



International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 3, March 2018

Customer Behaviour Analysis in E-Commerce

Ayush Kumar¹, Anand Mohan², Anjali Bhatia³, Pallavi Gadade⁴, Swapnali B. Ware⁵

Department of Information Technology, Sinhgad Institute of Technology, Lonavala, India

ABSTRACT: In today's world, there is enormous amount of data and there is a need to use that data for better business. In Every E-commerce website, customer interaction plays a very crucial role and having a system that would analyse behaviour is very useful. This includes analysing the behaviour in terms of predicting the customer's next move and knowing what they want before they do. With the use of python and applying suitable data science algorithms, such systems can be made and implemented on any e-commerce platform. This is an example of B2C (Business to Customer) relationship. As a result of this, a better understanding of the customers and better experience to them would be provided.

KEYWORDS: Data Mining, Data analysis, Data Forecasting, Time Series, ARIMA

I. INTRODUCTION

In today's world E-Commerce is the first industry to recognize the importance of predictive analytics and started to implement it. In today's competitive business environment understanding what are the customer requirement and offering them the right quality products at right time is the key of any successful business. Due to high growth of internet, online shopping is becoming most interesting and popular activities for the consumers.

Data analytics has quickly emerged as valuable tools for using current reporting capabilities to uncover and understand hidden and complex pattern in vast data base these patterns are then used in models that predict individual behaviour with high accuracy. The result of data mining helps in decision making helps in Customer Relationship Management (CRM) it also affect the cost and production of the business. There are a wide variety of data mining applications available, particularly for business uses, such as Customer Relationship Management (CRM). [4]

The main objective of his paper is to increase customer relationship by giving relevant suggestions and also by using effective framework and data mining and predictive methods, forecast the data to the production line and help the manufacturer produce only the required products based on the market trends.

II. DATA MINING

Data mining is used to discover patterns in larger datasets and making it useful for further work by creating appropriate structure for further use. Data mining uses mathematical algorithms to distribute the data into segments and evaluate probability of future events .Data mining model is created by specific algorithm .Other than raw analysis of datasets it also involves database and data management aspects, model and post processing of discovered structure.

The patterns found from input datasets are further used for machine learning and predictive analytics. For example data mining identifies multiple patterns in dataset, which can be further used for prediction.

The business applications of data mining is increasing as it can be used to predict valuable data ,like customer actions, their buying patterns and industry needs.[4]

III.DATA ANALYTICS

Data analytics is the process of inspecting, cleansing, transforming and modelling data with a motive of discovering useful information, suggesting conclusions and support decision making. The proposed system requires data analytics to be performed well. Data analytics needs data as inputs analysis. Data collections for m various sources is done. Data collected is processed or organised. After cleaning many techniques are applied to analyse the data. Data models are created from above inputs. The outputs of data models are outputs of data analysis.



International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijircce.com

Vol. 6, Issue 3, March 2018

IV. PREDICTIVE ANALYSIS

Predictive analytics deals with extracting information from historical data and using it to predict future trends and behaviour patterns. The core of predictive analytics relies on capturing relationships between explanatory variables and the predicted variables from past occurrences, and exploiting them to predict the unknown outcome.

Predictive analytics works on previous data. It first defines the project outcomes, deliverable, scope of the effort, business objectives, identify the data sets that are going to be used. It captures the data, then reporting or analysis is done which includes what happened and why that happened. After that monitoring of data is done which includes what is happening now. Finally, predictive analysis is performed which means to let the user know what is going to happen in future. Predictive analytic contains these following processes define project, data collection, data analysis, statistics, modelling, deployment, model monitoring.

V. FORECASTING

Forecasting is the process of making predictions of the future based on past and present data and most commonly by analysis of trends. If there are no data available, or if the data available are not relevant to the forecasts, then **qualitative forecasting** methods must be used. These methods are not purely guesswork there are well-developed structured approaches to obtaining good forecasts without using historical data. Quantitative forecasting can be applied when two conditions are satisfied:-numerical information about the past is available and it is reasonable to assume that some aspects of the past patterns will continue into the future. Most quantitative forecasting problems use either time series data (collected at regular intervals over time) or cross-sectional data (collected at a single point in time). Risk and uncertainty are central to forecasting. Several factors on which selection of right prediction technique depend are the context of forecast, the relevance and availability of historical data, the degree of accuracy desirable, the time period to be forecast, the benefit of forecast to company and the time available for analysis.

A. *Time series forecasting*

A time series can be defined as a chronological sequence of observed data from any periodical task or behaviour, or activity in fields like engineering, biology, economy, or social sciences, among many others [2]. Different approaches to time series forecasting are available, which are categorized as follows. Traditional approaches with the use of ARIMA which is linear and another is nonlinear technique .In this forecasting technique ARIMA is widely used for predicting future value.

VI. PROPOSED SYSTEM

The Proposed system is supposed to work on any e-Commerce website and it can be integrated with its database and generate models and perform prediction of user behavior and forecast the data for the production line so that only the required instances of certain products be produced at current stage which might be in more demand in the upcoming period of time. This can lead to business development and a quality relationship between the customer and the company. The forecasted data can increase the revenue of the company and make a lead on the global market and lead among the peers.

International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijirccce.com

Vol. 6, Issue 3, March 2018

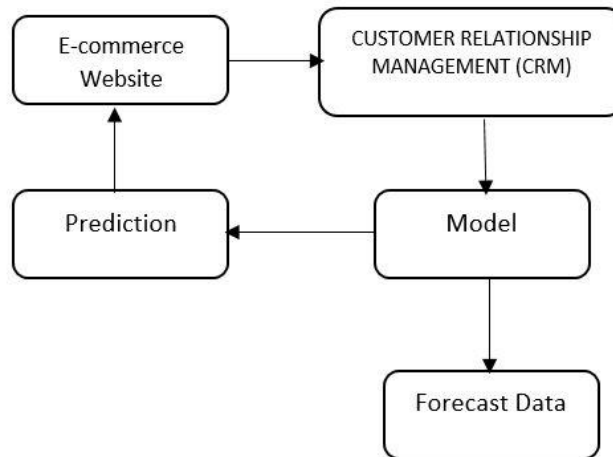


Fig 1. Generalized flow of the system

As stated above in the figure, it depicts the flow of architecture of the system. At first, there is the e-commerce website where the users will interact by logging in, registration, viewing products, short listing, buying, etc. The users will get suggestions based on their recent activities such as when they search for some product or shortlist them or purchase them. The idea is also to implement a system where one can get suggestions like “Frequently bought together”. For example, If someone is about to place an order of a phone, then they could get suggestions of buying a phone cover along with the phone. With the help of training dataset and the transaction table, a model can be prepared and after training it on suitable amount of data, the model will be ready to make predictions as product suggestions for end-user and forecast data for the company as popular products in demand right now.

The model can be prepared with the help of suitable classification algorithm, regression algorithm and time-series algorithm.

VII. EXPERIMENT AND RESULTS

We gathered a dataset of sales of a product over the course of three years. So, now the objective of the experiment was to apply ARIMA rolling model forecasting and obtain a forecast of the sales of that product. ARIMA is an acronym that stands for Autoregressive Integrated Moving Average. It is a generalization of the simpler Autoregressive Moving Average (ARMA) and adds the notion of integration. The simple plot of the dataset showed the following graph:

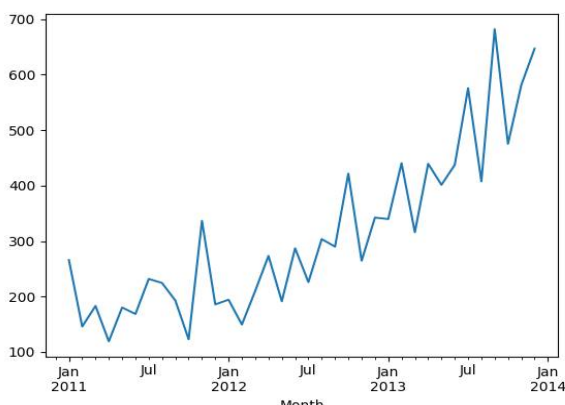


Fig 2. Plot of Sale Dataset using Graph

predicted=290.313926,	expected=264.500000
predicted=349.117718,	expected=342.300000
predicted=306.512988,	expected=339.700000
predicted=387.376451,	expected=440.400000
predicted=348.154272,	expected=315.900000
predicted=386.308755,	expected=439.300000
predicted=356.082067,	expected=401.300000
predicted=446.379447,	expected=437.400000
predicted=394.737360,	expected=575.500000
predicted=434.915543,	expected=407.600000
predicted=507.923470,	expected=682.000000
predicted=435.482930,	expected=475.300000
predicted=652.743730,	expected=581.300000
predicted=546.343483,	expected=646.900000
Test MSE: 6508.896	
Test RMSE 80.678	

Fig 3. Predicted vs. Expected Result



International Journal of Innovative Research in Computer and Communication Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijirce.com

Vol. 6, Issue 3, March 2018

After applying time series ARIMA (p,d,q), we obtain the result as the predicted and expected values of the number of units sold. The Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) is obtained.

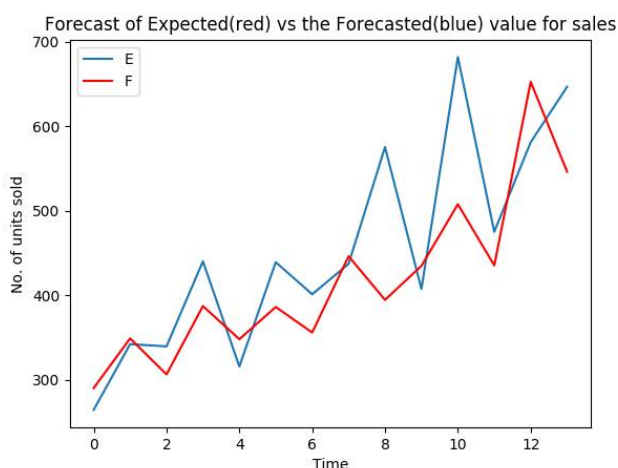


Fig 4. Forecasted Result

The process is repeated until either a desirable level of fit is achieved on the in-sample or out-of-sample observations. Now, with further tuning of this model, a better forecast can be obtained.

VIII.CONCLUSION

Through this paper retailers are able to gather the information of customer purchases in advance so that they can analyse those purchase details and discover for important purchase patterns and user behaviour, where the retailer can use the results to improve their sales.

It also focuses on Time series algorithm to give better results in sales by evaluating expected and predicted value and focus towards the predictive analytics, data mining techniques and forecasting. Predictive analytics are more efficient in choosing marketing methods and helpful in social media analytics.

REFERENCES

1. kavya.v1, arumugam.s2, "a review on predictive analytics in data mining", international journal of chaos, control, modelling and simulation (ijccms) vol.5, no.1/2/3, september 2016
2. sohelia mehrmolaei, mohammad reza keyvanpour, "time series forecasting using improved arima", ARTIFICIAL intelligence and robotics(iranopen), april 2016
3. nethmi deshani hettiarachchi, sobhani umanga pilapitiya, nirmal sankalpa jayasinghe, himash deemantha, sudheera vitharana" i shopping: intelligent shopping and predicate analysis system using data mining", iee international conference on data science and data intensive systems-2015
4. aditya kumar gupta, chakit gupta."Analyzing customer behavior using data mining techniques: optimising relationships with customer",indian institute of professional studies-june 2010.
5. huang lan, zhou chun-guang , zhou yu-qin , wang zhe," research on data mining algorithms for automotive customers' behavior prediction problem", seventh international conference on machine learning and applications-2008.