machine learning
ML
E
T
P
T_P
P_E
supervised learning
T
f
$x_i \in \mathcal{X}$
$y_i \in \mathcal{Y}$
E
N
$\mathcal{D} = \{(x_n, y_n)\}_{n=1}^{N}$
training set
classification
C
classes
$\mathcal{Y} = \{1, 2, \cdots, C\}$
pattern recognition
Empirical risk minimization
misclassification rate

$$\mathcal{L} \frac{1}{N} \sum_{n=1}^{N} I(y_n \neq f(x_n; \theta)) \tag{1}$$

loss function
$\ell(y, \hat{y})$
??
empirical risk

$$\mathcal{L}(\theta) \frac{1}{N} \sum_{n=1}^{N} \ell\left(y_n, f\left(\mathbf{x}_n; \theta\right)\right) \tag{2}$$

zero-one loss

$$\ell_{01} = I(y \neq \hat{y}) \tag{3}$$

modle fitting
training

$$\hat{\theta} = \arg\min_{\theta} \mathcal{L}(\theta) = \arg\min_{\theta} \frac{1}{N} \sum_{n=1}^{N} \ell(y_n, f(x_n; \theta)) \tag{4}$$

Uncertainty
epistemic uncertainty
model uncertainty
aleatoric uncertainty