

第一章，统计学习方法概论

1. 经验风险最小化，当数据量过小时，学习能力相对强，会产生过拟合的现象；为解决过拟合问题，提出了结构风险最小化（在经验风险的基础上加入正则项）
2. 正则项一般和权重矩阵的模有关
3. 交叉验证
 1. 简单交叉验证（2折交叉验证）
 2. K折交叉验证
 3. 留一法
4. 当样本容量增加时，泛化误差趋于零，当假设空间变大时，泛化误差增大。
5. 统计学习3要素：模型，策略，方法
6. 经验风险最小化是指训练集中的平均误差，期望风险是指所有的数据（不限于训练集中的，通常用X和Y的分布函数来表示，一般不可得），结构风险就是在经验风险的基础上加正则项。
7. 生成模型（朴素贝叶斯和隐马尔科夫模型等）和判别模型（k近邻，感知级，决策树，支持向量机等。）的区别？
8. 评价指标

$$\text{精确率} \quad P = \frac{TP}{TP + FP}$$

$$\text{召回率} \quad R = \frac{TP}{TP + FN}$$

9. 标注问题，分类问题（预测问题），回归问题。
- 10.