Stock Market Forecasting Technique using Arima Model

Bijesh Dhyani, Manish Kumar, Poonam Verma, Abhisheik Jain

Abstract: In financial market, various shares, bonds, securities or currencies are traded on the daily basis, thus making most of the datasets as time series data where price is plotted against a time series. There are many techniques and analysis technique that can be used with time series data like ARIMA model, exponential smoothing, neural networks or simple moving average. However, ARIMA model is commonly used to understand time series analysis in order to extract meaningful characteristics of the data and help in the prediction of the stock prices since it helps to understand what happened in past and past behavior of data can help to predict future. Time series is a special property and different set of predictive algorithms. There are three variants of the ARIMA model namely basic, trend-based and vector-based. In this paper, key components of time series data have been discussed and implemented using ARIMA model for forecasting the daily data of Nifty50 index and wants to predict the future value of the stock.

Keyword: ARIMA Model, Stock Market, Prices, Time Series Data

I. INTRODUCTION

Different investment comes with different level of risk and profitability but one who invest in stock market should know the higher return he expects comes in price of higher uncertainty and risk. As investment is growing people companies and economists analyse different prediction model for forecasting to minimize risk and to build the portfolio. Market Index gives an idea how a given stock market is performing so for trading in India there are two exchanges Bombay Stock Exchange (sensex) and National Stock Exchange (NSE). Many companies individual investor utilizes the forecasting technique for investment and report preparations. In financial market specially in stock market people buy and sell shares but there are other securities which are also traded like ETF (Exchange Traded Funds) corporate bonds, currency are also traded. ARIMA model are generally applied where there is evidence that data is non-stationary. In this paper, we hereby have the potential to propose the utilization of ARIMA model for the prediction of stocks' costs. An ARIMA, or autoregressive integrated moving average may be an applied mathematics analysis model that uses statistic information data to either perceive the data set or to predict future trends.

II. LITERATURE SURVEY

A lot of research utilizing the ARIMA model for the prediction purposes has been undergone. Some of the research carried out by different authors have been stated in this section. Authors conferred a hybrid intelligent system exploiting ARIMA and Neural Networks, to forecasts the demand of the merchandise, with a dramatic reduction within the sales failure than the previous solutions in context to the Chilean market. Variants of ARIMA Model include Basic, Trend Based and Wavelet Based. Hybrid ARIMA and ANN have been used for the prediction of stock market. ARIMA model cannot easily handle the nonlinear series, thus SVM can be implemented along with the ARIMA Model to forecast stock prices, which is non linear in nature. Stock Market indices predictions have been further carried out using Bacterial Foraging Optimization and Adaptive Bacterial Foraging Optimization techniques and compared with the Genetic algorithm based techniques. A hybrid intelligent data mining technique based on genetic algorithm and support vector machine is proposed to explore the forecast of the stock market. Further, many authors have also proposed a PHM (Proposed Hybrid Model) with weights determined by the Genetic Algorithm. Stock market price has been forecasted using the combination of the chaotic mapping, firefly algorithm and support vector regression, which outperformed variants of the genetic algorithm based support vector regression models. Hybrid ARIMA and ANN have been used for the prediction of the stock data. One of the base paper for this paper utilized Time Series Analysis for Stock Trend Prediction using ARIMA Model for Nifty Midcap 50 that was carried out to help people for their investment decisions.

III. OBJECTIVE AND STUDY

To predict the movement of Nifty 50 index of Indian stock market. The NIFTY 50 is the index on the NATIONA STOCK EXCHANGE OF INDIA Ltd. Which helps to follow the ongoing trend of the blue chip companies and reflects the true scenario of the Indian Stock Market on the daily basis. NIFTY 50 helps to offer exposure to the Investors and is self sufficient for the benchmarking and index based derivatives.
DATA AND METHODOLOGY
This analysis involves monthly information on the closing stock indices of NIFTY covering the period from 2014 to 2018 having a total number of 1222 observations. On the idea of this information, we have tried to determine an acceptable likelihood model to enable the ARIMA model to predict the future longer term unobserved price of NIFTY.

3.1 Data Decomposition
Data Decomposition is the main factor considered for analysis. Time series data is cleave into different parts: seasonality, trends and random oscillate, that can be used to identify trend from past data [15].

From the above diagram we can conclude that there exists seasonality in time series data.

3.2 ADF
After decomposing the data the next step is to determine visually whether data is stationary or not. As to use ARIMA effectively stationary is an important concept. A stationary series has constant mean and variance over time, a stationary dataset will allow our model to predict that the mean and variance will be same in future trends. There are mathematical test which can be used to test stationarity in dataset one commonly used is ADF(Augmented Dicky Fuller Test). It’s important to test the stationarity of data to apply fitting ARIMA model. In performing the test is assumed that data is non stationary and for performing ADF test following are p values for different time period
For whole series the value of p=0.99, for 0-3 months p=0.1803, for 3 to 6 months p=0.6713, for 6 to 12 p=0.7001 and for 12 + months p=0.8331 [15]. Data need to be stationary either visually or mathematically. To make data suitable for analysis one simple way can be differencing also but we carry out Augmented Dickey Fuller Test which is considered unit root test for stationarity causes result which are not known earlier for time series. This is why data needs to be stationary. Now after the data is stationary we can proceed further by plotting the autocorrelation and partial auto-correlation.

3.2. A Autocorrelation
An Autocorrelation plot shows the correlation of time series with itself, lagged by x time units so y axis so x axis is correlation value and x axis is number of units of lag. Our main priority here is to figure out whether we will use AR or MA components for ARIMA model or both and to also decide how many lags we should use. When we actually apply AR or MA the value of p and q need to be set. If the autocorrelation plot show positive autocorrelation at the first lag(lag - 1) then it is recommended to use Autoregressive(AR) in relation to lag. If this plot shows negative correlation at the first lag the one can use Moving Average model. This will allow us to decide what actual value of p,d,q to provide to ARIMA model. The value of p is decided by number of lag observations included in the model, value of d is decided by number of times raw information is differenced and q which is the size of moving average also called order of moving average [14]. It is denoted by ACF and ranges from -1 to +1.
1. Non-seasonal ARIMA

Non seasonal ARIMA models are typically denoted ARIMA(p,d,q) where p,d,q are considered to be non negative integers. AR(p). Its Autoregression that utilizes the dependant relationship between a current observation over previous time frame. Its a basic regression task. I(d). Its the Integrated portion of ARIMA model which carries out the differencing of observation(subtracting an observation from an observation at previous time step) in order to make series static. MA(d). It’s a Moving Average which uses dependency between an observation and a residual error from a moving average model applied to lagged observations [8].

2. Seasonal ARIMA

The seasonal ARIMA is for seasonal data model incorporates both non-seasonal and seasonal factors in a multiplicative model. ARIMA[p,d,q] x [P,D,Q]S with p = non-seasonal AR order, d = non-seasonal differencing, q = non-seasonal MA order, P = seasonal AR order, D = seasonal differencing, Q = seasonal MA order, and S = time span of repeating seasonal pattern[16]. In this study we have applied non-seasonal ARIMA.

RESULT AND ANALYSIS

The prediction for the stock data is given in the table as described below.

<table>
<thead>
<tr>
<th>Date</th>
<th>Actual Price</th>
<th>Predicted Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018-11-26</td>
<td>10628.599609375</td>
<td>10519.1969056294</td>
</tr>
<tr>
<td>2018-11-27</td>
<td>10685.599609375</td>
<td>10635.1627452067</td>
</tr>
<tr>
<td>2018-11-28</td>
<td>10687.267454215</td>
<td>10688.4766817159</td>
</tr>
<tr>
<td>2018-11-29</td>
<td>10858.7001953125</td>
<td>10730.5745566998</td>
</tr>
<tr>
<td>2018-11-30</td>
<td>10876.75</td>
<td>10867.3363402093</td>
</tr>
</tbody>
</table>

IV. CONCLUSION AND FUTURE SCOPE:

As described above in the paper, we have analysed the data of the time series in the stock market using the ARIMA Model. As shown in the results, the predicted price using the ARIMA Model is very close to the Actual price on the given dates. This indicates the successful implementation of the ARIMA Model on the time series data. In the future, we hope to analyze the dynamic data of the stock market. We would like to propose an optimized ARIMA model for the dynamic data in the future.

REFERENCES

Stock Market Forecasting Technique using Arima Model


AUTHORS PROFILE

Bijesh Dhyani, Pursuing Ph.D, from Graphic Era Deemed to be University and has completed his MBA from WBUT technical University, Kolkata. His interest of research is Blockchain ,Machine and Data Science

Dr Manish Bisht, Associaite professor in Graphic era Universtity has completed MTech(CS) MBA and PhD(Management) and his area of research is Big data machine learning

Poonam Verma, Pursuing Ph.D, from Graphic Era Deemed to be University and has completed her M.tech(CS) from UP technical University, Lucknow. Her interest of research is Blockchain and Machine Learning.

Abhishek Jain, Pursuing Ph.D, from Graphic Era Hill University completed M Tech(CS) from Graphic Era university.His intreset of research is Big data Machine Learning and Network Security.