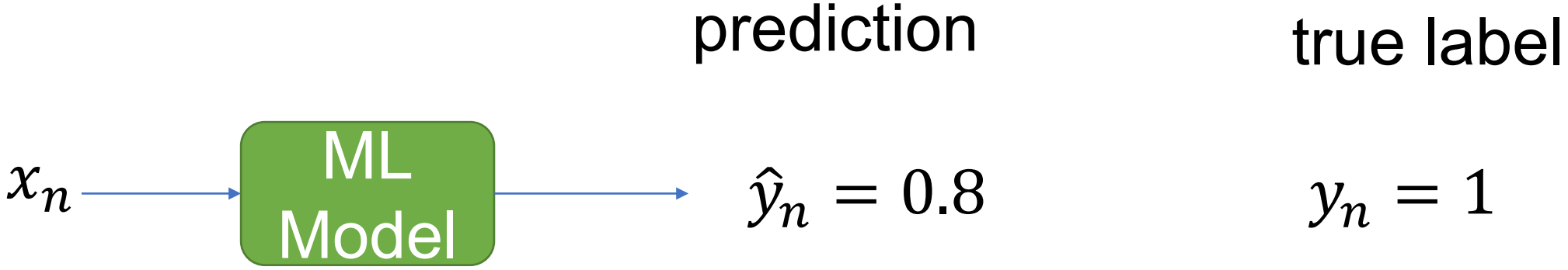
Binary-class classification



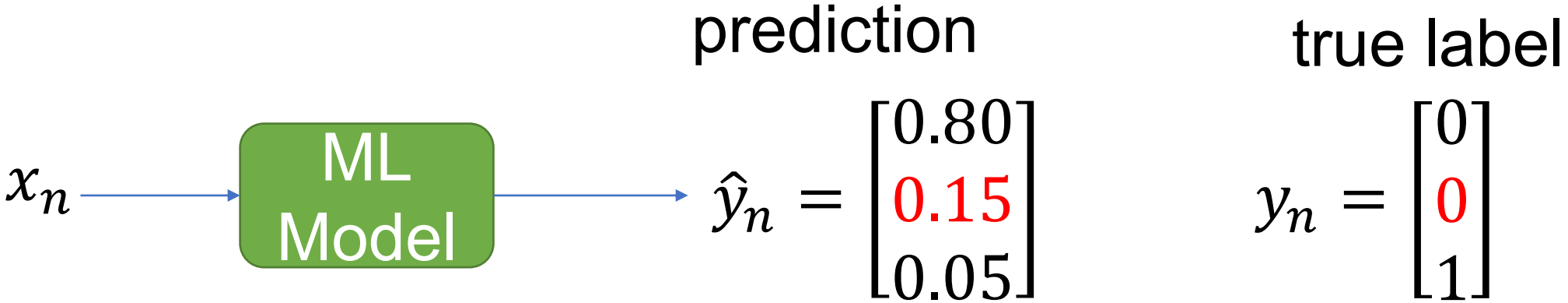
The binary cross-entropy loss of this sample is

$$L(y_n, \hat{y}_n) = -[1 \times \log(0.8) + 0 \times \log(1 - 0.8)]$$

$\log$ is the natural log

y_n could be 0 or 1

Multi-class classification



The cross-entropy loss of this sample is

$$L(y_n, \hat{y}_n) = -[0 \times \log(0.80) + 0 \times \log(0.15) + 1 \times \log(0.05)]$$

log is the natural log

y_n is a one-hot vector

The cross-entropy loss of a sample is $L(y_n, \hat{y}_n)$

The cross-entropy loss of the training samples is

$$Loss = \frac{1}{N}\sum_{n=1}^{N} L(y_n, \hat{y}_n)$$