

EXP-6-Python Lab Question

Objective: Use the provided dataset to design and evaluate multiple regression models.

Dataset: The dataset contains 1000 instances, with six continuous attributes (`Feature_1` to `Feature_6`) and one target attribute (`Target`).

Tasks:

1. Data Loading and Preprocessing:

- Load the dataset.
- Check for missing values and handle them if necessary.
- Split the data into training and testing sets (80% training, 20% testing).

2. Model Design:

- Design and implement the following regression models using the training data:
 - Linear Regression
 - Polynomial Regression (Degree 2)
 - Stepwise Regression (Implement a basic version using forward selection)
 - Ridge Regression
 - Lasso Regression
 - ElasticNet Regression

3. Model Evaluation:

- Evaluate each model on the test data.
- Display the following performance metrics for each model:
 - Mean Absolute Error (MAE)
 - Mean Squared Error (MSE)
 - R-squared (R^2)
- Plot the predicted vs actual values for each model.

4. Comparison:

- Summarize the performance of all models in a comparison table.
- Discuss which model performed best based on the evaluation metrics and explain why.

5. Visualization:

- Create visualizations comparing the performance of the different models.
- Plot the coefficients (if applicable) for the different models to compare how each model handles the feature importance.

Expected Output:

- Code implementation with detailed comments.
- Performance metrics and visualizations for each model.
- A final comparison table summarizing the performance of all models.
- Insights based on the comparison of models.

Note: Utilize libraries such as `pandas`, `numpy`, `scikit-learn`, `matplotlib`, and `seaborn` for data handling, modeling, and visualization.

This lab exercise will help you understand the nuances of different regression techniques and how to evaluate and compare their performance effectively.

Dr.Trilok Nath Pandey

SCOPE,VIT,Chennai