Problem Set 2 – Supervised Learning

DS598 B1 – DL4DS

Spring, 2024

Problem 2.1

To walk "downhill" on the loss function (equation 2.5), we measure its gradient with respect to the parameters ϕ_0 and ϕ_1. Calculate expressions for the slopes $\partial L/\partial\phi_0$ and $\partial L/\partial\phi_1$.

Assuming that L = MSE….

$$L = \frac{1}{N} * \sum_{i=1}^{N} (y_i - (\phi_0 + \phi_1 * x_i))^2$$

Then, for calculating the gradience the slopes, we would need to calculate the partial derivative with respect to omega_0, the intercept of the line, which shows how it will be affected by the Loss function….

$$\frac{uL}{u\phi_0} = \frac{u*\left[\frac{1}{N}*\sum_{i=1}^{N}(y_i-(\phi_0+\phi_1*x_i))^2\right]}{u\phi_0} \Rightarrow = \frac{1}{N} * \sum_{i=1}^{N} [- 2* (y_i - (\phi_0 + \phi_1 * x_i))] \Rightarrow$$

$$\frac{uL}{u\phi_0} = \frac{-2}{N} * \sum_{i=1}^{N} (y_i - (\phi_0 + \phi_1 * x_i))$$

Then, for calculating the gradience the slopes, we would need to calculate the partial derivative with respect to omega_1, the slope of the line, which will show how it will be affected by the Loss function….

$$\frac{uL}{u\phi_1} = \frac{u*\left[\frac{1}{N}*\sum_{i=1}^{N}(y_i-(\phi_0+\phi_1*x_i))^2\right]}{u\phi_1} \Rightarrow = \frac{1}{N} * \sum_{i=1}^{N} [- 2x_i* (y_i - (\phi_0 + \phi_1 * x_i))] \Rightarrow$$

$$\frac{uL}{u\phi_1} = \frac{-2}{N} * \sum_{i=1}^{N} x_i * (y_i - (\phi_0 + \phi_1 * x_i))$$

Problem 2.2

Show that we can find the minimum of the loss function in closed-form by setting the expression for the derivatives from problem 2.1 to zero and solving for ϕ_0 and ϕ_1.

We will need to set both of the gradient expressions equal to 0….

Starting with omega_0…

$$0 = \frac{-2}{N} * \sum_{i=1}^{N} (y_i - (\phi_0 + \phi_1 * x_i)) \Rightarrow \sum_{i=1}^{N} (y_i - (\phi_0 + \phi_1 * x_i)) \Rightarrow (\sum_{i=1}^{N} * y_i) - (N\phi_0 + \phi_1 * \sum_{i=1}^{N} * x_i) \Rightarrow$$

$$N\phi_0 = (\sum_{i=1}^{N} * y_i) - \phi_1 * \sum_{i=1}^{N} * x_i \Rightarrow$$

$$\phi_0 = \frac{(\sum_{i=1}^{N} * y_i) - \phi_1 * \sum_{i=1}^{N} * x_i}{N}$$

Now, omega_1

$$0 = \frac{-2}{N} * \sum_{i=1}^{N} x_i(y_i - (\phi_0 + \phi_1 * x_i)) \Rightarrow \sum_{i=1}^{N} x_i(y_i - (\phi_0 + \phi_1 * x_i)) \Rightarrow (\sum_{i=1}^{N} x_i y_i) - (\sum_{i=1}^{N} x_i \phi_0 + \sum_{i=1}^{N} x_i \phi_1 x_i) \Rightarrow$$

$$0 = (\sum_{i=1}^{N} x_i y_i) - (\sum_{i=1}^{N} x_i \phi_0 + \sum_{i=1}^{N} x_i \phi_1 x_i) \Rightarrow 0 = (\sum_{i=1}^{N} x_i y_i) - \sum_{i=1}^{N} x_i(\phi_0 + \phi_1 x_i) \Rightarrow$$

$$\sum_{i=1}^{N} x_i(\phi_0 + \phi_1 x_i) = (\sum_{i=1}^{N} x_i y_i) \Rightarrow \sum_{i=1}^{N} x_i^2 \phi_1 = (\sum_{i=1}^{N} x_i y_i) - \sum_{i=1}^{N} x_i \phi_0 \Rightarrow$$

$$\phi_1 = \frac{(\sum_{i=1}^{N} x_i y_i) - \sum_{i=1}^{N} x_i \phi_0}{\sum_{i=1}^{N} x_i^2}$$