Overview

Here in this capstone project, the objective is to build an ML model that can predict where a new Airbnb user is likely to book his/her stay.

Client & Data

In this scenario, the client would be Airbnb which is an online marketplace for arranging or offering lodging, primarily homestays, or tourism experiences. The company does not own any of the real estate listings, nor does it host events; it acts as a broker, receiving commissions from each booking.

The desired data is provided by the client while hosting a Data Science competition on Kaggle.com in 2015.

Approach

Data Wrangling:

The very first step will be to import the data into a jupyter notebook and explore it so that the anomalies can be identified and dealt with. Further analysis and model building will be done on the cleaned datasets.

EDA using visualization & statistical methods:

Once the data has been cleaned, it will be explored to understand relationships between the variables using inferential statistics. Multiple hypotheses will be drawn and validated in this step which will shape the procedure for initial predictive model building.

Model building:

Multiple ML methods will be used to build a predictive model and will be improved by testing against the test dataset provided by Airbnb.

Data story and results:

This step will involve creating a story around the initial problem, the problem it aims at solving and the insights gained from the data. Along with that, it will explain the challenges faced during model building as well as the results with accuracy.

Deliverables

Deliverables will include code, documentation, and graphics in a Jupyter notebook.