the more precise forecasts of economic activities. Moshiri and Cameron

(2000) have evaluated the back propagation neural network (BPNN) and conventional econometric techniques to forecast Canadian inflation. They have stated that ANN modeling can solve the complex problems and there is no need to make assumptions of linearity as in traditional models. They have concluded that hybrid BPNN outperform econometric techniques in some cases.

Haider and Hanif (2009) have forecasted inflation on the base of monthly data for Pakistan through ANN and simply compared with ARIMA and AR(1) out-of-sample forecast which showed ANN was more precise. They build the architecture of ‘feedforward with backpropagation’ by using standard levenberg marquadt algorithm. To reduce the error volatility 12 hidden layers were trained and model learning rate was kept 0.25.They evaluated the forecast by calculating RMSE. Burney *et al.* (2005) have forecasted Karachi stock exchange shares by ANN using pre-processed data. Levenberg marquardt algorithm is used through which weights are adjusted during the back error propagation. it proved fairly accurate forecast when compared with weighted exponential method and ARIMA model.

Samin *et al.* (2004) have used ANN for different metals hazardous in ground water contamination forecasting for Faisalabad city in Pakistan. The obtained results then compared with actual values and as well as World health organization (WHO) standards. They founded that ground water is fully contaminated with effluents and is for above the safe water standards. Awan *et al.* (2012) have proposed a hybrid non-linear Autoregressive exogenous model (NARX) established with feed Forward Network (FFNN), SVR and Neural Network Models. They have forecasted the long term industrial load by using these models comparison was made between stated three techniques based on MAPE. All of three models showed accuracy in results with acceptable MAPE in which proposed hybrid model NARX based FFNN remain more attractive in forecasting.

2.1.3. Literature Review Support Vector Machine (SVM)

Cortes and Vapnic (1995) have extended the SVMs to nonlinear regression problems by introducing the idea of mapping input vectors into higher dimensional space. Linear decision surface is constructed in this space which results to higher generalization ability. SVMs use structural risk minimization instead of empirical risk minimization which minimize the generalization error by minimizing its upper bound resulting to better generalization as compared with conventional techniques. Cao and Tay (2001) have evaluated the SVR as promising alternative to time series forecasting. They have noted that SVMs are better when compared with multilayer perceptron trained with back propagation (BP) based on criteria of normalized root mean square error (NRMSE) and mean absolute error (MAE). SVMs is faster and have small number of parameters as compared to BP. Calveria *et al.* (2015) have examined the regional forecasting for tourism using Support vector regression (SVR) with three different kernels and two ANN models of Radial Basis Function (RBF) and multilayer perceptron. SVR with Gaussian kernel outperformed other models of ANN and SVR with linear and polynomial kernels. The authors have concluded that the choice of kernel is important in SVR for better results and machine learning technique are better suitable for long term forecasts.

Guajardo *et al.* (2006) have used wrapper method for feature selection in SVR and then updated the model using proposed methodology to achieve better forecasting performance. They have stated that because of the complexity of data mining applications like regression, it is necessary to select the most valuable features to construct the respective model. The proposed methodology was applied to sales forecasting problem and its performance was compared with standard ARMAX method. The proposed methodology showed little advantage in forecasting performance. Many authors (Mukhergee, 1997; Kim, 2003; Cao and Tay, 2003) have done financial and time series forecasting in order to attain satisfactory results. They have stated that SVMs received an increasing interest from its earlier application in pattern recognition to other fields such as regression analysis due to its outstanding generalization ability. They have concluded that SVMs is advantageous over other machine learning techniques and a promising alternative to financial forecasting's.

2.1.4 Literature Review of Hybrid Models (HM)

Zhang (2001) has stated that hybrid methodology of ARIMA and ANN could be a productive way to improve forecasting performance by their unique feature of linearity and nonlinearity. He has proposed the hybrid methodology by estimating first linear part by ARIMA and then its residuals through ANN and compared the results based on mean square error (MSE). He has inferred that hybrid model may perform worse than ANN and ARIMA in some data points but its overall performance is better. Pai and Lin (2005) have stated hybrid methodology of ARIMA and SVM showed satisfactory results in stock price forecasting. They have used the real data sets of ten stocks in order to inspect the accuracy of proposed methods. In which they have used four statistical indices such as mean absolute error (MAE), MSE, MAPE and RMSE to measure the performance of the proposed model. Chen and Wang (2007) have proposed hybrid model of SARIMA and SVM and concluded that hybrid methods give reliable results as compared to individuals methods. They have forecasted the Taiwan's machinery industry production values to investigate the proposed methodology. Normalized mean square error (NMSE) and mean absolute percentage error (MAPE) are used to check the accuracy of presented methodology.

Khashei and Bijari (2010) have stated that ANN modeling is a universal approximator and flexible computing framework which can be applied to a wide range of time series forecasting. But for some real time series, it cannot give satisfactory results. So, there is a need of integration of some different models to improve predictive performance in an effective way. They have proposed a hybrid model of ANN and ARIMA and concluded that a hybrid model can be used as an alternative to traditional ANN to obtain high degree of accuracy in forecasting. Zhu and Wei (2013) have made three different hybrid models of ARIMA and least square support vector machine (modification of SVMs) to forecast carbon prices. To determine the optimal parameters of LSSVM in order improve the forecasts ability particle swarm optimization (PSO) was used. Two main future carbon prices are used to compare the forecasting performance based on the criteria of root mean square error (RMSE) and D stat of the European climate exchange market. The results showed that only one hybrid perform better but not all hybrid models are superior than their single ones.

Babu and Reddy (2014) have stated that it is not good to apply simply ARIMA model on a series and then its residuals modeled by ANN instead they have used moving average filter model to decompose the data into linear and nonlinear parts and then apply the respective models. Many authors have combined statistical models and suggested that hybrid statistical modeling improve predictive performance as compared to stand-alones (Stone, 1974; Breiman, 1996; Leblence and Tibshirani, 1996 and Mojirsheibani, 1999). Papatla *et al.* (2012) have presented two classes of hybrid models linear and nonlinear and suggested mixed NN has higher probability in performance. Kumar (2014) has made hybrid model of SVM, ANN, Random Forest with ARIMA and compared all with each other and their single ones, in which SVM-ARIMA hybrid model took first place. He has measured the results on RMSE, MAE and NMSE based criteria for the purpose of to identify the best hybrid model. Among others ANN-ARIMA performed better in forecasting the stock index returns. He has suggested that ARIMA- SVM could be a worth to the forecasting improvement and assure profit-making returns for policy makers in forecasting economic and financial data. Khandelwal *et al.* (2015) have proposed Discrete Wavelet Transform (DWT) to split the data into linear and nonlinear components. Then ARIMA and ANN models have been used independently to recognize and forecast the reconstructed detailed and approximate components. The proposed DWT hybrid model has outperformed other models based on MSE and MASE for four real world time series.

Al-hnaity and Abbod (2016) have explored the predictability of stock index time series. They have stated that a single classical model will not yield the accurate prediction results. In their study, they have used most famous data mining techniques such as SVR, SVM, and BPNN. In which they have combined these three models and made a hybrid model. The weights of models were determined by genetic algorithm and to improve the prediction performance of single SVR and SVM Quantization factor was used. They have concluded that proposed hybrid model outperformed all other single models and the bench mark traditional model AR based on MSE, RMSE and MAE. Zhu *et al.* (2017) have forecasted the Air Quality Index (AQI) by using the hybrid models of Empirical Mode Decomposition Support Vector Regression (EMD-SVR) and EMD-Intrinsic Mode Functions (IMFs). To validate the proposed models forecast performance as compared to all other models employed in this paper, famous loss errors like MAE, RMSE, MAPE, MSE, absolute relative error and index of agreement are utilized. The AQI forecasts for empirical research has showed that two proposed hybrid models were superior to all other models like ARIMA, SVR, GRNN, EMD-GRNN, Wavelet-GRNN and Wavelet-SVR.

Kim and Won (2018) have examined a new hybrid long short-term memory (LSTM) which was combined with various generalized autoregressive conditional heteroscedasticity (GARCH)-type models to forecast stock price volatility. The proposed models performance were compared with different existing methodologies such as GARCH, exponential GARCH, exponentially weighted average, deep feedforward neural network (DFN), LSTM and as well as DFN combining with one GARCH type model. They have concluded that proposed hybrid models have lowest prediction errors based on MAE, MSE, heteroscedasticity adjusted MAE and heteroscedasticity adjusted MSE. Karathanasopoulos and Osman (2019) have proposed a hybrid model to forecast the Dubai Financial general index in which they have combined the momentum effect with a novel methodology of deep belief networks. They have compared the empirical results based on MAE, MSE, MAPE and RMSE with other three linear models named moving average convergence divergence, naïve strategy and ARMA model. They have concluded that proposed hybrid model outperformed all other models significantly and provides auspicious results for further usage in financial forecasts.

2.2 Literature Gap

After reviewing the literature we come to conclude that different techniques have different features regarding their ability of forecasting. But all of these cannot capture everything lonely. Therefore, there is need to make hybrid model by taking advantage of their specific abilities.

To our best knowledge, there has been no evidence that SVM and hybrid modeling of ARIMA with ANN and SVM is used for time series forecasting of macro variables in Pakistan.

CHAPTER 3

METHODOLOGY

This chapter discusses, the methodological framework employed in this study and data description. In section 3.1 the methodology of ARIMA model is explained. ANN model is discussed in section 3.2. Section 3.3 described the methodology of SVM. While section 3.4 explained the hybrid modeling in detail. Data description is given in the last section of 3.5.

3.1 Autoregressive Integrated Moving Average (ARIMA) Modeling

Box and Jenkins (1976) introduced the ARIMA model one of the most popular approaches to forecast time series data. In an ARIMA model, the future value of a variable is assumed to be a linear combination of several past values and past errors. Mathematically, it can be written as:

$$y_t = \sigma + \varphi_1 y_{t-1}, +\varphi_2 y_{t-2}, + \cdots +\varphi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1}, -\theta_2 \varepsilon_{t-2}, - \cdots -\theta_q \varepsilon_{t-q} \qquad 3.1$$

Where y_t represents the series at time t, ε_t represents the error terms at time t, φ, θ are the coefficients and p, q are lag lengths of AR and MA respectively. Box and Jenkins methodology is followed to choose the appropriate univariate model of ARIMA for forecasting. Box-Jenkins gives a systematic procedure for choosing ARIMA model which involves iterative steps such as identification, estimation and diagnostic checks on model adequacy. The detail of these three iterative steps are discussed below respectively.

3.1.1 Model Identification

Assuming for the moment that series is stationary and there is no seasonal variations. The initial tentative model identification starts from the plots of autocorrelation and partial autocorrelation functions. Frequently, it is possible to detect one or several possible models for a given time series by matching empirical autocorrelation patterns with theoretical ones. The relevant properties of the correlation patterns are given below 3.1.

Table 3.1 Model Identification through ACF and PACF

	ACF	**PACF**
AR(p)	Infinite spikes.	Finite spikes and cuts off after p lags
MA(q)	Finite spikes and cuts off after q lags	Infinite spikes
ARMA(p, q)	Infinite spikes	Infinite spikes

This process encompasses the subjective element at the identification stage which can be an advantage because it permits non-sample information to be taken into account. Frequently, it is considered suitable to discern the magnitude of large autocorrelation and partial autocorrelation coefficient. An autocorrelation to be statistically significant must be at least $\frac{2}{\sqrt{N}}$ in absolute value. When time series data has time varying mean and variance or both then data is considered to be non-stationary and in case of ARIMA modeling there is preliminary requirement for analysis to do it stationary by a method of differencing the data or transformations. The Augmented dickey-fuller (1981) is used to determine the null hypothesis of unit root problem in the data set under investigation. The auxiliary regression for the detection of unit root is:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \theta_1 \Delta y_{t-1} + \cdots + \theta_p \Delta y_{t-p} + u_t \qquad 3.2$$

Where α and β are coefficients of constant and time trend and p is lag order of the autoregressive process. The null hypothesis $\gamma = 0$ indicate series has unit root against the alternative hypothesis of $\gamma < 0$ series has no unit root. If the calculated value is less than the critical value one can reject the null hypothesis that series has unit root. Test statistic for calculated value is:

$$DF_{cal} = \frac{\check{\gamma}}{SE(\check{\gamma})} \qquad 3.3$$

The critical values at 5% level with constant and if series has deterministic trend are -2.86 and -3.41 respectively. The amount of differencing and the inclusion of a constant in the model determine s the long term behavior of the model.

3.1.2 Model Estimation

Once a preliminary model is identified, the model estimation is simple. The parameters are estimated in such way that total measure of errors is minimized which can be done through a nonlinear optimization process. The frequently used method to estimate the ARIMA model is maximum likelihood estimation (MLE) presented in Box-Jenkins (1976). Which attempts to maximize the log-likelihood for given values of p, d and q and finding those values of parameters that would lead to highest probability of finding the data we have observed.

3.1.3 Diagnostic Checking

There are two types of diagnostic checks in which first extra coefficients are fitted and then tested for their significance. In the second residuals of the fitted model are examined to determine if they are white noise. The checking is accomplished by inspecting the autocorrelation plots of the residuals to verify there is no further structure can be found i.e. residuals are white noise. To check the white noise behavior of residuals of estimated model, portmanteau test is applied. The Ljung-Box test is used to diagnose the autocorrelation among error terms and to examine the heteroskedasticity problem same test is applied on the square of residuals. JB test is used to determine the normal distribution of the residuals. This procedure of inspection the residuals and modifying the values of p and q continues till the resulting residuals hold no extra information. Brock-Dechert-Scheinkman (BDS) test is applied to the estimated residuals to check the nonlinear dependence in all data series. It is a nonparametric test developed by Brock et al. (1987 & 1996) which is a powerful tool that is not only used to detect deterministic chaos and nonlinearities in stochastic time series, but also it can be used as general test for model misspecification. The null hypothesis for BDS is that series are linearly dependent. Statistically BDS test values can be obtained by using given formula as:

$$BDS_{\epsilon,m} = \frac{\sqrt{n}(c_m(\epsilon) - c_1(\epsilon)^m)}{\sqrt{V_{\epsilon,m}}}$$

Where ϵ is tolerance distance, m is embedding dimension, n is sample size, $c_m(\epsilon)$ referred as joint probability of each pair of points in the set satisfying the epsilon condition and $c_1(\epsilon)^m$ is the product of individual probabilities of each pair. While $\sqrt{V_{\epsilon,m}}$ is the standard deviation that varies with dimension m. If $BDS_{\epsilon,m} > 2$ the null hypothesis is rejected with a confidence level of 95 percent and if $BDS_{\epsilon,m} > 3$ the null hypothesis is rejected with 99 percent confidence interval. Once the appropriate model is nominated then the model may be used to produce future forecasts.

3.1.4 Seasonal Box Jenkins

If time series have a seasonal component, in case of seasonal fluctuations in time series data model would become seasonal ARIMA (SARIMA). Mathematically the general expression of the SARIMA model can be inscribed as follow:

$$\phi_p(B)\Phi_P(B)(1-B)^d(1-B^s)^D y_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t \qquad 3.4$$

Where P and Q are seasonal lag lengths, D is the order of seasonal differencing and s is the number of seasons per year. In case of seasonal series it would process the same cycle of Identification, Estimation and Diagnostic checking followed by forecasting. If the series is seasonal the autocorrelation will have spikes at the seasonal frequency. If seasonal differencing is required then the auto correlogram must be re-estimated for the seasonal series. For the nonseasonal case, identification of D proceeds the same way. The selection of p, q, P and Q tentatively identified from the autocorrelation and partial autocorrelation functions in a somewhat similar way as in the nonseasonal model. Identification of P and Q are done by looking at the autocorrelation and partial autocorrelation at lags s, 2s, 3s and so on to multiples of the seasonal frequency. The procedure set out in the Table 3.2 below.

Table 3.2 Model identification for Seasonal Box Jenkins

Properties	**Inferences**
SAC dies down, SPAC has spikes at SL, 2SL,…, PSL and cuts off after PSL	Seasonal AR of order P
SAC has spikes at lags SL, 2SL,…, QSL and SPAC dies down	Seasonal MA
SAC has spikes at lags SL,2SL, …, PSL SPAC has spikes at lags SL, S2L, …, QSL and both die down	Use either ▪ Seasonal AR of order P or ▪ Seasonal MA of order Q
No seasonal spikes	P = Q = 0
SAC and SPAC die down	Possible P = Q = 1

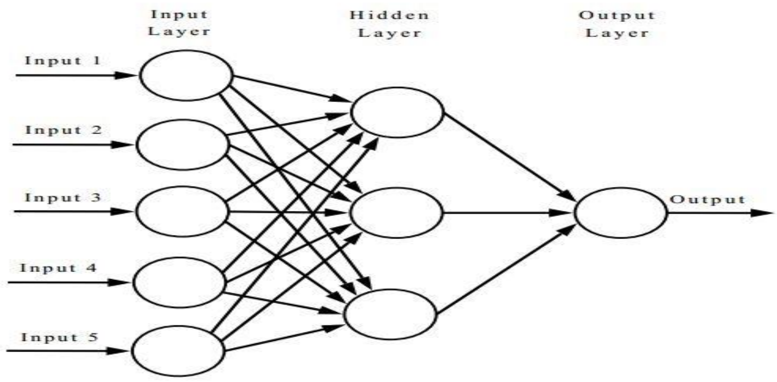
Source: Frain (1992)

3.2 Artificial Neural Networks (ANN) Methodology

A computational model for neural networks based on mathematics and algorithms called the threshold logic have been created by Mcculluch and Pitts (1943). This model covers the process for neural network study to split into two methods. One method focused on biological process in the brain whereas the second concentrated on the application of neural networks to artificial intelligence. Connectionist systems in ANN are computing system that are motivated by biological neurons works like human brain such as learn by experience and then analyze it. Such system "learn" (i.e. gradually upgrade performance on) tasks by evaluating examples, usually without task specific programming.

ANN established on a number of attached units or nodes termed artificial nodes. Each connected unit between artificial nodes can transfer a signal from one to another. In common ANN executes the signal at a connection among artificial neurons is real number, and the output of each artificial is computed by a nonlinear function of the summation of its inputs. ANN nodes and units consist of layers usually three layer perceptron is enough to train a model but it can be changed depending on the structure of data set. Most of the time second hidden layer is used to detect the relationship between variables and discontinuities. ANN works in a very analogous way, it takes several inputs, process it into and out of multiple neurons from multiple hidden layers and yields the result using an output layer. This outcome estimation procedure is technically known as "Forward propagation". The basic objective of ANN is to solve problems in the same as human mind do and to conduct different tasks more quickly and in better way than the traditional systems. Structure of an artificial ANN is given in Figure 3.1. Figure consists of three multilayer feed-forward network with five inputs neurons where each next neurons receives inputs from the previous layer and also outputs of neurons in one layer are inputs to the next layer.

Figure 3.1 ANN Architecture



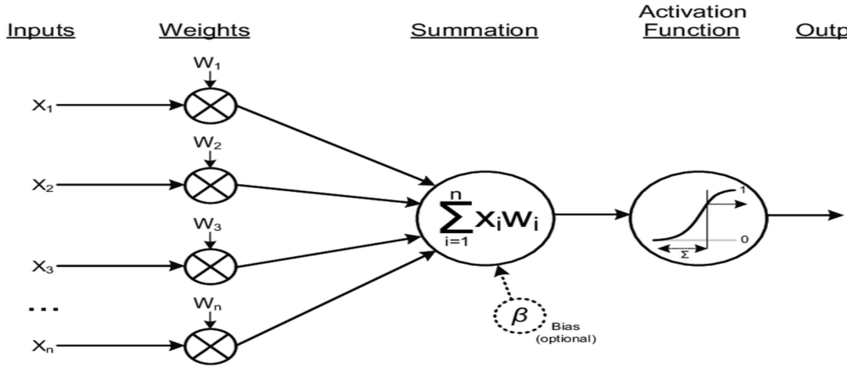
Source : Sharma and Bhardwaj (2015)

Artificial neurons and networks generally have weights and biases that are modified as learning proceeds. The weights increases or decreases the robustness of the signal at a connection. The target is to make the output to neural networks as close as possible to desired output. The mathematical association among the output (y_t) and the inputs ($y_{t-1}, \dots, y_{t-p}$) has the following representation:

$$y_t = w_0 + \sum_{j=1}^{q} w_j . g\,(w_{0.j} + \sum_{i=1}^{p} w_{i.j} . y_{t-p}) + e_t \qquad (3.5)$$

Where $w_{i.j} = (i = 0,1,2, \dots, p\ j = 1,2,3, \dots, q)$ and $w_j = (\ j = 0,1,2, \dots, q)$ are model constraints also called neuron weights; p represent the input nodes and q is the number of hidden nodes. The different activation functions are used to transmit an input signal of a neuron to an output signal. Most commonly used activation functions as the hidden layer transfer functions are Sigmoid, Tanh and ReLu. The neural network structure of activation function is showed in Figure 3.2:

Figure 3.2 ANN Architecture with a sigmoid activation function



Source : Patterson (2012)

These neurons are contributing of some error to final output individually. The value of neurons those are contributing more to the error are minimized and this happen while moving back to the neurons of the neural network and detecting where the errors lies. This process is known as "back propagation". Back propagation distributed the error terms back through the layers by updating the weights and each neuron. The neural networks uses a common algorithm known as "Gradient Descent" in order to reduce the numbers of iterations while minimizing the error. Which helps the task quickly and efficiently, one round of onward and back propagation iteration is known as one training iteration also known as Epoch.

3.2.1 Selection of Input Parameters

ANN modeling for time series analysis is considered good alternative to ARIMA forecast. ANN form of Multilayer Perceptron (MLP) is best and most generally used neural network architecture. The most decisive part of the ANN modeling is the selection of input lags because these determine (non-linear) autocorrelation form of the time series. Neural networks has no hard and fast rule for parameter selection. So determining the number of input lags p and number of hidden nodes q are data dependent. There exist many approaches for parameter selection in AAN modeling but none of these techniques give promise to perform best for actual time series analysis (Khashei and Bijari, 2010). Therefore, there is no rule of thumb or theory that give exact direction for parameter selection. For that reason many experiments are done to select the suitable p and q parameters that minimize the general criteria for accuracy like mean square error (Zhang, 2001). Once an appropriate ANN parameters are selected for a data set then the model is ready for final forecast analysis.

3.2.2 Training of Artificial Neural Networks Architecture

As specification of input lags is very crucial in training of ANN architecture, so there is need of universal methodology in specifying ANN architecture. Because a data generating process may show a diversity of stochastic and deterministic time series patterns of single or multiple seasonality, cycles, trends, pulses and structural breaks depending on time frequency. Therefore our selection of lags laboriously relies on automated network specification by using combined filter and wrapper approach for feature evaluation and a methodology that combines a novel iterative neural filter based on multilayer perceptron (Crone and Kourentzes, 2010; Kourentzes and Crone, 2010). Both automated methodologies are fully data driven and select input lags wisely without any expert intervention. The selection of hidden nodes is also data dependent and there is no systematic rule for deciding these parameters (Khashei and Bijari, 2010). To make optimum selection of hidden nodes validation test is considered in which 20% of time series validation sample is used within training data. Then chosen architecture based on discussed methods is again updated by modifying input lags and hidden nodes manually by error and trail method for the sake of parsimonious model. The final selected architecture is that which gives the minimum Mean square error (MSE).

3.2.3 Estimation of parameters and activation function

Once an appropriate ANN architecture with input lags and hidden nodes is selected then parameters are estimated through back propagation. This is the process in which weights of neurons are updated to traveling back through layer by layer and yielding the ultimate output. This process of backward pass is done until the desired output is achieved. Resilient back propagation (RPROP) is a well-known gradient descent algorithm and is used to train a neural network. RPROP only uses the sign of gradients to compute the parameter updates. If the sign of parameter remain in the same direction for several iterations then the step size of update value will be increased otherwise in case of oscillation it will be minimized. The main advantage of RPROP over standard back propagation is that it does not need any free parameter value like learning rate and momentum term and is much faster. RPROP also has separate step size for each weight which mean if one weight is much closer to optimal value while other needs more updates then it has no big issue like other gradient descent variants which can cause problem in this situation. But the drawback of RPROP is that it is more complex to apply than standard back propagation. Activation function basically is used to transfer input signal of a node to an output signal in ANN architecture. Activation function sum the multiplication of inputs and their respective weights and then feed it as input to the next layer of neurons through transfer function. The sigmoid or logistic function is used as transfer function and typically it can be written as:

$$Sig(x) = \frac{1}{1 + e^{-x}} \qquad (4.5)$$

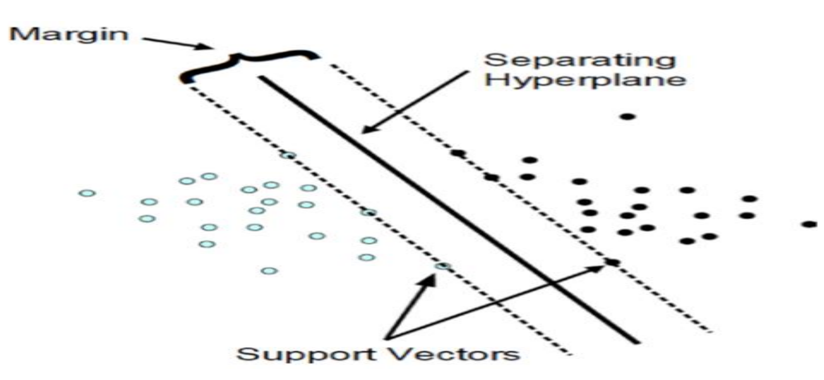
The all three combination operators like mean, median and mode can be used to produce ANN forecasts with their specific qualities depending on the nature of data. But median combination operator performs very well and converge to reliable and better ANN forecasts (Barrow and Kourentzes, 2018).

3.3 Support Vector Machines (SVM)

Support vector machine (SVM) developed in 1990's by Vapnic *et al.* SVM is basically a classification method that executes classification tasks by formulating hyperplane in a multidimensional space that splits the cases of unalike class labels. It is a most popular supervised machine learning which can be used for both regression and classification tasks and can handle continuous and categorical variables. SVM works by transforming the data to a higher dimensional space, and then performing separation in resulting hyperplane. Now we discuss how SVM works, theory behind it and its mathematical representation as follows:

In SVM a line is constructed that best separates the points to their class 0 or 1 in input space variable. The best or optimal hyperplane is the line which separates the two classes by largest margin, and the margin is the distance between the neighboring data points of that hyperplane. Margin is calculated as the vertical distance from line to only the closest points. These points are important in explaining the line and in constructing the classifier. These are the points that called support vectors and they define or support the hyperplane. Below Figure 3.3 further clarify the idea.

Figure 3.3: Hyperplane

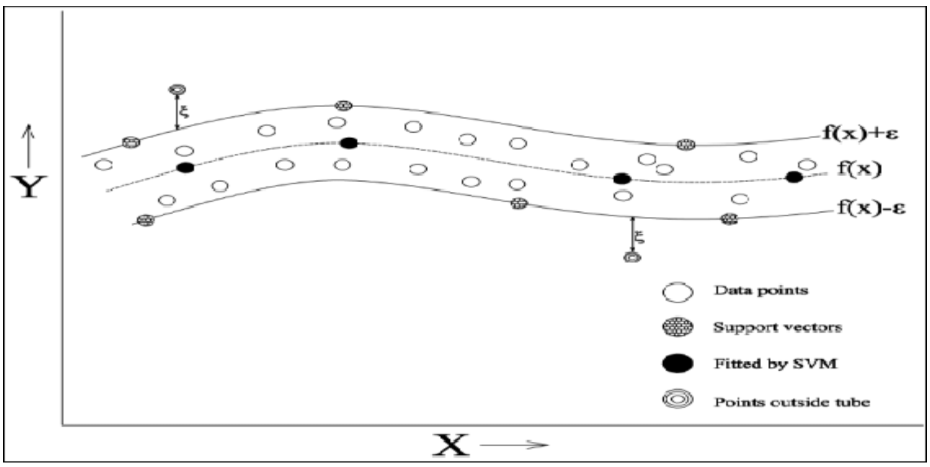


Source :Meyer and Wien (2015)

Mostly we relax the constraints of hyperplane that is called soft margin hyperplane because in practice the real data is complex and messy. Furthermore, the slack variables are added which give more wiggle room to the margin in each dimension. The magnitude of the wiggle allowed across all the dimensions is defined by introducing a tuning parameter C which describes the violation in the margin. If C=0 then there is no violation and if there is more the value of C means more the violation and vice versa. Because in reality it is not easy job to classify a complex, nonlinear and chaotic data effortlessly so SVM uses kernel functions which transform the data from input space to higher dimension space to classify data accurately.

First linear separable case is considered, then the soft margin support vector regression (SVR) is discussed and ultimately nonlinear case would be explained. Precisely, the ε -insensitive SVR would be used for forecasting. In ε -insensitive SVR, the main goal is to find a function f(x) that has an ε -deviation from the actually acquired target y_i for whole training data and simultaneously as flat as possible. Given below Figure 3.4 simply sketch the SVR.

Figure 3.4: Detailed epsilon tube with slack variables



Source : Kuntoji et al (2017)

Assume f(x) takes the following form as under taken by Smola and Scholkopf (2004):

$$f(x) = wx + b \qquad w \in x, b \in R \tag{3.6}$$

Here w is the regularizing term and give optimization problem central over the flatness of the solution and b is the bias. Then by solving the given problem:

$$\min \frac{1}{2} ||w||^2 \tag{3.6.1}$$

Subject to,

$$y_i - wx_i - b \leq \mathcal{E} \tag{3.6.2}$$

$$wx_i + b - y_i \leq \mathcal{E} \tag{3.6.3}$$

In case where the constraints are infeasible and to protect against outliers also to discern how many points can be tolerated outside the tube. One can insert slack variables $\xi_{i\,,}\xi^*$ in this case called soft margin formulations according to the Vapnic (1995) and is explained by the following problem (also see Figure 3.4 for slack variables which lie outsides the $\mathcal{E}$ -insensitive tube).

$$\min \frac{1}{2} ||w||^2 + c \sum_{i=1}^{n} (\xi_{i+} \xi_i^*) \tag{3.7}$$

Subject to,

$$y_i - wx_i - b \leq \mathcal{E} + \xi_i \tag{3.7.1}$$

$$wx_i + b - y_i \leq \mathcal{E} + \xi_i^* \tag{3.7.2}$$

$$\xi_{i\,,}\xi_i^* \geq 0$$

$$c \succ 0$$

Where c is trade of among the flatness of the f(x) and the number up to which deviations larger flatness apart from $\mathcal{E}$ are tolerated. That is called the $\mathcal{E}$-insensitive loss function $|\xi|_{\mathcal{E}}$ and is discussed by following equation,

$$|\xi|_{\mathcal{E}} = \begin{cases} 0 & if \quad |\xi| \quad \leq \mathcal{E} \\ |\xi| - \mathcal{E} & if \quad \xi > \mathcal{E} \end{cases} \tag{3.8}$$

According to Smola and Scholkopf (2004) By constructing the Lagrangian function as explained in Fletcher (1989), it can be deduced the dual problem. Specifically,

$$L = \frac{1}{2}||W||^2 + C\sum_{i=1}^{n}(\xi + \xi^*) - \sum_{i=1}^{n}\lambda_i(\mathcal{E} + \xi - y_i + wx_i + b)$$
$$- \sum_{i=1}^{n}\lambda_i^*(\mathcal{E} + \xi^* + y_i - wx_i - b)$$
$$- \sum_{i=1}^{n}(\eta_i\xi_i + \eta_i^*\xi_i^*), \quad (3.9)$$

In equation (3.9) all dual variables have to satisfy positivity constraints that is λ_i, $\lambda_i^*, \eta_i, \eta_i^* \geq 0$. At the optimal solution, it have

$$\frac{\partial L}{\partial w} = w - \sum_{i=1}^{n}(\lambda_i - \lambda_i^*)x_i = 0 \quad (3.9.1)$$

$$\frac{\partial L}{\partial b} = \sum_{i=1}^{n}(\lambda_i - \lambda_i^*) = 0 \quad (3.9.2)$$

$$\frac{\partial L}{\partial \mathcal{E}} = C - \lambda_I^* - \eta_I^* = 0 \quad (3.9.3)$$

According to Trafalis and Ince (2000), We will obtain the dual problem by substituting values of (3.9.1), (3.9.2) and (3.9.3) into (3.9) which is as follows:

$$\max \quad -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\lambda_i - \lambda_i^*)(\lambda_j - \lambda_j^*)x_i x_j$$
$$-\varepsilon\sum_{i=1}^{n}(\lambda_i + \lambda_i^*) + \sum_{i=1}^{n}y_i(\lambda_i - \lambda_i^*) \quad (3.10)$$

Subject to $\sum(\lambda_i - \lambda_i^*) = 0$

$\lambda_i, \lambda_i^* \ \epsilon(0, C)$

Solving (3.9.1) for w, it can be obtained:

$$w^* = \sum_{i=1}^{n} (\lambda_i - \lambda_i^*) x_i$$

Substituting the value of w in (3.6), the function will become as follows:

$$f(x) = \sum_{i=1}^{n} (\lambda_i - \lambda_i^*) x_i x + b^* \qquad (3.11)$$

The optimal value of b could be computed from the complementary slackness conditions Trafalis and Ince (2000) , precisely,

$$\lambda_i(\varepsilon + \xi_i - y_i + w^* x_i + b) = 0$$

$$\lambda_i^*(\varepsilon + \xi_i^* + y_i - w^* x_i - b) = 0 \qquad (3.12)$$

$$(c - \lambda_i)\xi_i = 0$$

$$(c - \lambda_i^*)\xi_i^* = 0$$

Only samples (x_i, y_i) with respective $\lambda_i = C$ exist outside the ℇ -insensitive tube around f. The set of dual variables λ_i, λ_i^* can never be nonzero simultaneously, therefore if λ_i is in (0,C) then relative ξ is zero. So b could be calculated as follows:

$$b^* = y_i - w^* x_i - \varepsilon \qquad for\ \lambda_i \epsilon (0, C)$$

$$b^* = y_i - w^* x_i + \varepsilon \qquad for\ \lambda_i^* \epsilon (0, C)$$

Till now it was dealt with in input space assuming f(x) is linear, now it can be looked at the nonlinear case briefly. First, there is need to map input space into feature space where data will be mapped into higher dimensional space named kernel space and try to discover a hyperplane in the feature space to achieve higher accuracy. Using Cortes and Vapnik (1995) the trick of kernel functions, one have the following quadratic problem.

$$\max \quad -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\lambda_i - \lambda_i^*)(\lambda_j - \lambda_j^*)K(x_i, x_j)$$

$$-\varepsilon\sum_{i=1}^{n}(\lambda_i + \lambda_i^*) + \sum_{i=1}^{n} y_i(\lambda_i - \lambda_i^*) \qquad 3.13$$

Subject to $\sum(\lambda_i - \lambda_i^*) = 0$

$\lambda_i, \lambda_i^* \ \epsilon(0, C)$

At the optimal solution we have,

$$w^* = \sum_{i=1}^{n}(\lambda_i - \lambda_i^*)K(x_i) \qquad (3.14)$$

And putting the values of (3.14) in (3.6) the general equation becomes,

$$f(x) = \sum_{i=1}^{n}(\lambda_i - \lambda_i^*)K(x_{i,}x) + b \qquad (3.15)$$

Where K(.,.) represent the kernel function, any symmetric semi-definite function, which fulfills the Mercer’s conditions can be used as a kernel function in the SVMs situation Cortes and Vapnik (1995). Different kernel functions can be used to achieve better generalization for specific problem. However one cannot say specific kernel outperform others so some validation techniques can be used to fix good kernel. Commonly used kernel functions are polynomial, gaussian and RBF. For more detail see tutorial on SVM by Smola and Scholkopf (2004) and SVMs for classification and regression (Gunn, 1998).

3.3.1 The Selection of Input Lags and Hyperparameters for SVM model

In literature many authors have selected different input lags depending on the nature and frequency of the data. We have used twelve lags for our analysis as it make sense because of monthly data and would be appropriate to next month forecasts. The selection of hyper parameters pay a crucial role in any analysis of SVM which are sort of kernel function, regularization constant C and the maximum allowable loss function ε. Many authors like (Kumar, 2014; Cao and Tay, 2003; Chen and Wang, 2007) found that gaussian radial basis function is superior to other kernel types as they take more time to train the model and gives adverse results as compared to gaussian radial basis function. Therefore, we have used Gaussian radial basis function because of its better prediction performance. In tuning of SVM model for parameter selection kernel parameter γ and regularization parameter C play an important role of model performance. Improper selection of these parameters can lead to under and overfitting of the trained model as too large value of constant C can cause overfitting of the train data while small values can underfit and vice versa for gamma parameter. So combination of optimal C and γ parameters are chosen through tuning the SVM model which can test several different values and yield the ones which has minimum error for 10-fold cross validation. Where 10-fold cross validation means splitting data randomly into ten equal parts in which each fold is used as testing set at some point. For example in 1st iteration, the first fold used to test the model while remaining are used to train the model and for second fold repetition, 2nd fold is used as testing set and the rest are used as the training set. This procedure is repeated until every fold of the 10 folds have been used as the testing set. In order to attain optimal parameters a suitable range of parameter C and γ is provided to training data while tuning of the model which chose best parameter values based on cross validation. A reasonable value for ε is insensitive to SVM modeling (Kumar, 2014). Now we will follow the discussed process to select the best performing model to make forecasts for our data sets.

3.4 Hybrid Model

Basically hybrid modeling means mixing or combining the two or more different models with their unique characteristics to obtain accuracy and precision in results within a given situation. The inspiration for mixing the models occurs from the consideration that a single model may not be enough to identify the true data generating process or cannot capture entire the features of the time series (Khashei and bijari, 2010), so the idea behind hybrid methodology was to acquire advantages of individual models best traits to gain best possible results (Mojirsheibani, 1999; Paptla *et al.*, 2002).

Let's assume that we have a series y_t consisting of linear and nonlinear components of L_t and N_t given as:

$$y_t = L_t + N_t \qquad (3.16)$$

We run ARIMA model first on given data sets to obtain linear forecasts. Then using the residuals acquired from the ARIMA model are estimated through ANN. After that combining the resulted output from ARIMA and ANN yields the final required output. The same procedure is repeated for ARIMA and SVM. If e_t denotes residuals from the linear model at time t then,

$$e_t = y_t - \hat{L}_t \qquad (3.17)$$

Where $\hat{L}_t$ represents the forecasting value by ARIMA model at time t, and then the residuals of ARIMA are estimated through SVM and ANN as follow:

$$e_t = f(e_{t-1}, e_{t-2}, \dots \dots\ e_{t-n}) + \mathcal{E}_t \qquad (3.17.1)$$

Where f is nonlinear function computed by ANN and SVM and $\mathcal{E}_t$ is random error. The resulted combined forecast is as given below:

$$\hat{y}_t = \hat{L}_t + \hat{N}_t \qquad (3.18)$$

Where $\hat{N}_t$ is the forecast value of equation (3.17.1).

3.5 Data Description

For empirical analysis stock return, exchange rate and inflation data set is used. The detail description of all the variables along their sources are given in Table 3.3.

Table 3.3: Data Description and Source

Variables	Definition	Frequency	Time period	Source
Stock return	$S_t = logp_t - logp_{t-1}$	Monthly	1994-2018	Business Recorder & KSE
Exchange rate	The rate at which currency of two countries can be exchanged. $E_t = \frac{Rs}{Dollar}$	Monthly	1990-2018	SBP Monthly Statistical Bulletin
Inflation	$\pi_t = \frac{CPI_t - CPI_{t-1}}{CPI_{t-1}}$	Monthly	1990-2018	SBP Monthly Statistical Bulletin

CHAPTER 4

RESULT AND DISCUSSION

In this chapter data set comprises of stock return, inflation and exchange rate is used to assess the forecasting ability of discussed techniques in case of Pakistan. Data points from all these variables are divided into two parts in which last 24 observations of total data points characterized as the test set while remaining all initial observations represented as training set. So that authentic out of sample forecasts can be obtained and compared with the actual values. The discussion is started from the statistical properties of data and assumptions of data normality which is preliminary requirement for ARIMA forecasting. It is also useful to make data normal for ANN and SVM to attain more accuracy in forecasting but due their nonparametric properties, it is not considered necessary to make data normal for these models. After discussing the nature of data proposed models will be applied in order to get out of sample forecast.

4.1 Statistical and Graphical Description of Data

Every different variable has distinct historical pattern and statistical characteristics, therefore various techniques perform dissimilarly on these different data sets. Jarque-Bera Test (JB), descriptive statistics and time plot has been used to elaborate the historical pattern and statistical characteristics of three type of data sets such as inflation, stock return and exchange rate. Jarque and Bera (1987) gives two sided test based on the skewness and kurtosis coefficients to check either data came from a normal distribution or not. The null hypothesis for JB test is that the data is from normal distribution and the alternative is that the data is not normal distributed. The statistic of JB test is given as below:

$$JB = n\left(\frac{S^2}{6} + \frac{K^2}{24}\right) \qquad 4.1$$

Where n specifies the sample size, S is the skewness and K is the kurtosis. Statistical description of all data sets is discussed in Table 4.1.

Table 4.1: Data Summary Statistics

Data	Mean	Std. Dev	Skewness	Kurtosis	JB	Probability	Obs
π_t	0.007	0.007	0.441	0.415	13.9	0.001	343
E_t	4.408	0.495	-0.386	-0.933	21.05	0.001	348
ΔE_t	0.005	0.013	1.825	5.932	712	0.001	347
S_t	0.009	0.084	-0.972	4.697	323.1	0.001	294

In Table 4.2 JB test indicate that none of the data set is normal. Inflation data has positive value of skewness its mean data is skewed to right and kurtosis is less than 3 which implying that data distribution is platykurtic. Exchange rate at level with negative skewness value having kurtosis below than 3 depicts non normal distribution of data that is left skewed and have leptokurtic distribution. While exchange rate at 1^{st} difference shows opposite data statistics with positive value of skewness which indicate that data has a lengthy right tail. The distribution with kurtosis greater than 3 is said to be leptokurtic or fat tailed. Stock return also shows abnormal distribution with negative value of skewness which implies that distribution has long left tail or skewed to the left. The kurtosis value greater than 3 is supposed to be leptokurtic distribution. In general mostly time series data give more information than skewness and kurtosis statistics and have more importance in time series analysis.

The line plots of all the data sets are provided in Figure 4.1, 4.2, 4.3 and 4.4 to judge the pattern of data either it shows stationary or non-stationary behavior. The basic concern for time plot is to check the fluctuations and trend in time series data which might be seasonal or nonseasonal.

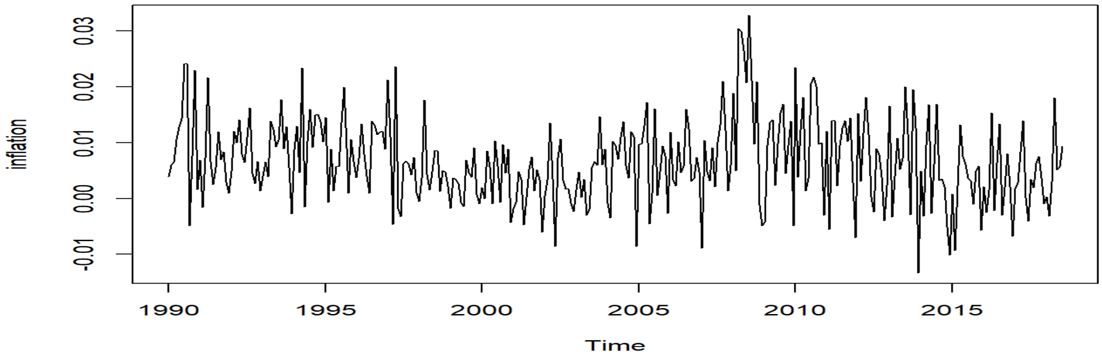


Figure 4.1: Inflation Rate

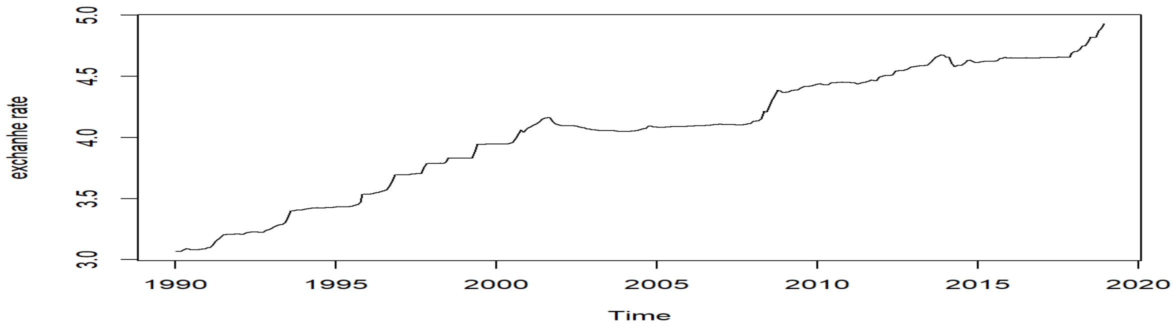


Figure 4.2: Exchange Rate at level

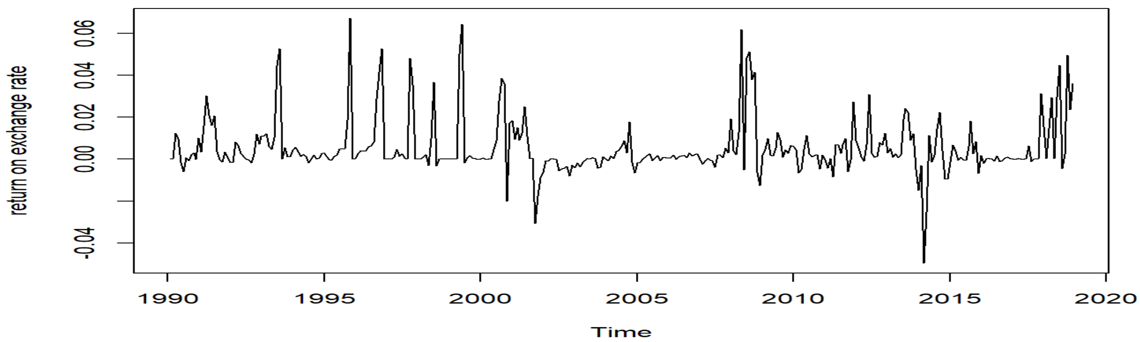


Figure 4.3: Exchange Rate after 1st difference

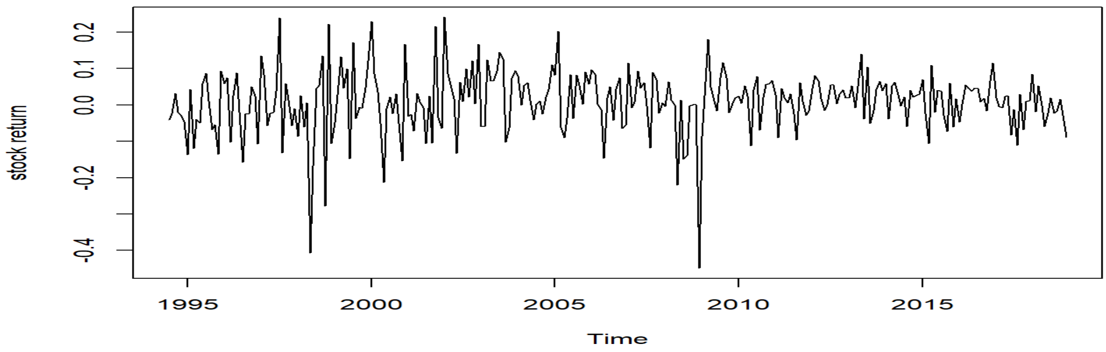


Figure 4.4: Stock return

It can be seen from plots of inflation and stock return in Figure 4.1 and 4.4 respectively that these series are fluctuating around constant mean level. So we consider both of these data sets are stationary with little bit unexpected variation. Also no structural break has been found in inflation but in case of stock return it exhibits little troublesome. Figure 4.2 exhibits a line plot of log of exchange rate series which showing an upward trend and it can be easily deduced that exchange rate is nonstationary at level. Whereas in Figure 4.3 1st difference of log of exchange rate is fluctuating around zero mean level. ADF statistic is further implemented on the series to get the statistical counter part of stationarity and non-stationarity.

Table 4.2: ADF test statistics for all series

Series	Deterministic part	Lags	T_{cal}	Integration order
π_t	None	6	-5.38	$I(0)$
E_t	Drift	7	-2.24	$I(1)$
ΔE_t	None	7	-5.42	$I(0)$
S_t	None	6	-6.08	$I(0)$

As in the Table 4.2 calculated values for inflation, and stock return at level are less than critical values, so we reject the null hypothesis and conclude that series has no unit root at 5% significance level. But in case of exchange rate it has unit root at level because its calculated value is greater than its critical value but after taking the 1st difference it become stationary. Now after data interpretation, next step is modeling of all discussed data sets and we will proceed from ARIMA then go as ANN, SVM and Hybrid models.

4.2 Analysis of series through ARIMA modeling

The analysis of ARIMA modeling is proceeded by following the iterative steps of Box and Jenkins methodology accordingly.

4.2.1 Inflation series

Identification step of ARIMA model starts from autocorrelations (ACF) and partial autocorrelations (PACF) plots of the series which are used to identify the significant lags of the model. Plots for ACF and PACF are given below in Figures 4.5 and 4.6 respectively.

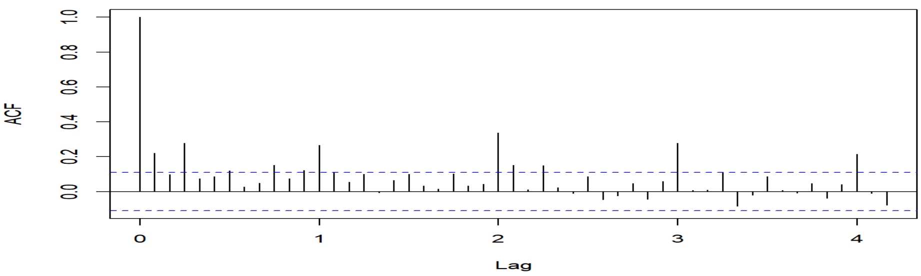


Figure 4.5: ACF of inflation series

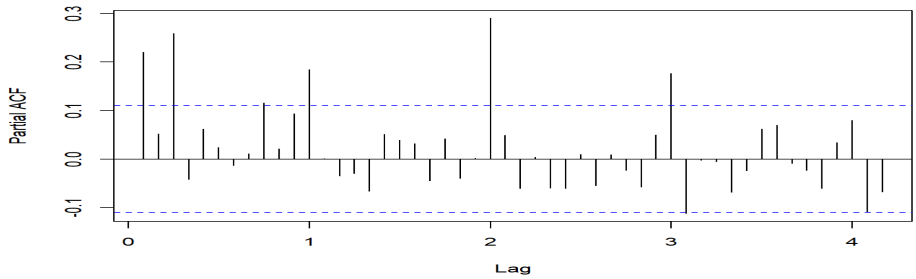


Figure 4.6: PACF of inflation series

Identification of seasonal and nonseasonal lags through given plots of correlogram are made as following. ACF and PACF of inflation series up to lags 50 are presented in

Figure 4.5 and 4.6 respectively. If spikes in ACF and PACF at specific lags lie outside the dotted line which shows $\pm\frac{2}{\sqrt{T}}$ confidence interval, it would be declare as significant lag otherwise not. In order to identify seasonal lags, spikes at seasonal lags will be considered. It can be seen from Figure 4.5 and 4.6 that there is seasonal behavior in the data set of inflation because both plots for ACF and PACF exceeds the significance bound at lags 12, 24, and 36. In addition with ACF which is signicant at 48 lag and also shows significancy to some extent at 13, 25, and 37 lags. Which implies the significant seasnonal autocorrelation and therefore, our model beome seasonal ARIMA.

In case of nonseasonal behaviour it can be seen from plots that spikes at lags1, 3 and 9 exceeds the significant level in both functions. Also it can be noted from plots of ACF and PACF that no other lags are much significant beyond the 9th lag. General model for ARIMA based on correlogram become as $ARIMA(9,0,9)(1,0,2)_S$. Excluding the insignificant lags by using general to specific methodology, the specific ARIMA model is $(9,0,1)(1,0,2)_S$.Subsequent step is to estimate specific equation and after diagnostic checking, it can be used for forecasting. In next step estimation of the selected ARIMA model for inflation is done. The statistics of the estimated model are given below in Table 4.3.

Table 4.3: Model results of inflation

Lags	**Coefficients**	**Standard errors**	**t-statistics**
AR(1)	0.584	0.142	4.112
AR(3)	0.170	0.077	2.208
AR(9)	0.113	0.047	2.404
MA(1)	-0.399	0.160	-2.493
SAR(1)	0.967	0.023	42.043
SMA(1)	-1.008	0.065	-15.508
SMA(2)	0.205	0.068	3.015
σ2 = 3.737e-05, Log Likelihood = 1167.14, AIC = -2318.27			

Where sigma^2 is the value of constant variance which model assumed, the value of log likelihood shows the probability of deriving the data we have observed and AIC represents the performance of selecting the best possible model among others based on its minimum value. Mathematically, the estimated model can be written as:

$$(1-0.584L)(1-0.17L^3)(1-0.113L^9)(1-0.967L^{12})\pi_t$$
$$= (1-0.399L)(1-1.008L^{24})\varepsilon_t \qquad (4.2)$$

Next step is diagnostic checking to discern as there is no information left in error terms (i.e. error terms are white noise).

After the estimation of the model diagnostic checking is made and statistics of applied portmanteau test and JB test are given below in the Table 4.4.

Table 4.4: ARIMA model fit $(9,0,1)(1,0,2)_{12}$ for inflation

Test	**Chi-square**	**Lags**	**P-Value**
JB Test	1.504	-	0.4713
Autocorrelation LM Test	4.477	12	0.9732
Heteroscedasticity LM Test	60.406	12	0.000

JB test (p-value >.05) in Table 5 shows that residual are normally distributed and LM test (p-value>.05) implies that we accept the null hypothesis i.e. there is no autocorrelation among residuals. But LM test for heteroskedasticity has p-value<.05 which indicate that there is problem of heteroscedasticity. Except the LM test for heteroskedasticity all other tests meet the assumptions of having no patterns in the residuals, after diagnose checks the next step is forecasting through estimated model. BDS test is applied to check the nonlinear dependencies among the residual and in case of nonlinear presence, the obtained residuals from ARIMA will be further used for nonlinear forecasting through hybrid models. The statistics of BDS test are given in Table 4.5 as:

Table 4.5: BDS test results for residuals of inflation series

m	p-values	z-statistics
2	0.026	2.226
3	0.000	3.816
4	0.000	4.193
5	0.000	4.690

Table 4.5 for all embedding dimension levels show that the null hypothesis of iid is rejected and there direct evidence of nonlinear dependence in inflation series. The 24 months ahead forecast for inflation series are done from ARIMA$(9, 0, 1)(1, 0, 2)_{12}$. 1 to 24 period ahead forecasts and actual values on different forecast horizon with their errors and absolute errors (AE) are given below in Table 4.6.

Table 4.6: Actual versus forecasted values of inflation from ARIMA

Months	Actual values	Forecast values	Error	AE
1	-0.00296	0.00833	-0.01129	0.01129
4	0.002082	0.00168	0.00040	0.00040
8	0.008369	0.00660	0.00176	0.00176
12	0.003448	0.00966	-0.00621	0.00621
16	0.003633	0.00249	0.00115	0.00114
20	0.003008	0.00537	-0.00236	0.00236
24	0.009319	0.00984	-0.00052	0.00052

Above Table 4.6 consist of forecasts and actual values with their corresponding errors, which shows that ARIMA model provided relatively good forecasts. Now we check how graphical representation of forecasted and actual values looks like which are given as in Figure 4.7:

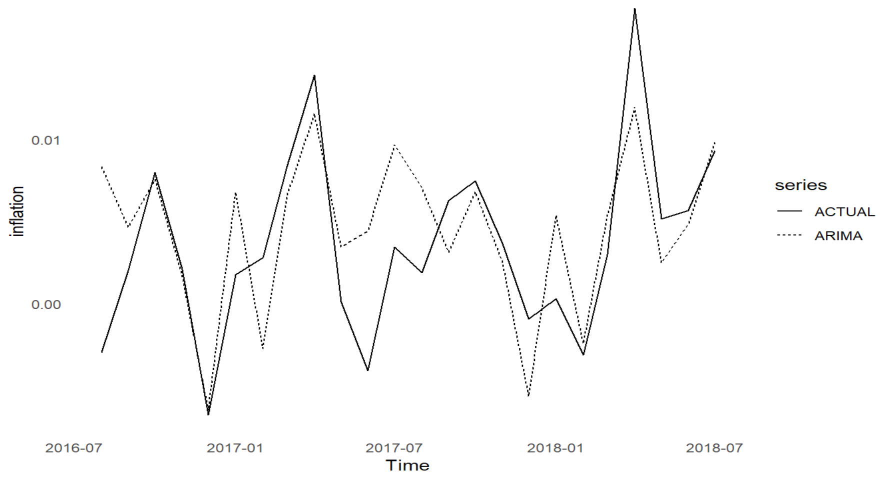


Figure 4.7: Actual versus forecasted values of inflation by ARIMA

In Figure 4.7 dash line shows forecast values whereas solid line represents actual values. The graphical representation indicate that selected ARIMA model tried its best to capture the actual values.

4.2.2 Return on exchange rate

For Identification of the model, plots for ACF and PACF of exchange rate at 1st difference known as return on exchange rare are given below in Figure 4.8 and 4.9 respectively. Which are used for model identification purpose.

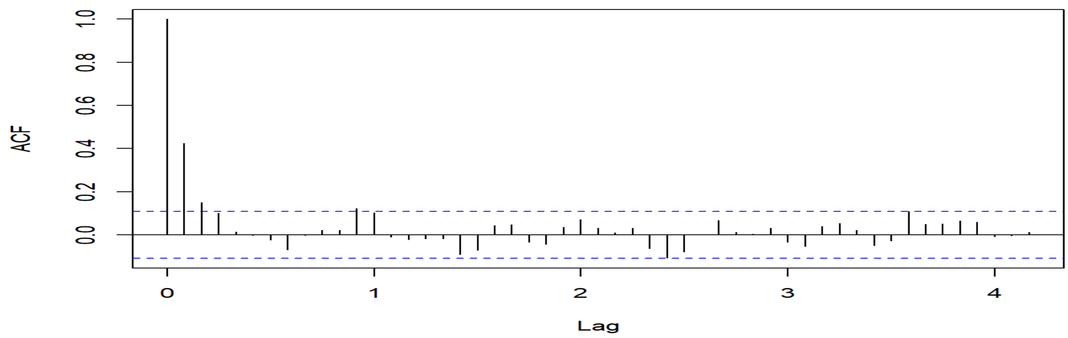


Figure 4.8: ACF of return on exchange rate

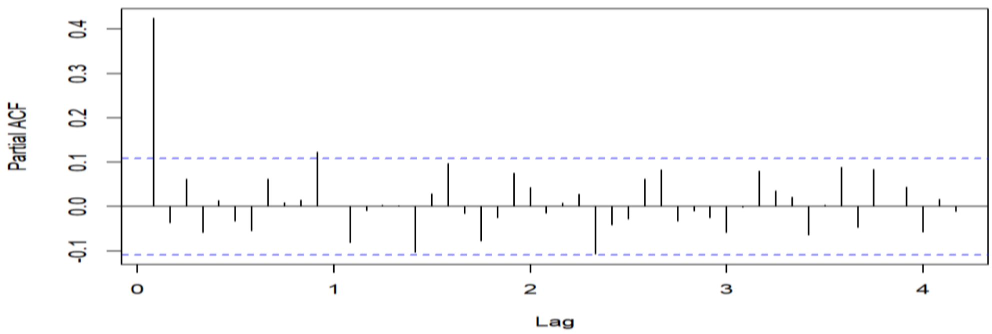


Figure 4.9: PACF for return on exchange rate

ACF and PACF of return on exchange rate series up to 50 lags are presented in Figure 4.8 and 4.9 respectively for Identification of seasonal and nonseasonal lags. As discussed earlier in case of inflation series that dotted line shows $\pm\frac{2}{\sqrt{T}}$ confidence interval, So any specific lag spikes either it would be seasonal or nonseasonal exceeds from this line will be considered significant. It can be seen from Figure 4.8 that spikes at nonseasonal lags 1, 2 and 11 and one seasonal lag at 12 seems to be significant. Whereas in case of PACF in Figure 4.9, it can be seen that only nonseasal spikes at lag 1 and 11 shows significant behaviour. According to the correlogram our general model for ARIMA will become as $ARIMA(11,0,11)(0,0,1)_S$ with intercept. Excluding the insignificant lags one by one the remaining significant lags are estimated in next step of estimation. The next stage is estimation of the selected model for return on exchange rate series. Table 4.7 consist of the statistics of the estimated specific model which is given as below.

Table 4.7: Model results of return on exchange rate

Lags	Coefficients	Standard errors	t-statistics
Intercept	0.005	0.001	4.08
AR(1)	0.422	0.049	8.67
AR(11)	-0.348	0.126	-2.78
MA(11)	0.478	0.131	3.65
SMA(1)	0.212	0.070	3.02
σ^2 = 0.00013: log likelihood = 983.85, AIC = -1955.69			

Mathematical equation for chosen ARIMA model can be written as:

$$(1 - 0.422L)(1 + 0.348L^{11})E_t$$
$$= 0.005 + (1 + 0.478L^{11})(1 + 0,212L^{12})\varepsilon_t \qquad (4.3)$$

Next step is diagnostic checking to determine as there is no information left in error. The test statistics of residuals for estimated model are given below in Table 4.8 as:

Table 4.8:ARIMA Model Fit $(\mathbf{11}, \mathbf{0}, \mathbf{11})(\mathbf{0}, \mathbf{0}, \mathbf{1})_{\mathbf{12}}$ for return on exchange rate

Test	Chi-square	Lags	P-Value
JB test	1.504		0.0001
Autocorrelation LM Test	4.477	12	0.9222
Heteroscedasticity LM Test	60.406	12	0.0001

From Table 4.8, it can be seen that there is no autocorrelation among residuals but it has the problem of heteroskedasticity and error terms are not normally distributed. Further BDS test is applied to check the nonlinear dependencies among the residual. The statistics of BDS test are given in Table 4.9 as:

Table 4.9: BDS test results for residuals of return on exchange rate

$\boldsymbol{m}$	p-values	z-statistics
2	0.000	10.06
3	0.000	9.905
4	0.000	9.673
5	0.000	4.690

BDS test results from Table 4.9 clearly suggest that reject the null hypothesis that series are linearly dependent. As many pairs of specific models was went through diagnostic check but the selected $ARIMA(11,0,11)(0,0,1)_{12}$ model remained best among all others, next step is of forecasting.

1 to 24 months ahead forecasts are done, forecast against actual values with their respective errors and absolute errors are given in Table 4.10 as:

Table 4.10: Actual versus forecasted values of return on exchange rate by ARIMA

Months	Actual values	Forecast values	Error	AE
1	0.00005	0.00422	-0.00417	0.00417
4	0.00005	0.00486	-0.00480	0.00480
8	-0.00111	0.00437	-0.00548	0.00548
12	0.03116	0.00506	0.02610	0.02609
16	0.02948	0.00442	0.02505	0.02505
20	-0.00454	0.00451	-0.00906	0.00906
24	0.03596	0.00499	0.03098	0.03098

Above Table contains the forecasts and actual values with their corresponding errors, which shows its error increases as forecasts horizon increases from 1 to 24. The graphical representation of actual and forecasted values is given as Figure 4.10:

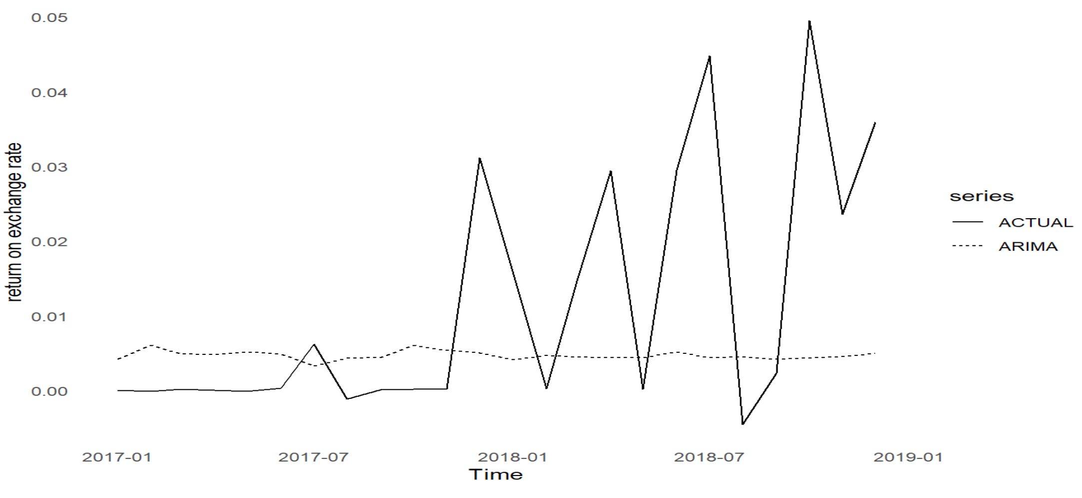


Figure 4.10: Actual versus forecasted values of return on exchange rate by ARIMA

In Figure 4.10 dash line shows forecast values and solid line shows actual values. The graphical representation indicate that selected ARIMA model do not capture actual direction of real values. Because straight solid line shows that its forecast values almost remain consistent over the period from 1 to 24.

4.2.3 Stock return series

Plots for ACF and PACF of exchange rate at level are given below in Figure 4.11 and 4.12 respectively, which are used for model identification.

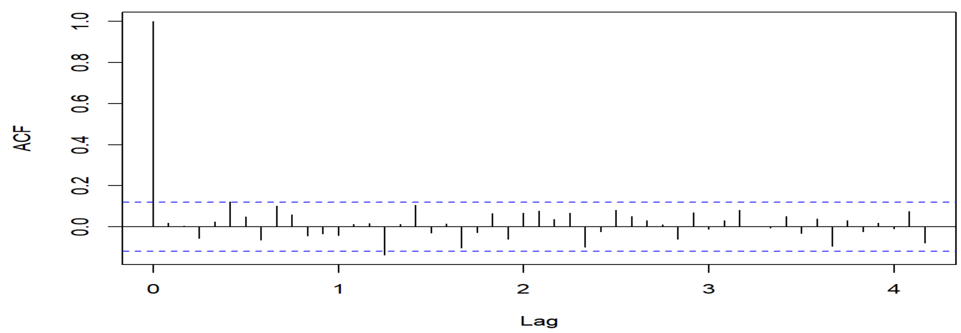


Figure 4.11: ACF for stock return

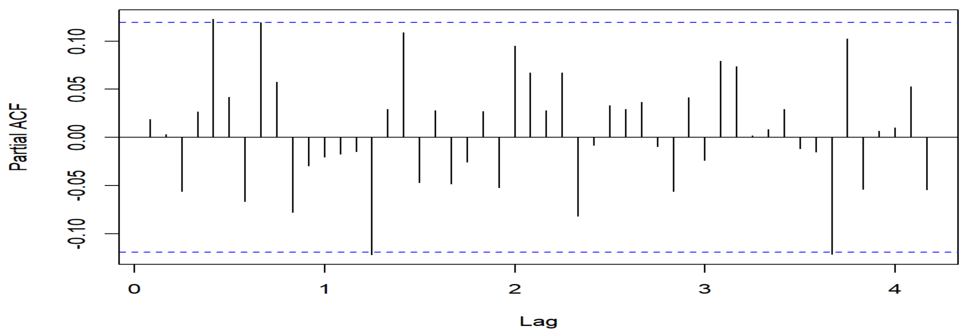


Figure 4.12: PACF for stock return

For the identification of seasonal and nonseasonal lags, plots for ACF and PACF of stock return series up to 50 lags are presented in Figure 4.11 and 4.12 respectively. It can be seen from Figure 4.11 that spikes at nonseasonal lags 5 and 15 exceeds the significant bound. While in case of PACF in Figure 4.12, it can be seen that nonseasal spikes at lag 5, 8 and 15 are seems to be significant. According to the correlogram our general model of ARIMA become as $ARIMA(15,0,15)$ with intercept and removing the insignificant lags one by one the specific model is given as $ARIMA(15,0,0)$ with intercept.

In next step estimation of the selected ARIMA model for stock return is done and the statistics of the estimated specific model are given below in Table 4.11.

Table 4.11: Model results of stock return

Lags	Coefficients	Standard errors	t-statistics
Intercept	0.011	0.005	2.46
AR(15)	-0.137	0.06	-2.82
σ^2 = 0.0074: log likelihood = 279.02, AIC =-552			

It can be written in mathematical equation form as:

$$S_t = 0.011 - S_{t-15} + \varepsilon_t \qquad (4.4)$$

Subsequent step is diagnostic checking to determine as there is no information left in error terms. The test statistics of residuals for estimated model are given below in Table 4.12 as:

Table 4.12: ARIMA Model Fit (15, 0, 0) with intercept on stock return

Test	Chi-square	Lags	P-Value
JB test	238.71	-	0.0001
Autocorrelation LM Test	18.85	12	0.54
Heteroscedasticity LM Test	15.88	12	0.2

From Table 4.12, it can be noted that there is no problem of serial correlation and heteroskedasticity as well in the selected model. But the only problem which could not be fixed during choosing the specific model is the normal distribution of error terms.

Now BDS test is applied to check the nonlinear dependencies among the residual of stock return obtained through ARIMA model. The statistics of BDS test are given in Table 4.13 as:

Table 4.13: BDS test results for residuals of stock return

m	p-values	z-statistics
2	0.003	2.959
3	0.002	3.134
4	0.000	3.368
5	0.000	4.846

From Table 4.13, it can be concluded that stock return series is not linearly dependent. Now forecasting of the stock return are done from 1 to 24 months ahead by selected model after diagnostic check. Forecasts and actual values from 1 to 24 period ahead on different forecasts horizon with their errors and absolute errors are given below in Table 4.14.

Table 4.14: Actual versus forecasted values of stock return by ARIMA

Months	Actual values	Forecast values	Error	AE
1	0.01969	0.00474	0.01495	0.01495
4	0.02349	0.01934	0.00416	0.00416
8	-0.11026	0.00767	-0.11793	0.11793
12	0.01145	0.01037	0.00109	0.00109
16	-0.00157	0.01221	-0.01378	0.01378
20	-0.022976	0.01114	-0.03412	0.03412
24	-0.088486	0.01198	-0.10046	0.10046

Above Table contains the forecast and actual values with their corresponding errors, it can be seen from Table 4.14 that overall error increased as forecasts horizon move from

1 to 24. The graphical representation of actual and forecasted values are given as in Figure 4.13:

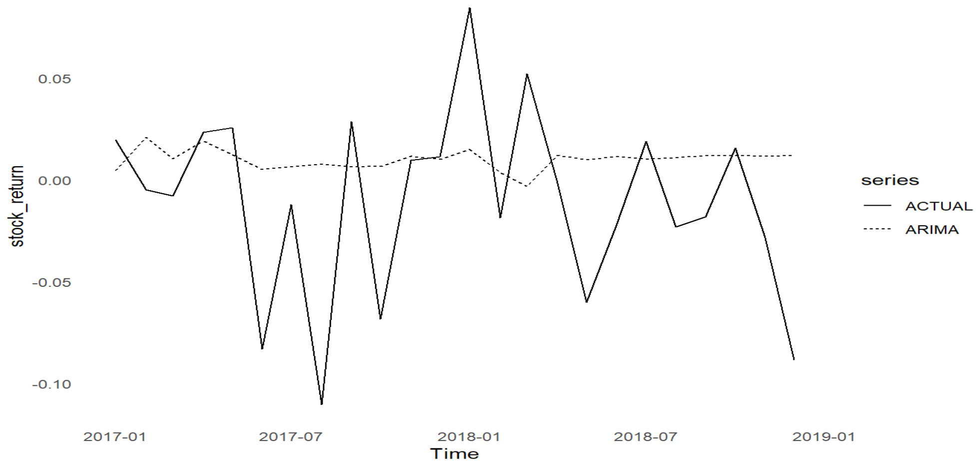


Figure 4.13: Actual versus forecasted values of ARIMA for stock return

In Figure 4.13 dashed line shows forecast values and solid line shows actual values. The graph shows straight line of forecasts values on average represents the real values but do not succeed to capture direction of the return.

4.3 Analysis of all Series through ANN Modeling

The analysis of ANN modeling is proceeded by following the discussed ANN methodology in section 3.2 accordingly.

4.3.1 ANN Modeling for Inflation

The plot of ANN architecture is given in Figure 4.14 for inflation series which was selected after a careful consideration with minimum MSE (0.000) by following all discussed steps in section 3.2.2:

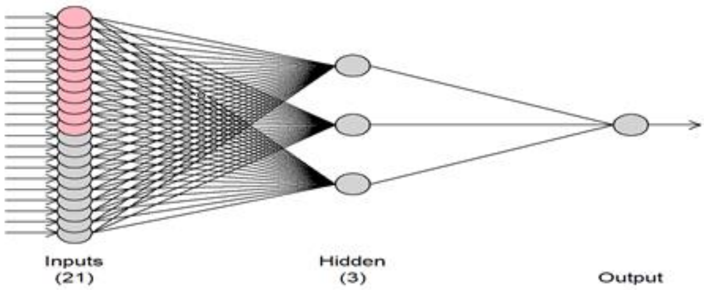


Figure 4.14: ANN Architecture for inflation

The diagram of multilayer perceptron indicate 21 input variables in which 10 nodes represents autoregressions or lags and 11 nodes shows deterministic dummies with 3 hidden nodes and 100 repetitions. Mathematically, it can be written as:

$$\pi_t = w_0 + \sum_{j=1}^{3} w_j . g\,(w_{0.j} + w_j(w_1\pi_{t-1} + w_2\pi_{t-2} + w_3\pi_{t-3} + w_4\pi_{t-4} + w_5\pi_{t-5} + w_6\pi_{t-6} + w_7\,\pi_{t-7} + w_8\pi_{t-8} + w_9\pi_{t-9} + w_{12}\pi_{t-12} + \sum_{i=1}^{11} w\delta_i D_i)) + e_t \quad (4.6)$$

Equation (4.6) is the specific equation of (3.5) general equation for ANN modeling which shows the number of inputs and hidden nodes used in estimation. Next step is of future forecasts and it can be done by using selected best suited model. The forecast values from 1 to 24 points ahead are given in Table 4.15 on different steps with their error and absolute errors.

Table 4.15: Actual versus forecasted values of inflation by ANN

Months	**Actual values**	**Forecast values**	**Error**	**AE**
1	-0.00296	0.003928	-0.00688	0.006884
4	0.002082	0.002698	-0.00062	0.000615
8	0.008369	0.005752	0.002617	0.002617
12	0.003448	0.008345	-0.0049	0.004896
16	0.003633	0.003225	0.000408	0.000408
20	0.003008	0.005595	-0.00259	0.002587
24	0.009319	0.009536	-0.00022	0.000217

Above Table contains the forecasts and actual values with their respective errors which shows ANN forecasts are good based on errors as errors are not much wide, visualization of these values are given in Figure 4.15 to check the gap between actual and forecasts values.

Figure 4.15: Actual versus forecasted values of inflation through ANN

Figure 4.15 represents the dashed line of forecast values while solid line indicating the actual values. It can be seen from Figure that ANN model remain good in capturing the direction of inflation series and giving overall good performance in forecasting.

4.3.3 ANN Modeling for Return on Exchange Rate

The plot of ANN architecture is given in Figure 4.16 for return on exchange rate which was selected after a careful consideration with minimum MSE (0.0001) by following all discussed steps:

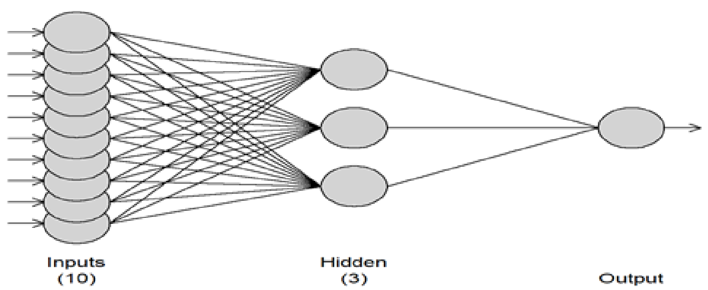


Figure 4.16: ANN Architecture for return on exchange rate

The diagram of multilayer perceptron indicate 10 input variables that represents autoregressions or lags with 3 hidden nodes and 100 repetitions. Mathematically, it can be written as:

$$E_t = w_0 + \sum_{j=1}^{3} w_j . g\ (w_{0.j} + w_j(w_1 E_{t-1} + w_2 E_{t-2} + w_3 E_{t-3} + w_4 E_{t-4} + w_5 E_{t-5} + w_6 E_{t-6} + w_7 E_{t-7} + w_8 E_{t-8} + w_9 E_{t-9} + w_{10} E_{t-10} + e_t \quad (4.7)$$

Equation (4.7) is the specific form of general equation of (3.5) which is extracted for the estimation of return on exchange rate through ANN. The next step is of forecasting for return on exchange rate through ANN modeling.

1 to 24 step ahead forecasts are done while Table 4.16 possess the actual and forecasts values of return on exchange rate on different steps with their corresponding errors and absolute errors.

Table 4.16: Actual versus forecasted values of return on exchange rate by ANN

Months	Actual values	Forecast values	Error	AE
1	0.00005	0.00178	-0.00173	0.00173
4	0.00006	0.00559	-0.00554	0.00554
8	-0.00111	0.00877	-0.00988	0.00988
12	0.03116	0.01227	0.01889	0.01889
16	0.02948	0.01591	0.01357	0.01357
20	-0.00454	0.01974	-0.02428	0.02428
24	0.03596	0.02332	0.01265	0.01265

Above Table 4.16 contains the forecasts and actual values with respective forecast errors, it can be seen from Table that error increases with successive forecasts. The visualization of actual versus forecasted values are given in below in Figure 4.17. Which further portrait the gap among actual and forecasts values.

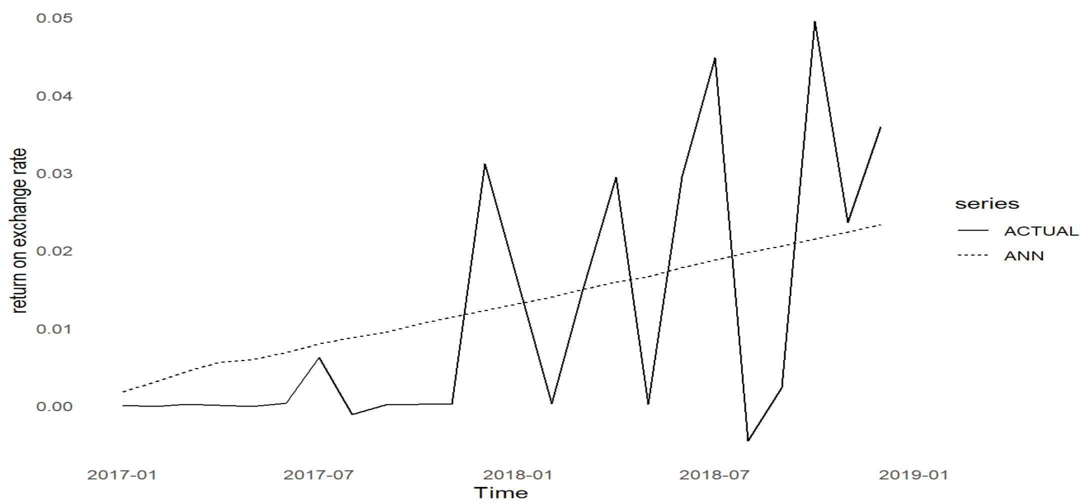


Figure 4.17: Actual versus forecasted values of return on exchange rate by ANN

Figure 4.17 shows that ANN model make a upward regression line that indicates it just averages the actual values. It can be concluded from Figure 4.17 that ANN do not capture the direction of actual values which means it produced the poor forecasts for exchange rate.

4.3.3 ANN Modeling for Stock Return

The plot of ANN architecture is given in Figure 4.18 for stock return series which has minimum MSE (0.0075) by following all discussed steps:

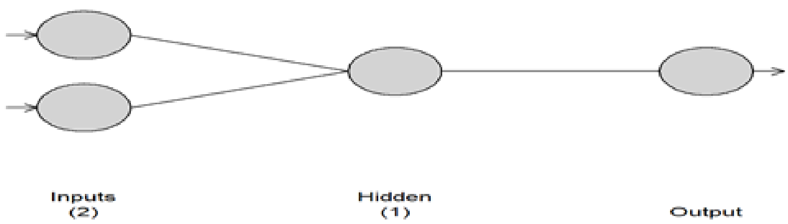


Figure 4.18: ANN Architecture for stock return

The diagram of multilayer perceptron is given with 2 input variables that includes 5th and 8th lags with 1 hidden nodes and 100 repetitions. Mathematically, it can be written as:

$$S_t = w_0 + \sum_{j=1}^{1} w_j . g\ (w_{0.j} + w_j (w_5 S_{t-5} + w_8 S_{t-8}) + e_t \quad (4.8)$$

Forecasts are made from 1 to 24 points ahead and Table 4.17 displays the actual and forecasts values on different steps.

Table 4.17: Actual versus forecasted values of stock return by ANN

Months	Actual values	Forecast values	Error	AE
1	0.01969	0.01535	0.00433	0.00434
4	0.02350	0.01930	0.00420	0.00420
8	-0.11026	0.02683	-0.13709	0.13709
12	0.01146	0.01486	-0.0034	0.00340
16	-0.00157	0.01454	-0.01611	0.01611
20	-0.02298	0.01374	-0.03672	0.03672
24	-0.08849	0.01358	-0.10208	0.10208

Above Table 4.17 contains the forecasts and actual values with respective forecast errors, which shows that size of errors increased as forecasts move from 1 to 24 steps ahead. Visualization of these values are given in Figure 4.19 to check the gap between actual and forecasts values.

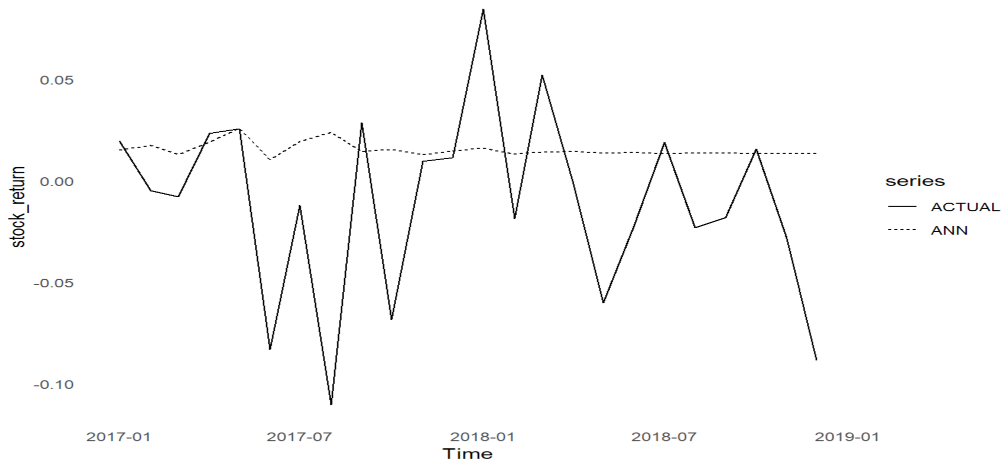


Figure 4.19: Actual versus forecasted values of stock return through ANN

Figure 4.19 depicts that ANN model make a straight regression line that indicates it took average of the actual values. Figure 4.19 represents that ANN model did not perform well in case of stock return series.

4.4.1 SVM Modeling on Inflation

Tuning of SVM parameters are done according to discussed procedure in section 3.4. The optimum selection of parameters is made by tuning the SVM model by providing a range of values to parameters based on 10 fold cross validation. Then that model is selected which made best performance based on minimum training mean square error (MSE). The graph for tuning of optimal hyper parameters is given as in Figure 4.20:

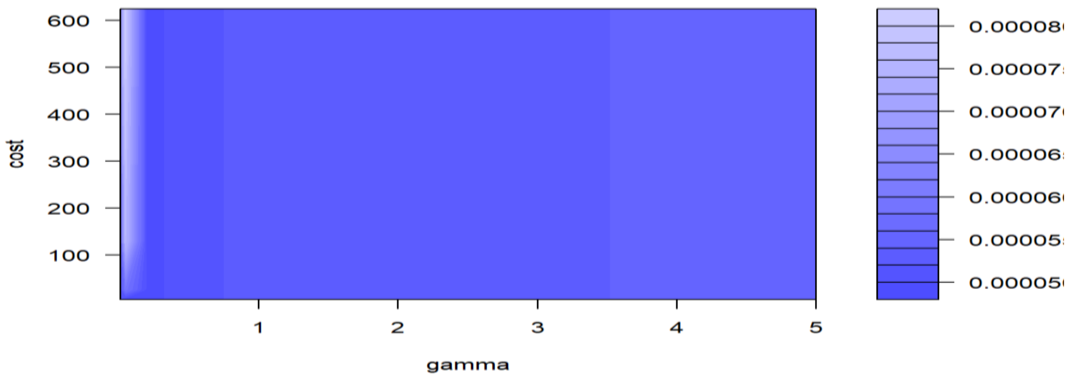


Figure 4.20: Performance of SVM model on inflation series

Figure 4.20 shows the performance of different models by using color coding in which the most darker region indicates the best model. Mean square error is exhibited on the right side of the Figure 4.20 by legend. One can again tune the model by narrowing rang on any darkest region in the plot to obtain further accuracy and lowest MSE. The best optimal values of parameters which obtained by this tuning are c=5, γ=0.04 and ε=0.1 on which SVM model give minimum training error for inflation data set. After selection of suitable SVM model, the next step is forecasting from that selected model.

The forecast values of 1 to 24 steps ahead are obtained by SVM model whose performance remained best among tuning. Actual and forecasts values on different steps are given along with their errors and absolute errors in Table 4.18 as following:

Table 4.18: Actual versus forecasted values of inflation by SVM

Months	Actual values	Forecast values	Error	AE
1	-0.00296	0.00248	-0.00543	0.00543
4	0.00208	0.00352	-0.00144	0.00144
8	0.00837	-0.00105	0.00942	0.00942
12	0.00345	0.00869	-0.00525	0.00525
16	0.00363	0.00479	-0.00116	0.00117
20	0.00301	0.00513	-0.00213	0.00213
24	0.00932	0.01264	-0.00332	0.00332

This Table shows how error and absolute error between actual versus forecasts values shrink and widened among different time horizons. Now the graphical representation in Figure 4.21 will show further detailed of gap among actual and forecasts values over the time horizon. In which dashed line represents the SVM forecasts values while solid line indicate actual values.

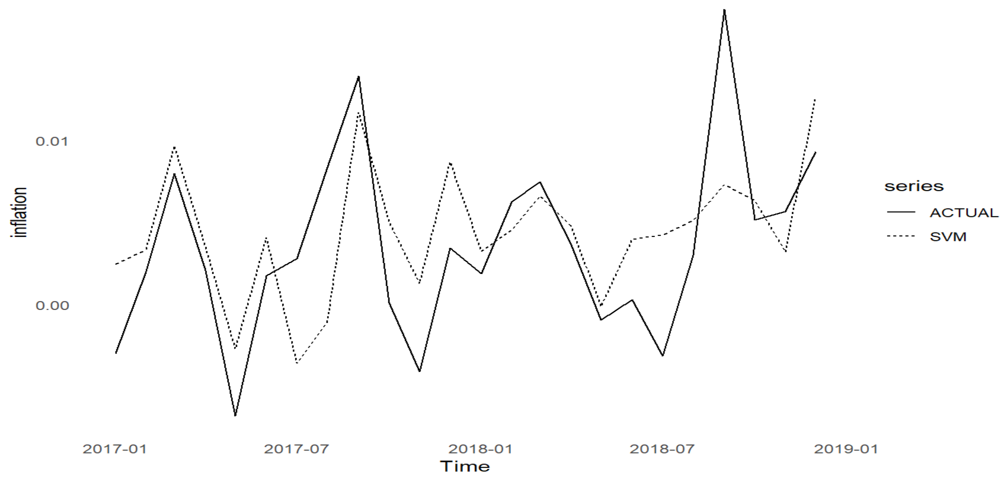


Figure 4.21: Actual versus forecasted values of inflation through SVM

Figure 4.21 shows that SVM model almost remain good in capturing the direction of inflation series and provided overall good performance in forecasting.

4.4.2 SVM modeling on return on exchange rate

The graph for SVM parameters performance on different range of values while tuning the model is given below in Figure 4.22.

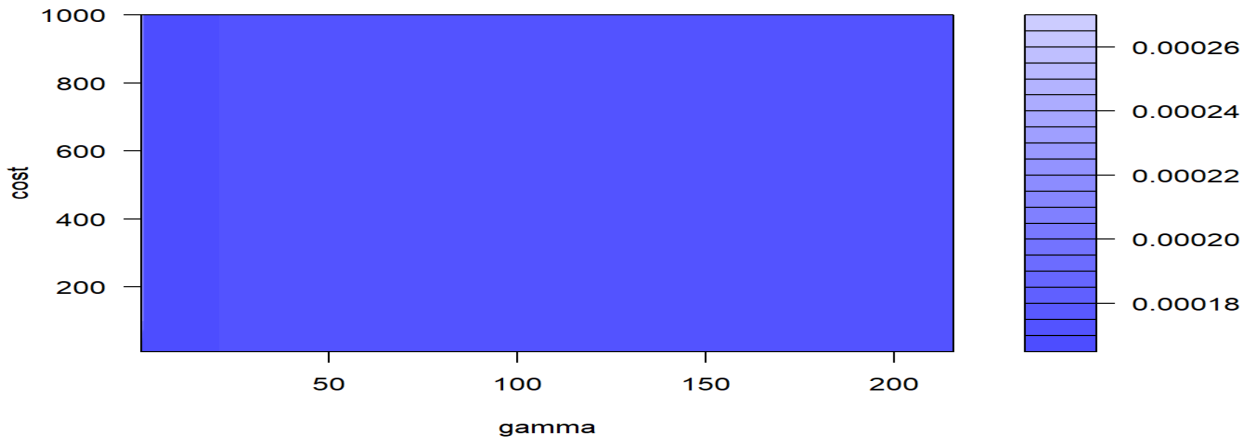


Figure 4.22: Performance of SVM model on return on exchange rate series

Above graph 4.22 shows the tuning performance of SVM model which select the optimal values for hyperparameters based on 10 fold cross validation. The finest values of cost and gamma were obtained by narrowing the rang according to darkest region of the plot. The more darker region means the more accurate model would be obtained by lowest MSE. The best chosen optimal parameters for return on exchange rate are c=10, γ=0.17 and ε=0.1 on which SVM model give minimum training error. Once best model is selected while tuning then forecasts can be done through this model.

Forecasts are done by the best SVM model and their values on different time periods are compared with actual values. Actual values and forecasts values on different forecast steps with their corresponding errors are given below in Table 4.19:

Table 4.19: Actual versus forecasted values of return on exchange rate by SVM

Months	**Actual values**	**Forecast values**	**Error**	**AE**
1	0.00005	0.00165	-0.0016	0.0016
4	0.00005	0.00118	-0.00113	0.00113
8	-0.00111	-0.00143	0.00032	0.00032
12	0.0312	0.00008	0.0311	0.03108
16	0.0295	0.0204	0.00904	0.00904
20	-0.00454	0.0132	-0.0177	0.01772
24	0.036	0.0063	0.0297	0.02966

Above Table 4.19 contains the actual and forecast values on different time horizons with their respective errors. Table shows forecast errors relatively increased from beginning to end. Now below given Figure 4.23 will depict the gap between actual and forecasts values over the forecasts horizon.

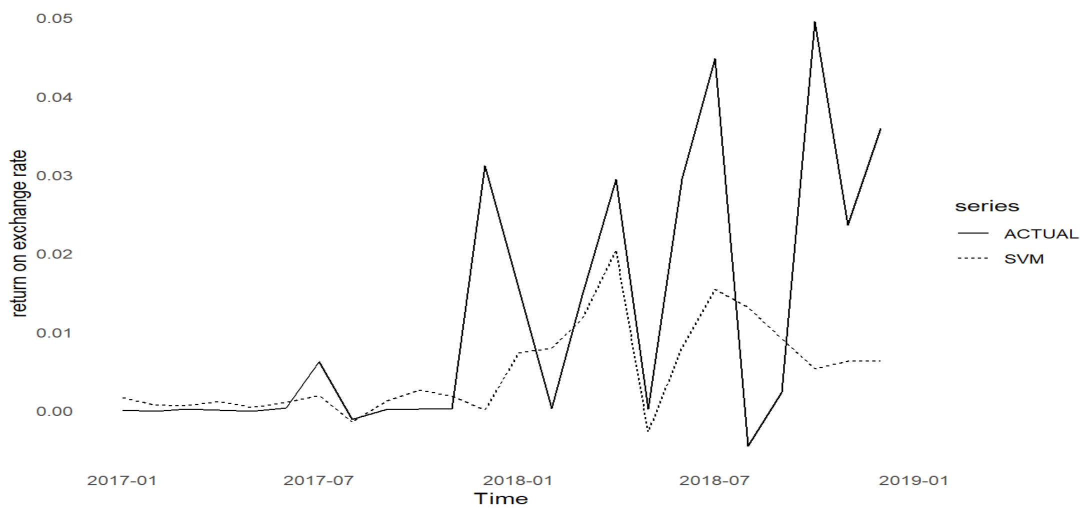


Figure 4.23: Actual versus forecasted values of return on exchange rate by SVM

Figure 4.23 shows SVM model tried to capture the trend of return on exchange rate which remain good in beginning but become poor as moving forward from 1 to 24 points ahead.

4.4.3 SVM Modeling on Stock Return

The best optimal parameters are c=10, γ=0.2 and ε=0.1 for stock return on which SVM model give minimum training error. The graph for SVM parameters performance on different range of values while tuning the model is given in Figure 4.24 as:

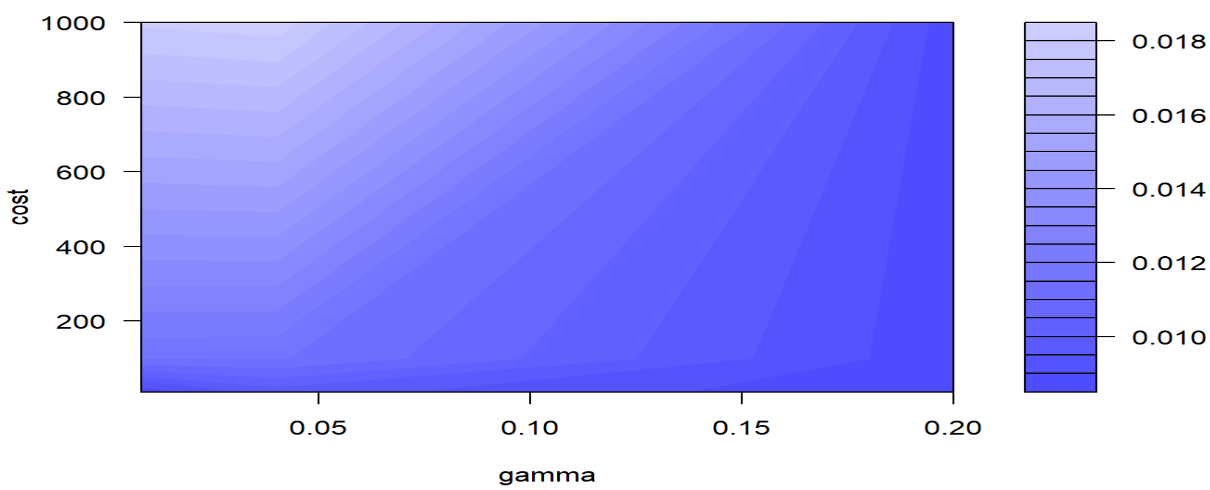


Figure 4.24: **Performance of SVM model on stock return series**

Above graph 4.24 shows the tuning performance of SVM model on different values which selects the optimal values for hyperparameters based on 10 fold cross validation. It can be seen in the Figure 4.24 that value of gamma remain round about 0.20 for different ranges of cost function which providing darkest blue region throughout the plot.

Forecasts by chosen best SVM model and their values on different time periods are compared with actual values. Actual values and forecasts values with their corresponding errors are given below in Table 4.20:

Table 4.20: Actual versus forecasted values of stock return by SVM

Months	Actual values	Forecast values	Error	AE
1	0.01969	0.03363	-0.01394	0.01394
4	0.02349	0.00061	0.02289	0.02286
8	-0.11026	0.06734	-0.17761	0.17761
12	0.01146	-0.04845	0.05991	0.05991
16	-0.00157	0.01669	-0.01826	0.01826
20	-0.02298	0.02564	-0.04862	0.04862
24	-0.08849	-0.00206	-0.08643	0.08643

Above Table 4.20 contains the actual and forecast values on different time horizons with their respective errors which shows error are increased as moving from 1 step to 24 step. Below Figure 4.25 shows its graphical representation.

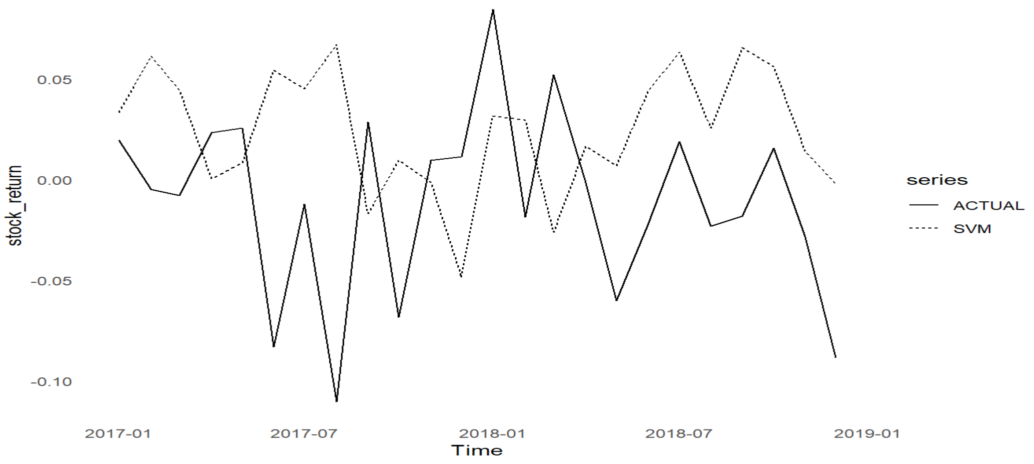


Figure 4.25: Actual versus forecasted values of stock return through SVM

In Figure 4.25 dashed lines depicts forecast values while solid line represents actual values. Figure shows SVM model tried to capture the direction of stock return to some extent but overall its performance remain poor in case of stock return.

4.5 Analysis of all Series by Hybrid Models

Our fundamental goal of hybrid modeling is to improve the forecast accuracy by using the unique features of both linear and nonlinear models as discussed earlier in chapter no 3. Because it is considered that none of the model can perform well in all circumstances, therefore proposed hybrid approach with both linear and nonlinear qualities can give good alternative to time series forecasts. Zhang's (2003) strategy is used to combine the proposed models.

4.5.1 Analysis of all Series through Hybrid ARIMA-ANN

As hybrid model is a sum of linear forecasts from ARIMA and forecasts of its residuals from ANN. The linear forecasts from ARIMA model were provided under section 4.2, now in order to obtain the nonlinear forecasts of its residuals the procedure of ANN is as followed.

4.5.1.1 Estimation of Inflation by Hybrid ARIMA-ANN

The ANN architecture for residuals of ARIMA from inflation data set with 0 MSE is given in Figure 4.26 as:

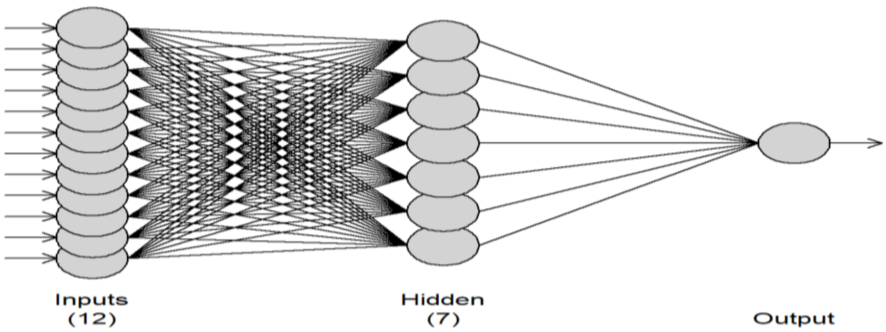


Figure 4.26: ANN Architecture for residuals of inflation

12 input lags with 7 hidden nodes and 100 repetitions are used for residuals forecasts. Mathematically, it can be written as:

$$e_t = w_0 + \sum_{j=1}^{7} w_j . g\ (w_{0.j} + w_j (\sum_{i=1}^{12} w_i e_{t-i}) + \mathcal{E}_t \qquad (4.9)$$

Where equation (4.9) represents the inputs and hidden nodes used for the forecasting of residuals from ARIMA. Forecasts of residual are obtained from fitted ANN model, now combining the both forecasts from ARIMA and ANN which would have equation as:

$$\hat{y}_t = \hat{L}_{t,ARIMA} + \hat{N}_{t,ANN} \qquad (4.10)$$

Actual versus forecasts values on different horizons are given in following Table 4.21 with their respective errors and absolute errors.

Table 4.21: Actual versus forecasted values of inflation by ARIMA-ANN

Months	Actual values	Forecast values	Error	AE
1	-0.00297	0.00247	-0.00542	0.00542
4	0.00208	0.00512	-0.00304	0.00304
8	0.00837	0.00612	0.00225	0.00225
12	0.00345	0.00955	-0.0061	0.00610
16	0.00363	0.00365	-0.00002	0.00002
20	0.00301	0.00482	-0.00181	0.00181
24	0.00932	0.00979	-0.00047	0.00047

It can be concluded from the Table 4.21 that hybrid model of ARIMA-ANN perform well for inflation series as corresponding errors are not huge. The graphical illustration of actual versus forecasted values are given below in Figure 4.27 as:

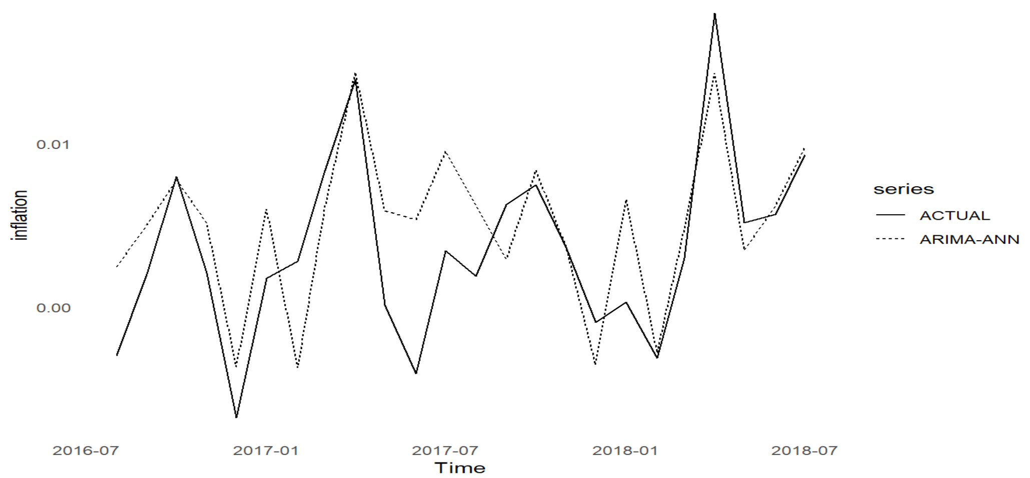


Figure 4.27: **Actual versus forecasted values of inflation from ARIMA-ANN**

It can be seen from Figure 4.27 that hybrid ARIMA-ANN model provided the good forecasts of inflation and it captured the direction of inflation well.

4.5.1.2 Estimation of return on exchange rate by Hybrid ARIMA-ANN

The ANN architecture for residuals of ARIMA from return on exchange rate having MSE 0.0001 is given below in Figure 4.28.

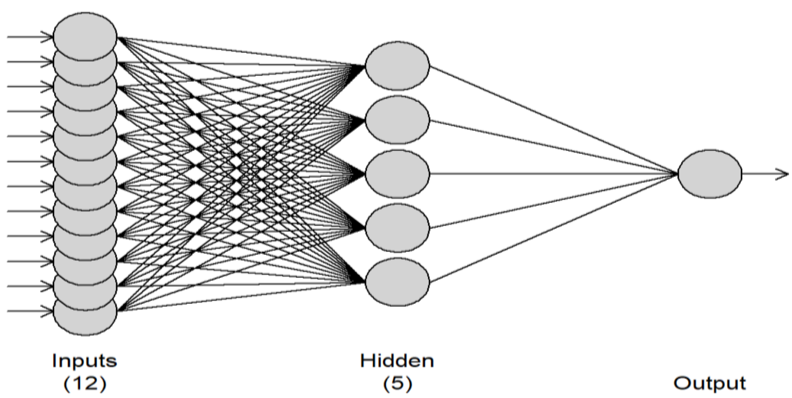


Figure 4.28: ANN Architecture for residuals of return on exchange rate

12 input lags with 5 hidden nodes and 100 repetitions are used for residuals forecasts. Mathematically, it can be written as:

$$e_t = w_0 + \sum_{j=1}^{5} w_j . g\left(w_{0.j} + w_j\left(\sum_{i=1}^{12} w_i e_{t-i}\right) + \mathcal{E}_t\right) \quad (4.11)$$

The next step is of combined forecasting in which forecasts of residual obtained from ANN model are combined with the forecasts values of ARIMA, here actual versus forecasts values are given in Table 4.22 on different steps with their errors:

Table 4.22: Actual versus forecasted values of return on exchange rate by ARIMA-ANN

Months	Actual values	Forecast values	Error	AE
1	0.00005	0.00367	-0.00362	0.00362
4	0.00005	0.00618	-0.00613	0.00613
8	-0.00111	0.00559	-0.00671	0.00671
12	0.031159	0.00817	0.02299	0.02299
16	0.029478	0.00853	0.02095	0.02095
20	-0.00454	0.01023	-0.01478	0.01478
24	0.035964	0.01199	0.02397	0.02397

Table 4.22 shows forecast errors are small in beginning but become large as the time horizon increased. Figure 4.29 comprises of graphical depiction of actual versus forecasted values as:

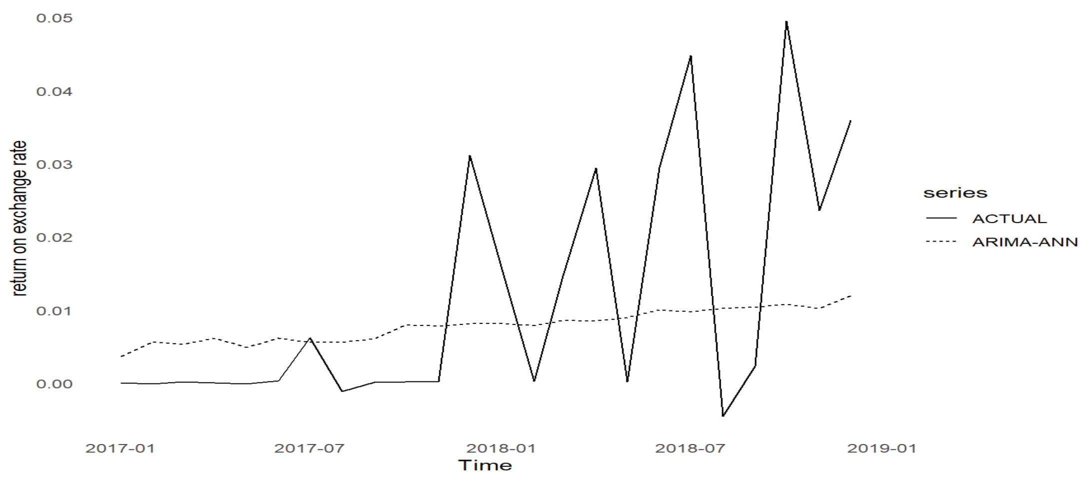


Figure 4.29: Actual versus forecasted values of return on exchange rate from hybrid ARIMA-ANN

It can be seen from Figure 4.29 that hybrid ARIMA-ANN model did not captured the trend of return on exchange rate well as it gives the upward straight line of forecast values. So it can be concluded that hybrid ARIMA-ANN produced the poor forecasts for return on exchange rate series.

4.5.1.3 Estimation of stock return by hybrid ARIMA-ANN

The ANN architecture for residuals of ARIMA from stock return series that have training MSE 0.005, is given below in Figure 4.30:

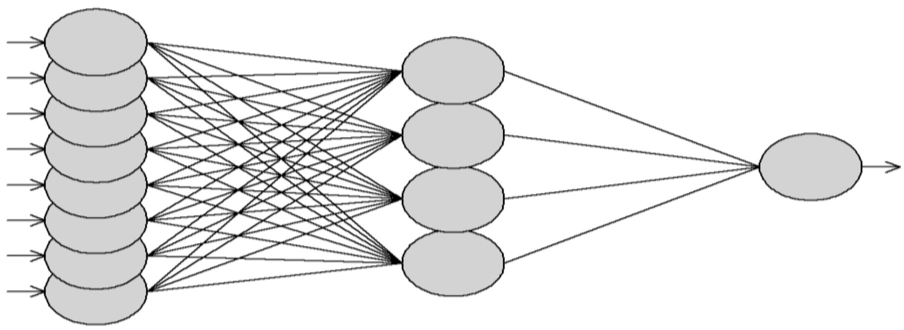


Figure 4.30: ANN Architecture for residuals of stock return

The architecture has 8 inputs with 4 hidden nodes. Mathematical equation for this architecture could be as:

$$e_t = w_0 + \sum_{j=1}^{4} w_j . g\,(w_{0.j} + w_j (\sum_{i=1}^{8} w_i e_{t-i}) + \mathcal{E}_t \qquad (4.12)$$

Where equation (4.12) represents the input and hidden nodes used for the estimation of residuals of ARIMA by ANN model.

The combined forecasts values of stock return from 1 to 24 steps ahead are obtained from hybrid ARIMA-ANN model. Table 4.23 indicates the actual and forecast values on different steps with their respective errors:

Table 4.23: Actual versus forecasted values of stock return by ARIMA-ANN

Months	Actual values	Forecast values	Error	AE
1	0.01969	0.03592	-0.01623	0.01623
4	0.02349	0.03241	-0.00891	0.00891
8	-0.11026	0.03104	-0.1413	0.14129
12	0.01146	0.02569	-0.01423	0.01423
16	-0.00157	0.01566	-0.01723	0.01723
20	-0.02298	0.02642	-0.0494	0.04939
24	-0.08849	0.01746	-0.10595	0.10595

Table 4.23 specifies that overall errors increased as forecasts horizons move from 1 to 24 steps. The graphical representation of actual versus forecasted values are given below in Figure 4.31 as:

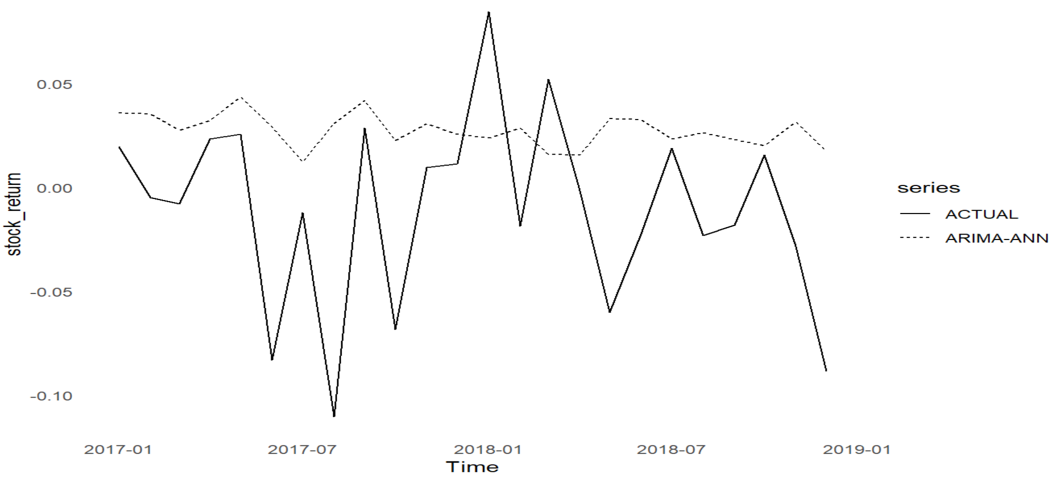


Figure 4.31: Actual versus forecasted values of stock return from hybrid ARIMA-ANN

It can be concluded from Figure 4.31 that forecasts performance of hybrid ARIMA-ANN remained poor as it failed to capture the direction of stock return.

4.5.2 Analysis of all Series through Hybrid ARIMA-SVM

The hybrid model of ARIMA-SVM has similar procedure as discussed for ARIMA-ANN in previous section. The forecasts of ARIMA and its residuals forecasts from

SVM is to be combined to obtain hybrid forecasts. Mathematically, it can be presented as:

$$\hat{y}_t = \hat{L}_{t,ARIMA} + \hat{N}_{t,SVM} \qquad (4.13)$$

Where $\hat{y}_t$ is the forecasted value at time t from hybrid ARIMAPSVM model, $\hat{L}_{t,ARIMA}$ represents the linear forecasts from ARIMA and $\hat{N}_{t,SVM}$ represents the nonlinear forecasts of residuals of ARIMA by SVM model.

4.5.2.1 Estimation of Inflation by Hybrid ARIMA-SVM

As ARIMA forecasts already available in section 4.2, now its residuals forecasts are obtained by following the modeling procedure of SVM thoroughly. Tuning of SVM model is done to obtain the model which has MSE. The best optimal parameters are c=5, γ=5 and ε=0.1 for residuals of inflation from ARIMA, on which SVM model give minimum training error. The plot of the tuning model is given in Figure 4.32 to see the performance of all the models together:

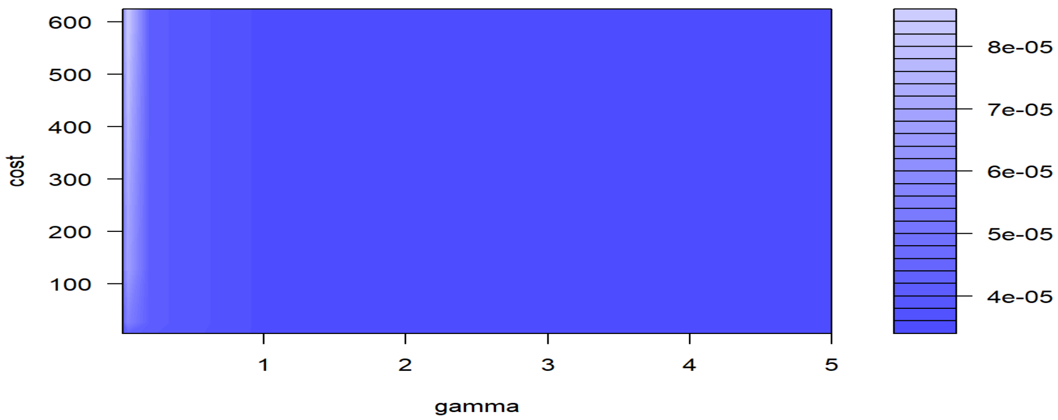


Figure 4.32: SVM performance on residuals of inflation

The 1 to 24 steps ahead hybrid forecasts from ARIMA-SVM are made and actual versus forecasted values on different forecast horizons with their particular errors are given in Table 4.24 as:

Table 4.24: Actual versus forecasted values of inflation by ARIMA-SVM

Months	Actual values	Forecast values	Error	AE
1	-0.00296	-0.00236	-0.00059	0.00059
4	0.00208	0.00168	0.00039	0.00039
8	0.00837	0.00778	0.00059	0.00059
12	0.00345	0.00404	-0.00059	0.00059
16	0.00363	0.00304	0.00059	0.00059
20	0.00301	0.00360	-0.00059	0.00059
24	0.00932	0.00984	-0.00052	0.00051

Table 4.24 indicate that forecasts error between actual and forecasts values did not changed much over the time and short values for errors shows that hybrid ARIMA-SVM performed good in forecasting inflation series. The graphical representation of forecasted versus actual values is given below in Figure 4.33:

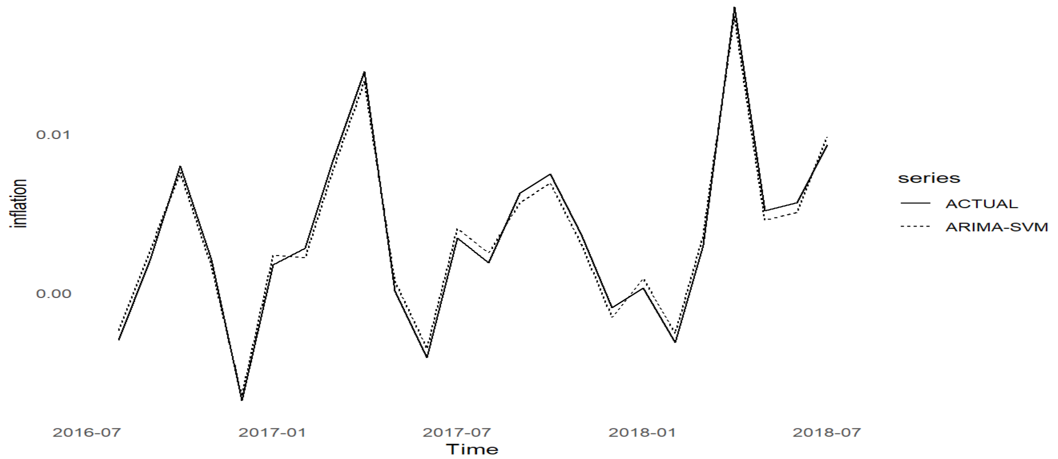


Figure 4.33: Actual versus forecasted values of inflation from hybrid ARIMA-SVM

It can be noted from Figure 4.33 that hybrid ARIMA-SVM captured the direction of inflation very well which shows ARIMA-SVM model could perform well in case of forecasting time series data up to maximum level.

4.5.2.2 Estimation of Return on Exchange Rate by Hybrid ARIMA-SVM

The best optimal parameters are c=10, γ=10 and ε=0.1 for residuals of exchange rate from ARIMA, on which SVM model give minimum training error after tuning of the model. The graph for SVM parameters performance on different range of values while tuning the model are given in Figure 4.34 as:

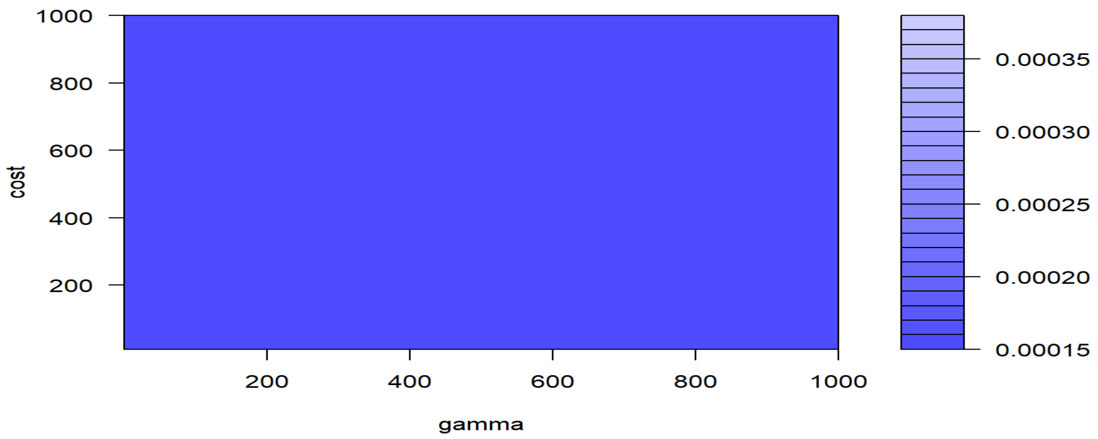


Figure 4.34: SVM performance on residuals of return on exchange rate

The interpretation of the above plot is same as other plots of SVM performance explained earlier. Darkest the blue region in plot means there are more possibility of obtaining the better model by narrowing the range gamma and cost values around the most darkest region.

The hybrid forecasts on different horizons with their particular errors are given in Table 4.25 as:

Table 4.25: Actual versus forecasted values of return on exchange rate by ARIMA-SVM

Months	Actual values	Forecast values	Error	AE
1	0.00005	0.00127	-0.00122	0.00122
4	0.00005	0.00127	-0.00122	0.00122
8	-0.00111	0.00011	-0.00122	0.00122
12	0.03116	0.02993	0.00123	0.00123
16	0.02948	0.02826	0.00122	0.00122
20	-0.0045	-0.0033	-0.00122	0.00122
24	0.03596	0.03474	0.00122	0.00122

Table 4.25 shows ARIMA-SVM hybrid do good in forecasting as it has little error throughout the forecasted step horizons.

The graphical portray of forecasted versus actual values is given as in Figure 4.35:

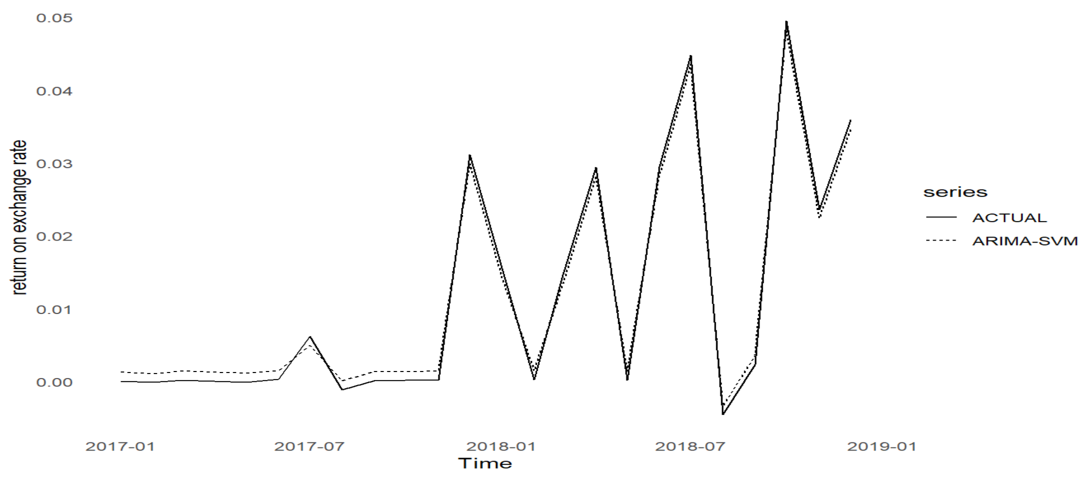


Figure 4.35: Actual versus forecasted values of return on exchange rate from hybrid ARIMA-SVM

Figure 4.35 indicates that hybrid ARIMA-SVM model performed good in case of return on exchange rate as it has little variance between the forecasted and actual values. It can be also seen from the Figure that it captured the right trend of return on exchange rate.

4.5.2.2 Estimation of Stock Return by hybrid ARIMA-SVM

The best optimal parameters are c=10, γ=1 and ε=0.1 for residuals of exchange rate from ARIMA, on which SVM model give minimum training error after tuning of SVM parameters. The graph for SVM parameters performance on different range of values while tuning the model is given below in Figure 4.36 as:

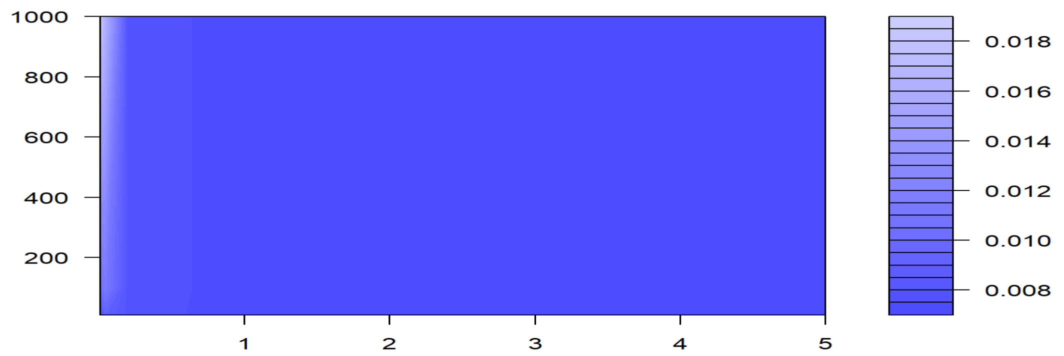


Figure 4.36: **SVM performance on residuals of stock return**

Forecasts from 1 to 24 step ahead are made by selected best SVM model and then combining these forecasts to earlier obtained forecasts of ARIMA, hybrid forecasts are obtained for stock return. The hybrid forecast on different horizons with their particular errors are given below in Table 4.26 for stock return as:

Table 4.26: Actual versus forecasted values of stock return by ARIM-SVM

Months	Actual values	Forecast values	Error	AE
1	0.01969	0.01247	0.00722	0.00722
4	0.02349	0.01953	0.00397	0.00397
8	-0.11026	-0.10184	-0.00842	0.00842
12	0.01146	0.00819	0.00327	0.00326
16	-0.00157	0.00683	-0.0084	0.00840
20	-0.02298	-0.01455	-0.00843	0.00843
24	-0.08849	-0.08002	-0.00847	0.00847

Likewise the other series ARIMA-SVM hybrid has also good forecast performance in stock return as the Table 4.26 indicate, it has relatively low errors throughout the whole time horizon. The graphical representation of forecasted versus actual values is given in Figure 4.37

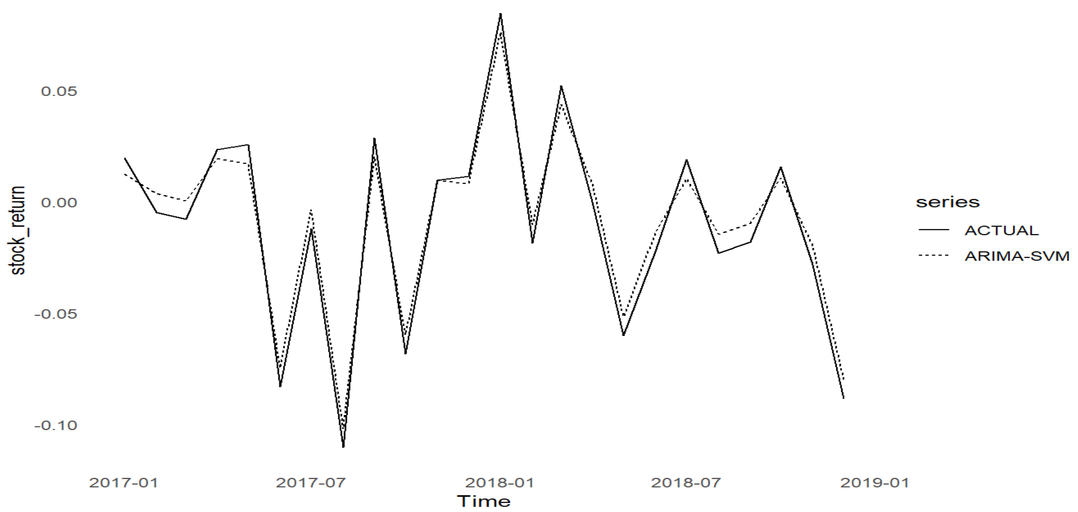


Figure 4.37: Actual versus forecasted values of stock return from hybrid ARIMA-SVM

The above Figure 4.37 shows hybrid model of ARIMA-SVM perform well in capturing the right direction of stock return. Figure indicates that there is very little difference among the forecasted and actual values which shows the preciseness of ARIMA-SVM hybrid model in capturing the right direction for stock return.

The estimated results from all proposed models are discussed one by one on all data sets such as inflation, return on exchange rate and stock return so for. Furthermore to analyze the preciseness of forecasted values against actual values graphical comparison is made as well as loss functions such as Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) are used. The statistical equations for used loss functions are given as:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(A_t - F_t)^2}{n}} \qquad (4.13)$$

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|A_t - F_t| \qquad (4.14)$$

4.6 Forecasts performance comparison of all used models

In this section performance of all models is discussed systematically for every data set used in analysis, in which RMSE and MAE are used to check the overall accuracy of a model forecasting ability. While graphs signifies that how much a model succeeded in capturing the right trend or direction of a data set. The Table 4.27 contains the used loss functions values for inflation series based on test data from all models is given below.

Table 4.27:Loss errors of all models for inflation series

Models	ARIMA	ANN	SVM	ARIMA-ANN	ARIMA-SVM
RMSE	0.004367	0.003654	0.004495	0.003977	0.000567
MAE	0.003363	0.003038	0.003603	0.003143	0.000562

Table 4.27 shows ARIMA-SVM hybrid model perform excellent as compared to other models based on RMSE and MAE. ANN has minimum RMSE and MAE after ARIMA-SVM model while the hybrid model of ARIMA-ANN remain at 3rd position. Then ARIMA shows it has better performance then SVM based on used loss functions. The graphical representation of inflation series is given in Figure 4.38.

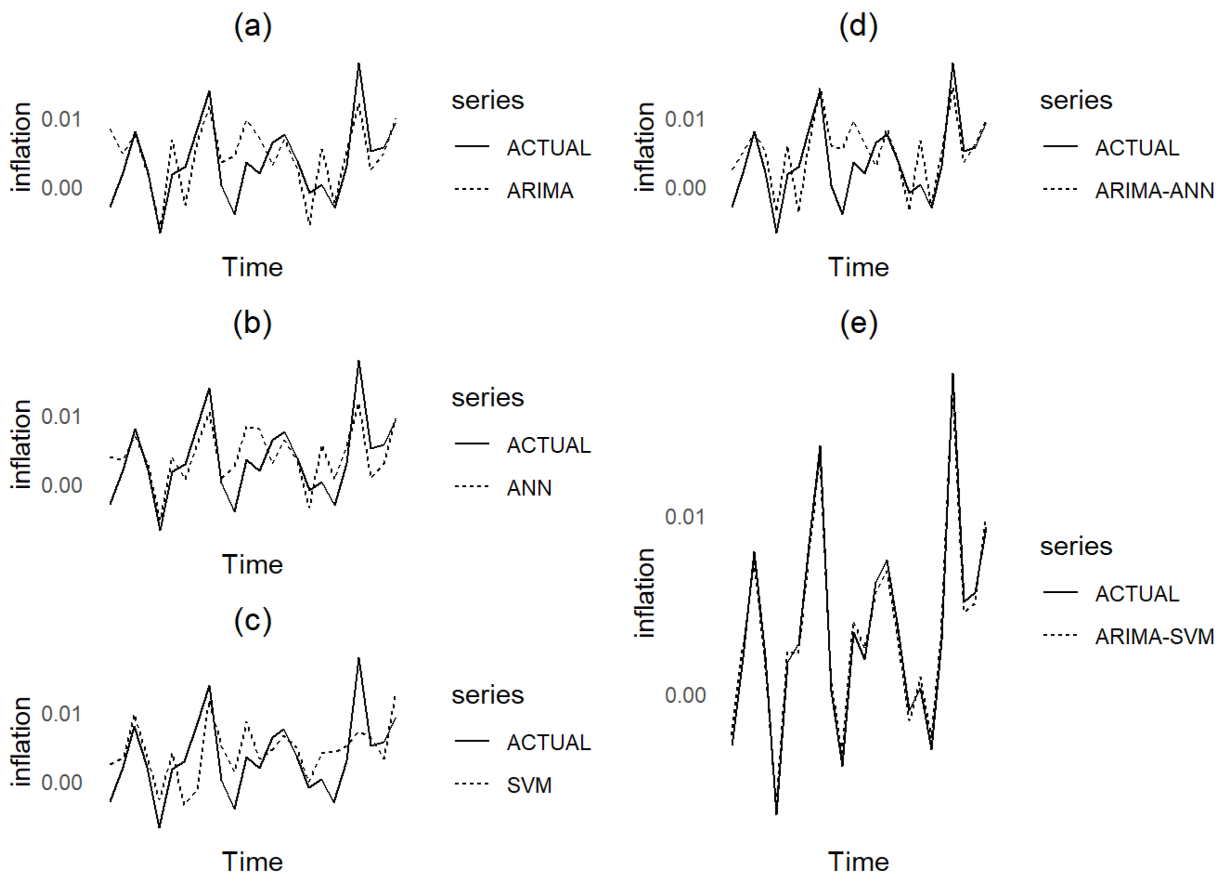


Figure 4.38: Actual and forecast values of inflation for all models

In the Figure 4.38 solid lines indicate the actual values and dashed lines represents the forecast values for respective models. It can be seen from the Figure that all model seems to be good in capturing the direction of the actual values of solid line. But the ARIMA-SVM Hybrid model outperform all other models because it almost overlapped the actual values. Also hybrid model for ARIMA-ANN doing good in catching the trend. After these hybrid models, SVM do good then ANN and ARIMA remain at last position. One important thing can be noted that ANN and ARIMA both perform poor in catching the real direction of actual values in start than SVM model, but has lower loss errors than SVM.

Next series for forecast performance evaluation of various models used is return on exchange rate. The Table 4.28 for its loss functions are given below as:

Table 4.28: Loss errors of all models for exchange rate series

Models	ARIMA	ANN	SVM	ARIMA-ANN	ARIMA-SVM
RMSE	0.017707	0.013309	0.015868	0.015421	0.001221
MAE	0.012744	0.010720	0.010139	0.011975	0.001221

Above Table shows that ARIMA-SVM hybrid model has lowest loss errors as compared to other models and ANN stands at second place because it has comparatively least RMSE and MAE compared to other models. The hybrid ARIMA-ANN and SVM relatively have mixed results as one performed better in RMSE while other has good performance in case of MAE. However, ARIMA has the highest loss errors as comparison with all other models. It can be concluded from the Table 4.28 that a suitable hybrid model can provide the better forecasts based on RMSE and MAE for return on exchange rate series than most of linear and nonlinear models. The next graphical picture of return on exchange rate by all used models is given in Figure 4.39 a:

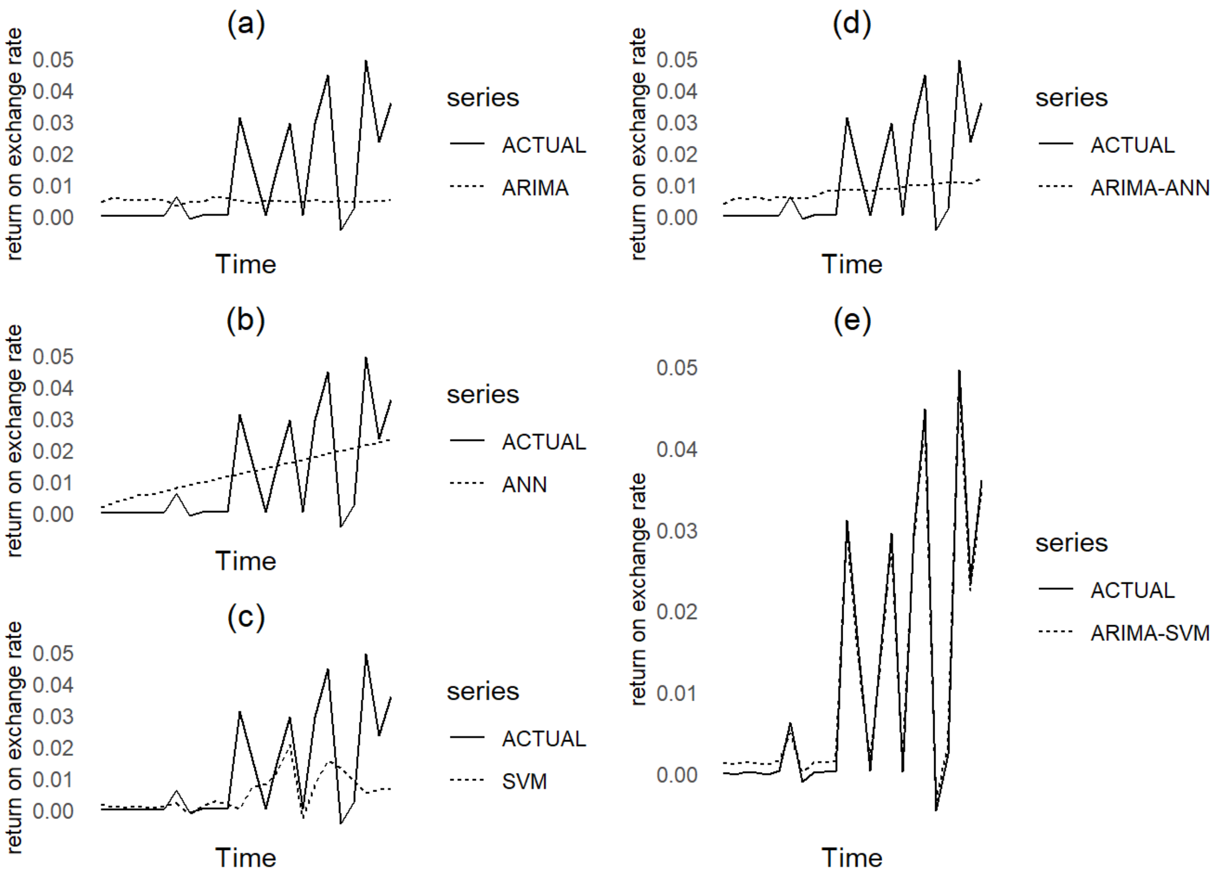


Figure 4.39: Actual versus forecast values of return on exchange rate for all models

The hybrid model of ARIMA-SVM perform very well in capturing the real direction of return on exchange rate in which the dashed line that remain throughout close to actual values of solid line can be seen in part (e) of the Figure 4.39. After this single model of SVM can be seen in part (c) of the Figure which captured the actual direction of real values to some extent. However all other models perform very poor in catching the real trend of return on exchange rate, part (a) represents straight dashed line of ARIMA, while the other models like ANN and hybrid ARIMA-ANN in part (b) and (d) have the upward straight line. One more thing can be noted that SVM model comparatively remain good in capturing the real trend of return on exchange rate than other models that have lower loss errors as compared to SVM except hybrid ARIMA-SVM.

The different models forecasts performance for stock return series is discussed as following. The Table 4.29 has the loss functions of all models for stock return are as:

Table 4.29: Loss errors of all models for stock return series

Models	ARIMA	ANN	SVM	ARIMA-ANN	ARIMA-SVM
RMSE	0.049153	0.051883	0.069622	0.059136	0.007787
MAE	0.036776	0.374887	0.058995	0.046278	0.007502

It can be seen from above Table 4.29 that ARIMA-SVM hybrid model has minimum loss errors which shows it performed better than others. The RMSE and MAE of ARIMA is lowest after hybrid ARIMA-SVM but from Figure 4.39, it can be seen that ARIMA performed worse in capturing the direction of return than any other models as in part (a) of the Figure 4.39 straight dashed line represent ARIMA model. While in case of SVM model, it has higher loss errors when compared to other models except it superseded ANN in case of MAE but perform better after ARIMA-SVM hybrid in capturing the direction of stock return. Graphical representation of actual versus forecast values of stock return for all models is given in Figure 4.40

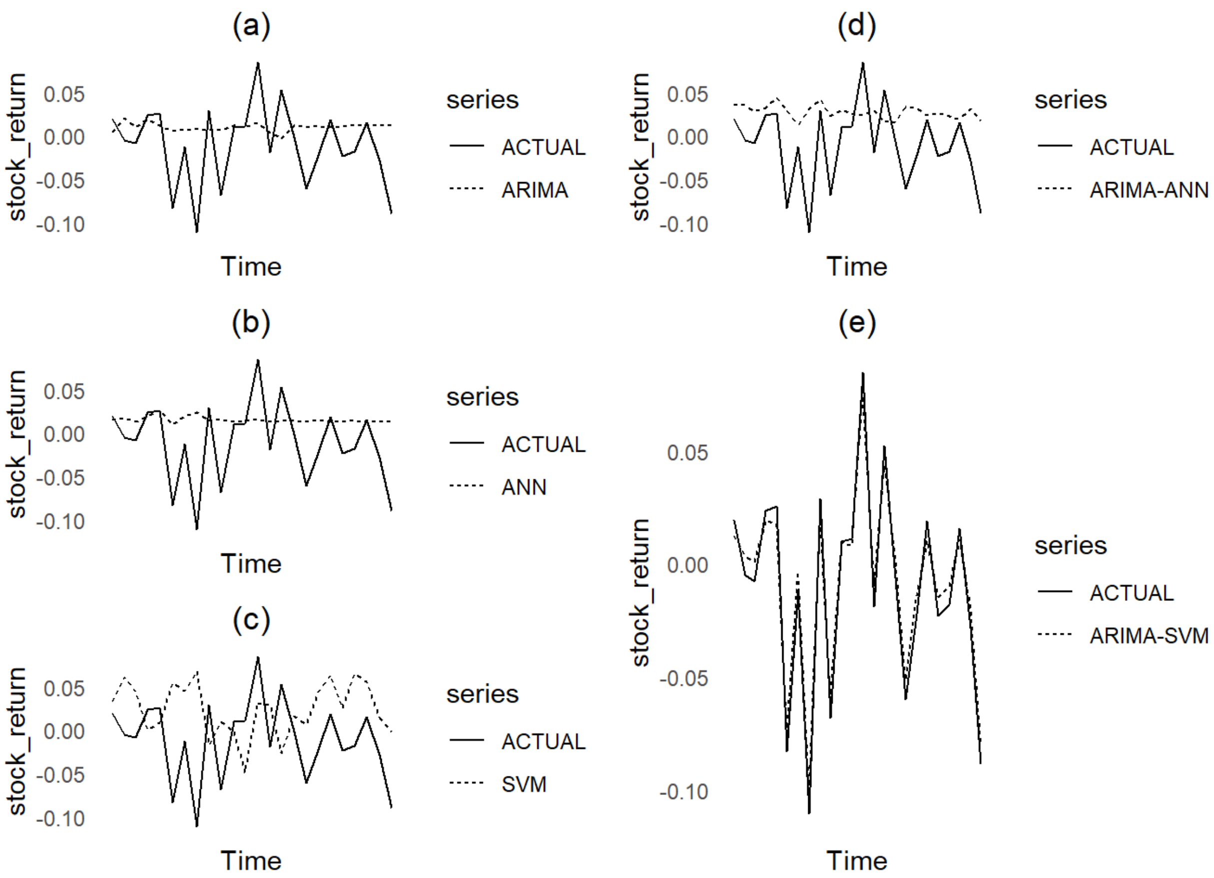


Figure 4.40: Actual and forecast values of stock return for all models

Figure 4.40 indicate that like other two series hybrid model of ARIMA-SVM also outperformed other models in this data series as the dashed line of ARIMA-SVM in part (e) overwhelming the actual values of the stock return. All other model has poor performance while capturing the direction of return except SVM model which performed better as compared to others but also has weak performance. After detailed forecast performance evaluation based on two loss errors and graphical representation, it was noted that hybrid ARIMA-SVM performed well and surpassed all other models either ARIMA has heteroskedasticity problem or not. However, ARIMA performed relatively better and improved its position if there is no problem of heteroskedasticity but did not made a significant impact over the hybrid of ARIMA-SVM.

CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

Time series forecasting is an active area of interest in many applications for research and policy process. The nature of data from different applications is unknown and complex. The ARIMA model has been dominated in many time series applications with the efforts of Box-Jenkins (1976). Latterly, ANN and SVM shows their well applicability in forecasting the time series with their nonlinear modeling applications. All of three model has an effective flexibility in forecasting time series data but none of these individually give always satisfactory performance. A lot of existing theoretical and empirical literature suggests that integration of two dissimilar models would be an effective way of achieving high degree of accuracy. Therefore in this study an attempt is made to explore the joint models with their individuals, in which linear ARIMA is combined with the nonlinear models like ANN and SVM. The fundamental objective of hybrid model is to take the advantage of both linear and nonlinear strength and to capture the diverse forms of relationship in time series data. The forecast performance of two hybrid models with their three independent models is evaluated through RMSE, MAE and graphical representation. Also forecast accuracy measured through absolute error on different time horizon. Three types of real data sets, namely Inflation, Exchange rate and Stock return are used for empirical analysis. Empirical results clearly indicate that ARIMA-SVM hybrid model outperformed all other models used in isolation and the other hybrid model of ARIMA-ANN built in this study. This comparison is made in terms of RMSE, MAE and graphical representation which is used to validate the directional prediction. It can be noted from results that all other models has the mixed results in terms of said accuracy measures including the ARIMA-ANN hybrid model. So our study contributes in some way as in one hand it validates the existing literature that combination of different models having unique and dissimilar qualities perform better than model used in isolation. And reduces the model uncertainty by unstable and changing pattern in the data which normally occur in statistical interpretation and time series forecasting. On the other hand, it contradicts on existing findings that all hybrid models can perform better in all cases because performance of hybrid ARIMA-ANN model cannot make satisfactory impact than individual models. One should make careful consideration while making the hybrid model. In our findings one thing also can be noted that this may be due to high volatility in our data set. So removing the volatility in data set an appropriate filter could be applied in order to achieve required outcomes. The credibility of ARIMA-SVM hybrid model also validated on different data sets in many existing literature like (Kumar 2014, Chen an Wang 2007 and Pai and Lin 2007). The outclass achievement of ARIMA-SVM above other models is due to SVM applies structural risk minimization principle which minimizes an upper bound of the generalization error rather than minimizing the training error. Which ultimately leads to superior generalization performance as compared to other nonlinear models. The hybrid ARIMA-SVM model superiority shows that it could may be utilized by policy makers and investors in forecasting economic and financial data.

It is recommended that the performance of ANN may be improved by applying some different methods of preprocessing and input selection and also using the alternative available training algorithms instead of RPROP. By experimenting different algorithms and input selection methods it would be interesting way to remove the volatility in data due to our current model fail to capture real trend of data. On other hand GARCH modeling could be used to solve the volatility problem in hybrid modeling instead of ARIMA, because it has limitations in resolving the volatility and it would be our next research purpose to explore the hybrid modeling with GARCH. The forecast performance of used models also be improved by adding the relevant significant regressor terms with the lagged variables. Autoregressive Moving Average with Exogenous Variables (ARIMAX) model could be used for this purpose. ARIMAX model is an expansion of ARIMA model which works like a multivariate regression that take the advantage of ARIMA terms with additional independent variables to improve the forecast accuracy.

REFERENCES

Al-hnaity, B., & Abbod, M. (2016). Predicting financial time series data using hybrid model. In Intelligent Systems and Applications (pp. 19-41). Springer, Cham.

Awan, S. M., Khan, Z. A., Aslam, M., Mahmood, W., & Ahsan, A. (2012). Application of NARX based FFNN, SVR and ANN Fitting models for long term industrial load forecasting and their comparison. In Industrial Electronics (ISIE), IEEE International Symposium on (pp. 803-807).

Babu, C. N., & Reddy, B. E. (2014). A moving-average filter based hybrid ARIMA–ANN model for forecasting time series data. Applied Soft Computing, 23, 27-38.

Barrow, D., & Kourentzes, N. (2018). The impact of special days in call arrivals forecasting: A neural network approach to modelling special days. European Journal of Operational Research, 264(3), 967-977.

Breiman, L. (1996). Stacked regressions. Machine learning, 24(1), 49-64.

Brock, W. A., Dechert, W. D., Scheinkman, J. A., & LeBaron, B. (1986). A test for independence based on the correlation integral. Technical report. University of Wisconsin-Madison.

Broock, W. A., Scheinkman, J. A., Dechert, W. D., & LeBaron, B. (1996). A test for independence based on the correlation dimension. Econometric reviews, 15(3), 197-235.

Burney, S. M. A., Jilani, T. A., & Ardil, C. (2005). Levenberg-Marquardt algorithm for Karachi Stock Exchange share rates forecasting. International Journal of Computational Intelligence, 1(3), 144-149.

Cao, L. J., & Tay, F. E. H. (2003). Support vector machine with adaptive parameters in financial time series forecasting. IEEE Transactions on neural networks, 14(6), 1506-1518.

Cao, L., & Tay, F. E. (2001). Financial forecasting using support vector machines. Neural Computing & Applications, 10(2), 184-192.

Chen, K. Y., & Wang, C. H. (2007). A hybrid SARIMA and support vector machines in forecasting the production values of the machinery industry in Taiwan. Expert Systems with Applications, 32(1), 254-264.

Claveria, O., Monte, E., & Torra, S. (2015). Regional Forecasting with Support Vector Regressions: The Case of Spain.

Cocianu, C. L., & Grigoryan, H. (2015). An artificial neural network for data forecasting purposes. Informatica Economica, 19(2), 34.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine learning, 20(3), 273-297.

Farooqi, A. (2014). ARIMA model building and forecasting on Imports and Exports of Pakistan. Pakistan Journal of Statistics and Operation Research, 10(2), 157-168.

Fletcher, R. (1989). Practical Methods of Optimization.

Frain, J. (1992). Lecture Notes on Univariate Time Series Analysis and Box Jenkins Forecasting. Economic Analysis, Research and Publications.

Guajardo, J., Weber, R., & Miranda, J. (2006). A forecasting methodology using support vector regression and dynamic feature selection. Journal of Information & Knowledge Management, 5(04), 329-335.

Gunn, S. R. (1998). Support vector machines for classification and regression. ISIS technical report, 14(1), 5-16.

Haider, A., & Hanif, M. N. (2009). Inflation forecasting in Pakistan using artificial neural networks. Pakistan economic and social review, 123-138.

Jafri, Y. Z., Sami, M., Waseem, A., Murtaza, G., & Akbar, S. (2012). Stochastic approaches for time series forecasting of rate of dust fall: A case study of northwest of Balochistan, Pakistan. International Journal of Physical Sciences, 7(4), 676-686.

Karathanasopoulos, A., & Osman, M. (2019). Forecasting the Dubai financial market with a combination of momentum effect with a deep belief network. Journal of Forecasting.

Khandelwal, I., Adhikari, R., & Verma, G. (2015). Time series forecasting using hybrid ARIMA and ANN models based on DWT decomposition. Procedia Computer Science, 48, 173-179.

Khashei, M., & Bijari, M. (2010). An artificial neural network (p, d, q) model for timeseries forecasting. Expert Systems with applications, 37(1), 479-489.

Khashei, M., & Bijari, M. (2011). Which methodology is better for combining linear and nonlinear models for time series forecasting?. Journal of Industrial and Systems Engineering, 4(4), 265-285.

Kim, H. Y., & Won, C. H. (2018). Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. Expert Systems with Applications, 103, 25-37.

Kim, K. J. (2003). Financial time series forecasting using support vector machines. Neurocomputing, 55(1-2), 307-319.

Kumar, M., & Thenmozhi, M. (2014). Forecasting stock index returns using ARIMA-SVM, ARIMA-ANN, and ARIMA-random forest hybrid models. International Journal of Banking, Accounting and Finance, 5(3), 284-308.

Kuntoji, Geetha & Rao, Subba & Mandal, Sukomal. (2017). Application of Support Vector Machine Technique for Damage Level Prediction of Tandem Breakwater. International Journal of Earth Sciences and Engineering.

LeBlanc, M., & Tibshirani, R. (1996). Combining estimates in regression and classification. Journal of the American Statistical Association, 91(436), 1641-1650.

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. The bulletin of mathematical biophysics, 5(4), 115-133.

Meyer, D., & Wien, F. T. (2001). Support vector machines. R News, 1(3), 23-26.

Meyer, D., & Wien, F. T. (2015). Support vector machines. The Interface to libsvm in package e1071, 28.

Meyler, A., Kenny, G., & Quinn, T. (1998). Forecasting Irish inflation using ARIMA models.

Mojirsheibani, M. (1999). Combining classifiers via discretization. Journal of the American Statistical Association, 94(446), 600-609.

Moshiri, S., & Cameron, N. (2000). Neural network versus econometric models in forecasting inflation. Journal of forecasting, 19(3), 201-217.

Mukherjee, S., Osuna, E., & Girosi, F. (1997). Nonlinear prediction of chaotic time series using support vector machines. In Neural Networks for Signal Processing. Proceedings of the IEEE Workshop (pp. 511-520).

Pai, P. F., & Lin, C. S. (2005). A hybrid ARIMA and support vector machines model in stock price forecasting. Omega, 33(6), 497-505.

Papatla, P., Zahedi, M., & Zekic-Susac, M. (2002). Leveraging the strengths of choice models and neural networks: A multiproduct comparative analysis. Decision Sciences, 33(3), 433-461.

Patterson, J. C. (2012). Managing a real-time massively-parallel neural architecture (Doctoral dissertation, The University of Manchester (United Kingdom)).

Petruseva, S., Zileska-Pancovska, V., Žujo, V., & Brkan-Vejzović, A. (2017). Construction costs forecasting: comparison of the accuracy of linear regression and support vector machine models. Tehnicki Vjesnik-Technical Gazette, 24(5), 1431-1438.

Saeed, N., Saeed, A., Zakria, M., & Bajwa, T. M. (2000). Forecasting of wheat production in Pakistan using ARIMA models. International Journal of Agricultural Biology, 2(4), 352-353.

Shamim, M. A., Ghumman, A. R., & Ghani, U. (2004, December). Forecasting groundwater contamination using artificial neural networks. In 1st International Conference on Water Resources and Arid Environments (pp. 5-8).

Sharma, A., & Bhardwaj, K. (2015). Identification of normal and abnormal ECG using neural network. Int. J. Inf. Res. Rev, 2, 695-700.

Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. Statistics and computing, 14(3), 199-222.

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. Journal of the royal statistical society. Series B (Methodological), 111-147.

Sultana, K., Rahim, A., Moin, N., Aman, S., & Ghauri, S. P. (2013). Forecasting inflation and economic growth of pakistan by using two time series metho ds. International Journal of Business and Economics Research, 2(6), 174-178.

Tkacz, G., & Hu, S. (1999). Forecasting GDP growth using artificial neural networks (pp. 99-3). Ottawa: Bank of Canada.

Trafalis, T. B., & Ince, H. (2000). Support vector machine for regression and applications to financial forecasting. In Neural Networks, 2000. IJCNN

2000, Proceedings of the IEEE-INNS-ENNS International Joint Conference on (Vol. 6, pp. 348-353).

Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing, 50, 159-175.

Zhang, G., Hu, M. Y., Patuwo, B. E., & Indro, D. C. (1999). Artificial neural networks in bankruptcy prediction: General framework and cross-validation analysis. European journal of operational research, 116(1), 16-32.

Zhu, B., & Wei, Y. (2013). Carbon price forecasting with a novel hybrid ARIMA and least squares support vector machines methodology. Omega, 41(3), 517-524.

Zhu, S., Lian, X., Liu, H., Hu, J., Wang, Y., & Che, J. (2017). Daily air quality index forecasting with hybrid models: A case in China. Environmental pollution, 231, 1232-1244