**Name: ABDULRAHMAN ALMYMAN Student Number:**

**1.IDE:**

Jupyter notebook(easy to use) & Google Collab for working online.

1. **libraries used to implement the models:**

-Pandas: to read the data easily.

-Sklearn(include: sklearn.preprocessing, sklearn.model\_selection, sklearn.metrics, sklearn.ensemble, etc. )

1. **Description of the chosen algorithms including all the selected hyperparameters:**

SVR() because it has the best score

**Selected Hyperparameters:**

default hyperparameters

**4. Description of Data Preprocessing Applied:**

HW3:

- aqardata\_2

1- Separate features and target variable

2- Encode categorical variables with LabelEncoder

3- # Initialize the KFold object with shuffle and random\_state

# class sklearn.model\_selection.KFold(, \*, shuffle=true, random\_state=24)

4- # Initialize a list to store the mean squared errors, mean absolute errors, RMSE, and R2 scores for each model

5- Define the models to evaluate

6 - Iterate over the models and perform cross-validation & Initialize lists to store the evaluation scores for each fold

7 - # Perform cross-validation & Split the data into training and testing sets for this fold

8 - # Fit the model on the training data

9- # Make predictions on the testing data

10 - # Calculate the evaluation scores for this fold

11 - # Append the evaluation scores to the respective lists

12 - Calculate the mean of the evaluation scores for this model

13 - # Append the average evaluation scores to the list

14 - # Sort the models based on their average mean squared error in ascending order

15 - # Get the best performing model and its corresponding evaluation scores

16- # Print the best performing model and its evaluation scores

merging solution(not yet):

-aqardata\_2

-DocRealestateSale

-DocRealestateSale2023Q3

-Transactionssaleforrealestate

1. **Description of all other steps that have done to get a final result:**

* Selected relevant features for training the model:[

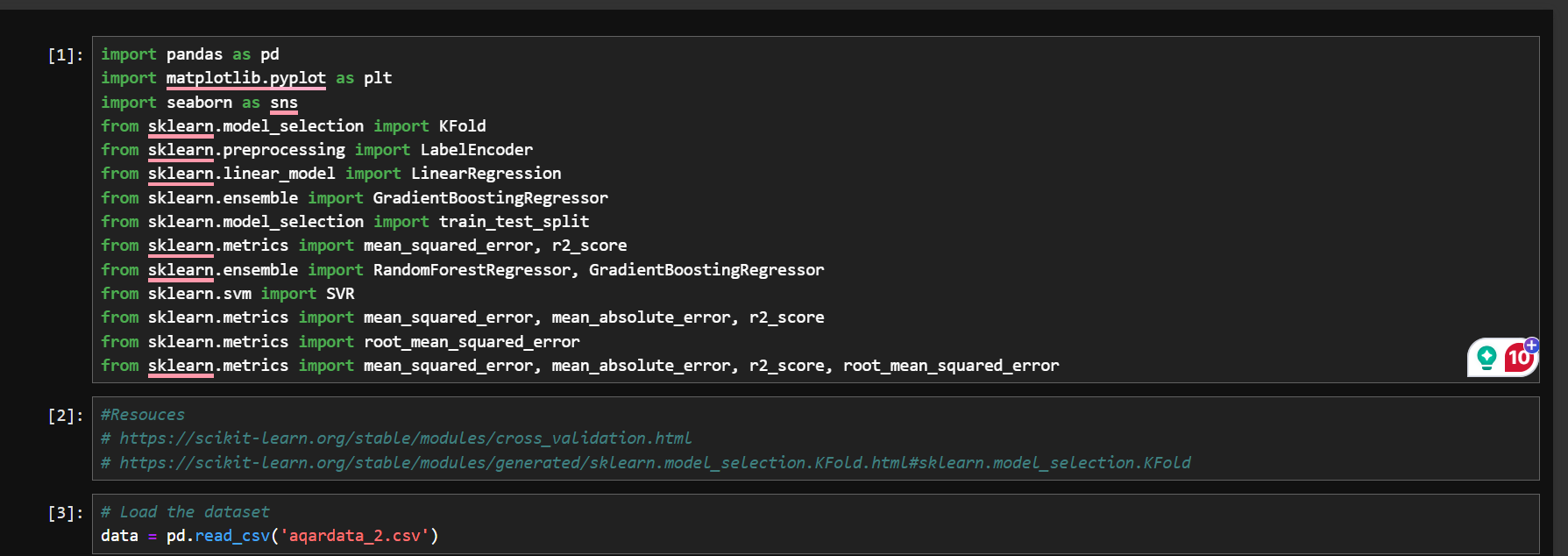
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| mainlocation | sublocation | neighborhood | frontage | purpose | streetwidth | size |

* ] and make the Target Pricepm
* After the all steps we choose svr because it has highest score.

1. **Evaluation results:**

The results for HW3:

1. Best Model: Support Vector Regression
2. Mean Squared Error: 7081681.56
3. Mean Absolute Error: 1338.34
4. Root Mean Squared Error: 2661.14
5. R² Score: 0.00%
6. **Screenshots of all the code & evaluation results from the platform was chosen:**



A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

THE RESULT FOR merging solution: