# 大作业介绍

1. 数据集RE\_PV\_panel\_datasets是本次作业1中要用到的太阳能光伏板沾灰数据集，其中将其分为三类：
2. 没有灰尘



1. 有些灰尘



1. 布满灰尘



根据提供的图片数据集设计卷积神经网络：

1、将数据集分为训练集、验证集和测试集

2、训练模型，提供至少3个不同模型参数的组合

3、对模型进行评价，包括准确率、精确率、召回率、F1分数、ROC曲线

4、用PyQT/gradio设计UI界面，可进行一定的交互功能，包括但不限于超参数设置、模型参数设置、训练功能

5、评价指标显示、数据集加载/选择等

1. 数据集情感分类数据集是本次大作业二所用到的数据集：

1、根据提供的情感分类数据集设计自然语言处理分析模型

2、其余要求同一。

三、用大语言模型微调训练一个跟自己专业相关的智能问答小助手

1、根据自己的专业选择一本教材

2、选择一个开源的大模型，比如 DeepSeek、qwen、Llama、GLM等

3、构建问答数据集，并分为训练集、验证集和测试集，至少500条

4、进行指令微调训练模型，至少比较2个不同的基座模型

5、通过测试集对模型进行模型性能评价

6、用PyQT/gradio设计UI界面，可进行一定的交互功能，包括但不限于超参数设置、模型参数设置、训练功能、

7、评价指标显示、数据集加载/选择等