

中国科学院大学

试题专用纸

课程编号: 09135404211

课程名称: 模式识别与机器学习

任课教师: 黄庆明、兰艳艳、郭嘉丰、山世光

注意事项:

1. 考试时间为 120 分钟, 考试方式 闭卷。

2. 全部答案写在答题纸上。

3. 考试结束后, 请将本试卷和答题卡、草稿纸一并交回。

1. (8 分) 试阐述线性判别函数的基本概念, 并说明既然有线性判别函数, 为什么还需要非线性判别函数? 假设有两类模式, 每类包括 6 个 4 维不同的模式, 且良好分布。如果它们是线性可分的, 问权向量至少需要几个系数分量? 假如要建立二次的多项式判别函数, 又至少需要几个系数分量? (设模式的良好分布不因模式变化而改变)

2. (8 分) 简述 SVM 算法的原理。如果使用 SVM 做二分类问题得到如下结果, 分别应采取什么措施以取得更好的结果? 并说明原因。

(1) 训练集的分类准确率 90%, 验证集的分类准确率 90%, 测试集的分类准确率 88%;

(2) 训练集的分类准确率 98%, 验证集的分类准确率 90%, 测试集的分类准确率 88%。

3. (8 分) 请从两种角度解释主成分分析 (PCA) 的优化目标。

4. (8 分) 请给出卷积神经网络 CNN 中卷积、Pooling、RELU 等基本层操作的含义。然后从提取特征的角度分析 CNN 与传统特征提取方法 (例如 Gabor 小波滤波器) 的异同。

5. (10 分) 用线性判别函数的感知器赏罚训练算法求下列模式分类的解向量, 并给出相应的判别函数。

$$\omega_1: \{(0 \ 0)^T, (0 \ 1)^T\}$$

$$\omega_2: \{(1 \ 0)^T, (1 \ 1)^T\}$$

6. (10 分) 试述 K-L 变换的基本原理, 并将如下两类样本集的特征维数降到一维, 时画出样本在该空间中的位置。

$$\omega_1: \{(-5 \ -5)^T, (-5 \ -4)^T, (-4 \ -5)^T, (-5 \ -6)^T, (-6 \ -5)^T\}$$

$$\omega_2: \{(5 \ 5)^T, (5 \ 6)^T, (6 \ 5)^T, (5 \ 4)^T, (4 \ 5)^T\},$$

其中假设其先验概率相等, 即 $P(\omega_1)=P(\omega_2)=0.5$ 。

(12 分) 请解释 AdaBoost 的基本思想和工作原理, 写出 AdaBoost 算法

8. (12分) 选择埃尔米特多项式, 其前几项的表达式为

$$H_0(x)=1, \quad H_1(x)=2x, \quad H_2(x)=4x^2-2,$$

$$H_3(x)=8x^3-12x, \quad H_4(x)=16x^4-48x^2+12$$

试用二次埃尔米特多项式的势函数算法求解以下模式的分类问题

$$\omega_1: \{(0 \ 1)^T, (0 \ -1)^T\}$$

$$\omega_2: \{(1 \ 0)^T, (-1 \ 0)^T\}$$

9. (12分) 已知以下关于垃圾邮件的8条标注数据, A、B为邮件的2个特征, Y为类别, 其中 Y=1 表示该邮件为垃圾邮件, Y=0 表示该邮件为正常邮件。请依此训练一个朴素贝叶斯分类器, 并预测特征为“A=0, B=1”的邮件是否为垃圾邮件。

序号	1	2	3	4	5	6	7	8
A	0	0	1	1	1	1	1	1
B	0	0	0	0	0	0	1	1

10. (12分) 假设有3个罐子, 每个罐子里都装有红、黑两种颜色的弹珠。按照下面的方法取弹珠: 开始, 以概率 π 随机选取1个罐子, 从这个罐子以概率 B 随机取出一个弹珠, 记录其颜色后, 放回; 然后, 从当前盒子以概率 A 随机转移到下一个盒子, 再从这个盒子里以概率 B 随机抽出一个球, 记录其颜色, 放回; 如此重复3次, 得到一个弹珠的颜色观测序列: $O=(\text{红}, \text{黑}, \text{红})$ 。请用前向传播算法计算生成该序列的概率 $P(O|\{A, B, \pi\})$ 。

$$\pi=[0.4, 0.4, 0.2]^T \quad A = \begin{matrix} & \begin{matrix} \text{罐子1} & \text{罐子2} & \text{罐子3} \end{matrix} \\ \begin{matrix} \text{罐子1} \\ \text{罐子2} \\ \text{罐子3} \end{matrix} & \begin{bmatrix} 0.3 & 0.5 & 0.2 \\ 0.2 & 0.3 & 0.5 \\ 0.1 & 0.4 & 0.5 \end{bmatrix} \end{matrix} \quad B = \begin{matrix} & \begin{matrix} \text{红} & \text{黑} \end{matrix} \\ \begin{matrix} \text{罐子1} \\ \text{罐子2} \\ \text{罐子3} \end{matrix} & \begin{bmatrix} 0.7 & 0.3 \\ 0.5 & 0.5 \\ 0.4 & 0.6 \end{bmatrix} \end{matrix}$$

2. 根据结果来看，SVM 在进行二分类时采用的是软间隔，其中 1 和 2 的训练集分类准确度不同，是因为 C 值设置的问题，C 表示对分类错误的惩罚程度，C 越大分类器就越不会允许出现分类错误现象，此时对应 2，C 越小分类器就越不会在乎训练集上的分类错误，此时对应 1，所以应该采取对 1 来说增大 C 值，对 2 来说减小 C 值。另外 C 值和间隔宽度有着互斥关系，C 越大导致间隔宽度表小。