Indian Stock Market Predictive Efficiency using the ARIMA Model

Manish R Pathak

Asst.Professor
Head Vidyabharti College of MCA (MBA),Umrakh

Jimmy. M. Kapadia

Professor and Director
S.R.Luthra Institute of Management, Surat

Abstract: Stock Market prediction is an important topic in finance and economics which has encouraged the interest of researchers over the years to develop better predictive models. The autoregressive integrated moving average (ARIMA) models have been explored in literature for time series prediction. This paper presents extensive process of building stock indices predictive model using the ARIMA model. Published stock indices data obtained from National Stock Exchange (NSE) are used with stock indices price predictive model developed. At first the stationarity condition of the data series are observed by ACF and PACF plots, then checked using the statistics such as Ljung-Box-Pierce Q-statistic and Dickey-Fuller test statistic. Results obtained revealed that the ARIMA (1,1,2) model has a strong potential for short-term prediction and Adjusted ARIMA gives more accurate forecasting compare to ARIMA.

Keywords: ARIMA model, Stock Price prediction, Adjusted ARIMA model, Short-term prediction

Introduction

Predicting index price in stock market is quite a challenging task due to the uncertain factors involved that influence the market price on each day like economic conditions, investor's sentiment, political events, pandemic, etc. The National Stock Exchange (NSE) was set up in November 1992. NSE is the first exchange in India to provide an automated screen-based electronic trading structure. NSE was set up by a group of leading Indian financial institutions at the request of the Government of India to bring precision to the national capital market. The Nifty_50 (NSE) is an indicator of several sectors like Automobiles, Bank, Realty, Media, Pharmaceuticals, Information Technology, Energy, etc.

Literature Review

Uma Devi, Sundar and Alli (2013) did the research on An Effective Time Series Analysis for Stock Trend Prediction Using ARIMA Model for Nifty Midcap-50. Top four companies having max Midcap value have been chosen for analysis. The historical stock data for the past five years has been collected and trained using ARIMA model with different parameters. The test criterions like Akaike Information Criterion & Bayesian Information Criterion (AIC/BIC) were applied to predict the accuracy of the model. The performance of the trained model is analyzed and tested to find the trend and the market behavior for future forecast.

Jiban Chandra Paul. S Hoque and Mohammad Rahma (2013) did the research on selection of Best ARIMA Model for Forecasting Average Daily Share Price Index of Pharmaceutical Companies in Bangladesh. They Found that ARIMA (2, 1, 2) is found as the best model for forecasting the SPL data series. PREETHI and

SANTHI (2014) studied Neural Networks and Neuro-Fuzzy systems that are identified to be the leading machine learning techniques in stock market index prediction area. They set algorithm and find out that ARIMA forecasting method provides a better accuracy than traditional methods.

Ayodele, Aderemi and Charles (2014) present extensive process of building stock price predictive model using the ARIMA model. Stock data obtained from publication of New York Stock Exchange (NYSE) and Nigeria Stock Exchange (NSE) are used with stock price predictive model developed. Results obtained revealed that the ARIMA model (1, 0, 1) has a strong potential for short-term prediction and can compete favourably with existing techniques for stock price prediction.

(Madhu, 2014) studied A time series modeling approach (Box-Jenkins' ARIMA model) has been used in this study to forecast sugarcane production in India. The order of the best ARIMA model is (2,1,0). Further, efforts were made to forecast, as accurately as possible, the future sugarcane production for a period up to five years by fitting ARIMA (2,1,0) model to our time series data. The results showed that the annual sugarcane production will grow in 2013, then will take a sharp dip in 2014 and in subsequent years 2015 through 2017, it will continuously grow with an average growth rate of approximately 3% year-on-year.

(K, 2017) did forecasting of National Stock Price using ARIMA model. The Nifty_50 stock market prices were analysed, and predicted the trend of upcoming trading days stock market fluctuations using Box-Jenkins methodology. They observed that influence of R-Square value is (94%) high and Mean Absolute Percentage Error is very small for the Fitted model [ARIMA (0, 1, 1)]. Therefore, the forcasting accuracy is more suitable of Nifty 50 closing stock price in ARIMA (0, 1, 1). They concluded that decreasing fluctuations trend for upcoming trading days. (Fattah, 2018) and Bijesh Dhyani et al (2020) have done the research on Stock Market Forecasting Technique using Arima Model. They found ARIMA (1,0,1) model proves that the model could be utilized to model and forecast the future demand in food manufacturing.

Swapnil Jadhav et al (2015), Vijay Pandey and Abhishek Bajpai (2019) attempt to identify superior combinations in ARIMA model as well as ANN model for predicting Indian Stock Market namely NSE Nifty 50 daily data of 10 years period. Through the simultaneous use of AAE, RMSE, MAPE and MSPE statistical tools, the predictive accuracy of ARIMA (p d q) and ANN model have been compared. The results indicate ARIMA (2,1,2) and ANN (4-10-1) with both train functions GDX and BFG are best predictors with ANN dominating over ARIMA model.

Research Gap

The literature review shows a gap that ARIMA model used for forecasting of stock price and Market Index. It has been observed that when time frame and market had changed, findings are not consisteit with past studies. Therefore prediction ability of ARIMA model also changes. Therefore, there is a scope that Adjusted ARIMA model can be used for increased forecasting ability.

Research Methodology

An important objective of the time series analysis is to study the past behaviour of the available data and then forecast with fitting a suitable model with the help of econometric or statistical techniques.

The objective of study is to forecast the National Stock Exchange closing stock price of Nifty 50 using ARIMA model in Time Series Analysis.

To achieve this descriptive research, secondary financial data of Nifty 50 for the past five years from 1st Sept-2015 to 1st Sept-2020 has been used. The tool used for implementation is Eviews software version 9.

To determine the best ARIMA model the criteria used in this study include Relatively small of BIC (Bayesian or Schwarz Information Criterion), Relatively small standard error of regression (S.E. of regression), Relatively high of adjusted R2 and

Q-statistics and correlogram show that there is no significant pattern left in the autocorrelation functions (ACFs) and partial autocorrelation functions (PACFs) of the residuals, it means the residual of the selected model is white noise.

ARIMA models

ARIMA models provide approach to time series forecasting. Its aim is to describe the autocorrelations in the data. It is a "stochastic" modeling approach that can be used to calculate the probability of a future value lying between two specified limits.

Overall Time series Analysis and Forecasting Process

To build a time series model issuing ARIMA, we need to study the time series and identify p,d,q

- Ensuring Stationarity:Determine the appropriate values of d
- Identification: Determine the appropriate values of p & q using the ACF, PACF, and unit root tests and p is the AR order, d is the integration order, q is the MA order
- Estimation: Estimate an ARIMA model using values of p, d, & q you think are appropriate.
- Diagnostic checking: Check residuals of estimated ARIMA model(s) to see if they are white noise; pick best model with well behaved residuals.
- Forecasting: Produce out of sample forecasts or set aside last few data points for in-sample forecasting.

ARIMA is also known as Box-Jenkins approach. It is popular because of its generality; It can handle any series, with or without seasonal elements, and it has well-documented computer programs

Result and Discussion

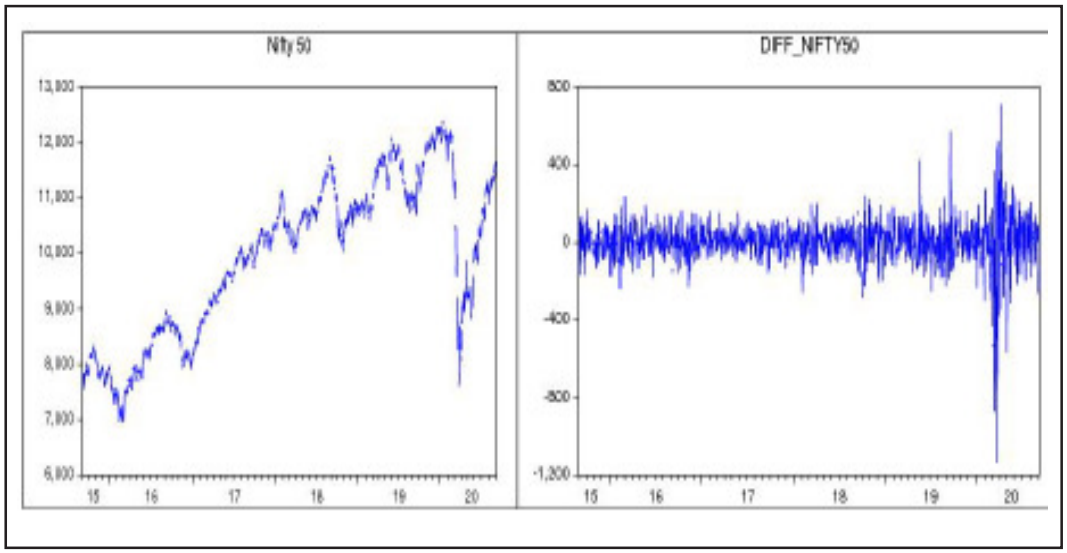

Figure 1: Source : Eviews Ouput

Figure 2 : Source: Eviews Output

Figure 1 shows original pattern of Nifty_50 series from 2015 to 2020, whether the time series is stationary or not. From the graph we can interpret that time series have random walk pattern. From the graph, the ACF dies down extremely slowly which simply means that the time series is non-stationary. If the series is not-stationary, it is converted to a stationary series by differencing. Figure 2 shows 1st difference of Nifty_50 series. After the first difference, the series "DNifty_50" of Nifty_50 becomes stationary.

Date: 10/28/20 Time: 09:54
Sample: 9/01/2015 9/01/2020
Included observations: 1224

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.996	0.996	1216.4	0.000
		2	0.992	0.026	2423.7	0.000
		3	0.987	-0.056	3620.9	0.000
		4	0.982	-0.041	4807.3	0.000
		5	0.977	0.000	5982.9	0.000
		6	0.972	-0.065	7146.4	0.000
		7	0.967	0.103	8300.1	0.000
		8	0.962	-0.067	9442.4	0.000
		9	0.957	-0.003	10574.	0.000
		10	0.952	0.008	11694.	0.000
		11	0.946	-0.089	12802.	0.000
		12	0.941	0.057	13898.	0.000
		13	0.935	-0.024	14983.	0.000
		14	0.930	-0.017	16055.	0.000
		15	0.924	0.000	17116.	0.000
		16	0.919	-0.022	18164.	0.000
		17	0.913	0.019	19200.	0.000
		18	0.908	0.015	20225.	0.000
		19	0.902	0.034	21239.	0.000
		20	0.897	-0.002	22243.	0.000
		21	0.892	-0.022	23235.	0.000
		22	0.887	0.025	24218.	0.000
		23	0.882	-0.004	25190.	0.000
		24	0.877	-0.007	26152.	0.000
		25	0.872	0.034	27104.	0.000
		26	0.867	-0.030	28046.	0.000
		27	0.863	0.017	28979.	0.000
		28	0.858	-0.014	29902.	0.000
		29	0.853	-0.025	30816.	0.000
		30	0.848	0.019	31719.	0.000
		31	0.843	-0.013	32613.	0.000
		32	0.838	0.008	33498.	0.000
		33	0.834	0.039	34373.	0.000

Source : Eviews Output

Date: 10/28/20 Time: 09:58
Sample: 9/01/2015 9/01/2020
Included observations: 1223

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.044	-0.044	2.3991	0.121
		2	0.067	0.065	7.8944	0.019
		3	0.037	0.043	9.5400	0.023
		4	-0.017	-0.018	9.8877	0.042
		5	0.125	0.119	29.028	0.000
		6	-0.140	-0.132	53.068	0.000
		7	0.117	0.097	68.947	0.000
		8	-0.007	0.006	70.007	0.000
		9	-0.015	-0.015	70.292	0.000
		10	0.133	0.113	92.046	0.000
		11	-0.087	-0.051	101.31	0.000
		12	0.092	0.039	111.84	0.000
		13	-0.019	0.011	112.30	0.000
		14	-0.023	-0.037	112.98	0.000
		15	0.049	0.019	115.94	0.000
		16	-0.061	-0.011	120.59	0.000
		17	0.018	-0.042	120.88	0.000
		18	-0.072	-0.040	127.41	0.000
		19	0.027	0.024	128.33	0.000
		20	0.037	0.017	130.01	0.000
		21	-0.028	0.011	131.02	0.000
		22	0.018	-0.018	131.44	0.000
		23	-0.015	0.006	131.74	0.000
		24	-0.059	-0.076	136.04	0.000
		25	0.039	0.041	137.91	0.000
		26	-0.046	-0.022	140.56	0.000
		27	0.034	0.023	142.04	0.000
		28	-0.003	0.019	142.05	0.000
		29	-0.003	-0.003	142.06	0.000
		30	0.036	0.014	143.70	0.000
		31	-0.030	0.001	144.78	0.000
		32	-0.016	-0.050	145.13	0.000
		33	-0.022	-0.007	145.72	0.000

The correlogram view of the residuals (forecast errors) shows the autocorrelations of the residuals. These are the correlation coefficients of values of the residuals k periods apart.

Nonzero values of these autocorrelations indicate omitted predictability in the dependent variable. Here If they die off more or less geometrically with increasing lag, k, it is a sign that the series obeys a low-order autoregressive process. After taking Lag 1 we found low autoregressive process in Nifty_50 Series.

Null Hypothesis: D(NIFTY_50) has a unit root			
Exogenous: Constant			
Lag Length: 6 (Automatic - based on SIC, maxlag=22)			
		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-11.95593	0.0000
Test critical values:	1% level	-3.435514	
	5% level	-2.863708	
	10% level	-2.567974	

The fundamental requirement of ARIMA model is to know the integration of the series under study. The table number 1 represents unit root test results for Augmented Dickey Fuller test of DNifty_50 under study from 2015 to 2020. We can see from the ADF t value, all are less than ADF 5% critical value that means we reject the null hypothesis and interpret that unit root does not exists. This can be confirmed from respective probability value of less than 0.05.

STATISTICAL RESULTS OF DIFFERENT ARIMA PARAMETERS FOR Nifty_50			
ARIMA	**BIC**	**Adjusted R2**	**S.E. of Regression**
(1,1,1)	12.21866	0.000629	108.7179
(1,1,2)	12.21524	0.004042	108.5321
(2,1,1)	12.21528	0.004007	108.534
(2,1,2)	12.21714	0.002147	108.6353

Here, we checked whether the association between the response and each term in the model is statistically significant or not. If the p-value is less than 0.05 in model (1,1,2), we can conclude that the coefficient is statistically significant. In this case, ARIMA (1,1,2) was selected as the best model for Nifty50 index after several adjustment of the autoregressive (p=1) and moving average (q=2). Parameters used to select best model is the AIC and BIC together with the adj-R2 of the estimated models to detect which model is the parsimonious one (i.e. the one that minimizes AIC and BIC and has the highest adj-Rz) and Minimizes S.E of Regression.

Dependent Variable: D(NIFTY_50)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 10/29/20 Time: 14:21
Sample: 9/02/2015 9/01/2020
Included observations: 1223
Convergence achieved after 88 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.010019	2.920924	1.030502	0.3030
AR(1)	-0.448782	0.060926	-7.365985	0.0000
MA(1)	0.421951	0.064321	6.560125	0.0000
MA(5)	-0.139421	0.013351	-10.44273	0.0000
SIGMASQ	11504.80	180.8227	63.69522	0.0000

R-squared	0.026449	Mean dependent var	3.012592
Adjusted R-squared	0.023252	S.D. dependent var	108.7521
S.E. of regression	107.4804	Akaike info criterion	12.19570
Sum squared resid	14070371	Schwarz criterion	12.21758
Log likelihood	-7453.286	Hannan-Quinn criter.	12.20456
F-statistic	8.272453	Durbin-Watson stat	1.971340
Prob(F-statistic)	0.000001		

Inverted AR Roots	-.45			
Inverted MA Roots	.66 -.44+.61i	.30+.61i -.81	.30-.61i	-.44-.61i

As we discussed, ARIMA (1,1,2) model is selected as the best model for Nifty50 index after several adjustments of the autoregressive (p=1) and moving average (q=2) . In the figure of correlogram, (ACF and PACG) shows lag 6 plays an important role in forecasting Nifty50 index because lag 6 peak is higher than any other lags. So we applied Adjusted ARIMA using D(Nifty_50) C AR(1) MA(1) MA(6) estimation. The above result shows adjusted ARIMA (1,6) provides accurate result compared to ARIMA(1,1,2). Adjusted ARIMA increased. Adjusted R2, BIC and decreased S.E of regression so we can increase forecasting ability using adjusted ARIMA in Box-Jenkins approach.

Conclusion

As described above in the paper , we have analysed the data of the time series in the stock market using the ARIMA Model. Results obtained revealed that the ARIMA model has a strong potential for short-term prediction and can compete favourably with existing techniques for stock price prediction. We would like to propose an optimized ARIMA (1,1,2) model for forecasting Nifty_50 index price in short term. Adjusted ARIMA gives more accurate forecasting compared to ARIMA using Box-Jenkins approach.

Bibliography

Fattah, J. a. (2018, 10). Forecasting of demand using ARIMA model. *International Journal of Engineering Business Management*, 184797901880867

K, M. A. (2017, January). FORECASTING NATIONAL STOCK PRICE USING ARIMA MODEL. *Global and Stochastic Analysis* , 77-81

Madhu, K. M. (2014). An Application Of Time Series Arima Forecasting Model For Predicting Sugarcane Production In India. *Studies in Business and Economics , 9*, 81 - 94

Devi Uma Sundar and P. Alli (January 2013) An Effective Time Series Analysis for Stock Trend Prediction Using ARIMA Model for Nifty Midcap-50. *International Journal of Data Mining & Knowledge Management Process (IJDKP)* Vol.3, No.1, PP 65-78

Adebiyi A , Aderemi & Charles Ayo (2014) Stock Price Prediction Using the ARIMA Model. UKSim-AMSS 16th International Conference on Computer Modelling and Simulation- 978-1-4799-4923-6 PP 105-111

Dhyani B et al (March 2020) Stock Market Forecasting Technique using Arima Model, *International Journal of Recent Technology and Engineering (IJRTE)* Volume-8 Issue-6

Jun Zhang, Rui Shan and Wenfang Su, (2009), "Applying Time Series Analysis Builds Stock Price Forecast Model", Modern Applied Science, 3(5), 152-157

Mohamed Ashik. A and Senthamarai Kannan. K (2017), *"Forecasting National Stock Price Using ARIMA Model",* Global and Stochastic Analysis, 4(1), 77-81

Datta K. (2011) "ARIMA forecasting of Inflation in the Bangladesh Economy", *The IUP journal of bank management,* voI.X, No.4, pp-7-15

Swapnil Jadhav et al (2015) Indian Share Market Forecasting with ARIMA Model, *International Journal of Advanced Research in Computer and Communication Engineering.* Vol. 4, Issue 11, PP 334-336

Bijesh Dhyani et al (2020) Stock Market Forecasting Technique using Arima Model, *Jain International Journal of Recent Technology and Engineering (IJRTE)* ISSN: 2277-3878, Volume-8, Issue-6, PP 2694-2697