



## Sheet 4 Regression

### Questions

1. Consider a perceptron model consisting of two input neurons, each with two weights and a bias term, that feed into a single output neuron with one weight and a bias term. Provide a mathematical proof that illustrates the linear collapse and the benefits of using a non-linear activation function
2. Suppose you are studying the relationship between a student's high school GPA and their first-year college GPA. You collect data from four students and obtain the following information:

Student	High School GPA	First-Year College GPA
1	2.5	2.8
2	3.0	3.5
3	3.2	3.3
4	3.8	3.9

You want to use linear regression to predict a student's first-year college GPA based on their high school GPA. You can assume that the relationship between high school GPA and first-year college GPA is approximately linear.

- Compute the sample mean and sample standard deviation of the high school GPAs and first-year college GPAs.
  - Compute the slope and intercept of the least-squares regression line of first-year college GPA on high school GPA.
  - Use the regression equation to predict the first-year college GPA of a student who had a high school GPA of 3.5.
3. A researcher wants to investigate the relationship between the price of a house (in thousands of dollars) and its size (in thousands of square feet). She collects data from 4 houses and obtains the following information:
    - Compute the sample mean and sample standard deviation of the house prices and sizes.
    - Compute the slope and intercept of the least-squares regression line of house price on size.



House	Price	Size
1	120	2.0
2	140	2.5
3	130	2.2
4	150	2.7

- Use the regression equation to predict the price of a house that has a size of 3.0 thousand square feet.
4. Suppose you are analyzing a dataset of patients who have undergone a medical procedure, and you want to predict whether each patient experienced a positive outcome (1) or a negative outcome (0) based on their age (in years) and a binary indicator variable for whether they smoke (1 = smoker, 0 = non-smoker). You have data for five patients, as shown in the table below:

Patient	Age	Smoking	Outcome
1	45	0	1
2	52	1	0
3	27	1	1
4	31	0	0
5	48	1	1

- Compute the logistic regression coefficients for the model  $y = b_0 + b_1x_1 + b_2x_2$ .
  - Use the logistic regression equation to predict the probability that a 35-year-old non-smoker will have a positive outcome.
  - Compute the log-odds ratio and odds ratio for the effect of smoking on the probability of a positive outcome, holding age constant at its mean value.
5. Suppose you have a logistic regression model with two coefficients,  $B_0$  and  $B_1$ , and you are using mean squared error (MSE) as the loss function.
- Write down the expression for the overall loss in terms of  $B_0$  and  $B_1$
  - Derive the partial derivatives of the loss with respect to each coefficient
  - Explain the meaning of these partial derivatives in the context of training the logistic regression model and how they can be used to update the values of the coefficients during training
6. Consider the same logistic regression model in question 4 but assume that the model is trained using cross entropy loss as the loss function.



$$L(\boldsymbol{\theta}) = y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)$$

- Write down the expression for the overall loss in terms of B0 and B1
- Derive the partial derivatives of the loss with respect to each coefficient

## Deliverable

- This sheet is to be solved **Individually**.
- You are required to submit a ZIP file named **ID\_FirstName\_LastName\_sheet4.pdf**, any other naming format will not be accepted and file will be discarded.
- Any copied sheets will be immediately zeroed and other penalties may be applied.

**Good Luck**