Homework 3 Release On Feb.011.

Submission Instructions

1. All of your code.
2. One output pdf file named with your first and last name (Answer all Q1 and Q2 questions (Non-coding Questions)).
3. Submit all of above files in a single ZIP file named with your first and last name.
4. Due Date 2024 Feb.18 11:59pm

Q1 Brief Questions

1. Linear Regression: Explain the concept of linear regression and its primary objective. (5')

2. Logistic Regression: Why do we need logistic regression compared with Linear Regression? (5 ')

3. Explain the differences and aims among the following figures: (20')

Figure of Model Fitting Data Fig.1: This figure typically shows the scatter plot of the actual data points along with the regression line or curve fitted by the model.

Figure of Cost Function versus Parameters (weights and bias) Fig.2: This figure displays the cost function (such as mean squared error for linear regression or cross-entropy loss for logistic regression) as a function of the model parameters (weights and bias).

Figure of the Learning Curve Fig.3: The learning curve shows the performance of the model (such as training and validation error) as a function of the number of training iterations or the size of the training dataset.

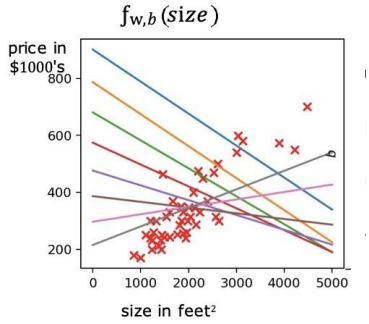

Fig.1

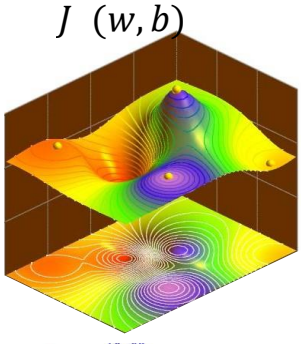

Fig.2

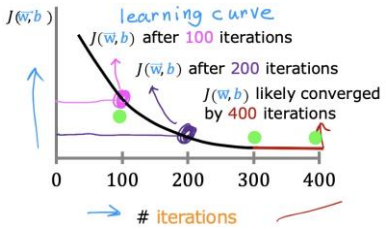

Fig.3

Q2 Calculation Question

You have the option to provide your answers in writing or to implement the solution through coding.

Consider the following dataset depicting the correlation between the number of hours studied (X) and the exam scores (y) of three students. Your task is to compute the slope (w) and bias/intercept (b) of the regression line using gradient descent. Model: y_est = wx+b

Hours Studied x	Exam Score y
2	68
3	75
4	82

Begin with initial values of w and b set to 0.1 and 0.1, respectively. Utilize a learning rate α of 0.01 and execute gradient descent for 3 iterations. Employ the Mean Squared Error Cost Function.

1. Illustrate the step-by-step process of updating w and b via the gradient descent algorithm. Output w, b, estimation on Exam score and J (w,b) at each iteration. (20')

2. Ultimately, utilize the computed values of w and b to forecast the exam score for a student who studies for 5 hours. (10')

Q3 Coding Questions:

Provide the following table:

Hours Studied (X1)	**Tutoring Hours (X2)**	**Previous Exam Score (X3)**	**Sleep Hours (X4)**	**Pass Exam Result (y)**
2	1	70.8	7	0
3	2	65.2	6	0
4	1	80.4	8	1
5	3	75.5	7	1
6	2	85.7	8	1
3	1	68.8	6	0
4	2	72.9	7	0
6	3	78.0	8	1
7	2	85.3	8	1
5	1	62.4	6	1
3	1	65.5	7	0
4	2	70.6	6	1
6	3	80.7	7	1

7	2	75.7	8	1
4	1	68.8	7	0
5	2	75.9	6	1
2	1	72.2	8	0
3	2	80.4	7	0
5	3	85.2	6	1
6	2	70.6	8	1

1. How many features does the dataset contain? (5’)

2.How many samples are included in the dataset? (5’)

3.Please proceed with writing your own code to implement logistic regression using gradient descent. Please find the dataset in the zip file. (20’)

Utilize a learning rate (α) of 0.01 and execute gradient descent for 100 iterations.
Employ the Cross-Entropy Error Cost Function.

Feel free to utilize libraries such as NumPy for mathematical computations. Customize the initial weights and bias based on your preference

4. Plot the learning curve for your logistic regression model. Then, whether the chosen learning rate is appropriate or not? If it's suitable, discuss why. If not, analyze the reasons behind it. (10’)