

2015

LOVE FOREVER...

5分

1. 关于ROC曲线 (1) "ROC"代表三个英文单词首字母, 请顺序写出这三个单词的全称 (3分) (2) 针对两类分类问题, 请描述"ROC"曲线的绘制步骤 (2分)

15分

2. 关于线性分类器 (1) 请给出Fisher线性判别分析的主要计算步骤和分类决策规则 (10分) (2) 请给出线性不可分情形下的支持向量机学习模型 (5分)

8分

3. 关于Boosting方法 假定有 n 个 m 维空间中的训练样本 $\{x_1, \dots, x_n\} \in \mathbb{R}^m$, 进一步假定这些样本属于两个不同的类别. (1) 基于这些样本, 采用Boosting方法学习一个两类分类器的步骤, 并指出拟采用的弱分类器 (4分) (2) 基于两类Boosting方法, 请设计一个针对多类分类问题的分类器学习框架 (4分)

8分

4. 关于聚类 (1) 请给出ISODATA方法的主要计算步骤 (4分) (2) 请指出ISODATA算法和C均值算法的不同点

8分

5. 关于主成分分析 (1) 请简述主成分分析方法的原理 (4分) (2) ^{采用}核技巧, 可以将主成分分析方法发展为核主成分分析方法, 请结合主成分分析方法解释核技巧的工作原理 (4分)

14分

6. 关于神经网络 (1) 针对多层前馈神经网络, 请给出反向传播算法的工作原理和训练步骤 (10分) (2) 请分析在前馈神经网络中, 隐含层数越多对分类预测可... 的影响 (4分)

10分

7. 关于特征选择 给定 n 个 m 维空间的训练样本 $\{x_1, \dots, x_n\} \in \mathbb{R}^m$, 且每个样本的类别标签是已知的. 请用迹地值判据 $J = \frac{\text{tr}(S_b)}{\text{tr}(S_w)}$ 设计一个完整的特征选择算法.

★ 10分

8. 现有一个二维空间中的两类分类问题, 记两个类分别为 w_1 和 w_2 . 假定 w_1 和 w_2 出现的先验概率相等, 且 w_1 和 w_2 的类条件概率密度均为正态分布. 其中 w_1 的类条件概率密度的均值为 $\mu_1 = [-1, 0]^T$, 协方差矩阵 $\Sigma_1 = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$, w_2 的类条件概率密度均值为 $\mu_2 = [1, 0]^T$, 协方差矩阵为 $\Sigma_2 = \begin{bmatrix} 1 & -0.5 \\ -0.5 & 1 \end{bmatrix}$, 其中上标 T 表示向量转置. 请写出贝叶斯对数似然比决策规则.

14分

9. 现有9个训练样本, 分别属于3个不同的类别. 第一类的样本点集为 $[1, 0]^T$ $[2, 0]^T$ $[1, 1]^T$, 第二类为 $[0, 1]^T$ $[-1, 0]^T$ $[-1, 1]^T$, 第三类为 $[0, -1]^T$ $[-1, -1]^T$ $[0, -2]^T$. 假定类先验概率均相等. 请计算类间散度矩阵和总类内散度矩阵.

8分

★ 10. 关于非线性特征提取 (1) 请描述 IsoMap 方法的主要计算步骤. (3分)
(2) 现有一个房间, 其四周墙壁均为白色, 且室内光照明亮均匀. 在该房间的一张白色的桌子上平放着一把彩色茶壶. 茶壶表面均绘有不同的花纹. 另室内有一个照相机可沿着一个圆移动. 该圆所在平面与桌面平行. 茶壶的质心与该圆的圆心重叠. 在照相机运动的过程中, 照像机的焦距等成像参数保持不变, 但其拍摄... 始终对准茶壶, 且能将茶壶定位于图像中央, 并使整个茶壶大致占满整个图像. 假设相机沿该圆匀速运动一周, 共连续拍摄到400张不同的图像. 现将 IsoMap 方法应用这些图像数据, 并假定最终输出一个二维特征提取结果. 如果将这400个二维特征在二维笛卡尔坐标系中进行绘制, 请分析可能呈现出的形状, 并给出理由. (5分)