

assignment3

Xiaoma

2022 年 10 月 23 日

题目 1.

解： 线性判别分析能够解决多分类问题, 而 SVM 只能解决二分类问题。线性判别分析能将数据以同类样例间低方差, 不同样例中心之间大间隔来投射到一条直线上, 但是如果样本线性不可分, 那么支持向量机和线性判别分析就不能有效进行。当两类样本线性可分时, 且处理二分类问题时等价。

题目 2.

解： 1、SVM 的基本形态是一个硬间隔分类器, 它要求所有样本都满足硬间隔约束 (即函数间隔要大于 1)。

2、当数据集中存在噪声点但是仍然满足线性可分的条件时, SVM 为了把噪声点也划分正确, 超平面就会向另外一个类的样本靠拢, 这就使得划分超平面的几何间距变小, 从而降低了模型的泛化性能。

3、当数据集因为存在噪声点而导致已经无法线性可分时, 此时就使用了核技巧, 通过将样本映射到高维特征空间使得样本线性可分, 这样就会得到一个复杂模型, 并由此导致过拟合 (原样本空间得到的划分超平面会是弯弯曲曲的, 它确实可以把所有样本都划分正确, 但得到的模型只对训练集

有效), 泛化能力极差。

4、所以说, SVM 对于噪声很敏感。因此, 提出了软间隔 SVM 来防止由于噪声的存在而导致的过拟合问题。

题目 3.

解: 核对应率回归模型:

$$\ell(\beta) = \sum_{i=1}^m (-y_i \beta^T \hat{x}_i + \log(1 + e^{\beta^T \hat{x}_i}))$$

其中 $\hat{x}_i = (x : 1)$, $\beta = (w : b)$, $w^T x_i + b = \beta^T \hat{x}_i$

SVM 模型:

$$\min_{w,b} \frac{1}{2} \|w\|^2 \quad s.t. \quad y_i(w^T x_i + b) \geq 1 \quad i = 1, \dots, m$$

软间隔 SVM 模型:

$$\min_{w,b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \ell_{0/1}(y_i(w^T \phi(x_i) + b) - 1)$$

其中 $\ell_{0/1}$ 是 0/1 损失函数。

令 $Z = y_i(w^T \phi(x_i) + b)$, 使用对率损失:

$$\ell_{\log}(Z) = \log(1 + e^{-Z}) = \log(1 + e^Z) - Z$$

则

$$\min_{w,b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m (-Z + \log(1 + e^Z))$$

$$h(x) = w^T \phi(x) = \sum_{i=1}^m a_i k(x, x_i)$$

题目 4.

解: