

Course:IC272

Assignment: Non Linear Regression

Dataset Description: You are given a CSV file “Non_Linear_data.csv” containing 1000 samples. Each sample has *input_feature* as independent variable and *Target* as dependent variable. Function governing the relationship between input and output is given by a polynomial function of degree p .

Question: Write a Python program to read the given data file (using Pandas) and perform the following tasks.

- 1) Visualise the data and Divide the data into training and testing data. Train, and test split should be 80% and 20% respectively.
- 2) Perform Polynomial Curve Fitting using simple non linear regression on the data for different values of $p=1$ to $p=5$.
- 3) Calculate the Mean Square Error(MSE) for training and testing data for all the cases.
- 4) For what value of p the MSE is minimum.
- 5) Compare the plots for actual and predicted labels for training as well as testing data for all the cases.
- 6) Inferences on the plots and inferences on the results observed..