TOURISM RECOMMENDATION SYSTEM: EMPIRICAL INVESTIGATION

Biljana PETREVSKA
Goce Delcev University, Faculty of Tourism and Business Logistics, Stip, Macedonia
biljana.petrevska@ugd.edu.mk

Saso KOCESKI
Goce Delcev University, Faculty of Computer Science, Stip, Macedonia
saso.koceski@ugd.edu.mk

Abstract
The paper makes an attempt to justify the necessity of implementing recommendation system which will assist tourists in identification of their ideal holiday. The proposed recommendation system based on collaborative filtering notes positive impulses in the case of Macedonia. A software module is developed being capable to generate a personalized list of favorable and tailor-made items. The research outcomes indicate that the designed national tourism web portal can provide satisfactory performance and may be of high importance to all key-tourism actors in the process of identifying measures necessary for creating competitive tourism product.

Key words: Collaborative filtering; Tourism web-portal; Tourist profiling; Recommendation system.

JEL Classification: L83, C02, L86

INTRODUCTION
As one of the most dynamic world industries, tourism is constantly facing numerous challenges which affect its development. Given the fact that it influences the world economy by benefiting various sectors, it is scheduled in the top priority agenda of the national governments in order to gain positive trends from the variety of tourism impacts. In this respect, Macedonia identified tourism as a strategic priority for enhancing overall economic development (Petrevska, 2010). Moreover, tourism is perceived as a mean for generating various micro and macro-economic effects. Up-to-date, tourism performances are positive resulting with 1.7% participation in the gross domestic product, 3.1% participation of tourism employees in the total workforce and 1% net tourism inflows (Petrevska, 2011a, 2011b). Such condition indicates high potential to increase tourism effects in the economic activity.

However, attracting a bigger number of tourists is not a trouble-free process, particularly in times of ever-changing travel preferences. Despite the variety of options regarding tourist destination or attraction, tourists frequently are not capable to cope with such a huge volume of choice. So, they need advice where to go and what to see. In a tourism domain, recommendations may indicate cities to go to, places to visit, attractions to see, events to participate in, travel plans, road maps, options for hotels, air companies, etc. Such scope of work very often is not a trivial task. In this respect, recommendation systems assist tourists by facilitating personal selection and prevent them from being overwhelmed by a stream of superfluous data that are unrelated to their interest, location, and knowledge of a place. So, the way out is detected in application of recommendation systems as a promising way to differentiate a site from the competitors.

Based on the growth and spread of internet penetration and usage, the last ten years have seen an unprecedented rise in online travel - from ‘looking’ (research into travel and destination options) to booking. In this line, the internet penetration has grown from 0.4% of the global population (16 million users) in 1995 to 30% (2 billion) in 2011 (WTTC, 2011). Consequently, the numerous changes were noted, like: shorter lead-time for bookings, making last-minute decisions, tailoring own packages from a suite of options etc.


Generally, the contribution of this paper lies in the fact that it enriches the poorly-developed empirical academic work within this scientific area in Macedonia. Additionally, the empirical investigation may alarm the relevant tourism-actors in the country, that the time has changed and that the online experience has shifted from searching and consuming to creating, connecting and exchanging. Previously passive consumers and web surfers are now generating content, collaborating and commenting. So, this research proposes development of national tourism recommendation system since only if being prepared
in due time, one may struggle the unexpected challenges. The reminder of the paper is organised as follows: Section 2 provides a critical overview of theoretical and empirical literature on tourism recommendation systems. Section 3 provides explains the applied methodology within the research. The development of the suggested web-portal is presented in Section 4. Section 5 deals with the system accuracy, while the most interesting conclusions and future challenges are presented in the final Section 6.

LITERATURE REVIEW

The issue of importance and effectiveness of applying recommendation systems in tourism has attracted much interest in academia, and practitioners as well. Namely, Wang (2008) underlines the inevitable relationship between tourists and information and notes the widely-recognized fact that information and decision-making have become the foundation for the world economy. So, due to such importance of tourism industry, the recommendation systems applied in tourism have been a field of study since the very beginnings of artificial intelligence. In this respect, it is a matter of identifying a class of intelligent applications that offer recommendations to travelers, generally as a response to their queries. They mostly leverage in-built logical reasoning capability or algorithmic computational schemes to deliver their recommendation functionality. Consequently, the recommendation systems are an attempt to mathematically model and technically reproduce the process of recommendations in the real world.

Numerous researchers have put an accent on various aspects. In this respect, Mirzadeh et al. (2004), McSherry (2005) as well as Jannach (2006) elaborated the need for developing intelligent recommendation systems which can provide a list of items that fulfill as many requirements as possible. Furthermore, Ricci and Werthner (2002) and later Wallace et al. (2003) introduced a recommender system dealing with a case-based reasoning in order to help the tourist in defining a travel plan. However, as the most promising recommendation systems in the tourism domain are the knowledge-based and conversational approaches (Ricci et al., 2002; Thomson et al., 2004). Yet, another group of researchers propose some other variants of the content-based filtering and collaborative filtering like knowledge-filtering, constraint-based and casebased approaches (Kazienko and Kołodziejski, 2006; Ricci and Missier, 2004; Zanker et al., 2008). Simultaneously, a recommendation system based on a text mining techniques between a travel agent and a customer through a private Web chat may easily find an application (Loh et al., 2004).

Some recent researches are notable since bringing more sophisticated outcomes. So, Hinze et al. (2009) introduced a personalized tourist information provider as a combination of an event-based system and a location-based service applied to a mobile environment. The investigation on sources and formats of online travel reviews and recommendations as a third-party opinion in assisting travelers in their decision making during the trip planning was brought by Zhang et al. (2009). Interesting findings regarding development of a web site in order to enable Internet users to locate their own preferred travel destinations according to their landscape preferences were raised by Goossen et al. (2009). Furthermore, Vansteenwegen and Wouter (2011) elaborated the usage of the orienteering problem and its extensions to model the tourist trip planning problem as efficient solution for number of practical planning problems. It is evidently that the research area is extending from a few exceptions (Niaraki and Kim, 2009; Charou et al., 2010).

METHODOLOGY

The authors’ main objective is to propose recommendation system based on novel algorithms and methodology. Specifically, the paper makes an attempt to develop a national tourism web portal that relies on efficient and accurate personalized recommendation system. So, the travelers and tourists who intend to visit Macedonia will be assisted and supported in identification of certain relevant tourism objects by matching with their personal interests, preferences and desires.

To this purpose, a several step methodology is developed. The first introductory step is modeling the tourist types and profiling tourism objects. The tourist profile indicates the degree to which tourists identify themselves with the given types following the Yiannakis and Gibson (1992) methodology. Typically, individual tourist cannot be characterized by only one of these archetypes but has unique combination of these personalities, although to varying degrees. Thus, tourists’ generic interests are modeled in an abstract form using 12 dimensional vectors. This means that each dimension in the tourist profile vector corresponds to a certain tourist type while the value indicates how much the tourist identifies him- or herself with the corresponding type.

According to our methodology tourist profiling is considered as a two-step process which involves creating the profile and then reviewing the profile to make any necessary adjustments. The initial tourist profile for each system user is created by the user himself during the process of registration, by determining the degree of membership to each of the tourist types. Considering the fact that the human preferences change over time due to various factors, the tourists might change their behavior too. To make the system capable to cope with these changes, we
have enabled tourist profile adjustment. It is based on the ratings the tourist give for each tourist object that he visits after his journey and according to the Eq. 1.

$$\overline{U_{i,t+1}} = \overline{U_{i,t}} + R_{ik} \cdot w \cdot \overline{O_k}$$ (1)

where $\overline{U_{i,t+1}}$ represents the profile vector of the $i$-th user in the moment of time $t$ and $U_i \in U$, $U$ is the complete set of users registered to the system. $\overline{O_k}$ represents the profile vector of the $k$-th object in the set of all objects $O$ registered in the system $\overline{O_k} \in O$, $w$ is the weighting factor and $R_{ik}$ is the rating of the $k$-th tourist object given by the $i$-th user. The weighting factor in the Eq.1. is used simply to prevent significant change of the tourist profile from a single rating.

Similarly, to tourist profiles every tourism object is modeled through a vector as well. This vector describes in a quantitative way how much the object is related to the given tourist types. For example, the Memorial house of Mother Teresa dedicated to the humanitarian and Nobel Peace Prize laureate Mother Teresa and located in her hometown Skopje, might be highly relevant for sightseeing tourists but not for such kind of tourists that would like to do some risky activities.

In the developed system a manual process to link the given tourist types to appropriate tourism objects is proposed. Therefore, for each of the tourism objects, the degree of relationship to each of the tourist types is specified by domain experts.

In order to prevent information overload of the tourist and provide only relevant information, the system should recommend a subset of tourism objects according to the personal experiences individual tourist desire and those he/she prefer to avoid. This in turn might lead to an increase of the tourist’s satisfaction of experiencing a relaxed sightseeing trip.

According to this, the next step of the proposed methodology aims to match tourist profiles against the set of tourism objects on the basis of tourist types, thus producing a ranked list of objects for each given tourist and reducing the set of objects. If a tourist profile matches the characteristics of an object, this object should be recommended to the respective tourist. Therefore, the matchmaking algorithm has to examine whether they share similar structures.

The more similarities they have in common, the more contributes the tourism object to the tourist’s satisfaction and therefore should be ranked higher. To estimate the similarity degree between tourist profiles and tourism objects, the system contains a special module based on a vector-based matchmaking function, whereby a given profile and each tourism object constitute vectors and are compared in a vector space model. A common method to obtain the similarity is to measure the cosine angle between two vectors. The dimensions of the vector space model correspond to selected tourists types found in scientific tourism literature (Gibson and Yiannakis, 2002), such that each distinct tourist type (e.g., adventure or cultural type) represents one dimension in that space. The implemented matchmaking function has the following form:

$$SIM_{cos}(Ui,Oj) = \frac{\sum_{k=1}^{N} U_{ik} \cdot O_{jk}}{\sqrt{\sum_{k=1}^{N} U_{ik}^2 \cdot \sum_{k=1}^{N} O_{jk}^2}}$$ (2)

where $U_{ik}$ is the degree of membership of the $i$-th user to the tourist type $T_i$, $O_{jk}$ is the degree of membership of the $j$-th tourist object to the tourist type $T_j$, and $N$ is the number of tourist types. According to the previous equation, the degree of similarity between tourist profiles and tourism objects will be calculated. The degree of appropriateness of a particular tourist object to the tourist profile of the given tourist is calculated according to the following equation:

$$R_{1_{u,j}} = r_{Oj} + SIM_{cos}(Ui,Oj), \text{ for } \forall Oj \in O$$ (3)

where $r_{Oj}$ is the average rating of the object $Oj$, and is used as an universal measure for object attractiveness.

In our methodology, we have considered another very important fact related with the behavior of the people while planning a vacation or trip. In everyday life, people also rely on recommendations from reference letters, news reports, general surveys, travel guides, and so forth. In addition, they desire personal advice from other people with similar preferences or people they trust. In fact, over 80% of travelers participating in a TripAdvisor.com survey agree that “reading other travelers’ online reviews increases confidence in decisions, makes it easier to imagine what a place would be like, helps reduce risk/uncertainty, makes it easier to reach decisions, and helps with planning pleasure trips more efficiently” (Gretzel, 2007).

Experimental findings show that there exists a significant correlation between the trust expressed by the users and their similarity based on the recommendations they made in the system; the more similar two people are, the greater the trust between them (Ziegler and Golbeck, 2007). Similarity can be interpreted in several ways such as similarity in interests or ratings or opinions. Different methodologies can be used to calculate the similarity between the users in the system.

As one of the most prevailing and efficient techniques to building recommender systems, collaborative filtering (CF) implements the idea for automating the process of “word-of-mouth” by which people recommend items to one another. It uses the known preferences of a group of users who have
shown similar behavior in the past to make recommendations of the unknown preferences for other users. CF is facing many challenges, among which the ability to deal with highly sparse data and to scale with the increasing numbers of users and items, are the most important in order to make satisfactory recommendations in a short time period. Sparsity of ratings data is the major reason causing poor recommendation quality. The sparsity problem occurs when available ratings data is rare and insufficient for identifying the similar neighbors. This problem is often very significant when the system is in its early stages. On the other hand, when numbers of existing users and items grow tremendously, traditional CF algorithms will suffer serious scalability problems, with computational resources grown nonlinearly and going beyond practical or acceptable levels.

To reduce the dimensionality of data and avoid the strict matching of attributes in similarity computation the cloud-model CF approach has been adopted. It is constructing the user’s global preference based on his perceptions, opinions and tastes, which are subjective, imprecise, and vague (Palanivel and Siavkumar, 2010), and it seems to be an appropriate paradigm to handle the uncertainty and fuzziness on user preference.

The main goal of the cloud model CF is to construct the global preference for each user by calculating a triple of three digital characteristics $\vec{V} (= (Ex, En, He))$. The expected value $Ex$ represents the typical value of user ratings, that is, the average of user ratings. The entropy $En$ represents the uncertainty distribution of user preference, which is measured by the deviation degree from the average rating. The hyper-entropy $He$ is a measure of the uncertainty of the entropy $En$, which is measured by the deviation degree from the normal distribution. Given a set of ratings data for a user $u_i$, $r_{u_i} = (r_{u_1}, r_{u_2}, \ldots, r_{u_n})$, the three characteristics can be defined as (Zhang et al, 2009):

$$Ex = \frac{1}{n} \sum_{i=1}^{n} r_{u_i}$$

$$En = \sqrt{\frac{\pi}{2} \times \frac{1}{n} \sum_{i=1}^{n} (r_{u_i} - Ex)^2}$$

$$He = \sqrt{S^2 - \frac{1}{3} En^2}, \text{where } S = \frac{1}{n-1} \sum_{i=1}^{n} (r_{u_i} - Ex)^2$$

The $k$ similar (neighbor) users, for an active user are selected based on the cloud model similarities between the active user and the users that already rated the object $Oj \in O$. A likeness similarity method based on cloud model using the cosine measure was proposed in Zhang et al, 2007. Given two cloud models in terms of the characteristic vectors $\vec{V}_{u_i} = (Ex_{u_i}, En_{u_i}, He_{u_i})$, the similarity between them are defined as

$$SIM(\vec{V}_{u_i}, \vec{V}_{u_j}) = \cos(\vec{V}_{u_i}, \vec{V}_{u_j})$$

$$(5)$$

$$SIM(\vec{V}_{u_i}, \vec{V}_{u_j}) = \frac{Ex_{u_i}Ex_{u_j} + En_{u_i}En_{u_j} + He_{u_i}He_{u_j}}{\sqrt{Ex_{u_i}^2 + En_{u_i}^2 + He_{u_i}^2 \sqrt{Ex_{u_j}^2 + En_{u_j}^2 + He_{u_j}^2}}}$$

Considering this similarity metric, a subset of $k$ most similar users to the observed user $Ui$ is created. Recommendation function based on the cloud model is defined as:

$$R_{2u_{i,j}} = \frac{\sum_{u \in \Omega(u_i)} (r_{u_{i,j}} - \bar{r}_{u_i}) \times SIM(\vec{V}_{u_i}, \vec{V}_{u_j})}{\sum_{u \in \Omega(u_i)} SIM(\vec{V}_{u_i}, \vec{V}_{u_j})}$$

where $\Omega(u_i)$ is the subset of $k$ most similar users to active user $Ui$ and $\bar{r}_{u_i}$ and $\bar{r}_{u_j}$ are the average rating of users $Ui$ and $Uk$, respectively. The value of rating $r_{u_{i,j}}$ is weighted by the similarity of users $Ui$ and $Uk$; the more similar the two users are, the more weight $r_{u_{i,j}}$ will have in the computation of the recommendation function.

Total recommendation function for a given tourist object ($O$), is calculated using a weighted average of the functions given by equations (3) and (6):

$$Frec_{i,j} = \frac{w_1 \times R_{1u_{i,j}} + w_2 \times R_{2u_{i,j}}}{w_1 + w_2}$$

(7)

By calculating the recommendation value for each object according to Eq. 7, the objects for a particular tourist will be ordered in a list

$$LQ = \{O1, O2, \ldots, On\},$$

where $Frec_{1,j} > Frec_{1,2} > \ldots > Frec_{i,n}$

and will be further clustered into five tourist specific categories in the following way:

$$\text{Cat}_i = \{k, \forall Oj \in LQ, 1 \leq j \leq 0.1 \times n \}$$

$$\text{Cat}_2 = \{k, \forall Oj \in LQ, 0.1 \times n < j \leq 0.2 \times n \}$$

$$\text{Cat}_3 = \{k, \forall Oj \in LQ, 0.2 \times n < j \leq 0.3 \times n \}$$

$$\text{Cat}_4 = \{k, \forall Oj \in LQ, 0.3 \times n < j \leq 0.4 \times n \}$$

$$\text{Cat}_5 = \{k, \forall Oj \in LQ, 0.4 \times n < j \leq n \}$$

(9)
DEVELOPING WEB PORTAL

The suggested national tourism web portal is structured in the form of a social network. So, the portal is a significant improvement on existing travel websites and provides tourists with a customized, unique, and enriching travel experience. Moreover, it incorporates some standard plugins typical for social networks like Facebook. But, it advances the concept by including custom plugins, like the recommended objects plugin which is based on the proposed methodology and represents the core of the portal. It is using the Google Map of Macedonia to visualize both: static tourism objects (object that are not temporary, like churches, museums, archeology localities, etc.) and dynamic object (object that have limited time duration, like events, expositions, etc.). They are displayed on the map according to their geographical location. Moreover, they are geographically classified into 84 municipalities and grouped into eight regions.

Municipalities are classified into five groups, according to the number of tourist objects related to the municipality and are recommended to the user in the form of circles displayed on the map (Figure 1). The size of the circle indicates the tourist’s affinity for the municipality; therefore, a larger circle indicates a municipality that matches better the tourist profile and contains more tourism objects with higher recommendation marks. The radius of the circle for a particular municipality Mj as seen by the tourist Ui is defined as:

$$R_{i,j} = 10^{9} \cdot \text{Int} \left\{ \sum_{k=1}^{5} \text{Ncat}_k (k) \cdot (5 - k + 1) \right\}$$

(10)

where by Ncat, is given the number of objects placed in the municipality Mj, belonging to one of the five categories, as defined by Eq. 9. By displaying the tourist’s affinity through the size dimension of the circle, tourists can easily observe which municipality is of most interest to them.

Highly relevant objects (i.e. those classified into category 1 according to Eq. 9) are also clearly marked on the map. When the map is zoomed in the objects are represented by icons. The icon indicates the type of tourism objects such as a museum, church or restaurant. The size of the icon indicates how closely the object meets the user’s interests. When the icon of a tourism object is clicked an information window appears (Figure 2). The information window usually includes the name and picture of the object, an icon of an umbrella indicating that the attraction is accessible in the rain, and tags. The information window also displays a general idea of time consumption of the attraction, friends who have visited the attraction and an option to view multimedia materials either in video, audio, or text format. Through this window, the user can also rate the object. This operation is recommended to be done after visiting the object and according to the personal experience and satisfaction. The goal of this operation is two-fold: to help updating the user profile, and to make the process of recommendation more accurate.

SYSTEM ACCURACY

The research is based on proprietary database collected by the mixed research group composed of researchers from Faculties of Computer Science and Tourism at the “Goce Delcev” University. It contains 9420 ratings from 194 users for 380 tourism objects. Each user rated at least 30 objects, and each object has been rated at least once.
In order to measure accuracy in a more precise manner, we use information-retrieval classification metrics. Namely, we evaluate the capacity of the recommender system by suggesting a list of appropriate objects to the user. With such metrics it is possible to measure the probability that the recommender system takes a correct or incorrect decision about the user interest for an item. When using classification metrics, we distinguish four different kinds of recommendations (Table 1).

If the system suggests an interesting tourist object to the user we have a true positive (TP), otherwise the object is uninteresting and we have a false positive (FP). If the system does not suggest an interesting tourist object we have a false negative (FN). In case when the system does not suggest an interesting tourist object, we have a true negative (TN). The most popular classification accuracy metrics are the recall and the precision. These metrics can be calculated by counting the number of test objects that fall into each cell in the following table (Table 1) and according to equations 11 and 12.

Table 1 - Classification of possible result

<table>
<thead>
<tr>
<th></th>
<th>Recommended</th>
<th>Non recommended</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interesting</td>
<td>True-positive (TP)</td>
<td>False-negative (FN)</td>
</tr>
<tr>
<td>Uninteresting</td>
<td>False-positive (FP)</td>
<td>True-negative (TN)</td>
</tr>
</tbody>
</table>

Precision = \( \frac{TP}{TP + FP} \) \hspace{1cm} (11)

Recall (True Positive Rate) = \( \frac{TP}{TP + FN} \) \hspace{1cm} (12)

Recall measures the percentage of interesting objects suggested to users, with respect to total number of interesting objects, while precision measures the percentage of interesting objects suggested to the users, with respect to the total number of suggested objects. In the line of understanding the global quality of a recommender system, we combine recall and precision by means of the F-measure

\[ F - measure = \frac{2 \cdot \text{recall} \cdot \text{precision}}{\text{recall} + \text{precision}} \] \hspace{1cm} (13)

In evaluating the quality of the recommendation, we use these metrics. To evaluate the system, a methodology which uses the k-fold and the leave-one-out together with classification metrics recall and precision was used. According to the k-fold, users in the dataset are partitioned into k parts: k - 1 parts represent the and are used to construct the model, the remaining part represents the testing set. The model created with the k - 1 partitions is tested on the remaining partition by means of the following algorithm:

Step 1: One user in the testing set is selected (the active user);

Step 2: One rated tourist object (the test object) is removed from the profile of the active user;

Step 3: An order list of recommended tourist objects is generated; and

Step 4: If the test item is in the top-3 categories (according to the Eq. 9) of recommended objects, either the true positive or false positive counter is incremented, depending whether the user liked or disliked the test item.

We considered two distinct user groups. Group A contained all users who have rated 30–50 objects (the few raters user group), while Group B contained all users who have rated 51–100 objects (the moderate raters user group). Step 1 of the proposed algorithm was repeated for all the users in both groups. Steps 2-4 are repeated for all the objects rated by the active user.

In order to understand if a user likes or dislikes a rated tourism object, we suppose that an object is interesting for the user if it satisfies the two following conditions:

\[ Rate_{i,j} \geq 3 \land Rate_{i,j} \geq \bar{Rate}_i \] \hspace{1cm} (14)

where \( Rate_{i,j} \) is the rate given by the user \( i \) for the tourism object \( j \) and \( \bar{Rate}_i \) is the mean of ratings for user \( i \). The first constraint reflects the absolute meaning of the rating scale, while the second the user bias. If a rating does not satisfy conditions given by Eq.14 we assume the item is not interesting for the user.

In order to measure accuracy in a more precise manner, we use information-retrieval classification metrics. Namely, we evaluate the capacity of the recommender system by suggesting a list of appropriate objects to the user. With such metrics it is possible to measure the probability that the recommender system takes a correct or incorrect decision about the user interest for an item. When using classification metrics, we distinguish four different kinds of recommendations (Table 1).

If the system suggests an interesting tourist object to the user we have a true positive (TP), otherwise the object is uninteresting and we have a false positive (FP). If the system does not suggest an interesting tourist object we have a false negative (FN). In case when the system does not suggest an interesting tourist object, we have a true negative (TN). The most popular classification accuracy metrics are the recall and the precision. These metrics can be calculated by counting the number of test objects that fall into each cell in the following table (Table 1) and according to equations 11 and 12.

Table 1 - Classification of possible result

<table>
<thead>
<tr>
<th></th>
<th>Recommended</th>
<th>Non recommended</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interesting</td>
<td>True-positive (TP)</td>
<td>False-negative (FN)</td>
</tr>
<tr>
<td>Uninteresting</td>
<td>False-positive (FP)</td>
<td>True-negative (TN)</td>
</tr>
</tbody>
</table>

Precision = \( \frac{TP}{TP + FP} \) \hspace{1cm} (11)

Recall (True Positive Rate) = \( \frac{TP}{TP + FN} \) \hspace{1cm} (12)

Recall measures the percentage of interesting objects suggested to users, with respect to total number of interesting objects, while precision measures the percentage of interesting objects suggested to the users, with respect to the total number of suggested objects. In the line of understanding the global quality of a recommender system, we combine recall and precision by means of the F-measure

\[ F - measure = \frac{2 \cdot \text{recall} \cdot \text{precision}}{\text{recall} + \text{precision}} \] \hspace{1cm} (13)

In evaluating the quality of the recommendation, we use these metrics. To evaluate the system, a methodology which uses the k-fold and the leave-one-out together with classification metrics recall and precision was used. According to the k-fold, users in the dataset are partitioned into k parts: k - 1 parts represent the and are used to construct the model, the remaining part represents the testing set. The model created with the k - 1 partitions is tested on the remaining partition by means of the following algorithm:

Step 1: One user in the testing set is selected (the active user);

Step 2: One rated tourist object (the test object) is removed from the profile of the active user;

Step 3: An order list of recommended tourist objects is generated; and

Step 4: If the test item is in the top-3 categories (according to the Eq. 9) of recommended objects, either the true positive or false positive counter is incremented, depending whether the user liked or disliked the test item.

We considered two distinct user groups. Group A contained all users who have rated 30–50 objects (the few raters user group), while Group B contained all users who have rated 51–100 objects (the moderate raters user group). Step 1 of the proposed algorithm was repeated for all the users in both groups. Steps 2-4 are repeated for all the objects rated by the active user.

In order to understand if a user likes or dislikes a rated tourism object, we suppose that an object is interesting for the user if it satisfies the two following conditions:

\[ Rate_{i,j} \geq 3 \land Rate_{i,j} \geq \bar{Rate}_i \] \hspace{1cm} (14)

where \( Rate_{i,j} \) is the rate given by the user \( i \) for the tourism object \( j \) and \( \bar{Rate}_i \) is the mean of ratings for user \( i \). The first constraint reflects the absolute meaning of the rating scale, while the second the user bias. If a rating does not satisfy conditions given by Eq.14 we assume the item is not interesting for the user.

In order to measure accuracy in a more precise manner, we use information-retrieval classification metrics. Namely, we evaluate the capacity of the recommender system by suggesting a list of appropriate objects to the user. With such metrics it is possible to measure the probability that the recommender system takes a correct or incorrect decision about the user interest for an item. When using classification metrics, we distinguish four different kinds of recommendations (Table 1).

If the system suggests an interesting tourist object to the user we have a true positive (TP), otherwise the object is uninteresting and we have a false positive (FP). If the system does not suggest an interesting tourist object we have a false negative (FN). In case when the system does not suggest an interesting tourist object, we have a true negative (TN). The most popular classification accuracy metrics are the recall and the precision. These metrics can be calculated by counting the number of test objects that fall into each cell in the following table (Table 1) and according to equations 11 and 12.

Table 1 - Classification of possible result

<table>
<thead>
<tr>
<th></th>
<th>Recommended</th>
<th>Non recommended</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interesting</td>
<td>True-positive (TP)</td>
<td>False-negative (FN)</td>
</tr>
<tr>
<td>Uninteresting</td>
<td>False-positive (FP)</td>
<td>True-negative (TN)</td>
</tr>
</tbody>
</table>

Precision = \( \frac{TP}{TP + FP} \) \hspace{1cm} (11)

Recall (True Positive Rate) = \( \frac{TP}{TP + FN} \) \hspace{1cm} (12)

Recall measures the percentage of interesting objects suggested to users, with respect to total number of interesting objects, while precision measures the percentage of interesting objects suggested to the users, with respect to the total number of suggested objects. In the line of understanding the global quality of a recommender system, we combine recall and precision by means of the F-measure

\[ F - measure = \frac{2 \cdot \text{recall} \cdot \text{precision}}{\text{recall} + \text{precision}} \] \hspace{1cm} (13)

In evaluating the quality of the recommendation, we use these metrics. To evaluate the system, a methodology which uses the k-fold and the leave-one-out together with classification metrics recall and precision was used. According to the k-fold, users in the dataset are partitioned into k parts: k - 1 parts represent the and are used to construct the model, the remaining part represents the testing set. The model created with the k - 1 partitions is tested on the remaining partition by means of the following algorithm:

Step 1: One user in the testing set is selected (the active user);

Step 2: One rated tourist object (the test object) is removed from the profile of the active user;

Step 3: An order list of recommended tourist objects is generated; and

Step 4: If the test item is in the top-3 categories (according to the Eq. 9) of recommended objects, either the true positive or false positive counter is incremented, depending whether the user liked or disliked the test item.

We considered two distinct user groups. Group A contained all users who have rated 30–50 objects (the few raters user group), while Group B contained all users who have rated 51–100 objects (the moderate raters user group). Step 1 of the proposed algorithm was repeated for all the users in both groups. Steps 2-4 are repeated for all the objects rated by the active user.

In order to understand if a user likes or dislikes a rated tourism object, we suppose that an object is interesting for the user if it satisfies the two following conditions:

\[ Rate_{i,j} \geq 3 \land Rate_{i,j} \geq \bar{Rate}_i \] \hspace{1cm} (14)

where \( Rate_{i,j} \) is the rate given by the user \( i \) for the tourism object \( j \) and \( \bar{Rate}_i \) is the mean of ratings for user \( i \). The first constraint reflects the absolute meaning of the rating scale, while the second the user bias. If a rating does not satisfy conditions given by Eq.14 we assume the item is not interesting for the user.
user. Once computed recall and precision, we synthesize them with the f-measure, as defined in (Eq. 13).

Upon the conducted evaluation the results for system precision, recall and f-measure were averaged for each of the groups (Table 2).

**Table 2 - Average values for recommendation system (%)**

<table>
<thead>
<tr>
<th>Group</th>
<th>Precision</th>
<th>Recall</th>
<th>F-measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A</td>
<td>73.67</td>
<td>78.12</td>
<td>75.83</td>
</tr>
<tr>
<td>Group B</td>
<td>79.43</td>
<td>83.44</td>
<td>81.39</td>
</tr>
</tbody>
</table>

According to the result outcomes, the developed national tourism web portal with its collaborative recommender system seems to be robust as it achieves good results in both scenarios (users with few and moderate ratings). It also accomplishes a good trade-off between precision and recall, a basic requirement for all recommendation systems. Experimental results show that the proposed approach can provide satisfactory performance even in a sparse dataset.

**REFERENCES**


**CONCLUSION**

Although being in its initial phase of development, the suggested designed national tourism portal is rich in accurate recommendations and guidelines. Hence, the tourists and travelers willing to visit Macedonia may apply it in identification of their ideal trip and holiday. Due to the fact that tourism is identified as one of the most economically-oriented world-wide industries, it can be used as a mean for enhancing and strengthening the national economy. So, the development of such software module contributes generally to increasing the awareness of tourist destination that is capable of fulfilling travelers’ preferences, and respectfully the tourism competitiveness of the country.

The national web portal “MyTravelPal” assists all interested parties in planning their travel on more intelligent way by generating a personalized list of favorable and tailor-made items. Since this portal assists tourists and travelers in identification of their ideal holiday place within Macedonia, it contributes to improvement of tourism demand in qualitative and quantitative manner. Hence, this empirical investigation underlines the high priority importance of creating this kind of tourism recommendation system which will consequently enable the country to create competitive tourism product.