

Supervised Learning Assignment 2



Cairo University, Faculty of Computers
and Artificial Intelligence

FACULTY OF COMPUTERS AND AI, CAIRO UNIVERSITY

Supervised Learning Year 2024-2025 Second Semester

Assignment #2



Assignment #2

Objective

In this assignment, you will implement **Batch Gradient Descent (BGD)**, **Stochastic Gradient Descent (SGD)**, and **Mini-Batch Gradient Descent (MBGD)** for both **Linear Regression** and **Logistic Regression** from scratch. You will then compare their performance against Scikit-Learn's built-in models.

Tasks

Task 1: Implement Linear Regression

- **Generate synthetic data**
 1. Use `numpy` to generate a dataset based on $y = 4x - 2 + \text{noise}$ (or use ready dataset from the internet)
 2. Split the dataset into **training (80%)** and **testing (20%)**
- **Implement three gradient descent methods from scratch:**
 1. Batch Gradient Descent (BGD)
 2. Stochastic Gradient Descent (SGD)
 3. Mini-Batch Gradient Descent (MBGD)
- **Train and evaluate models**
 1. Train each model and compute the **Mean Squared Error (MSE)** on the test set
 2. Compare results with Scikit-Learn's built-in `LinearRegression`

Task 2: Implement Logistic Regression(for binary case)

- **Generate synthetic classification data**
 1. Use numpy to generate **binary classification data** where labels are 0 or 1 ,Use sklearn.datasets.make_classification() (or manually create a separable dataset or use ready dataset from the internet)
 2. Split the dataset into **training (80%) and testing (20%)**
- **Implement three gradient descent methods from scratch:**
 1. Batch Gradient Descent (BGD)
 2. Stochastic Gradient Descent (SGD)
 3. Mini-Batch Gradient Descent (MBGD)
- **Train and evaluate models**
 1. Train each model and compute the **accuracy** on the test set
 2. Compare results with Scikit-Learn's built-in Logistic Regression

Requirements

- ✓ Implement **gradient descent manually (without using Scikit-Learn)**
- ✓ Use **matplotlib** for visualization
- ✓ Compare with Scikit-Learn's LinearRegression and LogisticRegression
- ✓ Evaluate performance using **MSE (for Linear Regression) and Accuracy (for Logistic Regression)**
- ✓ Each team consists of 3 students