IJHRMOB

# International Journal of HRM and Organizational Behavior

www.ijhrmob.com

editor@ijhrmob.com

International Journal of
HRM and Organizational Behavior
ISSN 2454 - 5015

# A DECISION TREE BASED RECOMMENDATION SYSTEM FOR TOURISTS

DOGUPARTHI LALITHA RANI*1, M. KRANTHI*2, THELLA MAMATHA*3, GAMINI PAVAN KUMAR*4, KOTHAPALLI VISHNU VARDHAN*5

* 1,3,4,5 B. Tech Students, *2 Assistant Professor
Dept. of Computer Science and Engineering,
RISE Krishna Sai Prakasam Group of Institutions

## ABSTRACT

Choosing a tourist destination from the information that is available on the Internet and through other sources is one of the most complex tasks for tourists when planning travel, both before and during travel. Previous Travel Recommendation Systems (TRSs) have attempted to solve this problem. However, some of the technical aspects such as system accuracy and the practical aspects such as usability and satisfaction have been neglected. To address this issue, it requires a full understanding of the tourists' decision-making and novel models for their information search process. This paper proposes a novel human-centric TRS that recommends destinations to tourists in an unfamiliar city. It considers both technical and practical aspects using a real world data set we collected. The system is developed using a two-steps feature selection method to reduce number of inputs to the system and recommendations are provided by decision tree C4.5. The experimental results show that the proposed TRS can provide personalized recommendation on tourist destinations that satisfy the tourists.

## I. INTRODUCTION

### 1.1 Problem statement :

Choosing a tourist destination from the information that is available on the Internet and through other sources is one of the most complex tasks for tourists when planning travel, both before and during travel. Previous Travel Recommendation Systems (TRSs) have attempted to solve this problem. However, some of the technical aspects such as system accuracy and the practical aspects such as usability and satisfaction have been neglected..

### 1.2 MOTIVATION:

To address this issue, it requires a full understanding of the tourists' decision-making and novel models for their information search process. This paper proposes a novel human-centric TRS that recommends destinations to tourists in an unfamiliar city. It considers both technical and practical aspects using a real world

ISSN 2454 - 5015

data set we collected. The system is developed using a two-steps feature selection method to reduce number of inputs to the system and recommendations are provided by decision tree C4.5. The experimental results show that the proposed TRS can provide personalized recommendation on tourist destinations that satisfy the tourists.

### 1.3 Objective:

a tourist destination from the information that is available on the Internet and through other sources is one of the most complex tasks for tourists when planning travel, both before and during travel. Previous Travel Recommendation Systems (TRSs) have attempted to solve this problem. However, some of the technical aspects such as system accuracy and the practical aspects such as usability and satisfaction have been neglected. To address this issue, it requires a full understanding of the tourists' decision-making and novel models for their information search process.

#### 1.3.1 Proposed System:

The proposed DM framework consists of four phases including data acquisition, data pre- processing, data analysis, and result interpretation. (1) For data acquisition, the designed questionnaire, which has four parts, is distributed and collected from Chiang Mai, Thailand. (2) The collected data is pre-processed using several data pre-processing techniques involving data cleaning, data transformation, and feature selection methods. (3) The third phase involves the data analysis processes using a decision tree C4.5 as classifier. The aim of the third phase is to identify suitable features and find personalized systems have not been a focus of RS research.

To overcome from above problem author is asking to use C4.5 decision tree algorithms which take experiences of previous users and then build a model and if new user enter his requirements then decision tree will predict best location based on his given input. Decision tree don't need new users past experience data.

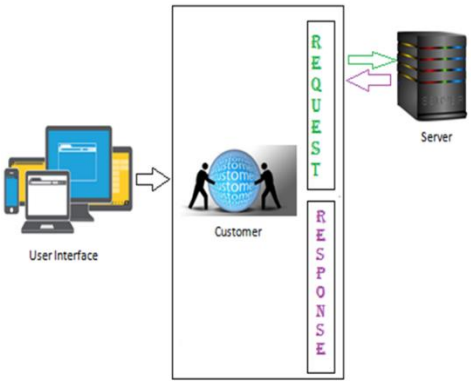
To implement decision tree model, we need to have dataset and this dataset sometime will have empty or garbage values and this values will put bad effect on decision tree model so we can remove such empty or garbage values by applying pre-process techniques.

Sometime to predict or build model no need to use all columns (attributes) values from dataset and these unnecessary attributes can be remove by apply features selection algorithms and here we are using MRMR features selection algorithms to remove unnecessary attributes to reduce

ISSN 2454 - 5015

execution time of building model and to increase system accuracy.

## II. System Architecture




## III. Module description

Our application consists of three modules

1. Customer module

**Customer**

This module describes all about customers, by using this module any customer can perform operations like the upload dataset preprocess & MRMR Feature Selection Generate C4.5 Decision Tree Model Tourist Recommendation features Selection Graph.

### A. Data acquisition

To understand tourist's search behaviour in assessing travel information and decision-making processing for destination choice, we use a questionnaire as a data collection method due to its effective mechanism for collecting information from tourists. Pre-study on variety of factors that influence tourist's preferred destinations were identified for questionnaire design. The questionnaire design contains four parts containing a set of factors related to tourist's preferred destinations as following:

1) Trip characteristics: These variables are the most important variables when tourists select their destinations. This includes trip length, travel purpose, trip composition, and etc.

2) Tourist characteristics: These variables include pychological, congitive and socioeconomic status varaibles that influence on the tourist destination choice process.

3) Travel motivations: Travel or tour motivation is one of the important factors we have found from literature reviews when tourists are selecting their destinations. This variable describes the reason that a tourist chooses to visit a destination.

4) Tourist sociodemographic information: The individual demographics may influence the

information seeking behaviour.

## IV. IMPLEMENTATION AND RESULTS

In this paper author is implementing C4.5 decision tree algorithm with MRMR features selection to recommend travel areas to tourist by using dataset from past tourist experiences. All existing algorithms such as collaborative or content filtering algorithms uses current user past experience data to recommend him new locations. These algorithms will not work if this current user has no past experiences data.

To overcome from above problem author is asking to use C4.5 decision tree algorithms which take experiences of previous users and then build a model and if new user enter his requirements then decision tree will predict best location based on his given input. Decision tree don't need new users past experience data.

To implement decision tree model we need to have dataset and this dataset sometime will have empty or garbage values and this values will put bad effect on decision tree model so we can remove such empty or garbage values by applying pre-process techniques.

Sometime to predict or build model no need to use all columns (attributes) values from dataset and this unnecessary attributes can be remove by apply features selection algorithms and here we are using MRMR features selection algorithms to remove unnecessary attributes to reduce execution time of building model and to increase system accuracy.

Below are the dataset columns or attributes taken from previous users to build model.

This data set is populated by crawling TripAdvisor.com. Reviews on destinations in 10 categories mentioned across East Europe are considered. Each traveller rating is mapped as Excellent (4), Very Good (3), Average (2), Poor (1), and Terrible (0) and average rating is used against each category per user.

Dataset columns and values

userid,art_galleries,dance_clubs,juice_bars,restaurants,museums,resorts,parks_picnic_spots,beaches,theaters,religious_institutions,location

Above are the column names and below are the column values

User 1,0.93,1.8,2.29,0.62,0.8,2.42,3.19,2.79,1.82,2.42,Amsterdam_Heining_2
User 2,1.02,2.2,2.66,0.64,1.42,3.18,3.21,2.63,1.86,2.32,Amsterdam_Jachthaven_ijbur
User 3,1.22,0.8,0.54,0.53,0.24,1.54,3.18,2.8,1.31,2.5,Amsterdam_Bert_Haanstra_Kad
User 4,0.45,1.8,0.29,0.57,0.46,1.52,3.18,2.96,1.57,2.86,Amsterdam_Ruigoord_Ker

In above values first column is USER_ID and second column is ART_GALLERIES and third is DANCE CLUB etc and for each column user had given rating from 4 to 0 and 4 means Excellent service.

Now using above values we can build C4.5 decision tree and prediction will be done using below test values

'User 122',0.93,1.8,2.29,0.62,0.8,2.42,3.19,2.79,1.82,2.42,?
'User 222',1.02,2.2,2.66,0.64,1.42,3.18,3.21,2.63,1.86,2.32,?
'User 3222',1.22,0.8,0.54,0.53,0.24,1.54,3.18,2.8,1.31,2.5,?
'User 4222',0.45,1.8,0.29,0.57,0.46,1.52,3.18,2.96,1.57,2.86,?

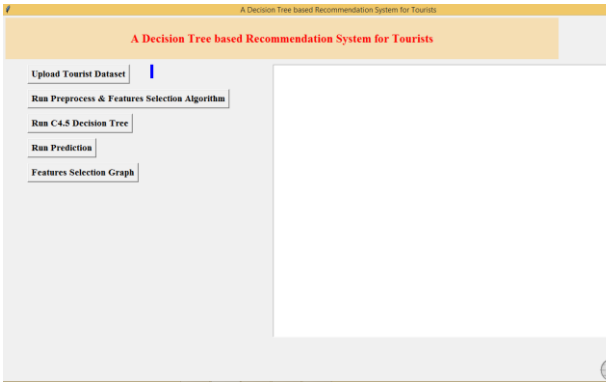
ISSN 2454 - 5015

'User 522',0.51,1.2,1.18,0.57,1.54,2.02,3.18,2.78,1.18,2.54,?

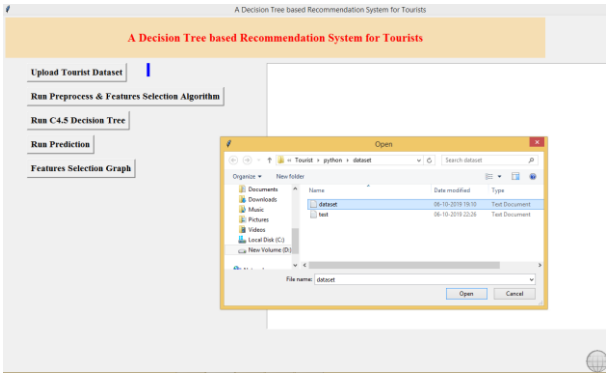
In above test values new user has given values to look for location which has above service rating but new user don't know which location provides such services so he will put question mark and when we upload above test values to decision tree then it will take decision and predict best location and inform to user.

Screen shots

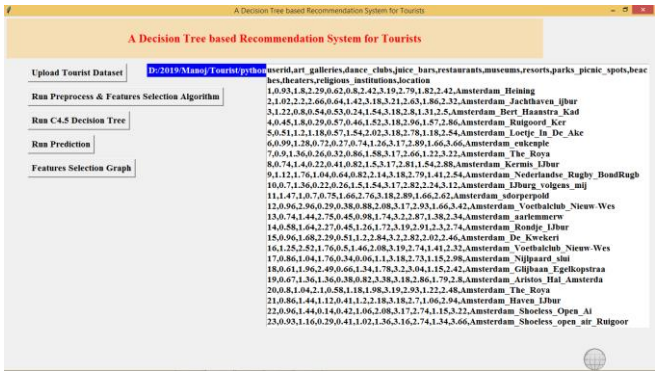
Double click on 'run.bat' file to get below screen



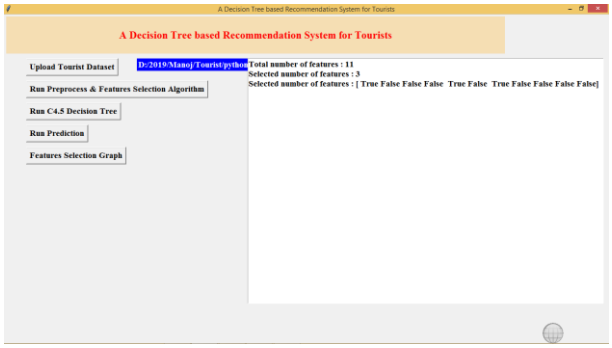

In above screen click on 'Upload Tourist Dataset' button and upload dataset file



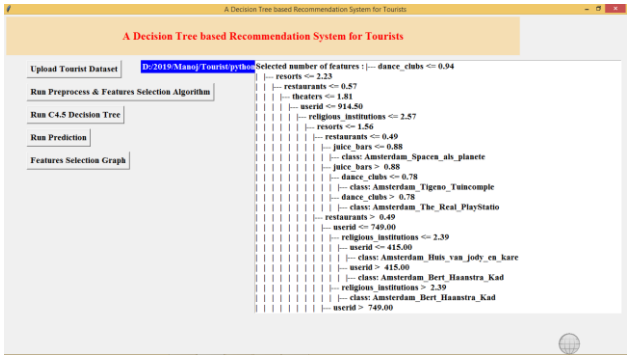

After file upload will get below screen with all dataset details




In above screen all users past experience dataset loaded and total 12 attributes are there in the dataset. Now click on 'Run Preprocess & Feature Selection Algorithm' button to remove empty values and reduce attributes size.



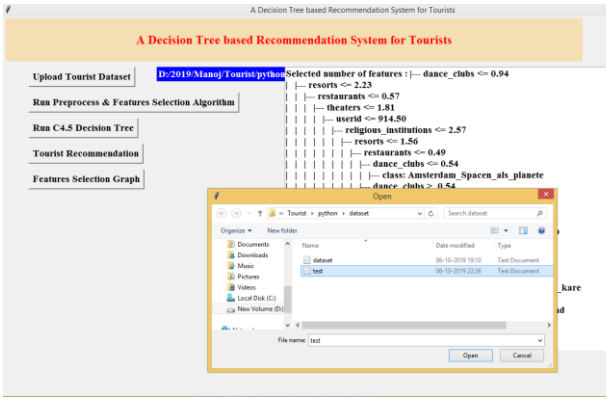

In above screen after applying MRMR features size reduces to 3 and only those attributes will be used whose column is TRUE and FALSE column will be ignore. Now click on 'Generate C4.5 Decision Tree Model' to build model



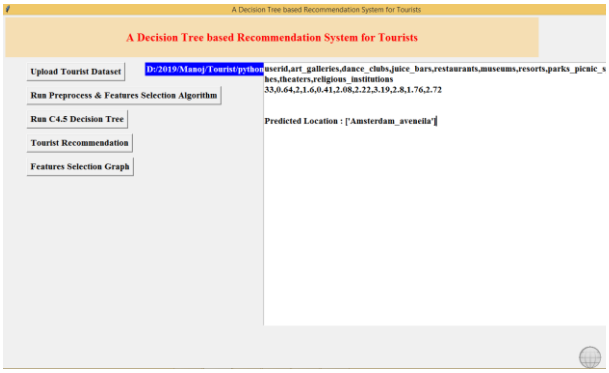

In above screen we can see using IF and ELSE statement decision tree has generated model. If > it will choose some decision if < it will choose some other

ISSN 2454 - 5015

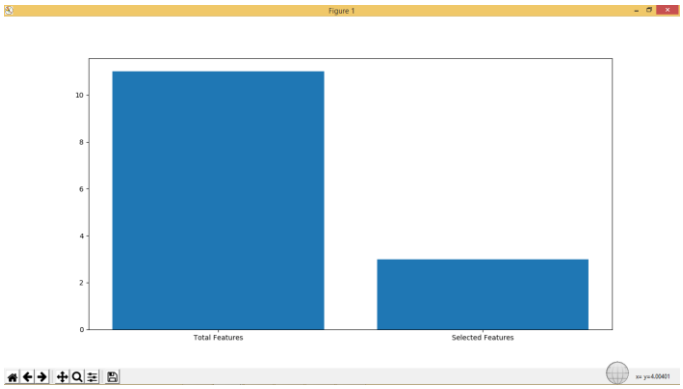
decision. Now click on 'Tourist Recommendation' button to upload test file with no location name and application will predict it




In above screen i am uploading test file now click open to get predicted or recommended location. In test file location name is not there application will give




In above screen after uploading test data we can see all values are there in test data but it not has location name and base on test values application predicted or recommend location name.
Now click on Features Selection Graph button to get below graph




In above graph x-axis contains total features and MRMR selected features and y-axis represents count of features and in above graph we can see after applying MRMR technique features size reduces to 3.

## V. CONCLUSION & FUTURE WORK

a decision tree based tourist recommendation system has been presented in attempt of solving the current challenge of the destination TRS. The data set has been decomposed into two sub data sets using relevant tourism domain knowledge. This was done to increase classification accuracy rate and to reduce the complexity of the decision tree. The optimal decision trees from NMIFS with the highest accuracy rate and simplicity (i.e. less number of leaf and tree size) have been constructed for destination choice. The decision rules from decision trees were extracted. It can be seen that NMIFS is the optimum method because it uses fewer number of feature than MRMR for both of the data sets. Finally, the experimental results confirm applicable of

the proposed a TRS. The proposed TRS satisfies the tourists' requirements who plan to visit or during their visit the city of Chiang Mai.

## VI. REFERENCES

[1] J.Chiverton, "Helmet Presence Classification with Motorcycle Detection And Tracking",IET Intelligent Transport Systems,Vol. 6, Issue 3, pp. 259–269, March 2012.

[2] Rattapoom Waranusast, Nannaphat Bundon, Vasan Timtong and Chainarong Tangnoi, "Machine Vision techniques for Motorcycle Safety Helmet Detection", 28th International Conference on Image and Vision Computing New Zealand, pp 35-40, IVCNZ 2013.

[3] Romuere Silva, Kelson Aires, Thiago Santos, Kalyf Abdala, Rodrigo Veras, Andr´e Soares, "Automatic Detection Of Motorcyclists without Helmet", 2013 XXXIX Latin America Computing Conference (CLEI).IEEE,2013.

[4] Romuere Silva, "Helmet Detection on Motorcyclists Using Image Descriptors and Classifiers", 27th SIBGRAPI Conference on Graphics, Patterns and Images.IEEE, 2014.

[5] Thepnimit Marayatr, Pinit Kumhom, "Motorcyclist"s Helmet Wearing Detection Using Image Processing", Advanced Materials Research Vol 931- 932,pp. 588-592,May-2014.

[6] Amir Mukhtar, Tong Boon Tang, "Vision Based Motorcycle Detection using HOG features", IEEE International Conference on Signal and Image Processing Applications (ICSIPA).IEEE, 2015.

[7] Abu H. M. Rubaiyat, Tanjin T. Toma, Masoumeh Kalantari-Khandani, "Automatic Detection of Helmet Uses for Construction Safety", IEEE/WIC/ACM International Conference on Web Intelligence Workshops(WIW).IEEE, 2016.

[8] XINHUA JIANG "A Study of Low-resolution Safety Helmet Image Recognition Combining Statistical Features with Artificial Neural Network".ISSN: 1473-804x

[9] Kunal Dahiya, Dinesh Singh, C. Krishna Mohan, "Automatic Detection of Bike-riders without Helmet using Surveillance Videos in Real-time", International joint conference on neural network(IJCNN). IEEE, 2016.