

## Recommender system for personalised travel itinerary

Tanuja Choudhary B, Tulasi B

Department of Computer Science, CHRIST (Deemed to be University), India

---

### Article Info

#### Article history:

Received Feb 23, 2019

Revised Apr 17, 2019

Accepted Apr 26, 2019

---

#### Keywords:

Recommender system

Tourpedia

Travel

Tourism

---

### ABSTRACT

A recommender system is an approach to give an appropriate solution to a particular problem. This helps in recognising the pattern or behaviour of a user to suggest future possible likes of the user. Nowadays people like to travel during their spare time, it has become a rigid task to decide where to go. This paper represents a customised recommender system to help users in destining their itinerary. A model is designed to suggest the best places to visit in Rome. A questionnaire was prepared to get information about user's interest during their travel. The model generates the best five places to visit with respect to the choice picked by the user. The top five places for each category will be displayed to the user and the user was asked to pick a starting point for the itinerary. Then the model generates another set off a filtered list of places to enhance their travel experience. It includes displaying the top 5 restaurants to visit during their travel.

*Copyright © 2019 Institute of Advanced Engineering and Science.  
All rights reserved.*

---

### Corresponding Author:

Tanuja Choudhary B,

Department of Computer Science,

CHRIST (Deemed to be University),

No -17, Meera Bhavan, 2nd cross, Munnikrishnappa Layout, Adugodi, Bengaluru 560030, India.

Email: tanujachoudhary.choudhary@gmail.com

---

## 1. INTRODUCTION

A recommender system is a system used for information filtering. The system learns information from the past data to give a suggestion. It is a tool used to absorb detail of information in a specific environment. Each environment is a form of different domains such as e-commerce, tourism, social-media, advertisement and etc. With the explosive growth of social media and development of web 2.0, large amounts of travel information are being uploaded per minute on travel websites [1]. In a domain, there is a various number of factors which can affect in the way the recommender model behaves. One such need was in the field of e-commerce, due to the increase of growth in number of customers during the last four years only about 15% (2010-2014) [2]. This led to a significant need to build a recommender framework to choose which factors yield an accurate outcome.

Recommender systems are usually classified according to their approach to rating estimation [3]. This system is classified into the following types, Collaborative filtering, Content-based, Knowledge based and Hybrid recommenders. Collaborative filtering is probably the most familiar, most widely implemented and most mature of the technologies [4]. It takes into consideration the views/ratings of other people when deciding on recommendations; sometimes its narrowed down into a specific demographic with similar interests. This filtering has some challenges namely data sparsity and scalability [5]. Content-based recommenders make decisions based on what the user has previously rated or what the user is currently looking at [6]. Knowledge-based set out a suggestion on the basis of user's needs and tastes. It matches the need of the particular item with the need of a particular user to give a possible suggestion. The hybrid system is a combination of Content-based and Collaborative filtering. With the help of historical information and the new information of the user is considered and analysed to give the recommendation.

This paper focuses on the field of tourism. As a major part of the modern service industry, tourism has experienced rapid growth over the past decade [7]. The tourists of today are very demanding and have

complex, multi-layered desires and needs [8]. Travel and tourism is an area where the majority of people come under. There are many factors which affect to build an efficient recommender system. Due to this choosing relatively, important factors is difficult. In order to search and recommend tourist spots effectively, it is first necessary to characterise the tourist spots [9]. Many users prefer a need for personal recommender system which can give them a set of an itinerary on the basis of their preferences on traveling. To address this issue, In this paper, a recommender system is proposed to enhance users travel experience.

## 2. LITERATURE REVIEW

During leisure time users prefer to travel. Travel anywhere around the world. When a user decides on where to visit it always becomes a tedious task to decide where to go first. Nowadays there are multiple options available which help the user overcome this task. A recommender system is a solution. Recommender system helps end users to narrow down on which places to visit. In [10] a particular location was examined using a crowdsourcing approach. The main focus was on the amount of crowd present at a particular location. The proposed model not only considers information from the blogs, web pages, sensor readings, etc... but also takes in users viewpoint. With the help of TSRS (Tourist Spot Recommender System) and set of location suggested to the users the system performs just-in-time information enrichment for those selected set of location. The proposed system is designed for mobile users to enhance their travel experience. The growth of information is increasing drastically which leads to the concern of big data. Due to vast challenges in the field of big data [11], recommender system is solution for some. The paper uses Singular value decomposition (SVD) to outperform the challenges faced in the field of big data. SVD based recommender system is proposed to evaluate and learn the performance and accuracy in a distributed environment. Hadoop and spark are the tools used to perform the experiment.

Sometimes the textual details are just not enough to capture complete information about a rental place hence this paper [12] considers image details as well. Image and textual information are captured to generate a recommender system. Choosing a vacation rental was a challenging task. This paper uses three methods images-based cosine similarity calculation, textual description-based Jaccard similarity calculation and the fusion of both the methods. A hybrid recommender system was built to suggest preferable accommodation to the user. In spite of the fact that many travel proposal frameworks are worked with a mobile social network, the articles and conveyed messages between users are still not used well to assess users' inclinations. To eliminate this issue in [13] paper cloud-based recommender system is proposed. The ideology was to combine the latest cloud technology, analyse posted blogs on the mobile network system and identify users' pattern through sensors. The process was carried down in three different stages namely pre-touring, in-touring and post-touring. Each tour functions are different from the other. The proposed system initially constructs meta-group based on similar user preferences. Furthermore, with the help of CLOPE algorithm meta-group was classified. The result was well examined and suitable POIs was given to the users during touring.

There are different challenges faced to recommend travel itinerary to a user or group of users based on their tourist interest and choice. This paper [14] builds a recommender model to overcome the challenges. The proposed algorithm in the paper construct tourists past POI visit based on their geo-tagged photos and then construct a model of user's choice based on their time spent visiting each POI. The proposed model out-performs various factors to give a relatively better recommendation to users based on their preferences. In recent years the use of contextual data is become more powerful to give a recommendation. With the help of this data and considering the past information the following [15] paper constructs a recommender system. The recommender framework gets logical data by mining client reviews and joining them with client rating history to figure a utility capacity over an arrangement of things. In the framework, the setting derivation is displayed as a regulated subject demonstrating issue in which an arrangement of classes for a relevant trait establishes the point set. Words play an important role in the comment on a place or an object. These words can be used and helped in generating a recommender system. This paper [16] generates temporal feature vectors on passing objects to their proposed recommender system. The system evaluates by identifying the vocabulary related to the objects with the help of Wikipedia, finding the trend of all the objects with the help of Twitter and to feature the heaviness of words contained in each distinguished trend to get temporal feature vectors for each object. The resultant vectors precisely return the resemblance of a point of interests for appointed time duration.

Recommender system is helpful to give recommendation on bases of historical data but when it comes to new users it fails to perform at the maximum. To resolve this issue the paper of [17] proposed a goal-based approach. The paper uses content based filtering and collaborative filtering using K-nn. The combination of the two technique to give appropriate and relevant overcome to the new user profile. Goal based approach with K-nn, it compares the similarities between users to give suggestion. It was a

crucial task to collect appropriate data from a vast amount of information for a recommender system. The paper of [18] proposed the architecture of tourist support information system which includes VR contents to promote Iwate area in Japan and gathers content repository and training data to build regional specific recommender engine on the tourist support system. The authors in [19] utilise textual information from users reviews and use ad-hoc and regression-based recommendation measures to give a personalised rating value. The model proves that the acquired rating value gives a better score than the derived rating value given by the users. The work in [20] establishes a connection with the user by collecting information of users interests and uses his/her current context to recommend nearby spots based on their interest.

### 3. RESEARCH METHOD

Vast amount of information is on the internet, information about places to visit, restaurants, movies, shopping center, hotels and many more. It has become a necessity to go to the top place. This makes it difficult for users to decide or pick among the vast options. To deal with this issue a system is proposed named recommender system. The recommender system is a technique to filter out the information and give suggestion to end users. Following are the steps to build a recommender system:

Step 1: Gathering data

Step 2: Preprocess and transform data

Step 3: Build a model

Step 4: Evaluate the model and generate the outcome.

R tool is used to generate the model. This tool provides a set of statistical and graphical techniques. It includes regression, statistical inference, machine learning algorithm, data analysis to name a few. R provides various package through which various operation can be performed on the data. Some of the packages used to build the recommended model are googleway, dplyr, lmap to name few. For instance, lmap package helps in computing the geo distance between two points specified by latitude/longitude using Vincenty inverse formula for ellipsoids. Hence, it is a powerful tool to get appropriate insight into the raw data and helps in decision making.

#### 3.1. Dataset

The data is collected from [www.tour-pedia.org](http://www.tour-pedia.org). The link consists of data on 8 different European cities. Each city is categories with four types namely Accommodation, Restaurant, Point of Interest and Attraction. In this paper Point of Interest and Attraction of Rome is considered. Data from tourpedia consists of 12 attributes such as name, id, Category, location, latitude and longitude, polarity, address, subCategory, originalId, details, and review. With respect to the above attributes, the dataset consists of 28198 observation in total. The details of each place in Rome is collected using the Google Places API. This API is initiated using the R tool, in which googleway package is used. This package has the Google Places API method, which is used to retrieve the place details. The data is collected from end user's perspective on the basis of their preferences and choice for the travel.

#### 3.2. Pre-processing

Data preprocessing is the process of transforming data into a form that is fit for analysis. This is due to the various factors like data not being in the right datatype, unstructured format, erroneous data, etc. There are several techniques by which data can be transformed into the right format for analysis. Data-mining has techniques to preprocess the data and it also has a collection of various techniques to extract patterns and to build models from large data-sets [21]. The data collected for this research had erroneous and unwanted data which had to be filtered in order to create the recommender system.

##### 3.2.1. Data cleaning

The dataset from tourpedia consists of 12 attributes such as id, name, address, originalId, latitude, longitude, location, subCategory, Category polarity, details, and reviews. The name column had observations which had garbage values which had to be removed. This was done using R. A function was written to check and omit the observations which did not have characters from a-z and 0-9 in the name column.

##### 3.2.2. Data reduction

After data cleaning was performed the data was further reduced to consider only the name column. The place names were then passed to Google Places API to fetch the details of each place. For each place-name rating and types were considered. The columns which had multiple rating values were then reduced by considering the average rating. Among the value fetched there were few columns which had rating value has 0, these observations were omitted from the dataset.

#### 4. IMPLEMENTATION

This section provides details about the proposed recommender system. The system generates the top 5 places to visit during their stay in Rome. The suggestion of places was based on the answer to the category of the questions asked to user. The questionnaire contains a set of predefined questions. These questions are to be answered by users based on their choice and preferences. Among them three questions have one or more sub question and three questions have no sub question. The response was taken in the form of binary either yes or no. Depending on the user requirements the recommender system will generate output. The proposed model takes the input from the user and processes it. Figure 1 consist of some of the questions asked to the user. The output was generated on the basis of past database. The database consists of 1336 observations.

Each sample was differentiated from the type value. This type value consists of 17 different levels. The levels are art\_gallery, Historical\_place, museum, Monument, mosque, place\_of\_worship, synagogue, Roman\_temple, natural\_feature, park, zoo, campground, amusement\_park, book\_store, library, lodging and premise. There are 22 observation with rating 5 for art\_gallery. With the help of sample function in R at each run random five places are displayed to the users. There are only 14 observations under the category of historical place. These observations fall under rating of 4.6, 4.7 and 4.9. At each run any five random places are displayed. With respect to monument there are 12 observations and few samples are distributed into 4.7 and and few are distributed into 4.6 rating. A set of 5 random places are displayed from the set of 12 observation. In total there are 136 observation with the type value museum. The data was distributed well across rating ranging from 5 to 2. From the total a single sample is randomly picked from the following rating value 5, 4.9, 4.8, 4.7 and 4.6.

Rome is famous for its churches due to this fact there are 596 observation under place\_of\_worship. Among them 39 observations have rating 5 from which top 5 is randomly picked. Under mosque type there are 6 observation which fall in the following rating value such as 4.7, 4.5, 4.2 and 4.1. From these set, random 5 places of mosque are generated and displayed to the user. There are only 3 synagogue places two of which is rated 4.4 and the other is rated 3.6. All the three places was suggested to the users. In the dataset there are only two places of the type roman temple one of which is 4.7 and the other is 4.6. Both of these places are suggested to the users on their visit only if they are interested in visiting roman temple in Rome. For the users who are nature lovers and interested in visiting local parks at Rome a single sample is randomly picked from the rating value between 4.4 to 4.8 of natural\_feature and 4.3 to 4.7 of park. There are 4 observation with the type value zoo of which three have a rating of 4.3 and the other has a rating of 2.1. These places will be suggested to the users who are interested in visiting zoo in Rome. There are 8 observations which have the type value campground and if the user select yes from these five random places are picked and displayed to the user.

There are 20 observations having the type amusement\_park, among these 20 top five are randomly picked which have the rating 4 or above. For those users interested in bookstores, there are 11 observations from which the top five places are randomly chosen on the basis of rating value where rating value is either 4 or above. And lastly, for those users who would like to visit libraries, there are 11 observations which have 5 rating under library type, five places are picked randomly among 11 and displayed to the user. After picking the top places the user was asked to pick a starting point to start there itinerary, as shown in Figure 2.

**Tourist Form**

Are you interested in art galleries?  
 Yes  
 No

Are you interested in historical places?  
 Yes  
 No

Do you like visiting Museum?  
 Yes  
 No

Do you like visiting Monument?  
 Yes  
 No

Do you enjoy going to amusement parks?  
 Yes  
 No

Figure 1. Questionnaire design

**Tourist Form**

Choose your starting point ?  
 Galleria Tartaglia Arte ▼

**SUBMIT**

Figure 2. Starting point

The recommender system will then generate a distance matrix has shown in Table 1. With the help of distance matrix, the original list of places to visit was narrowed down. The final outcome will show all the places within 8km radius from the starting point. Additionally, top 5 restaurants from the starting place is suggested to the user. Google Places API method, which is used to retrieve the restaurant details. Apart from the questions asked as shown in the Figure. 1, premise and lodging are some additional places which the proposed system is going to recommend. From premise and lodging top five places is randomly picked which has rating 4 or above to all the users.

Table 1. Distance matrix

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Galleria Trincia	0	2	1	8	4	4	1	4	1	1	6	6	4	4	24	4	2
Galleria Rilievi Ro	2	0	2	7	2	2	4	3	3	2	5	4	3	3	26	2	1
Galleria d'Arte Na	1	2	0	8	1	3	2	3	1	4	6	6	5	5	24	1	1
Galleria Tartaglia	8	7	8	0	8	5	7	9	9	8	7	4	5	5	33	8	8
Galleria Bordini	3	2	1	8	0	3	1	3	1	3	6	6	4	4	24	3	1
Pia Unione Madre	4	2	3	5	3	0	2	4	4	3	4	2	3	3	27	3	3
Casa Generalizia	1	4	2	7	1	2	0	4	2	1	5	5	3	3	26	1	1
Azione Cattolica I	4	3	3	9	3	4	4	0	4	3	3	5	7	7	24	4	2
Ente Provincia Ro	1	3	1	9	1	4	2	4	0	1	7	7	5	5	23	1	2
Curia Generalizia	1	2	4	8	4	3	1	3	1	0	6	6	4	4	24	4	1
Masjeed E Rome	6	5	6	7	6	4	5	3	7	6	0	3	7	7	27	6	5
Grande Moschea	6	4	6	4	6	2	5	5	7	6	3	0	5	5	29	6	5
Islamic Cultural C	4	3	5	5	4	3	3	7	5	4	7	5	0	5	29	4	5
Muslim Cultural I	4	3	5	5	4	3	3	7	5	4	7	5	5	0	28	4	5
Mosque Al Aqsa	24	26	24	33	24	27	26	24	23	24	27	29	29	28	0	25	24
Pantheon	4	2	1	8	4	3	1	4	1	4	6	6	4	4	25	0	1
House of the Vest	2	1	1	8	1	3	1	2	2	1	5	5	5	5	24	1	0

5. RESULTS AND ANALYSIS

The framework is customised to one European city named Rome. This generates the result of the places name to be from Rome. The output of the model gives a set of specific place names depending on the answer to a specific question. The outcome was solely dependent on users' interest and choices. The result is picked from the past collection of a dataset on Rome. The result shows the top 5 places for each question asked to the user. These top 5 places are the mostly preferably places to visit during the stay in Rome. Once the user picks the starting point of the journey all filtered list of places as shown in Figure 3 is displayed. This is the final outcome. The result shows the places name within 8km radius from the starting point to visit during their travel. Along with the places, the top 5 restaurants from the starting point was also recommended to the users as shown in Figure 4.

THESE ARE THE TOP ART GALLERIES IN ROME  
 Galleria Trincia  
 Galleria Rilievi Roma Arte Contemporanea  
 Galleria d'Arte Nardi  
 Galleria Tartaglia Arte  
 Galleria Bordini

THESE ARE THE TOP CHURCHES IN ROME  
 Pia Unione Madre Santissima Della Visitazione  
 Casa Generalizia Della Societa Del Sacro Cuore  
 Azione Cattolica Italiana - Diocesi di Roma  
 Ente Provincia Romana Frati Minori Cappuccini  
 Curia Generalizia Marianisti

THESE ARE THE TOP MOSQUES IN ROME  
 Masjeed E Rome (Mosque)  
 Grande Moschea di Roma  
 Islamic Cultural Center and Mosque of Italy  
 Muslim Cultural Institute and Mosque  
 Mosque Al Aqsa

THESE ARE THE ROMAN TEMPLE IN ROME  
 Pantheon  
 House of the Vestals

Figure 3. Top places

These are the following places within 8kilometers radius from the starting point

Grande Moschea di Roma  
 Pia Unione Madre Santissima Della Visitazione  
 Islamic Cultural Center and Mosque of Italy  
 Muslim Cultural Institute and Mosque  
 Galleria Rilievi Roma Arte Contemporanea  
 Casa Generalizia Della Societa Del Sacro Cuore  
 Masjeed E Rome (Mosque)

Here are some top restaurants you can visit during your travel around the listed places

MILE PIGNETO  
 Ma.Ma. Ristorante Feudo Lucano  
 Le Petit Bio  
 Ambrosia Rooftop Restaurant & Bar  
 UNA Restaurant

Figure 4. Nearby places and restaurants

## 6. CONCLUSION

People usually prefer traveling in the free time of their everyday life. It has become a tedious task to decide on the itinerary. This leads to the need of a recommender system. A recommender system helps in choosing the top places and generates an appropriate outcome on which place to visit. Though there are many types of recommender systems, a personalised one is always preferred. A customised model helps to narrow down on the top places to visit. The proposed model generates places on the basis of users preferences and choice. Currently, the proposed structure is narrowed down to one European city which can be further utilised to achieve the same among the other cities.

## REFERENCES

- [1] J. Shen, J. Shen, T. Mei, X. Gao, "Landmark Reranking for Smart Travel Guide Systems by Combining and Analyzing Diverse Media," *IEEE Transactions on Systems Man and Cybernetics: Systems*, vol. 46(11), pp. 1492-1504, 2016.
- [2] Sfenrianto S., Saragih M. H., Nugraha B., "E-Commerce Recommender For Usage Bandwidth Hotel," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9(1), pp. 227-233, 2017.
- [3] G Adomavicius, A. Tuzhilin, A. Tuzhilin, "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17(6), pp. 734-749, 2005.
- [4] Robin Burke, "Hybrid Recommender Systems: Survey and Experiments," *User Modeling and User-Adapted Interaction*, vol. 12(4), pp. 331-370, 2002.
- [5] Veena Ch, Bahu B.V, "A User-Based with a Scalable Machine Learning," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 5(5), pp. 1153-1157, 2015.
- [6] K. Meehan, T. Lunney, K. Curran, A. McCaughey, "Context-aware intelligent recommendation system for tourism", *IEEE PerCom*, vol. 17, pp. 328-331, 2013.
- [7] H. Ding, L. Ke, Z. Geng, "Route planning in a new tourist recommender system: A fireworks algorithm-based approach", *Evolutionary Computation (CEC) 2016 IEEE Congress on*, 2016, pp. 4022-4028.
- [8] Tobias Berka and Manuela Plößnig, Salzburg Research Jakob, "Designing Recommender Systems for Tourism". *Salzburg Research*, pp. 1-8, 2003.
- [9] K. Oku, K. Ueno, F. Hattori, "Mapping geotagged tweets to tourist spots for recommender systems," *Proceedings of the IIAI 3rd International Conference on Advanced Applied Informatics*, 2014, pp. 789-794.
- [10] S. Tiwari and S. Kaushik, "Information enrichment for tourist spot recommender system using location aware crowdsourcing," in *Proc. IEEE 15th Int. Conf. Mobile Data Manage*, 2014, pp. 11-14.
- [11] Almohsen KA, Al-Jobor H., "Recommender Systems in Light of Big Data," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 5(6), pp. 1553-1563, 2015.
- [12] Pai Chet Ng, James She, Ming Cheung and Alexander Cebulla, "An Images-Textual Hybrid Recommender System For Vacation Rental," *2016 IEEE Second International Conference on Multimedia Big Data*, 2016, pp. 60-63.
- [13] Yuan-Tse Yu, and Chung-Ming Huang, Yun-Tz Lee, "An Intelligent Touring Systems based on Mobile Social Network and Cloud Computing for Travel Recommendation," *2014 28th International Conference on Advanced Information Networking and Applications Workshops*, 2014, pp. 19-24.
- [14] Kwan Hui Lim, "Personalized Recommendation of Travel Itineraries based on Tourist Interests and Preferences," *July 2016 International Conference on User Modelling, Adaptation, and Personalization*, jul 2016.
- [15] Hariri N., Mobasher B., Burke R., Zheng Y., "Context-aware recommendation based on review mining," In *Proceedings of the 9th Workshop on Intelligent Techniques for Web Personalization and Recommender Systems in IJCAI'11*, 2011, pp. 30-36.
- [16] G.S. Fang, S. Kamei, S. Fujita, "Automatic Generation of Temporal Feature Vectors with Application to Tourism Recommender Systems," *Proc. 4th Int. Symp. On Computing and Networking (CANDAR)*, 2016, pp. 676-680.
- [17] Chughtai M W, Selamat A, Ghani I., "Goal-based Hybrid Filtering for User-to-user Personalized Recommendation," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 3(3), pp. 329-336, 2013.
- [18] Go Hirakawa, Goshi Sato, Kenji Hisazumi, Yoshitaka Shibata, "Data gathering system for recommender system in tourism," *18th International Conference on Network-Based Information System Systems*, 2015, pp. 521-525.
- [19] Gayatree Ganu, Noémie Elhadad and Amélie Marian, "Beyond the Stars: Improving Rating Predictions using Review Text Content," *WebDB*, pp. 1-6, 2009.
- [20] van Setten M., Pokraev S., Kooldwaaij J., "Context-Aware Recommendations in the Mobile Tourist Application COMPASS ". In *De Bra P.M.E., Nejd W. (eds) International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems. AH 2004. Lecture Notes in Computer Science*, vol. 3137, pp. 235-244, 2004.
- [21] S Girish, B Ramamurthy, T Senthilnathan, "Mining the Web Data for Classifying and Predicting Users' Requests," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 8(4), pp. 2390-2398, 2018.