Tourist Place Recommendation System

Pravinkumar Swamy
Information Technology
Vidyavardhini's College Of Eng.

Sandeep Tiwari
Information Technology
Vidyavardhini's College Of Engg.

Kunal Pawar
Information Technology
Vidyavardhini's College Of Engg.

Prof.Bharati Gondhalekar
Information Technology
Vidyavardhini's College Of Engg.

***Abstract*—Tourism, these days involves mass availability and mass participation of peoples in holidays for enjoyment purpose.But many times, a tourist cannot decide which place to visit, or where to stay,also the cost associated,point of interest of each user and many such factors. So,we are going to propose a website which will recommend places of attractions to the users. For Dataset,we have scrapped holidify.com and collected data of various places and build a data-set around 1 Lakh attributes. Places were recommended based on two factors, similar users and similar contents. Collaborative Filtering was used to recommend places based on similar users and Content Based Filtering was used to recommend places based on similar content. The Machine learning model was trained and pickled so that can be used in Front-end,also data was stored in My-sql Database.**

***Index Terms*—Collaborative filtering algorithm,Content Based Filtering,Web Scrapping,MySql.**

I. Introduction

In recent years, smartphone has been an inalienable part of people's life. People can obtain all the information they need at their fingertips every day.Tourism has become an important sector that has an impact[3] on the development of the country's economy as on this date. Tourism has become one of the biggest industry for gaining wealth in the world many countries like UAE,Japan,USA gain lot of wealth from the tourism industry And the most crucial task, to carry[2] this out efficiently is to plan the trip. Conventionally this planning was done by travel agents . For example, a particular travel company would plan the tours of a particular place, showing the prime tourist attractions. For example a person want to travel from Mumbai to Delhi to the agent will list down all the important tourist places and accordingly plan his trip like first he should visit Akshardham temple then red fort but does every one like to visit temple.this trip won't satisfy everyone. If a user is interested in places like museums and aquariums, then they would have to plan such a trip on their own.

Nowadays, mobile phones are a necessary part of people's life. There is a high amount of rise in the number of mobile computing applications which are centered on people's day-to-day life. One of the areas in which the user can benefit from smartphone applications is tourism traveling.So the main aim of the project is to make an mobile application where the tourist [8] will be recommended some places according his or her interests and based on that he or she can his or her own trip. Choosing a tourist destination from the information which is top rated by other user, is one of the most difficult tasks for tourists who are making travel plans, both before and during their travel.There are lot of websites and applications that provide the top rated places from various cities which may not be the place user likes.Also,some websites based on nearest location,try to recommend places to the nearest location of the user.So,to overcome this, similar contents and similar set of users need to be analyzed and accordingly build an optimized recommendation system.

II. BASIC TERMINOLOGIES USED

A. Collaborative Filtering Algorithm

Collaborative Filtering is used in recommendation systems. There are two types of collaborative filtering systems, Item-based and Content-based Collaborative Filtering.[6] In general, Collaborative Filtering is the process of filtering information or patterns using techniques involving collaboration among multiple agents, viewpoints, data sources, etc. There are various reasons why Collaborative Filtering is used like it can be used with large datasets. As compared to content-based filtering, the accuracy of Collaborative Filtering is much more.
Using Collaborative Filtering, the whole process of recommending tourists places can be divided into three steps:
a. The representation of tourist information, wherein the travel style and reviews of the user are checked.
b. Next, the similarity of tourists can be computed according to the visiting history[7] data and the Collaborative Filtering algorithm presented by us.
c. The generation of attraction recommendations. On the basis

of similarities with other similar users, the top Attractions, Restaurants, and Hotels are recommended.

B. Content Based Filtering Algorithm

Content-based filtering methods focus on both the profile of the user's preferences and the item description for recommending items that are similar to the items that are rated very high by others in the past.[1] This makes it easier to scale to a large number of users. The model can capture the specific interests of a user, and can accordingly recommend items where in other users are also interested in. Disadvantages are Since the feature representation of the items are hand-engineered to some extent,[4] this technique requires a lot of domain knowledge. Therefore, the model can only be as good as the hand-engineered features. The model can only make recommendations based on existing interests of the user. In other words, the model has certain limitations to work on the users' existing interests.

C. Web Scrapping

Web scraping is a technique of extracting useful data from large websites.Web scraping helps to build a dataset which can be used by algorithms for further computations.Web scrapping is illegal to a certain extent.In holidify.com various attributes such as placIn Python,web scrapping is performed through BeautifulSoup Library using requests method. Also,Selenium can be used for performing web scrapping ,when the data is rendered in the client side elsewhere python can be used.

III. RELATED WORK

In the literature, there are several studies about users' habits and routines in a city using digital traces. Some of them analyzed GPS [11]data and cellular footprints of users to understand, for instance, their usual trajectories. A

	Google Goggles	**GuideMe**	**Our Proposed System**
Input	Image	Image/ Text	Image, Text
Output	i) Recognize image ii) Give Name iii) Web Search Results	i) Recognize image ii) Give Name iii) Give Details	i) Recognize image ii) Suggest nearby places in the given city based on users' interests.
Image Category	Any Image	Images of Monuments or famous buildings	Images of 1. Hotels 2. Attractions 3. Restaurants

Fig. 1. Comparison of various similar systems for place recognition.

Recommendation System is a personalization tool that offers users with a list of items that best fit their individual taste. A Recommendation System analyses the available data and recommends the items according to user's interest. The more the system understands the [9]interest of the users, the better recommendations, it can perform.

Recommendation System was proposed to cope with the problem of information overload. There are so many traditional methods available for recommending the items to the users. For example, we have content-based filtering, Collaborative based Filtering, and knowledge-based filtering but some of these methods have some shortcomings. In recommendation is totally based on the evaluation in which we will ask the questions to the user and observe the answer. In the past decades, many studies have been conducted to develop methods for recommender systems and improve their accuracy. In, the authors proposed a fuzzy c-means approach for a collaborative user-based filtering system. They used the MovieLens datasets to compare the different techniques of clustering. In , the authors investigated the applicability of the cluster ensemble approaches for recommender systems. They utilized k-means and Self-Organizing Maps (SOM) [10]as baseline clustering techniques, and the multiple clustering ensemble technique to combine the results of clusters.

In the authors proposed a keyword-aware service recommendation method, named KASR, to indicate users' preferences and generate appropriate recommendations on MapReduce for big data applications. In , Lee et al. proposed an adaptive recommendation algorithm, ACFSC, that is focused on scalable clustering. They addressed the problem of scalability by composing neighborhood based on reducing time complexity. They also addressed the problem of sparsity by making items' and users' feature vectors incrementally learning. In , the authors proposed a typical model that integrates collaborative filtering, clustering techniques, and social network analysis (SNA) in order to enhance the perdition accuracy results in recommender systems. Their model uses SNA to identify the people who are most influential on social networks and then uses these people to conduct clustering analysis. Following that, the model focuses on cluster-index collaborative filtering to make accurate recommendations. Additionally, Tian et al. developed a new method for improving recommendation quality by formalizing trust relationships in online social networks. In [12], authors proposed two varieties of algorithms for developing an effective recommender system. The first one uses the improved kmeans clustering technique while the second one uses the improved k-means clustering technique coupled with principal component analysis to enhance the recommendation accuracy

IV. METHODOLOGY

A. Place Recommendation System

The system has a MySQL database, a Web Server, and a front-end application. There is a GUI for the user which takes basic information from the users and saves the

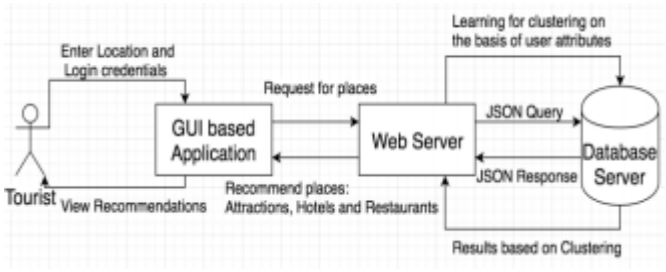

Fig. 2. Recommendation system architecture.

answers given by user in the database and then it analyze the answers given by the users and upon that analysis it recommends new places to the users using various algorithms.

1) Data Cleaning and Data Pre-Processing: Data collec by scraping and downloading data from trip-advisor has lo missing values and lot wrong data. hence it was importan a cleaning and removing redundant or missing data from datasets. Moreover, it was also necessary converting all the data into categorical data.

2) Recommendation by content based and Collaborative filtering: Initial collected data about the user are fed into these two algorithms and based on the input given new places are recommended to the users in order to increase the efficiency of the algorithm the size of the dataset is increased after checking the efficiency of the algorithm we have found out the efficiency as 78%

Formula for collaborative filtering is as follow:

$$r_{ij} = \frac{\sum_k Similaries(u_i, u_k) r_{kj}}{number\ of\ ratings}$$

Without knowing anything about items and users themselves, we think two users are similar when they give the same item similar ratings . Analogously, for Item-based CF, we say two items are similar when they received similar ratings from a same user. Then, we will make prediction for a target user on an item by calculating weighted average of ratings on most X similar items from this user. One key advantage of Item-based CF is the stability which is that the ratings on a given item will not change significantly overtime, unlike the tastes of human beings.

Formula for Content Based Filtering:-
By applying the definition of similarity, this will be in fact equal to 1 if the two vectors are identical, and it will be 0 if the two are orthogonal. In other words, the similarity is a number bounded between 0 and 1 that tells us how much the two vectors are similar.

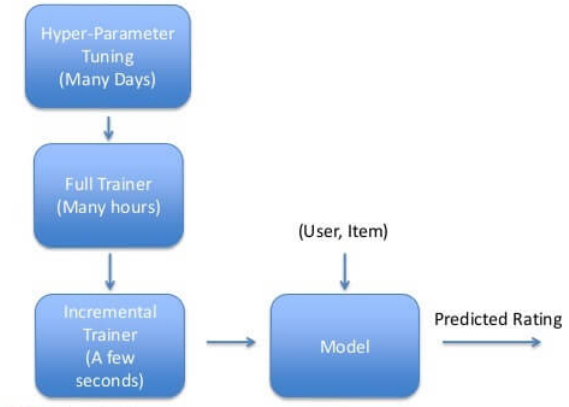

Fig. 3. Collaborative filtering Architecture

$$similarity = cos(\theta) = \frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|} = \frac{\sum_{i=1}^{n} u_i v_i}{\sqrt{\sum_{i=1}^{n} u_i^2}\sqrt{\sum_{i=1}^{n} v_i^2}}$$

V. CONCLUSION

As a result of the rapid growth in the numbers of tourists who are travelling, the Internet is becoming increasingly populated with travel information. When selecting their preferred destinations before or during their travel to an unfamiliar city, tourists can therefore easily be overwhelmed. Destination recommendation systems (DRSs) are recognised as a valuable decision-support tool for online travel as well as for tourism marketing. A model-based DRS and an ensemble-based DRS with an adaptive, responsive and interactive user interface has been successfully developed and implemented. The DRS aims to assist tourists plan before or during their visit to an unfamiliar city. Both technical and practical aspects were considered, including data sparsity, scalability, transparency, system accuracy, usability and user acceptance.

VI. FUTURE WORK

The main concentration of the content-based systems is to recommend things to the users based on their profile, by generating the user specific results which may fascinate the user. In the area of tourism, this may be considered as an essential issue of travel recommender systems. Few recommender systems try to promote new spots or new activities, which makes the recommendations ineffective. A good travel recommender system provides broad suggestions to the users and allows them to choose their routes with activities. The list of user's attractions should be synchronized with rating limit in order to maintain the quality of recommendations. The usage of multiple techniques to filter the activities for the recommendation generation is a new scope in this domain. The clustering of items into groups with similar attributes is an effective mechanism in the recommendation process to build a good list of suggestions by utilizing every cluster of items for a user specific proposition of their tourist interests.

A unique feature of tourism domain is the area where the recommenders have been used, as it adapts to the users and helps them through generated suggestions in different places and in different moments. These travel recommender systems have begun to fuse context aware mechanisms with it. The accomplishment of this methodology is because of the far reaching utilization of Smartphones. Numerous tourism recommenders keep running mobile devices, so the location of the user may be utilized in the filtering process of items to be demonstrated. In contrast to the existing systems, the enhancement of location consideration has to be modified. The present location of the user is important. But along with it, the significance of other spots visited already should also be analyzed and used in the recommendations. In tourism recommender systems, different highlights are considered as related data for an event. For example, the present climate is analyzed to choose appropriate outdoor or indoor activities to be recommended.

VII. RESULTS

	title	ratings
0	Marine drive| Mumbai	4.5
1	Gateway of India	4.4
2	Colaba Causeway	4.3
3	Juhu Beach	4.4
4	Siddhivinayak Temple	4.3
5	Haji Ali Dargah	4.3
6	Elephanta caves	4.3
7	Girgaum Chowpatty	4.3
8	Bandra Worli Sea Link	4.4
9	Chhatrapati Shivaji Terminus	4.4
10	Prince of Wales museum	4.3
11	Essel World	3.6
12	Fashion Street| Mumbai	3.8
13	Film City Mumbai	3.8
14	National gallery of Modern Art| Mumbai	3.8
15	Sanjay Gandhi National park	3.8
16	Global Vipassana Pagoda	3.8
17	Mumbai Zoo	3.8

Above is the Dataset obtained after scrapping holidify.com and also performing Data-processing on the same Data.The dataset contains record of around 1 Lakh places all over the world.

```
Marine drive| Mumbai
Juhu Beach
Girgaum Chowpatty
Water Sports in Mandwa Beach
Gorai Beach
Marve Beach
```

After Performing Content Based filtering ,the above result were obtained which states that , if ”Marine drive” is liked by someone,then the places that matches the content of ”Marine Drive” will be recommended which is proved from the above result. After training the model again and again the accuracy was found to be 78percent and after building the machine learning model

REFERENCES

[1] Badaro, G., Hajj, H., El-Hajj, W., Nachman, L., 2017. A hybrid approach with collaborative filtering for recommender systems, in: 2017. 9th International Wireless Communications and Mobile Computing Conference (IWCMC). Presented at the 2017.

[2] Al Shalabi, L., Shaaban, Z., 2017. Normalization as a Preprocessing Engine for Data Mining and the Approach of Preference Matrix, in: International Conference on Dependability of Computer Systems, 2017. DepCos-RELCOMEX ’17. Presented at the International Conference on Dependability of Computer Systems, 2017. DepCos-RELCOMEX.

[3] Abbaspour, R.A., Samadzadegan, F., 2017. Time-dependent personal tour planning and scheduling in metropolises. Expert Systems with Applications 38,12439–12452.

[4] Ardissono, L., Goy, A., Petrone, G., Segnan, M., Torasso, P., 2018. Intrigue: Personalized Recommendation of Tourist Attractions for Desktop and Hand Held Devices. Applied Artificial Intelligence 17, 687.

[5] Anacleto, R., Figueiredo, L., Almeida, A., Novais, P., 2018. Mobile application to provide personalized sightseeing tours. Journal of Network and Computer Applications.

[6] Cheng W, Yin G, Dong Y, Dong H, Zhang W (2018) Collaborative Filtering Recommendation on Users’ Interest Sequences. PLoS ONE 11(5): e0155739. https://doi.org/10.1371/journal.pone.0155739.

[7] Chu. W, Park. T.S. et al, (2018), ”Personalized Recommendation on Dynamic Content Using Predictive Bilinear Models” Yahoo! Labs.,2821 Mission College Blvd, Santa Clara, CA95054.

[8] Dang, T. A., Viennet, E. (2018, January). Collaborative filtering in social networks: A community-based approach. In Computing, Management and Telecommunications (ComManTel), 2018 International Conference on (pp. 128-133).

[9] Farhan Hassan Khan, Saba Bashir. Collaborative Filtering based Online Recommendation Systems: A Survey. International Conference on Information Communication Technologies, At Karachi, Pakistan, Volume: 7th (2018).

[10] Rana, C., Jain, S. K. (2018). Building a Book Recommender system using time based content filtering. WSEAS Transactions on Computers, 11(2), 2224-2872.

[11] Kyo-Joong Oh, Zaemyung Kim, Hyungrai Oh, Chae-Gyun Lim, and Gahgene Gweon, “Travel Intention-Based Attraction Network for Recommending Travel Destinations”.

[12] Yuan Q, Cong G, Ma Z, et al. Time-aware point-of-interest recommendation[C]// International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 2013:363-372.