zhang_ella_final_report

Ella Zhang

December 19, 2020

Introduction

It seems utterly strange to link tires with food, but Michelin does a great job on this task. Michelin, originally and still a tire company, publishes a restaurant reference guide annually, which always shocks the culinary world. The guide awards up to three Michelin stars for excellence to a select few restaurants in the world. The stars are awarded as follows:

For my data science project, I chose to analyze Michelin starred restaurants in the United States. To be specific, I want to answer these two questions. What factors make it more or less likely for a restaurant to be in the Michelin guide and to get a Michelin star.

A successful project will first produce a clean dataset that allows people to search for a Michelin restaurant and its Yelp rating easily. Based on my experience, I noticed a vast Yelp rating difference among one-starred restaurants. If I ever wanted to go to a one-starred restaurant with a high Yelp rating, I have to copy and paste each one-starred restaurant name into Yelp, search for the rating, and compare all the ratings to find the highest one. The process was too complicated and time-consuming, especially when you have a connection problem with Yelp. Therefore, I think a clean dataset will be extremely beneficial to foodies in the United States.

Second, a successful project should allow me to give recommendations grounded in the concrete analysis. For example, when people come to me and ask for Michelin restaurant recommendations in Washington, D.C., I should suggest them to choose a specific cuisine, because based on my project analysis, that cuisine has the highest rating among starred restaurants.

Problem Statement and Background

Just in case that you are not a foodie and have no idea about what a Michelin star is, here is some background information. Michelin is a French tire manufacturer. In fact, it is actually the second largest tire manufacturer in the world, ranked only after Bridgestone. Back in the 1900s, Michelin launched its first guidebook to encourage road tripping in France, and introduced the star rating system for fine dining restaurants.

Michelin awards up to 3 stars to a select few restaurants in the world. With 1 star indicating

a good place to stop on your journey. And 2 stars for restaurants worth a detour. And 3 stars for restaurants worth a special journey. These restaurants probably have exceptional cuisine where diners eat extremely well, often superbly. Or the restaurant can be in the guide but receives no stars. This simply means that it is a good restaurant, recommended, but not good enough for a star.

To this day, Michelin relies entirely on its full-time staff of anonymous reviewers to discuss which restaurants will be in the guide and awarded the stars. In this way, the Michelin stars are very different than Yelp ratings. It also takes longer for the staff to review restaurants around the world. In the United States, Michelin has only reviewed five cities: D.C., Chicago, New York, San Francisco, Los Angeles. In 2019, the guidebooks of the latter two cities merged into one for the entire state of California.

So, this project is actually a classification problem. I want to predict the likelihood of being on the Michelin guide and get a star. Some factors or predictors that I considered are: Yelp ratings and reviews, cuisine of the restaurant, and etc.

There are two sets of predicted outcomes that I want to investigate on: The first is whether or not the restaurant will be in the guide And the second is whether or not the restaurant will receive a star

Data

Kaggle: There is a dataset called Michelin Restaurants, which contains ratings, cuisine, and geographic information about Michelin restaurants worldwide. The dataset was lastly updated a year ago, so it doesn't include the 2020 ratings.

Wikipedia: There are pages for List of Michelin starred restaurants in New York City, Chicago, Washington, D.C., San Francisco Bay Area, and Los Angeles, each contains a table of restaurants and their star ratings since first awarded.

Michelin Guide Website: The website has separate links for each region. Every starred restaurant has its page, which contains some necessary information and a couple of pictures.

Yelp: Most starred restaurants have already been added on Yelp, accompanied by ratings and reviews.

Analysis

Results

Discussion