CS 422 Assignment 3 Report

Name(s) of the student(s) completing the assignment:

Ivan Biacan

The dataset you used, its source and characteristics:

https://www.kaggle.com/aungpyaeap/fish-market

Input Features of the dataset:

- Length 1
- Length 2
- Length 3
- Height
- Width

Output Feature of the dataset:

- Weight

The solution for both the OLS and the gradient descent algorithms.

Linear Regression Solution:

- Coefficients/Weights: [44.79159408 20.27127003 -37.89429046
  28.12434825 20.62305808]
- Intercept/Bias: [-513.21828267]

Gradient Descent Solution:

- Coefficients/Weights: [6.5895988 6.14817817 2.97524141 2.99147721
  2.03460054]
- Intercept/Bias: [-5.61527433]

The learning rate(s) you used for gradient descent and how many iterations it took for gradient descent to converge.

Learning Rates/Parameters:

- max_iter = 1000000
- alpha = .000001
- learning_rate = 'constant'
- eta0 = 1e-7

Iterations

- 9855

Relevant evaluation metrics (MSE, MAE, R2) for the training dataset for both algorithms.

Linear Regression:

- Mean Squared Error: 15437.23547451023
- Mean Absolute Error: 93.91140535396316
- R^2: 0.8797518697479778

Gradient Descent:

- Mean Squared Error: 48794.773604544425
- Mean Absolute Error: 194.75102188627872
- R^2: 0.6293806641806203

Relevant evaluation metrics (MSE, MAE, R2) for the test dataset with for both algorithms.

Linear Regression:

- Mean Squared Error: 12109.803906070745
- Mean Absolute Error: 88.3250802746941
- R^2: 0.8928702867801026

Gradient Descent:

- Mean Squared Error: 37048.47133591687
- Mean Absolute Error: 169.16317971803326
- R^2: 0.6634344946989937