Bon Voyage

A Trip Recommendation System

Ankit Garg
ECEN
Texas A&M University
College Station
ankitgarg304@tamu.edu

Harmeet Sethi
ECEN
Texas A&M University
College Station
harmeet_sethi@tamu.edu

Mayuresh Hooli
CSCE
Texas A&M University
College Station
mayurhooli@tamu.edu

Praveen Venugopal
CSCE
Texas A&M University
College Station
razorvine@tamu.edu

ABSTRACT

Recommender systems have become an active research topic during the last two decades, thus giving rise to several approaches and techniques. They have also become increasingly popular among practitioners and used in variety of areas including movies, news, books, research articles restaurants, garments, financial services, insurance, social tags and products in general. Tourism is an important sector for economic development and a potential application area of use of recommender systems.

In this paper, we present an overview of one stop recommender for both travel destination and modest accommodation ideal for novice travelers. Section 1 will give a brief introduction, Section 2 illustrates Background/Related Work in this domain, in section 3 we give our Proposed Solution/ Methodology, Section 4 is our Evaluation Criteria, Section 5 is the Conclusion and Section 6 are References.

KEYWORDS

vector, vacation, history, recommender, travel, housing, accommodations, content-based recommendation, collaborative filtering

1 Introduction

The modern life is hectic. People need respite from their day-to-day work as the lifestyle gets monotonous after a certain point of time. One of the best remedies from this hectic monotonous life is to get away from it all for some time. That is why a vacation to a different place is necessary.

The main goal of a vacation is to relax. As such, the idea of a perfect vacation differs from person to person. Some people may like to travel the World alone, explore different things, while some others prefer to travel alone or enjoy with their friends. Some people prefer going on short trips as they are rejuvenated after a shorter period, while some people may go on a long trip which at time may last for a month.

Humans are creatures of habit. Thus, people may prefer certain types of cities to other cities. If a person visits a city for its beaches, chances are that the person will visit other cities that have beaches. If any person is interested in hiking around a hilly region, he will look for places with lots of hills. Another important factor for people to visit places is the region of the location. People tend to prefer certain regions on the planet over the other regions. Some people would prefer to travel around Europe for the 19th Century Architecture, while some others would like to travel to the Middle East for the Desert Safari and the sprawling concrete jungle. Another important aspect for a vacation is the time of the year. Not all places around the world are good for travel all the time. As an example, London isn't a good place to visit in Winter, but once Summer comes along, it has a lot of attractive features which one may consider for travel.

There are many other factors that directly impact the decision of the trip. The travel expense and the total cost of the trip are two more important factors people would consider while planning. Since there are a lot of variables involved, planning a vacation is extremely difficult and can be a cause for stress as the amount spent for the vacation is huge.

Recommendation Systems, thus, could be a valuable help while preparing a trip or searching a service among many destinations, numerous attractions and activities. The use of Recommendation

Figure 1: The 66 Travel Destinations across the World

System could help tourists save time and energy while planning a vacation that match their preferences and interests.

Our solution to this exasperating situation is to assist tourists by recommending best possible destination cities for their travel along with suitable housing accommodations that are based on based on user preferences. To do this, we use the Airbnb Data from 66 of the most frequently visited cities and the travel history of 300 users. Based on the user travel history, we create a theme for every user based on the most commonly traveled places by each user. Once we get this, we identify the similar users to this user through multiple features and create a suggestion pool of cities where our target user has not travelled and recommend most similar cities. Now using Airbnb data set for these selected cities and certain user features, we identify best places for the stay. We strongly feel by merging our proposed solution with existing trip planner ecosystems, companies can provide a more personalized experience to their users and improve their valuation.

2 Related Work

Travel Recommendation has bee

In tourism field, existing recommender systems aim to match the characteristics of tourism and leisure resources or attractions with the user needs. Unlike Recommendation Systems in other domains, current travel recommendation systems combine multiple types of recommendation techniques for example SigTur/E-Destination employs many recommendation techniques, such as the use of stereotypes (standard tourist segments), content-based and collaborative filtering techniques, personalized and ontology-based approaches. However, the special characteristics of tourism items generate continuous appearance of new problems and the need to develop new techniques.

Critically, these tourism recommendation systems have no implementational proof due to lack of proper datasets. Even though housing recommendation for Airbnb has been worked out in notable Kaggle competitions and have in fact given resounding successful results, but its success is specific to the company and does not generalize to travel planning. For instance, the Airbnb dataset is divided into 66 sets each corresponding to a specific city and recommendations are in place of a city only. This can't be a holistic travel recommender, since it conveniently ignores city

features like terrain, historical or religious significance, culture and lifestyle, etc. With our proposed recommender model, we will consider user's thematic preferences while finding potential travel destinations and further recommend housing accommodations.

3 Proposed Solution

3.1 Data Sets

Here we are considering two types of datasets:

3.1.1 Simulated Dataset based on User Vacation History. We randomly populate vacations and extract thematic users from the random pool. Themes represents the user preferences i.e. in which time of year user prefers to go on vacation, his feature of attractions (beach, nightlife, modern architecture etc.), trip expenses (how much money user willing to spend for his trip) etc. Considering all these themes, we create a vacation vector denoting a single vacation for a specific user. Likewise, User Travel History includes collection of user vacation vectors v1, v2, v3 and so on.

3.1.2 Airbnb Dataset. Inside Airbnb is an independent, non-commercial set of tools and data that allows us to explore how Airbnb is really being used in cities around the world. By analyzing publicly available information about a city's Airbnb's listings, Inside Airbnb provides filters and key metrics so you can see how Airbnb is being used to compete with the residential housing market. Using this dataset, we recommend housing accommodation to the users on their selected city. We extract preferences (Expense category and Stay Length) from user history and filter Airbnb housings w.r.t those preferences.

3.2 Methodology

We believe that a single stop recommender for both travel destination and modest accommodation would be of great assistance for novice travel planners. Based on this approach, we divide our implementation in two phases as shown below:

3.2.1 Phase I: User based Collaborative Filtering. Collaborative filtering is based on the similarity between the users. This filtering recommends items appreciated by users who have previously made choices like those of the current user. Collaborative-filtering RS determines the utility of an item based on the feedback (ratings, likes.) of similar users. The idea here is not to focus specifically on the new item that would be likely appreciated by the user, but to look to which items have interested other users who are close to the current user. Collaborative filtering technique start by building a database (user-item matrix) of preferences for items by users. Then it matches the items with the users based on the outcome of the similarity test between their profiles. This approach is the most mature, the most common and the most referenced in literature. It is the most implemented since it often gives good results and does not require much data preparation to start with.

In phase I, our main aim is to recommend top cities to the target user based on its prior vacation history. The prior history consists of multiple vacation vectors for each user. We extract tags from vacations and create User Vector from weighted sum of tags. To determine the other users who shared same preferences, we calculated their cosine similarity score. Then we create a suggestion pool of cities from the history of similar which were not visited by target users. From this pool, we return the top 5 highly recommended cities.

3.2.2 Phase II: Content Based Recommendation System (RS). Content-based RS are content oriented, which means the content of users interests and the content of the features of items play an essential role in the recommendation process. RS using content-based filtering approach base their evaluations on ratings given by a user on a set of items Unlike Collaborative filtering method, content-based filtering determines which items are likely to be useful or interesting to a given user by analyzing the content or the descriptions of items. Most content-based recommendation systems identify items like those that a given user has appreciated. This approach is based on the similarity between the different objects: the objects are recommended to users based on their feedback on similar objects.

Here, we are using TF-IDF Vectorizer and Cosine Similarity approach to extract user preferences and filter housing based on these preferences. Initially we choose a random housing from our filtered list. "Summary" or "Description" for each housing contains descriptive text. After removing the stop words, we apply TF-IDF approach to vectorize the text. Then we implement cosine similarity on this vector representation to find similar housing.

Then we implement cosine similarity on this vector representation to find similar housing.

4 Evaluation and Analysis

The evaluation for our recommendation system is split into two phases.

4.1 Phase-1: City Recommendation

For a given user vector, we choose top 6 similar users provided by our recommendation system and document similarities between their prior travel destinations. Similarity is based on parameters like season of the year, feature of attractions, user travel and expenditure preferences. Figure 1 depicts the cosine similarity between the selected user and all other users in the dataset.
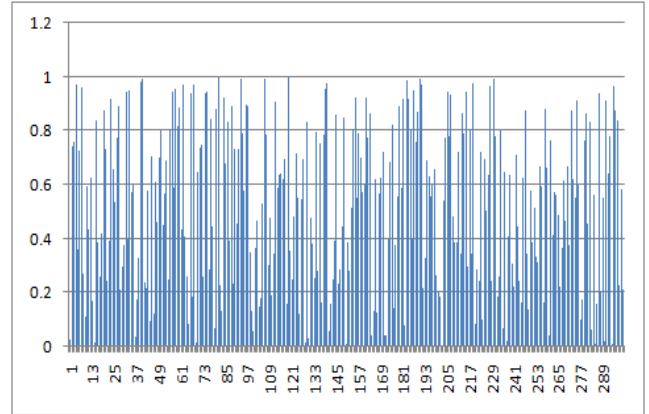

Figure 2: Cosine Similarity vs Users

Target user history:
['Vienna', 'San Francisco', 'Berlin', 'Vancouver']

Recommended user's histories:
['Vienna', 'Geneva', '**Dublin**', 'Melbourne']
['Copenhagen', '**Dublin**', 'Brisbane', 'Paris']
['Sydney', 'Mallorca', 'Berlin', 'San Francisco']
['Sydney', 'Hawaii', 'Mallorca', '**Dublin**']
['Tasmania', '**Dublin**', 'Cape Town', 'Milan']
['Hawaii', 'Cape Town', 'Berlin', 'Canberra', '**Dublin**']

Suggestion pool contains highest occurrences for **Dublin**, which is then fed as input to phase II of the models.

4.1 Phase-2: Accommodation Recommendation

Once we have a target city, we find top 4 housing accommodations which consider thematic preferences of the selected user. Figure 2 gives a plot showing the recommended scores for these selected housings.
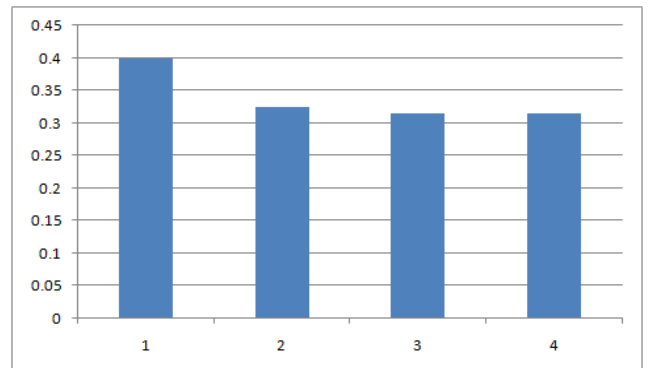

Figure 3: Recommended Score vs Housing

Dublin

Recommending 4 products like Spacious City Centre Apartment

Description: Ideally located in one of Dublin's most popular areas, this home is a **pleasant stroll into the main streets**. The neighborhood is perfectly located to provide you with...

Recommended: **Georgian Luxury**

Description: Georgian Period 3 bed home in central zone of Dublin. **Walk everywhere**. Nestled on a historical leafy street in Dublin, this traditional Georgian home is a pleasant s...
(score:0.3995059096180262)

Recommended: **Central Penthouse Apartment along the River Liffey**

Description: Our apartment is in a great location as the city center is on your doorstep and is only a few minutes' **walk to all the major sight**s including Guinness Storehouse, T...
(score:0.32391805845147104)

Recommended: **Stylish and Modern Two Bedroom Dublin Apartment**

Description: My lovely and modern 2-bedroom apartment in Fitzwilliam Quay, Dublin 4, is the perfect place for a group of between 2 and 4 to stay whilst visiting this beautiful c...
(score:0.31491069768084357)

Recommended: **Central location next to St. Stephen's Green**

Description: The neighborhood is perfectly located to provide you with all you need particularly those great "local" feels that many of us seek while abroad. It's the ideal place...
(score:0.3141840138813951)

5 Conclusion and Discussion

Planning a trip is a complex decision process taking in account all the variables on tourism items and users, recommendation approaches need more involvement and more use of all the features of the items. For Example: the opening and closing times of the attractions, or the time needed to go from one point of interest to another. More importantly, the user's characteristics his experiences, his behavior and his interactions on social media are to be taken in each steps of recommendation building. Therefore, to catch user needs and generate a satisfying recommendation especially in tourism domain is a hard work, and for better results will be more interested in data models in future work to have a better understanding of the tourism and travel items.

In this paper, some of the approaches used to produce tourism recommender systems were overviewed. Firstly, the classical approaches used generally in recommender system, mainly divided in two categories: User based collaborative filtering methods and Content-based filtering. Secondly, the new approaches, which integrate, personalized recommendation, the context and the semantic knowledge about the users and items. In addition, the steps to build a user profile within a recommender system to identify the items matching their preferences were presented. Fruitful attempts were made to simulate vacation history for 300 users with emerging theme for each user.

In the future works, the focus will be to include additional factors such as circumambience, family size, gender etc. while recommending housing. Moreover, the cold start problem in phase 1 can be fixed by implementing Item based Collaborative Filtering. Likewise, we can include GIS data to provide housing location and its distance from the airport.

REFERENCES

[1] Context-Aware recommender systems Gediminas Adomavicius, Alexander Tuzhilin in F. Ricci, et al. (Ed.), Recommender Systems Handbook, 2011, pp. 217–253.
[2] Moreno, A., Valls, A., Isern, D., Marin, L., & Borràs, J. (2013). SigTur/e-destination: Ontology-based personalized recommendation of tourism and leisure activities. Engineering Applications of Artificial Intelligence, 26, 633–651.
[3] Kaggle Airbnb content-based recommendation system https://www.kaggle.com/rdaldian/airbnb-content-based-recommendation-system.
[4] Tourism Recommender Systems: An Overview of Recommendation Approaches Larbi Kzaz, Dounia Dakhchoune, Dounia Dahab, February 2018- International Journal of Computer Applications (0975 – 8887).
[5] About Inside Airbnb: Adding data to the debate http://insideairbnb.com/about.html