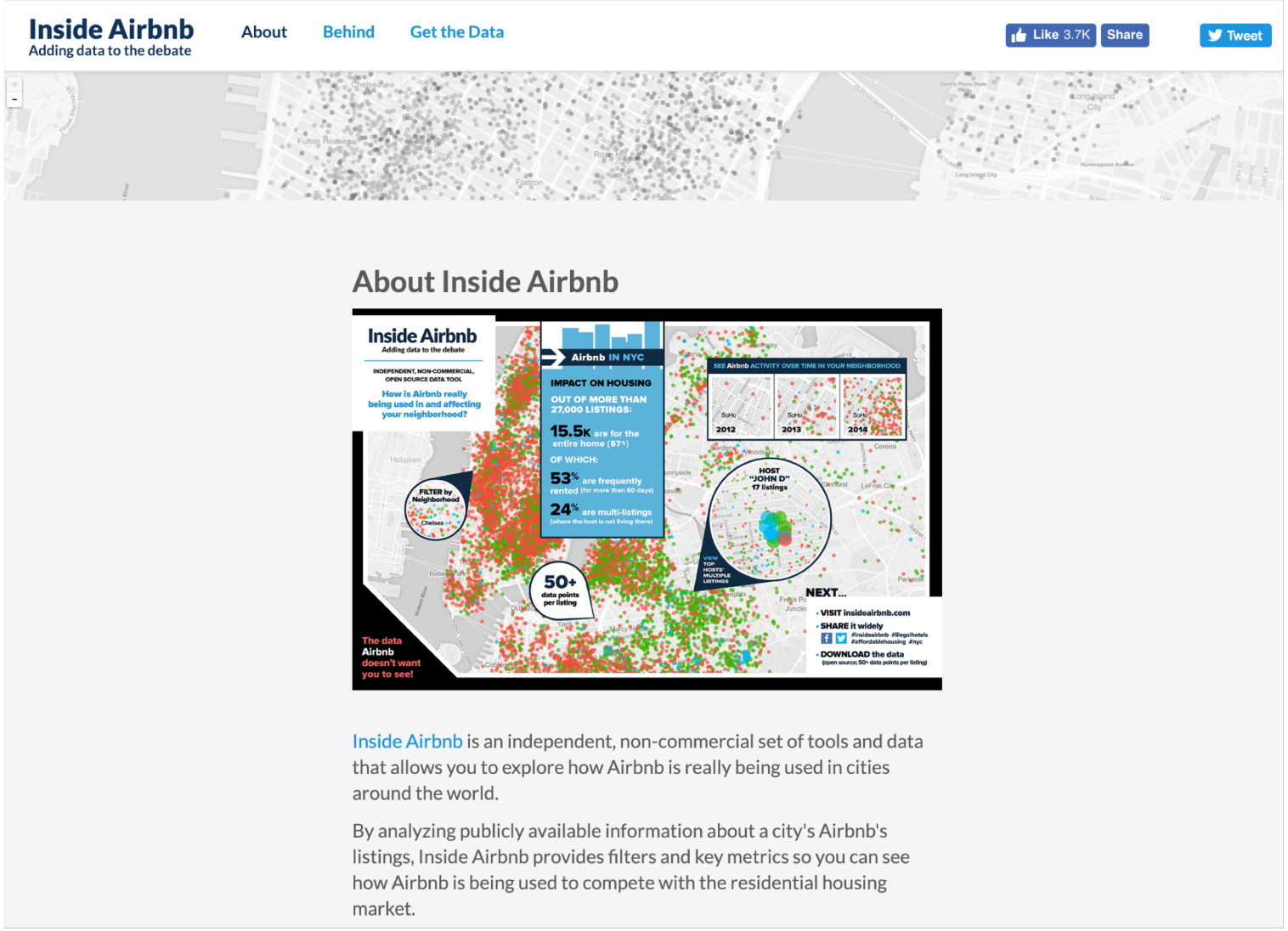
VIEW FROM MY AIRBNB PLACE IN HALLSTATT, AUSTRIA, 2016 SUMMER.

Airbnb Yield Prediction

Chao Li

PhD Candidate at NYU Chemistry

Where to get data?



http://insideairbnb.com/get-the-data.html

Exploratory data analysis
Data cleaning

↓

Web scraping photos
Photo quality assessment (both technically and aesthetically)

↓

Text mining on description

↓

Build machine learning models
Fine tuning the model

https://github.com/cl3080/Machine_Learning_Models_for_Airbnb_Yield_Prediction

Neural Image Assessment (NIMA) on Photos of the listings

NIMA Score: 5.268

NIMA Score: 4.519

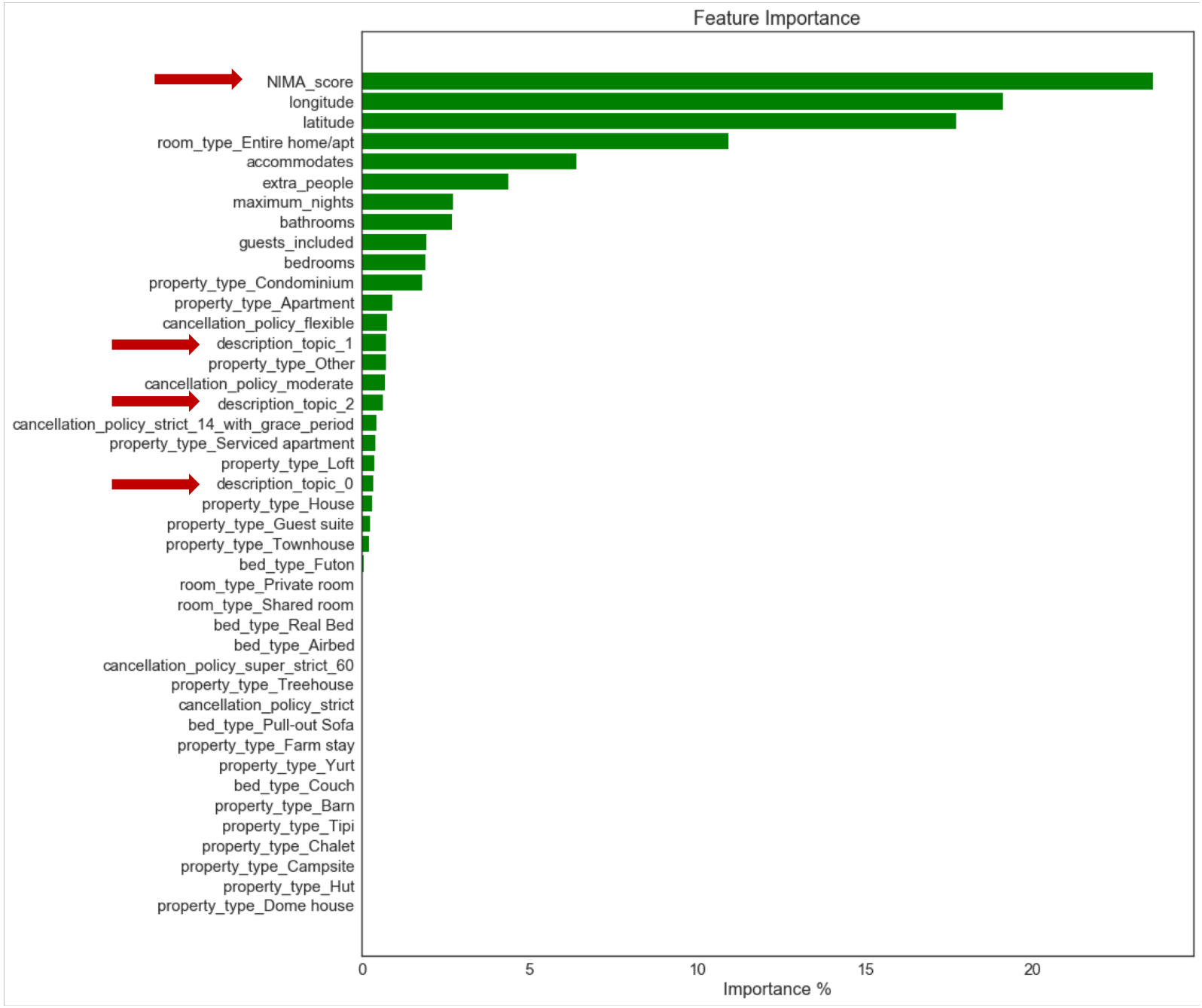
NIMA Score: 4.826

NIMA Score: 3.788

Sentiment Analysis on Description of the listings

http://htmlpreview.github.io/?https://github.com/cl3080/Machine_Learning_Models_for_Airbnb_Yield_Prediction/blob/master/lda_description.html

Feature Importance



Summary

- The goal for this project is to build machine learning models to give potential hosts some insights on how much they can earn from listing their beloved houses on Airbnb.

- There are definitely lots of information can be obtained from web photos and descriptions. My preliminary data shows that incorporating these information into machine learning model can improve the performance.

Thank You!