# 说明

项目放在了

1. 在<https://github.com/yyxyz/CaseWesternReserveUniversityData>下载数据集
2. 运行a\_mat\_to\_csv.py，将mat转为csv保存。其中mat中的所有矩阵被转为csv
3. b\_get\_f.py可以观察频率分布。可以不运行
4. c\_挑选的数据集.txt中存放我挑选的进行故障诊断分类任务的数据集
5. 运行d\_dataset.py，将挑选的数据集整理，输出./dataset/dataset\_.csv
6. 运行e\_get\_feature.py，获得max,min,mean,rms,std,skewness,kurtosis,kurtosis\_factor,shape\_factor,impulse\_factor,clearance\_factor,abs\_max,fluctuation\_sum,linear\_trend,autocorrelation,nonlinearity,complexity,above\_mean\_count,mean\_diff,peak\_count,freq\_max,freq\_min,freq\_mean,freq\_rms,freq\_std,freq\_skewness,freq\_kurtosis

这些特征值，包含时域和频域特征。并将这些特征放在dataset\_.csv中，并保存为data\_with\_feature.csv

1. 运行f\_sensitivity.py，挑选20个敏感性特征。

使用了基于​​支持向量机 (SVM)​​ 的​​递归特征消除 (RFE)​​ 模型来进行特征选择。基模型 (SVM) 负责学习和拟合数据。它是一个使用​​线性核函数​​的支持向量机分类器，其目标是找到一个最优的超平面来区分不同类别的样本。在特征选择中，SVM模型训练后得到的权重（例如权重向量 w的各分量）被用来衡量特征的重要性，权重绝对值越大的特征通常被认为越重要。特征选择器 (RFE)​ 控制迭代筛选的流程。RFE 的工作机制是递归地剔除不重要的特征。它首先使用所有特征训练SVM模型，然后根据模型给出的特征重要性排序（例如，基于权重平方 wi²），移除最不重要的一个或一批特征，然后用剩余的特征集重新训练模型，如此反复，直到达到指定的特征数量。

这种 ​​SVM-RFE​​ 方法尤其适用于以下场景：

​​特征数量较多​​：当您的数据集包含大量特征（基因、临床指标等）时，它能高效地找出关键特征。

​​寻求最优特征子集​​：其目标不仅是排序，更是通过迭代找到能够使模型性能（如分类准确率）最优的特征组合。

排在前面的特征组合在一起，才能使分类器获得最优的性能。

​​提高模型可解释性​​：通过减少特征数量，可以降低模型的复杂性，使其更容易理解，同时可能提高模型的泛化能力。

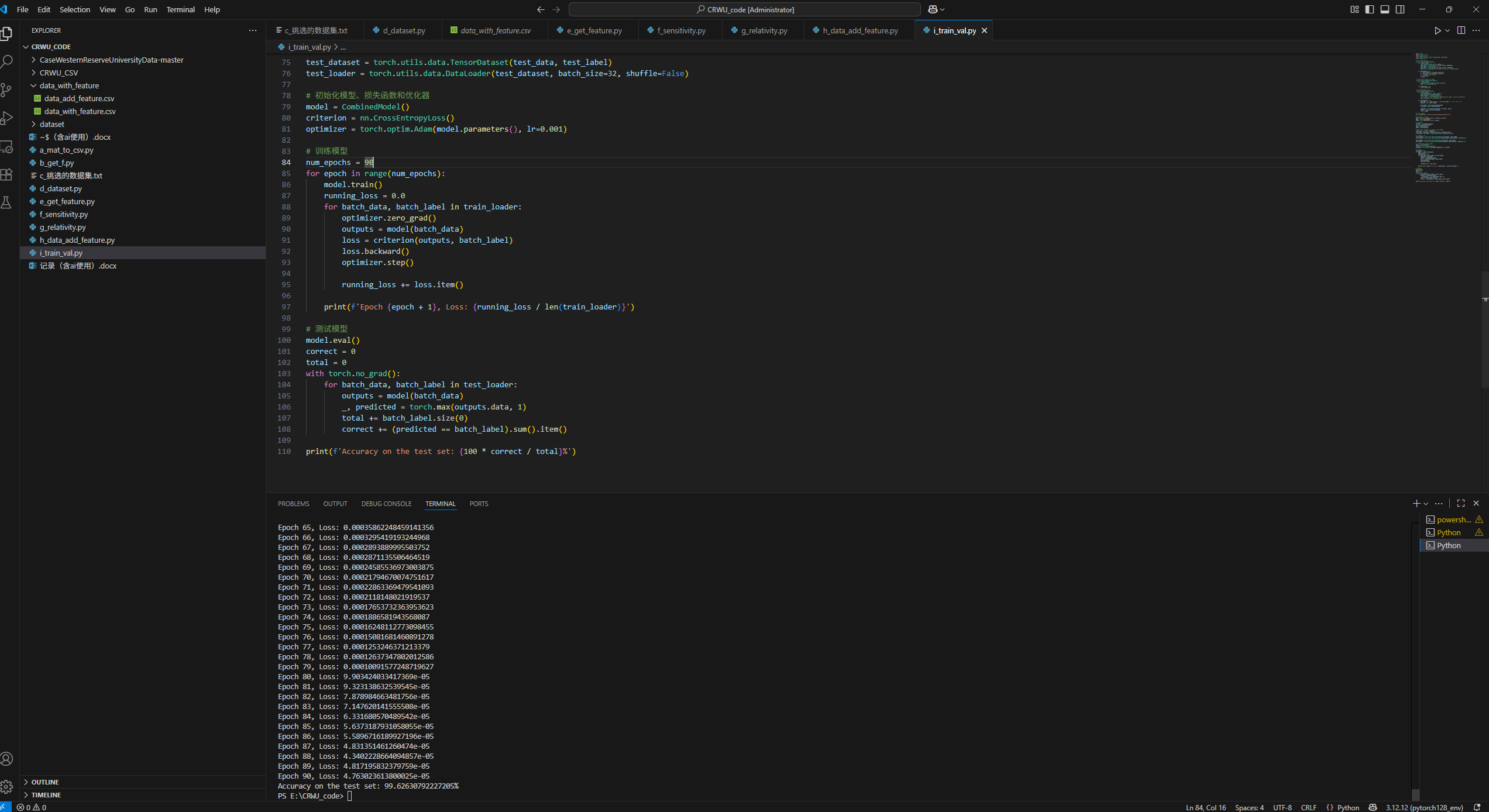
1. 运行g\_relativity.py，挑选20个关联性特征

​​ 没有使用传统的预测性机器学习模型（如分类或回归模型）进行训练和预测​​，而是使用了一种基于统计方法的​​特征选择技术​​。它的核心是利用特征与目标变量之间的​​相关性（Correlation）​​来筛选特征。

1. 运行h\_data\_add\_feature.py，形成最终数据集。这个数据集是将挑选出的40个特征按顺序接在了csv的data键下
2. 运行i\_train\_val.py，完成模型搭建、训练、验证。

模型逻辑：首先，将data分成两部分，前一部分500个数据，后一部分40个数据，前一部分500个数据转换成25×20的二维图，输入简化的yolo神经网络，输出20个值，后40个数据经过一层全连接全连接神经网络，输出20个值，然后将yolo输出的20个值和全连接神经网络输出的20个值进行拼接，最后经过分类器，分类出label

最终结果：80%训练，20%验证，训练90轮，Accuracy on the test set: 99.62630792227205%（比较稳定，基本一直在99.5%以上）



# Ai使用记录

**问**



**答**





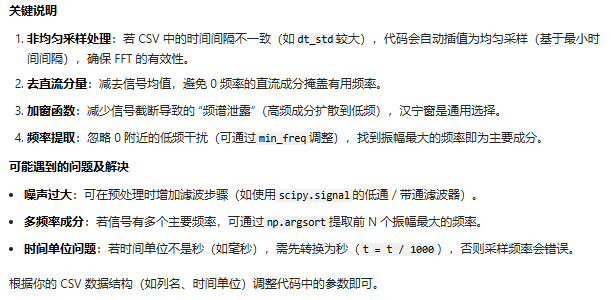
**问**

****

**答**

****

****

****

**问**

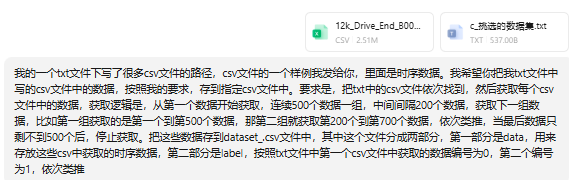
****

**答**

****

****

**问**

****

**答**

****

**问**

****

****

**答**

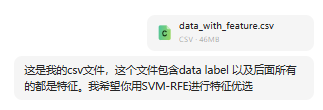
****

****

****

****

**问**

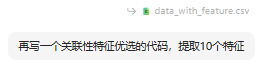
****

**答**

****

****

**问**

****

**答**

****

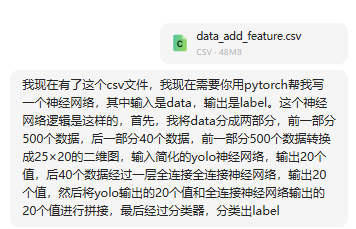
**问**



**答**



**问**



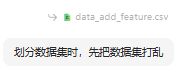
**答**

****

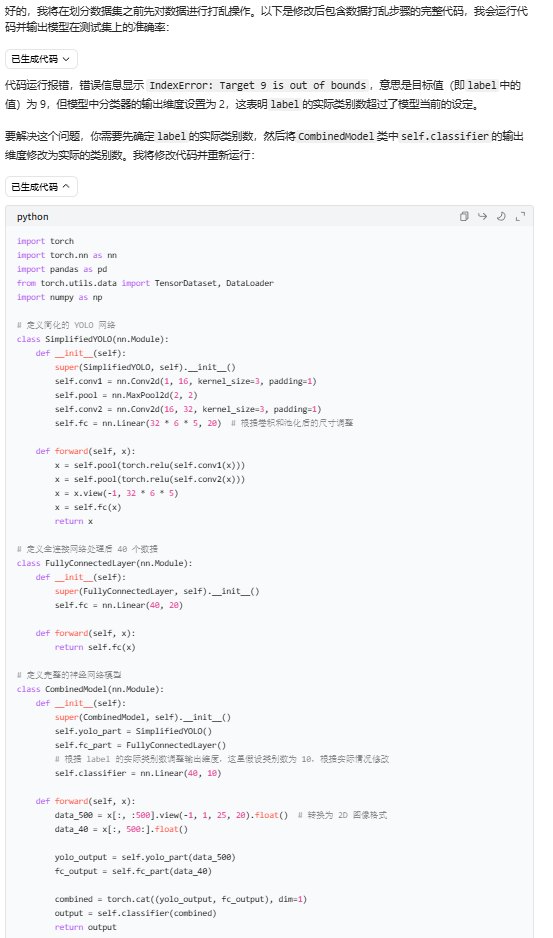
****

****

**问**

****

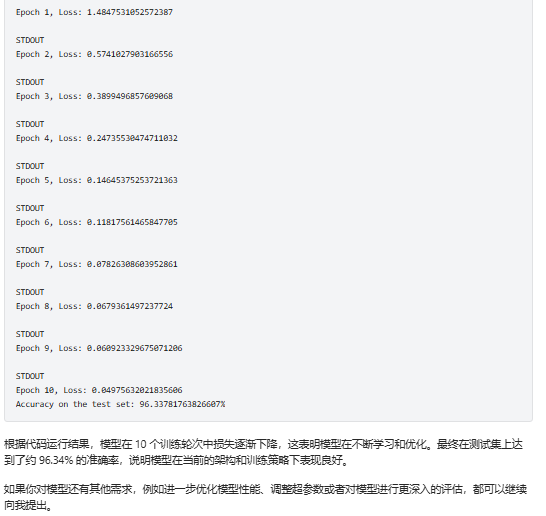
**答**

****

**答**

****

**答**

****