Capstone Project Proposal

For my Capstone Project, I will be working on the following competition from Kaggle:

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

The goal of this problem is to predict the country in which a new user will make his/her booking. All the users are from USA. This is a list of all the possible outcomes -  'US', 'FR', 'CA', 'GB', 'ES', 'IT', 'PT', 'NL','DE', 'AU', 'NDF' (no destination found).

The client is Airbnb. They will use this information to share more personalized content, decrease the average time to first booking and better forecast demand.

Data is provided by Kaggle. The data consists of details such as the demographics, web session records and some summary statistics.

To solve this problem, I will be going through the steps and execute the things that I have learnt in this course. I will start with a logistic regression model to come up with my first model. However I will try several machine learning algorithms to solve this problem.

I will be looking at the metrics such as Accuracy and ROC to evaluate my model. I will also be testing my model on a validation sample to ensure that it does not overfit and can hold well on other independent samples.

For my deliverables, I will include the Python code that I have created for this project. I will also write a paper for my project with all the relevant details.