

统计学习方法总结

- 一共10种学习方法：感知机，K近邻法，朴素贝叶斯法，决策树，逻辑斯蒂回归与最大熵模型，支持向量机，提升方法，EM算法，隐马尔可夫模型，条件随机场。

方法	适用问题	模型特点	模型类型	学习策略	学习的损失函数	学习算法
感知机	二类分类	分离超平面	判别模型	极小化误分类点到超平面距离	误分点大超平面距离	随机梯度下降
K近邻法	多类分类, 回归	特征空间, 样本点	判别模型	-----	-----	-----
朴素贝叶斯法	多类分类	特征与类别的联合概率分布, 条件独立假设	生成模型	极大似然估计, 极大后验概率估计	对数似然损失	概率计算公式, EM算法
决策树	多类分类, 回归	分类树, 回归树	判别模型	正则化的极大似然估计	对数似然损失	特征选择, 生成, 剪枝
逻辑斯蒂回归与最大熵模型	多类分类	特征条件下类别的条件概率分布, 对数线形模型	判别模型	极大似然估计, 正则化的极大似然估计	逻辑斯蒂损失	改进的迭代尺度算法, 梯度下降, 拟牛顿法
支持向量机	分类	分离超平面, 核技巧	判别模型	极小化正则化合页损失, 软间隔最大化	合页损失	序列最小最优优化算法(SMO)
提升方法	二类分类	弱分类器的线性组合	判别模型	极小化加法模型的指数损失	指数损失	前向分步加法算法
EM算法	概率模型参数估计	含隐变量概率模型	--- --- -	极大似然估计, 极大后验概率估计	对数似然损失	迭代算法
隐马尔可夫模型	标注	观测序列与状态序列的联合概率分布模型	生成模型	极大似然估计, 极大后验概率估计	对数似然损失	概率计算公式, EM算法

方法	适用问题	模型特点	模型类型	学习策略	学习的损失函数	学习算法
条件随机场	标注	状态序列条件下观测序列的条件概率分布，对数线性模型	判别模型	极大似然估计，正则化极大似然估计	对数似然损失	改进的迭代尺度算法，梯度下降，拟牛顿法