Housing Estimator

By: Team Test Data

Carol Buckinger, John Gutierrez, Shruti Bansal, and Megan Staley

Overview

In this project, we attempt to predict the median price of various houses in California depending on house-related features such as location, total rooms, and bedrooms as well as neighborhood-associated features such as median income of homeowners, number of homes within a specific range, and population. We will use multiple machine learning models such as Random Forest Classification, Neural Networking, and Multivariate Linear Regression.

Sources:

- Kaggle: https://www.kaggle.com/datasets/camnugent/california-housing-prices
- Kaggle:https://www.kaggle.com/datasets/camnugent/california-housing-feature-engineering?select=cal_populations_city.csv
- https://api.census.gov/data/1990/cbp?get=GEO_TTL,EMP,ESTAB&for=county:*&in=state:06&key=
- http://openweathermap.org/

Tools: Technologies, languages, tools, and algorithms used throughout the project

- Python in Jupyter Notebook
- Pandas, numpy, citypy
- Mlenv enviorment
- Multivariate Linear Regression
- Random Forest Classifier
- AWS Relational Database System
- Tableau

GitHub

- https://github.com/jsguti323/Housing_Estimator

Cleaning Data:

1.