

§ 1 绪论

1. 基本术语

dataset $\begin{cases} \text{instance 实例, sample 样本} \\ \text{attribute 属性, feature 特征} \\ \text{dimensionality 维度} \end{cases}$

training data/set \leftrightarrow test data/set

prediction $\begin{cases} \text{离散型 classification} \\ \text{连续型 regression} \end{cases}$ $\begin{cases} \text{binary} \\ \text{multi-class} \end{cases}$ supervised learning (labeled data)

clustering 聚类

半监督学习 强化学习

\rightarrow generalization 泛化能力

unsupervised learning.

2. 假设空间

归纳 induction 演绎 (deduction)

泛化 generalization 特化 specialization

归纳的偏置 inductive bias — Occam's razor 奥卡姆剃刀