Capstone Project:

Walking Tour

Data

The walking tour algorithm will select venues from a list provided by Foursquare. Foursquare stores data for venues which can be accessed using various API endpoints. The SEARCH endpoint returns a list of venues within a given radius of a location that match the given parameters. The response of the SEARCH endpoint provides the id, name, location, and categories for each venue matching the search parameters. A category parameter can be included, which limits the results to the selected categories. Table 2, below, outlines the SEARCH parameters used in this project.

Table 2: SEARCH Endpoint Parameters

SEARCH Endpoint Parameters		
Name	**Description**	**Value**
ll	comma separated latitude and longitude values for the location of the search	The latitude and longitude of the start location
intent	intent of the search	The following intents will be explored to determine the optimal intent. checkin: venues a user is likely to checkin to at the current location and time browse: Finds venues in the area without considering the distance from ll
radius	radius, in meters, from the desired search area	The radius will be calculated using the distance the user is willing to travel and the number of stops they want to make
limit	number of results to return	The maximum number of results Foursquare returns is 50
categoryId	comma separated list of category Ids to search for	The categories the user is interested in

To better understand the Foursquare data and the SEARCH endpoint, we will look at an example. Let's search for venues within 300 meters of the Sheraton Dallas Hotel in Dallas, TX in the category 'Dessert Shop'. We will first look at the results for a search with a checkin intent, then compare those results to the same search with a browse intent. Table 3 outlines the parameters used in our first search. Refer to Table 2 for a description of these parameters.

Table 3: Example SEARCH Endpoint Parameters

Example 1: SEARCH Endpoint Parameters	
Name	**Value**
ll	32.785150, -96.794982
intent	checkin
radius	300
limit	50
categoryId	4bf58dd8d48988d1d0941735

Once we have defined the parameters we can conduct our search. From the response field we will extract relevant information and create a data frame. Table 4 shows us what this data frame will look like for our example. Each column represents a property and each row a venue.

Table 4: Example SEARCH Endpoint Results (intent=checkin)

	Name	Category	Category ID	Latitude	Longitude	Distance from Start
0	Yumi Yogurt	Ice Cream Shop	4bf58dd8d48988d1d0941735	32.7869	-96.7956	208
1	Chill Frozen Yogurt	Frozen Yogurt Shop	4bf58dd8d48988d1d0941735	32.7854	-96.7953	41
2	Kist Kitchen	Frozen Yogurt Shop	4bf58dd8d48988d1d0941735	32.7822	-96.7973	395

Notice we searched for 'Dessert Shops', but the category column doesn't contain that category name. Foursquare classifies venues according to a hierarchical scheme. When you search using a category id, Foursquare returns all venues within that category and all sub-categories. The first category in the response is the lowest level category ascribed to that venue. This information will give us, and the user, more information about the venues used in the algorithm. The category ID field in Table 3 is the category id that was used in the search rather than the id of the first category returned. This field is to help train the algorithm to create a loop with variety in the categories of the venues visited.

The 'Distance from Start' is the distance in meters from the location of the search, in our example, the distance from the hotel. This will be used to train the algorithm to select venues that can loop back to the start. We can notice from this field that Kist Kitchen is 395 meters from the hotel, which is not with in

our radius. This is because the checkin intent does not consider the search radius, only venues a user is likely to checkin to for that location and time.

Now, let us look at the results of the request using the browse intent. In Table 5 we can observe the difference between the two search intents. The browse intent only returns venues within our given radius, making it the better choice for our algorithm.

Table 5: Example SEARCH Endpoint Results (intent=browse)

	Name	Category	Category ID	Latitude	Longitude	Distance from Start
0	Chill Frozen Yogurt	Frozen Yogurt Shop	4bf58dd8d48988d1d0941735	32.7854	-96.7953	41
1	Yumi Yogurt	Ice Cream Shop	4bf58dd8d48988d1d0941735	32.7869	-96.7956	208

<small>Example SEARCH Endpoint Results (intent=browse)</small>