

Capstone Project # 1

10/26/2019

Anu Kashyap

Springboard Data Science Career Track

Overview

Airbnb provides an online marketplace where house owners can rent their houses to travellers looking for a short stay. It lets hosts share the full details their properties on an online platform where guests can search and select the property of their choice.

Problem statement: In which country will a new guest book their first travel experience?

Client name: Airbnb

Objective: The objective is to help Airbnb with the following:

- Better forecast demand
- Share customized content with client
- Decrease the average time to first booking

Specifications

Data: Data will be taken from the Kaggle competition held 4 years ago. It contains the following datasets:

- Test data
- Train data
- Other relevant data

The link to the competition is:

<https://www.kaggle.com/c/airbnb-recruiting-new-user-bookings/data>


Approach

This is a supervised learning algorithm because we already have training data. Also this is a classification problem because there are fixed number of categories that need to be predicted in the response variable.

Use sklearn module in Python to look at various multiclass classification algorithms

The variable to predict is destination country.

The predictor variables which can be used are as follows:



User's current country, age, gender, month, distance between their country of residence and destination country, user's native language and language of the destination country

Project Deliverables

The project will have the following final deliverables:

1. Code
2. Presentation containing an executive summary and the approach followed
3. Detailed Project documentation

Milestones

Data Cleaning and Processing: Nov 3rd 2019

Exploratory Data Analysis: Nov 17th 2019

Implementation of classification Algorithms: Nov 24th 2019

Results and Summary: Dec 15th 2019

Project Presentation and Documentation: Dec 31st 2019