

Proposal for Project 4

Project name: The Prediction models for the Popularity and Success of Uber Eats Restaurant and Menus

Group Members: Yuanfeng Xu, Alexis Hernandez, Mary Jane Rafol and Taib Diallo

Backgrounds:

<https://www.ubereats.com> - An online food ordering and delivery platform launched by Uber in 2014. Users can read menus, reviews, ratings, order, and pay for food from participating restaurants using an application on the iOS or Android platforms, or through a web browser. Users are also able to tip for delivery. Payment is charged to a card on file with Uber. Meals are delivered by couriers using cars, scooters, bikes, or foot. It is operational in over 6,000 cities across 45 countries.

Data Source:

- <https://www.kaggle.com/datasets/ahmedshahriarsakib/uber-eats-usa-restaurants-menus>
- This dataset contains lists of Restaurants and their menus in the USA that are partnered with Uber Eats.
- Data was collected via web scraping using python libraries.

Project Objectives:

1. Create predictive models for popular restaurants in Uber Eater delivery system.
2. Create predictive models for popular menus in Uber Eater delivery system.
3. Create a model to determine whether a restaurant will be successful or not (online delivery segment).
4. Determine what parameter contributes to a restaurant's success (Ratings, Score, and Price Range)
5. Create a visualization that continues to learn where clusters lie based on ML.

Methodology:

1. Unsupervised Machine Learning
2. Supervised Machine Learning: Logistic Regression and Random Forest Classifier
3. Deep learning: Neuron Network
4. Details:

Figure out the top-20 categories, use commas to separate, use key word dictionary—if key word exist, 1 (minimize columns),

Top 20—nested for loop, column name is the key (Sushi, Asian), if key is in category add 1, do for all rows (create column)

From menu, add the cheapest, average and most expensive plate—merge to the main file (check possible correlation to the price range)

Feature engineering

Extract states from the address, top states, others

How to increase your position in Uber eat search results?

Assumption: The higher the position the better chance of success

1. Apply linear regression to determine which feature correlates better to the position? (Taib)

- Ratings
- Count of ratings
- Category (Yuanfeng, NLP, feature engineering-Mary Jane & Alexis)
- Price range

This also answers the question whether there is any relationship between the price level and the popularity of the restaurant.

1. Perform clustering (DBSCAN/KMEANS) via unsupervised learning to see whether our dataset has some logical groupings or not. (Yuanfeng)
2. For feature that correlates better, use them as features for our supervised learning model (Mary Jane)
 1. Classification model (logistic regression—probability of event from occurring (answers the questions yes or no), neural networks—deep learning that allows machine to recognize patterns to solve a problem (answers the question about success, efficiency-time based)
3. Apply J3 Leaflet or PowerBI or tableau for an interactive map of where the highest-rated restaurants (drop pin, heatmap) (Alexis)
 1. Location
 2. Ratings

Tools:

1. Jupyter Notebook: Python Pandas, Python Matplotlib
2. Tableau