

User Preference and Reviews Analysis with Neural Networks for Travel Recommender Systems

S.Bairavel^{1*}, Dr. M. Krishnamurthy²

¹Assistant Professor, Selection Grade, Department of CSE, KCG College of Technology, Karapakkam, Chennai, Tamilnadu, India.

²Professor, Department of CSE, KCG College of Technology, Karapakkam, Chennai, Tamilnadu, India

¹ORCID: 0000-0003-1903-9143

Abstract

These days' people like to explore numerous places all over the world predominantly which are highly recommended. For instance, people who are new to a particular location regularly choose places to visit manually by typing some of them wished by the user working on search engine applications. It becomes difficult whenever to search manually and plan accordingly which might be not accurate. To overcome the above issues, the travel recommendation system applies sentiment analysis comparing user preferences, number of days, number of people and delivers the top recommended places to the user considering the pre-existing reviews of the places which are recommended. While processing it calculates the similarity between user preferences/inclinations and reviews using sentimental attributes (positive and negative) to match the similarity. Travel recommendation system takes sentiment polarity and similarity value as parameters on the whole, also the days to be visited and suggests esteem places.

Keywords: sentiment analysis, reviews, sentiment polarity, similarity value, user preferences, Pre-existing reviews.

1. INTRODUCTION

Older people used to prescribe puts by social affair data from different individuals on different destinations which is beyond the realm of imagination all an opportunity to visit physically and suggest. This has been changed because of the fast development in innovation where the suggestion framework overwhelms. The proposed framework assumes a significant job in this day and age as though everyone needs exact data concerning their desires. This framework works on sentiment analysis [1] by taking the client inclinations and previous audits [2] as information and presents the top spots. Processing incorporates computing the similarity value [3] and sentiment polarity [4] to match the similarity. It thinks about the set of past reviews and inclinations which are given by the client to create a numerical value. An incentive with the most noteworthy numerical rate is prescribed to the clients.

Recommendation system has been utilized in a variety of area services like Netflix, YouTube, and Spotify, product recommenders for services such as Amazon, or content recommenders for social media platforms such as Facebook and Twitter. This framework was characterized as a method

for helping and enlarging the social procedure of utilizing proposals of others to settle on decisions when there is no adequate individual information or experience of the other options.

Finding the similarity value is another significant viewpoint to prescribe the high rated spots. Here the similarity value is found using Jaccard and cosine similarity. Jaccard similarity is characterized as the number of basic attributes that are isolated by the number of characteristics that exist in any event one of the two articles. While in cosine similarity, the number of common attributes are partitioned by the total number of possible attributes.

The recommendation system helps a user discover products and content by predicting the user's rating of every item and showing them the things that they'd rate highly. Recommendation systems play an important role in helping users find products and content they care about. The most widely recognized use for a proposed framework is positioning items by how much a client might want them. On the off chance that a client is perusing or looking for items, we need to show them the items they might want generally first in the rundown.

2. RELATED WORK

(Babak Maleki Shoja et al.,2019) [5] Dealing with the unstructured data is difficult. Natural Language Processing is used to analyze the natural language data while performing the tasks. In recommender systems, the integration of customer reviews is in either aspect based or aspect aware recommendations system. Collaborative filtering is one of the methods used by the recommender system. Aspect based Recommender system proposed a method called Aspect based Latent Factor model (LFM) and Aspect aware Matrix Factorization (MF)model. FM is used to predict the rating by using the user's and item's information and review the collected data. MF model learns the latent factors from review and rating and effectively combines without having one to one mapping constraint. MF model has aspect rating and assigns a weight to those aspects. Some recommender systems use semantic analysis. After collecting the preference from the user, it starts to predict the sentiment of the user's item and identify the user's most valuable item. The recommender system integrates opinion mining. Opinion mining uses parts

of speech to extract the user's valuable item.

(Xi Shao et al.,2019) [6] The research cluster is mainly focused on building complex information models and has been focused on both healthy subjects and patients. The patients who are affected by diet-related disease. To generate health care recommendations, they have used flowchart answered by users as well as the use of ontologies to maintain menus and recipes to the user. Already available nutrition information services tend to identify the second research cluster it mainly focused on processing the nutritional information, instead of prioritizing the data. By the way, the menu planning scenario treated as an optimization scenario and its recommended as an optimization problem for the health menu generation. Some approaches don't consider the optimization approach, they are based in adhoc heuristic of health care generation. Some of the researchers focused on restaurant menu recommendations, ranking dishes based on medical conditions, and the user's preferences on rank settings.

Some of the other small researchers proposed live personalized medical recommendations by using GPS location, pedometer, and barometer output to calculate the live estimate of the user's daily menu, which is used to rank the menu based on how to fulfill they're the individual needs for nutrition. In mobile platforms, the food recommendation system also developed which not only offers recipe recommendations it also offers dietary menu, individual nutritional recommendations menu and fine-grained food preferences and assumes a basic strategy to calculate the rank. To create a content-based recommendation system, manages a personalized weekly nutritional menu by considering some criteria like vegetarian and non-vegetarian. Suppose the user is non-vegetarian they need the separation of meat and fish, limitation in the recipient of foods. The current research is related to the collection of user information, gathering of nutritional food and their recipes and changing of eating behaviors.

According to data, the travel recommendation system comes under the two approaches GPS based approach and travelogues-based approach. In GPS based approach, obtain

the user's preferences for the travel by utilizing the GPS data. For some more personalized recommendations to users, it proposed a map-based conversational mobile recommender system by integrating GPS data and electronic map technologies. An electronic map is used for outlining the location of the interests of users and producing compatible places to visit. The drawback of GPS based approach is GPS data is not available for each and every place. To overcome the problem, some researchers consider the user's content with trip-related knowledge. In travelogues-based approach, to estimate the location of photos with the tags and features of the places, it categorizes the photos into trip patterns, the traveler can see the photos whatever they prefer to go.

The travelogues-based approach considers the traveler preference, they didn't suggest the place for attraction domain. In order to facilitate travel recommendations, the above-mentioned approaches use latent semantic distribution but the major methods are drawback due to themes of the attraction. Personalized travel recommendations systems interact with the system and their preference to visit. It took the sentiments of attraction themes and exploited multi-model data but failed to overall opinion for an attraction theme into opinions of the users.

3. METHODOLOGY

This section discusses about the proposed Travel Recommendation System, which is based on the user preferences specific to topics related to the theme of places they want to visit. In view of this, the functional architecture of the Travel Recommendation System is shown in Fig.1. To provide recommendations for relevant places for user preferences, the system first collects the reviews given for places. Then calculates the similarity between user preference and reviews, and associates a similarity value for each place which will be used in further steps. Finally, the system tries to capture the correlation between user preferences and places theme spaces for recommendation.

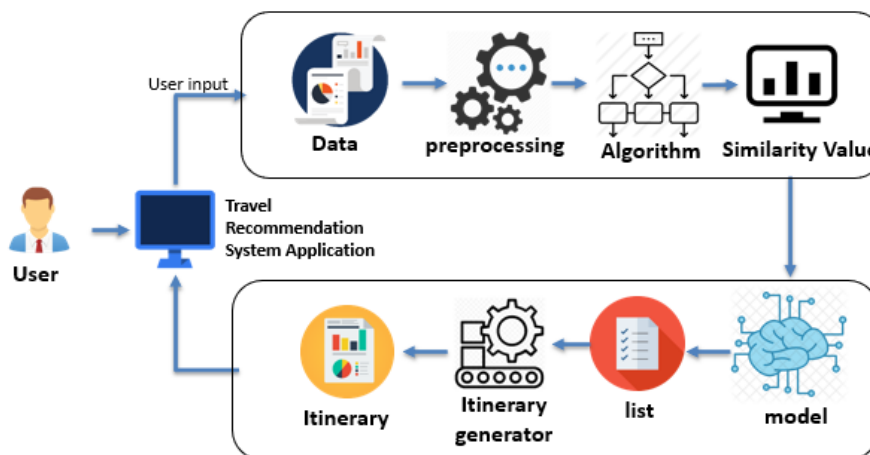


Fig 1: Functional block diagram

Table 1. Sample Reviews

Name	Reviews
Guindy National Park	Parking facility is good. I had a memorable time with my family and friends. This is the suitable place for hangouts. Could see variety of wild life. the park is very clean and neat. Good maintenance. Nice place to visit during the weekends.
Gandhi Mandapam	Very quiet place. You can go in the afternoon for a pleasant nature stay. No noise. Close to nature. You can see deer there in open jungle. It is in Guindy. Beside Guindy children's park
Arignar Anna Zoological Park	If you are planning for a perfect picnic then this is the place in Chennai. It's recommendable that you bring your own food and snacks as the quality of food is not that great inside. But water you have quite a few outlets and quality are ...
Chennai Marina Lighthouse	Beautiful place... View from the top is thrillingly beautiful... The charge to entry in the light house is ₹ 20 for an adult and ₹ 50 for foreigners and ₹ 10 for a child (less than 14yrs of age) ... Nearest railway station is Light House railway station (on Chennai Beach to Velachery rail route) ... Visit this place at the time of evening around 5pm... Overall, it's a clean and nice place to visit with your friends and family... Enjoy the scenic beauty of the Bay of Bengal from the top...
ISKCON	Peaceful place. Best place to visit every morning. Parking, veg restaurant and small shopping place is available. A book store is present inside the temple. Kind and helpful staff. Arrange your own transport at evening times because transit ...
VGP Universal Kingdom	A wonderful place to spend your day. A 1000 bucks' ticket can give you access to amusement, water and marine park. It's very reasonable in that aspect. From good rides, an amazing water park, a small zoo, good picturesque backdrops to good ...

3.1 Review Dataset

The initial step of the work is data collection, in this work review data is gathered from Google maps review. As discussed earlier, the developed Travel Recommendation system needs to use large amount of reviews. So, the utilized review dataset consists of large volume of review data which contains sentiment attributes, places theme and so on. From the review's dataset preprocessing such as cleaning, removal of non-ascii values will be done for further procedure. The sample reviews are tabulated in the table 1.

3.2 Similarity Calculation

The next and major step of the work is calculation of the similarity between user preference which obtained through the user interface of the recommendation application. The system uses combination of two types of similarity techniques namely Jaccard similarity index [7] and cosine similarity index [8] along with consideration of sentiment attributes [9] to match sentences with correct meaning.

For example, good for kids and not good for kids both texts were completely opposite meaning but contains similar words hence consideration of sentiment attribute words are important. The formula for calculation of Jaccard similarity (1) and cosine similarity (2) are given below.

Jaccard Similarity

$$\text{Jaccard Similarity} = \frac{|A \cap B|}{|A \cup B|} \quad (1)$$

Cosine Similarity

$$\text{similarity} = \cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\| \|\mathbf{B}\|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}} \quad (2)$$

The similarity value will be calculated for each review and cumulated value will be taken into consideration for each place's similarity values. The fig 2 shows the rating graph for places based on past travelers' ratings, using this alone is not sufficient for the recommendation since many places have high rating it's not possible to recommend more personalized results to the user.

Fig 2 shows the rating for the sample review dataset, where x axis represents places and y axis represents ratings of the places. Graph clearly shows that many places comes under closer values hence recommendation based on this value alone leads to a list of places which are good to visit but hardly satisfy the user preference and interest over the trip. Therefore, the system needs an additional value which implies the user preference that value is given by similarity index.

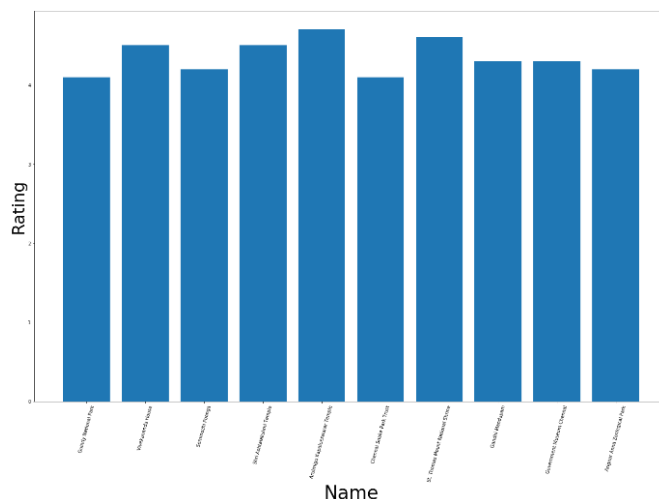


Fig 2: Relation between places and rating

It compares the user preferences with the list of pre-existing reviews and generates a similarity value. The Fig 3 clearly shows there are peaks in graph those peaks represents places which are more similar to the user preference unlike Fig 2, since Fig 2 shows many places in close range rating and it is not personalized to user need. But Fig 3 shows variations in the rating with similarity value, it is easy to recommend the places more relevant to user preference [10].

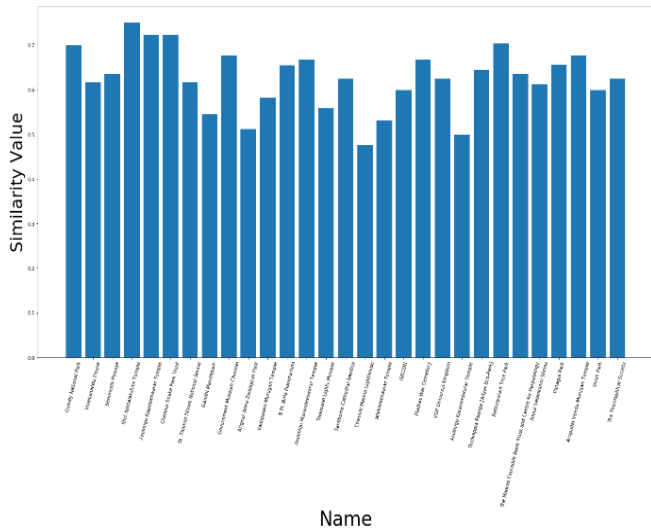


Fig 3: Relation between places and similarity value

3.3 Itinerary Generator

From the similarity value and sentiment polarity value top relevant places will be chosen using neural network. Which uses two values as parameter to choose the place as like the human deciding over number options. After the selection of top relevant places, an itinerary is created considering best time to visit the respective places and the itinerary includes estimated budget consist of travel, stay and food expenses.

4. RESULTS AND DISCUSSIONS

The user expectation is taken as the text sentence and it is compared with each review of the place to provide the similarity score, average value is taken as the similarity value for that place. For example let us consider a user input expectation text which is unstructured “im planning to have picnic and need more spacious place to spend ..as well as silent one travelling with kids”. This is the user input which is not formed properly and this review compared with the review of the places let us consider the first place from the table 1 Guindy National park’s review and produce the similarity value of 0.7 similarly the similarity value for all the places will be calculated. Table 2 shows the similarity value of all places with respect to the given user input.

Table 2: Similarity value

S.No	Name of the Places	Similarity value
1	Shri Ashtalakshmi Temple	0.75
2	Arulmigu Kapaleeswarar Temple	0.724137931
3	Chennai Snake Park Trust	0.724137931
4	Kotturpuram Tree Park	0.703703704
5	Guindy National Park	0.7
6	Arupadai Veedu Murugan Temple	0.677419355
7	Government Museum Chennai	0.677419355
8	Arulmigu Marundeeswarar Temple	0.666666667
9	Madras War Cemetery	0.666666667
10	Panagal Park	0.65625

The above table 2 consist of the top ten places according to the similarity values. This similarity value alone not enough to provide more efficient recommendations, hence sentiment polarity value is used. Table 3 shows the sentiment polarity value for all the places in the table 2, sentiment polarity indicates general goodness of the places. If the place got more positive value then that place got good opinion from travelers irrespective of the preference.

Table 3: Sentiment polarity

S,No	Name of the Places	Sentiment polarity
1	Shri Ashtalakshmi Temple	0.5875
2	Arulmigu Kapaleeswarar Temple	0.361111111
3	Chennai Snake Park Trust	0.4
4	Kotturpuram Tree Park	0.5125
5	Guindy National Park	0.518095238
6	Arupadai Veedu Murugan Temple	0.246944444
7	Government Museum Chennai	0.385
8	Arulmigu Marundeeswarar Temple	0.455333333
9	Madras War Cemetery	0.0296875
10	Panagal Park	0.25

The sentiment polarity used to break the tie to choose the places in the close range the figure 4 shows how to select the places contains close similarity values.

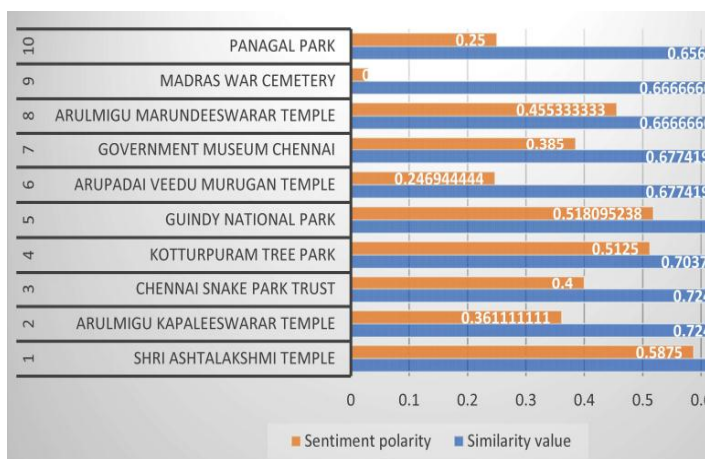


Fig 4: Top places

The Figure 4 shows Shri Ashtalakshmi Temple have the highest similarity value as well as good sentiment polarity, but moving to next positions 4 places falls under close range of 0.7 here sentiment polarity plays vital role. Guindy park and Kotturpuram park will be consider as second and third places even though their similarity value is lesser than other two places. This principle is trained to the neural network to select the places according to the above both constraints to prepare more accurate and personalized recommendation.

5. CONCLUSION

This framework helps the utilizations to plan travel in a like manner based on their inclinations. utilizing cosine similarity and Jaccard similarity travel proposal framework gives practically precise top positioned spots to the client. The framework has numerous preferences as it considers the inclinations of each individual and takes under thought numerous elements to suggest travel sequences like the similarity between the client's prerequisite and previous audits.

For the future research directions, the prefect period for the user requested area will also be included in the recommendation and also it includes mode of travel, hotels to stay and more specific to the user preference by making alternate suggestions. Addition, removal and editing of the plan.

REFERENCES

[1] K. Venkata Raju , M. Sridhar “Sentimental Analysis Inclination A Review” International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC) 8-9 Sept. 2017 IEEE DOI: 10.1109/CTCEEC.2017.8455135.

[2] Periakaruppan Sudhakaran ,Shanmuga sundaram Hariharan , Joan Lu “Classifying Product Reviews from Balanced Datasets for Sentiment Analysis and Opinion Mining” 6th International Conference on Multimedia, Computer Graphics and Broadcasting 20-23 Dec. 2014

DOI: 10.1109/MulGraB.2014.14.

[3] Yung-Shen Lin , Jung-Yi Jiang , Shie-Jue Lee “A Similarity Measure for Text Classification and Clustering” IEEE Transactions on Knowledge and Data Engineering (Volume: 26 , Issue: 7 , July 2014).

[4] Ruth Cobos , Francisco Jurado , Alberto Blázquez-Herranz “A Content Analysis System That Supports Sentiment Analysis for Subjectivity and Polarity Detection in Online Courses” IEEE-RITA (Volume: 14 , Issue: 4 , Nov. 2019).

[5] Babak Maleki Shoja , Nasseh Tabrizi “ Customer Reviews Analysis With Deep Neural Networks for E-Commerce Recommender Systems “ IEEE Access (Volume: 7) 26 August 2019.

[6] Xi Shao , Guijin Tang , Bing-Kun Bao “Personalized Travel Recommendation Based on Sentiment-Aware Multimodal Topic Model “ IEEE Access (Volume: 7) 13 August 2019.

[7] Hangyu Yan , Yan Tang “ Collaborative Filtering Based on Gaussian Mixture Model and Improved Jaccard Similarity “ IEEE Access (Volume: 7) 21 August 2019

[8] H. Mohana , M. Suriakala “Integrated Cosine And Tuned Cosine Similarity Measure To Alleviate Data Sparsity Issues For Personalized Recommendation “ 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions 20-22 Dec. 2018 DOI: 10.1109/CSITSS.2018.8768763.

[9] Shaoliang Shi , Yunpeng Li , Yimin Wen , Wu Xie “Adding the sentiment attribute of nodes to improve link prediction in social network “ 12th International Conference on Fuzzy Systems and Knowledge Discovery 15-17 August 2015 DOI: 10.1109/FSKD.2015.7382124.

[10] Pavel Stakhiyevich , Zhao Huang “Building user Profiles Based on user Interests and Preferences for Recommender Systems “IEEE International Conferences on Ubiquitous Computing & Communications (IUCC) and Data Science and Computational Intelligence (DSCI) and Smart Computing, Networking and Services 21-23 Oct DOI: 10.1109/IUCC/DSCI/SmartCNS.2019.00101.