

AI - Assignment - 4

Swati Sharma
2021568

DATE _____
PAGE _____

1. a. $y = \omega_1 x_1 + \omega_2 x_2 + \omega_3 x_3 + \omega_4 x_4 + \omega_5 x_5$

Mean Squared Error (MSE):

$$L(\omega; \{x^i, y^i\}) = \frac{1}{n} \sum_{i=1}^n (\omega^T x^i - y^i)^2 \quad i = 1, 2, \dots, N$$

(x^i, y^i) : labeled data samples available for optimization
 y^i : predictions from linear model

$$\Rightarrow L(\omega; \{x^i, y^i\}) = \frac{1}{n} \sum_{i=1}^n \left(\sum_{j=1}^5 \omega_j x_{j,i} - y^i \right)^2$$

$$\frac{\partial L}{\partial \omega_j} = \frac{1}{n} \sum_{i=1}^n \left(\frac{\partial L}{\partial e_i} \cdot \frac{\partial e_i}{\partial \hat{y}^i} \cdot \frac{\partial \hat{y}^i}{\partial \omega_j} \right)$$

$$\begin{aligned} e_i &= y^i - \hat{y}^i \\ \hat{y}^i &= \omega^T \cdot x^i \end{aligned}$$

$$\frac{\partial L}{\partial \omega_j} = \frac{1}{n} \sum_{i=1}^n (2(y_i - \hat{y}^i) \cdot (x_{j,i}))$$

* GRADIENT UPDATE RULE:

$$\textcircled{1} \quad \frac{\partial L}{\partial \omega_1} = -\frac{2}{N} \sum_{i=1}^N (y_i - \hat{y}_i) x_{1,i}$$

$$\textcircled{2} \quad \frac{\partial L}{\partial \omega_2} = -\frac{2}{N} \sum_{i=1}^N (y_i - \hat{y}_i) x_{2,i}$$

$$\textcircled{3} \quad \frac{\partial L}{\partial \omega_3} = -\frac{2}{N} \sum_{i=1}^N (y_i - \hat{y}_i) x_{3,i}$$



$$\textcircled{4} \quad \frac{\partial L}{\partial \omega_4} = -\frac{2}{N} \sum_{i=1}^N (y_i - \hat{y}_i) x_{4,i}$$

$$\textcircled{5} \quad \frac{\partial L}{\partial \omega_5} = -\frac{2}{N} \sum_{i=1}^N (y_i - \hat{y}_i) x_{5,i}$$

* VECTOR FORM :

Linear regression in vector form : $y = \omega^T x + b$

$$\text{MSE} : L(\omega; \{x^i, y^i\}) = \frac{1}{N} \|X\omega - y\|^2$$

$$\nabla_{\omega} L = \frac{2}{N} X^T (X\omega - y)$$

$$\nabla_b L = \frac{2}{N} \sum_{i=1}^N (\omega^T x^i - y^i)$$

* GRADIENT UPDATE RULE :

a) Weight ω :

$$\omega' = \omega - \alpha \nabla_{\omega} L \quad [\alpha : \text{learning rate}]$$

$$\omega' = \omega - \frac{2\alpha}{N} X^T (X\omega - y)$$

b) Bias b :

$$b' = b - \alpha \frac{\partial L}{\partial b}$$

$$= b - \frac{2}{N} \sum_{i=1}^N (\omega^T x^i - y^i)$$