AML Question 4 (1)

Thursday, September 22, 2022 3:59 PM

Stochastic Gradient Descent Derivation to update weights for the given model:

$$y = w_0 + w_1 e^{-x_1} + w_2 x_1 + w_3 x_1 x_2$$

Sum of Squared Residuals of the equation

$$(y - w_0 - w_1 e^{-x_1} - w_2 x_1 - w_3 x_1 x_2)^2$$

Derivatives wrt to the weights

(1) d_0 derivative wrt w_0

$$d_0 = -2 (y - w_0 - w_1 e^{-x_1} - w_2 x_1 - w_3 x_1 x_2)$$

(2) d_1 derivative wrt w_1

$$d_1 = -2 e^{-x_1} (y - w_0 - w_1 e^{-x_1} - w_2 x_1 - w_3 x_1 x_2)$$

(3) d_2 derivative wrt w_2

$$d_2 = -2 x_1 (y - w_0 - w_1 e^{-x_1} - w_2 x_1 - w_3 x_1 x_2)$$

(4) d_3 derivate wrt w_3

$$d_3 = -2 x_1 x_2 \cdot (y - w_0 - w_1 e^{-x_1} - w_2 x_1 - w_3 x_1 x_2)$$

Updating the coefficients

$$\text{new } w_0 = (\text{learning rate}) \times d_0$$
$$\text{new } w_1 = (\text{learning rate}) \times d_1$$
$$\text{new } w_2 = (\text{learning rate}) \times d_2$$
$$\text{new } w_3 = (\text{learning rate}) \times d_3$$

Model Equation with Regularization

$$(y - w_0 - w_1 e^{-x_1} - w_2 x_1 - w_3 x_1 x_2)^2 + \alpha (w_0^2 + w_1^2 + w_2^2 + w_3^2)$$

Derivation of the additional term that needs to be added to update the weights respectively

$$2 \times \alpha \times w_0 \;,\; 2 \times \alpha \times w_1 \;,\; 2 \times \alpha \times w_2 \;,\; 2 \times \alpha \times w_3$$