A Decision Tree Based Recommender System for Backpackers Accommodations

Maizan Mat Amin[1]*, Jannifer Yep Ai Lan[1], Mokhairi Makhtar[1] , Abd Rasid Mamat[1]

[1] *Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Besut Campus,*
22200, Besut, Terengganu, Malaysia
*Email: maizan@unisza.edu.my

Abstract

Backpackers often travel for a longer period of time, have their own budgets and requirements on accommodations. The existing systems do not offer personalized recommendation criteria and some proposed inefficient recommender system (RS) for users. Moreover, other than information searching from websites and bloggers, only limited systems were specifically designed for backpackers' accommodations recommender system. An observation and online survey was conducted to get the information from backpackers regarding their preferences while 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.

Keywords: *Backpackers' Accommodations, Classification Model, Data Mining, Decision Tree Technique, Recommender System*

1. Introduction

The problem of information overload happened as the explosive growing of information [1]. The need for effective information retrieval and filtering tools became essential for easy access to relevant information. Technological benefits 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 untangle the problem of information overhead [2].

Recommender systems (RSs) 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 behavior about item. It 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]. More and more recommender systems (software) have been implemented with the development of recommendation approaches and techniques, and many real-world recommender system applications have been developed.

The aim of this study is to develop backpackers' accommodations recommender system using a decision tree model. Most backpackers get accommodations information and depend 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 reasonable travelling budget and requirements on convenient accommodations [6]. Hence, they al-ways 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, other than information searching from websites and bloggers, only limited systems were specially designed for backpackers' accommodations recommender system. Hence, this study was carried out to construct a decision tree model and implemented in 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.

2. Materials and Methods

This section explains the materials and methods that are used in developing backpackers' accommodations recommender system based on constructed decision tree model [7]. The steps involved in this study shown in Fig. 1. The steps is regarding the process of data mining: Data Pre-processing, Data Analysis, and Result Interpretation.
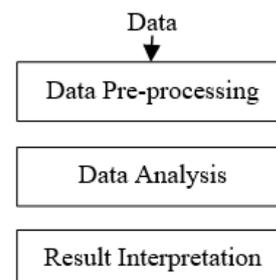

Fig. 1: Main step and methods used

Data

Data defines as a collection of objects (also known as record, item, point, sample, observation, or instance) and their attributes (also known as property, variable, field, characteristic, or feature). An observation and an online survey have been done to get to highlight the criteria to be considering for backpackers' preferences in making decision in choosing the best accommodations [7].

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

The reviews and evaluation made on existing systems are considered as 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.

An online survey also was conducted to get the information from backpackers regarding their preferences when considering 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 analysed and were classified using decision tree technique.

Data Pre-processing

Real-life data typically needs to be pre-processed (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 characteristics. Sampling is used both in the pre-processing 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 the task of assigning objects to one of several predefined categories. 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 (Refer to Figure 2). 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 algorithm should both fit the input data well and correctly predict the class labels of unknown records.
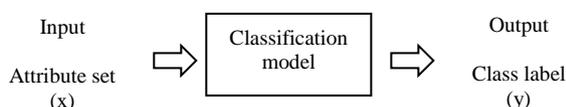
Fig. 2: Classification as the task of mapping an input attribute set x into its class label y.

In this proposed system, decision tree is used as a classification technique. Few categories of backpackers' accommodations have been classified based on criteria 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. Supervised classification has been used where a set categories is known in advance and we have a set of labelled examples which constitute a training set.
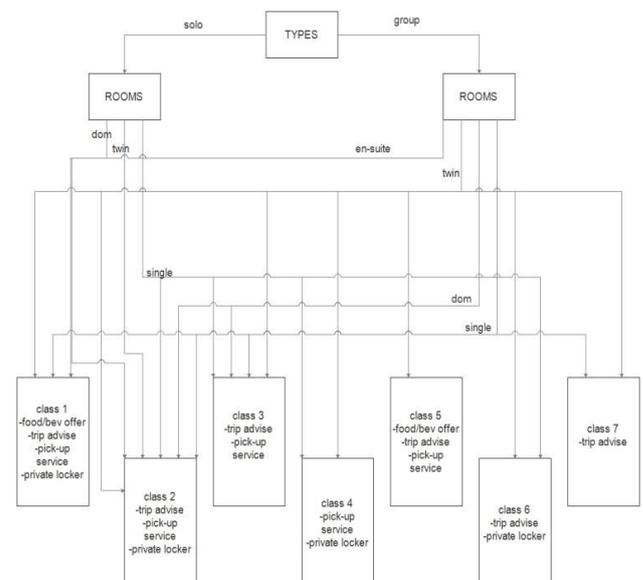
Decision tree is one of the classification technique used in data mining, decision support system and machine learning process [8]. Decision tree uses a divide-conquer technique to split the problem search space into subsets. The most important feature of decision tree classifier is their ability to break down a complex decision making process into collection of simpler decision, thus providing solution which is easier to interpret [9][10].

Decision trees 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.

C4.5 Algorithm has been chosen to be used in this study. 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 classes as shown in Fig. 3. Then, the trained trees convert into sets of if-then rules as shown in Fig. 4.

Result Interpretation

The proposed model of classification was implemented in the backpacker's accommodation system. 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.

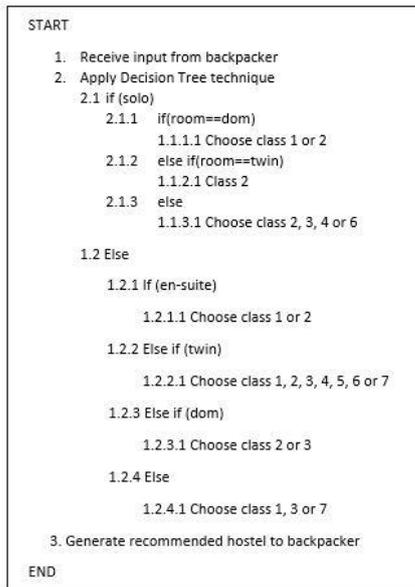**Fig. 3:** Classification of backpacker's accommodation preferences using decision tree technique

START

1. Receive input from backpacker
2. Apply Decision Tree technique
   2.1 if (solo)
      2.1.1 if(room==dom)
         1.1.1.1 Choose class 1 or 2
      2.1.2 else if(room==twin)
         1.1.2.1 Class 2
      2.1.3 else
         1.1.3.1 Choose class 2, 3, 4 or 6

   1.2 Else
      1.2.1 If (en-suite)
         1.2.1.1 Choose class 1 or 2
      1.2.2 Else if (twin)
         1.2.2.1 Choose class 1, 2, 3, 4, 5, 6 or 7
      1.2.3 Else if (dom)
         1.2.3.1 Choose class 2 or 3
      1.2.4 Else
         1.2.4.1 Choose class 1, 3 or 7
3. Generate recommended hostel to backpacker

END

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

3. Results and Discussion

The development of the backpackers' accommodations' recommender system can be presented through framework as shown in Fig. 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. The proposed model and classification was implemented in this system.

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.
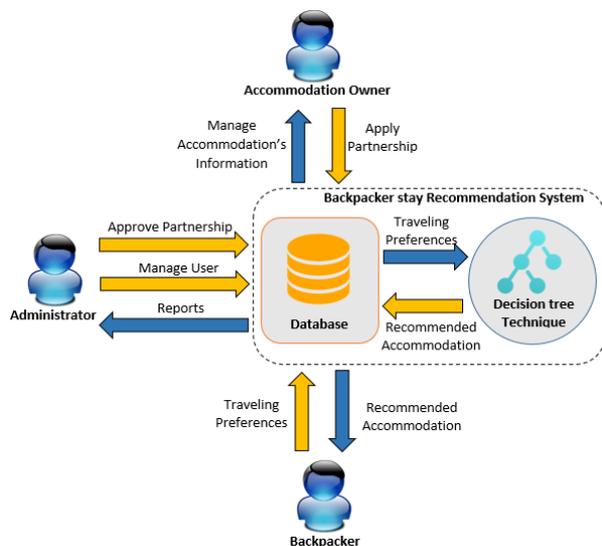

Fig. 5: Framework of Backpackers' accommodation recommendation system.

Fig. 6 shows the main page for Backpackers' Accommodation Recommender System (BRS). There are buttons Home, Be a Host,

Sign Up/Login and Contact. Be a Host button is for Accommodation Owner to join the partnership in this system. Sign Up button is for a new user and Login button is for registered user. Contact button provided to get connected to the admin of the system.

After the user login to the system, homepage for backpackers appear as shown in Fig/ 7. Backpackers may search by filling in the form given for accommodations based on their preferences. By clicking Check Now button, backpackers are directed to Recommended Accommodation page (Fig. 8). In this page, the best suitable accommodation is listed based on the criteria selected by backpacker. By clicking on the accommodation name, backpacker may view the accommodation details as shown in Fig. 9. Fig.10 and Fig. 11 also shows the page of highest review and the cheapest accommodations available.

Fig. 6: Main page of Backpackers' Accommodation Recommender System (BRS)

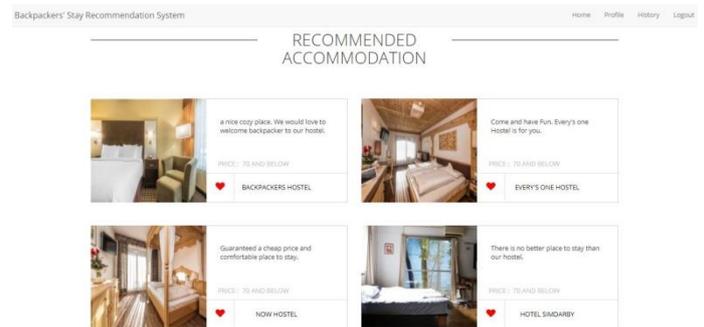**Fig. 7:** Home page for Backpackers.

Fig. 8: Recommended Accommodation page
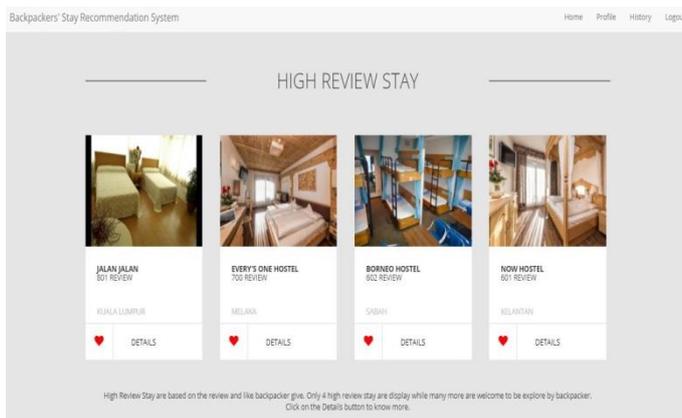

Fig. 9: Accommodation Details page.

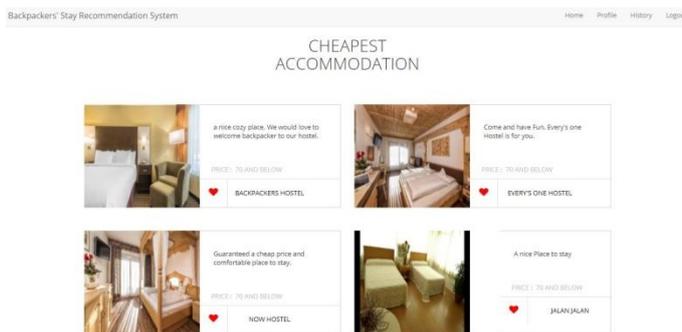**Fig. 10:** High Review Stay page.

Fig. 11 : Cheapest Accommodation page.

4. Conclusion

In order to alleviate the problem of information overwhelming on the Internet, there is a need to filter, prioritize and efficiently deliver relevant information. 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. In future, backpackers' accommodations recommender system may consider more criteria for more accurate recommendations and add reservation or booking functions as well.

Acknowledgement

References

[1] J. A. Konstan and J. Riedl, "Recommender systems : from algorithms to user experience," pp. 101–123, 2012.

[2] V. V. RVVSV Prasad, Kumari, "A C ATEGORICAL R EVIEW OF R ECOMMENDER S YSTEMS," vol. 3, no. 5, pp. 73–83, 2012.

[3] F. Ricci, L. Rokach, and B. Shapira, *Recommender Systems : Introduction and Challenges*. 2015.

[4] P. Resnick, H. R. Varian, and G. Editors, "Recommender Systems," vol. 40, no. 3, pp. 56–58, 1997.

[5] B. A. Isinkaye, F. O., Folajimi, Y. O., & Ojokoh, "Recommendation systems: Principles, methods and evaluation," *Egypt. Informatics J.*, vol. 16(3), pp. 261–273, 2015.

[6] L. Mccarthy and R. V. P. D, "How Travelers Use Online and Social Media Channels to Make Hotel-choice Decisions How Travelers Use Online and Social Media Channels to Make Hotel- ," vol. 10, pp. 6–18, 2010.

[7] M. M. Amin, J. Yep, A. Lan, and M. Makhtar, "A Decision Tree Model for Backpackers ' Accomodations Recommender System," *World Appl. Sci. J.*, vol. 35, pp. 153–158, 2017.

[8] X. Wu, V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, G. J. Mclachlan, A. Ng, B. Liu, P. S. Yu, Z. Z. Michael, S. David, and J. H. Dan, *Top 10 algorithms in data mining*. 2008.

[9] J. R. Quinlan, "Induction of Decision Trees," pp. 81–106, 2007.

[10] O. Rokach, L., & Maimon, *Data mining with decision trees: theory and applications*. 2014.