Tutorial 6

Question 1 (Linear Regression)

Imagine you're working on a basic sentiment analysis task. You have collected a small dataset where each data point consists of the number of positive words in a movie review and the corresponding rating given by the reviewer on a scale of 1 to 10. Now use a simple linear regression model to predict the movie rating based on the number of positive words. Your model should be defined as: $y = w \times x + b$, where x = number of positive words. Use the following data to derive

1) The optimal weight w and bias b using the least squares method.
2) Once you have the parameters, what would be the predicted rating for a review with 4 positive words?

Number of Positive Words	Movie Rating
1	3
3	6
5	8
6	9

Question 2 (Logistic Regression)

You're trying to predict whether a student will pass or fail an upcoming exam based on the number of hours they've studied. You have the following data:

Hours studied (x)	Pass (1) or Fail (0)
1	0
2	0
4	1
5	1

Now use binary logistic regression to formalize this problem.

1) Write the logistic regression function given the input x.
2) Compute the logistic loss for the first data point (x=1, y=0) given that w=1 and b=–3.
3) Derive the gradient of the logistic loss with respect to w and b. Then compute the gradient for the first data point using the same w and b values.

Coding exercises (for your own interest)

- Linear Regression: https://colab.research.google.com/drive/12QpBf7x_Jt6-zypN4OrUFFHXz1u6CmYe?usp=sharing
- Logistic Regression: https://colab.research.google.com/drive/1nTrYW5dUu6WO9cx7SGEvP9oX7qRbsGJk?usp=sharing