iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 Issue: X Month of publication: October 2021

DOI: https://doi.org/10.22214/ijraset.2021.38575

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue X Oct 2021- Available at www.ijraset.com

SARIMA – A Model for Forecasting Product order demand

Pratiksha Rajendra Dharmadhikari
Bharati Vidyapeeth

***Abstract:** Product analysis is the most important part for any working manufacturing. It provides the sales record of their currently manufactured product and also it helps to predict its performance in the future. For this analysis, a SARIMAX model has been used with Time series forecasting. This paper will explain the need of such model instead of using a simple regression model to predict the order demand. This study analyses and presents a forecasting model to predict an order demand for the Product over the time period. Demand in Product is a main component for planning all processes in supply chain, and therefore determining Product demand is a great interest for supply chain. Mean forecasting for product order demand was carried out using SARIMA model, by using the past data from the period of 2011 to 2017. The model with the least value of Akaike Information Criterion (AIC) was selected as the appropriate model for forecasting mean Error. Test for normality of residuals were performed to see the adequacy of the chosen model. SARIMA (1, 1, 1) (0, 1, 1) (12) was selected as the best model for mean product order demand forecast. The results obtained will prove that the model could be utilized to forecast the future demand in the Product manufacturing industry. These results will help the manufacturers for manufacturing reliable guidelines in making decisions.*
***Keywords:** ARIMA, AIC, S-ARIMA, Regression*

I. INTRODUCTION

In continuous evolving manufacturing market adapting the changes in supply & demand is a prime importance. Currently organizations are more inclined towards developing most effective demand supply chain model. The manufacturing market has involved with customers for demanding and discriminating the supplier what products they want and when they need them. Inventory management has a difficult time forecasting demand. Inventory management levels depend on demands that coming from the customers. The incorrect evaluation for demand can cause huge loss to the manufacturing companies, which proves that the process is not correct. However, large investments requires in inventories. you cannot define when there are no demand and have continues demands in stock. There is time when the product has several periods of no demand these will leads the difficulties for traditional time demand forecasting methods to predict the orders demand for future. For this time series model are used which considers time as main properties for Predictions. In the time series analysis, the datapoints are sequentially measured at particular time period. This method is trying to understand the content of the data points, or make a prediction about the future values of those particular data points. Many methods can be used for forecasting, but there is challenges of surety and availability of the Product which are needed at all time. Therefore, it needs demand prediction for Product orders an inventory to assure the Product in hand. The main Reason for this study is for manufacturers to manufactures the product according to the demand and make the decision so there will be no chances for any losses. This study deals past order data recorded for product order demand from year 2011 to 2017. Seasonal Autoregressive Integrated Moving Average model was developed to obtain the forecast of 2017 monthly root mean error. The aim of this study is to model and forecast the Product demand by using Box–Jenkins's time series approach, with SARIMA model. To achieve this goal, large and historical demand data from January 2011 until January 2017 has been used. There are many ARIMA models were developed and evaluated by the performance types AIC and SBC. The new historical order demand data is used for validating the best model in same conditions.

II. LITERATURE REVIEW

A. *Forecasting Demand*

The forecasts are the main for survival of business as in today's organizations, the changes are affecting rapidly and vastly to the well-established structures of organizations. Here, all business market needs the exact and practical correct reading for the future. In forecasting the predicting the future are done using the variables, and these variables can demand, supply or price. the most the time these variables are demand based. In forecasting by studying these variables the future values are assumed.

The most issues in inventory management are forecasting demand while productions. These is often considered while operation planning, in assembly process, in capacity planning and gained used product. In supplying, the demand forecast is considered as a base for supply planning. There are processes which can called as pull Processes for customer demand and Push processes for considering or hoping for the customer demand. Choosing the best method for forecasting is the most important decision a company has to make. These company must consider the Pull and push process while the selecting the method. There are mainly four types of forecasting model are used. Which can be used for qualitative, time series, causal, and simulation.
In these the time series forecasting model are used for prediction of order demands[5]. In Time series model the order of time in the past are observed. The past data over a time period are used for forecasting the demand. In many studies it is stated that the development of future will depend on its past data that's why the historical data are need to be used for future demands. The demand forecasting studies are done in many different regions such as for food product sales, travelling, repair parts, electricity, automobile, temperature and some other parts. In the time series analysis, the accuracy for forecasting demand are depends on the attributes of time. we can assume that if curves in the graph shows the non-moving and consistency in pattern then there is a high accuracy forecasting and if the curves shows the moving/non stable and irregular patterns then this will not be assumed for high accuracy forecasting.

B. Autoregressive Integrated Moving Average

In ARIMA the time series are analysed by past values which are found by the model and forecasting that time series. As AR is an acronym that stand for Autoregressive model that depends on the relationship between the observations and some lagged observations, I for integrated to make the time series stationary also supports the data with its trends and MA is the Moving average model that apply on the lagged observation to use the dependency between the observation and residual errors[2]. The equation of ARIMA is used for forecasting the values in future by lagged forecast errors. The model can use the non-seasonal time series which does not having the white noise and showing the patterns. The classical ARIMA approach are most expensive in many cases because when the seasonal adjustment is done if seasonal order arrangement is high or they did not show the series in stationary pattern. ARIMA model static parameters are used for forecasting the seasonal demands also the model requires the huge number of observations for calculating the best fit model of data series. ARIMA model each component are defined as an ARIMA (p, d, q), where p is autoregressive terms, d is differences, q is moving averages. These variables are taken as a integers for the evaluation of best fitting model. The main issue with ARIMA model is that it doesn't support time series which has repeating values cycle.

C. SARIMA

SARIMA model is an extended ARIMA model which supports the time series data with seasonal trends**.** when there are seasonal patterns, and need to add a seasonal term then SARIMA model are used. The ARIMA model are like linear regression model that uses its own lags for predictions. These models can greatly perform when the predictors independent with each other and they are not correlated. when the previous value of a series is subtracted form the current value then series can be made as stationary. These differencing of the series are needed as the complexity of the series. for stationary series d elements are considered the minimum number of differencing as required. And if in case d value is 0 then there is no need for making the time series stationary as it is considered as a stationary series already. The SARIMA added new hyperparameters for defining the ARIMA model with seasonal and trends elements. The SARIMA components contain seasonal ARIMA (p, d, q) (P, D, Q) m as the P, D, Q, m are trends elements where p is trend for autoregression order is for moving average order and d is for differencing order. These elements are the same as in the ARIMA model. The four seasonal elements are used in the SARIMA model which are not part in ARIMA are P, D, Q, m where P is seasonal autoregressive order, Q is for seasonal moving average order, D is for seasonal differencing order and m is for time steps numbers for single seasonal period. Here for Prediction the external Prediction variable is used that is also called as exogenous variable. By using these variables, the model is also known as the SARIMAX model[2].

D. Dataset

The dataset used in this study is "Historical Product Demand" and collected From Kaggle website[1]. It contains a daily data of products orders demand from Jan 2011 to Jan 2017.The dataset consists of above 10 lakh data. These datasets consist data samples from years to daily and weekly basis including various data, which are the product number, the product category type, the warehouse, the Date and order demand. In this the Product category type are used to differentiate the product, whereas warehouse is defined product from which it is manufactured and sold and Date at which a date the particular product was manufactured and sold. The dataset consists of 1,048,575 dated entries of different products for different warehouses with different Product code.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue X Oct 2021- Available at www.ijraset.com

There are 2,160 different products that shows to Product code, 4 warehouses, and 33 different product categories. All these data are considered as to build the forecasting model for order demand to calculate the number of products that sold and their categories. the dataset are used for predicting changes need to be done for the product in the future using the time series forecasting. Here as the dataset are having the random values of each variable so these will leads to create an relationship between the variables for a successful model and to accessing the sale. The dataset are very complex that's why the simple regression methods cannot be apply on them to make a reliable model.

E. Implementation Model

The dataset contain the data from 2011 to 2017.The dataset are verify if there is null values present or not in them and then null values are removed from the dataset. Here while visualizing the data its also verify that why a linear model without any transformation would fail to capture the elements of the data.

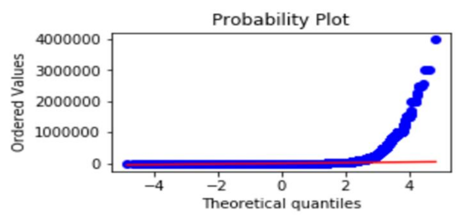


Fig A. distribution of order demand   Fig B. Probability plot of order demands

Some different patterns appear in the results. The overall demand of the products category wise, ranging from 1 to 28. Next, is the how much demand is coming from the customers for each of the warehouse that is providing the products to the market.

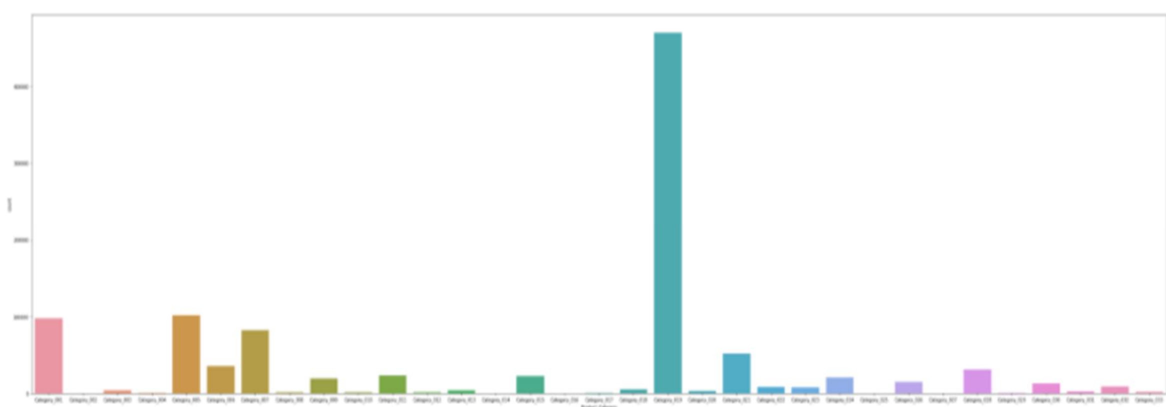


Fig no. (3) order demand by product category

While visualizing the dataset some discontinuous patterns are seen in the graph. The data are prepared as the raw data and make it suitable for the model. i.e. Creating the test and train data frame from the dataset according to the date format. Test and train data frames are used for predicting the order in demand. The test and train data are mainly divided by train to 70% and for test 30% data from the dataset. When the preprocessing is done, while exploring the dataset you can see the product order is always low from the beginning of the year. As in from the graph the highest peak in product order demand is in the mid of 2015. In 2014 to 2016 the observed trend shows that orders were highest Peak and then it is reducing down slowly.

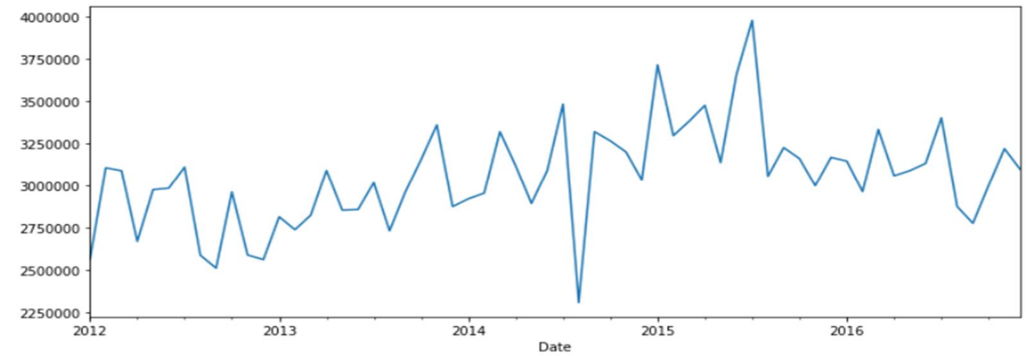


Fig no. (4) time series for order demand

The Following graphs show the overall observed changes over the period by decomposing into trend, seasonal and residual patterns.

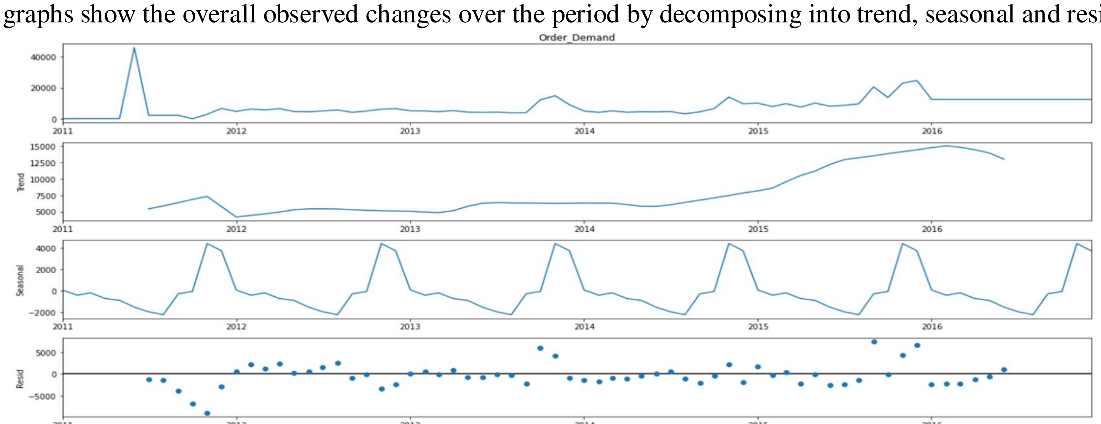


Fig no. (5) Decomposition of time series

Box and Jenkins approach is used for developing an order to perform SARIMA models[3]. The Box–Jenkins firstly check the model identification, then the parameter evaluation, and properties checking. They must have some theoretical autocorrelation. By matching the theoretical and empirical autocorrelation patterns, it easy to make it possible to identify one or several correct models for the given time series. To identify the order of the SARIMA model, Box and Jenkins recommended to use the autocorrelation function (ACF) and the partial autocorrelation function (PACF) of the sample data, there is residuals in a time series model that defines what is left after the model are fitted. differencing the observation and similar fitted values show the residuals of the model.

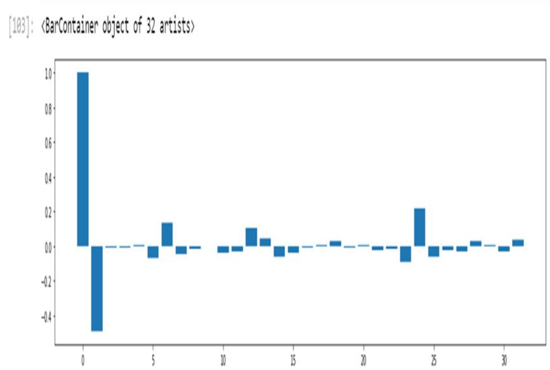


Fig.(6)ACF

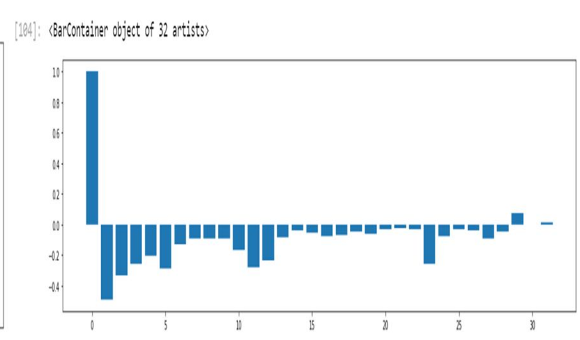


fig.(7)PACF

The AR and MA model seasonal part will be seen in the seasonal lags of the PACF and ACF. The most significant spike at lag 1 in the ACF suggests a non-seasonal MA component, and the significant spike at lag 24 in the PACF suggests a seasonal MA component. Here Both the ACF and PACF show significant spikes at lag 0, displaying some additional non-seasonal terms that need to be added in the model. From the ACF and PACF the parameters for the Arima to fit model are used.

The Akaike information criterion (AIC) are used for relative quality of time series models dataset. AIC measures and tells that how well a model fits the data while taking into account the overall complexity of the model. SARIMA Model - (1, 1, 1)x(0, 1, 1, 12) are selected as the best model. Where (1,1,1) are trends order and (0, 1, 1, 12) are seasonal order that are fitting into the model will give the analysis of Co-efficient and Standard Error by going through the model. Co-efficient check for weight that how each feature effects the time series and P-value for significance of each feature weight to ensuring for statistically significant or not. By default the null hypothesis as 'there is no relationship between them' are choose.

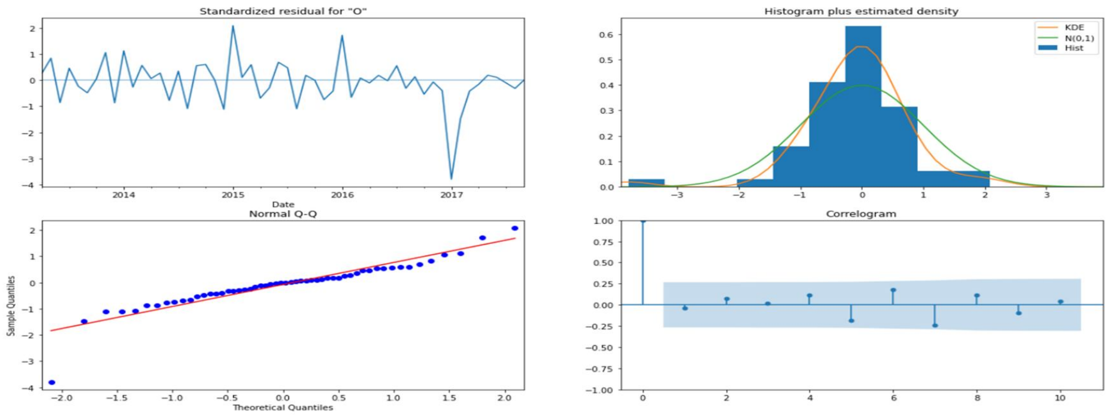
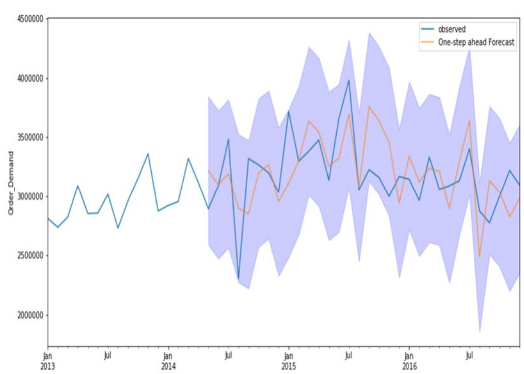


Fig A.standard residual for order demand B.Histogram for density C.Normal QQ D.Correlogram

The Histogram graph show the standard deviation observed, which gives the data response and fitted response. The KDE i.e. Kernel Density Estimation line must be closely matched with green colored N(0,1) line With a mean of 0 and a standard deviation of 1, this is the conventional notation for a normal distribution.The Normal QQ Graph shows the ordered distribution of residuals follows the definite trend of the samples taken from a standard normal distribution with N(0, 1). Here the standard residuals doesn't display any seasonality and appear to be white noise. The autocorrelation plot on the bottom right, shows the time series residuals have low correlation with its own lagged versions. Model accuracy are calculated by mean root squared error. In which the lowest the mean root squared error are considered to be the best accuracy for the model. The mean Error for the model SARIMA(1, 1, 1)(0,1,1,12). The model gives root Mean squared error are 57.01 which are used to check the accuracy. The graphs in below shows the output of all the analysis done till now. In upcoming years from the graph can be observed the changes in demands. How will the sales go through over period of time which transition can be observed and studied from the presented output graphs using the SARIMA model in time series analysis.



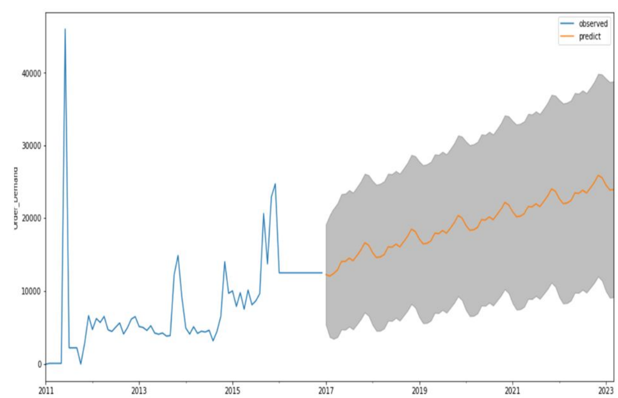


Fig. no. (9) observed trend vs one step ahead of forecasting Fig. no. (10). observed trend vs forecasting for next 5 years.

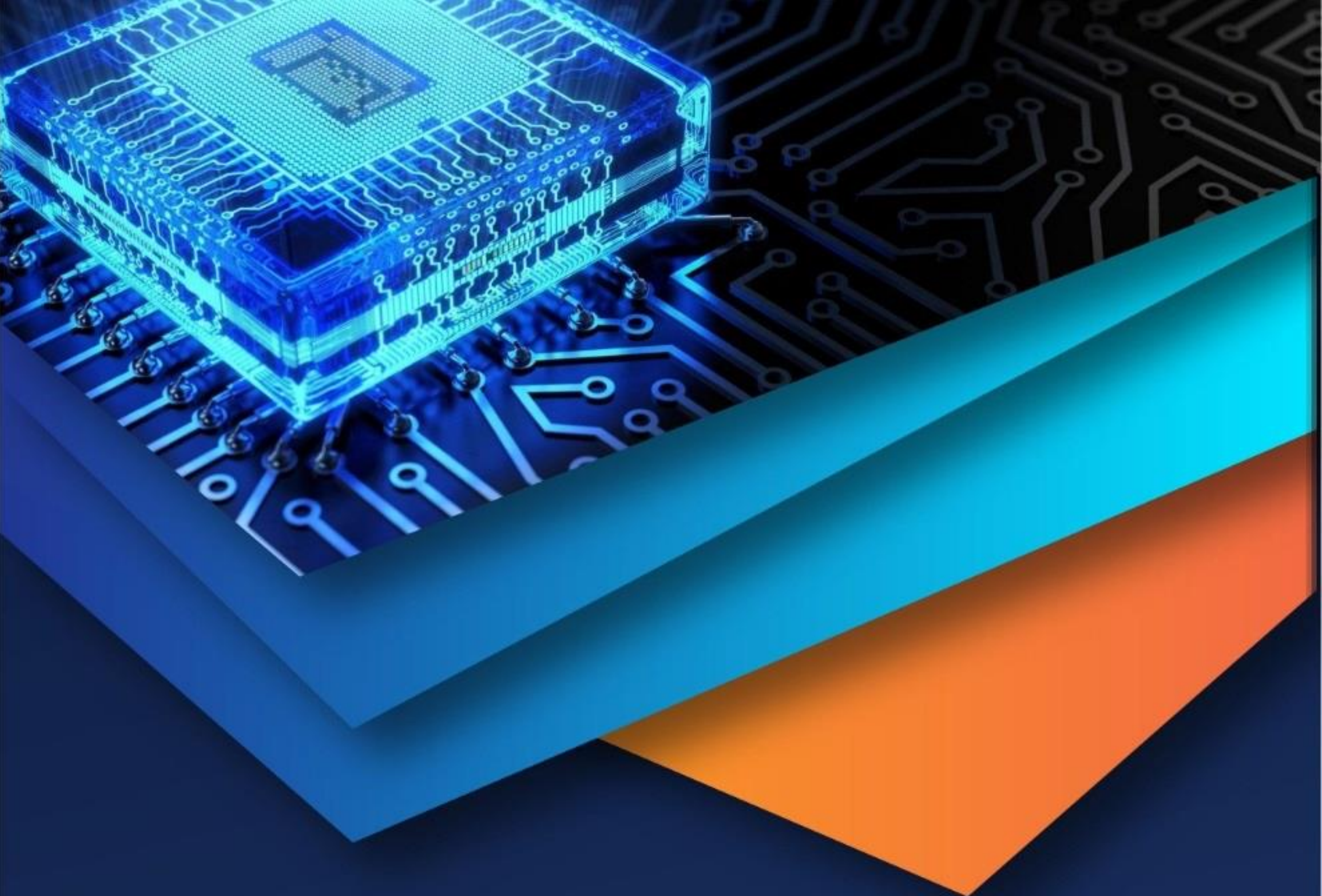
The SARIMA Model are Compared with three more time series model.so to choose the best model for Prediction for product orders demands. The model which are used are NAÏVE Method, simple average and simple exponential smoothing methods of time series models. These time series models can be considered as checking the prediction for order demands.by these method the root mean squared error are calculated. The Naïve method gives 148.56, simple average has 233.9 and simple exponential smoothing has 146.2 Root Mean squared error.

III. CONCLUSION

The SARIMA model is feasible to predict the behavior of businesses for the futuristic product demand. The model is very best suitable for the companies to predict the life cycle of a product i.e., points of growth, maturity and decline. The model predicts more accurately with the accurate data while lowering the time intervals in algorithms. The results show that SARIMA is the best suitable method for predicting the product order demands over the Naïve, simple average and Exponential smoothing method.

REFERENCES

[1] https://www.kaggle.com/felixzhao/productdemandforecasting
[2] https://www.machinelearningplus.com/time-series/arima-model-time-series-forecasting-python
[3] Box G., Jenkins G., Reinsel G., Time Series Analysis Forecasting and Control. 3rd ed., Prentice Hall, 1994
[4] Ruchir Kulkarni1, Milind Rane,'Pattern Recognition - Product Sales Analysis Using SARIMA Model in Time Series Forecasting' ,2018
[5] Box GEP and Jenkins G. Time series analysis, forecasting and control. San Francisco: Holden-Day, 1970.
[6] XiaoxinZhu1· GuanghaiZhang1· BaiqingSun2,A comprehensive literature review of the demand forecasting methods of emergency resources from perspective artificial intelligence,2019
[7] Gerald Reiner1 , Johannes Fichtinger2 ,Demand forecasting for supply processes in consideration of pricing and market information ,2009



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY