人工智能基础作业 8

傅申 PB20000051

1.

以如下方式构造决策树. 每一层都只对同一个属性进行划分, 子结点对应该属性所有可能的取值, 依次遍历所有的属性, 这样得到的树的每个叶结点都与一个属性取值一一对应. 去掉训练集中不存在的分支后, 由于训练集中没有冲突数据, 所以每个叶结点都只对应一个类别, 该类别就是训练集中对应样本的类别. 这样就得到了与训练集一致的决策树.

2.

(1) 因为数据中不同的特征存在的误差程度是不同的, 所以规范化项应该对不同的特征有不同的权重, 因此选择 $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{w}$. $\boldsymbol{D}$ 的对角元素体现了对应特征的数据中存在的误差程度, $\boldsymbol{D}_{ii}$ 越大, 说明特征 i 的误差越大.

(2) 令 $\ell = (\boldsymbol{X}\boldsymbol{w}-\boldsymbol{y})^{\mathrm{T}}(\boldsymbol{X}\boldsymbol{w}-\boldsymbol{y}) + \lambda\boldsymbol{w}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{w}$, 则

$$\frac{\partial \ell}{\partial \boldsymbol{w}} = 2\boldsymbol{X}^{\mathrm{T}}(\boldsymbol{X}\boldsymbol{w}-\boldsymbol{y}) + 2\lambda\boldsymbol{D}\boldsymbol{w} = 0 \Rightarrow \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X} + \lambda\boldsymbol{D}\right)\boldsymbol{w} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{y}$$

得到闭式解

$$\boldsymbol{w}^* = \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X} + \lambda\boldsymbol{D}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{y}$$

3.

(1) $\forall (i,j), K_{i,j} = K\left(\boldsymbol{x}_i, \boldsymbol{x}_j\right) = \varphi(\boldsymbol{x}_i)\cdot\varphi\left(\boldsymbol{x}_j\right) = \varphi\left(\boldsymbol{x}_j\right)\cdot\varphi(\boldsymbol{x}_i) = K\left(\boldsymbol{x}_j, \boldsymbol{x}_i\right) = K_{j,i}$, 因此 K 是对称矩阵.

(2) 记 $\Phi = (\varphi(\boldsymbol{x}_1), \cdots, \varphi(\boldsymbol{x}_n))^{\mathrm{T}}$, 则 $K = \Phi\cdot\Phi^{\mathrm{T}}$, 因此, $\forall \boldsymbol{z}\in\mathbb{R}^n$, 有

$$\boldsymbol{z}^{\mathrm{T}}K\boldsymbol{z} = \boldsymbol{z}^{\mathrm{T}}\Phi\cdot\Phi^{\mathrm{T}}\boldsymbol{z} = \left(\boldsymbol{z}^{\mathrm{T}}\Phi\right)^2 \geq 0$$

所以 K 是半正定矩阵.

4.

K-means 算法一定会收敛. 考虑 K-means 算法的目标函数 $\ell = \sum_{i=1}^{k}\sum_{\boldsymbol{x}\in C_i}\left|\boldsymbol{x}-\boldsymbol{\mu}_i\right|^2$ 和算法的主要两步:

(1) 第一步中, 固定了 $\boldsymbol{\mu}$ 而优化 C, 使得 ℓ 减小;

(2) 第二步中, 固定了 C 而优化 $\boldsymbol{\mu}$, 使得 ℓ 减小.

即执行这两步后, ℓ 一定减小, 且 ℓ 有下界 0. 在 K-means 算法执行过程中, 目标函数值单调递减且有下界, ℓ 一定会收敛, 即 K-means 算法一定会收敛.