**DESCRIPTION OF THE HOSUING PRICES PREDICTION MODEL USING ML**

DATASET IMPORTED DIRECTLY FROM SKLEARN LIBRARY USING THE BELOW COMMAND

***from*** *sklearn.datasets* ***import*** *fetch\_california\_housing*

* **HOUSING-PRICES-PREDICTION-MODEL**
* **PROBLEM STATEMENT**:

USING THE CALIFORNIA HOUSING DATASET –

1. Try replacing GridSearchCV with RandomizedSearchCV.
2. ii) Create a single pipeline that does the full data preparation plus the final prediction.
3. iii)Try adding a transformer in the preparation pipeline to select only the most important attributes.

* **Housing Dataset Analysis**

This repository contains the analysis of the housing dataset using scikit-learn. The primary goal is to replace GridSearchCV with RandomizedSearchCV in a single pipeline for data preparation and final prediction. Additionally, a transformer is added to select the most important attributes.

* **Folder Structure**

data/: Contains the housing dataset files.

notebooks/: Jupyter notebooks for data preprocessing, modeling, and analysis.

* **README.md:**

This file providing an overview of the project.

* **Dataset Description**

The housing dataset consists of features related to housing in Boston, aiming to predict the median value of owner-occupied homes. The dataset includes various attributes such as crime rate, pupil-teacher ratio, etc.

* **Task Overview**

Replacing GridSearchCV with RandomizedSearchCV We've implemented RandomizedSearchCV in a single pipeline for data preparation and prediction. This process includes:

1.Loading and preprocessing the housing dataset.

2.Constructing a pipeline for data preparation and model prediction.

3.Using RandomizedSearchCV to tune hyperparameters efficiently.

4.Feature Selection Transformer In this analysis, we've integrated a transformer into the pipeline to select the most important attributes for model training. This helps in improving model performance and reducing computational overhead.

* **Conclusion**

The implementation of RandomizedSearchCV in the pipeline, along with the feature selection transformer, enhances the efficiency of model tuning and data preparation for predicting housing prices.

* **References**

Scikit-learn Documentation Housing Dataset Source