

Loss functions (MSE & Cross-Entropy)

A loss function tells the model how wrong its prediction is.
Training means adjusting parameters to reduce this loss.

- Mean Squared Error (MSE)
used for regression.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

- y_i = true value
- \hat{y}_i = predicted value
- Squaring penalizes large errors more

Goal: minimize squared distance between prediction and truth.

- Cross-Entropy Loss (Binary)
Used for classification.

$$\mathcal{L} = -\frac{1}{n} \sum_{i=1}^n [y_i \log(\hat{p}_i) + (1 - y_i) \log(1 - \hat{p}_i)]$$

- $y_i \in \{0, 1\}$
- \hat{p}_i = predicted probability
- High confidence + wrong prediction \rightarrow very large loss.

Goal: maximize probability of the correct class

MSE measures numerical error,

Cross-Entropy measures probability error.