总结：

1 老的两份试卷遗传占一半，新的占1题；考察内容逐渐从超纲变成场景题，意味着哪怕我们只学前100页，仍然可以出很难，而且随时间推移越来越灵活，作为第一题经常是整张卷最难的存在。

2 毛kezhi所负责的ML出题非常非常呆板，完完全全的背多分，因此完全有可能为了区分度第一题只能出更难。

**SEMESTER 2 2021-2022 EE6227**

Q1 遗传算法

场景题，4小问，个性化定制防护头盔，根据头盔参数构建目标函数以供优化算法最小化，比较实数编码和二进制编码的优缺点。

（范围外：实数编码）

Q2 遗传算法  
6小问

1 自己的语言解释NP问题和TSP问题

2 计算10个点的距离矩阵

3 用排列编码验证TSP的两个解（范围外：排列编码）

4 基于最长循环应用循环交叉操作，判断后代的适应度是否得到提高。（范围外：排列编码）

5 定义连通性适应度，表达连通性适应度方程，计算连通性适应度（范围外：排列编码）

6 自己构造适应度函数（范围外：排列编码）

Q3 第三章 贝叶斯分类 第四章LDA 第七章分类器评估

3小问

给出两个类别均值向量和协方差矩阵

1 设计一个贝叶斯决策规则，并使用该规则对以下样本进行分类

2 设计一个Fisher线性判别分类器，并使用分类器对以下样本进行分类

3 讨论可用于评估模式分类器的过程和指标

Q4 第九章 特征选取和提取 第十章 聚类

3小问

1 画出过滤（filter）和包装（wrapper）2种监督特征选择算法的框图，并讨论它们的主要区别以及各自的优缺点。

2 列出三种聚类方法，并为每种方法给出一个示例算法。讨论每种聚类方法的应用场景。

3 描述DBSCAN聚类算法的步骤。

**SEMESTER 2 2022-2023 EE6227**

Q1 遗传算法

场景题，7小问，假设每人拥有固定数量心跳，用完生命结束，如何优化？基于给出的心率公式

1 在表格中填入目标函数的值 自由选择20个合适x值，计算出近似的最佳锻炼时间。 给出每日最佳锻炼时间的分钟数

2 写出一个适当的适应度函数(x)，以将适应度值从0到1归一化

3 从优化的角度，用你自己的语言解释求x的最优值的难度。

对问题类型的分类（优化，model，仿真）

4567

Q2遗传算法

4小问，给定一个复杂的需要优化的多维函数

1 给几个x值计算y值

2 应用算术重组公式为每次交叉操作生成两个子代。

3 染色体可以表示为二进制字符串，写出最优两个染色体的18位二进制字符串。

4 改参数，求出X\*的18位二进制字符串。

Q3 第九章 特征选取和提取 第四章LDA 第七章 分类器评估

1 对于一个模式分类问题的数据集，建议一种特征选择方法来解决该问题，并描述所建议方法的过程

2 为该二类分类问题设计一个Fisher线性判别分类器。

3 简要描述k-折交叉验证方法，并讨论其原理。

Q4 第十章 聚类

1 使用凝聚层次聚类算法对样本进行聚类，并绘制树状图

2 选择一种方法解决该聚类问题，解释选择该方法的理由，并描述所选方法的过程

**SEMESTER 1 2023-2024 EE6227**

Q1 遗传算法

场景题，题干超长的2小问

12城市闭环旅行商问题，引入新方法“超重组”并介绍（全新的完整步骤）

1 确定图1中两位父代经过超重组后选择的两个子代，展示你的全部步骤和计算

2 简要讨论这种重组方法的优缺点。

Q2 第二章 数据预处理（最放水的一题）

1 在机器学习问题中，我们可能遇到哪些数据类型？用示例解释

2 讨论我们如何处理异常值。

3 讨论我们如何处理缺失值。

4 讨论我们如何处理具有不同范围、测量单位或数量级的特征。

Q3 第四章 LDA 第三章 贝叶斯分类

两类样本，给均值向量和散布矩阵

1 使用训练数据集设计一个线性判别分类器，并预测以下测试样本的类别标签

2 使用训练数据集设计一个朴素贝叶斯分类器，说明任何假设，并预测以下测试样本的类别标签

Q4 第十章聚类 第十一章聚类指标

1 为图2中显示的数据选择一个合适的聚类算法，解释选择理由，并描述所选聚类算法的过程。

2 为图3中显示的数据选择一个合适的聚类算法，解释选择理由，并描述所选聚类算法的过程。

3 列出聚类的5个主要评估指标，并简要解释每个指标。

**SEMESTER 2 2023-2024 EE6227**

Q1 遗传算法

场景题，2大题5小问，

给出比例-积分-微分（PID）控制器的公式，详细描述使用遗传算法调整PID控制器的实现：

1 用实数编码表示

2 二进制表示

15种家庭物品，三个现金券，游戏规则

1 填表，如何达到最省钱

2 描述并展示一个遗传算法的基因型表示方案，用于优化现金券的使用。使用你在第(i)部分中的答案作为表示的示例。

3 简要描述如何构建适应度函数。

Q2 第五章 SVM 第四章LDA

1 假设训练样本是线性可分的，构建线性SVM的原始问题，并证明你的答案。

2 已知对偶问题，给支持向量，给拉格朗日乘子，给标签，确定线性SVM分类器的公式，并使用该分类器对以下测试样本进行分类

3 讨论Fisher线性判别分析（LDA）与线性支持向量机（SVM）之间的主要区别。

Q3（Quiz2） 第七章分类器评估 第九章 特征选取和提取

1 根据给出的混淆矩阵，计算分类器的准确率、精确率、召回率和F1分数。

2 机器学习中的峰值现象（Peaking Phenomenon）是什么意思？讨论峰值现象如何指导训练数据的收集和特征选择。

3 列出并简要描述三种特征子集选择方法，并讨论它们的优缺点。

Q4 第十章 聚类

1 列出并简要描述四种聚类方法，并讨论其各自的应用场景。

2 使用凝聚层次聚类算法对以下数据进行聚类，并绘制树状图。