

Trippler Tourism Management System Using Data Mining

Fonseka P.D.K.P, Kaluarachchi S.J, Jayaneththi D.D.M, Thennakoon D.B.S.P. N Vithana

Sri Lanka Institute Of Information Technology

Abstract- Annually a large number of tourists are arriving to Sri Lanka from different places in the world, to visit so many places in Sri Lanka. Therefore the tourism industry is becoming a great portion of national economy and it opens vast number of occupations for the Sri Lankan community throughout the various tourism companies. So the tourism companies should be more effective and enrich with enough facilities to keep all the customers which coming for them and serve the customers as they expect. But still the tourism industry couldn't overcome this challenge due to the many tourism companies have not a real understanding about who are the most frequent customers? From where, when and how they come, what are their preferences and expectations? If the tourism companies could gather this information, they could be manage and utilize their services and resources in a very efficient and effective way in order to serve all the customers perfectly. So this research is focusing to developing a web based solution to find out most frequent and potential customers who arrive to the tourism companies and important details about them. The data mining techniques will be applied to find out this information using the data which collected from tourism companies. And discovered knowledge will be presented in graphs and send this information Tourism Company management. So this information is more valuable to change their businesses into effective ways.

Index Terms- Sri Lankan Tourism industry; Web application; tourism business strategy; data mining; graphical data representations; find frequent customers; proper business navigation; Time series algorithm

I. INTRODUCTION

Data mining is commonly termed for a set of techniques that help to analyze and recognize significant facts, relationships, trends, patterns, exceptions and anomalies that might otherwise go unnoticed that exists in large amounts of data. Although Data Mining (Data Mining) is still in its infancy, many developed countries have successfully applied Data Mining tools and techniques in different domains and have gaining useful knowledge from historical data. Data Mining incorporates a multitude of techniques from a variety of fields including databases, statistics, data visualization, machine learning and others. Large amounts of data are collected, stored and analyzed to obtain knowledge which is sometimes crucial in decision making. Data mining have been successfully applied in various domains such as Business, Finance, Biology, Medicine, Text Mining, Web Mining, Marketing and Financial Forecasting [1].

Sri Lanka is one of the most popular tourist destinations in the region. But the tourists which coming for hotels and other tourism companies, are not properly served and thrilled customers as they expected. Incapability to manage resources and services in effective and efficiently way is the main reason for this. Sometimes the resources will be inefficient for tourists such as rooms, foods etc. Servants in the hotel may be not enough to serve all the tourist properly. So those are reasoning to reduce the quality of services and it will be highly affect for the business profit. Unexpected customer arrivals is root for this problem. If those companies could find about most frequent and potential customers and important details about them (when and how they come, their economy level, their interesting places, rooms basis, taxies and room packages), the business could be customize into an attractive and interesting way for those tourists. This research is focusing to discover the most frequent and potential customers and important details about them. Those data will be graphically represented and alert relevant details hotel wise using the web application.

In order to conduct a Data Mining experiment, the project team recommend considering the following questions to provide a direction to the experiment with less guesswork.

Identify what types of knowledge useful for tourism industry

This was a very initial step of this research. This question means what are the most important and valuable information in the tourism companies as well as tourist in the Sri Lanka. What were the most important knowledge which useful for making the desersions in tourism companies.

Identify what data to select for the experiment

The discovered result should be more accurate, because the business plans and management decisions are highly affected by those results. In order to increase the accuracy of discovered knowledge, the data set should be large. Large number of tourists are arriving to tourism agencies. So these data in past years (at least one or more) are most relent to this project.

Find How to predict most important tends which useful for tourism industry

After determined what are predicted and collected data to predict this question arise. The knowledge which discovered should be accurate and effective for the tourism industry in the Sri Lanka. So a proper methodology was needed for this project.

II. BACKGROUND

There are many researches done in the field of data mining.

An article which reviewed is discussing the advantages of using network data in forecasting demand for China travel. Mapping information organization and network of relationships to build demand in the tourism sector between the author, and describes the use of methods of data mining Web, analyzes and forecasting of tourism demand, Based on the principle of building knowledge and explanations web data mining process Analysis and forecasting. In the case of the use of Shanghai as a city to do an empirical analysis Points. [2]

The aim of this study is to find a time, travelers are looking for guests from the travel search behavior. This study focuses on the use of mobile devices as a destination recognized travel options, such as visitor accommodation, tourist attractions, travel services, restaurant and gift shop tools. They use data mining and association rules technology approach to analyze the relationship between information and transactions between the travelers. Knowledge found in the database can be used as a set of rules that provides flight information via mobile travelers. Our system is designed as a complementary educational knowledge. The results show that: the use of data from the intelligent analysis of the tourism sector can increase the chances of a competitive tourism company to effectively respond to the needs of travelers. [3]

Japan has become an important tourist destination. Arrivals of tourists has increased significantly in the last decade. According to the 2000 Japan accounted for passengers to travel JNTO is 4,757,146, a total of 2012 the number of Japanese visitors to 8,358,105, and in 2013 and 24% in the same period last year, As a result of our analysis, They found that the satisfaction of the two different groups, the most important factor is the price of the ticket, and Japanese cuisine, shopping, how many visits Japan ions past experience. In addition, satisfaction and non-Asian groups in Asia, two different groups, different preferences and purpose of the visit. The behavioral intentions favorite experience some aspects, such as the future of Japan's hot springs are used to return the incentives. The tourists mainly travel with your family and the ticket price is an important factor. [4]

Every day, millions of people around the world travel business, Vacation, travel, or other reasons. Astronomical Money spent on tickets, accommodation, food intake, Transportation and entertainment. According to the World TourismBureau, tourism is about Global gross domestic product (GDP), 11 percent (of GDP) (Werthner and Ritchie, 2004). Tourism is based on information Service, wherein there are two types of traffic. One the flow of information from the provider to the consumer or tourists. This is about the goods Tourist information Consumption, such as tickets, hotel rooms, entertainment and so on. Go ahead. Another flow of information, which should be the opposite Line consists of a summary of relevant visitors Service providers. In this chapter, they will discuss the second Tourist information in the form of flow behavior. When gathering data about your visitors presented in the right direction, by using the correct algorithm, and provide analysis in the right hand, it can be translated into Italian Tourist information services produced important decisions Vendors increase revenue and profits. Data

mining can Analysis of data related to travel a very useful tool. [5]

In recent years, tourism has become one of the fastest growing sectors of the global economy, it is idly regional and national economic development was recognized contribution. Travel Product Design and development of the region's growing number of foreign sources / countries have become important activities and household income. On the other hand, it is a strong contender for customer relationship management Companies need to integrate their customers and need to stay focused on strategy the organization customer-centric approach. Therefore, They use Appropriate algorithm One for data mining association rules and clustering analysis method, which is due to the implementation of Case Firm, Phoenix International Travel customer knowledge to dig in Taiwan. Know how Knowledge extraction from data mining results mode, rules and knowledge described in the map In order to make the new product development and customer case for the company to provide advice and solutions Relationship management. [6]

High competitive pressure in the global manufacturing industry makes efficient, effective and continuously improved manufacturing processes a critical success factor. And Indication-based Manufacturing Optimization as a novel data mining-driven approach for process optimization provided by the Advanced Manufacturing Analytics Platform. And this research paper defines conceptual use cases and described implementation details, Indication-based Manufacturing Optimization goes beyond existing analytics in manufacturing, which focus on manual reporting and OLAP functions using isolated data extracts and it discussed pre-defined data mining use cases are applied to identify hidden data patterns for the optimization of the whole manufacturing process, from the creation of the production order until the finishing of the product [7]

Another article reviews the application's data Mining Processing travel papers Industry, whether it is from both demand and supply of the web. Especially important in the field of tourism science, SCOPUS Published in the journal in 1995 – 2013 Search and use proper keywords. Literature search and found 88 papers presentation Tourism data mining applications. Keywords and The concept of network analysis conducted Sixth device using Wordle and Lanet. Paper from Tourism-related magazines and journals were related to ICT Separate analysis. In order to explore the historical trend, These two periods were analyzed Conclusions 2005 and since 2006, Paper is a step in the way of the development of tourism, as Two people driven and data-driven, and therefore need to use As a method to improve the use of data mining Competitiveness and profitability. [8]

An article detailed Indication-based Manufacturing Optimization as a novel data mining-driven approach for process optimization provided by the Advanced Manufacturing Analytics Platform. Defined conceptual use cases and described implementation details. [9]

The System, which use to have a more efficient plan for future year by analyzing past data. The proposed system is uses Data Mining technology to analyze the data. In the literature review, the team has discussed and searched about the existing deployed systems in the hotel management systems around the world and related solutions and researches have been done and developed by using data mining also the literature review has

covered about the data mining technology and how to apply them and how has applied.

III. METHODOLOGY

Data mining denotes a rather difficult and specific field. A complicated and fixed approach is necessary for the use of data mining in order to help organizations use the data mining. For this project, Prototypes are expected and there will be very critical risk assessment when going through the project because the team does not have much experience within the domain. The spiral model emphasizes risk analysis. Remove all potential risks through careful analysis and, if necessary, by constructing a prototype before data is being mined several steps have to follow

Data integration: First collected and integrated all the data from different sources

Data selection: Then selected the data that can be used for data mining

Data cleaning: Most of the time, data gathered are not clean and may contain errors, missing values, noisy or inconsistent data. So, they have to remove. This process was done manually

Data transform: This step is done for get more efficient results and to gain results in understandable manner. Mainly normalization was done for the data

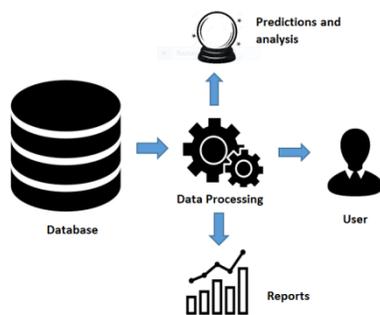
Data mining: In this step the algorithms were used for mine the database. The prediction was done by using Time Serious algorithm

Pattern Evaluation and Knowledge Presentation: First some of the results that are not related were removed fine tune other results well to understand by the users of the system.

Using Time Serious algorithm can forecast most tourist attracted hotels, destinations, room types and board basis

Which useful for management decisions in tourism companies. The predicted results were representing in web application in bar charts. Any user can view those predicted results using this web application. The results are getting through a mining model in Microsoft Business Intelligence using DMX queries.

Figure 1: High Level



PREDICTIONS AND ANALYSIS

IV. DESIGN OF THE SYSTEM

This chapter describes the expected results and actual outcome for each test case of testing each module. Some of the expected results of our system shown below one by one descriptively.



Figure 2.1: Home Page

This is the home page. This is the startup page of the web site and from this page; user can be navigated to the appropriate page which contains the prediction reports (hotel prediction, destination prediction, room class prediction, board basis prediction). Further navigated user can be to the pages which can view more details about the web application (using about us button) and contact details (using telephone receiver symbol).



Figure 2.2: Hotel Prediction

This page displays the report which contains details of hotel predictions. The bar chart at the left side of the page is demonstrating the top 10 hotels against the population in next twelve months. The x-axis represents the hotel name and the y-axis represents the population (number of tourist which can be come in to the hotel). And all the hotels will be listed according to the ranks. The rank will be considered according to the population.



Figure 2.3: Destination Prediction

This page displays the report which contains the destination prediction details. The bar chart at the left side of the page is demonstrating the top 10 destinations against the population in next twelve months. The x-axis of the bar chart is representing the destination name and the y-axis representing the population (number of tourist which can be come in to the destination). And all the destinations will be listed according to the ranks at the right side of the page. The destination rank will be considered according to the population.



Figure 2.4: Room Class Prediction

This page displays the report which contains the room class prediction details. The bar chart at the left side of the page is demonstrating the top 10 room classes against the population. The x-axis of the bar chart is representing the room class and the y-axis representing the population (number of tourist which can be using the room class). And all the room classes will be listed according to the ranks at the right side of the page. The room class rank will be considered according to the population.



Figure 2.5: Board Basis Prediction

This page displays the report which contains the board basis prediction details. The bar chart at the left side of the page is demonstrating the top 10 board bases against the population. The x-axis of the bar chart is representing the board basis and the y-axis representing the population (number of tourist which can be use the board basis). And all the board bases will be listed according to the ranks at the right side of the page. The board basis rank will be considered according to the population.

V. DISCUSSION

The “Tripper Web Application” gives the accurate and important knowledge for the tourism companies in Sri Lanka. Final product is controlled by web application. Project team used latest technologies to the system when developing the project. This system predicts most tourist attracted places, hotels, room types and board basis for the next five months.

When developing this project team members felt with some difficulties. First one is clear the data. Data was collected from verity of sources and team had to form these data to one format which acceptable for the prediction algorithm. Another problem was the software incompatibility. The data mining tool and the web server was not connecting properly. So team members had to deep study about MSSQL and xamppweb server connection to overcome this problem. The Tripper Web Application is demonstrate the mostly tourist attracted hotels, destinations, room types and board basis in bar charts using predicted results in the mining models. Further user can see the results in next five months. In every bar charts the y-axis is demonstrate the tourist’s population. So the population can be conceded as the tourist attraction.

The bar chart is not demonstrating all the predicted results. It demonstrates only top ten results. So the ranking list shows the all predicted results in a descending order according to the tourist’s population. So user can check the tourist attraction level of a given factor (destination, hotel, room type, board basis) When user wants to see the results in a particular month, there is a dropdown list to select the month. When the month was selected the update the bar chart according to the new results of selected month. For this task the team had to implement this web application using Ajax technologies.

VI. CONCLUSION

Tripper Web Application is a very useful and effective application in tourism industry in Sri Lanka. The predicted results by this Tripper Web Application are very important and valuable for the tourism companies as well as tourist which hopes to travel in Sri Lanka.

When the management is taking decisions in tourism companies, the predicted results of the Tripper Web Application will be very useful. So this web app can be considered as a decision support system also. When a tourist wants to travel in to Sri Lanka he/she can the most attracted places, hotels, room types and board basis in upcoming months. So those results will be very important for panning the trips to the Sri Lanka for the tourists also.

When a student wanted to get an idea about currently most tourists attracted thing in the Sri Lanka, this web application will be very useful. The government also can be use this web application to make their future planes to develop the tourism industry in the Sri Lanka. So when consider above thing the Tripper Web Application is a very important and effective system in the tourism industry in the Sri Lanka

As a main limitation in Tripper Web Application is the Tripper gives predictions only for Sri Lanka. So this application is valid for only the Sri Lankan Tourism industry. The another limitation is this application gives the predictions about next five months only. User can't get results more than five months. User can get predations for monthly wise and user cannot get predictions for a particular date. So this also a limitation of this web application.

The user can predicts only most popular places, hotels, room type and board basis only. User cannot get predictions about other factors in the tourism industry such as most popular transport method, food, drinks etc. Tripper Web Application is update mining models monthly wise only. So that also can be consider as a limitation of this system.

VII. FURTHER WORK

The Tripper Web Application is currently predicts only the several tends in the tourism industry (mostly tourists attracted hotels, places, room type and board basis). But this application can be expand to predict more thing in tourism industry such as mostly tourists attracted fool items, transport methods, drinks etc. Currently this application get prediction for the next five months only. So this web application can be more developed to predict results more than next five months. Currently Tripper gives predictions for months only, so it also can be expanded to get prediction for an upcoming particular date.

REFERENCES

- [1] "World Health Organization", World Health Organization, 2016. [Online]. Available: <http://www.who.int/en/>. [Accessed: 05- May- 2016].

- [2] "Mobile phone users worldwide 2013-2019 | Statistic", Statista, 2016. [Online]. Available: <http://www.statista.com/statistics/274774/forecast-of-mobile-phone-users-worldwide/>. [Accessed: 05- May- 2016].
- [3] "Mobile phone users worldwide 2013-2019 | Statistic", Statista, 2016. [Online]. Available: <http://www.statista.com/statistics/274774/forecast-of-mobile-phone-users-worldwide/>. [Accessed: 05- May- 2016].
- [4] Delgado, Armando Roy; Picking, Rich and Grout, Vic (2002) "Remote-Controlled Automation
- [5] Systems with Different Network Technology " CAIR,2002 [Online] .Available: <http://www.glyndwr.ac.uk/groutv/papers/p5.pdf> /. [Accessed: 05- May- 2016].
- [6] Dhawan S. Thakur and Aditi Sharma (2013)"Voice recognition wireless Home Automation
- [7] System Based On Zigbee",IOSR-JECE,2013. [Online] .Available: <http://aircse.org/journal/jwmn/6114jwmn04.pdf> /. [Accessed: 10- May- 2016].
- [8] NSriskanthan,FTan,AKarande(2002)"Bluetoothbasedhomeautomationsystem",IJSER,2002.[Online] .Available:<https://pdfs.semanticscholar.org/a001/5665acb00ddf1491aec436e47cfb75835b82.pdf>/. [Accessed:09-April-2016].
- [9] LS Humphries, G Rasmussen, DL Voita (2003) "HomeAutomationSystem",IJSR,2013.[Online] Available:<http://research.ijcaonline.org/volume116/number11/pxc3902601.pdf> /. [Accessed: 11May-2016].
- [10] Luca Piccini, Sergio Parini, Luca Maggi and Giuseppe Andreoni (2005) "Home Automation System For Disabled People using BCI " ,IOSR-JCE,2016. [Online]. Available: <http://iosrjournals.org/iosr-jce/papers/ICAET-2014/volume-3/16.pdf?id=7557/>. [Accessed: 07- May- 2016].

AUTHORS

- First Author** – Fonseka P.D.K.P, Sri Lanka Institute Of Information Technology
- Second Author** – Kaluarachchi S.J, Sri Lanka Institute Of Information Technology
- Third Author** – Jayaneththi D.D.M, Sri Lanka Institute Of Information Technology
- Fourth Author** – Thennakoon D.B.S.P, Sri Lanka Institute Of Information Technology
- Fifth Author** – N Vithana, Sri Lanka Institute Of Information Technology