USA Housing Price Predictor

Find Your Nook—Your Dream Home Awaits!

MSiA 423 Final Presentation
Selina Song
June 8, 2020

Motivation

➢ Housing price information online is kind of scattered. It takes lots of time and effort to collect and evaluate the price of a specific house that fits users' need.

➢ This web app is born in order to:

1) enable potential homebuyers to efficiently find the up-to-date countrywide housing price information based on their preferences

2) help house sellers better understand the housing market before setting the price

➢ This web app will enable users to input their preferences on housing features and get an estimated price in one second.
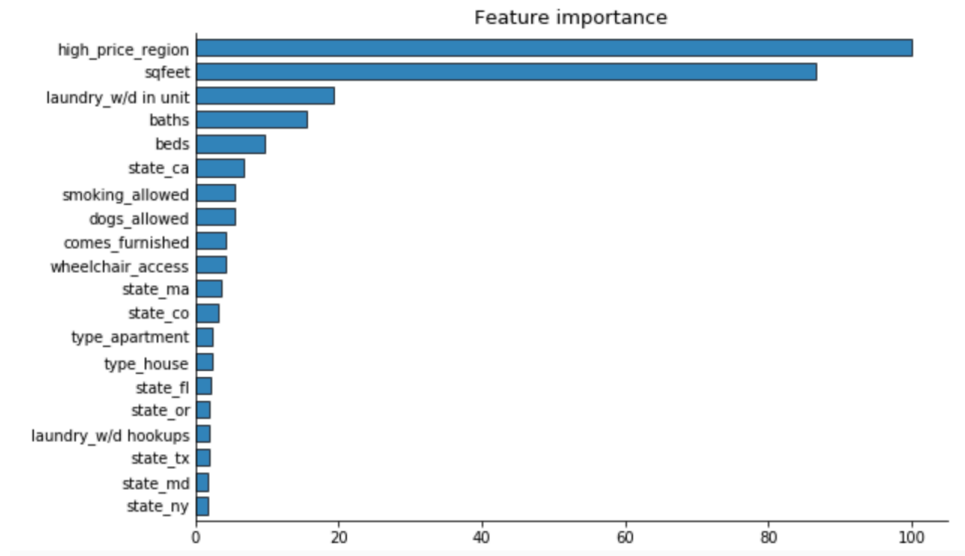
Data

➤ Kaggle's "USA Housing Listings" data that scrapped from Craigslist (the world's largest collection of privately sold housing options) in Jan. 2020

➤ In total 384,977 housing listings

➤ Target variable: housing price (per month)

➤ 11 predictors used in web app: State, Region, Type, Square Footage, Num of Beds, Num of Baths, Comes Furnished, Laundry Options, Smoking Allowed, Dogs Allowed, Wheelchair Access

Model

➢ Random Forest Regressor (n_estimators=100)

➢ 74 features after one-hot encoding

➢ Train/test split: 75/25

➢ Success Criteria:

   ❖ *Machine Learning*: Root Mean Square Error (RMSE) less than $200

   ❖ *Business*: users who have visited the web app to predict housing price will be contacted later asking for their experience—the satisfaction rate should be greater than 70%

➢ ML result: RMSE of $251.167

Insight

- ➤ high_price_region (mean price > $1,650—85th quantile of all price) and sqfeet dominate feature importance

Feature importance

Region	State	Mean Price
Santa Barbara	CA	$2,942
SF Bay Area	CA	$2,769
New York City	NY	$2,553
Florida Keys	FL	$2,461
Long Island	NY	$2,407
Los Angeles	CA	$2,388

THANK YOU!

✉ yanmengsong2020@u.northwestern.edu

in https://www.linkedin.com/in/yanmeng-song/