

2013级模式识别复习课

模式识别复习

考试题型:

- 一、填空题（10个空，每空2分，共20分）
- 二、简答题（2题，每小题10分，共20分）
- 三、计算分类题（2题，每题12分，共24分）
- 四、证明题/设计题（1题，12分）
- 五、综合应用设计题（1题，24分）

考试范围：九个PDF课件文件(打*不考)，教材第1章、第2章、第4章、第5章、第6章、第9章(九个PDF课件中要求的内容)

模式识别复习

笔试形式：开卷考试，120分钟。

按照开卷考试的统一规定，开卷考试可以带书、纸质版资料、计算器，但开卷考试不能使用手机、平板电脑、笔记本电脑等电子设备查阅电子资料。

根据学校的规定，考试笔试成绩须在55分以上才有可能及格，否则按不及格处理。

模式识别复习

考试范围：

一、李弼程编《模式识别原理与应用》教材。

第1章 绪论

第2章 Bayes决策理论

第4章 线性判别分析(4.1 线性判别函数; 4.2 线性分类器：4.2.1-4.2.4)

第5章 特征提取和选择(5.1; 5.2.1, 5.2.2; 5.3.2; 5.4 主成份分析 PCA)

第6章 聚类分析(6.1; 6.3; 6.4.1, 6.4.2; 6.5)

第9章 神经网络模式识别(此章内容见“模式识别(九).PDF”课件的要求部分)

二、九个PDF课件文件。

PDF课件中打“*：了解”的部分不考，下面“九个PDF课件文件说明”中黑体字部分不考；

九个PDF课件文件说明：

1.模式识别(一).PDF

模式识别复习

九个PDF课件文件说明Cont. :

2.模式识别(二).PDF : 2.4 广义线性判别函数 , 不考

**3. 模式识别(三).PDF : 2.6 分段线性分类器设计、
2.7 非线性分类器设计 , 不考**

4.模式识别(四).PDF

**5.模式识别(五).PDF : 3.9 序贯分类、 3.10
Bayes分类器编程举例 , 不考**

模式识别复习

九个PDF课件文件说明Cont. :

6.模式识别(六).PDF : 4.5 聚类分析编程举例 , 不考

7.模式识别(七).PDF : 5.1 引言中的灰度共生矩阵、
5.3 类别可分性判据中基于熵函数的可分性判据、
5.4 基于可分性判据的特征提取中的基于概率
密度函数可分性判据的特征提取方法、 5.5 补充
PCA人脸图像的预处理方法及编程(编程, 不考)、
5.6 快速PCA及实现、 5.7 基于PCA的人脸特
征提取及实现(实现, 不考)

模式识别复习

九个PDF课件文件说明Cont. :

8.模式识别(八).PDF :

9.模式识别(九).PDF :

考试题型举例：

一、填空题举例

1、神经网络的特性及能力主要取决于网络拓扑结构及()。

答案：学习方法

2、()可以利用多个分类器之间的互补性，有效地提高分类的准确度。

答案：多分类器融合/多分类器集成

3、欧氏距离具有()和()。

- ①平移不变性 ②尺度缩放不变性
- ③旋转不变性 ④不受量纲影响的特性

答案：①，③

说明：一般是从PDF课件或从教材中找出正确答案，少数填空题是根据自己的理解填写出合适的答案。

二、简答题举例

1、简述统计模式识别方法。

答：统计模式识别把观察对象表达为一个随机向量(即特征向量), 将模式类表达为由有穷或无穷个具有相似数值特性的模式组成的集合。

识别是从模式中提取一组特性的度量, 构成特征向量来表示模式, 然后通过划分特征空间的方式进行分类。

统计模式识别系统构成：主要由信息获取、预处理、特征提取和选择以及分类器4部分组成；其中，分类器包括分类器设计和分类决策。

说明：一般是从课件或从教材中找出相关答案整理成段落、算法步骤等，也有可能根据自己的理解加工成段落、算法步骤等。

三、计算分类题

考试范围：读懂模式识别课件中的所有计算例题与思考题(如：**Bayes**公式在**Bayes**决策中的运用、感知器固定增量法、**Fisher**判别分析、**K-L**变换/主成份分析/**PCA**/主分量分析、聚类分析等)。

重要提醒1：线性代数中矩阵求逆的方法----特别是有时候会用到二阶矩阵求逆计算公式可加快二阶矩阵求逆的计算速度。

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

重要提醒2： 在Bayes公式在Bayes决策中的运用中，要加深对最小风险Bayes决策概念的理解：

(1)拒绝判决的处理：若允许拒绝判决，可将拒绝类看成独立的一类纳入到最小风险Bayes决策中；

(2)两类问题最小风险Bayes决策，可直接采用似然比形式的判决准则计算公式：

$$\frac{P(\mathbf{x} | \omega_1)}{P(\mathbf{x} | \omega_2)} > \frac{P(\omega_2)(\lambda_{21} - \lambda_{22})}{P(\omega_1)(\lambda_{12} - \lambda_{11})}, \text{则} \mathbf{x} \text{判为} \omega_1$$

$$\frac{P(\mathbf{x} | \omega_1)}{P(\mathbf{x} | \omega_2)} < \frac{P(\omega_2)(\lambda_{21} - \lambda_{22})}{P(\omega_1)(\lambda_{12} - \lambda_{11})}, \text{则} \mathbf{x} \text{判为} \omega_2 [\text{见教材} P13 \text{公式} (2-22)]$$

。 。 。 。 。 。

计算题举例：有两类样本

$$\omega_1=(x_1,x_2)=\{(1,0,1)^T,(0,1,1)^T\}$$

$$\omega_2=(x_3,x_4)=\{(1,1,0)^T,(0,1,0)^T\}$$

设初始权值 $W_1=(1,1,1)$ ， $\rho_k=1$ ，试用感知器固定增量法求判别函数，画出决策面。

解答：见模式识别课件(三).PDF，此略。

四、证明题/设计题

证明题举例：试证明欧氏距离具有平移不变性和正交旋转不变性。

证明：

$$\begin{aligned}\because d(\mathbf{x}, \mathbf{y}) &= \|\mathbf{x} - \mathbf{y}\| = \|(\mathbf{x} + \mathbf{u}_0) - (\mathbf{y} + \mathbf{u}_0)\| \\ &= d(\mathbf{x} + \mathbf{u}_0, \mathbf{y} + \mathbf{u}_0)\end{aligned}$$

\therefore 欧氏距离具有平移不变性.

∴ 正交变换矩阵 A 具有性质 $A \cdot A^T = I$

$$d(A\mathbf{x}, A\mathbf{y}) = \|A\mathbf{x} - A\mathbf{y}\|$$

$$= (A\mathbf{x} - A\mathbf{y})^T (A\mathbf{x} - A\mathbf{y})$$

$$= (\mathbf{x} - \mathbf{y})^T A^T A (\mathbf{x} - \mathbf{y})$$

$$= (\mathbf{x} - \mathbf{y})^T (\mathbf{x} - \mathbf{y}) = d(\mathbf{x}, \mathbf{y})$$

∴ 欧氏距离具有正交旋转不变性.

证毕.

五、综合分析应用题

主要范围：数字及各种图案的识别及综合应用(如：飞机、艇船等军事或民用的目标识别，计算机车牌识别与自动收费管理与分类计算机控制的自动化流水线中的邮政编码识别与分类，手写字符识别等)，人脸识别及其综合应用(如人脸识别门禁、人脸识别考勤、人脸识别在安全领域中的应用)，语音识别应用(如手机语音自动拨号、语音交互与机器翻译等)，模型识别的其他综合应用(如文本分类、垃圾邮件识别、网页分类等，相关内容可参考教材第11章)等。

说明：若开展语音模式识别或图像模式识别研究，还需要分别有针对性地进行语音处理和图像处理等相关基础课程的学习。

综合分析应用题举例：

在某选举投票过程中需要采用计算机对选票中的($\sqrt{\quad}$, \times , \bigcirc)三种“图像图案”进行自动识别分类统计，“ $\sqrt{\quad}$ ”表示投票通过，“ \times ”表示投票未通过，“ \bigcirc ”表示弃权（假设所有3种“图像图案”已经从所采集的图像选票中分割出来,但各“图像图案”大小不一）。试运用你所学的模式识别方法设计自动分类统计“ $\sqrt{\quad}$ ”-投票通过的人数、“ \times ”-未通过的人数和“ \bigcirc ”-弃权的人数的技术方案。要求：

1. 给出总体设计方案；
2. 给出对“图像图案”进行预处理、特征提取与选择以及识别分类统计等关键设计步骤。

答案不唯一，关键点回答正确即可得分，此略。

关于模式识别编程(本届不考):

可使用自己熟悉的任何一种计算机语言编程(研究和开发计算机模式识别软件对数学能力要求较高, 本课程的一大特色是采用第4代计算机语言MATLAB软件工具讲授模式识别编程技术, 能加深对模式识别基本概念、原理和方法的理解和掌握, 激发对模式识别问题的学习和研究兴趣)。编程是算法的具体表现, 它有助于帮忙学习者加深对模式识别算法的理解, 复习过程中可有选择地看课件中的模式识别问题MATLAB编程, 包括编程思考题 (往届考过编程题, 本届考试不考编程题)。

编程题举例(*:本届不考，仅供学习参考)

对一幅森林起火图像，希望把着火部分划分出来，可以采用以下法：在该图像中分别在着火部分与森林部分画出一个窗口，把在这两个窗口中的像素数据作为训练集，用线性判别方法求得分类器参数，再用该分类器对整幅图进行分类。试用你熟悉的编程语言针对你所选的线性判别方法编写分类器程序。

解答：若采用Fisher分类器，可参考Fisher分类器程序(lecture3_2.m)，详见模式识别课件(三).PDF，此略。

推荐复习方法：

建议将九个PDF课件要求的部分打印出来，至少复习两遍，重点练习本复习课中的计算分类题型。

因为本课程是开卷考试，请大家考试时务必携带所打印的课件和教材。