Scientific Computing with Python: Description of Projects

Kamini Garg
University of Applied Sciences and Arts of Southern Switzerland, Manno
Email: kamini.garg@supsi.ch
SkypeID: kaminigarg.86

Istvan Lengyel
EIT, New Zealand
Email: ilengyel@eit.ac.nz

1 Project 1: Time Distance-based Location Recommendation System using FourSquare Data

In this project, you need to create a software that recommends new location based on time-distance i.e. the location should be recommended based on the time taken to reach that particular location. The following data will be provided to build your recommendation system (Please refer Readme.txt file provided along with data set for detailed description):

- Location history of users
- Social network of users

More specifically, you need to answer following questions in this project:

1. Recommend a list of 10 unvisited locations for any userID as input. Each unvisited location is scored based on the travel time from the last visited location of the user.
2. Recommend a list of 10 unvisited locations for any userID as input. Each unvisited location is scored based on the travel time from the last visited location of the user and must be a place visited by any 2-hop friends of the given user.
3. For a given userID, starting point and end point; provide a travel plan that contains at max 10 locations with shortest travel time to visit them. Further, also show this travel plan on Google maps.

2 Project 2: IMDB Software of Hollywood Actors and Actresses

In this project, you need to create a user-friendly software that stores and extracts information about the top 50 popular Hollywood actor and actresses http://www.imdb.com/list/ls053501318/.

Your IMDB software should provide following functionality:

1. List of all available actors and actresses
2. About the actor/actresses
3. All time movie names and years
4. Awards to actor/actresses in different years
5. Movie genere of actor/actresses

6. Average rating of their movies (overall and each year)
7. Top 5 movies, their respective years and genere

Please remember, your software needs to to be user-friendly and well-formatted.

3 Project 3: Interest-based Location Recommendation System using FourSquare Data

In this project, you need to create a software that recommends new location based on user interests. Interests are reflected through the type of places they visit like food, travel, bar, shopping etc. To build your recommendation system, the following data will be provided (Please refer Readme.txt file provided along with data set for detailed description):

- Location history of users
- Semantic meaning (FourSquare Category ID) of VenueID i.e. bar, park etc.

More specifically, in this project you need to address following questions:

1. Recommend a list of 10 unvisited locations for any userID and CatagoryID as input. The new suggested venues must be of similar category to the one given as input.
2. Find the list of top 10 most similar users (i.e. their interests match) in the entire data set for any userID as input.
3. For a given set of 5 userIDs and their randomly selected locations from data, recommend a place where these 5 people can meet. The recommended place must be relatively close to each one of the 5 userIDs.

4 Project 4: BIXI: Toronto Bike Station Live Feed

In this project, you need to process the live feed of BIXI in Toronto city and answer the following questions: https://www.bixi.com/en/page-27.

1. Given the current location of a person and number K as input, find K-nearest bike stations based on available bikes.
2. Given the current location of a person who has a bike and number K as input, find K nearest bike stations where docks are available.
3. Given a source and destination location in Toronto, present the route on Google maps of a person using BIXI bike. Please remember that for traveling from source to destination, you can only use your foot and BIXI bike from live feed.

5 Project 5: Twitter Data Analysis

In this project, you need to fetch the data from Twitter of your favorite event e.q. #football by creating an account and utilizing the twitter API and answer the following questions:

1. Derive the sentiment of each tweet using Python module (no need to create your Algorithm)

2. Top 10 hash tags and users based on their number of tweets in your data set.
3. Get the followers of a given twitter user from your acquired data set.
4. Given a twitter user, obtain the tweets and profiles of all followers of the user and show it.