

# HOMWORK 1: MANUAL IMPLEMENTATION OF SIMPLE AND MULTIVARIATE REGRESSION

(Duration: 4 days)

## Learning Objectives :

- ✓ Understand and manually implement simple and multivariate regression models.
- ✓ Implement both the least squares method and gradient descent manually without using the scikit-learn library.

## Exercise 1: Computing Simple Linear Regression Coefficients Manually using Least Squares

### Dataset:

You are given the following dataset of study hours () and exam scores ():

Hours Studied	Exam Score
1	10
2	11.50
3	13.50
4	14
5	15

1. Compute the slope and intercept manually, then implement the calculations in Python.
2. Use the computed and to predict the exam score for a student who studied 6 hours.

## Exercise 2: Implementing Simple Linear Regression using Gradient Descent

### Dataset:

Use the same dataset from Exercise 1.

1. Compute the gradient descent manually for 2 iterations. Set initial values:
  - $\theta_0 = \theta_1 = 0$
  - Learning rate:  $\alpha = 0.01$
  - Number of iterations: 2(Show for each iteration, the values of  $\theta_0$  and  $\theta_1$ )
2. Implement gradient descent in Python using loops. Set initial values:
  - $\theta_0 = \theta_1 = 0$
  - Learning rate:  $\alpha = 0.01$
  - Number of iterations: 1000
3. Plot the cost function over iterations to check for convergence.
4. Compare the final values of  $\theta_0$  and  $\theta_1$  with the least squares solution from Exercise 1. What do you observe and recommend?

5. Additional Question: Generate synthetic data using a known linear equation with random noise and apply gradient descent to estimate the parameters.

### Exercise 3: Computing Multivariate Regression Coefficients Manually (Least Squares)

#### Dataset:

You are given data about students' study hours and sleep hours, and their corresponding exam scores:

Hours Studied ( $x_1$ )	Sleep Hours ( $x_2$ )	Exam Score ( $y$ )
1	8	10
2	7	11.50
3	6	13.50
4	6	14
5	5	15

1. Compute the slope and intercept manually, then implement the calculations in Python.
2. Use the computed values to predict the exam score of a student who studied for 6 hours and slept for 5 hours.

### Exercise 4: Implementing Multivariate Regression using Gradient Descent.

#### Dataset:

Use the same dataset from Exercise 2.

1. Compute the gradient descent manually for 2 iterations. Set initial values:
  - $\theta_0 = \theta_1 = \theta_2 = 0$
  - Learning rate:  $\alpha = 0.01$
  - Number of iterations: 2(Show for each iteration, the values of  $\theta_0$ ,  $\theta_1$  and  $\theta_2$ )
2. Implement gradient descent in Python using loops. Set initial values:
  - $\theta_0 = \theta_1 = \theta_2 = 0$
  - Learning rate:  $\alpha = 0.01$
  - Number of iterations: 10000
3. Compare the results with the least squares solution from Exercise 3.
4. Additional Question: Effect of Learning Rate ( $\alpha$ ):
  - a. Implement gradient descent with different values of  $\alpha$  (e.g., 0.001, 0.01, 0.1, 1).
  - b. Plot the cost function over iterations for each value.
  - c. What do you observe when  $\alpha$  is too small or too large?
5. Additional Question: Effect of Number of Iterations:
  - a. Train the model with different iteration values (e.g., 100, 1000, 10,000).
  - b. Plot the cost function for each case.
  - c. How do you decide the optimal number of iterations?
  - d. How can you stop training automatically when the model has converged?

.../...