Thinkful DataScience Final capstone Proposal

Yuta Katsuyama

1. **What is the problem you are attempting to solve?**

   The goal of this project is to help small restaurants to find the best pickup spot for their food. COVID-19 crisis has a great impact on restaurants. Since the stay at home order started in several cities around the world, many restaurants are struggling to make a profit. Many restaurants started using the 3rd party delivery service like UberEats, GrubHub, DoorDash, and Postmates. However, there is a big issue using these services. Their commission rates are very high, it's around 30-40%. For most of the restaurants, it's almost impossible to make a sufficient profit with this rate. Therefore, some restaurants are now delivering by themselves to several pickup spots around the city (Not delivering to each user's house. Because of the limited delivery capacity). This causes another problem. They don't know where are the best places for pickup. To solve this issue, we need to help restaurants to predict the best location to deliver.

2. **How is your solution valuable?**

   It helps small restaurants to choose best pickup spots in the city. It can also provide other useful information like popular cuisines in the city, how much other restaurants charge for the delivery cost, etc. Also, by getting feedback from users, we can improve the prediction algorithms.

3. **What is your data source and how will you access it?**

   I'll use the Zomato API. It's free to use up to 1000 calls/day. Also, since I am in Chicago, I'll use Chicago's data. There are data about foodies, that shows where people with a strong interest in food are. It helps us to select the best location to set up pickup spots. Also, there is data about the restaurant type. So, if you are a Japanese restaurant owner, you can find the area with strong competition.
   https://developers.zomato.com/api

4. **What techniques from the course do you anticipate using?**

   Data visualization, statistics, supervised learning (e.g. KNN clusters, Random forest, SVM) for predicting the spots, and try to use Tensorflow and Keras.

5. **What do you anticipate to be the biggest challenge you'll face?**

   Creating a dataset from API may be challenging. Also, improving the model can be a challenge because of the limited samples, and missing values.