Using SARIMA Method to Forecast Office Supplier Demand

Muhammad Nadeem[1], Bilal Shah[2], Danish Nabeel[3], Muhammad Abdul Mohsee[4], Muhammad Azam Chughtai[5]
[1]Assistant Professor, Department of Computer Science & IT, Sir Syed University of Engineering & Technology, Karachi, Pakistan
[2,3,4,5] Department of Computer Science & IT, Sir Syed University of Engineering & Technology, Karachi, Pakistan

Abstract: **It becomes problematic specifically for the organizations that provide services of office supplying, to consider the conditions of marketplaces. The anticipated investigation tends to forecast the requirements associated with services as per auction or sale in the commercials of office suppliers so that there is no deficiency of the numeral units on the portion of the catalog (inventory). This study emphasized demand prediction as an instrument for demonstrating the precise numeral of services to be manufactured or produced. In this research, we have created a web-centered system of demand forecasting that builds on hugely and generally models named as SARIMA model (Seasonal Autoregressive Integrated Moving Average) for estimating the service and products regular demand (monthly or daily) for the attained dataset. This system for demand prediction (forecasting) has been imposed utilizing the model of flask web technology and the language of Python programming. The framework is flexible, forceful, and comfortably attainable and comprises the capability to investigate a very huge quantity of data.**

Keywords: Office Supplier, Time Series, SARIMA, Forecasting, Demand.

I. INTRODUCTION

Commercials providing services of office supplying have located estimating (forecasting) entirely challenging. Prediction of sales provides the supplier with a huge framework. They attain a normal concept of the incoming period so that they can be supported to develop their aims to increase their progress and profit. Precise forecast of demands can direct to better sales of services, storing management, presentation associated with a business decision, the fulfillment of requirements of consumers and reduce in expenses thereby stopping deficiency or blockage of service on the catalog side. It is due to huge development in IT (Information Technology) in several sectors. It gets as a rough, challenging, and sensitive task to accomplish with an efficacious model of sales forecasting as instability of services sales has risen regularly.

II. RELATED WORK

A study conducted by Kilimci et al., [4] compacts with methods of Regression and Time Series. In the given framework, nine sorts of algorithms including ARIMA (Autoregressive Integrated Moving Average), MA (Moving Average), Holt-Winters, and MA (Moving Average), and three distinct models of Regression were imposed. They also comprise the utilization of SVR (Support Vector Regression) which involves regression imposition of SVM (Support Vector Machine) for cataloging and forecasting of incessant variables. The profound method of learning deliberated here is the MLFANN (*Multilayer Feedforward artificial neural network*) in which calculations and information flow just in a forward direction without any assistance. This was skilled utilizing "*stochastic gradient descent*" centered on backward promulgation. The machine-learning algorithm of Holt-Winters is utilized for estimating the sales of Walmart organization in his research by Harsoor, Anushree Patil, and Anita S. [5]. The trend seasonality and residual randomness were estimated in the given algorithm and the forecasting of sales was accomplished through a proficient dataset.

YU, Jian-hong, and Xiao-Juan LE [6], in which they investigated three methods of prediction on a dataset of Amazon, conducted an investigation. The initial was Exponential Smoothening by Winter which was placed into the understanding of both seasonal patterns and trends when the procedure of smoothing was imposed. The subsequent model was a decomposition model for time series with the practice of Box Jenkins. This was utilized to evaluate a direct trend of information, alongside a cyclical factor and seasonality index. The conclusive methodology for prediction was ARIMA where multiple models with the least RMSE were imposed for estimating the sales.

A study conducted by Babai et al., [7], in which they utilized a model of ARIMA which demonstrates an association among MSE [Mean Square Error] and catalog expenses in a supply-chain of two-stage comprising a retainer fronting a non-stationary ARIMA (0,1,1) requirement procedure and one producer.

III. METHODOLOGY AND IMPLEMENTATION

The dataset comprises more than ten thousand accesses of service purchasing over four years, with three chief groupings. The functionalities associated with the dataset includes Order ID, Order Date, Ship Date, Row ID, Customer Region, Customer ID, Ship Model, Customer Name, State, Country, Segment, Product ID, Postal Code, Region, Product Name, Sub-Category, Category, Discount, Sales, Quantity, and Profit.

A. Data Preprocessing

The dataset comprised of many functionalities, few of which were either unrelated or contained unnecessary influences on requirements of services. Therefore, such sort of functionalities was detached just Quantity sold and Order date. Then concentrated dataset was examined for any absent or misplaced data and collected on numerical values for a date. Since the objective was to attain regular (monthly) prediction, data were re-sampled through the utilization of

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 02, February-2022

average regular numerical data of sales for that period and utilizing the initiating of the period (month) as a timestamp.

B. Evaluation

The models are related centered on arithmetical evaluation – RMSE (Root mean squared error)

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}(y_j - \hat{y}_j)^2} \qquad (1)$$

In the above formulae, yi is forecasted measure; yi is forecasted quantity comprised of variables measured n times.

C. Model Deployment

Lastly, we need our model to be attainable for end operators so that they can create utilization of it. Model deployment is considered as one of the ending stages of any project of data science. Utilizing Flask to impose a prediction model. Flask is the structure of a web application written in Python.

D. System Architecture Diagram of Web Application

The model illustration given below demonstrates a general model for requirement or demand prediction in the existing system displayed in Figure 1.

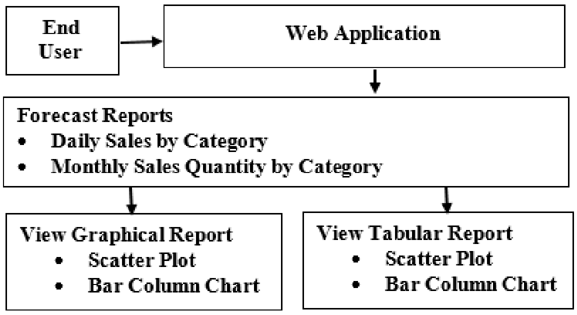


Figure 1: System Architecture Diagram of Web Application

IV. FORECASTING MODEL

This portion comprises a summary of prediction techniques utilized by SARIMA. Once pre-processing of data is accomplished, it is uploaded in given models, which are equipped through the utilization of Python.

SARIMA

Centered on category and dates, Seasonal ARIMA (SARIMA) was imposed for training the model. It is considered as an addition of ARIMA which also comprises a component of seasonality, because of this season SARIMA is utilized rather than ARIMA as information is located to display seasonal aspects (frequently every 12 months). The autoregressive portion is for the aspect of development (growth) or decline in date, its ratio of varying is accounted through integrated segment and any sound among consecutive segments of data is accounted through a segment of moving average. SARIMA comprises about seven hyper-parameters (p,d,q) – the non-seasonal segment and seasonal segment includes (P, D, Q,s). The basic formulae for SARIMA (p,q,r) x (P,D,Q,s) is given as;

$$\phi_P(B^s)\varphi(B)\nabla_s^D\nabla^d x_t = \Theta_Q(B^s)\theta(B)w_t \qquad (2)$$

Here $\phi_P(B^s)$ is declared as seasonal AR section of edict P, $\Theta_Q(B_s)$ the seasonal affecting average machinist of edict Q, ∇^DS and ∇^d are seasonal and regular variances respectively, w_t is a procedure of white sound, $\varphi(B)$ and $\theta(B)$ are imposed as the polynomials identifying autoregressive and moving typical section of the edict (order) p and q respectively. B is declared as backshift mechanist or operator and is numeral of time stages for a solely period of season (in our instance 12 as the series recurrences every 12 months). The model comprises the capability of predicting demand of every category of service for imminent months centered on aspects detected in about four years of training data utilized. Time-series as the sequence of well-defined information evaluated at unchanging era over a while.

It is also to be considered that gathered data over an unrealistic or irregular period does not create a time series. Expressive functionalities and statistics of data can be found through the investigation of time series. That data can be little beneficial for monitoring and predicting the points associated with data through adjusting the precise model to it. Since the numerical data comprises several distinct categories and sub-categories and dates can be complicated to operate with, the general of regular sales for every month have been imposed. Values pf p,d, and q are selected through the utilization of an estimator AIC. The least of value in AIC executes a more optimal solution. It is given as.

V. RESULT

The web application was created and written in Python through the assistance of the framework of Flask web, the chief border of web application executed in given figure 2. The application tends to predict in three categories including office supplies, furniture, and technology. It tends to forecast quantity and sales. Sales can be forecasted regularly (daily) and monthly while quantity can be forecasted monthly.

Figure 2: User Interface.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 02, February-2022

Figure 3: Report Forecasting Option

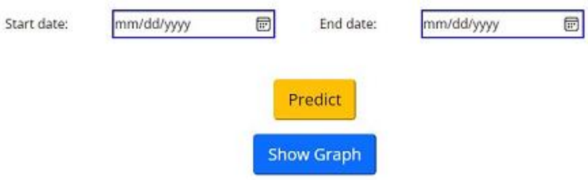
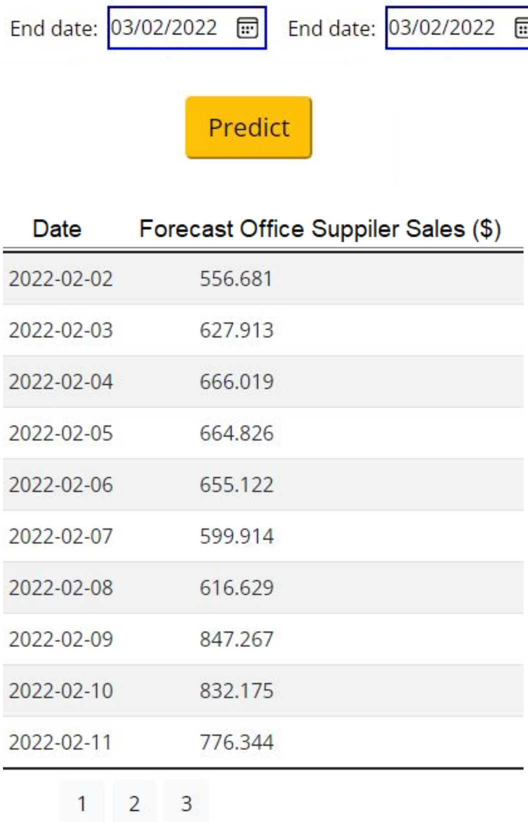


Figure 4: Forecast between the start date and end date



Date	Forecast Office Suppiler Sales ($)
2022-02-02	556.681
2022-02-03	627.913
2022-02-04	666.019
2022-02-05	664.826
2022-02-06	655.122
2022-02-07	599.914
2022-02-08	616.629
2022-02-09	847.267
2022-02-10	832.175
2022-02-11	776.344

1 2 3

Figure 5: Forecast Result in tabular form

Figure 6: Forecast Result in Graphical form.

Figure 7: Forecast Result in Scatter Plot.

Figure 8: Forecast Result in Bar Chart.

VI. CONCLUSION

A web database application is considered as an application that can be retrieved or accessed through the involvement of a web browser. The database and software exist on a basic or central source (server) despite being connected to the desktop screen of the computer system and is connected over a certain network connection. In this study, we have done a development procedure that comprises a simplified system: system design, system analysis, system maintenance, and system implementation. We initially ongoing through analyzing the existing manual system for gathering previous needed data and utilizing them to calculate upcoming demand data. After that, we continued to design of application system by considering three general and hugely utilized models for forecasting with less square technique, general operating (moving) average, and exponentially weighted smoothing average. Afterward, we created and imposed a web-centered demand system for prediction through utilizing the technology model of flask web.

Lastly, we have deliberations that how we documented the whole procedure and created recommendations on how to precisely connect, evaluate, provide the system into manufacturing and then sustain the system under-utilization. Though there were certain few constraints in creating this demand predicting system, this system has been verified at the position of design and it has been located to be efficacious and dependable. This web database application manages a very huge numeral of past sales or historical data. Web-centered applications are the final path to attain benefits of existing technology to explore productivity and efficiency associated with commercials. The web-centered application tends to deliver you a chance to connect your business information from anywhere in the domain of any instance. It also delivers support for saving time and expenses and creating betterments in interactivity with consumers and partners.

VII. RECOMMENDATIONS

For this predicting system for demand to operate efficaciously and endure to accomplish requirements of producing commercials and the general community, to comprise to be precisely connected or installed, processed and sustained. This investigation delivers recommendations that the computer system that would involve in utilization to development of this application framework should include similar general capabilities and functionalities as the regional machine that was utilized in evolving the application.

REFERENCES

[1] Jason Brownlee (July 21, 2016) Deep learning in time series. https://machinelearningmastery.com/time-series-prediction-lstmrecurrent-neural-networks-python-keras

[2] Susan Li (Jul 9, 2018) An End-to-End Project on Time Series Analysis and Forecasting with Python. https://towardsdatascience.com/an-end-to-end-project-on-time-seriesanalysis-and-forecasting-with-python-4835e6bf050b

[3] Mupparaju, Kalyan, Anurag Soni, Prasad Gujela, and Matthew A. Lanham. "A Comparative Study of Machine Learning Frameworks for Demand Forecasting." (2018).

[4] Kilimci, Zeynep Hilal, A. Okay Akyuz, Mitat Uysal, Selim Akyokus, M. Ozan Uysal, Berna Atak Bulbul, and Mehmet Ali Ekmis. "An improved demand forecasting model using deep learning approach and proposed decision integration strategy for the supply chain." Complexity 2019, July 18-20, New Brunswick, New Jersey, USA

[5] Harsoor, Anita S., and Anushree Patil. "Forecast of sales of Walmart store using big data applications." International Journal of Research in Engineering and Technology 4, no. 6 (2015): 51-59.

[6] YU, Jian-hong, and Xiao-Juan LE. "Sales forecast for amazon sales based on different statistics methodologies." DEStech Transactions on Economics, Business and Management iceme-ebm (2016), June 24-26, Guangzhou, China

[7] Babai, Mohamed Ziad, Mohammad Mojiballah Ali, John E. Boylan, and Aris A. Syntetos. "Forecasting and inventory performance in a two-stage supply chain with ARIMA (0, 1, 1) demand: Theory and empirical analysis." International Journal of Production Economics 143, no. 2 (2013): 463-471.