### Assignment 2: Regression Task with evaluation methods

**Objective**:

Train a regression model on a real-world dataset, emphasizing both prediction accuracy and interpretability.

**Dataset**:

Use the [Boston Housing dataset](https://www.kaggle.com/code/prasadperera/the-boston-housing-dataset) from the UCI repository. This dataset measures various factors about houses in Boston suburbs and aims to predict the median value of owner-occupied homes.

**Data Exploration & Preprocessing**:

* Load the dataset.
* Handle missing values if any.
* Visualize the distribution of the target variable (MEDV - Median value of homes).
* Explore relationships between predictors and the target variable using scatter plots or correlation matrix.

**Data Splitting**:

Divide the dataset into a training set (70%) and a test set (30%).

**Model Selection & Training**:

* Choose a regression algorithm of your choice (e.g., Linear Regression, Decision Trees, SVM regression, etc.).
* Train the model using the training data.

**Model Evaluation**:

* Use the test set to evaluate the model.
* Calculate the following evaluation metrics on the test set:
  + Mean Absolute Error (MAE)
  + Mean Squared Error (MSE)
  + Root Mean Squared Error (RMSE)
  + R-squared (Coefficient of Determination)
  + Adjusted R-squared

**Feature Importance**:

* Depending on the chosen model, determine the importance of each feature.
* Discuss the significance of each feature in predicting the target variable.

**Improvement (Bonus)**:

* Apply at least one technique to improve the model's performance, such as:
  + Feature engineering.
  + Polynomial regression.
  + Regularization techniques (L1/L2).
* Re-evaluate the model using the metrics and compare with the initial model.

**Submission Guidelines**:

* Submit a Jupyter Notebook or a Python script containing all the code used for the assignment.
* The code should be well-commented to explain your reasoning at each step.
* Include visualizations for data exploration, feature importance, and results.
* A report (1-2 pages) summarizing your findings, the model's performance metrics, and any conclusions drawn from the exercise.

**Evaluation Criteria**:

1. Data Preprocessing: Clean handling and transformation of data.
2. Implementation: Correctness and clarity of code.
3. Evaluation: Proper and correct computation of metrics.
4. Interpretation: Insight into feature importance and model performance.
5. Improvement: Effectiveness and clarity of the improvement technique.

*By the end of this assignment, you should have a solid understanding of regression tasks, the intricacies of feature interactions, and the significance of model interpretability.*