

Regression task:

➔ You will find house_data dataset on the drive, use it to solve the assignment.

1. Dataset Exploration:

- Load the dataset and explore its structure.
- Identify the target variable and potential features for linear regression.

2. Data Preprocessing:

- Handle missing values: Decide on an appropriate strategy (imputation, removal, etc.).
- Explore and address outliers if present.
- Check for and handle any categorical features using encoding techniques.

3. Exploratory Data Analysis (EDA):

- Visualize the distribution of the target variable.
- Examine the relationships between features and the target through pair plots or correlation matrices.

4. Feature Scaling:

- Apply feature scaling if needed to ensure that all features are on a similar scale.

5. Train-Test Split:

- Split the dataset into training and testing sets.

6. Linear Regression Model:

- Build a linear regression model using a suitable library (e.g., scikit-learn).
- Train the model on the training dataset.

7. Model Evaluation:

- Evaluate the model's performance on the testing set.
- Use appropriate regression metrics (e.g., Mean Squared Error, R-squared).

8. Visualization:

- Visualize the predicted values against the actual values using scatter plots or regression plots.

9. Conclusion:

- Summarize the findings from the linear regression task.
- Reflect on the model's performance and potential areas for improvement.