**四川大学期末考试试题（闭卷）**

**（2020——2021学年第 2 学期） B卷**

课程号：304072020 课序号： 课程名称：模式识别引论 任课教师：卢晓春、赵启军 成绩：

适用专业年级：计算机学院2019级 学生人数：81 印题份数： 学号： 姓名：

|  |
| --- |
| **考 生 承 诺**  我已认真阅读并知晓《四川大学考场规则》和《四川大学本科学生考试违纪作弊处分规定（修订）》，郑重承诺：  1、已按要求将考试禁止携带的文具用品或与考试有关的物品放置在指定地点；  2、不带手机进入考场；  3、考试期间遵守以上两项规定，若有违规行为，同意按照有关条款接受处理。  **考生签名：** |
| 一、填空题（本大题共3小题，每空2分，共6分）  1. 模式识别是\_\_\_研究用计算机来实现人类模式识别能力的一门学科\_\_\_\_\_\_。  2. 两类别样本都服从一维正态分布且具有相同的方差，从概率分布角度来说，这两类别样本完全不可分的说法\_\_\_\_\_\_错误\_\_\_\_\_\_\_\_\_\_\_（正确、错误）。  3. 已知两类训练样本，[1,2]T属于ω1 ，[-1,0]T属于ω2，请对样本的增广样本向量进行规范化 [1,2,1] T 和 [1,0,-1] T 。  **二、问答题（本大题共5小题，每题10分，共50分）**  1. 解释“模式”与“模式类”的含义？“王小二是学生”这句话里，谁是模式，谁是模式类？  答：在模式识别学科中，把所见到的具体事物称为模式，而将它们归属的类别称为模式类。因此，王小二是模式，学生是模式类。  2. 以两类别样本为例，简述在最小错分类贝叶斯准则情况下，分界面由什么决定？如果想要尽可能避免把第二类样本错分为第一类样本的情况，可以采用什么办法？  答：分界面由先验概率和类条件概率密度决定。如果要避免，可以采用两种方式解决，一是采用调整风险系数的做法，加大把第二类错分为第一类的惩罚，二是采用限定把第二类的分类错误率为极小值的情况下，再去寻求第一类分类错误率的极小值。  3. 线性判别函数是一种较为简单的分类器，适用于任何情况。这个说法是否合适？请进行论述。  答：错误。论述答案不唯一，可自行展开，比如两类别样本的分界面呈现圆环形状时就不适用。  4. ISODATA算法与C均值算法相比有何优势？  解答：C均值算法受初始聚类中心的选择影响大，而且类别数相对不能改变。  ISODATA算法在C均值算法的基础上考虑了类别的分裂与合并，因此有了自我调整类别数的能力。  5. 特征选择是模式识别中的一个关键问题，但是一般情况下，穷举策略不可行。任选一种特征选择的搜索算法进行介绍。  答：非标答题，考生可任选最优搜索算法（分支定界法）或次优搜索算法（顺序前进、顺序后退、增 l减r法、遗传算法等）来进行阐述。  **三、计算题（本大题共2小题，每题12分，共24分）**  1. 已知二维空间中有两类共9个训练样本分别为： ω1：[-1,0]T, [-2,0] T, [-2,1] T, [-2,-1] T；  ω2： [1,1]T, [2,0] T, [1,-1] T, [2,1]T, [2,2] T，试分别用最近邻法和K近邻法求测试样本[0,0]T的分类，取K=3。根据所给训练样本，按照最近邻法绘制分界面。  答：  1）样本点到各个训练样本之间的距离依次为：  ω1：  ω2：  显然，按照最近邻法，最近邻样本点为(-1,0)T分类为1类，按照K近邻法，近邻样本为(-1,0)T, (1,1)T,(1,1)T则分类为2类。（5分）  2）分界面示意图（3分），由点的分布可知，分界面主要由(-1,0)T, (1,1)T,(1,1)T这三个点即可确定。 |

第 1 页，共 2 页

试卷编号：

|  |
| --- |
| 2. 设某类样本总体服从一维正态分布N(μ,δ)，参数未知。现有独立同分布样本x1，x2,…,xn，。试用最大似然估计来估算其均值μ和方差σ2。(要求：需给出详细步骤。)  解：令    （3分）  取对数似然函数    对上式关于θ求导可得：      （5分）  解方程可得：  （4分）  **五、综述题（本大题共1小题，每题20分，共20分）**  1. 如下所示道路图，试分别使用监督分类方法和非监督分类方法来实现车道线的识别？ |

第 2 页，共 2 页