World Applied Sciences Journal 35 (New Advancement of Research & Development in Computer Science): 153-158, 2017
ISSN 1818-4952
© IDOSI Publications, 2017
DOI: 10.5829/idosi/wasj.2017.153.158

A Decision Tree Model for Backpackers' Accomodations Recommender System

Maizan Mat Amin, Jannifer Yep Ai Lan, Mokhairi Makhtar,
Abd Rasid Mamat, Siti Dhalila Mohd Satar and Azilawati Rozaimee

Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Malaysia

Abstract: As backpackers travel for a longer period of time, they often have their own budgets and requirements on accommodations. The existing systems do not provide personalized recommendation criteria and some proposed inefficient recommender system (RS) for users. Moreover, only limited systems were specially designed for backpackers' accommodations recommender system other than information searching from websites and bloggers. An observation and online survey was conducted to get the information from backpackers regarding their preferences when looking for the accommodations. Fifty (50) respondents were involved in the survey and the data have been analyzed and were classified to build a decision tree. The decision tree model then implemented in the Backpackers' accommodations Recommender System (BRS). BRS offers a convenient way and solution for backpackers by including decision tree technique in the system to suggest best accommodations suit to backpacker's preferences.

Key words: Model · Recommender System · Decision Tree Technique · Backpackers' Accommodations

INTRODUCTION

Explosive growing of information has resulted in the problem of information overload [1]. The need for effective information retrieval and filtering tools became vital for easy access to relevant information. Technological aids such as information retrieval systems, information filtering systems, intelligent agents, ranking algorithms, clustering techniques, categorization techniques, data mining techniques, web mining techniques, personalization and recommender systems to unravel the problem of information overhead [2].

Recommender systems are information filtering systems that deal with the problem of information overload by filtering vital information piece out of big volume of dynamically generated information according to user's preferences, interest, or observed behaviour about item. Recommender systems (RSs) have become progressively popular in recent years and are exploited in a diversity of areas comprising search queries, research articles, books, social tags, services and products in general. The recommender systems are necessary to make selections without adequate personal experience of the alternatives and also improve decision making process and quality [3]. With the development of recommendation approaches and techniques, more and more recommender systems (software) have been implemented and many real-world recommender system applications have been developed.

The aim of this study is to construct a decision tree model for backpackers' accommodations recommender system. Most backpackers get accommodations information and rely on the recommendation from experienced one, either by word of mouth or review printed in newspapers [4] and through websites or blogs [5]. As backpackers often travel for long period of time, they definitely have plans on their travelling budget and requirements on accommodations [6]. Hence, they always search and decided for the best planning possible including the accommodation suits to their preferences. The existing systems do not provide personalized recommendation criteria and some proposed inefficient recommender system (RS) for users. Moreover, only limited systems were specially designed for backpackers' accommodations recommender system other than information searching from websites and bloggers. Hence, this study was carried out to construct a decision tree model for backpackers' accommodation recommender system. The remainder of this paper is structured as follows: Following section describes the Materials and Methods and next, is Results and Discussions and finally is Conclusions.

Corresponding Author: Maizan Mat Amin, Faculty of Informatics and Computing, Besut Campus, 22200, Terengganu, Malaysia

MATERIALS AND METHODS

This section explains the materials and methods that are used in constructing the decision tree model for backpackers' accommodations recommender system. Decision tree is one of the classification technique used in data mining, decision support system and machine learning process. A decision tree is a predictive modeling technique uses in classification, clustering and predictive task. Decision tree uses a divide-conquer technique to split the problem search space into subsets. The most important feature of decision tree classifier is their ability to break down a complex decision making process into collection of simpler decision, thus providing solution which is easier to interpret. In this study, the steps involved in this study shown in Figure 1. The steps is regarding the process of data mining: Data Preprocessing, Data Analysis and Result Interpretation.

Data: Data defines as a collection of objects (also known as record, item, point, sample, observation, or instance) and their attributes (also known s property, variable, field, characteristic, or feature). An observation and an online survey have been done to get to point out the criteria to be considering for backpackers' preferences in making decision in choosing the best accommodations.

Observation: An observation from selected websites and blogs has been done to discover the criteria commonly required by backpackers in searching accommodations. Based on Table 1, only few criteria are consider in existing websites for filtering users preferences such as destination, check-in date,check-out date, guest number. This study propose more criteria to take into concern such as destination, price, types of traveller, types of room, food and beverage offer, private locker offer, pick-up service and a trip package/adviser. The reviews and evaluation made on existing systems will be one of the guidelines in developing model for backpacker's accommodation recommender system. Most of the evaluations made were intended to improvise the quality of the existing accommodation Recommender Systems (RS) to cope with backpackers' preferences on RS.

Online Survey: An online survey also was conducted to get the information from backpackers regarding their preferences when looking for accommodations. The questionnaire consists of twenty four (24) questions regarding preferences for choosing backpackers accommodations. Fifty (50) respondents were involved in the online survey. The data of backpacker's accommodation preferences have been analyzed and were classified using decision tree technique. The details explain further in the next section.

Data-Preprocessing: Real-life data typically needs to be preprocessed (e.g. cleansed, filtered, transformed) in order to be used by the machine learning techniques in the analysis step. In designing a RS, first step is to review different similarity or distance measures and then the issue of sampling, a way to reduce the number of items in very large collections while preserving its main

Table 1: Characteristic of existing and proposed accommodation RS

Characteristics	Website 1 [7]	Website 2 [8]	Website 3 [9]	Proposed Backpackers RS
Designed for backpackers	No	Yes	Yes	Yes
As recommender only	Yes	Yes	No	Yes
Less navigation process	Yes	No	Yes	Yes
Price filtering option	Yes	No	No	Yes
Offer more than three criteria	Yes	No	No	Yes
Criteria offered	Destination, check-in date, check-out date, guest number	Destination	Check-in date, number of night	Destination, price range, travel types, rooms types, private locker, food/beverage, trip package/advise, pickup service

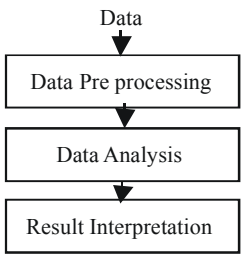

Fig. 1: Main step and methods used

characteristics. Sampling is used both in the preprocessing and final data interpretation steps. Sampling may be useful because processing the entire data set is computationally too costly. It can also be used to produce *training* and *testing* datasets. The training dataset is used to learn the parameters or configure the algorithms used in the analysis step, while the testing dataset is used to evaluate the model or configuration obtained in the training phase, making sure that it performs well or generalizes with previously unseen data.

Data Analysis: Classification is a pervasive problem encompasses diverse applications. It is the task of assigning objects to one of several predefined categories (Refer to Figure 2). In classification, a classifier is a mapping between a attribute set and a class label, where the attribute represent characteristics of the elements to classify and the labels represent the classes. Examples include decision tree classifiers, rule-based classifiers and neural networks. Each technique employs a learning algorithm to identify a model that best fits the relationship between attribute set and class label of the input data. Therefore, the model generated by a learning algoritm should both fit the input data well and correctly predict the class labels of unknown records.

A backpackers accomodations RS, for example, can be implemented by classifing backpackers accomodations into few categories based on a number of features such as destination, price, types of traveller, types of room, food and beverage offer, private locker offer, pick-up service and a trip package/adviser. For this study, supervised classification has been used where a set categories is known in advance and we have a set of labeled examples which constitute a training set.

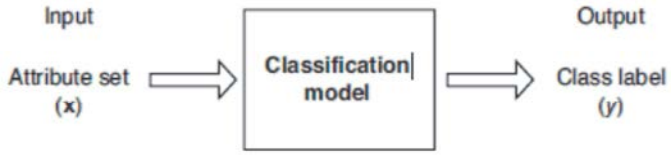

Fig. 2: Classification as the task of mapping an input attribute set x into its class label y

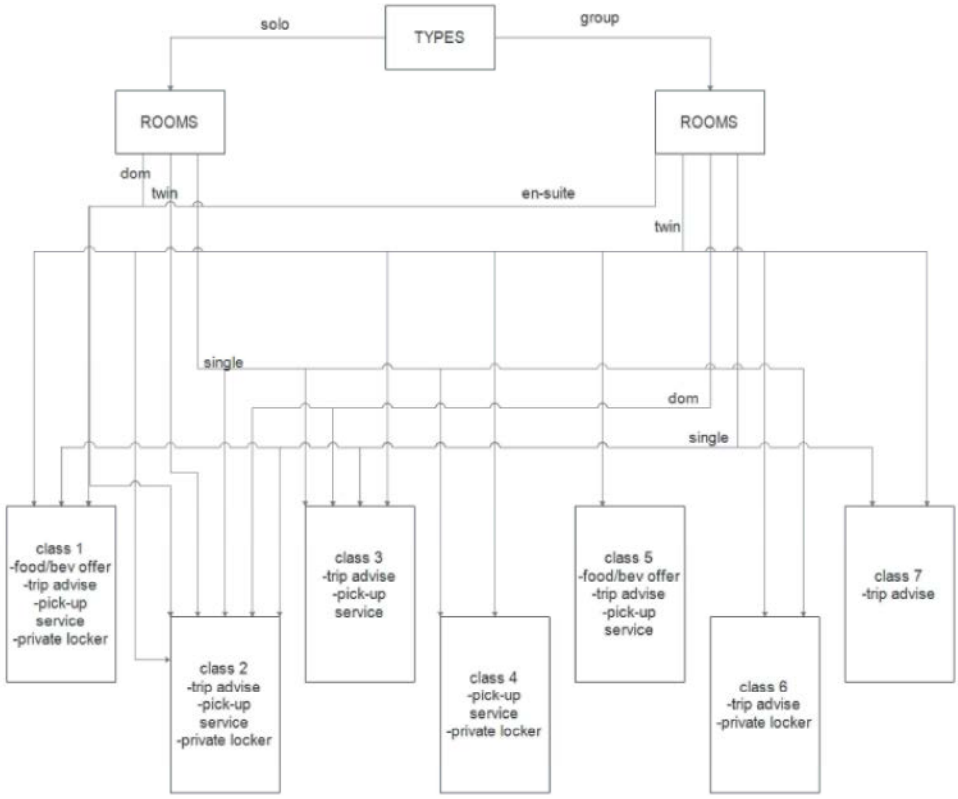

Fig. 3: Classification of backpacker's accommodation preferences using decision tree technique

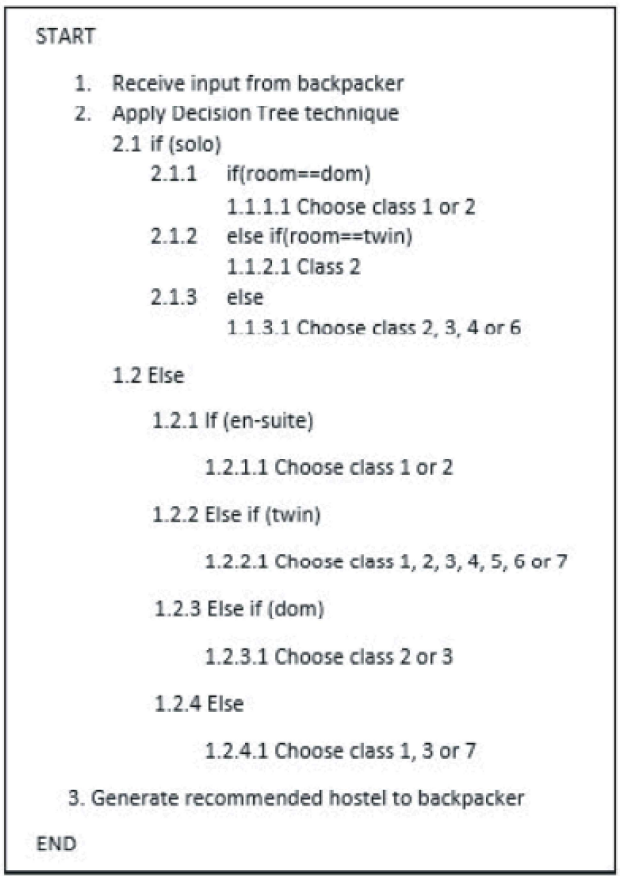

Fig. 4: If-then rules for classification of backpacker's accommodation preferences using decision tree technique

Decision Tree: Decision trees [(Quinlan, 1986; Rokach, L., & Maimon, 2014)] are classifiers on a target attribute or class in the form of a tree structure. The observations or items to classify are composed of attributes and their target value. The nodes of the tree can be: a) *decision nodes*, in these nodes a single attribute-value is tested to determine to which branch of the subtree applies.or b) leaf nodes which indicate the value of the target attribute.

There are many algorithms for decision tree induction: Hunt's Algorithm, ID3, C4.5, SLIQ, CART and so on. C4.5 Algorithm has been chosen to be used in this study. It is the successor to ID3 and removed the restriction that features must be categorical by dynamically defining a discrete attribute (based on numerical variables) that partitions the continuous attribute value into a discrete set of intervals. Decision tree induction stops once all observations belong to the same class. C4.5 converts the trained trees into sets of if-then rules. This accuracy of each rule is then evaluated to determine the order in which they should be applied. Pruning is done by removing a rule's precondition if the accuracy of the rule improves without it. From the data in this study, it has been classified to seven class as shown in Figure 3. Then, the trained trees converts into sets of if-then rules as shown in Figure 4.

Result Interpretation: The proposed model of classification will be implemented in the system to generate list of recommended accommodations based on backpacker's preferences. The model use to filter and propose users or backpackers' according to their preferences in searching for suitable accommodations. The user has more personalized and accurate searching result based on their preferences.

RESULTS AND DISCUSSION

Based on an observation and online survey, there are few criteria that backpackers normally take into consideration when searching and choosing accommodations via online. The criteria are different compare to hotel or homestay accommodation recommender system. The criteria or preferences consist more than just destination and price, where also consider

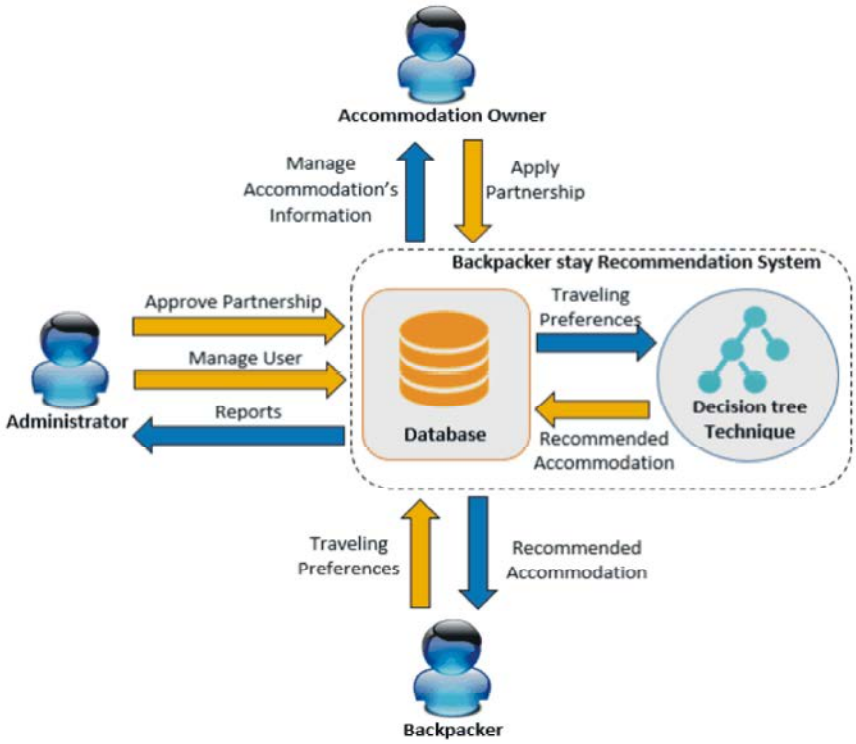

Fig. 5: Framework of Backpackers' accommodation recommendation system

the types of traveler, types of room, food and beverage offer, private locker offer, pick-up service and a trip package/advisor.

The proposed model and classification will implement in backpackers' accommodations' recommender system as shown in Figure 5. In the framework, the users for the system are Administrator, Accommodation Owner and Backpackers. Administrator will manage the database and approving the partnership. Accommodation owner will request for partnership and may manage the information regarding accommodations. Users or backpackers may search for the suitable accommodation based on their preferences. All the information such as partnership request, traveling preferences, entities personal details were store in system's database. The traveling preferences from backpackers were then used to generate the accommodations' recommendation using the proposed model. Meanwhile, reports and entities details can be retrieving by the users for update and review.

CONCLUSIONS

As more information overwhelming on the Internet, there is a need to filter, prioritize and efficiently deliver relevant information in order to alleviate the problem of information overload. Recommender systems solve the problem by searching through great volume of vigorously generated information to provide users with personalized content and services. In this study, the proposed model using decision tree technique as it is simple to understand and to interpret, require little data preparation, able to handle both numerical and categorical data and possible to validate a model using statistical tests. The backpackers' accommodation recommender system (BRS) with the employed model may help backpackers in getting the best accommodations based on their preferences. By using this model, the BRS may bring out potential benefit including easier to find for specific accommodations, effective searching and thus time saving for efficiently the process.

REFERENCES

1 Konstan, J.A. and J. Riedl, 2012. Recommender systems?: from algorithms to user experience, pp: 101-123.
2. RVVSV Prasad, V.V. and Kumari, 2012. A C Ategorical R Eview of R Ecommender S Ystems, 3(5): 73-83.
3. Ricci, F., L. Rokach and B. Shapira, 2005. Recommender Systems?: Introduction and Challenges.

4. Resnick, P., H.R. Varian and G. Editors, 1997. Recommender Systems, 40(3): 56-58.
5. Isinkaye, B.A., F.O. Folajimi and Y.O. Ojokoh, 2015. Recommendation systems: Principles, methods and evaluation, Egypt. Informatics J., 16(3): 261-273.
6. Mccarthy, L. and R.V.P.D, 2010. How Travelers Use Online and Social Media Channels to Make Hotel-choice Decisions How Travelers Use Online and Social Media Channels to Make Hotel, 10: 6-18.
7. "HotelsCombined," 2017. [Online]. Available: https://www.hotelscombined.com/.
8. "Backpacker Hostel," 2017. [Online]. Available: http://www.backpackerhostel.net/.
9. "Gilligans Backpackers Hotel and Resort," 2017. [Online]. Available: https://www.gilligans.com.au/.
10. Quinlan, J.R., 1986. Induction of decision trees, Mach. Learn., 1(1): 81-106.
11. Rokach, O. and L. Maimon, 2014. Data Mining with Decision Trees: Theory and Applications.