



AL2002 – Artificial Intelligence Lab

Lab Task # 11

Note:

- Plagiarism will not be **tolerated!!**
- Use comments wherever applicable.
- Please ensure to submit both a PDF document and a Python file containing your code on the classroom platform.

Problem: 1 - Linear Regression Analysis on the Boston Housing Dataset

Explore the renowned '**Boston Housing**' dataset, containing information about various factors affecting housing prices in Boston suburbs. Delve into the dataset's features, such as crime rate, nitric oxides concentration, and more,

1. Dataset Download: Obtain the 'Boston Housing' dataset, renowned for its comprehensive information on factors affecting housing prices in Boston suburbs.

2. Data Preprocessing: Prior to analysis, preprocess the dataset. Handle missing values, outliers, and consider necessary transformations to ensure data quality and suitability for analysis.

3. Statistical Model Application: Apply linear regression to predict the median value of owner-occupied homes using the preprocessed dataset. Split the data into 80% training set and 20% test set. Calculate the mean squared error (MSE) on the test set to evaluate model performance. Determine and display the y-intercept and slope of the best-fitted line.

4. Visualization: Draw a chart representing the best-fitted line generated by the linear regression model. Visualize the relationship between predictors and the target variable to enhance understanding and interpretation of the model's performance.