Restaurant Selector

1 Method:

For this assignment, I have utilised two ML approaches.

1.1 Using back propagation to fill the missing rating for the self-reviewer.

a. From the ratings data, we can observe that the reviewer "you" have not rated and/or visited some of the restaurant. Therefore, first objective is to fill the missing ratings for this reviewer.
b. To fill the missing ratings, I have utilised backpropagation algorithm to predict the missing ratings.
c. The parameters are:
    i. Average ratings of each restaurant by the other users.
    ii. Number of users rated that particular restaurant.
    iii. Standard deviation of the ratings for that restaurant
d. Output: ratings for those missing ratings by the reviewer "you."
e. Training set are those data where rating by "YOU" is available.
f. Predicting a rating is for those restaurant where the rating is unavailable (Not rated by "you").
g. Since I have taken the average, total count of the users who have rated the restaurant, and standard deviation in their ratings, I have ignored ratings that were missing by the other reviewers.
h. This is because, I can approach and select a restaurant just by knowing my own rating and the distance I need to travel.

1.2 Now, we select a restaurant.

i. From the above step 1, we obtain ratings for all the restaurants – some are given and some are predicted by neural network.
j. Now, the distance to each of the restaurants is calculated by utilising the coordinates given.
k. However, to calculate the distance, we need our origin. Therefore, we provide the x and y coordinates of the reviewer "you" as an argument.
l. The classifier used is SVM kernel classifier.

2 Observation:

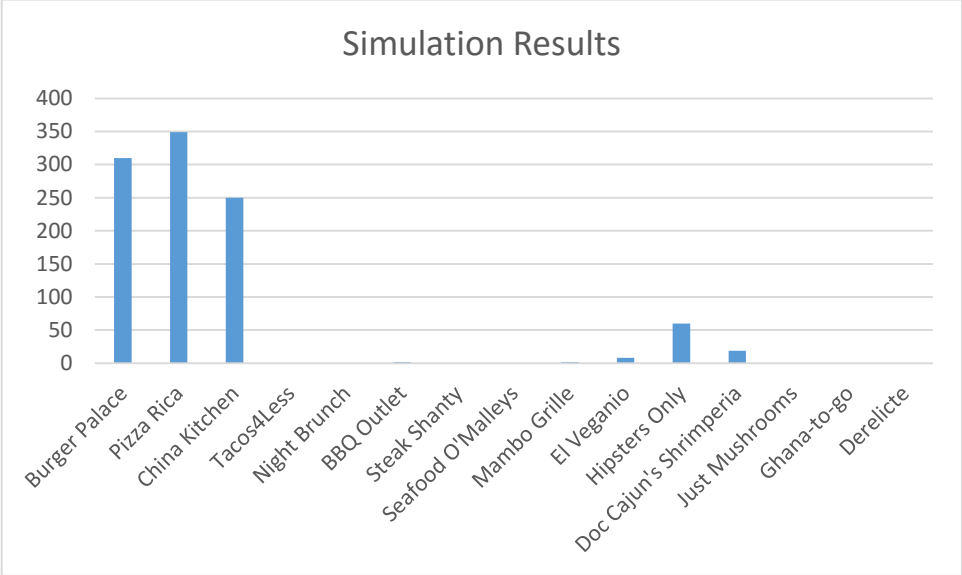
2.1 Back propagation:

The algorithm works well and has the accuracy of 90% or above. Since the data is less, accuracy may vary on large dataset.

2.2 SVM-Regression Classifier:

The output by SVR had a balanced selection of restaurants. Even when the distance is close but the rating is low, the classifier successfully picks a best candidate ahead of picking the restaurant closer to home.

The SVM classifier when tried with linear and precomputed kernels, the result was inclined mostly towards better rating and the distance had less weightage in the restaurant selection.



ID	Res	Occurrences	Rating before Neural Network
A	Burger Palace	310	5
B	Pizza Rica	349	4
C	China Kitchen	250	4
D	Tacos4Less	0	**NA**
E	Night Brunch	0	1
F	BBQ Outlet	2	**NA**
G	Steak Shanty	0	2
H	Seafood O'Malleys	0	**NA**
I	Mambo Grille	2	**NA**
J	El Veganio	8	3
K	Hipsters Only	60	4
L	Doc Cajun's Shrimperia	19	**NA**
M	Just Mushrooms	0	1
N	Ghana-to-go	0	1
O	Derelicte	0	1

3 References:

http://scikit-learn.org/stable/modules/svm.html

http://docs.scipy.org/doc/scipy-0.16.0/reference/index.html