Project Proposal 8.6.3

Goal: Predict the price of real state in California

Steps:

1. Choose an official website where the data is public accessible and continuously updated, i.e. Zillow, Redfin and etc.
2. Write code to collect data with API for later manipulation in jupyter notebook.
3. Find a pre-trained model which we can further improve, e.g. a pre-trained model in Kaggle.
4. Find out the features within the model and make sure if more data is needed with this model.
5. Split our data to three sets: training set, test set and validation set.
6. Write code to train the model with our training set and then check if the model is underfitting or overfitting with test set/validation set by evaluating ML metrics (accuracy, recall or precision)
7. Improve the model with our knowledge to make it applicable for all sets to achieve a reasonable accuracy.
8. Make some predictions for future price in some certain locations in CA.