

Airbnb price prediction for NYC

Introduction

New York city (NYC) is one of the largest cities in the United States (US) as well as in the entire world. It is also the most populous and one of the most dense cities in US. Among many factors, a vibrant economy and the rich multi-cultural and multi-ethnic make up of the city, have popularized NYC as a popular destination for tourists. Hence, finding affordable accommodation is a challenge here. Airbnb helps alleviate this challenge to some extent by adding options for accommodation offered by private parties.

Pricing the listings is an important factor associated with the listings in Airbnb. Here lies in the challenge from two perspectives. Owners need to find appropriate price to make their listing competitive and the tourists (renters) want to find the best value for their expense. Precisely, predicting/determining the price of the listing is the business problem that I set forth to solve in this project.

In this project, I will build explainable predictive models to determine the price of the listings based on the current market rates. In order to do so multiple attributes of the listing need to be taken into account such as number of rooms, location, nearby venues, accessibility and so on. Hence, Foursquare API can provide valuable location and neighborhood information related to the listings that can help the predictive model.

Data

I decided to use a publicly available dataset of the listings in NYC on Airbnb platform. It can be accessed using this [link](#). The dataset has the following columns:

- id : listing id
- name : name of the listing
- host_id : id of the host
- host_name : name of the host

- neighbourhood : the area where the listing is located
- latitude
- longitude
- room_type : type of the room advertised
- price : price of the listing in US dollars

Using the coordinates and the neighbourhood information, I will query Foursquare API to get the nearby venues for food, entertainment and transport. For example, I will use the Venues search and explore endpoints in Places API to get the relevant information.

Note: there are about 48.9k listings in the dataset. Because of the rate limit of Foursquare free API, I will consider only a small sample of the dataset in this project.