**Network Based Recommender System**

Human generated network data consists of plenty of information which can be classified as event id, event classification, device id, network data. Network data consists of bssid information, network name, person name, locality information etc.

All this information can be helpful to decide customer behavior, user/uses classification based on network provider/ location/ social interest.

Using locality and area information, by using dependency graphs one can predict the transport uses of a customer or interest of ventures (shopping malls, tourist places etc.) which could be useful for various service providers or companies for their products.

**Dataset Provided:**

**1. WiFi data:**

1. Timestamp
2. Range
3. Bssids
4. Req type
5. Device name
6. Latitude
7. Longitude

**2.** Word\_list

**Dataset Analysis:**

**(i)** From dataset 2 of wordlist we can have entity detection which comprise of noun entities having attributes as person name, geopolitical entity, place name, city, state and movie etc.

Where project titled as **“proper\_noun\_classification”** proposes a solution where we have used recurrent neural network approach to predict the named entities by training a model. On the above mentioned project we have predicted a word as entity named as drug, place, movie or person name where our training model initially performed at 85% accuracy which could be improved further.

**(ii)** From dataset 1 of wifi related data as we explained it includes many fields. Timestamp gives us month, year, day, hour, minute, second where all these are important for data clustering based on season or spent duration for a stipulated event. Range of device can be further classified as probably category of device used by a user like mobile/ pc/ broadband connection etc. Bssids can be used for reverse lookup of location related fields and information if tracable or for device general identification. Latitude and longitudes are most important feature which is used to track location and related geo data which can further included with time for current whether information.

So after getting location from latitude, longitude we can extract following information from it:

By getting location/ city/ state of user we can use timestamp for location to get weather information for the time and traffic between two points.

**Uses of analytics done:**

After reviewing all information, we have following features:

1. Timestamp (year, month, day, hour, second)
2. Device information (type, information)
3. Latitude and longitude (useful for mapping on google map)
4. Location (city, state, country)
5. Weather information
6. Traffic related data
7. Entity recognition as place/ person name/ movie name/ theater/ club etc.

**Report:**

1. Using above features if a user/user-group or device(s) visits through multiple places and uses wifi data then we can make a trace for the customer on map and can recommend a user for offers related to visited places which will be a recommender system and it better works with rich set of data and its accuracy depends on size of data which is proportional.
2. Many other types of clustering and uses based on features can be included for data which can be provided to client as api service so it can be visualized.

Skillset Required: Machine learning/analytics modeling, Rest API’s, Python, Django/MongoDB, Postman, Anaconda etc.