**内容概述：**

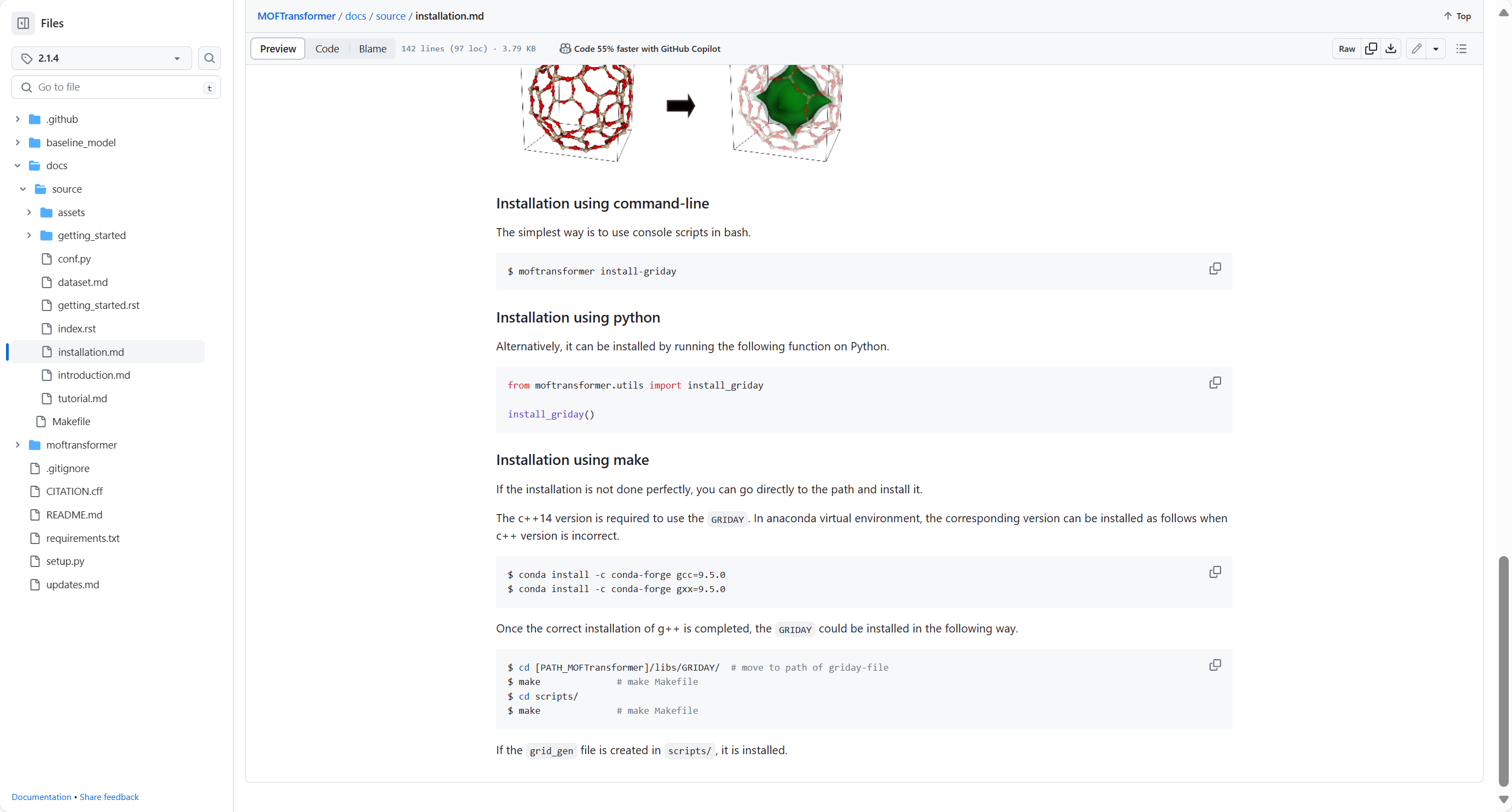
1. 解决了MOFTransformer服务器环境配置问题；
2. 利用甲烷存储COFs数据库微调MOFTransformer预训练模型；
3. 设计了属性预测的完整流程；
4. 写好了GCMC计算氧气吸附的脚本

**具体内容：**

1. **解决了MOFTransformer环境配置问题**

**问题描述：**搭建MOFTransformer的计算环境时，由于程序中某个功能需要调用一个C可执行程序grid.gen，然后我们的服务器C编译器的版本和该可执行程序需要的不同，导致动态链接库加载不了，程序运行失败。

**问题解决方案：**



通过查看github库里面的doc目录，其中有提到如何安装grid计算程序，已经如何编译，作者提供了三种方法：

1. 通过python可执行程序moftransformer进行安装
2. 通过python脚本进行安装
3. 通过cmake进行安装

我尝试了前面两种方法，当使用第二个方法的时候成功安装了grid。

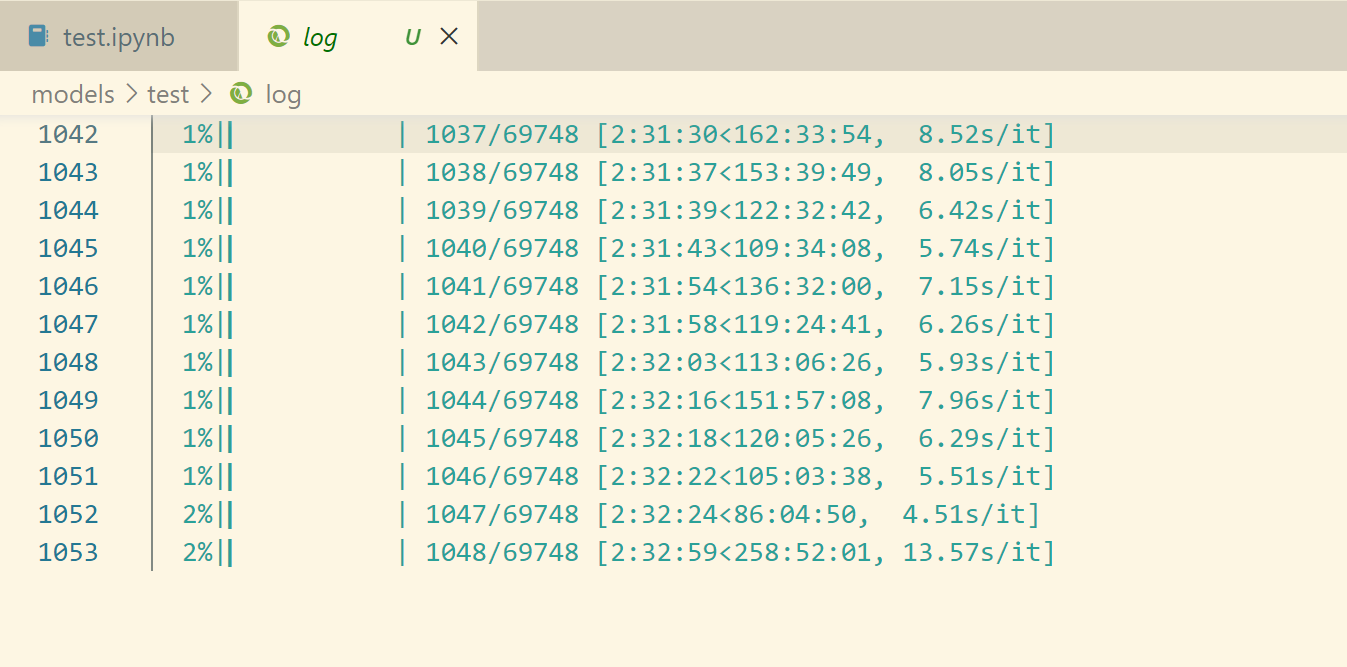
1. **利用甲烷存储的COFs数据微调MOFTransformer模型**

**任务内容：**

调试好moftransformer环境后，就可以使用它提供的函数进行微调了，微调任务包含两个步骤，第一步：将原始的cif文件转换为moftransformer需要的格式；第二步，调用run()函数进行微调。

**第一步：**

Moftransformer程序中提供了一个prepare()函数，只需要设置好待处理的原始文件的目录格式，再调用prepare()函数就可以实现计算。



从日志文件中可以看出，计算单个COFs大约需要花费5-10s，甲烷存储的COFs一共包含7万个，全部计算完毕大概需要150个小时，大约下周可以计算完毕。

附：第一次计算在这周三就已经开始进行了，但是由于中途程序因未知原因突然结束，周日才发现问题，所以又开始了重新计算。

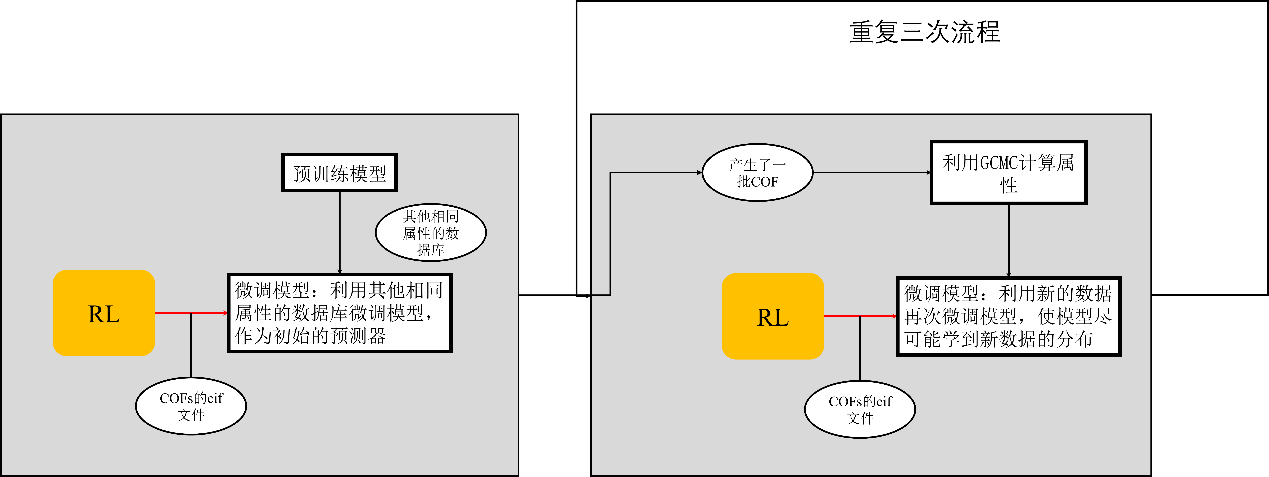
**第二步：**

Moftransformer提供了run()函数用于进行模型的微调，根据第一步计算好的数据，可以直接进行计算，由于第一步还没有完成，目前没有测试第二步可能需要的时间。

**结果：第一版模型cgcnn在计算COFs甲烷吸附属性上存在较大误差，moftransformer则在论文中提到能够实现较小的误差，所以该任务完成后应该会得到一个预测效果较好的第二版模型。**

1. **设计了属性预测完整流程**

**流程图：**



**设计目标：**

由于缺少RL生成的COFs的样例，从别的地方找来的数据库，大多和RL生成的COF类型不一致，这可能导致分布不同，从而预测效果不理想。为了解决这个问题，首先需要一个能力较强的模型（即在大多数数据库上都取得了不错的效果，并且迁移能力强），这里采用moftransformer，利用这个模型在其他数据库上进行训练，得到一个初始的预测器（目的是防止RL在第一轮的探索过程中由于预测器效果太差而导致的错误），这个预测器能够预测目标属性，但是对新的数据（RL产生的数据）预测的精度较差；RL第一轮探索完成后，会产生一批COFs数据，对这批数据运行GCMC计算，得到的结果用于再次训练模型；重复几次之后，模型在预测RL生成的COF属性时误差应该会更小。

1. **写好了GCMC计算氧气吸附的脚本**

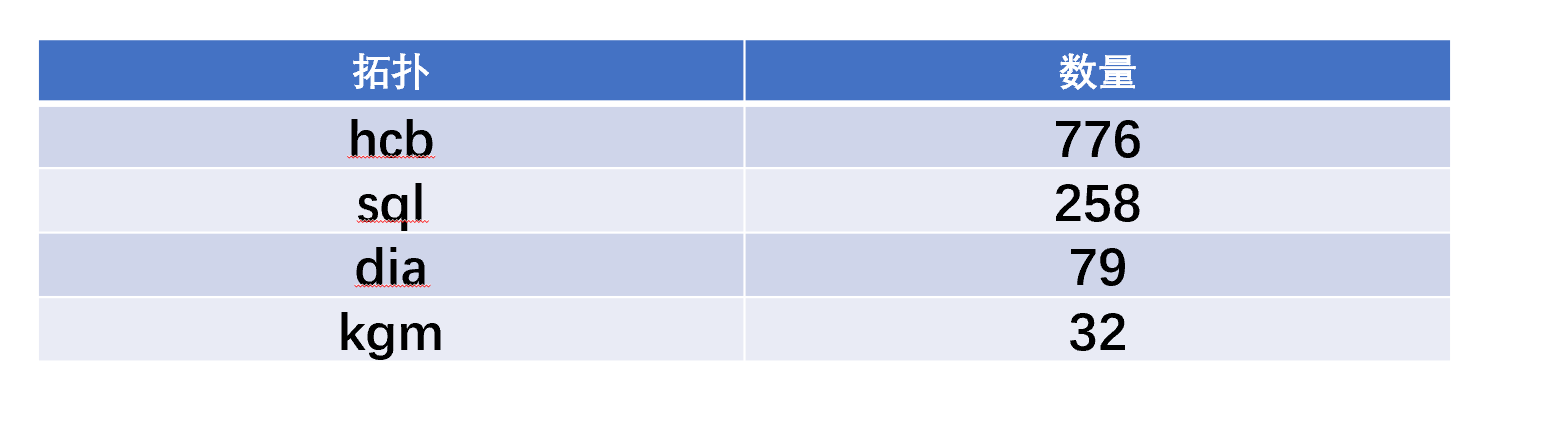
**任务描述：**

RL那边的目标属性时氧气吸附量，但是目前不存在COFs氧气吸附的数据，导致3中流程图中的第一步初始的预测器没法得到，为了解决这个问题，需要自己动手计算一些氧气吸附数据。

**具体方案：**

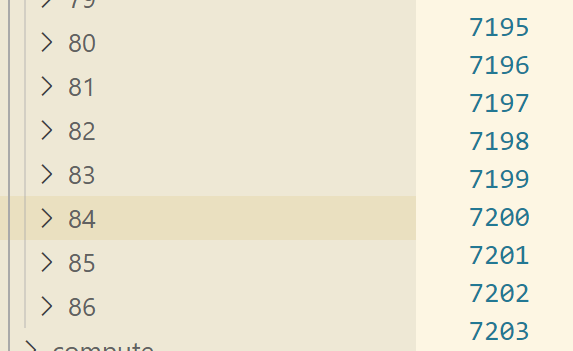
CoreCOF中包含了1242个实验合成的COFs结构，并且提供了cif文件，所以我将初始了氧气吸附训练数据目标放在了这个数据集上。

首先统计了一下CoreCOF中每种拓扑包含的结构数量：



发现hcb拓扑数量很多，而且目前RL产生的COF也是hcb拓扑，所以利用这些hcb拓扑的结构进行GCMC计算

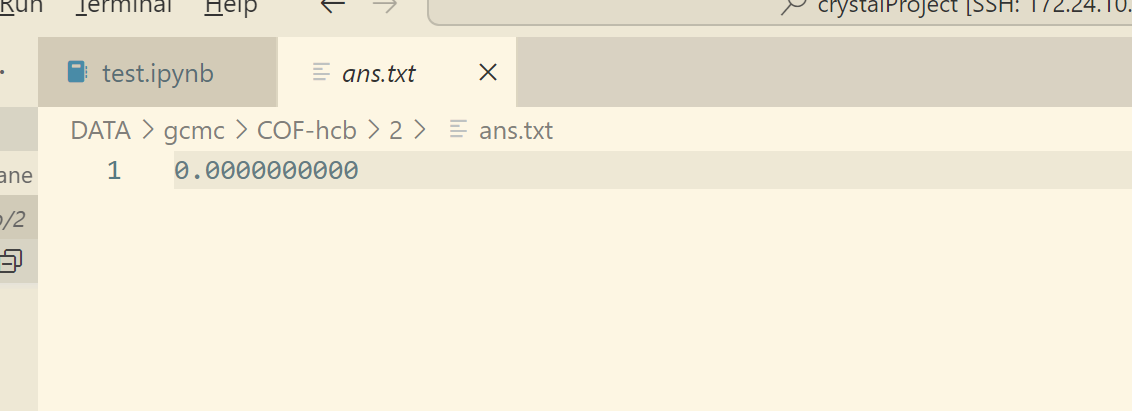
为了执行GCMC计算，首先需要需要安装raspa软件，直接通过github利用pip安装就可以了，然后就是具体的计算脚本，这个工作我和浩宇已经完成了，他写了单个结构的计算脚本，我在他的基础上完成了批处理脚本。目前程序已经开始运行。



目前已经计算了86个结构，大约20-30min可以计算一个结构，全部计算完成需要两周。

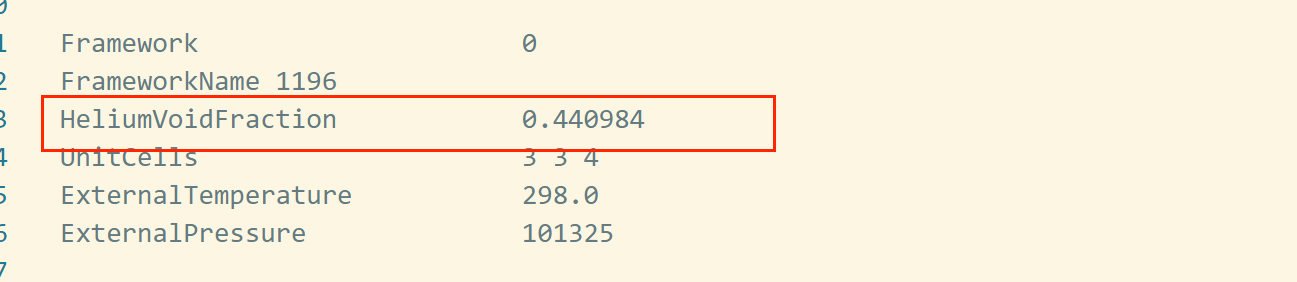
**附：由于服务器cpu数量较少，目前只能开16核计算，每次只计算一个结构，如果核数增加的话，速度可以提升至少四倍。**

**存在问题：**



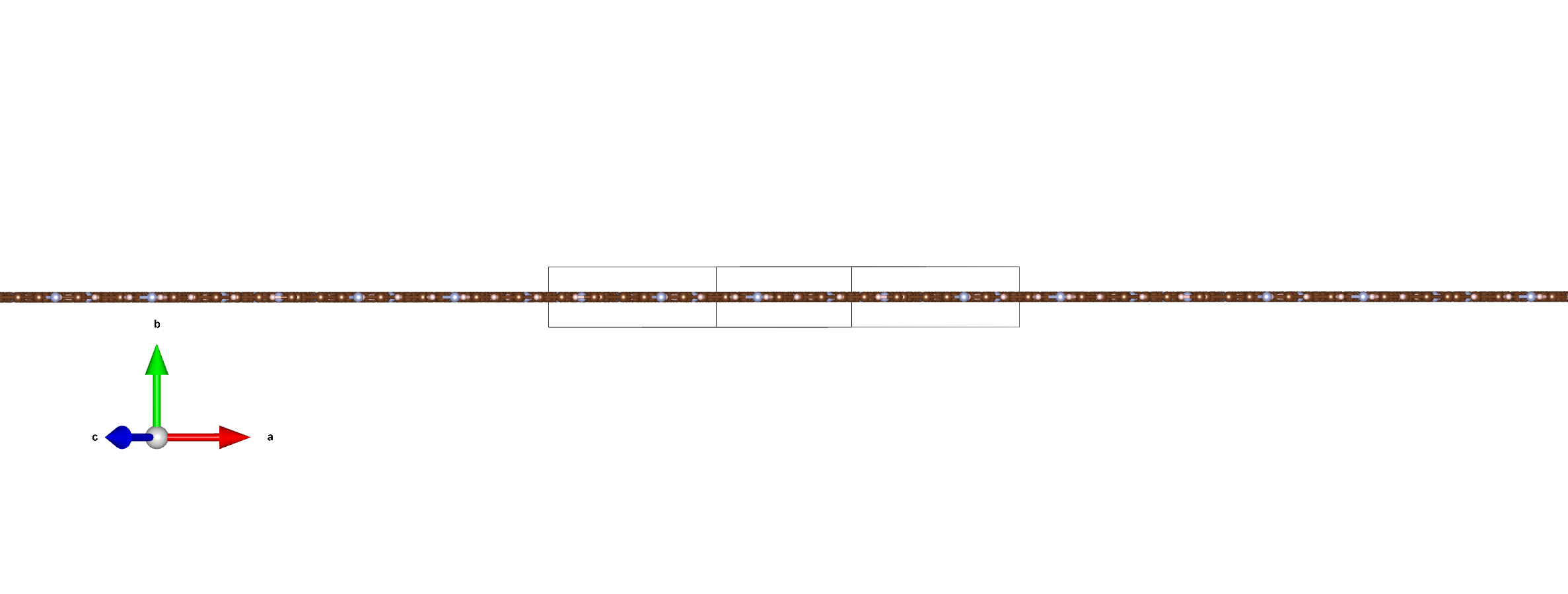
发现其中有部分COF计算完成后吸附量的结果是0。

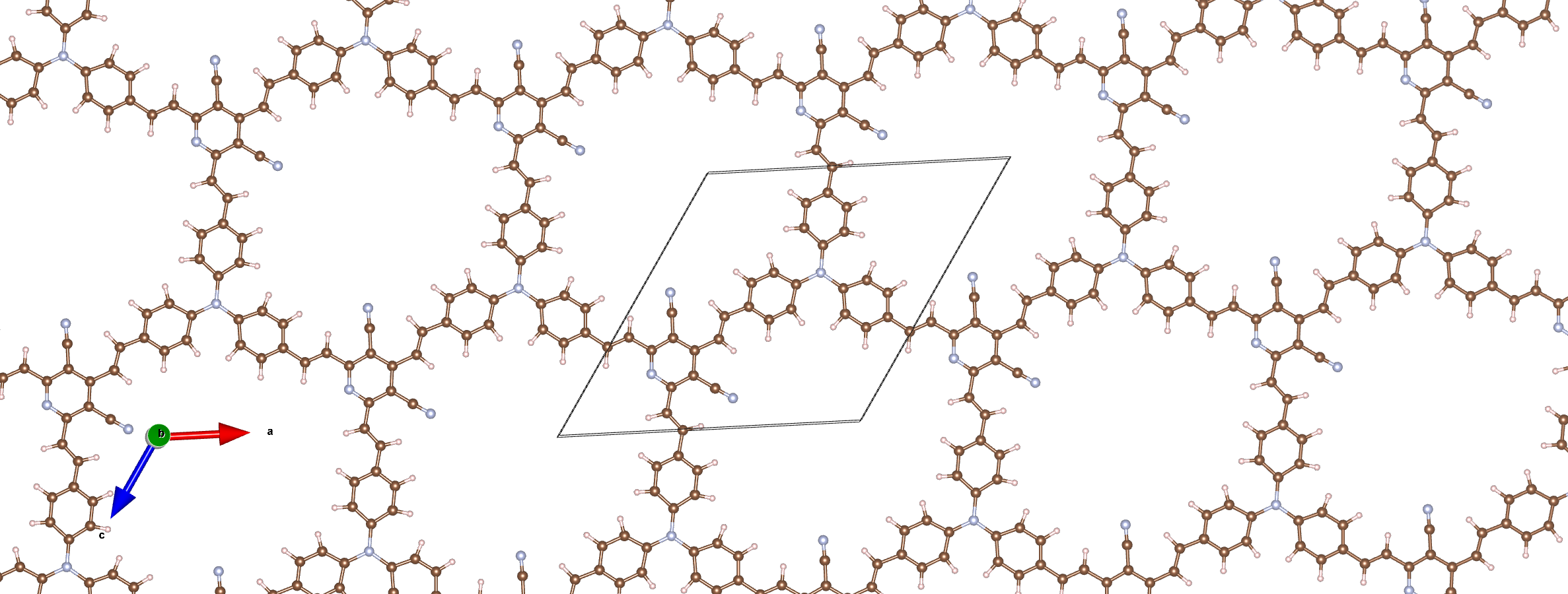
首先，我和吴优师兄认为是第一步计算出现了问题（氧气吸附一共包含两步计算，第一步计算孔隙率，第二步才是计算氧气吸附）。



但是第一步的孔隙率计算是正常的。

然后，我查看了一下出问题的COFs结构，发现这些COFs的cif文件对b轴和c轴式反过来的





可能是这个原因导致了计算出现了问题。

**下周安排：**

1. **利用甲烷存储的COFs数据微调MOFTransformer模型**

**下周数据预处理应该可以完成，之后就可以进行第二步的微调了。**

1. **Core-COF氧气吸附属性计算**
2. **解决部分COF氧气吸附属性值为0的问题**