Introduction/Business Problem

Point of interest (POI) recommendation helps people find useful and interesting places, and location-based social networks (LBSNs), for example, Foursquare, promote the service rapidly. Foursquare provides the check-in, like, tip, and other functions for users, which makes it possible to construct a list of users' like and even the preference of each user. Therefore, an accurate personalized POI recommendation can be made according to users' tastes. Some preceding studies suggest the recommender system based on user preference, social influence and geographical influence (Ye, Mao; Yin, Peifeng; Lee, Wang-Chien; Lee, Dik-Lun (2011-01-01). Exploiting Geographical Influence for Collaborative Point-of-interest Recommendation. Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information). User preference is determined by the similarity in behavior of other users and social influence by the taste of mutual connections. Geographical influence is affected by the distance between the target place and the typical places that a user frequently visits. Other context factors also play an essential role in providing the suggestion, such as location, time, weather, reachability, which usually make the recommendation complex. The recommendation system not only lifts the burden of users in searching location information but also provides values for third-party companies to advertise and to forecast service demand.

Data

In this study, a simple recommendation system for places in Chicago will be proposed. The venue information is acquired from the Foursquare API. Due to the restriction of free account in the Foursquare, one user with a profound profile is selected, and the taste of the user is analyzed by a user-created like-list rather than by the user's friends' taste. The geographical influence is simplified by dividing Chicago into 246 neighborhoods, and hence a nearby location can be selected.

First, a list of neighborhoods in Chicago on Wikipedia is scraped by the library BeautifulSoup, and a dataframe of 246 neighborhoods is constructed. Then, the latitude and the longitude information are searched by the library geopy.geocoders, and the geographical information of 155 of 246 neighborhoods are found. The Foursquare API explores the venues information as JSON format within the limit of each neighborhood. Venue name, category, and ID within each neighborhood are collected in a dataframe. Each neighborhood is analyzed, and the top 5 most common categories for each neighborhood are obtained. The venues classified as a coffee shop are selected, and the rating for each coffee shop is acquired by the Foursquare API. The list of the coffee shops is sorted by the rating score, and the coffee shop with the highest score is then explored into the tips. The tip with the highest agree count is filtered, and the user who left the tip is selected as the target user. The user is then explored, and a like-list is used to find his/her taste. The items in the list are categorized, and the user's favorite categories are found. The venues in Chicago are sorted based on the favorite categories and the rating. The top venues are then recommended to the user.

Methodology

Wikipedia provides the list of neighborhoods in Chicago. BeautifulSoup scapes the table on the website, and the texts in the table are written in a list by a for loop. The list is then formatted into a dataframe with the shape of (246, 2), shown in the following table.

Table 1. Neighborhoods in Chicago.

	Neighborhood	Community area
0	Albany Park	Albany Park
1	Altgeld Gardens	Riverdale
2	Andersonville	Edgewater
3	Archer Heights	Archer Heights
4	Armour Square	Armour Square

The neighborhood name and community area are used as the address in the geopy library, and the function geolocator.geocode is used to acquire the latitude and longitude information for each neighborhood. A list of geographical information is constructed by a for loop, and the list is then added into the dataframe, shown in Table 2. The neighborhoods without latitude or longitude are dropped.

Table 2. Neighborhoods with latitude and longitude in Chicago

	Neighborhood	Community area	latitude	longitude
0	Albany Park	Albany Park	41.980269	-87.718485
1	Altgeld Gardens	Riverdale	41.654864	-87.600446
2	Andersonville	Edgewater	41.977139	-87.669273
3	Archer Heights	Archer Heights	41.807623	-87.727406
4	Armour Square	Armour Square	41.840311	-87.631996

The venues nearby each neighborhood are explored by the Foursquare API. The version is 20190801, the radius is 500 meters, and the limit is 100 for each neighborhood due to the quota restriction from the Foursquare. The JSON file is requested containing the venues' information, and the venue name, location, category, and ID are returned for each venue. A for loop is used to obtain all the venues for each neighborhood, and a dataframe is constructed, shown in Table 3.

Table 3. Venues in the neighborhoods of Chicago

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Venue ID
0	Albany Park	41.980269	-87.718485	Subway	41.981375	-87.718396	Sandwich Place	4d51b489747f6dcb70fdc3d4
1	Albany Park	41.980269	-87.718485	7-Eleven	41.983368	-87.714196	Convenience Store	4b748851f964a52042e22de3
2	Albany Park	41.980269	-87.718485	Café Descartes	41.981552	-87.718156	Coffee Shop	4c4082e6cc410f4794b4a961
3	Albany Park	41.980269	-87.718485	Campus Aquatics	41.978416	-87.717182	Gym Pool	4d66fc5e9792b1f7479b381f
4	Albany Park	41.980269	-87.718485	E.Leaven Food Truck	41.979290	-87.715994	Food Truck	4f1dacbee4b083823465eff6

The dataframe is grouped by the neighborhood, and the number of venues in each neighborhood is calculated. The number of unique categories for all venues is then determined. One hot encoding is used to transfer 316 unique venue categories into a dataframe by the get_dummies method. The categories are grouped by the neighborhood, and the mean value for each category is calculated. The dataframe is shown in Table 4.

Table 4. Venue categories proportions for each neighborhood

	Neighborhood	Yoga Studio	ATM	Accessories Store	Adult Boutique	African Restaurant	Airport	Airport Lounge	Airport Service	American Restaurant	...	Video Game Store	Video Store	Vietnamese Restaurant	Vin
0	Albany Park	0.000000	0.000000	0.00	0.000000	0.000000	0.00	0.0	0.0	0.000000	...	0.000000	0.000000	0.000000	
1	Altgeld Gardens	0.000000	0.000000	0.00	0.000000	0.000000	0.00	0.0	0.0	0.000000	...	0.000000	0.000000	0.000000	
2	Andersonville	0.000000	0.000000	0.00	0.011765	0.000000	0.00	0.0	0.0	0.011765	...	0.000000	0.000000	0.000000	
3	Archer Heights	0.000000	0.000000	0.00	0.000000	0.000000	0.00	0.0	0.0	0.000000	...	0.000000	0.050000	0.000000	
4	Armour Square	0.000000	0.000000	0.00	0.000000	0.000000	0.00	0.0	0.0	0.000000	...	0.000000	0.000000	0.000000	
5	Ashburn	0.000000	0.000000	0.00	0.000000	0.000000	0.00	0.0	0.0	0.000000	...	0.000000	0.000000	0.000000	
6	Avalon Park	0.000000	0.066667	0.00	0.000000	0.000000	0.00	0.0	0.0	0.000000	...	0.000000	0.000000	0.000000	
7	Avondale	0.000000	0.032258	0.00	0.000000	0.000000	0.00	0.0	0.0	0.000000	...	0.000000	0.032258	0.000000	
8	Back of the Yards	0.000000	0.000000	0.00	0.000000	0.000000	0.00	0.0	0.0	0.000000	...	0.000000	0.033333	0.000000	

The venues are then sorted in descending order, and the top 5 venues for each neighborhood are created in a dataframe, shown in Table 5.

Table 5. The top 5 venues for each neighborhood

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Albany Park	Bus Station	Convenience Store	Asian Restaurant	Food Truck	Gym Pool
1	Altgeld Gardens	Park	Fast Food Restaurant	Women's Store	Factory	Eastern European Restaurant
2	Andersonville	Italian Restaurant	Coffee Shop	Pet Store	Lounge	Sandwich Place
3	Archer Heights	Mexican Restaurant	Bakery	Gas Station	Grocery Store	Pharmacy
4	Armour Square	Chinese Restaurant	Sandwich Place	Cosmetics Shop	Breakfast Spot	Grocery Store
5	Ashburn	Clothing Store	Mexican Restaurant	Locksmith	Fried Chicken Joint	Park
6	Avalon Park	Business Service	Burger Joint	Fast Food Restaurant	Food	Grocery Store
7	Avondale	Bakery	Mexican Restaurant	Bar	Donut Shop	Park
8	Back of the Yards	Mexican Restaurant	Pizza Place	Bank	Grocery Store	Clothing Store
9	Belmont Gardens	Fast Food Restaurant	Discount Store	Food	Entertainment Service	Flea Market
10	Belmont Terrace	Liquor Store	Salon / Barbershop	Home Service	Automotive Shop	Bank

Coffee shop is a common category for many neighborhoods, and therefore the venues classified as a coffee shop are sliced into a dataframe for all neighborhoods, shown in Table 6.

Table 6. Coffee shops in the neighborhoods of Chicago

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Venue ID
2	Albany Park	41.980269	-87.718485	Café Descartes	41.981552	-87.718156	Coffee Shop	4c4082e6cc410f4794b4a961
24	Andersonville	41.977139	-87.669273	La Colombe Coffee Roasters	41.976095	-87.668577	Coffee Shop	546634ba498e2e08a21d1ce0
52	Andersonville	41.977139	-87.669273	Starbucks	41.978163	-87.668505	Coffee Shop	4a7efa6df964a52064f21fe3
67	Andersonville	41.977139	-87.669273	Colectivo Coffee Roasters	41.980619	-87.668247	Coffee Shop	5ba8cecb033693002c6c407f
239	Belmont Terrace	41.937585	-87.834795	Starbucks	41.934900	-87.835768	Coffee Shop	4bbe59b982a2ef3b57352bd2

Ratings for each venue are acquired by the Foursquare API, and the results are added in the dataframe. The order is then sorted in descending by the rating, shown in Table 7.

Table 7. Top coffee shops in the neighborhoods of Chicago

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Venue ID	Rating
1684	Logan Square	41.928400	-87.706764	Sip of Hope	41.924656	-87.703974	Coffee Shop	5ac6e19d78782c6401c32429	9.0
622	Dearborn Park	41.866553	-87.628954	Stan's Donuts & Coffee	41.867516	-87.626402	Coffee Shop	569a2c27498e9defd6054348	9.0
839	Edgewater	41.983369	-87.663952	The Coffee Studio	41.984374	-87.669105	Coffee Shop	4a54e2b2f964a52062b31fe3	8.8
895	Edgewater Beach	41.983369	-87.663952	The Coffee Studio	41.984374	-87.669105	Coffee Shop	4a54e2b2f964a52062b31fe3	8.8
538	Clarendon Park	41.963275	-87.648840	Everybody's Coffee	41.965519	-87.653944	Coffee Shop	534d3962498e0eb23f7dfa5b	8.8

The coffee shop with the highest rating is explored by the Foursquare API, and the tips are filtered leaving the text, agree count, disagree count, and the user id. The tip with the most agree count is used to select the user id, and the user id is selected to make the recommendation.

Results and Discussion

The user's information is requested through the Foursquare API, and one of the user's features is the lists that the user created. Among the lists, a list of venuelikes is obtained, shown in the Figure 1.

```
results['response']['lists']['groups'][0]['items'][1]
```

```
{'id': '169953/venuelikes',
 'name': 'ben’s Liked Places',
 'description': '',
 'placesSummary': "Joe's Pizza, Lucky Dorr Patio & Tap, Otto's Place, Au Pied de Cochon",
 'type': 'liked',
 'user': {'id': '169953',
```

Figure 1. A list of venuelikes that the user created

Using the list_id, the items in the list are extracted, shown in Figure 2.

```python
ben_likes_list = []
ben_likes_list.append([(
    place['venue']['id'],
    place['venue']['name'],
    #place['venue']['location']['city'],
    place['venue']['categories'][0]['name']) for place in ben_likes])
ben_likes_list
```

```
[[('3fd66200f964a520dbea1ee3', 'The Blind Tiger', 'Beer Bar'),
  ('45ebc982f964a52091431fe3', "Joe's Pizza", 'Pizza Place'),
  ('59653657018cbb46a640bd78', 'Lucky Dorr Patio & Tap', 'Beer Bar'),
  ('4bdc54552a3a0f477c54b2b6', 'Flying Saucer', 'Breakfast Spot'),
  ('40b28c80f964a52029fc1ee3', "Delilah's", 'Whisky Bar'),
  ('4bc9e445937ca593cc86a692', 'Tiztal Cafe', 'Breakfast Spot'),
  ('4df4e6b763659543662d05a7', "Otto's Place", 'American Restaurant'),
  ('4adb95ebf964a5202d2921e3', 'Le Garde-Manger', 'Seafood Restaurant'),
  ('4b84801af964a520fd3831e3', 'Joe Beef', 'Steakhouse'),
  ('4ad4c06bf964a52046f920e3', 'Au Pied de Cochon', 'French Restaurant'),
  ('49ea7b74f964a5206e661fe3',
   "Pappy & Harriet's Pioneertown Palace",
   'BBQ Joint'),
  ('4b4e9dc9f964a520b2f226e3', 'Carnevor Steakhouse Moderne', 'Steakhouse'),
  ('5c09d27867af3a002cb81bbc', 'The Swill Inn', 'Pub'),
  ('4c1f8dd3b306c928029768b7', 'Girl & the Goat', 'New American Restaurant'),
  ('4b05865cf964a520d05e22e3', 'Proof on Main', 'Bar'),
  ('4e39e96588772c3bf11d150a', "Jimmy's Pizza Cafe", 'Pizza Place'),
  ('5a76331e0336935b3dcf3eab', 'SGD DUBU - @h mart', 'Korean Restaurant'),
  ('49e4e821f964a52067631fe3', 'Spacca Napoli Pizzeria', 'Pizza Place'),
  ('4b16cd4ff964a520b0bd23e3', "Kaufman's Bagel & Delicatessen", 'Bagel Shop'),
  ('4a2c3405f964a5202f971fe3', "Weegee's Lounge", 'Cocktail Bar'),
  ('5bd388af9ef8ef003953e234', 'Beermiscuous', 'Beer Bar'),
  ('4b69c83bf964a520e0b32be3', 'The Ivy Club', 'Lounge'),
  ('4db4670d316a3bec525cc66a', 'Scafuri Bakery', 'Bakery'),
  ('529918be498e59fe18712653', 'chair 9', 'Cocktail Bar'),
  ('55f43027498efda6b370df8c', 'Esphahan', 'Indian Restaurant'),
  ('4dfe2580d22d056d59a6cefb', 'Upre', 'Food'),
  ('515d96d2e4b07a8438b3936a', 'Sheesh Mahal At Leela', 'Indian Restaurant'),
  ('4bab1b3ef964a52051953ae3', 'Bukhara', 'North Indian Restaurant'),
  ('4ad8ac03f964a520921321e3',
   "Bombacigno's J & C Restaurant",
   'Italian Restaurant'),
  ('5ab54763b25fee59a956539f', 'Din Tai Fung', 'Chinese Restaurant')]]
```

Figure 2. Items in the list of venuelikes

Then, a dataframe containing the venue id, name, and category is constructed, shown in Table 8.

Table 8. Venues the user likes

```
df_ben_likes = pd.DataFrame([item for venue_list in ben_likes_list for item in venue_list])
df_ben_likes.columns = ['ID', 'Name', 'Category']
df_ben_likes
```

	ID	Name	Category
0	3fd66200f964a520dbea1ee3	The Blind Tiger	Beer Bar
1	45ebc982f964a52091431fe3	Joe's Pizza	Pizza Place
2	59653657018cbb46a640bd78	Lucky Dorr Patio & Tap	Beer Bar
3	4bdc54552a3a0f477c54b2b6	Flying Saucer	Breakfast Spot
4	40b28c80f964a52029fc1ee3	Delilah's	Whisky Bar
5	4bc9e445937ca593cc86a692	Tiztal Cafe	Breakfast Spot

The number of each category is determined using the dataframe, and the results shown in the Table 9.

Table 9. The number of each category in the list

```
df_ben_likes.Category.value_counts()
```

```
Beer Bar                   3
Pizza Place                3
Indian Restaurant          2
Steakhouse                 2
Breakfast Spot             2
Cocktail Bar               2
New American Restaurant    1
North Indian Restaurant    1
Bagel Shop                 1
French Restaurant          1
Korean Restaurant          1
Seafood Restaurant         1
Food                       1
BBQ Joint                  1
Italian Restaurant         1
Lounge                     1
Bar                        1
Pub                        1
Whisky Bar                 1
Chinese Restaurant         1
American Restaurant        1
Bakery                     1
Name: Category, dtype: int64
```

From the table, it is assumed that the beer bar and the pizza place are the user's favorite category. Therefore, some venues of the two categories in the neighborhoods of Chicago are recommended based on the ratings of these venues which are obtained through the Foursquare API. The results are shown in the Table 10 and 11.

Table 10. Recommended beer bar in Chicago

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Venue ID	Rating
11	Andersonville	41.977139	-87.669273	Hopleaf Bar	41.975813	-87.668509	Beer Bar	412a8500f964a520770c1fe3	9.1
2209	Near North Side	41.900033	-87.634497	Centennial Crafted Beer & Eatery	41.895807	-87.632539	Beer Bar	58c1b0879b7eac37c66519f5	9.1
3491	Wrigleyville	41.947022	-87.656477	Lucky Dorr Patio & Tap	41.948660	-87.656829	Beer Bar	59653657018cbb46a640bd78	8.8
77	Andersonville	41.977139	-87.669273	Meeting House Tavern	41.973295	-87.667916	Beer Bar	5b16e138db3aef0024361aac	7.4
2450	O'Hare	41.977928	-87.902955	Goose Island Beer Co.	41.977056	-87.907214	Beer Bar	581e34531338633a85a462a3	7.4

Table 11. Recommended pizza place in Chicago

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Venue ID	Rating
1132	Goose Island	41.910048	-87.655765	Pizzeria Bebu	41.909237	-87.651089	Pizza Place	58962beb375c4a79fe9ecdb0	9.1
325	Boystown	41.943883	-87.649267	Dimo's Pizza	41.945113	-87.654730	Pizza Place	4a3b2be5f964a52098a01fe3	8.8
625	Dearborn Park	41.866553	-87.628954	Flo & Santos	41.865398	-87.626123	Pizza Place	4bc7b41e92b376b08c4e503a	8.7
639	Dearborn Park	41.866553	-87.628954	Giordano's	41.864562	-87.624400	Pizza Place	5284eb42498e7b97a8388083	8.5
411	Buena Park	41.957810	-87.652833	Michael's Original Pizzeria & Tavern	41.956879	-87.651865	Pizza Place	4b0b3045f964a520552e23e3	8.4

The more specific recommended venues can be recommended if the current location of the user is known since the venue latitude and longitude can be used to calculate the distance between the user and the target. If quota permitted, all the list of the user can be explored. Then, the number of each category will be more representative. Also, more investigations can be performed on the other users and their like lists. Based on these lists, the users who have similar like venues can be retrieved. Then, the collaborative filtering can be used to recommend potential venues.

Conclusion

In this study, a simple recommendation system for places in Chicago is investigated. The Foursquare API is used to provide location information, the user information, and the rating list of each category of places. More than 150 neighborhoods in Chicago are explored, and the venues in these neighborhoods are obtained with the location, category, and rating. The taste of a user is determined based on the user-creating like list. Then, a list of beer bar and pizza place which the user preferred are recommended with the ratings. The recommendation system not only lifts the burden of users in searching location information but also provides values for third-party companies to advertise and to forecast service demand.