

Training proceeds by minimizing a loss function.

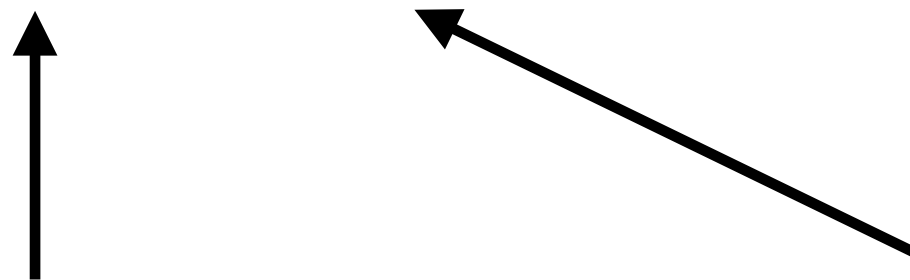
Typical loss functions

Squared error: $(y_i - \hat{y}_i)^2$

Cross-entropy: $-y_i \log(\hat{y}_i) - (1 - y_i) \log(1 - \hat{y}_i)$

True label (0 or 1)

NN output



The diagram illustrates the components of the cross-entropy loss function. It shows the formula $-y_i \log(\hat{y}_i) - (1 - y_i) \log(1 - \hat{y}_i)$. An arrow points from the text 'True label (0 or 1)' to the variable y_i in the first term. Another arrow points from the text 'NN output' to the variable \hat{y}_i in the first term.