

姓名\_\_\_\_\_ 学号\_\_\_\_\_ 成绩\_\_\_\_\_

- (10 分) 简述 Fisher 线性判别方法的基本思路，写出准则函数及对应的解。
- (12 分) 假设某个地区细胞识别中正常( $w_1$ )和异常( $w_2$ )两类的先验概率分别为：正常状态： $P(w_1) = 0.95$ ，异常状态： $P(w_2) = 0.05$ 。现有一待识别的细胞，其观察值为  $x$ ，已知  $p(x|w_1) = 0.2$ ， $p(x|w_2) = 0.5$ 。同时已知风险损失函数为：
$$\begin{pmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{22} & \lambda_{21} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 8 & 0 \end{pmatrix}$$
其中  $\lambda_{ij}$  表示将本应属于第  $j$  类的模式判为属于第  $i$  类所带来的风险损失。试对该待识别细胞用以下两种方法进行分类：
  - 基于最小错误率的贝叶斯决策，并写出其判别函数和决策面方程。
  - 基于最小风险的贝叶斯决策，并写出其判别函数和决策面方程。
- (10 分) SVM 可以借助核函数(kernel function)在特征空间(feature space)学习一个具有最大间隔的超平面。对于两类的分类问题，任意输入  $x$  的分类结果取决于下式：
$$\langle \hat{w}, \phi(x) \rangle + \hat{w}_0 = f(x; \alpha, \hat{w}_0)$$
其中， $\hat{w}$  和  $\hat{w}_0$  是分类超平面的参数， $\alpha = [\alpha_1, \dots, \alpha_{|SV|}]$  表示支持向量(support vector)的系数，SV 表示支持向量集合。使用径向基函数(radial basis function)定义核函数  $K(\cdot, \cdot)$ ，即  $K(x, x') = \exp(-\frac{D(x, x')^2}{2s^2})$ 。假设训练数据在特征空间线性可分，SVM 可以完全正确地划分这些训练数据。给定一个测试样本  $x_{far}$ ，它距离所有训练样本都非常远。
 

试写出  $f(x; \alpha, \hat{w}_0)$  在核特征空间的表达形式，进而证明： $f(x_{far}; \alpha, \hat{w}_0) \approx \hat{w}_0$
- (10 分) K-L 变换属于有监督学习(supervised learning)还是无监督学习(unsupervised learning)? 试利用 K-L 变换将以下样本集的特征维数降到一维，同时画出样本在该空间的位置。
$$\{(-5 - 5)^T, (-5 - 4)^T, (-4 - 5)^T, (-5 - 6)^T, (-6 - 5)^T, (5 5)^T, (5 6)^T, (6 5)^T, (5 4)^T, (4 5)^T\}$$
- (12 分) 过拟合与欠拟合。
  - 什么是过拟合? 什么是欠拟合?
  - 如何判断一个模型处在过拟合状态还是欠拟合状态?
  - 请给出 3 种减轻模型过拟合的方法。

6. (12 分) 用逻辑回归模型 (logistic regression model) 解决  $K$  类分类问题, 假设每个输入样本  $x \in \mathbb{R}^d$  的后验概率可以表示为:

$$P(Y = k|X = x) = \frac{\exp(w_k^T x)}{1 + \sum_{l=1}^{K-1} \exp(w_l^T x)}, \quad k = 1, \dots, K-1 \quad (1)$$

$$P(Y = K|X = x) = \frac{1}{1 + \sum_{l=1}^{K-1} \exp(w_l^T x)} \quad (2)$$

其中  $w_k^T$  表示向量  $w_k$  的转置。通过引入  $w_K = \vec{0}$ , 上式也可以合并为一个表达式。

- 1) 该模型的参数是什么? 数量有多少?
- 2) 给定  $n$  个训练样本  $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ , 请写出对数似然函数 (log likelihood function)  $L$  的表达式, 并尽量简化。

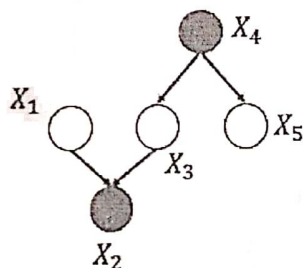
$$L(w_1, \dots, w_{K-1}) = \sum_{i=1}^n \ln P(Y = y_i | X = x_i)$$

- 3) 如果加入正则化项 (regularization term), 定义新的目标函数为:

$$J(w_1, \dots, w_{K-1}) = L(w_1, \dots, w_{K-1}) - \frac{\lambda}{2} \sum_{l=1}^K \|w_l\|_2^2$$

请计算  $J$  相对于每个  $w_k$  的梯度。

7. (10 分) 给定如下概率图模型, 其中变量  $X_2, X_4$  为已观测变量, 请问变量  $X_1$  和  $X_5$  是否独立? 并用概率推导证明之。



8. (12 分) 假设有 2 枚硬币, 分别记为 A 和 B, 以  $\pi$  的概率选择 A, 以  $1-\pi$  的概率选择 B, 这些硬币正面出现的概率分别是  $p$  和  $q$ 。掷选出的硬币, 记正面出现为 1, 反面出现为 0, 独立地重复进行 4 次试验, 观测结果如下: 1, 1, 0, 1。给定模型参数  $\pi = 0.4, p = 0.6, q = 0.5$ , 请计算生成该序列的概率, 并给出该观测结果的最优状态序列。
9. (12 分) 基于 AdaBoost 的目标检测需要稠密的扫描窗口并判断每个窗口是否为目标, 请描述基于深度学习的目标检测方法, 如 SSD 或 YOLO, 如何做到不需要稠密扫描窗口而能发现并定位目标位置?