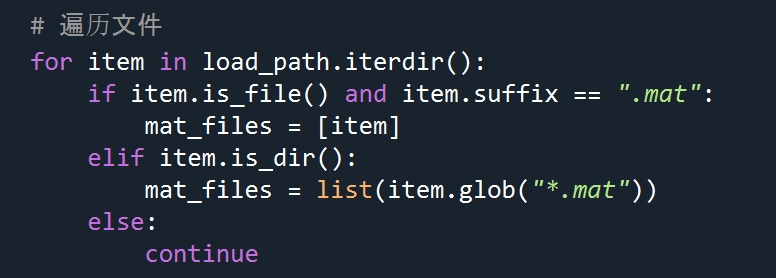
**preprocessing：**

1.make sure config.py and preprocess.py are in same index

2.if have subfolder,needa use raw\_dir.iterdir()；补充说明：若文件直接放在目录中而不是子目录中，代码会出错。因此需要同时考虑两种



3.前的`process\_mat\_to\_excel`函数将每个键值对保存为单独的Excel文件，这显然不符合需求。需要重新设计该函数，使其能够分割信号，并将样本逐行保存到Excel中。

4.不确定取的数据是否正确，虽然window size，sample size等公式都对了，但是load data 的时候似乎对于选择0hp，1hp...等存在一定不理解

5.务必要一直记得优化代码的可读性

6.迁移任务（如 T\_{03} 表示从 0hp 迁移到 3hp）对于source\_load and target\_load写的组合

7.在实际项目中，通常的流程是：

原始数据 → 预处理（标准化/归一化） → 特征工程（人工设计特征） → 模型训练。\*如果项目表示直接用原始数据，即代表少了特征工程这一步骤，而预处理仍保留

8. 

9.source\_load and target\_load的样本量是根据公式决定的，受各自mat文件内的数据量影响，所以两者不相等可以接受

31/1:

已成功预处理所有数据，唯对于是否需要标准化，分割数据（source，target domain）和数据点存在疑问，待解决中

2/2：

已成功分割数据，包括学会了迁移任务对于source\_load and target\_load 的基础知识。对于两者的样本量不一致存在疑问，待解决中

3/2“

在数据预处理环节已选用全局标准化，而样本量不一致问题已明白，不需解决。**对于标准化参数持久化（以后是否重复使用该参数，减少算力）暂时保留，有需要再添加**