Stock Price Prediction of RVNL L.T.D Using Time Series Analysis

Project to be submitted as a self project
By

Shreyasi Mondal

M.Sc. Statistics 1st Year Student in IIT Kanpur,
Completed B.Sc Statistics Hons. from St. Xavier's College (Autonomous), Kolkata

DECLARATION

I affirm that I have identified all my sources and that no part of my project paper uses unacknowledged materials.

Date : 29.07.2024 Shreyasi Mondal.

CONTENTS

INTRODUCTION

- What does Stock Price Prediction mean ?

— Stock Price Prediction is simply the act of trying to predict the future value of a company stock or other financial instrument traded on an exchange.The entire idea of predicting the stock price or forecasting the next price values is actually to gain significant amount of profit.Stock Price Prediction using machine learning algorithm helps you discover the future value of company stock and other financial assets traded on an exchange.

“The successful prediction of a stock's future price could yield significant profit. The efficient market hypothesis suggests that stock prices reflect all currently available information and any price changes that are not based on newly revealed information thus are inherently unpredictable.”

- Importance of Stock price prediction : Stock market prediction is a major challenge owing to non-stationary, blaring, and chaotic data, and thus, the prediction becomes challenging among the investors to invest the money for making profits.

DATA SET

Source of Data : Collected Data on Daily stock prices of RVNL NS (Rail Vikas Nigam Limited) from Yahoo Finance Website.

DETAILS OF THE DATA SET : In our data set we have daily opening and closing prices of RVNL from 15/07/2019 to 15/07/2024. Along with that this data set also contains information like daily highest , lowest stock prices and total volume of the stock traded that day.

In our data set we have a total of 1236 data points which means it has 252 trading day data points every year.

Suppose we consider variables Y, Xt,t,T. Descriptions of the variables are given below:

1. Yi : denotes the daily trading days of the ith year, where i=1(1)6
2. Xt : denotes the daily closing price of that stock
3. Tt : be the value of trend present in the data set

METHODOLOGY

Step (1) - DATA VISUALISATION

At first we will plot the data and start visualising the dataset. Then we will interpret from the graph that if any fixed pattern is present there and if trend , seasonality, cyclical pattern or irregular variation is present there.

Step (2) - FITTING OF AN APPROPRIATE TREND EQUATION

From the graph we will try to analyse if any pattern is present there or not.If any fixed pattern is present there then we will fit the corresponding trend equation there.

Step (3) - FITTING OF AR(1) , AR(2) , MA(1) , MA(2) , ARIMA MODEL

After that we obtained the residual series , by removing the systematic part. Then we fitted an AR(1) process , AR(2) process. We also fitted MA(1) , MA (2) and ARIMA models.

Step (4) - COMPARING THE GOODNESS OF FIT

Thereafter we compared the goodness of fit of the following models using residual sum of the squares i.e MAD (Mean Absolute Deviation) . We declared that process as the best one which had the lowest value of MAD.

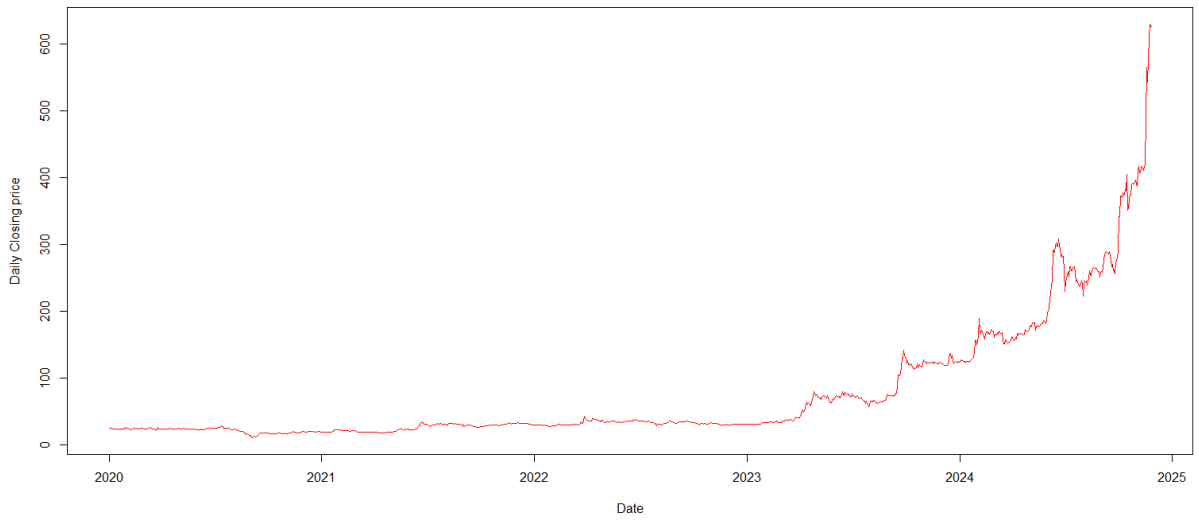
Step (5) - FORECASTING OF STOCK PRICE FOR NEXT 30 DAYS

Then using the best model we forecasted the next 30 days stock prices and plotted them.

RESULTS AND DISCUSSION

(1) Now at first we will plot the data and interpret our findings.

1. Time series plot of RVNL Daily closing price Data



Comment : From the graph we can clearly notice that there is an increasing trend present. Again from the graph we can observe that the graph is showing an Exponential Trend.

(2) Since we observed that an Exponential Trend equation would be appropriate, that's why we fitted Exponential Trend equation in our data set.

Now to detrend the data, we need to fit the exponential trend equation to the deseasonalized data.

Let, Tt=ab^t , $a \neq 0$

Here we got, $a = 47.17797$

$b = 1.001124$

Tt = (47.177797)*(1.001124)^t

Where t be the required year.

(3) Now we fit the AR(1) model first on the residual series which we obtained using the decompose function in R.

The model of AR(1) is given by,

$$Yt = \alpha(X(t-1) - \mu) + zt$$

After that we calculated the value of MAD for AR(1),

MAD_AR_1 = 0.9563167

Similarly we fitted AR(2) model , the model is given by,

$$Yt = \alpha 1(X(t-1) - \mu) + \alpha 2(X(t-1) - \mu) + zt$$

MAD value for AR(2) is given by ,

MAD_AR_2 = 0.9563514

(4) Mean Absolute Deviation (MAD) value is lowest for AR(2) process.

That means this is the most appropriate model.That's how we computed goodness of fit of the following models.

(5) Finally using the ARIMA model we forecasted the next few days share

prices. Now we fitted ARIMA MODEL and forecasted the next 30 days share prices.
The snapshot of the forecasted data is given below,

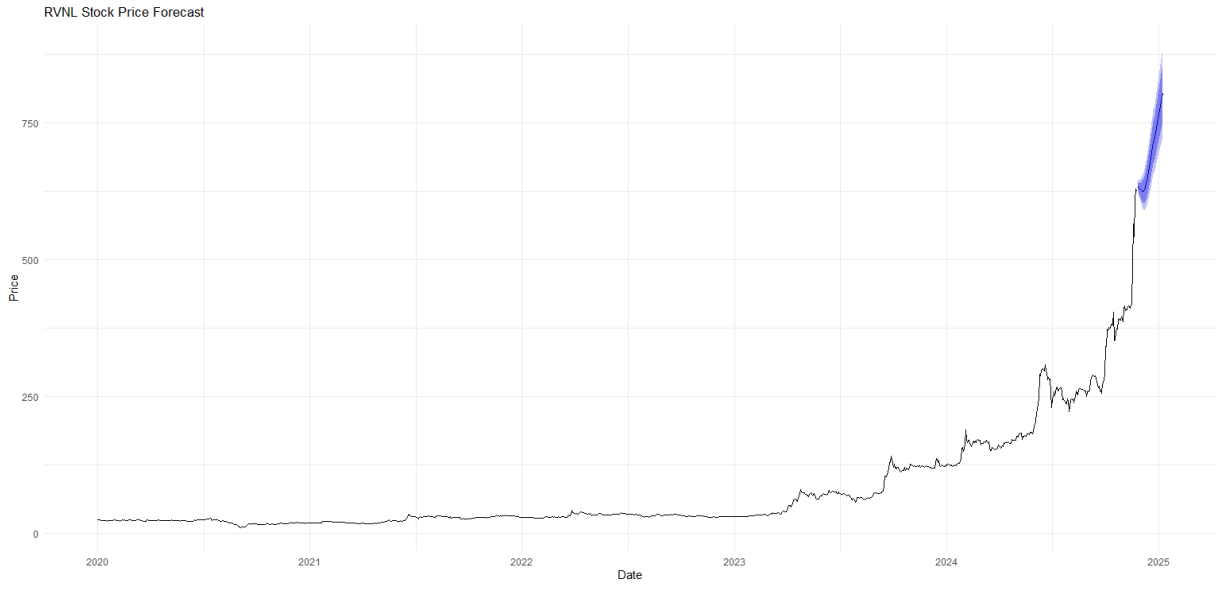
forecasted_values

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2024.905	634.3573	627.4768	641.2379	623.8345	644.8802
2024.909	631.1713	621.3879	640.9547	616.2089	646.1337
2024.913	627.3510	614.9112	639.7908	608.3259	646.3760
2024.917	626.1642	610.8729	641.4556	602.7782	649.5503
2024.921	625.5412	607.4066	643.6758	597.8068	653.2757
2024.925	624.7054	604.1756	645.2353	593.3078	656.1031
2024.929	625.7735	603.0759	648.4710	591.0606	660.4864
2024.933	628.5113	603.8451	653.1775	590.7876	666.2350
2024.937	632.4287	606.0161	658.8413	592.0341	672.8232
2024.940	637.5957	609.6345	665.5568	594.8328	680.3585
2024.944	644.0360	614.6684	673.4036	599.1221	688.9499
2024.948	651.4184	620.7599	682.0769	604.5303	698.3065
2024.952	659.5194	627.6597	691.3791	610.7943	708.2446
2024.956	668.1887	635.1878	701.1897	617.7182	718.6593
2024.960	677.2375	643.1299	711.3450	625.0745	729.4004
2024.964	686.4810	651.2825	721.6796	632.6495	740.3126
2024.968	695.7917	659.5005	732.0829	640.2891	751.2943
2024.972	705.0685	667.6689	742.4682	647.8707	762.2663

2024.976	714.2298	675.6951	752.7645	655.2960	773.1636
2024.980	723.2234	683.5192	762.9276	662.5010	783.9458
2024.984	732.0234	691.1099	772.9369	669.4516	794.5952
2024.988	740.6210	698.4554	782.7866	676.1342	805.1077
2024.992	749.0216	705.5602	792.4831	682.5531	815.4902
2024.996	757.2427	712.4423	802.0431	688.7264	825.7590
2025.000	765.3085	719.1279	811.4891	694.6814	835.9356
2025.004	773.2469	725.6476	820.8462	700.4501	846.0438
2025.008	781.0870	732.0336	830.1404	706.0663	856.1076
2025.012	788.8567	738.3173	839.3962	711.5633	866.1501
2025.016	796.5817	744.5275	848.6359	716.9716	876.1917
2025.020	804.2839	750.6892	857.8785	722.3180	886.2498

>

Plot of next 30 days forecasted value



Here the blue part is the forecasted values

CONCLUSION

From the above project we can conclude that ARIMA model fits best on the RVNL daily stock price data set.There fore we used this model to forecast the Next 30 days stock price values. Then we also plotted them.