

实验作业题（4 道，对应 后 4 次实验）

一、作业 1（对应实验 4：线性回归算法）—— 工业设备能耗预测

题目名称

基于线性回归的车间电机能耗预测

任务描述

利用线性回归模型，根据电机的运行参数（转速、工作温度）预测其能耗，掌握工业场景下线性回归的预处理与预测流程。

数据说明

- 训练数据：10 组电机样本，特征为「转速（r/min）、工作温度（℃）」，标签为「能耗（kW·h）」，数据如下：

特征 $x = \text{np.array}([[1000, 35], [1200, 38], [1100, 36], [1300, 40], [900, 32], [1400, 42], [1050, 34], [1250, 39], [1150, 37], [800, 30]])$

标签 $y = \text{np.array}([8.2, 9.5, 8.8, 10.1, 7.5, 10.8, 8.4, 9.8, 9.1, 7.2])$

- 待预测样本：2 组新电机参数 $xs = \text{np.array}([[1180, 37], [1500, 45]])$ （注：电机工作温度超过 43℃时，修正为 43℃，模拟工业设备温度上限）。

具体要求

- 对 xs 进行预处理：将温度 $> 43^\circ\text{C}$ 的样本修正为 43°C ；
- 用训练数据构建并训练线性回归模型；
- 预测预处理后的 xs 能耗，打印“原始参数 - 预处理后参数 - 预测能耗”；
- 验证第 1 个待预测样本的预测值是否在 $8.5 \sim 9.5 \text{ kW} \cdot \text{h}$ 范围内（合理区间）。

二、作业 2（对应实验 6：决策树算法）—— 工业设备故障类型判断

题目名称

基于决策树的机床故障类型判断

任务描述

利用决策树模型，根据机床的运行参数（振动值、噪音值）判断其故障类型，实现工业设备的交互式故障诊断。

数据说明

- 问卷（特征）：2 个工业参数问题 —— 「振动值（ $0 \leq 0.1\text{mm}$ ， $1=0.1\sim 0.3\text{mm}$ ， $2=>0.3\text{mm}$ ）、噪音值（ $0 \leq 60\text{dB}$ ， $1=60\sim 80\text{dB}$ ， $2=>80\text{dB}$ ）」；
- 训练样本：8 组数据 `answers = np.array([[0,0],[0,1],[1,0],[1,1],[2,0],[2,1],[2,2],[1,2]])`，对应故障标签 `labels = ['正常','轻微异响','轻微振动','轻微故障','严重振动','中度故障','严重故障','中度异响']`；

具体要求

1. 用 `answers` 和 `labels` 训练决策树模型；
2. 实现交互式输入：让用户依次输入 2 个参数（0/1/2），用 `try-except` 拦截非整数输入；
3. 输出用户输入的参数对应的故障类型；
4. 验证：输入「振动值 1、噪音值 0」时，预测结果是否为“轻微振动”。

三、作业 3（对应实验 7：KMeans 聚类算法）—— 工业传感器数据分群

题目名称

基于 KMeans 的车间传感器数据聚类分群

任务描述

利用 KMeans 算法，对车间 12 个传感器的监测数据（温度、压力）进行聚类，划分设备运行状态群（正常 / 关注 / 异常），掌握工业数据的聚类预处理与结果分析。

数据说明

- 传感器数据：12 组样本 `data = np.array([[25, 0.8], [26, 0.9], [24, 0.7], [30, 1.2], [32, 1.3], [29, 1.1], [35, 1.5], [36, 1.6], [34, 1.4], [27, 0.85], [28, 0.95], [31, 1.25]])`（特征：温度 $^{\circ}\text{C}$ 、压力 MPa）；

具体要求

1. 将 `data` 转换为 KMeans 要求的格式（无需 `reshape`，直接使用）；
2. 构建 KMeans 模型，设置 `n_clusters=3`（对应 3 种运行状态）；
3. 打印聚类中心数量和每个样本的聚类标签；
4. 分析结果：判断聚类中心温度最高的群是否为“异常状态群”。

四、作业 4（对应实验 8：Matplotlib 多图形绘制）—— 工业车间生产数据可视化

题目名称

车间月度生产数据多图形可视化

任务描述

利用 Matplotlib 绘制车间 3 条生产线的月度数据（产量、合格率），掌握工业数据的多类型图形展示方法。

数据说明

- 基础数据：月份 `month = ['1 月','2 月','3 月','4 月']`，生产线 A/B/C 的产量（件）`output = [500, 520, 490, 530]`、`output_b = [480, 490, 500, 510]`、`output_c = [520, 510, 530, 540]`；合格率（%）`qualify = [98.2, 98.5, 98.0, 98.8]`。

具体要求

- 配置中文显示（避免乱码）；
- 绘制 2×2 子图：
 - 子图 1（1 行 1 列）：生产线 A 的月度产量折线图（蓝色，圆形标记）；
 - 子图 2（1 行 2 列）：3 条生产线 4 月产量柱状图（颜色分别为绿 / 橙 / 紫）；
 - 子图 3（2 行 1 列）：生产线 A 的月度合格率散点图（红色，点大小 100）；
 - 子图 4（2 行 2 列）：3 条生产线 4 月产量占比饼图（显示 1 位小数百分比）；
- 保存图片为“车间生产可视化.png”，确保子图无重叠（用 `tight_layout()`）。