

Scientific Computing with Python: Description of Projects

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.

Project 2: 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 from this list <http://www.imdb.com/list/ls053501318/>.

Your 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 genre of actor/actresses
6. Average rating of their movies (overall and each year)
7. Top 5 movies, their respective years and genre

Please remember, your software needs to be user-friendly and well-formatted. You may use existing API's to help obtain data.

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 CategoryID 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 userID's 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 userID's.

Project 4: Transportation Station Live Feed.

In this project, you need to process the live feed of a transportation provider and answer the following questions: <https://bikeshare.metro.net/about/data/>.

1. Given the current location of a person and number K as input, find K-nearest station based on available devices, e.g. bikes, scooters, etc.
2. Given the current location of a person who has a bike/scooter and number K as input, find K nearest bike/scooter stations where docks are available.
3. Given a source and destination location for example Los Angeles, present the route on Google maps or another mapping product of a person using Metro bike. Please remember that for traveling from source to destination, you can only use your foot and station live feed.

Example data feed <https://bikeshare.metro.net/about/data/>

Project 5: Social Media Data Analysis

In this project, you need to fetch the data from a social media provider, example X, Instagram, Facebook, etc. of your favorite event e.g. #football by creating an account and utilizing the social media API and answer the following questions:

1. Derive the sentiment of each post using Python modules (no need to create your Algorithm)
2. Top 10 tags and users based on their number of posts in your data set.
3. Get the followers of a given social media user from your acquired data set.
4. Given a social media user, obtain the posts and profiles of all followers of the user and show it.