Loss Functions

Loss functions quantify how well (or not well!) a model is performing.

## Classification

In classification models, it measures the difference between the predicted probabilities and the actual labels.

**Binary Cross-Entropy Loss**

Also known as log loss or BCE, is a loss function commonly used in binary classification problems. It quantifies the difference between the two probability distributions; the true distribution (labels) and the predicted distribution (model output).

For each datapoint, the loss is calculated as:

Where:

* y is the true label (1 or 0)
* p is the predicted probability of the label being 1

The logic behind this is if y = 1 (positive class) the loss is – log(p). This means the loss grows larger as p approaches 0. If y = 0 (actual negative) the loss is -log(1-p). Meaning the loss grows larger as p approaches 1.

BCE is not a direct measure of accuracy, however. A lower BCE score shows that outputs are closer to labels which suggests better accuracy and vice versa. However, BCE loss is a continuous value. Even if a prediction is correct, if the probability is not high (around 0.5), BCE loss will still be high.

The loss is quantified for the entire distribution of values. This can be done as a sum or normally the mean average. Represented by the equation: