A Data-Driven Ontology Building Method for Tourism Recommender Systems

Rasoul Norouzi[a], Hamed Baziyad[a], Elham Akhondzadeh[b], Amir Albadvi[c,*]

[a] *M.Sc., Department of Information Technology, Faculty of Industrial and Systems Engineering, Tarbiat Modares University, Tehran, Iran*
[b] *Associate Professor, Department of Information Technology, Faculty of Industrial and Systems Engineering, Tarbiat Modares University, Tehran, Iran*
[c] *Professor, Department of Information Technology, Faculty of Industrial and Systems Engineering, Tarbiat Modares University, Tehran, Iran*

Abstract

The effect of information technology on tourism has been twofold. On the one hand, it has increased information and users' access to tourism services. Alternatively, it has caused information overload, which makes the decision-making process difficult. To this extend, personalization tools such as recommender systems (RS) can capture the user's needs and recommend his/her interested items. Hybrid RSs with ontology are important because, in addition to increasing RSs performance's accuracy, they can also solve common RSs problems such as cold-start and sparsity. This study has presented a data-driven method for building a three-layer ontology for tourism RSs. Zero layer refers to the domain name, layer one is tourism categories, obtained by factor analysis of the second layer data, and the third layer is the concept layer or tourism activities. The relationship between category and concept layers' elements is formulated with a multi-label classifier (MLC) algorithm. The users' interest in 18 known tourism activity index by scoring only seven categories. In comparison between MLC algorithms, binary relevance in joint Naive Bayes Classifier performed best in our method. In another scenario, this method in comparison with two collaborative filtering algorithms outperformed slightly better in all evaluated metrics. Besides, our method, like CF systems, does not need users' previous rating records for profile indexing, and it's done by reducing the level of users' interaction with 18 activities to 7 tourism categories.

Keywords: Ontology, Multi-Label Classification, Recommender System, Tourism, Data-Driven

*Corresponding author
Email address: albadvi@modares.ac.ir (Amir Albadvi)

1. Introduction

A typical trip plan consists of several steps, from selecting tourist attractions, choosing accommodations, 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 and Hwang, 2013). Besides, 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 concerning their accessibility and opening hours (Vansteenwegen et al., 2011). Over the last decade, the rapid 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 have realized 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 increases at a tremendous rate, and users are usually confronted with too many options to choose from (Gao et al., 2010). In 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 beneficial 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). Besides, around one-third of these users cannot explicitly express their travel needs and expectations (Zins, 2007). Hence, they require help for exploring and filtering out irrelevant information based on their specific preferences and needs identified as personalization.

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 mainly related to recommender systems (RS), which aim to filter irrelevant information and to provide personalized information to each particular user (Borràs et al., 2014). RS can be defined as a personalization tool that provides people with a list of items that best fit their individual preferences, restrictions, or tastes (Sebastia et al., 2009). The most crucial feature of RS is that it can guess a user's preferences and interests by analyzing the behavior of this user and/or other users' behavior to generate personalized recommendations (Lu et al., 2015). The RSs can automatically learn the user's preferences by analyzing their explicit or implicit feedback. Explicit data might be given by the user in different ways, requiring the user to fill a questionnaire about his or her preferences and interests. The system can infer implicit interests through the analysis of the user's behavior (Borràs et al., 2014). The main aim of developing RSs is to counteract the risk of information overload by assisting users searching for relevant information from a vast amount and variety of information (Grün et al., 2017). In recent years, RSs 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. Various RS techniques in tourism have been applied standalone or in combination with heuristic algorithms, machine learning, and context-aware methods (Batet et al., 2012; García-Crespo et al., 2009; 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 and Ricci, 2006; Montejo-Ráez et al., 2011; Rivolli et al., 2017; Sebastia et al., 2009; Yang and Hwang, 2013).

Commonly used recommendation techniques are knowledge based (KB), content based (CB), collaborative filtering (CF) techniques, or Hybrid RS that tries to use a combination of different methods to overcome the weaknesses of each of these methods (Lu et al., 2015). KB systems are based on understanding users' and items' features and their underlying relationship. Such systems recommend items that satisfy users' needs based on users' and items' features (Borràs et al., 2014). 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 analysis process's output 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 matching 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 (Lu et al., 2015). CF-based 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 list items. When the system finds out which people are closer to each other based on interests and choices, other users' favorites 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). Each of RSs has some drawbacks that prevent them from being suitable for all scenarios and conditions. Two of the most important problems, especially in CF based system, are sparsity and cold-start. Sparsity means many user-item matrix cells suffer from the lack of user giving rates, which makes it challenging to train machine learning models, especially in memory-based algorithms. By growing the number of items and users, the sparsity and the dimension of the user-item problems will be severe and more problematic.The cold-start problem refers to entering a new user into the system; since there is no record of the user interests and rates to the items, it is impossible to predict the user's interests. The same problem can exist for newly added items to the system, which in literature, it refers to items' cold-start (Thiengburanathum, 2018). In such situations, hybrid systems and expanding user/item profile attributes are employed, especially for dealing with the cold start problem. One approach is deploying ontology-based RSs to solve the cold start issue when the user or item's historical data is not available in the system (Kuznetsov et al., 2016).

Ontology conceptualizes a domain into a human-understandable, machine readable format consisting of entities, attributes, relationships, and axioms (Tho et al., 2006). One of the approaches to using ontology in the cold-start problem is to employ it alongside other methods as a standalone module. When a new user enters into the system, static information such as demographic information or even the user profile on social media is sent to the ontology module for finding and recommending matched items. In other cases, when there is historical data about user's interaction with items, the system uses conventional methods like CF (Meng et al., 2013; Dascalu et al., 2016). In other applications, ontologies can be used to reduce sparsity. For example, instead of measuring the distance between the user profile and the items' attributes, measure the distance between the user profile and the ontology top layers' elements. The point is that the number of elements in ontologies' top layers is much lower than the items' numbers, which can mitigate the sparsity problem (Kuznetsov et al., 2016). In addition to managing new items or users, ontologies in more recent approaches are used in combination with other methods and algorithms to increase the recommendations' accuracy and performance (Duzen and Aktas, 2016; Bagherifard et al., 2017; Bahrani et al., 2020). We have two options to use ontology in RS, using prebuilt ontologies or design a new one according to our domain's requirements. Prebuilt ontologies may not fully meet our requirements; for example, many elements of a city's tourism-ontology may not be available in another city. During this study, We found that there are many entertainments such as nightclubs in a country like Spain, which is not available in a country like Iran. Therefore, in such cases, the ontology must be adapted or redesigned according to the project requirements. Designing a practical ontology is challenging and time-consuming. Experts in that particular domain need to design layers and their connections with sub-layers. Also, the relationship between layers' elements is static and does not change; thus, it can consider only limited states and outputs (Kuznetsov et al., 2016).

In this regard, this study has provided a Method for designing a data-driven ontology in tourism, which consists of three layers. The zero layer or domain layer refers to tourism, the next layer is the category layer, and the second layer is the concept's layer or tourism's activities. From a data-driven viewpoint, there are two main steps in building this ontology: the first step is to build a categorical layer based on the concept layer elements (tourism activities). The second step is to connect the categorical layer elements and the concept layer dynamically for managing various states. To establish the relationship between the two-layer elements, we used multi-label classification (MLC) algorithms.

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). In this ontology, the user interacts only with the category layer elements. By rating each of the category elements, his/her interest in tourism activities in the concept layer is predicted. Therefore, in another view, it can be considered as a way to reduce the amount of explicit information received from the user. In our proposed Method, the user does not need to interact with all concept layer elements; only by rating the categories in the second layer of the ontology his/her profile will be indexed. The Method is evaluated with three metrics; results show that it can appropriately capture and predict users' interests. To evaluate the Method's ability to reduce explicit information received from the user, we transformed this problem into a CF problem and compared the Method results with two state-of-the-art algorithms in CF. The proposed Method results slightly outperform CF algorithms; this shows that it can properly reduce the amount of explicit information received from the user in addition to its slight superiority in the metric results.

The rest of the paper organized as follows: The literature section provides an overview of ontology in tourism and multi-label classifiers and their applications in RS; 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 result section, we review and analyze the method's ability in user interests capturing and predicting outcomes. Finally, in the conclusion section, we review our method and discuss this paper's achievements, research limitations, and future research suggestions.

2. Literature

This section will first define the ontology in more detail and then examine the ontology applications in tourism personalization. We will then introduce the MLC algorithms used in this study and briefly review their application in RSs.

2.1. Ontology

The information available on the web is often published by various travel information providers with vast backgrounds. Numerous providers might use different terms 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, due to this heterogeneity of the information, it is challenging to automatically integrate the information. 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 high-level communication. Much research has been conducted in applications of ontology in many research fields. Maillot and Thonnat (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, García-Sánchez et al. (2006) suggested an ontology-based recruitment system to provide intelligent matching between employer advertisements and the applicants' curriculum vitae. 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 generic aspects and 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: accommodation, contexts, and infrastructures. Then, Prantner et al. (2007) developed a tourism ontology, entitled Manteca, which included essential concepts of the tourism domain 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 it 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. Harmonize was one of the first ontologies that aimed to overcome tourism's interoperability problems, 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 associates 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 $|L| > 2$). However, 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 and Katakis, 2007). Problem transformation methods are divided into three main algorithms: 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 and Katakis, 2007; Spyromitros et al., 2008). Besides these approaches, Ensemble learning algorithms can learn from multi-label 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 algorithm candidates.

MLC has many applications in various 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 and Sane (2018) and (2009) Tsoumakas et al. (2009) papers. Moreover, MLC has leveraged its power into RSs world too. Carrillo et al. (2013) demonstrated the MLC ability to recommend and deal 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 MLC algorithms are more capable of recommending and predicting than base algorithms. Rivolli et al. (2017) used the MLC algorithms 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 to provide remedial actions to address students' shortcomings in Learning Outcome Attainment Rates. In their model, each instance is a student 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 the chain classification method with the decision tree algorithm gives the best results for the given dataset.

3. Research Methodology

This study attempts to provide a data-driven Method for building an ontology in tourism, which user by rating to category layer elements, his/her interest to the following layer elements (tourism activities) will be predicted. The city of Tehran in Iran has been selected for the case study, so the presented ontology is based on the tourism sites and activities in this city. 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 diverse learners' algorithms on the second dataset. All of the proposed Method's stages are represented graphically in Fig.1.

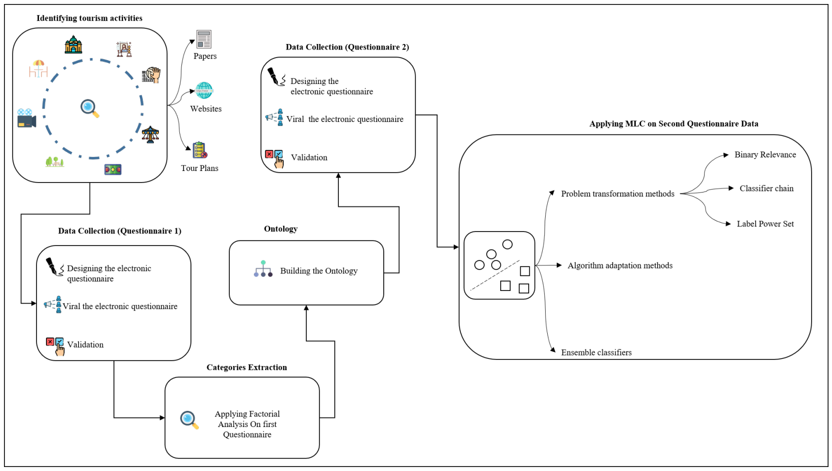

Figure 1: The stages' flow of the proposed Method

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[1], and the famous website Trip Advisor[2]. 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; Therefore, 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 can be difficult and tedious for the user. Thus, reducing these 18 activities to few and more interpretable categories makes the user more comfortable recognizing and categorizing content. We need data to build the ontology's category layer; therefore,

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

a questionnaire was 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 slightest interest in an activity, and the five score the most. To better guide users, we introduced several POIs 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 distributed randomly on social media like Facebook, Twitter, and messenger applications. Totally 272 questionnaires were collected, and the reliability of the designed questionnaire was proved by calculating Cronbach's alpha equal to 0.846, which was more than the cutoff required of 0.7 (Kopalle and Lehmann, 1997).

3.3. Extracting Tourism categories (Factorial Analysis of questionnaire 1)

Factorial Analysis (FA) 's primary goal is summarizing data for revealing relationships and patterns with regrouping variables into a limited collection of clusters based on shared variance. Indeed, FA utilizes mathematical methods to simply interrelated measures for discovering patterns in a set of variables (Child, 1990). FA was applied in various types of fields such as behavioral and social sciences, medicine, economics, and geography (Yong et al., 2013) and is divided into two main classes, namely Exploratory Factorial Analysis (EFA) and Confirmatory Factorial Analysis (CFA). EFA is used when the research goal is discovering the number of influencing variables or finding variables that go together (Child, 1990). 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 considering too many keys is easier and makes variables meaningful by clustering them into some clusters. In this paper, EFA was applied for accessing meaningful categories of variables. Our data's determinant score 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 et al., 2013). We used the Varimax rotation method with 30 iteration bases on the default value in SPSS software for rotation. To check the adequacy and suitability of the dataset for EFA, Kaiser-Meyer-Olkin Measure (KMO) and Bartlett's Test of Sphericity were 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 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). 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 Fig.2 (Scree Plot), the best number of factors after rotation for this dataset is 7. 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 1: 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

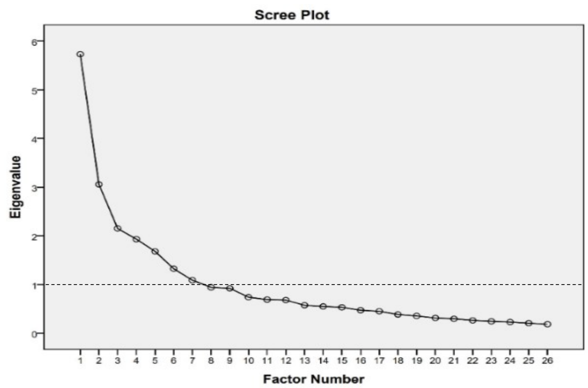

Figure 2: Scree Plot for determining the optimum number of factors, factors below eigenvalue one are dropped out

Table 2: Extracted factors name and their associates' variables

Factor	**Descriptor Variables**
Historical	Museums, Historical Sites
Fun	Restaurants, Cafe, Male
Ecotourist	Rivers and Waterfalls, Lakes, Mountains, Rural Places
Sportive	Sport Activities, Sport Events
Cultural	Music Events, Art and Book Exhibitions, Cinema, Theaters
Religious	Mosques and Churches, Holy Sites
Urban-related	Zoo, Public Gardens, Parks

3.4. Ontology

In this research, the ontology has three layers: zero-layer representing the domain name which is tourism, the first layer or category layer, representing the extracted categories, and the second layer or the concept layer is tourism activities. In this ontology, the relationship between the category layer and the concept layer is unknown; as mentioned, this study aims to provide a method that can map the category layer to the concept layer with a data-driven approach (Fig.3). For this purpose, MLC algorithms have been used, which in the next section, we will formulate the problem.

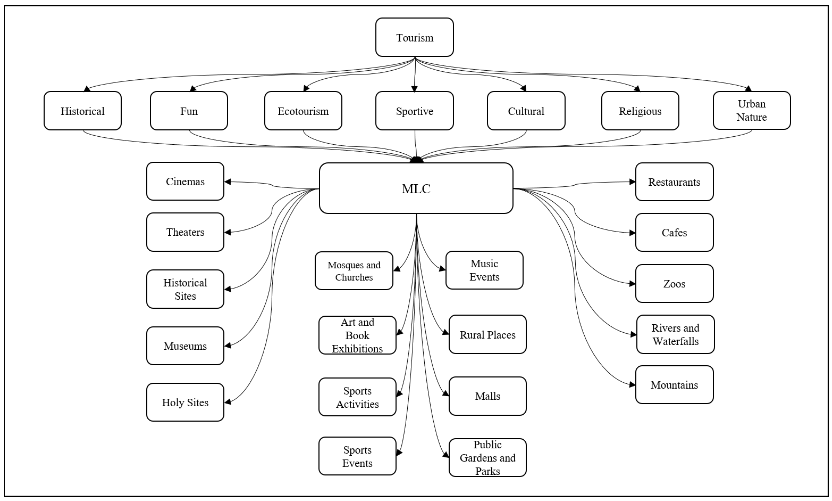

Figure 3: The proposed three-layer ontology

3.5. Data Collection (Questionnaire 2)

For building a data-driven connection between the category layer and the concept layer elements, data is required to train MLC algorithms. The advantage of this connection is that different states can be considered, and the user's interest in tourism activities can be predicted only by rating the seven categories extracted by factorial analysis. So, a dichotomous questionnaire was designed. The first part asked people to determine their interest rate for every category on a five-point scale Likert. The second part requested to fulfill every 18 items with binary values (0 means interested, and one means not interested regarded item). This electronic questionnaire was randomly broadcast through social media and messengers' applications for Tehran residents. 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 $\mathcal{L} = \lambda_1, \lambda_2, \ldots, \lambda_k$ be a finite set of labels or activities. A user $x \in \chi$, represented in terms of features vector

$x = (x_1, x_2, \ldots, x_m)$, 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 $Y = (y_1, y_2, \ldots, y_k)$, where $y_i = 1 \Leftrightarrow \lambda_i \in L . \mathcal{Y} = 0, 1^k$ is the set of all such possible labeling (Fig.4).

Features vector or given rates to each category *Associated subset of labels*

	$category_1$	$category_2$	...	$category_m$	$activity_1$	$activity_2$	...	$activity_k$
$user_1$	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,$(x_i \in \mathcal{X}, Y_i \in \mathcal{Y}) i.i.d^3$ drawn from an unknown distribution D, the goal of the multi-label learning is to produce a multi-label classifier $h : \mathcal{X} \rightarrow \mathcal{Y}(inotherwords, h : \mathcal{X} \rightarrow 2^{\mathcal{L}})$ that optimizes some specific evaluation function (i.e. loss function) (Sorower, 2010).

This study uses the second questionnaire data 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 in 18 activities will be predicted. In the transformation approach, all three algorithms (BR, LP, CC) with LR, DT, RF, SVM, KN classifiers are used. In the adaptation algorithm, BRKNN, and MLKNN and ensembles algorithm, RF and ET classifiers were used. To implement MLC algorithms, Python version 3.5 with scikit-learn and scikit-multilearn packages are used. In this study, all the classifiers' Hyperparameters are the package's default values.
An issue that should be considered is imbalanced labels' problems; as shown in Fig.5, the number of classes is not equal in any of the labels, and this could cause problems in some algorithms' learning process. To solve this problem, we used the Scikit Learn Package class weight balancing feature in all classifiers except the NB, MLKNN, and BRKNN; because the technique does not apply to such classifiers.

[3] independent and identically distributed

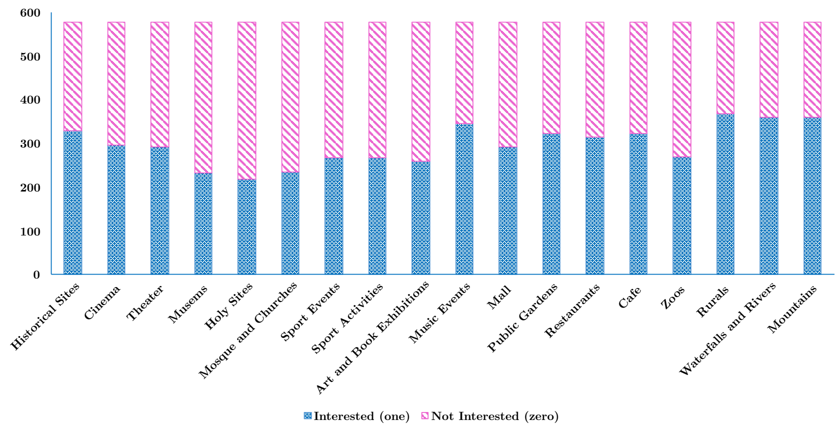

Figure 5: Labels imbalance, vertical pivot represents the number of samples, and the horizontal pivot shows the label names

3.7. Evaluation

In this research, the evaluation stage is vital in two ways; first, to measure and compare the performance of different MLC algorithms and approaches in capturing and predicting users' interests. The second evaluation measures the proposed ontology's ability to reduce explicit information received from the user and compare it with state-of-the-art FC algorithms. 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 to compare the proposed framework with state-of-the-art CF techniques. Thus, in this study, we select Precision, Recall, and F1-Score metrics for evaluation. Since the weight-balancing technique is not applicable for some classifiers (NB, MLKNN, BRKNN), we use the macro-average criteria that do not take label imbalance into account.

3.7.1. Metrics

Let T be a multi-label dataset consisting n multi-label examples $(x_i, Y_i), 1 \leq i \leq n, \left(xi \in \mathcal{X}, Yi \in \ \mathcal{Y} = \{0,1\})\ ^{k}\right)$, 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):

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

3.7.2. CF System vs. Proposed Method

In this part, the presented framework should be compared with a situation that the user interacted directly with some activities and his/her other interests were predicted by CF algorithms. The problem should be transformed into a CF problem to evaluate the CF's prediction ability with our proposed framework. CF systems generally have two approaches for the recommendation, memory-based and model-based. There are no assumptions on data in memory-based 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, the variance is low in the item-item method, and its bias is high, which could cause less personalization. On the other side, model-based systems try to assume data and learn a model that explains user-item matrix interaction (Fig.6).

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

$$\begin{cases} u_{ri} & \text{if such rating exist} \\ ? & \text{otherwise} \end{cases}$$

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

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

Figure 6: An example of user-item matrix

user interests on a test set (which is, of course, unknown during the training phase), that is to minimize:

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

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 languag (Hug, 2020); we also used the benchmark results to select the top two CF algorithms [4].

In a 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 the problem:

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

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).

[4] http://surpriselib.com/

In the model-based approach, we used Singular Value Decomposition (SVD), a Matrix factorization algorithm that 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 the mentioned algorithm is in the 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:

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

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. Result and Discussions

In this section, there are two parts of evaluation; in the first part, we will review MLC algorithms' results in the proposed ontology. In the next part, we will compare the results of the best MLC algorithm with the CF problem-solving approach.

4.1. MLC Results

The performance of MLC algorithms is shown in Fig.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 a classifier; also for Ensemble methods, we use "Ensemble-Classifier" form. The experimental results show that the best performance is related to BR algorithms, and on the other hand, the LP algorithm has lower metrics values than others. The closeness of the precision score and the Recall in most classifiers indicates that the class imbalance has not affected the learning process. Since we have 18 tourism activities, BR algorithms transform the problem into 18 separate problems, nor

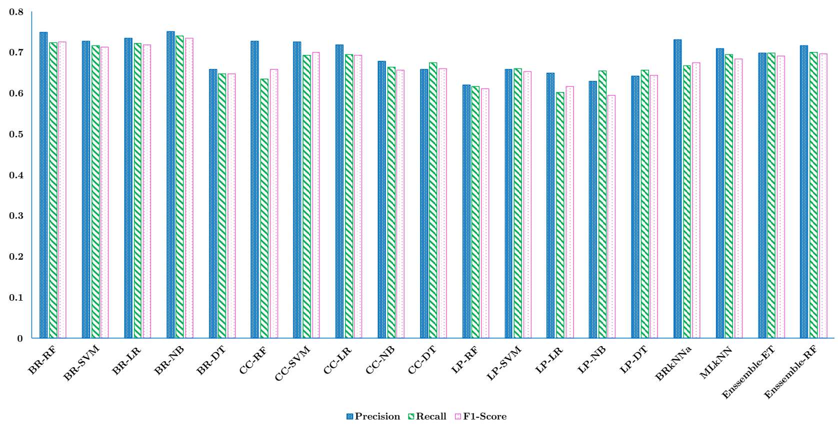

Figure 7: Comparison of the MLC algorithms performance

does it consider the interdependence of labels. This algorithm's good performance indicates that the detected activities are distinctive from each other, and the algorithm has been able to map the category layer to the concept layer space well. For example, the RF classifier with the BR algorithm had better results than other algorithms with the RF classifier. Nevertheless, the LP algorithms' disappointing results in our problem are 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, it is difficult to train such algorithms on this data set; however, algorithms' results and superiority may change as the number of data increases. Recall's high score refers to the classifier's success rate 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. On the other hand, Precision indicates what percentages of the activity recommended to the user were 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 the two mentioned metrics, can be a good criterion for comparing classifiers' performance. 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 are the ensemble algorithms.

4.2. BR-NB VS BSKNN and SVD

The BSKNN and SVD algorithms' output is in the range of zero to one, so a threshold was used to convert it to a binary output. Fig.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 Fig.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 both. Therefore, an α value of 0.5 was chosen for both algorithms. To compare the

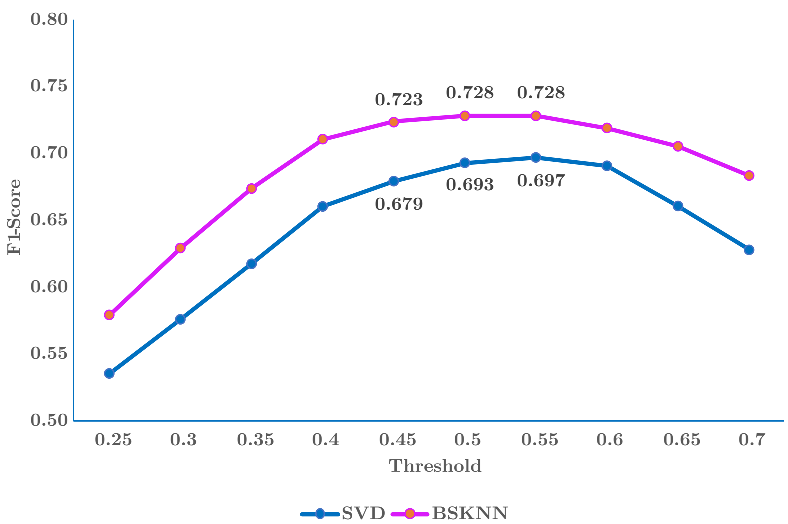

Figure 8: BSKNN and SVD F1-Scores for different α values

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. Fig.9 Shows Precision, Recall, and F1 scores for each of them. The weakest performance is related to the SVD algorithm because of such algorithms' learning process, which usually uses gradients decent to estimate their parameters, while this technique requires many data. The performance of BR-NB in the proposed method is superior to the BSKNN algorithm by nearly three percent in Precision, nearly two percent in Recall, and nearly one percent in F1-score. The proposed method's performance was slightly better than the CF algorithms. Still, it should note that in the CF systems, we only needed one step of data collection and did not need to extract tourism categories. Nevertheless, Our method's advantage is that the user interacts with only seven categories; in other words, we have reduced the problem's dimensions.

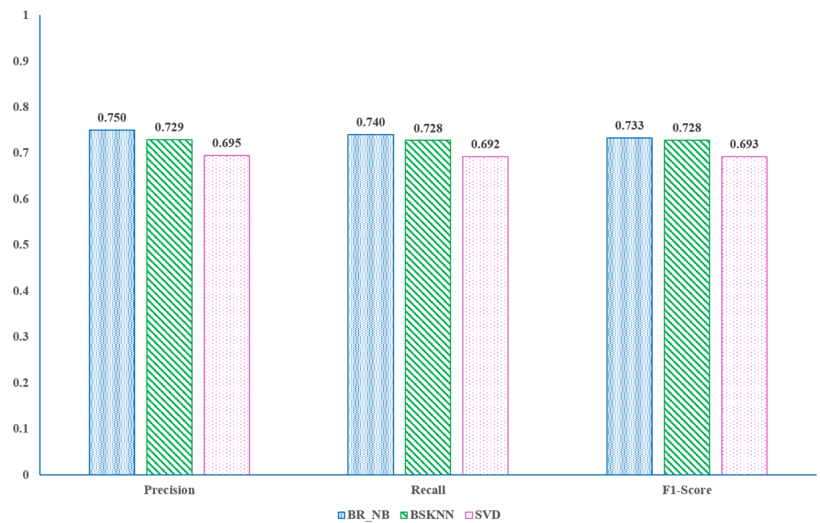

Figure 9: Comparison of BR-NB results in the proposed method in with two state of the art algorithms, BSKNN and SVD, in CF systems

Another critical point is that we also evaluated CF approach algorithms with the 5-fold cross-validation method. The total number of activities is 18; in each step of cross-validation 4-fold (for each user, nearly 15 given rate records) is considered as training data and one fold as test data(for each user, about three given rate records). This means that the CF algorithms' results are based on having 15 records of user interaction with tourism activities and only predicting user interests to nearly three activities. However, in the proposed method, the user interests in all 18 tourism activities (second layer's elements in the ontology) are predicted by rating $[1, 5]$ scale (explicit data) to each of the seven tourism categories without having any interaction records.

5. Conclusion and Future Research

This study proposed a data-driven method for building an ontology for tourism RSs that consists of three layers. The important point about this ontology is that the relationship between the category layer and the concept layer is data-driven, established by an MLC algorithm. In addition to being able to consider different states, this data-driven connection can reduce the user's interaction with the system for profile indexing. The first layer of the ontology was obtained by applying FA on the first questionnaire that measured users' interests in Identified tourism activities (the second layer elements in the proposed ontology architecture). We formulated the layers' connections into an MLC problem, and the BR-NB results performed best compared to other classifiers. According to the literature review section and the performance outcome of different MLC algorithms in this study, they can have different performance given to the problem and data; in other words, there is not any definite superiority. 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 outperformed the two mentioned algorithms, although this was a slight advantage. In addition to slight superiority in metrics, reducing the problem space from 18 activities to seven tourism categories is another advantage; means that the user does not need to interact with all 18 activities or have rating records in the CF method for the profile indexing.

One of our system's limitations is that each layer's elements are fixed; for future study, we suggest finding a solution for this problem. This study focused on building a data-driven ontology for tourism recommender, and designing RS was not our aim. Future research will combine this data-driven ontology with a context-aware approach to recommend items related to each activity according to the user location.

References

Abbaspour, R.A., Samadzadegan, F., 2011. Time-dependent personal tour planning and scheduling in metropolises. Expert Systems with Applications 38, 12439–12452.

Bagherifard, K., Rahmani, M., Nilashi, M., Rafe, V., 2017. Performance improvement for recommender systems using ontology. Telematics and Informatics 34, 1772–1792.

Bahrani, P., Minaei-Bidgoli, B., Parvin, H., Mirzarezaee, M., Keshavarz, A., Alinejad-Rokny, H., 2020. User and item profile expansion for dealing with cold start problem. Journal of Intelligent & Fuzzy Systems , 1–13.

Barta, R., Feilmayr, C., Pröll, B., Grün, C., Werthner, H., 2009. Covering the semantic space of tourism: an approach based on modularized ontologies, in: Proceedings of the 1st Workshop on Context, Information and Ontologies, pp. 1–8.

Batet, M., Moreno, A., Sánchez, D., Isern, D., Valls, A., 2012. Turist@: Agent-based personalised recommendation of tourist activities. Expert Systems with Applications 39, 7319–7329.

Belmonte, M.V., Pérez-de-la Cruz, J.L., Triguero, F., 2008. Ontologies and agents for a bus fleet management system. Expert Systems with Applications 34, 1351–1365.

Borràs, J., Moreno, A., Valls, A., 2014. Intelligent tourism recommender systems: A survey. Expert Systems with Applications 41, 7370–7389.

Büyüközkan, G., Ergün, B., 2011. Intelligent system applications in electronic tourism. Expert systems with applications 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 , 181–188.

Castillo, L., Armengol, E., Onaindía, E., Sebastiá, L., González-Boticario, J., Rodríguez, A., Fernández, S., Arias, J.D., Borrajo, D., 2008. Samap: An user-oriented adaptive system for planning tourist visits. Expert Systems with Applications 34, 1318–1332.

Chen, X., Vorvoreanu, M., Madhavan, K., 2014. Mining social media data for understanding students' learning experiences. IEEE Transactions on learning technologies 7, 246–259.

Chiang, H.S., Huang, T.C., 2015. User-adapted travel planning system for personalized schedule recommendation. Information Fusion 21, 3–17.

Child, D., 1990. The essentials of factor analysis. Cassell Educational.

Dascalu, M.I., Bodea, C.N., Mihailescu, M.N., Tanase, E.A., Ordoñez de Pablos, P., 2016. Educational recommender systems and their application in lifelong learning. Behaviour & information technology 35, 290–297.

Duzen, Z., Aktas, M.S., 2016. An approach to hybrid personalized recommender systems, in: 2016 International Symposium on INnovations in Intelligent SysTems and Applications (INISTA), IEEE. pp. 1–8.

Elhassan, A., Jenhani, I., Brahim, G.B., 2018. Remedial actions recommendation via multi-label classification: a course learning improvement method. International Journal of Machine Learning and Computing 8, 583–588.

Fesenmaier, D.R., Ricci, F., Schaumlechner, E., Wöber, K., Zanella, C., et al., 2003. Dietorecs: Travel advisory for multiple decision styles, in: ENTER, Citeseer. pp. 232–241.

Ganda, D., Buch, R., 2018. A survey on multi label classification. Recent Trends in Programming Languages 5, 19–23.

Gao, M., Liu, K., Wu, Z., 2010. Personalisation in web computing and informatics: Theories, techniques, applications, and future research. Information Systems Frontiers 12, 607–629.

Garcia, I., Sebastia, L., Onaindia, E., 2011. On the design of individual and group recommender systems for tourism. Expert systems with applications 38, 7683–7692.

García-Crespo, A., Chamizo, J., Rivera, I., Mencke, M., Colomo-Palacios, R., Gómez-Berbís, J.M., 2009. Speta: Social pervasive e-tourism advisor. Telematics and informatics 26, 306–315.

García-Sánchez, F., Martínez-Béjar, R., Contreras, L., Fernández-Breis, J.T., Castellanos-Nieves, D., 2006. An ontology-based intelligent system for recruitment. Expert Systems with Applications 31, 248–263.

Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. Machine learning 63, 3–42.

Grün, C., Neidhardt, J., Werthner, H., 2017. Ontology-based matchmaking to provide personalized recommendations for tourists, in: Information and Communication Technologies in Tourism 2017. Springer, pp. 3–16.

Hepp, M., Siorpaes, K., Bachlechner, D., 2006. Towards the semantic web in e-tourism: can annotation do the trick?, in: ECIS 2006 Proceedings. 15, pp. 2362–2373.

Hsu, F.M., Lin, Y.T., Ho, T.K., 2012. Design and implementation of an intelligent recommendation system for tourist attractions: The integration of ebm model, bayesian network and google maps. Expert Systems with Applications 39, 3257–3264.

Huang, S., Peng, W., Li, J., Lee, D., 2013. Sentiment and topic analysis on social media: a multi-task multi-label classification approach, in: Proceedings of the 5th annual ACM web science conference, pp. 172–181.

Huang, Y., Bian, L., 2009. A bayesian network and analytic hierarchy process based personalized recommendations for tourist attractions over the internet. Expert Systems with Applications 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., 1997. Alpha inflation? the impact of eliminating scale items on cronbach's alpha. Organizational Behavior and Human Decision Processes 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.

Kuznetsov, S., Kordík, P., Řehořek, T., Dvořák, J., Kroha, P., 2016. Reducing cold start problems in educational recommender systems, in: 2016 International Joint Conference on Neural Networks (IJCNN), IEEE. pp. 3143–3149.

Lee, C.S., Chang, Y.C., Wang, M.H., 2009. Ontological recommendation multi-agent for tainan city travel. Expert Systems with Applications 36, 6740–6753.

Lee, C.S., Kao, Y.F., Kuo, Y.H., Wang, M.H., 2007. Automated ontology construction for unstructured text documents. Data & Knowledge Engineering 60, 547–566.

Liu, S.M., Chen, J.H., 2015. A multi-label classification based approach for sentiment classification. Expert Systems with Applications 42, 1083–1093.

Loh, S., Lorenzi, F., Saldaña, R., Licthnow, D., 2003. A tourism recommender system based on collaboration and text analysis. Information Technology & Tourism 6, 157–165.

Lu, J., Wu, D., Mao, M., Wang, W., Zhang, G., 2015. Recommender system application developments: a survey. Decision Support Systems 74, 12–32.

Lucas, J.P., Luz, N., Moreno, M.N., Anacleto, R., Figueiredo, A.A., Martins, C., 2013. A hybrid recommendation approach for a tourism system. Expert systems with applications 40, 3532–3550.

Maillot, N.E., Thonnat, M., 2008. Ontology based complex object recognition. Image and Vision Computing 26, 102–113.

Meng, C., Cheng, Y., Jiechao, C., Peng, Y., 2013. A method to solve cold-start problem in recommendation system based on social network sub-community and ontology decision model, in: 3rd International Conference on Multimedia Technology (ICMT-13), Atlantis Press. pp. 159–166.

Montejo-Ráez, A., Perea-Ortega, J.M., García-Cumbreras, M.Á., Martínez-Santiago, F., 2011. Otiŭm: A web based planner for tourism and leisure. Expert Systems with Applications 38, 10085–10093.

Moreno, A., Valls, A., Isern, D., Marin, L., Borràs, J., 2013. Sigtur/e-destination: ontology-based personalized recommendation of tourism and leisure activities. Engineering applications of artificial intelligence 26, 633–651.

Neidhardt, J., Seyfang, L., Schuster, R., Werthner, H., 2015. A picture-based approach to recommender systems. Information Technology & Tourism 15, 49–69.

Ou, S., Pekar, V., Orasan, C., Spurk, C., Negri, M., 2008. Development and alignment of a domain-specific ontology for question answering, in: Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08), European Language Resources Association (ELRA), Marrakech, Morocco. pp. 221–228. URL: http://www.lrec-conf.org/proceedings/lrec2008/pdf/561_paper.pdf.

Pal, M., 2005. Random forest classifier for remote sensing classification. International journal of remote sensing 26, 217–222.

Pestian, J., Brew, C., Matykiewicz, P., Hovermale, D.J., Johnson, N., Cohen, K.B., Duch, W., 2007. A shared task involving multi-label classification of clinical free text, in: Biological, translational, and clinical language processing, pp. 97–104.

Prantner, K., Ding, Y., Luger, M., Yan, Z., Herzog, C., 2007. Tourism ontology and semantic management system: State-of-the-arts analysis. IADIS International Conference: IADIS .

Ricci, F., Arslan, B., Mirzadeh, N., Venturini, A., 2002. Itr: a case-based travel advisory system, in: European Conference on Case-Based Reasoning, Springer. pp. 613–627.

Ricci, F., Rokach, L., Shapira, B., 2015. Recommender systems: introduction and challenges, in: Recommender systems handbook. Springer, pp. 1–34.

Rivolli, A., Parker, L.C., de Carvalho, A.C., 2017. Food truck recommendation using multi-label classification, in: EPIA Conference on Artificial Intelligence, Springer. pp. 585–596.

Rokach, L., Schclar, A., Itach, E., 2014. Ensemble methods for multi-label classification. Expert Systems with Applications 41, 7507–7523.

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

Sebastia, L., Garcia, I., Onaindia, E., Guzman, C., 2009. e-tourism: a tourist recommendation and planning application. International Journal on Artificial Intelligence Tools 18, 717–738.

Sorower, M.S., 2010. A literature survey on algorithms for multi-label learning. Oregon State University, Corvallis 18, 1–25.

Spyromitros, E., Tsoumakas, G., Vlahavas, I., 2008. An empirical study of lazy multilabel classification algorithms, in: Hellenic conference on artificial intelligence, Springer. pp. 401–406.

Thiengburanathum, P., 2018. An intelligent destination recommendation system for tourists. Ph.D. thesis. Bournemouth University.

Tho, Q.T., Hui, S.C., Fong, A.C.M., Cao, T.H., 2006. Automatic fuzzy ontology generation for semantic web. IEEE transactions on knowledge and data engineering 18, 842–856.

Tidake, V.S., Sane, S.S., 2018. Multi-label classification: a survey. International Journal of Engineering and Technology 7.

Tsoumakas, G., Katakis, I., 2007. Multi-label classification: An overview. International Journal of Data Warehousing and Mining (IJDWM) 3, 1–13.

Tsoumakas, G., Katakis, I., Vlahavas, I., 2009. Mining multi-label data, in: Data mining and knowledge discovery handbook. Springer, pp. 667–685.

Vansteenwegen, P., Souffriau, W., Berghe, G.V., Van Oudheusden, D., 2011. The city trip planner: an expert system for tourists. Expert Systems with Applications 38, 6540–6546.

Venturini, A., Ricci, F., 2006. Applying trip@ dvice recommendation technology to www. visiteurope. com. Frontiers in Artificial Intelligence and Applications 141, 607.

Wang, J., Yang, Y., Mao, J., Huang, Z., Huang, C., Xu, W., 2016. Cnn-rnn: A unified framework for multi-label image classification, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2285–2294.

Wen, Z., 2008. Recommendation system based on collaborative filtering. CS229 lecture notes .

Yang, W.S., Hwang, S.Y., 2013. itravel: A recommender system in mobile peer-to-peer environment. Journal of Systems and Software 86, 12–20.

Yong, A.G., Pearce, S., et al., 2013. A beginner's guide to factor analysis: Focusing on exploratory factor analysis. Tutorials in quantitative methods for psychology 9, 79–94.

Zhang, J., Zhang, Z., Wang, Z., Liu, Y., Deng, L., 2018. Ontological function annotation of long non-coding rnas through hierarchical multi-label classification. Bioinformatics 34, 1750–1757.

Zheng, Y., Mobasher, B., Burke, R., 2014. Context recommendation using multi-label classification, in: 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT), IEEE. pp. 288–295.

Zins, A.H., 2007. Exploring travel information search behavior beyond common frontiers. Information Technology & Tourism 9, 149–164.