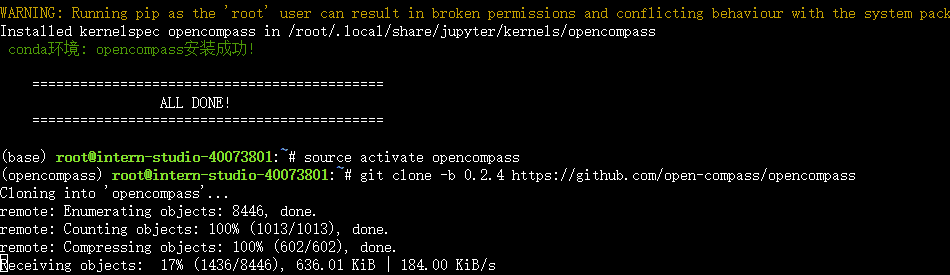
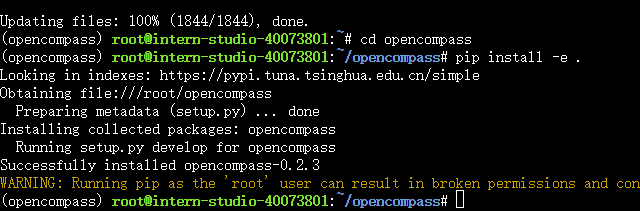
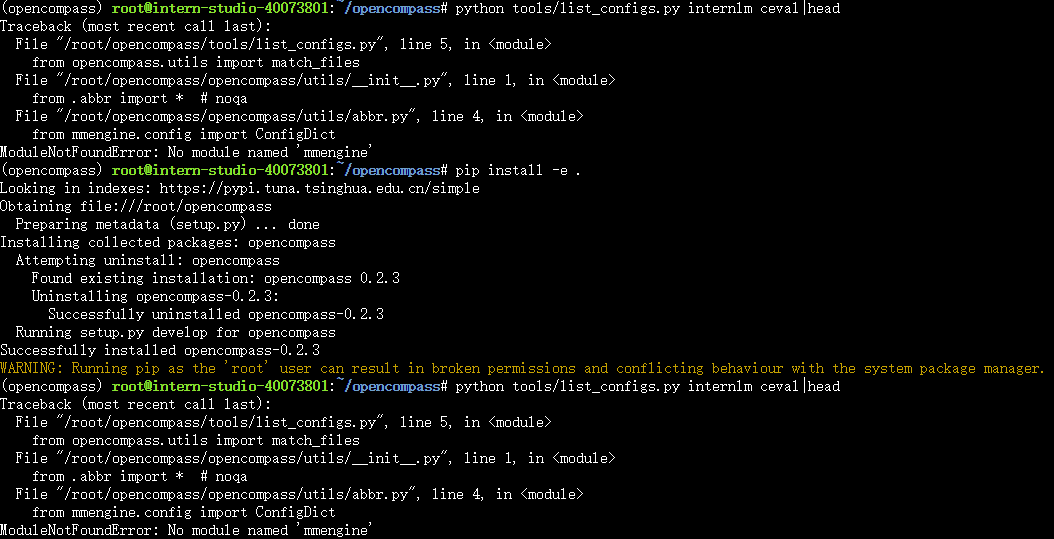
第七节作业

使用 OpenCompass 评测 internlm2-chat-1\_8b 模型在 C-Eval 数据集上的性能

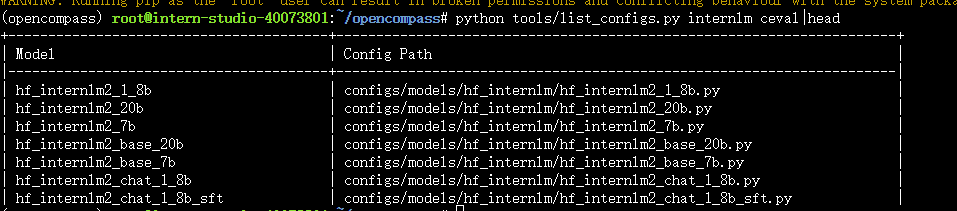
构建conda镜像和OpenCompass安装







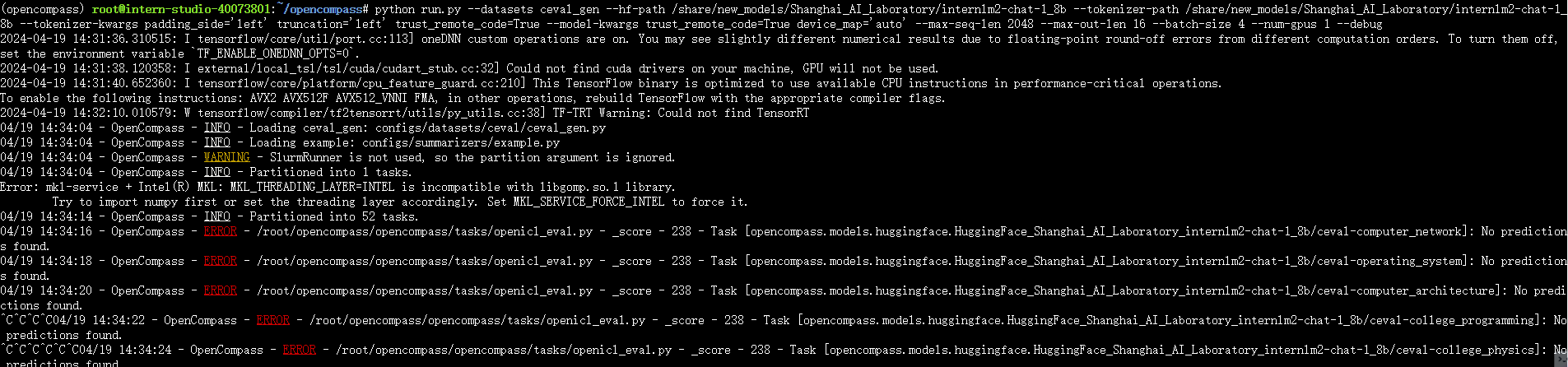
安装后可以正常看到数据集和模型信息



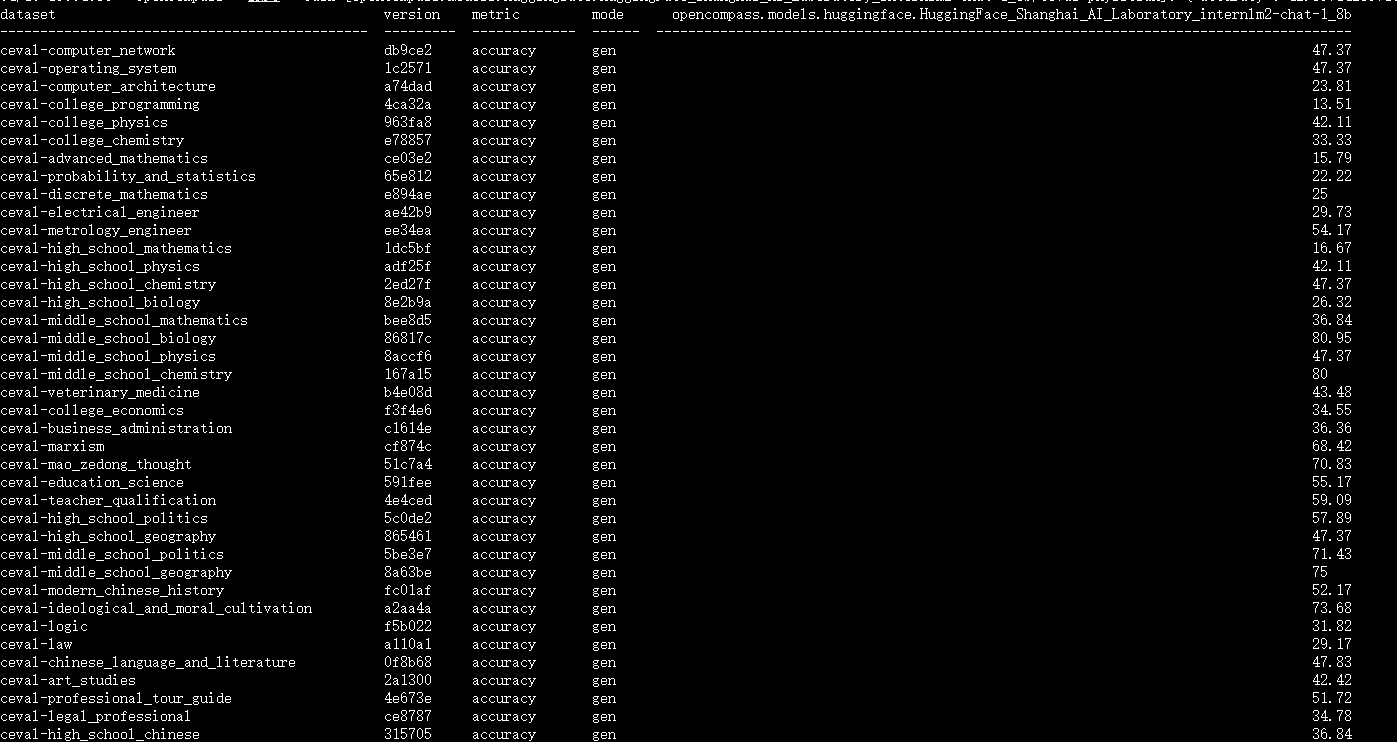
1.3 启动测评



和第六节一样，有一个线程冲突的bug



需要先声明export MKL\_SERVICE\_FORCE\_INTEL=1，之后再启动可以得到正常测评结果：



将自定义数据集提交至OpenCompass官网

