Airbnb First Booking Prediction

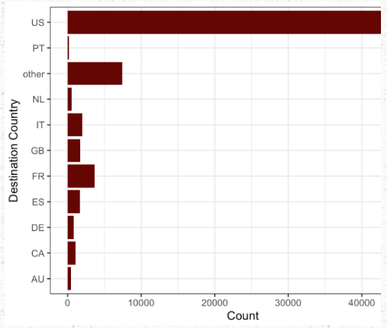
Killian Farrell, Suzana Iacob, Brandy Piao, Alexandru Socolov (MBAn '20)

Problem statement

- Users can book in **190+ countries**.
- **Goal**: **predict where users will book their first travel experience.**
- **Motivation**: Targeted advertisements and recommendations, enhanced user experience; increasing the likelihood of booking.

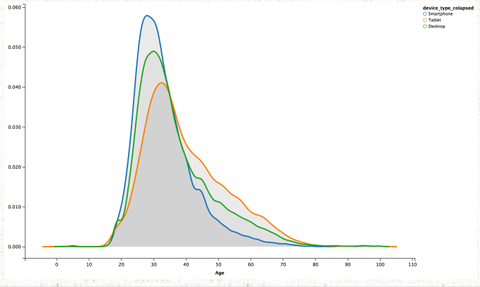
Exploratory Analysis

- **Dependent variable**: destination country (11 levels)
- **Predictors:** age, gender, first device & browser, sign-up method (e.g. Facebook), web session statistics, etc.



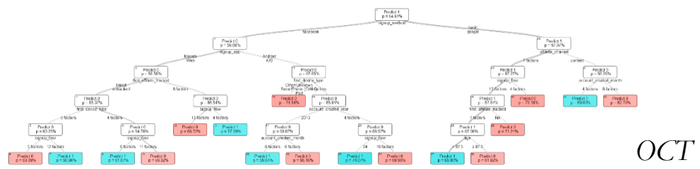
Destination Country

- **Feature engineering:** age groups (e.g. 30s), millennials, device type (e.g. tablet), day, month and year of the account creation.
- Due to imbalanced classes problem split into 3 steps



Distribution of the users over age and device types

1: Will the user book?



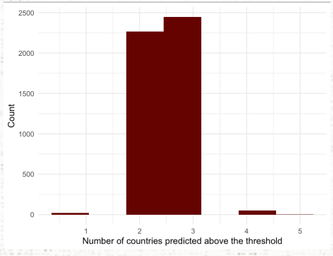
OCT

- **Binary prediction**: booking (yes/no)
- Best Performance: **Ensemble** of logistic regression and boosted trees: **AUC 0.68**
- **Feature importance**: age, gender, sign-up method, first browser

2: Will they book in the US?

- Given booking = yes, binary **prediction of US/Non-US destination**
- All users US based, **70% of destinations are US**
- Up/Down-sampling, loss matrix implementation
- **Important Features:** speaking Italian (+), web signup (+), Sign-up in June (+)

3: Where exactly will they book?



How many countries per user? (Mostly 2-3)

- Predict **a list of countries** for each user
- Aside from US, model predicts more France and Italy in general, as more users speak French and Italian.

Performance

Out-of-Sample Performances

AUC	Booking	US/Abroad	Country
CART	0.62	0.55	0.59
RF	0.66	0.54	0.61
GBM	0.67	0.56	0.66

Multi-Country Prediction – we calculate an accuracy score based on whether the country appears in the prediction list – **87% accuracy**

Managerial implications

Which users book?

- Show ads when the user is in the middle of a sign-up flow (stage 3 through 10)
- Facebook referrals correlated with higher chance of booking

Destination Country

- Since the goal is targeted ads, we predict **multiple countries per user –** top destinations **ranked** based on predicted probabilities
- Display multiple ads, starting with the most likely destination