

8.2 比较支持向量机、AdaBoost、逻辑斯谛回归模型的学习策略与算法

1. 支持向量机的学习策略与算法

7.2.4节的合页损失函数：

对于线性支持向量机学习来说，其模型为分离超平面 $w^* \cdot x + b^* = 0$ 及决策函数 $f(x) = \text{sign}(w^* \cdot x + b^*)$ ，其学习策略为软间隔最大化，学习算法为凸二次规划。

线性支持向量机学习还有另外一种解释，就是最小化一下目标函数： $\sum_{i=1}^N [1 - y_i(w \cdot x_i + b)] * + \lambda |w|^2$ 目标函数的第1项是经验损失或经验风险，函数 $L(y(w \cdot b + x)) = [1 - y_i(w \cdot x_i + b)] *$ 被称为合页损失函数，第2项是系数为 λ 的 w 的 L_2 范数，是正则化项。

7.4节的序列最小最优化算法：

SMO算法是一种启发式算法，其基本思路是：如果所有变量的解都满足此最优化问题的KKT条件，那么这个最优化问题的解就得到了。因为KKT条件是该最优化问题的充分必要条件。

整个SMO算法包括两个部分：求解两个变量二次规划的解析方法和选择变量的启发式方法。

- 支持向量机的学习策略：软间隔最大化、最小化由合页损失函数和正则化项组成的目标函数
- 支持向量机的学习算法：凸二次规划、SMO算法（序列最小最优化算法）

1. 核技巧的应用：

- 在实际问题中，线性分类可能无法很好地分离数据。这时可以使用核技巧将输入空间映射到高维特征空间，使得数据在高维空间中线性可分。常用的核函数包括线性核、多项式核、高斯核等。

2. 对偶问题：

- SVM的原始问题是一个凸二次规划问题，但通过拉格朗日对偶性质可以得到对偶问题，通过求解对偶问题可以得到支持向量和决策函数。

3. 核函数的选择：

- 选择合适的核函数是SVM模型中的关键一步。不同的数据集和问题可能适合不同的核函数。

2. AdaBoost的学习策略与算法

8.3节的AdaBoost算法的解释：

AdaBoost算法还有另一个解释，即可认为AdaBoost算法是模型为加法模型、损失函数为指数函数、学习算法为前向分步算法时的二类分类学习方法。

给定训练数据及损失函数 $L(y, f(x))$ 的条件下，学习加法模型 $f(x)$ 成为经验风险极小化即损失函数极小化问题：
$$\min_{\beta_m, \gamma_m} \sum_{i=1}^N L\left(y_i, \sum_{m=1}^M \beta_m b(x_i; \gamma_m)\right)$$

定理8.3 AdaBoost算法是前向分步加法算法的特例。这时，模型是由基本分类器组成的加法模型，损失函数是指数函数。

- AdaBoost的学习策略：极小化通过加法模型组成的指数损失函数
- AdaBoost的学习算法：学习加法模型的前向分步算法

1. 弱分类器的选择：

- AdaBoost算法的性能与弱分类器的选择密切相关。弱分类器通常是简单的分类器，如决策树桩（单节点决策树）。

3.逻辑斯谛回归模型的学习策略与算法

6.1.3节的模型参数估计：

逻辑斯谛回归模型学习时，对于给定的训练数据集 $T = (x_1, y_1), (x_2, y_2), \cdots, (x_N, y_N)$ ，其中 $x_i \in R^n$ ， $y_i \in 0, 1$ ，可以应用极大似然估计法估计模型参数，从而得到逻辑斯谛回归模型。

6.3节的模型学习的最优化算法

逻辑斯谛回归模型、最大熵模型学习归结为以似然函数为目标函数的最优化问题，通常通过迭代算法求解。常用的方法有改进的迭代尺度法、梯度下降法、牛顿法或拟牛顿法。

- 逻辑斯谛回归模型的学习策略：极大似然估计法
- 逻辑斯谛回归模型的学习算法：改进的迭代尺度法、梯度下降法、牛顿法或拟牛顿法

	学习策略	算法
支持向量机	软间隔最大化、最小化由合页损失函数和正则化项组成的目标函数	凸二次规划、SMO算法（序列最小最优化算法）
AdaBoost	极小化通过加法模型组成的指数损失函数	学习加法模型的前向分步算法
逻辑斯谛回归	极大似然估计法	改进的迭代尺度法、梯度下降法、牛顿法或拟牛顿法