**内容概述：**

1. RL+预测器接口修改为批处理
2. 拓扑神经网络实验
   1. 完成了整体项目架构
   2. 重写了cgcnn
   3. 确定了超图神经网络
3. 甲烷存储预测器训练（资源不足暂时没有启动）

**具体内容：**

1. **RL+预测器接口修改为批处理**

**任务来源：**RL修改了算法，现在一次性生成了一批数据，而不再是一个COF，由于之前写的接口没有考虑到这一点，所以需要修改一下。

**任务详情（留个记录）：**整个任务其实并不难，就是修改为批处理就可以了。

原先的接口设计的是前端传入一个COF的文件路径，然后后端根据这个路径去改找到对应的COF，创建一个临时的文件夹来存储本次预测任务的信息。现在传入这一批COF所在的目录，（要求：这个目录有且仅有这一批COF）。

1. **拓扑神经网络实验**

**背景：**考虑到多孔材料的复杂性，一般的图结构不能够描述材料的分层结构以及各组分之间的多元关系。所以需要使用更加一般的数据结构，这些数据结构包括——单纯复形、胞腔复形、超图和组合复形，尤其考虑组合复形。有了新的数据结构，需要设计新的深度学习模型从这些更加一般的数据结构上提取特征。类似于利用基于消息传递的图神经网络从图数据上提取特征，我们利用基于消息传递的拓扑神经网络从这种更加一般的数据结构（单纯复形、胞腔复形、超图和组合复形）上提取特征。提取好特征后再接一个MLP用于预测属性。

**实验驱动：**单纯复形、…、组合复形是图的推广，为了最终实现应用于多孔材料的拓扑神经网络，这里从目前应用于多孔材料的图神经网络——cgcnn开始，一步步将图数据推广到图数据。

**整体的路线**：

——cgcnn（重写cgcnn）

——超图神经网络（将多孔材料转换为超图，并使用超图神经网络提取特征，这一步会考虑到多孔材料中原子间的多体作用，相较于图神经网络会考虑更多信息）

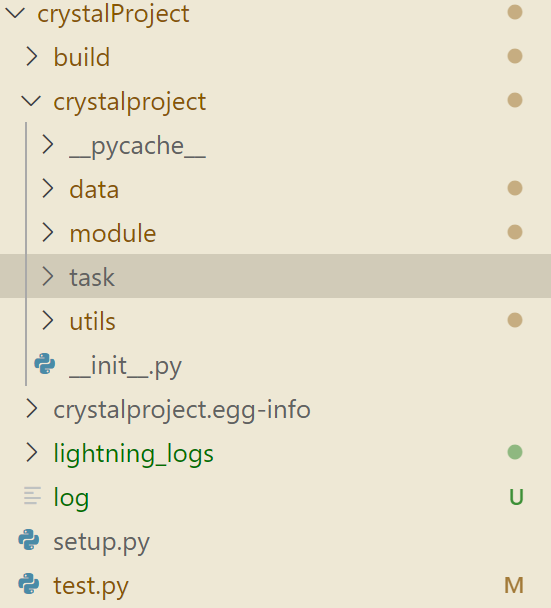
——构造多孔材料的一般化拓扑数据（使用组合复形这个更加一般的数据结构）（从晶体学拓扑的角度出发构造出多孔材料的拓扑数据，到这一步，不仅要考虑到多孔材料中的多元关系，还考虑多孔材料的层次结构，这些信息对于描述COF非常重要）

——利用组合复形神经网络（具体的模型是需要设计的，作者仅仅提供了一个设计组合复形神经网络的框架）

**实验内容：**

——整体的项目架构**（这里是为了应对即将到来的可复用性和可拓展性需要）**，由于整个实验流程中会设计多种模型——图神经网络、超图神经网络、组合复形神经网络，并且也可能涉及到多种不同的消息传递函数——基本的聚合、注意力等；数据集的构建存在多种——图结构、超图结构、组合复形等（可能还会考虑到对比其他不是图结构的数据结构——正弦库伦矩阵、描述符等）；

整个项目的框架如下：

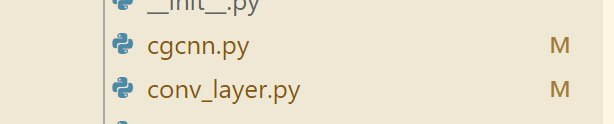


其中crystalProject是整个项目，crystalproject是整个项目的代码，data是数据集、数据处理的相关模块以及一些LightningDataModule模块，module是模型、功能模型以及一些LightningModule模块。

基于这种架构，当需要设计一个新的模型的时候只需要写一个模型的模块就可以了，其他部分可以复用。同样，设计一个新的数据结构只需要写一个数据集的模块。

——cgcnn重写

这个实验是为了使用更加一般的框架去重写cgcnn，这样有利于对于不同模型之间的不同点，便于改进。

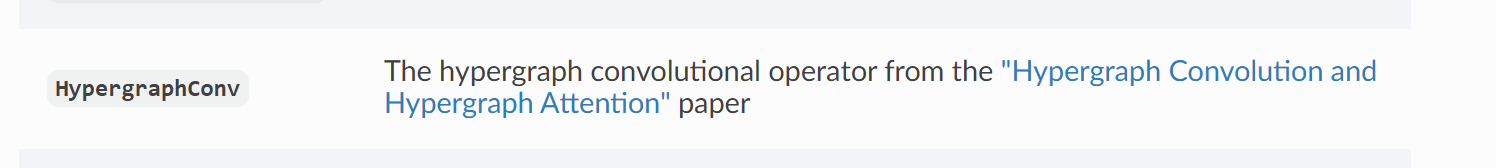


需要注意的是，cgcnn模型当初的代码并不是基于目前常用的pyg中所提出的图结构实现的，所以重写cgcnn还需要重写它的数据集模块，并且整个流程已经按照pyg重新实现了。

——超图神经网络

这个实验包含两个部分——设计多孔材料的超图数据、设计超图神经网络，其中有一个难点，分子的超图数据比较容易设计，但是多孔材料是一种晶体，需要考虑周期性边界条件，目前晶体的超图构建方法还没有被研究。

第二部份是设计超图神经网络，这里我查询了pyg的官方文档，发现已经提供了一个实现，这里我准备直接用。



1. 甲烷存储预测器训练

这是使用MOFT预测属性的方案，由于需要的资源较多，暂时实验室资源不足没有进行实验。

下周安排：

1. 用于COF的超图神经网络的实现
   1. 设计COF的超图数据（从分子的超图数据构造出发，加入周期性边界条件）
   2. 超图神经网络，使用pyg现成的代码
2. 氧气吸附预测器的第二次训练
   1. 参考周报9-10提到的训练框架，下周浩宇的实验结束后应该会产生一批结构，准备用这些结构再次微调一下模型