Tianyu Xu (Frederic) From Group 5

Self-Assessment

Roles played: Machine learning model training with unbalanced data. Using Random Forest Classifier to predict the recognition of Michelin Restaurant.

Challenges:

1. Visualization of the random tree. Not taught in the class so I have to google a lot to learn what module to import to generate the tree.
2. Downloads the Random Forest Tree model and re-use it for furthered predictions: as the model would take a very long reaction time and eats up all my disk storage. But finally, I have found a suitable number for the size of the model.
3. Unbalanced Data and Model Manipulation: Did not overcome this. In future I may get every feature classified into buckets.

Team-Assessment

Team Communication protocol:

1. Communication with slacks and zoom to set up at least twice per week.
2. My teammates are very supportive. Ashwin setting up the Github Page with main/branches very knowledgeably and he is always there to help. Julianna and Hanna have a precision sense on combining and cleaning the dataset. Geneille also works very hard on the machine model with further balanced data. I would really like to work with them again.

Summary:

In this project I have built up a Random Forest Classifier to predict the recognition of Michelin Restaurant (0 star /1 or more star) with an overall accuracy rate of 98%+. And the top 3 features are Total\_review\_count, Price\_level, Average\_rating. It could help the stack holder of a restaurant who is eagerly seeking for the potential of Michelin recognition to understand what feature his restaurant needs to strengthen.