Tarbiat Modares University

Faculty of Engineering

Information Technology Engineering Department

Thesis report

Ontology-Based Tourism Activity Recommender

(Case study: Tehran province)

Student

Rasoul norouzi

Supervisor

Dr. Amir Albadvi

Advisor

Dr. Elham Akhondzadeh

This manuscript is an under-preparation paper based on my master thesis.

Note that this version may have grammatical and structural defects.

Ontology-Based Tourism Activity Recommender

Abstract

Although the increase of information on the web has created good opportunities, on the other hand, it has caused serious problems such as information redundancy so that users have difficulty finding their desire information. This problem is more acute in the tourism industry, where the decision-making process is emotional for the user. In such cases, using Recommender Systems (RS) for personalization is a good way, but these systems have some problems: data sparsity and cold start. This study tries to solve the two mentioned problems by proposing an ontology-based method, collecting data with questionnaires, and recommending tourism activities. In this method, by the user's rating to tourism categories (ontology's first layer), his/her favorite tourism activities in ontology's second layer are predicted and recommended. Experimental results show that the proposed method has captured the user's interests well by the user scoring to seven tourism categories. Also, the comparison of this method with two state-of-the-art algorithms, BSKNN and SVD, in Collaborative Filtering (CF) systems shows our method's slight superiority.
Keywords**:** Recommender System, Tourism, Multi-Label Classification, Ontology

1. Introduction

Information technology and its integration with human life have led to a new way of human behavior and interactions. Information technology has increased the speed of processes, reduced costs, and made it easier to access services. These benefits are twofold, both the users and service providers can benefit from it. However, the diverse range of information on the web has led to information redundancy, which in turn has led to difficulties in the decision-making process and finding information that meets users' needs. As a result, the need for the systems that provide personalized and appropriate information to the user strongly felt (Grün et al., 2017).

Personalization is the ability to provide tailored content and services to users based on knowledge about their preferences and tastes (Gao et al., 2010). Personalization techniques are particularly related to recommender systems (RS), which aim to filter irrelevant information and to provide personalized information to each particular user (Borràs et al., 2014). A recommender system can be defined as a personalization tool that provides people with a list of information items that best fit their individual preferences, restrictions, or tastes (Sebastia et al., 2009). The most important feature of a recommender system is that it can guess a user's preferences and interests by analyzing the behavior of this user and/or the behavior of other users to generate personalized recommendations (Lu et al., 2015). The recommender systems can automatically learn the user's preferences through the analysis of his or her explicit or implicit feedback. Explicit data might be given by the user in different ways, for instance, requiring the user to fill a questionnaire about his or her preferences and interests. Implicit interests can be inferred by the system through the analysis of the behavior of the user (Borràs et al., 2014). The main aim of developing recommender systems is to counteract the risk of information overload by assisting users in searching for relevant information from a huge amount and variety of information (Grün et al., 2017). Commonly used recommendation techniques include knowledge-based (KB), collaborative filtering (CF), and content-based (CB) techniques or Hybrid RS, which combines the mentioned techniques (Lu et al., 2015). KB systems are based on understanding users and items' features and their underlying relationship. Such systems try to recommend items that satisfy users' needs based on users and items' features. Ontology is used for representing domain knowledge in knowledge-based systems. The domain's ontology can compute the logical similarity between items and users (Borràs et al., 2014). CF base systems help people to make their choice based on the opinions of people who are similar to them. The similarity between users is calculated based on the scoring rate they have given to the items on the list. When the system finds out which people are closer to each other based on interests and choices, the favorites of other users are suggested to the intended user. In this approach, feedback is necessary to find out which recommendation is favorable and which is not. CF systems use the user-item matrix to predict users' interest in items, in that each row represents a user, and each column represents an item, and each cell represents a user giving rate to an item (Ricci et al., 2015). Such systems have two main drawbacks, first cold-start problem refers to by entering a new user to the system since there is no record of the user interests and rates to the items, it is not possible to predict user's interests. The same problem can exist for newly added items to the system. Another problem is that the user-item matrix is too sparse and high dimensional, which many of the cells of this matrix suffer from the lack of the user giving rate to the items. This can make it challenging to train machine learning algorithms, and it causes false assumptions about the user's preferences (Thiengburanathum, 2018). In CB, the system measures the degree of similarity between the user and the items. This process is done by analyzing product features concerning user preferences. Therefore, it is assumed that both the user and the items selected for the recommendation have common features. The output of the analysis process usually displays the overall performance score, which indicates how much the user profile matches with recommended items. Selected items that get more points will have better performance, and their compliance rate is high. Sometimes this approach also deals with the user's scoring history. In this approach, the system must have accurate knowledge of the user to provide a recommendation. These systems suffer from the cold-start problem too (Lu et al., 2015).

A typical trip plan is made up of a number of steps, from selecting tourist attractions, choosing accommodations, to deciding routes, etc.(Huang and Bian, 2009). A tourist attraction is often the main motivation for a tourist to choose a particular destination for traveling (Huang and Bian, 2009). Therefore, selecting tourist attractions to visit is one of the most crucial decisions in trip planning, and it is strongly connected to the tourists' travel preferences and interests, which are not explicitly known (Hsu et al., 2012).Tourists are often confused about where to go when reaching new and unfamiliar places because there could be a wide variety of selections for consideration (Yang et al., 2013). In addition, they typically have a limited amount of time and budget available; thus, it is almost impossible to visit all tourist attractions during a trip, especially to large cities (Abbaspour and Samadzadegan, 2011). As a result, tourists have to select the most interesting points of interest (POIs) according to their preferences. Then, they plan a route among them, taking into account the time available to reach the different POIs with regard to their accessibility and opening hours (Vansteenwegen et al., 2011). Over the last decade, the quick development of Information and Communication Technologies (ICTs) and the global expansion of the Internet strongly influenced the tourism sector as it affected all the sectors of the economy (Büyüközkan and Ergün, 2011). People more and more have been realizing the advantages of ICTs for planning an agenda of recreational and leisure contexts (Garcia et al., 2011), to the extent that the Internet has become an integral and inseparable part of tourism (Neidhardt et al., 2015). However, the volume of information related to travel and tourism available on the web is increasing at a tremendous rate, and users are usually confronted with too many options to choose from (Gao et al., 2010). In the procedure of trip planning, the most posted and searched information is concerning travel destinations and their associated resources, such as tourist attractions, accommodations, restaurants, local gourmet food etc. to appeal to the tourists. Although all this information may be particularly useful for those users who plan to visit an unfamiliar place, the evaluation of this long list of options to select the one that fits better with a particular tourist needs is overwhelming and time-consuming (Borràs et al., 2014). Furthermore, despite the fact that almost two-thirds of all Internet users make use of this medium to search for travel-related information (European Travel Commission 2009)[1], around one-third of these users cannot explicitly express their travel needs and expectations (Zins and Tourism, 2007). Hence, they require help for exploring and filtering out irrelevant information based on their specific preferences and needs identified as personalization. In recent years, recommendation systems increasingly have caught the attention of scholars in the tourism field as a powerful tool for supporting information searching and decision making in this context (Batet et al., 2012, García-Crespo et al., 2009b, Ricci et al., 2002, Lucas et al., 2013, Chiang and Huang, 2015, Lee et al., 2009, Castillo et al., 2008, Hsu et al., 2012, Huang and Bian, 2009, Loh et al., 2003, Fesenmaier et al., 2003, Moreno et al., 2013, Neidhardt et al., 2015, Venturini et al., 2006, Montejo-Ráez et al., 2011, Rivolli et al., 2017, Sebastia et al., 2009, Yang et al., 2013).

As mentioned, the recommender systems suffer from two common problems of cold start and data sparsity. To mitigate these problems, this study has proposed an ontology-based method in which the user preference for each of the 18 identified activities will be predicted by only rating to seven tourism categories. These seven tourism categories were extracted based on a questionnaire and used instead of using demographic explicit data that can be sensitive to the user (Moreno et al., 2013, Schiaffino and Amandi, 2009). Our proposed ontology comprises three layers; the zero layer or domain layer, which refers to tourism. The first layer refers to the category layer, which contains tourism categories, and finally, the second layer or activity layer includes tourism activities. The proposed ontology can have deeper layers, like instance layer, which refers to related POI to each of the activities in the activity layer; however,

---

[1] http://www.newmediatrendwatch.com (Accessed in 2010).

this study focuses only on predicting the user's favorite activities in activity layer based on user rating to each of categories in the category layer. Therefore, proposing touristic places/destinations is not our aim in this research. Based on this method, touristic places recommendation can be made by post-processing tasks such as meta-heuristic or context-aware algorithms (Tung and Soo, 2004, Pashtan et al., 2003, Ruotsalo et al., 2013, Console et al., 2003, García-Crespo et al., 2009a, Grün et al., 2017). Thus, in this study, instead of using the user-item matrix, we will have the user-activity matrix, which also referred to the context in the literature. The user-activity matrix has two advantages; first, the user-activity matrix will have much smaller dimensions than the user-item matrix. Also, adding a new item that can be a new touristic place does not change the user-activity matrix structure, because the new item will be a subset of the identified tourism activities. Another advantage of this approach is the possibility of using such methods for news and tourism websites that display the desired content according to the user's preferences in tourism activities (Baltrunas et al., 2010, Zheng et al., 2014). Besides, this study has a user-category matrix that indicates users' interest rates to each category. We formulate this problem to a Multilabel Classification (MLC) problem. MLC is a kind of supervised machine learning algorithm, each instance can belong to multiple classes or, in other words, can have multiple labels (Ganda and Buch, 2018). MLC is used to find the mapping between the user-activity matrix and the user-activity matrix, in other words finding a space that can map the category layer into the activity layer in the ontology.

The rest of this paper in the research methodology section deals with how to identify tourism activities, how to extract tourism categories, proposing algorithms to predict the user's favorite activities, and how to evaluate this method. In the next section, we review and analyze the results, and in the last step, in the conclusion section, we review our method and discuss the achievements of this paper, research limitations, and suggestions for future research.

2. Literature

2.1 Ontology

The information available on the web is often published by various travel information providers with vastly diverse backgrounds. Different terms might be used by numerous providers to represent the same meaning or the same terms for different meanings (Huang and Bian, 2009). Furthermore, when users start trip planning, travelers' preferences are often hidden and are not explicitly known (Loh et al., 2003). Therefore, as a consequence of this heterogeneity of the information, it is difficult to integrate the information automatically. Ontology is a conceptualization of a domain into a human-understandable, machine-readable format consisting of entities, attributes, relationships, and axioms (Tho et al., 2006). Ontologies define areas of common understanding between multiple actors, easing their interoperability, and permitting a high-level communication (Berners-Lee et al., 2001). Much research has been conducted in applications of ontology in many research fields. Maillot et al. (2008) proposed an ontology-based cognitive vision approach for complex object recognition. Lee et al. (2007) proposed an automated ontology construction for unstructured text documents. In the year 2006, Francisco et al. suggested an ontology-based recruitment system to provide intelligent matching between employer advertisements and the curriculum vitae of the applicants. Belmonte et al. (2008) developed a multi-agent decision support system for the bus fleet management domain. Recently, different tourism ontologies have also been developed to facilitate information management. Some of them have reached a considerable level of consolidation, allowing the representation of not only generic aspects but also specific sub-domains that describe detailed scenarios (Moreno et al., 2013). The DERI e-tourism ontology (Hepp et al., 2006), developed in the On Tour project, covered three main subjects, namely accommodation, contexts, and infrastructures. Then, Prantner et al. (2007) developed a tourism ontology, entitled Manteca, which included important concepts of the tourism domain that are defined in the World Tourism Organization (WTO) thesaurus developed by the WTO. In the QALL-ME project, Ou et al. (2008) developed a domain-specific ontology and used for multi-modal and multi-lingual question-answering in the domain of tourism. Barta et al. (2009) developed a core Domain Ontology for Travel and Tourism (cDOTT) ontology, based on the Harmonize ontology. The Harmonize was one of the first ontologies that aimed to overcome the interoperability problems of tourism, concentrating on data exchange between organizations. The Harmonize covered four main topics of the tourism domain including attractions, events, food and drink, and accommodation.

2.2 Multi-label Classification

In machine learning, single-label classification is one of the commonly used methods in which each instance in the data set associate with a unique class label from a set of disjoint class labels L. Depending on the number of these classes, our problem can be binary calcification (when $|L| = 2$) or multi-class classification (when $|L| > 2$). But in the multi-labeling problems, each instance can be associated with multiple classes. In such algorithms, the goal is to learn from a set of instances to label each instance's class or classes in L (Sorower, 2010). MLC approaches categorize into a) problem transformation, and b) algorithm adaptation methods. In problem transformation, the multi-label classification problem transforms into one or more single-label classification problems. Therefore, it does not need any changes or adaptations to traditional algorithms, and those algorithms can apply to the problem (Tsoumakas et al., 2007). Problem transformation methods are divided into three main categories mainly Binary Relevance (BR), Label Power Set (LP), and Classifier Chain (CC), which this study applies five classifiers namely Support Vector Machine (SVM), Decision Tree (DC), Random Forest (RF), Naïve Bays (NB), and K-Nearest Neighbor (KNN) by these three algorithms. In adaptation algorithms, instead of transforming the problem, the algorithms are changed and modified to handle multi-label data. We use two adaptation algorithms namely Binary Relevance KNN (BRKNN) and Multilabel K Nearest Neighbors (MLKNN) (Tsoumakas et al., 2007, Spyromitros et al., 2008). Besides these approaches, Ensemble learning algorithms can learn from MLC data natively without any transformation in base algorithms or the problem. Ensemble methods are learning algorithms that construct a set of classifiers and then classify new data points by taking a (weighted) vote of their predictions (Rokach et al., 2014). This study uses Random Forest (RF) (Pal, 2005) and Extra Tree classifiers (ET) (Geurts et al., 2006) as ensemble algorithms candidates. MLC has many applications in such domains like text categorization and sentimental analysis (Pestian et al., 2007, Liu and Chen, 2015), image classification(Wang et al., 2016), bioinformatic (Zhang et al., 2018), genre classification (Sanden and Zhang, 2011), social media analysis (Chen et al., 2014, Huang et al., 2013). For more details, see Tidake et al. (2018) and Tsoumakas et al. (2009) papers. Moreover, MLC has leveraged its power into recommender systems domains too. Carrillo et al. (2013) demonstrated the MLC ability for the recommendation and dealing with RS common problems like data sparsity. Zheng et al. (2014) Have used MLC to recommend users' contexts. In this method, instead of recommending the item to the user, it predicts user-related contexts based on the items selected by the user and the ratings given to each item. To this end, they have transformed the problem into an MLC problem and have shown that the use of MLC algorithms is more capable of recommending and predicting than base algorithms. Rivolli et al. (2017) used the MLC method to recommend track foods. They obtained a set of data using a questionnaire. In this questionnaire, 21 questions ask the user, which are the attributes that describe the user. These 21 questions were considered predictive attributes. The second part of the questionnaire is 12 food alternatives in which the user is asked to specify their preferences to each of them. These alternatives associate with classes' labels or target attributes. Their results show that adopted method performance is weaker in comparison to transformation methods. Elhassan et al. (2018) Used MLC approaches to provide remedial actions to address students' shortcomings in Learning Outcome Attainment Rates. In their model, each instance is a student who is described by a set of characteristics such as field of study, academic level, grades, and so on. Related tags are also remedial actions for each student. The results show that for the given dataset, the chain classification method with the decision tree algorithm gives the best results.

3. Research Methodology

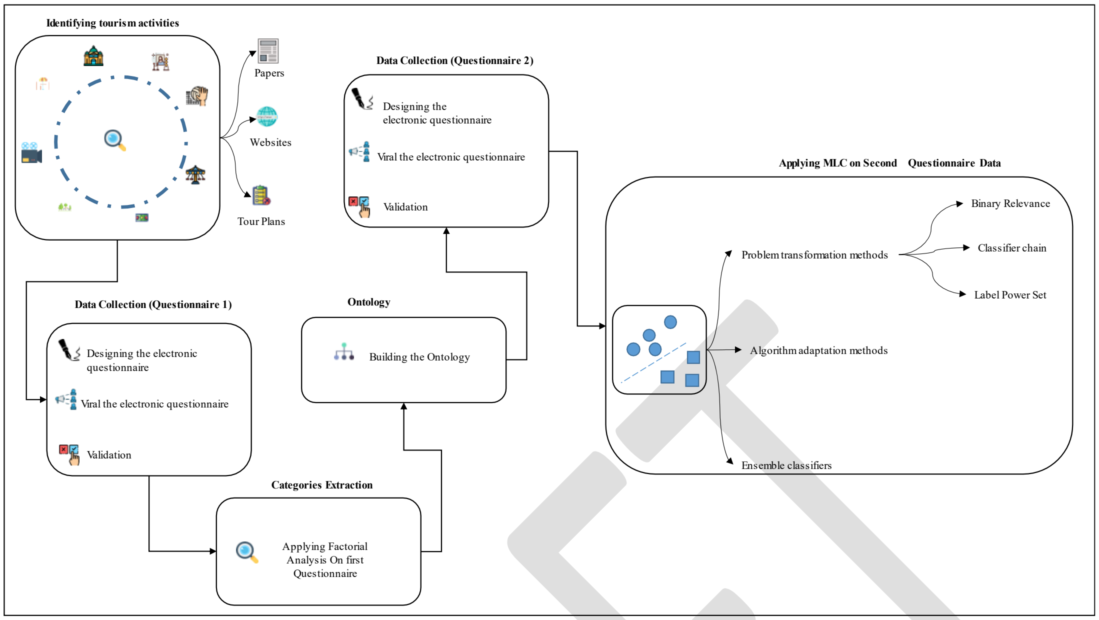

Figure 1 The stages' flow of proposed method

This research tries to predict which tourism activity the user is interested in based on the user rates to each of the extracted categories. In this section, we will address to tourism activities identification process, the first phase of data collection by questionnaire, applying factorial analysis on the first dataset to extract tourism categories, second phase data collection by questionnaire, training, and reporting the performance of different learners' algorithms on the second dataset. All of the proposed method stages are represented in a graphical manner in figure 1.

3.1 Identifying tourism activities

To identify tourism activities in Tehran, we used previous research (Moreno et al., 2013), analytical reports of the British Tourism Organization[2], as well as the famous website Trip Advisor[3]. The point to keep in mind was that many of the tourism activities mentioned in the paper and British Tourism Organization's analytical reports did not exist in Tehran, such as nightclubs or beach tours; Therefor by combining and modifying the activities mentioned in Previous references, we identified 18 types of related tourism activities in Tehran; namely going to Cinema, going to Theaters, visiting Museums, visiting Holy sites, visiting Historical Sites, going to Sports Events, going to Sports activities, visiting Art and Book Exhibitions, going to Music Events, going to Malls, visiting Public Gardens, going to Restaurants, going to Cafe, visiting Zoo, visiting Rural Places, visiting Rivers and Lakes, and going to Mountains.

3.2 Data Collection (Questionnaire 1)

Reviewing and rating each of these 18 tourism activities on a tourism news website can be a difficult and boring task for the user. Thus, reducing these 18 activities to a few less and more interpretable categories makes the process of recognizing and categorizing content more comfortable for the user. We need data to get tourism categories; therefore, a questionnaire is designed to measure users' interest in each of these 18 activities based on the Likert scale, a scale between one and five. One indicates the least interest in a activity, and the five score the most. To better guide users, we introduced a number of POI in Tehran for each of the mentioned activities as instances. As an example, for visiting the historical sites, Saadabad Palace and Negarestan Mansion were mentioned. After designing the questionnaire, it was randomly distributed in social media like Facebook, Twitter, and messenger applications. Totally 272

[2] https://www.visitbritain.org/archive-great-britain-tourism-survey-overnight-data
[3] https://www.tripadvisor.com/Attractions-g293998-Contexts-Iran.html

questionnaires were collected, and the reliability of the designed questionnaire proved by calculating Cronbach's alpha equal to 0.846, which was more than the cutoff requireed of 0.7 (Kopalle et al., 1997).

3.3 Extracting Tourism categories (Factorial Analysis of questionnaire 1)

The main goal of Factorial Analysis (FA) is summarizing data for revealing relationships and patterns with regrouping variables into a limited collection of clusters based on shared variance (Yong and Pearce, 2013). Indeed, FA utilizes mathematical methods to simply interrelated measures for discovering patterns in a set of variables (Child, 1990). FA was applied in different types of fields such as behavioral and social sciences, medicine, economics, and geography (Yong and Pearce, 2013) and is divided into two main classes namely Exploratory Factorial Analysis (EFA) and Confirmatory Factorial Analysis (CFA). Trying to confirm hypotheses is the goal of CFA and EFA attempts to discover complex patterns (Child, 1990). EFA is used when the research goal is discovering the number of influencing variables or finding variables go together (McDonald, 1985). FA is useful for studies such as questionnaires based on few or hundreds of variables which can be reduced to a smaller set of variables leading to simplify interpretations. So, focusing on a smaller set of variables rather than having to regard too many keys is easier and makes variables meaningful by clustering them into some clusters (Rummel, 1970). In this paper EFA was applied for accessing meaningful categories of variables. The determinant score of our data is 0.0000135, which is more than 0.00001, and this indicates a violation of the assumption of correlation of variables; in such case, to extract the factors, it is recommended to use Principal Axis Factor (Yong and Pearce, 2013).For rotation, we used the Varimax rotation method with 30 iteration bases on the default value in SPSS software. To check the adequacy and suitability of the dataset for EFA, Kaiser-Meyer-Olkin Measure (KMO), and Bartlett's Test of Sphericity applied. The minimum value of the KMO index for factor analysis is 0.5, which in our research is 0.76. The Bartlett test takes a statistical hypothesis, and its null hypothesis states that the correlation matrix is an identity matrix, so there is no significant relationship between the variables. As we can see in the table 1, the $p-value$ is not in the rejection area (the value of sig must be less than 0.05, which is zero for our data).

Table 1 The experimental results of FA on the first questionnaire

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.76
Bartlett's Test of Sphericity	Approx. Chi-Square	1833.255
	Df	325
	Sig.	0

To determine the number of significant factors, Kaiser's criteria states only factors should retain which their Eigenvalues are one or more than one. According to the Figure 2 (Scree Plot), the best number of factors after rotation for this dataset is 7.

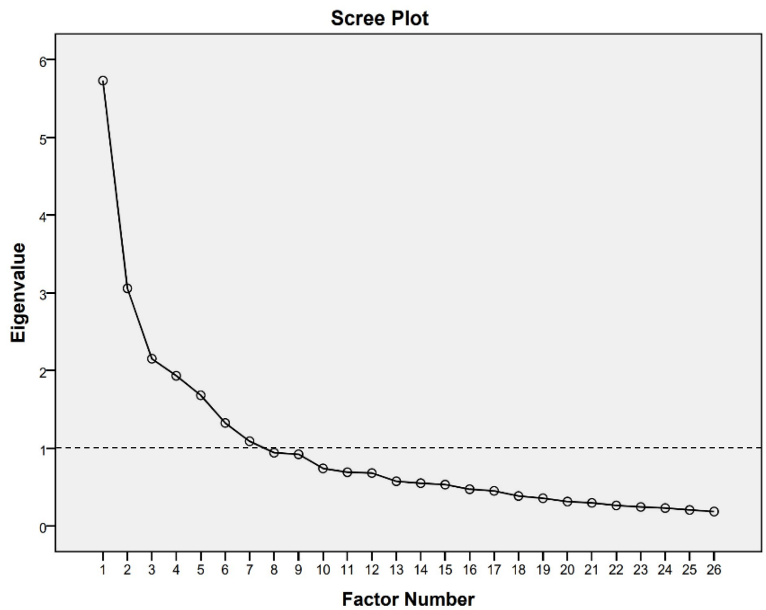

Figure 2 Scree Plot for determining the optimum number of factors, factors below eigenvalue one are dropped out

Factor naming does not follow a specific rule, and here we used to name each factor based on the associated variables that describe that factor (table 2).

Table 2 Extracted factors name and their associates' variables

	Factor	Descriptor Variables
1	Historical	Museums, Historical Sites
2	Fun	Restaurants, Café, Male
3	Ecotourist	Rivers and Waterfalls, Lakes, Mountains, Rural Places
4	Sportive	Sports Activities, Sports Events
5	Cultural	Music Events, Art and Book Exhibitions, Cinema, Theaters
6	Religious	Mosques and churches, Holy Sites
7	Urban-related	Zoos, Public Gardens, and Parks

3.4 Ontology

In this research, the ontology has three levels, the zero-level representing the domain name which is the tourism, the first level or category layer, represent the extracted categories, and the second layer or the activity layer is tourism activities. In this ontology, the relationship between the category layer and the activity layer is unknown, as mentioned, the purpose of this study is to provide a method that can map the category layer to the activity layer (figure 3). For this purpose, MLC algorithms have been used, which in the next section, we will formulate the problem.

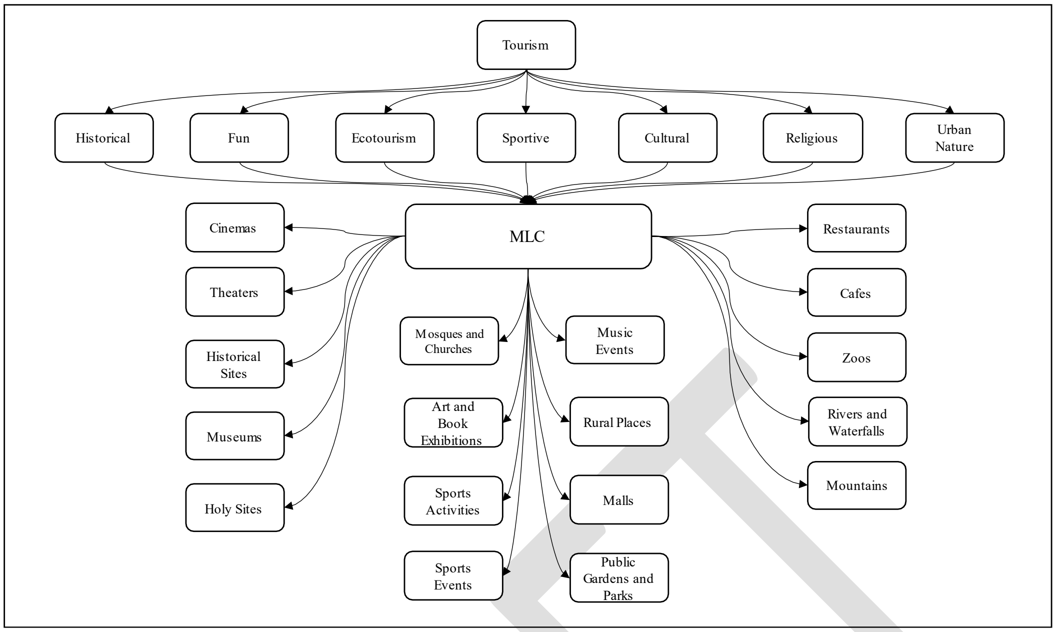

Figure 3 The proposed three-layer ontology

3.5 Data Collection (Questionnaire 2)

This paper's main goal is to predict users' interest in each of the 18 activities based on their scoring to every of seven tourism categories was extracted by factorial analysis. So, a dichotomous questionnaire was designed. The first part asked people to determine their interesting rate for every category on a five-point scale Likert. The second part requested to fulfill every 18 items with binary values (0 means dislike, and one means like regarded item). This electronic questionnaire was randomly broadcast through social media, and messengers' application for people who were residents of Tehran. Finally, 578 questionnaires were collected, and the calculated Cronbach's alpha was equal to 0.859.

3.6 Multi-Label Classification Problem Definition

Let χ be the Users-Category matrix, and $\mathcal{L} = \{\lambda_1, \lambda_2, \dots, \lambda_k\}$ be a finite set of labels or activities. A user $x \in \chi$, represented in terms of features vector $x = (x_1, x_2, \dots, x_m)$, which these features vector is referred to the given rates of a user to each of extracted categories; therefore, the user x is associated with a subset of labels $L \in 2^{\mathcal{L}}$. Notice that if we call this set L be the set of relevant labels of x, then we could call the complement $\mathcal{L} \backslash L$ to be the set of irrelevant labels of x. Let denote the set of relevant labels L with a binary vector $Y = (y_1, y_2, \dots, y_k)$, where $y_i = 1 \Leftrightarrow \lambda_i \in L. \mathcal{Y} = \{0,1\}^k$ is the set of all such possible labeling (figure 4).

	Features vector or given rates to each category				*Associated subset of labels*			
	$category_1$	$category_2$	···	$category_m$	$activity_1$	$activity_1$	···	$activity_k$
$user_1$	3	5		1	1	0		1
$user_2$	2	4		5	0	0		1
⋮	⋮	⋮		⋮	⋮	⋮		⋮
$user_n$	3	2		5	1	1		0

Figure 4 Part of a multi-label problem matrix

Therefore:

Given a training set, $S = (x_i, Y_i), 1 \leq i \leq n$, consisting n training instances, $(x_i \in \mathcal{X}, Y_i \in \mathcal{Y})$ $i.i.d$[4] drawn from an unknown distribution D, the goal of the multi-label learning is to produce a multi-label classifier $h: \mathcal{X} \rightarrow \mathcal{Y}$(in other words, $h: \mathcal{X} \rightarrow 2^{\mathcal{L}}$) that optimizes some specific evaluation function (i.e. loss function) (Sorower, 2010).

This study uses the second questionnaire as training data set for MLC algorithms; as a result, by entering a new user to the system, only by rating to each of seven categories in range 1 to 5, his/her interest to 18 activities will be predicted. In the transformation approach, all three algorithms (BR, LP, CC) with LR, DT, RF, SVM, KN classifiers used. In the adaptation algorithm, BRKNN and MLKNN and in ensembles algorithm, RF and ET classifiers were used. To implement MLC algorithms, Python version 3.5 with scikit-learn and scikit-multilearn packages were used. In this study, all the classifiers' Hyperparameters are the packages default values.

An issue should be considered is imbalanced labels problems; as shown in Figure 5, the number of classes is not equal in any of the labels, and this could cause problems in some algorithms' learning process.

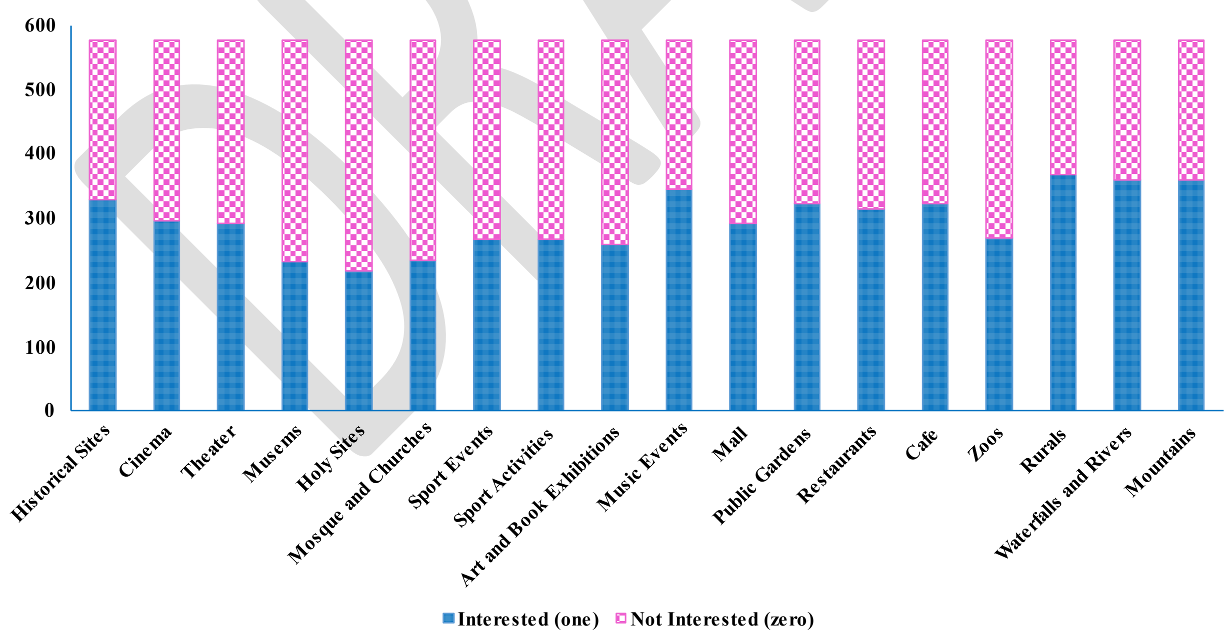

Figure 5 Labels imbalance, vertical pivot represents the number of samples and the horizontal pivot shows the label names

[4] independent and identically distributed

To solve this problem, we used the Scikit Learn Package[5] class weight balancing feature in all classifiers except the NB, MLKNN, and BRKNN; because the technique is not applicable to such algorithms.

3.7 Evaluation

In this research, the evaluation stage is important in two ways, first, to compare the performance of different MLC algorithms and approaches on present research dataset, and second, to compare our proposed method with conventional methods of CF recommenders. For this purpose, 5-fold cross-validation with tree metrics was used. Cross-validation can train algorithms with little data; Moreover, since all samples are used both as training data and as test data in the algorithm learning process, it is a good way to compare different algorithms performance in classification problems. There are some different metrics for MLC evaluation, but we should notice that metrics should be chosen that are usable for CF methods too; as mentioned, one of our goals in the evaluation stage is comparing the proposed method with state-of-the-art CF techniques. Thus, in this study, we select Precision, Recall and F1-Score metrics for evaluation. Since weight balancing technique is not applicable for some classifiers (NB, MLKNN, BRKNN), we use the macro-average criteria that does not take label imbalance into account.

Metrics

Let T be a multi-label dataset consisting n multi-label examples (x_i, Y_i), $1 \leq i \leq n$, $(xi \in \mathcal{X}, Yi \in \mathcal{Y} = \{0,1\})^{k})$, with a label set $\mathcal{L}, |\mathcal{L}| = k$. Let h be a multi-label classifier and $Z_i = h(x_i) = \{0,1\}^k$ be the set of label memberships predicted by h for the example x_i.

Therefore:

Precision (p): Precision is the proportion of predicted correct labels to the total number of actual labels, averaged over all instances. In our case, Precision indicates how much of the predicted activities are correct for the user.

$$Precision, P = \frac{1}{n}\sum_{i}^{n}\frac{|Y_i \cap Z_i|}{|Z_i|}$$

Recall: Recall is the proportion of predicted correct labels to the total number of predicted labels, averaged over all instances. In our case, Recall indicates how much the algorithm has been able to predict the user's favorite activities.

$$Recall, R = \frac{1}{n}\sum_{i}^{n}\frac{|Y_i \cap Z_i|}{|Y_i|}$$

F1-Score: Definition for precision and recall naturally leads to the following definition for F1-score (Sorower, 2010, Ganda and Buch, 2018):

$$F1 = \frac{1}{n}\sum_{i}^{n}\frac{2 * |Y_i \cap Z_i|}{|Y_i| + |Z_i|}$$

CF System vs. Proposed Method

CF systems generally have two approaches for the recommendation, memory-based and model-based. In memory-based, there are not any assumptions on data and are essentially based on the nearest neighbors' search to find the closest pairs of items or users. When the recommendation is based on measuring the similarity between items, it is an item-item method. When the recommendation is based on measuring users' similarity, it is a user-user method. For our problem, we focus on the user-user method because the number of items (activities) is few. When items are few, in item-item method variance is low, and its bias is high, which could cause less personalization. On the other side,

[5] https://scikit-learn.org/stable/modules/generated/sklearn.utils.class_weight.compute_class_weight.html

model-based systems try to make an assumption on data and learn a model that explains user-item matrix interaction (figure 6).

	$activity_1$	$activity_1$	...	$activity_k$
$user_1$	1	0		1
$user_2$	0	0		1
.	.	.		.
.	.	.		.
.	.	.		.
$user_n$	1	1		0

Figure 6 An example of user-item matrix

To formulate our problem for CF systems, we only need the user activity matrix (users' interest to activities) without user-category (their rating to each category). Thus, let we have N users and K activities, define the user-activity matrix $A \in \{0,1\}^{N \times K}$:

$$A_{ui} = \begin{cases} r_{ui}, user\ u\ interested\ on\ activity\ i & if\ such\ rating\ exicit \\ ? & if\ no\ such\ rating \end{cases}$$

CF systems try to replace all the “question marks” in A by some optimal guesses; the goal is to minimize the RMSE (root mean square error) when predicting the user interests on a test set (which is of course unknown during the training phase), that is to minimize:

$$rmse = \sqrt{\frac{1}{|S_{test}|} \sum_{(i,u) \in S_{test}} (r_{ui} - \hat{r}_{ui})^2}$$

where $(m, u) \in S$ test if User u interest activity i in the test set, $|S_{test}|$ is its cardinality, r_{ui} is the true rating and $\hat{r}_{ui}$ is the prediction based on the recommendation system (Wen, 2008).

The use of CF systems was performed using the surprise package in the Python programming language (Hug, 2020); we also used the benchmark results of this package to select the top two CF algorithms[6].

In memory-based approach we choose BSKNN that taking into account a baseline rating; A baseline estimate for an unknown rating r_{ui} is denoted by b_{ui} and accounts for the user and item effects:

$$b_{ui} = \mu + b_i + b_u$$

The parameters b_m and b_u indicate the observed deviations of user u and activity m, respectively, from the average and μ denotes the overall average rating. To predict $b_{m,u}$ that is to minimize problem:

$$\min \sum_{(u,i) \in \kappa} (r_{ui} - \mu - b_u - b_i)^2 + \lambda_1 (\sum_u b_u^2 + \sum_i b_i^2)$$

Therefore, the prediction $\hat{r}_{ui}$ is set as:

$$\hat{r}_{ui} = \mu_u + \frac{\sum_{v \epsilon N_i^k(u)} sim(u, v).(r_{vm} - \mu_v)}{\sum_{v \epsilon N_i^k(u)} sim(u, v)}$$

Where $sim(u, v)$ denotes similarity measurement between user u and v, and $N_i^k(i)$ only include neighbors for which the similarity measure is positive (Koren, 2010).

6 http://surpriselib.com/

In model-based approach we used Singular Value Decomposition (SVD), a Matrix factorization algorithm which tries to decompose the original sparse matrix to low-dimensional matrices with latent factors.

The prediction $\hat{r}_{ui}$ is set as:

$$\hat{r}_{ui} = \mu + b_u + b_i + q_i^T p_u$$

If user u is unknown, then the bias b_u and the factors p_u are assumed to be zero. The same applies for item i with b_i and q_i.

To estimate the unknown parameters, minimize the problem:

$$\sum_{r_{ui} \in R_{train}} (r_{ui} - \hat{r}_{ui})^2 + \lambda(b_i^2 + b_i^2 + \|q_i\|^2 + \|p_u\|^2)$$

For more details see Surprise package's document and Koren (2010).

Another problem is left, the output of mentioned algorithm is in range [0,1], but the desired output should be binary, which one and zero respectively denotes a user interest and dislike of an activity. To solve this problem, we suggest using a threshold α where:

$$output = \begin{cases} 1, & \hat{r}_{ui} > \alpha \\ 0, & else \end{cases}$$

For both selected algorithms, to choose the best threshold value, the F1-score is calculated for different α values, the α value having the best F1-score is chosen as the threshold.

4. Results and Discussions

The performance of MLC algorithms is shown in Figure 7. The three bars represent the precision, recall, and F-1 Score, respectively; the higher the bar values, the better the results. For MLC transformation algorithms, we use "Algorithm-Classifier" to denote the algorithm, e.g. BR-NB denotes we use binary relevance (BR) as transformation algorithm and choose Naïve Bayes as the classifier, also for Ensemble methods, we use "Ensemble_ Classifier" form.

Figure 7 Comparison of the MLC algorithms performance

The experimental results show that the best performance is related to BR algorithms and on the other hand LP algorithm has lower metrics values compared to others. The closeness of the precision score and the recall in most classifiers indicate that the imbalance of the classes has not been able to affect the learning process. Since we have 18 tourism activities, BR algorithms transform the problem into 18 separate problems, nor does it take into account the interdependence of labels. The good performance of this algorithm indicates that the detected activities are distinctive from each other, and the algorithm has been able to map the category layer to the activity layer space well. For example, the RF classifier with the BR algorithm had better results than other algorithms with the RF classifier. Nevertheless, the reason for the disappointing results of the LP algorithms in our problem is related to the problem-solving approach in such algorithms. LP turns the MLC problem into a multi-class classification problem with 2^L possible class values; since our dataset has few records and many labels, so it is difficult to train such algorithms on this data set; however, the results and superiority of algorithms may change as the number of data increases. The high score of Recall refers to the success rate of the classifier in identifying and proposing the user's favorite activities, and the higher this score is, the more the recommender has recommended the user's favorite tourism activities. Precision, on the other hand, indicates what percentages of the activity recommended to the user actually was the user's favorite activities. Of course, sometimes, a precision error can also be welcome, i.e., recommend an activity outside of the user's favorite activities, and the user feedback to that be positive; therefore, it can be a way to prevent over-personalization. If a recommender offers all the activities to the user, in such cases its recall score will be 100%, but its precision score will be low and also if it tries to suggest a smaller number of activities to the user, its recall will be low, and Its precision score will be high. In such cases, the F1-score, which is a harmonic average of two mentioned metrics, can be a good criteria for comparing the performance of classifiers. To select the best algorithm based on the F-score results, BR_NB has the best performance among the categories; its Precision, Recall, and F1 scores are 0.75, 0.74, and 0.733 respectively. The evaluation results of the two adaptive algorithms, BRKNN and MLKNN, were similar to each other and did not have a significant advantage over each other, as is the case with the ensemble algorithms.

4.1 BR_NB VS BSKNN and SVD

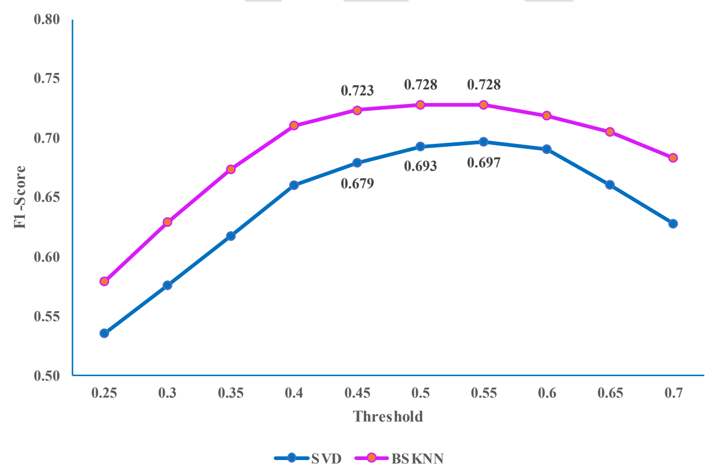

Figure 8 BSKNN and SVD F1-Scores for different α values

The output of the BSKNN and SVD algorithms is in the range of zero to one, so a threshold was used to convert it to a binary output. Figure 5 shows each of these algorithms' F1-Score for different α values. As expected from the performance of these algorithms, with the increase of α, the process of recommending activities to the user becomes more rigorous, and as a result, the amount of recall score decrease and on the other hand, the score of precision increase. For both algorithms, the best equilibrium point is the α value that gives the highest F-Score. As figure 8 reparents, the best value of F1-Score for both algorithms is in the range of 0.45 to 0.55, and the difference of F1 in this limit is negligible for each of the algorithms. Therefore, an α value of 0.5 was chosen for both algorithms.

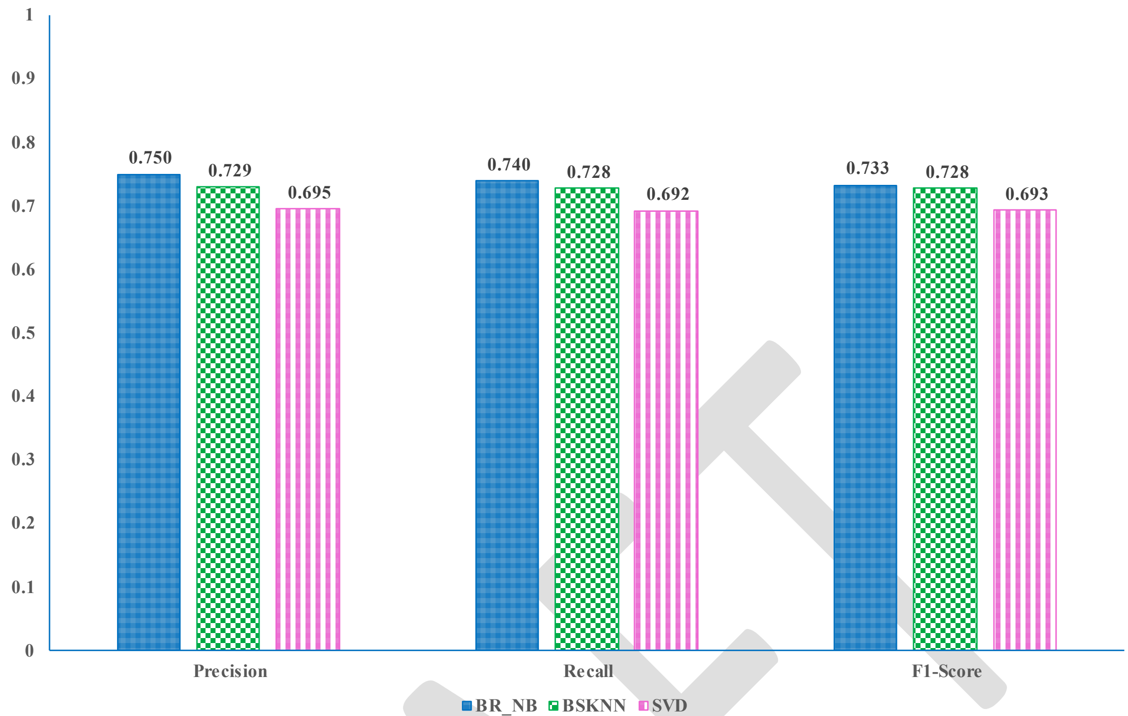

Figure 9 Comparison of BR_NB results in the proposed method in with two state-of-the-art algorithms, BSKNN and SVD, in CF systems

To compare the proposed method with the state-of-the-art CF algorithms, BSKNN and SVD, we selected the BR_NB classifier, which has the best performance among the MLC algorithms. Figure 9 Shows precision, recall, and F1 scores for each of them. The weakest performance is related to the SVD algorithm because of the learning process of such algorithms, which usually use gradients decent to estimate their parameters, while this technique requires many data. The performance of BR_NB in the proposed method is superior to BSKNN algorithm by nearly three percent in precision, nearly two percent in recall, and nearly one percent in F1-score. The performance of the proposed method was slightly better than the CF algorithms, but it should be noted that in the CF systems, we only needed one step of data collection and did not need to extract tourism categories. Nevertheless, the advantage of our method is the user instead of engaging and reviewing 18 tourism activities by only rating to 7 tourism categories, the system will recommend his/her desire activities.

5. Conclusion

In this paper, we presented a tourism activity recommender method that predicts and recommends the user desired activities in the second level of ontology with the user given rate to the tourism categories in the first level of the ontology. The first level of ontology was obtained based on the first questionnaire that measured users' interests to Identified tourism activities. We turned this into an MLC problem, and the BR_NB results performed best compared to other classifiers. We also compared the best classifier result in our proposed method with state-of-the-art CF algorithms, i.e., BSKNN and SVD. Our method performed better than the two mentioned algorithms, although this was a slight advantage. In addition to excellence in evaluation criteria, reducing the size of the problem space from 18 activities to seven tourism categories is another advantage; means that the user does not need to be engaged in 18 activities and related items. As mentioned, this study focuses on recommending tourism activities, which is one of the limitations of the present study. For future research, it is suggested to combine this method with a context-aware approach to suggest items related to each activity according to the user context. Data collection through a questionnaire is challenging and time-consuming, so another limitation of this research is low data volume. For future research, we will collect more data and use newer algorithms for implementing and evaluating.

References

ABBASPOUR, R. A. & SAMADZADEGAN, F. J. E. S. W. A. 2011. Time-dependent personal tour planning and scheduling in metropolises. 38**,** 12439-12452.

BALTRUNAS, L., KAMINSKAS, M., RICCI, F., ROKACH, L., SHAPIRA, B. & LUKE, K.-H. Best usage context prediction for music tracks. Proceedings of the 2nd Workshop on Context Aware Recommender Systems, 2010.

BARTA, R., FEILMAYR, C., PRÖLL, B., GRÜN, C. & WERTHNER, H. Covering the semantic space of tourism: an approach based on modularized ontologies. Proceedings of the 1st Workshop on Context, Information and Ontologies, 2009. 1-8.

BATET, M., MORENO, A., SÁNCHEZ, D., ISERN, D. & VALLS, A. J. E. S. W. A. 2012. Turist@: Agent-based personalised recommendation of tourist activities. 39**,** 7319-7329.

BELMONTE, M.-V., PÉREZ-DE-LA-CRUZ, J.-L. & TRIGUERO, F. J. E. S. W. A. 2008. Ontologies and agents for a bus fleet management system. 34**,** 1351-1365.

BERNERS-LEE, T., HENDLER, J. & LASSILA, O. J. S. A. 2001. The semantic web. 284**,** 34-43.

BORRÀS, J., MORENO, A. & VALLS, A. J. E. S. W. A. 2014. Intelligent tourism recommender systems: A survey. 41**,** 7370-7389.

BÜYÜKÖZKAN, G. & ERGÜN, B. J. E. S. W. A. 2011. Intelligent system applications in electronic tourism. 38**,** 6586-6598.

CARRILLO, D., LÓPEZ, V. F. & MORENO, M. N. 2013. Multi-label classification for recommender systems. *Trends in Practical Applications of Agents and Multiagent Systems.* Springer.

CASTILLO, L., ARMENGOL, E., ONAINDÍA, E., SEBASTIÁ, L., GONZÁLEZ-BOTICARIO, J., RODRÍGUEZ, A., FERNÁNDEZ, S., ARIAS, J. D. & BORRAJO, D. J. E. S. W. A. 2008. SAMAP: An user-oriented adaptive system for planning tourist visits. 34**,** 1318-1332.

CHEN, X., VORVOREANU, M. & MADHAVAN, K. J. I. T. O. L. T. 2014. Mining social media data for understanding students' learning experiences. 7**,** 246-259.

CHIANG, H.-S. & HUANG, T.-C. J. I. F. 2015. User-adapted travel planning system for personalized schedule recommendation. 21**,** 3-17.

CHILD, D. 1990. *The essentials of factor analysis*, Cassell Educational.

CONSOLE, L., TORRE, I., LOMBARDI, I., GIORIA, S. & SURANO, V. 2003. Personalized and adaptive services on board a car: an application for tourist information. *Journal of Intelligent Information Systems,* 21**,** 249-284.

ELHASSAN, A., JENHANI, I., BRAHIM, G. B. J. I. J. O. M. L. & COMPUTING 2018. Remedial actions recommendation via multi-label classification: a course learning improvement method. 8**,** 583-588.

FESENMAIER, D. R., RICCI, F., SCHAUMLECHNER, E., WÖBER, K. & ZANELLA, C. DIETORECS: Travel advisory for multiple decision styles. ENTER, 2003. 232-241.

GANDA, D. & BUCH, R. J. R. T. I. P. L. 2018. A Survey on Multi Label Classification. 5**,** 19-23.

GAO, M., LIU, K. & WU, Z. J. I. S. F. 2010. Personalisation in web computing and informatics: Theories, techniques, applications, and future research. 12**,** 607-629.

GARCÍA-CRESPO, A., CHAMIZO, J., RIVERA, I., MENCKE, M., COLOMO-PALACIOS, R. & GÓMEZ-BERBÍS, J. M. 2009a. SPETA: Social pervasive e-Tourism advisor. *Telematics and informatics,* 26**,** 306-315.

GARCÍA-CRESPO, A., CHAMIZO, J., RIVERA, I., MENCKE, M., COLOMO-PALACIOS, R., GÓMEZ-BERBÍS, J. M. J. T. & INFORMATICS 2009b. SPETA: Social pervasive e-Tourism advisor. 26**,** 306-315.

GARCÍA-SÁNCHEZ, F., MARTÍNEZ-BÉJAR, R., CONTRERAS, L., FERNÁNDEZ-BREIS, J. T. & CASTELLANOS-NIEVES, D. J. E. S. W. A. 2006. An ontology-based intelligent system for recruitment. 31**,** 248-263.

GARCIA, I., SEBASTIA, L. & ONAINDIA, E. J. E. S. W. A. 2011. On the design of individual and group recommender systems for tourism. 38**,** 7683-7692.

GEURTS, P., ERNST, D. & WEHENKEL, L. J. M. L. 2006. Extremely randomized trees. 63**,** 3-42.

GRÜN, C., NEIDHARDT, J. & WERTHNER, H. 2017. Ontology-based matchmaking to provide personalized recommendations for tourists. *Information and Communication Technologies in Tourism 2017.* Springer.

HEPP, M., SIORPAES, K. & BACHLECHNER, D. 2006. Towards the semantic web in e-tourism: can annotation do the trick?

HSU, F.-M., LIN, Y.-T. & HO, T.-K. J. E. S. W. A. 2012. Design and implementation of an intelligent recommendation system for tourist attractions: The integration of EBM model, Bayesian network and Google Maps. 39**,** 3257-3264.

HUANG, S., PENG, W., LI, J. & LEE, D. Sentiment and topic analysis on social media: a multi-task multi-label classification approach. Proceedings of the 5th annual ACM web science conference, 2013. 172-181.

HUANG, Y. & BIAN, L. J. E. S. W. A. 2009. A Bayesian network and analytic hierarchy process based personalized recommendations for tourist attractions over the Internet. 36**,** 933-943.

HUG, N. 2020. Surprise: A Python library for recommender systems. *Journal of Open Source Software,* 5**,** 2174.

KOPALLE, P. K., LEHMANN, D. R. J. O. B. & PROCESSES, H. D. 1997. Alpha inflation? The impact of eliminating scale items on Cronbach's alpha. 70**,** 189-197.

KOREN, Y. 2010. Factor in the neighbors: Scalable and accurate collaborative filtering. *ACM Transactions on Knowledge Discovery from Data (TKDD),* 4**,** 1-24.

LEE, C.-S., CHANG, Y.-C. & WANG, M.-H. J. E. S. W. A. 2009. Ontological recommendation multi-agent for Tainan City travel. 36**,** 6740-6753.

LEE, C.-S., KAO, Y.-F., KUO, Y.-H., WANG, M.-H. J. D. & ENGINEERING, K. 2007. Automated ontology construction for unstructured text documents. 60**,** 547-566.

LIU, S. M. & CHEN, J.-H. J. E. S. W. A. 2015. A multi-label classification based approach for sentiment classification. 42**,** 1083-1093.

LOH, S., LORENZI, F., SALDAÑA, R., LICTHNOW, D. J. I. T. & TOURISM 2003. A tourism recommender system based on collaboration and text analysis. 6**,** 157-165.

LU, J., WU, D., MAO, M., WANG, W. & ZHANG, G. J. D. S. S. 2015. Recommender system application developments: a survey. 74**,** 12-32.

LUCAS, J. P., LUZ, N., MORENO, M. N., ANACLETO, R., FIGUEIREDO, A. A. & MARTINS, C. J. E. S. W. A. 2013. A hybrid recommendation approach for a tourism system. 40**,** 3532-3550.

MAILLOT, N. E., THONNAT, M. J. I. & COMPUTING, V. 2008. Ontology based complex object recognition. 26**,** 102-113.

MCDONALD, R. P. 1985. *Factor analysis and related methods,* Psychology Press.

MONTEJO-RÁEZ, A., PEREA-ORTEGA, J. M., GARCÍA-CUMBRERAS, M. Á. & MARTÍNEZ-SANTIAGO, F. J. E. S. W. A. 2011. Otiŭm: A web based planner for tourism and leisure. 38**,** 10085-10093.

MORENO, A., VALLS, A., ISERN, D., MARIN, L. & BORRÀS, J. J. E. A. O. A. I. 2013. Sigtur/e-destination: ontology-based personalized recommendation of tourism and leisure activities. 26**,** 633-651.

NEIDHARDT, J., SEYFANG, L., SCHUSTER, R., WERTHNER, H. J. I. T. & TOURISM 2015. A picture-based approach to recommender systems. 15**,** 49-69.

OU, S., PEKAR, V., ORASAN, C., SPURK, C. & NEGRI, M. Development and Alignment of a Domain-Specific Ontology for Question Answering. LREC, 2008.

PAL, M. J. I. J. O. R. S. 2005. Random forest classifier for remote sensing classification. 26**,** 217-222.

PASHTAN, A., BLATTLER, R., ANDI, A. H. & SCHEUERMANN, P. 2003. CATIS: a context-aware tourist information system.

PESTIAN, J., BREW, C., MATYKIEWICZ, P., HOVERMALE, D. J., JOHNSON, N., COHEN, K. B. & DUCH, W. A shared task involving multi-label classification of clinical free text. Biological, translational, and clinical language processing, 2007. 97-104.

PRANTNER, K., DING, Y., LUGER, M., YAN, Z. & HERZOG, C. 2007. Tourism ontology and semantic management system: state-of-the-arts analysis.

RICCI, F., ARSLAN, B., MIRZADEH, N. & VENTURINI, A. ITR: a case-based travel advisory system. European Conference on Case-Based Reasoning, 2002. Springer, 613-627.

RICCI, F., ROKACH, L. & SHAPIRA, B. 2015. Recommender systems: introduction and challenges. *Recommender systems handbook.* Springer.

RIVOLLI, A., PARKER, L. C. & DE CARVALHO, A. C. Food truck recommendation using multi-label classification. EPIA Conference on Artificial Intelligence, 2017. Springer, 585-596.

ROKACH, L., SCHCLAR, A. & ITACH, E. J. E. S. W. A. 2014. Ensemble methods for multi-label classification. 41**,** 7507-7523.

RUMMEL, R. 1970. Applied factor analysis, Evanston, IL: Northwestern Univer. Press.

RUOTSALO, T., HAAV, K., STOYANOV, A., ROCHE, S., FANI, E., DELIAI, R., MÄKELÄ, E., KAUPPINEN, T. & HYVÖNEN, E. J. J. O. W. S. 2013. SMARTMUSEUM: A mobile recommender system for the Web of Data. 20**,** 50-67.

SANDEN, C. & ZHANG, J. Z. Enhancing multi-label music genre classification through ensemble techniques. Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval, 2011. 705-714.

SCHIAFFINO, S. & AMANDI, A. J. E. S. W. A. 2009. Building an expert travel agent as a software agent. 36**,** 1291-1299.

SEBASTIA, L., GARCIA, I., ONAINDIA, E. & GUZMAN, C. J. I. J. O. A. I. T. 2009. e-Tourism: a tourist recommendation and planning application. 18**,** 717-738.

SOROWER, M. S. J. O. S. U., CORVALLIS 2010. A literature survey on algorithms for multi-label learning. 18**,** 1-25.

SPYROMITROS, E., TSOUMAKAS, G. & VLAHAVAS, I. An empirical study of lazy multilabel classification algorithms. Hellenic conference on artificial intelligence, 2008. Springer, 401-406.

THIENGBURANATHUM, P. 2018. *An intelligent destination recommendation system for tourists.* Bournemouth University.

THO, Q. T., HUI, S. C., FONG, A. C. M., CAO, T. H. J. I. T. O. K. & ENGINEERING, D. 2006. Automatic fuzzy ontology generation for semantic web. 18**,** 842-856.

TIDAKE, V., SANE, S. J. I. J. O. E. & TECHNOLOGY 2018. Multi-label Classification: a survey. 7.

TSOUMAKAS, G., KATAKIS, I. & VLAHAVAS, I. 2009. Mining multi-label data. *Data mining and knowledge discovery handbook.* Springer.

TSOUMAKAS, G., KATAKIS, I. J. I. J. O. D. W. & MINING 2007. Multi-label classification: An overview. 3**,** 1-13.

TUNG, H.-W. & SOO, V.-W. A personalized restaurant recommender agent for mobile e-service. IEEE International Conference on e-Technology, e-Commerce and e-Service, 2004. EEE'04. 2004, 2004. IEEE, 259-262.

VANSTEENWEGEN, P., SOUFFRIAU, W., BERGHE, G. V. & VAN OUDHEUSDEN, D. J. E. S. W. A. 2011. The city trip planner: an expert system for tourists. 38**,** 6540-6546.

VENTURINI, A., RICCI, F. J. F. I. A. I. & APPLICATIONS 2006. Applying Trip@ dvice Recommendation Technology to www. visiteurope. com. 141**,** 607.

WANG, J., YANG, Y., MAO, J., HUANG, Z., HUANG, C. & XU, W. Cnn-rnn: A unified framework for multi-label image classification. Proceedings of the IEEE conference on computer vision and pattern recognition, 2016. 2285-2294.

WEN, Z. 2008. Recommendation system based on collaborative filtering. *CS229 Lecture Notes*.

YANG, W.-S., HWANG, S.-Y. J. J. O. S. & SOFTWARE 2013. iTravel: A recommender system in mobile peer-to-peer environment. 86**,** 12-20.

YONG, A. G. & PEARCE, S. J. T. I. Q. M. F. P. 2013. A beginner's guide to factor analysis: Focusing on exploratory factor analysis. 9**,** 79-94.

ZHANG, J., ZHANG, Z., WANG, Z., LIU, Y. & DENG, L. J. B. 2018. Ontological function annotation of long non-coding RNAs through hierarchical multi-label classification. 34**,** 1750-1757.

ZHENG, Y., MOBASHER, B. & BURKE, R. Context recommendation using multi-label classification. 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2014. IEEE, 288-295.

ZINS, A. H. J. I. T. & TOURISM 2007. Exploring travel information search behavior beyond common frontiers. 9**,** 149-164.