Lab 7

School of Computer Science Engineering and Technology

Course: B. Tech.	**Type: Core**
Course Code: CSET301	Course Name: Artificial Intelligence and Machine Learning
Year: 2025	Semester: Odd
Date: [Insert Date]	Batch: 2023-2027

CO-Mapping

	CO1	**CO2**	**CO3**	**CO4**	**CO5**
Lab		√	√		

Lab Assignment 7: **Comparative Study of Linear Regression and Polynomial Regression**

Objective:
To understand and compare the performance of linear regression and polynomial regression models on a dataset by evaluating key performance metrics including R^2 (coefficient of determination) and MAE (Mean Absolute Error).

Dataset:
You may use the publicly available Auto MPG Dataset :
"https://raw.githubusercontent.com/mwaskom/seaborn-data/master/mpg.csv"

Instructions:

1. Data Loading and Preprocessing:
    - Load the dataset.
    - Identify features (independent variables) and target (dependent variable).
2. Model Implementation:
    - Implement Linear Regression from scratch or using libraries.
    - Implement Polynomial Regression (choose polynomial degree 2 or 3 for comparison).
3. Training and Prediction:
    - Split the dataset into training and testing sets.
    - Train both models on the training data.
    - Predict outputs on the test set.
4. Performance Metrics Calculation:
   Calculate and compare the following metrics for both models:

- R^2 (Coefficient of Determination): Measures how well the regression predictions approximate the real data points.
- MAE (Mean Absolute Error): Measures the average magnitude of errors in predictions, without considering their direction.

5. Plotting:
    - Plot the data points along with both regression lines/curves.
    - Visualize and interpret the fit of each model.
6. Discussion:
    - Compare the strengths and weaknesses of linear vs. polynomial regression based on metrics and plots.
    - Discuss when polynomial regression might be preferred.
    - Relate findings with overfitting and underfitting concepts.