

COS30049

Computing Technology Innovation Project

Assignment 2

Housing Price Prediction:

The goal of the "House Price Prediction" project is to forecast home prices through the application of machine learning methods. Using well-known Python modules like NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn (sklearn), and XGBoost, this project offers a comprehensive solution for precise price estimation.

Project Overview:

The goal of the "House Price Prediction" project is to create a model that, utilising a variety of features, can anticipate house prices with accuracy. This prediction problem is of considerable significance in real estate and finance, enabling educated decision-making for buyers, sellers, and investors. With the use of carefully chosen datasets and machine learning techniques, this project offers a potent tool for property value estimation.

Dependencies:

pandas, scikit-learn, matplotlib, seaborn, numpy, and python 3.8+ , and Jupyter Notebook.

All dependencies can be installed by running:

```
bash
```

```
pip install -r requirements.txt
```

Alternatively, use the Conda environment file:

```
bash
```

```
conda env create -f environment.yml  
conda activate housing-ml
```

Datasets Used:

The primary dataset, `housing_data.csv`, includes attributes like:

The quantity of bedrooms

Area in square footage

Location Cost

year built

All of the datasets that have been used are from different areas of Melbourne.

Model Training:

Models of Regression and Classification

We've put into practice:

A house price prediction model utilising **linear regression**.

A dwelling type classification system using a **Random Forest classifier**.

How to Train the Model?

Open **the .ipynb** file in Jupyter notebook and run the cells or run the everything from top.

Model Evaluation:

The following measures are used to assess the models:

Regression: R2 score, Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE).

Classification: F1 score, recall, accuracy, and precision.

Getting started To Run the Project Locally:

To run this project locally, follow these steps:

1. Open [the .ipynb](#) file or download it and open the file in vs code in Jupyter and run it.
2. Install the required libraries: [If you're using Google Colab, you don't need to pip install](#). Just follow the importing the dependencies section.
3. Launch Google Colab: <https://colab.research.google.com/>
4. Open [the .ipynb](#) file and run the notebook cells sequentially.

Getting Started to Run the Visualization Project Locally:

To run the visualization.ipynb project locally, follow these steps:

1. Open the visualization.ipynb File:
2. Download the visualization.ipynb file and open it in VS Code using Jupyter or any other notebook environment.
3. Install the Required Libraries:
4. If you're using Google Colab, you don't need to install libraries manually. Just follow the importing dependencies section in the notebook.
5. Launch Google Colab:
6. Open Google Colab.
7. Open the visualization.ipynb File:
8. Upload or open the visualization.ipynb file and run the notebook cells sequentially to see the visualizations.

Conclusion:

This study effectively illustrates how machine learning techniques may be applied to categorise different types of properties and estimate house prices. Regression and classification models allowed us to examine important variables impacting the housing market and offer insightful analysis of market patterns. The model's performance was improved by the inclusion of several datasets and thorough data processing, enabling more precise predictions. For the purpose of further investigating and refining machine learning models in real estate analytics, this project provides a fundamental tool.