**Fundamentals of Machine Learning (Fall 2022)**

**Homework #1 (80 Pts, Due date: Sep 21, 2022)**

**Student ID**

**Name**

**(1)** Given training samples, , we want to find a constant that minimizes the following error function.

Assume that we have five training samples () such that .

**(a) [10 pts]** Calculate the optimal using an analytical solution.

**Answer:**

|  |
| --- |
|  |

**(b) [10 pts]** Explain the meaning of the optimal solution in terms of normal distribution .

**Answer:**

|  |
| --- |
|  |

**(2)** We provide all template code and datasets in Python. Write your code to implement linear regression. You need to install NumPy and Matplotlib libraries.

**(a) [5 pts]** Implement the util function “add\_bias” in ‘models/LinearRegression.py.’ You should add a column of ones for bias after the last column of the input matrix.

**Note: Fill in your code (EDIT HERE part). You also have to submit your code to i-campus.**

**Answer:**

|  |
| --- |
|  |

**(b) [5 pts]** Implement the training function “numerical\_solution” in ‘models/LinearRegression.py’ using the **batch gradient descent method**. The error function is defined as follows:

**Note: Fill in your code (EDIT HERE part). You also have to submit your code to i-campus.**

**Answer:**

|  |
| --- |
|  |

**(c)** **[10 pts]** Implement training function “numerical\_solution” in ‘models/LinearRegression.py’ using the **mini-batch stochastic** **gradient descent method**. The error function is defined as follows:

**Note: Fill in your code (EDIT HERE part). You also have to submit your code to i-campus.**

**Answer:**

|  |
| --- |
|  |

**(d) [10 pts]** Implement the training function “analytical\_solution” in ‘models/LinearRegression.py’ using the **normal equation**.

**Note: Fill in your code (EDIT HERE part). You also have to submit your code to i-campus.**

**Answer:**

|  |
| --- |
|  |

**(3)** Evaluate your code using two datasets, “Wave” and “Diabetes.”

**(a) [10 pts]** After training your model on the “Diabetes” dataset, fill the blank using the following metrics.

RMSE =

Write your opinion briefly on two solutions, the analytical solution and the gradient descent method. (Hyperparameter for gradient: Epoch = 10,000, Batch size = 32, learning rate = 0.01, optimizer=’SGD’)

**Answer: Fill the blank in the table.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  |  | **RMSE** |
| Initial value | 0.0 | 0.0 | 0.0 | 168.7240 |
| Gradient Descent |  |  |  |  |
| Analytic solution |  |  |  |  |

1. **[10 pts]** For the “Wave” dataset, draw the plots by adjusting learning rates, where the other hyperparameters are the same as (a). For each plot, the x-axis is # of iterations, and the y-axis is the error value. Try at least five different learning rates.

**Answer:**

|  |
| --- |
|  |

1. **[10 pts]** For the “Wave” dataset, draw the plots by adjusting batch sizes, where the other hyperparameters are the same as (a). For each plot, the x-axis is # of iterations, and the y-axis is the error value. Try at least five different values.

**Answer:**

|  |
| --- |
|  |