



Sales Prediction in Tourism Industry Using Data Mining

Shivam Kumar Shukla¹, Harsh Patel², Milan Kanzariya³, Ashwini Deshmukh⁴, Akshata Patil⁵.

U.G. Student, Department of Information Technology, Shah and Anchor Kutchhi Engineering College, Mumbai, Maharashtra, India^{1,2,3}

Associate Professor, Department of Information Technology, Shah and Anchor Kutchhi Engineering College, Mumbai, Maharashtra, India^{4,5}

ABSTRACT: The present forceful circumstance for Indian travel and the tourism industry is the need to raise their market and keep control their organizations to utilize data mining devices and methods to create, advertise the travel industry items and administrations. Despite the fact that there are many anticipating models for deciding deals in the tourism industry, data mining systems have been viewed as the best method for determining deals in the travel industry. Data mining is characterized as the way toward discovering valuable examples, connections, and rules, which are not known beforehand, by sifting through a lot of information put away in some repositories (database). Deals designs from sales information demonstrate advertise slants and can be utilized in anticipating which has incredible potential for dynamic, vital arranging and market rivalry. In this project we use multiple linear regression (MLR) technique to predict sales from historical data and get profit in tourism industry. So, this project aims that how to increase sales from tourism industry using data mining.

KEYWORDS: Prediction, Multiple linear regression, Travel and Tourism industry, data mining.

I. INTRODUCTION

In recent few year tourism industry become very important and main incoming source for India. The travel industry in India is significant for the nation's economy and is developing quickly. The World Travel and Tourism Council determined that travel industry produced ₹16.91 lakh crore (US\$240 billion) or 9.2% of India's GDP in 2018 and bolstered 42.673 million occupations, 8.1% of its absolute work. The area is anticipated to develop at a yearly pace of 6.9% to ₹32.05 lakh crore (US\$450 billion) by 2028 (9.9% of GDP)[2]. Sales prediction is the way toward assessing the amount of an item or administration that shoppers will buy. Request estimating fundamentally includes strategies including both casual techniques as conjectures, and quantitative strategies, for example, the utilization of recorded deals information or current information from test markets. Sales prediction might be utilized in settling on evaluating choices, in surveying future limit necessities, or in settling on choices on whether to enter another market. The estimate of the travel industry of uncommon significance since it is a marker of future interest, along these lines giving fundamental data to ensuing arranging and strategy making. The main prediction methods which are used in the tourism and recreation fields are time-series methods, multiple regression methods, multivariate methods and qualitative forecasting methods.[4]

II. LITERATURE WORK

In earlier studies, there are various techniques and methods can be used to predict sales and get accurate result. Some of them are can be referred by us during production of our project which can be explain in following section. These methods are used to give the accurate sales prediction model. These methods are going to give different result on different data set and in different situation. As technology emerging very rapidly, different methods and different datamining algorithms are tested. Due to increase in competition in tourism industry and growth of tourism in India many researcher are take in interest in these topic.

Aditya Joshi, Nidhi Pandey, (Professor) Rashmi Chawla, Pratik Patil are proposed a paper in which they used clustering association mining methods on stock data to classify and find association pattern on sale. From their experimental result these method is very efficient for large stock data and predicting factor of sale. But their method is very simple. Limitation of study is, it requires proper formatted data with specific attribute.[1]

Shini Renjith, A. Sreekumar, M. Jathavedan proposed a paper Evaluation of Partitioning Clustering Algorithms for Processing Social Media Data in Tourism Domain in they used different data mining algorithms to analyse and classify accuracy of that on predicting sales on destination reviews across South India published on holidayiq.com data set. they have used k-means, k-mediod, CLARA algorithms for predicting sales of given data set.[7]



Girish Kumar Sharma , Promila Sharma proposed a paper A Study of Data Mining Algorithm for Tourism Industry in which they used different data mining algorithm such as decision tree ,classification, clustering, association rule to access a pattern in tourism industry. This paper has overall goal of use different data mining algorithm for tourism industry.[2]

III. SCOPE

The main scope of our project is to increase sales of tourism using different data mining techniques according to its prediction upon historical or collected data set from different tourism website. Using our project tourism administrator can increase their sales. Because, using collected data and prediction on it gives idea about which packages are going to put to customer in different season and different location.

As competition in tourism industry increases customer satisfaction is very important to get increase in sales of tourism.Sales Prediction in tourism has a huge scope for the generations to come.

IV. PROPOSED WORK

In this chapter we have give the brief description of proposed work and examine different module of our system along with different models through which this system is understood and represented. Some of the techniques of data mining are also examine and study. Accordingly we implement our system.

Block Diagram

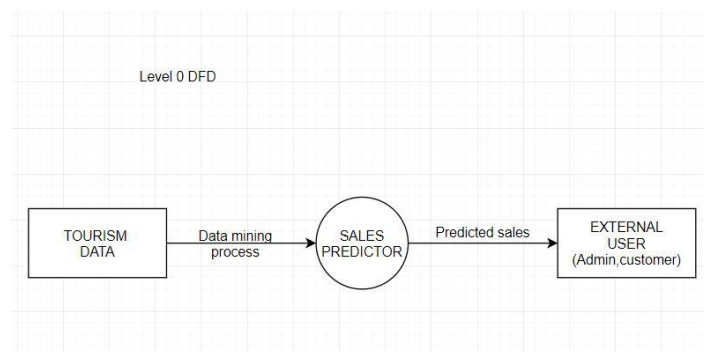


Fig 1

level 0 also known as Context level DFD, gives the overview of whole system in which the external entities like the tourism data and data mining techniques are process, which are processed for predicting the sales.

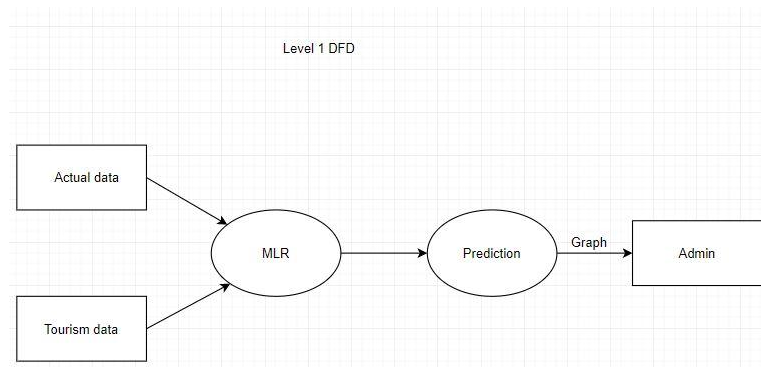


Fig 2

Represents collection of data, data mining technique, prediction visual. It is elaboration of level 0 DFD. Pre-processing method is done in traditional way .After the-processing the dataset is applied to the data mining algorithm MLR.



Architecture

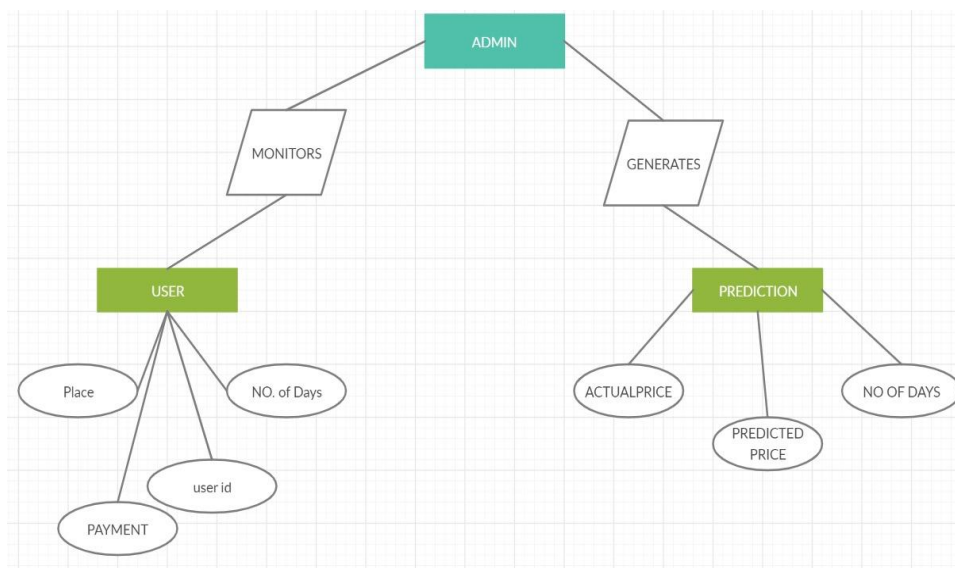


Fig 3

PREDICTION OR ADMIN MODULE

This module is our main focus of project where we have predicting module as well. According to our database and historical data set of the sales, system will predict sales for that season of tourism .In recent year tourism become very important aspect in India and competition is very high .So ,to attract customer and give good service becomes essential part. This prediction will help in letting the admin know that which package should remove and add in which period of time or season.

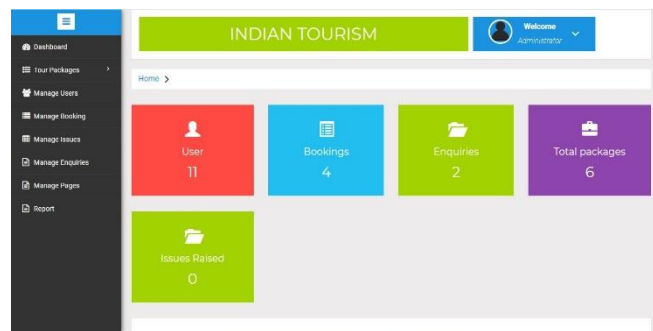


Fig 4 Admin module

USER MODULE

User module has simple system as every tourism website. This module contain information about tour packages and cost of them. In that module user can enquire about different tour packages and book for that packages with secure payment system. User also cancel their booking for any reason in few hours. Customer also give feedback about its experience and also give rating for that to give suggestion to other customer.



Fig 5 Signup Page

Fig 6 Login Page

Fig 7 Package List

Fig 8 Booking Page

Fig 9 Payment Page

DATA SET

In order to get data set we create data set up to 200 which contain different attribute for prediction using MLR algorithm. Such attribute are: Place, Customer name , Rating , Cost. Rating can be given in five different way such as:

1. Poor
2. Less poor
3. Moderate
4. Good
5. Excellent



Figure 4.3 Admin module

As shown in Fig. 4.4 according to prediction module graph are display to add or remove packages by admin to give correct choice to customer and grow their sales during that season.



Figure 4.4 represent graph for sales prediction

VI. CONCLUSION

In this project, we studied different data mining algorithms for sales prediction in tourism industry. According to our study we select best algorithm that is Multiple Linear Regression which can predict sales accurately than other algorithm. Multiple Linear Regression technique is suitable for our project because it is very efficient and accurate for large data set as for our project. On the bases of our assumption we create efficient and accurate system for predicting sales and according to that add and remove packages to increase their sales for in future. As our main project focus is for tourism industry admin that how they can increase their sales or growth in these competitive environment. It also has some limitation in this project that we cannot used sentiment analysis through which user verbal feedback is not analyse. Finally, according to our paper we successfully created predictive module from Multiple Linear Regression that predict any sales in tourism to increase sales.

ACKNOWLEDGMENT

The success and final outcome of this project required a lot of guidance and assistance from many people and we are extremely privileged to have got this all along the completion of our project. We respect and thank **Ms. Akshata Patil** and **Ms. Ashwini Deshmukh**, for providing us an opportunity to do the project work in data mining and giving us all the support and guidance which made us complete the project duly. We are extremely thankful to her for providing such good support and guidance. We owe our deep gratitude to her for taking keen interest in our project work and guiding us all along, till the completion of our project work by providing all the necessary information for developing a good system. We thank her for her encouragement and more over for her timely support and guidance which helped us in successfully completing our project work. We would also like to thank our principal **Dr. Bhavesh Patel** and **Ms. Swati Nadkarni**, I/c Head of Information Technology Department for providing us the opportunity to implement our project. We are really thankful to them. Finally we would also like to thank our parents and friends who helped us a lot in finalizing this project within the limited timeframe.



REFERENCES

- [1] Aditya Joshi, Nidhi Pandey, (Professor) Rashmi Chawla, Pratik Patil. " Use of Data Mining Techniques to Improve the Effectiveness of Sales and Marketing" IJCSMC, Vol. 4, Issue. 4, April 2015, pg.81 – 87.
- [2] Girish Kumar Sharma , Promila Sharma. "A Study of Data Mining Algorithm for Tourism Industry" International Journal of Latest Trends in Engineering and Technology (IJLTET), Vol 7 issue 1 May 2016.
- [3] Manoj B. Kathariya , Ravinder Singh Sakshi , Diler Singh Sakshi , Dhaval R. Kathariya. "Data Mining for Travels and Tourism", Journal of Information and Operations Management, , Volume 3, Issue 1, 2012, pp-114-118, 31 May 2014.
- [4] Shini Renjith1 , Anjali C2. " A Personalized Travel Recommender Model Based on Content-based Prediction and Collaborative Recommendation" , International Journal of Computer Science and Mobile Computing, ICMIC13, December- 2013, pg. 66-73
- [5] Mirjana Pejic Bach , Markus Schatten , Zrinka Marušić. "Data Mining Applications in Tourism: A Keyword Analysis" Central European Conference on Information and Intelligent Systems , September 18-20, 2013.
- [6] Shini Renjith, Anjali C. "A Personalized Mobile Travel Recommender System using Hybrid Algorithm" 2014 First International Conference on Computational Systems and Communication (ICCSC), 17-18 December 2014.
- [7] Shini Renjith , A. Sreekumar , M. Jathavedan. "Evaluation of Partitioning Clustering Algorithms for Processing Social Media Data in Tourism Domain" 2018 IEEE Recent Advances in Intelligent Computational Systems (RAICS), December 06 – 08, 2018