

Deep Learning: Loss Functions

Deep Learning: Loss Functions

Overview

- Evaluates model performance during training
- Gradual improvement due to optimizer
- Is minimized during training
- Multiple loss functions for one model possible (one for each output variable)

Regression

Classification

Deep Learning: Loss Functions

Regression Loss Functions

Regression Losses

- Mean Squared Error $MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}$
- Mean Absolute Error $MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n}$
- Mean Bias Error $MBE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n}$
- Output layer has 1 node
- Typical activation function: linear

Deep Learning: Loss Functions

Binary Classification Loss Functions

Binary Cross Entropy

- Applicable for binary classification
- Most common
- Output layer has 1 node
- Typical activation function: sigmoid

$$CE = -(y_i \log \hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)$$

Deep Learning: Loss Functions

Binary Classification Loss Functions

Hinge Loss

- Also called SVM loss
 - Applicable for binary classification
 - Used for maximum margin classifiers
 - Output layer has 1 node
 - Typical activation function: sigmoid
-
- $HingeLoss = \sum_{j \neq y_i} \max(0, s_i - s_{y_i} + 1)$

Deep Learning: Loss Functions

Multi-Label Classification Loss Functions

Multi-Label Cross Entropy

- Most common loss for multi-label classification
- Output layer has n nodes, where n is number of labels
- Typical activation function is softmax