

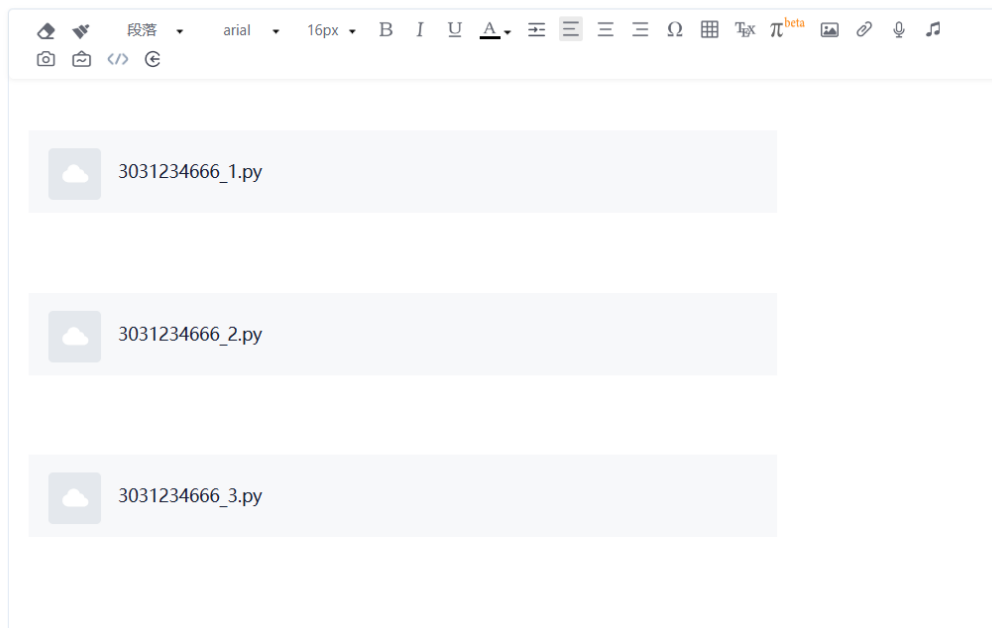
期末测试

总体要求：

- (1) 代码要适配任意大小数据集，例如在提取第二题的文件时，不要出现类似[:1001]这种提取前 1000 行的切片。
- (2) 提交时将三个 python 文件上传到学习通，不要打包。
- (3) 按照每题的要求命名和设置根目录，不然无法在教师端跑通。

提交示例：

1. (简答题) 测试



测试一：现有一眼底镜数据集 CHASE_DB1，包含左眼、右眼的眼底镜检查结果以及相应的分割掩膜，计算所有眼底镜检查结果分别在 R、G、B 三个通道上的均值和方差

注：

- (1) 不区分左右眼
- (2) 函数名为 statistics，返回值为两个列表，第一个列表为所有眼底镜检查结果依次在 R、G、B 通道的全局均值，第二个列表为所有眼底镜检查结果依次在 R、G、B 通道的全局方差（按 R、G、B 顺序），函数名和主函数已经在 test1.py 中给出，不要更改。

要求：

- (1) 将“./期末测试”设置为工作区根目录。
- (2) 将代码文件命名为“学号_1.py”提交，只用提交代码文件。（示例 3031234666_1.py）

测试二：现有一个磁悬浮电机旋在运动过程中转偏移量预测的数据集（包括历史数据-滤网.csv，历史数据-滤网.csv，历史数据-温度.csv，历史数据-运行.csv），目的是通过“历史数据-滤网”、“历史数据-温度”、“历史数据-运行”三个部分，预测“历史数据-轴心”的值，构建一个线性回归模型的机器学习方法解决问题。

注：

(1) 对“历史数据-滤网.csv”、“历史数据-温度.csv”、“历史数据-运行.csv”三个文件取样时，注意时间间隔与“历史数据-轴心.csv”文件不同。

(2) 读取文件如果报错：

UnicodeDecodeError: 'gbk' codec can't decode byte 0xb4 in position 8: illegal multibyte sequence

将 open 的参数 encoding 设置为 encoding="utf-8"。

要求：

(1) 使用前 80%的数据作为训练集，其他 20%的数据作为测试集合。

(2) 将代码封装为一个函数。函数名为 machine_learning，函数返回值为测试集上的 mse，函数名和主函数已经在 test2.py 给出，不要更改。

(3) 将“./期末测试”设置为工作区根目录。

(4) 将代码文件命名为“学号_2.py”提交，只用提交代码文件。（示例 3031234666_2.py）

测试三：现有一患者 CT 检查结果数据集 imgs，请你根据 imgs 的存储结构，构建一个嵌套的 json 字典结构，以保存患者、区域和图像的组织信息，并最终写入文件。

注：

(1) 最终输出的结果参考“example.json”，json 文件的结构组织形式参考“参考组织形式.txt”（每个字段的含义已在此文件中给出）。

(2) 准备写入的数据和将数据写入 json 文件的代码已经在主函数中给出，不要更改。

要求：

(1) 必须封装三个函数，三个函数的命名和 return 的组织形式、主函数已经在 test3.py 给出，主函数已经给出 json 的嵌套字典结构，不要更改。

(2) 将“./期末测试”设置为工作区根目录。

(3) 将代码中的 json 文件名部分改为“学号_3.json”。

(4) 将代码文件命名为“学号_3.py”提交，只用提交代码文件。（示例 3031234666_3.py）