第六次实验

58119125

蒋卓洋

1. 配置与调试：
   1. 问题解决：
      1. 虚拟机扩容失败后打不开了：重装虚拟机重新配环境
      2. 配置过程中忘记换环境，导致后续操作错乱：删除anaconda相关文件，从头开始配置。
   2. 样例测试结果：

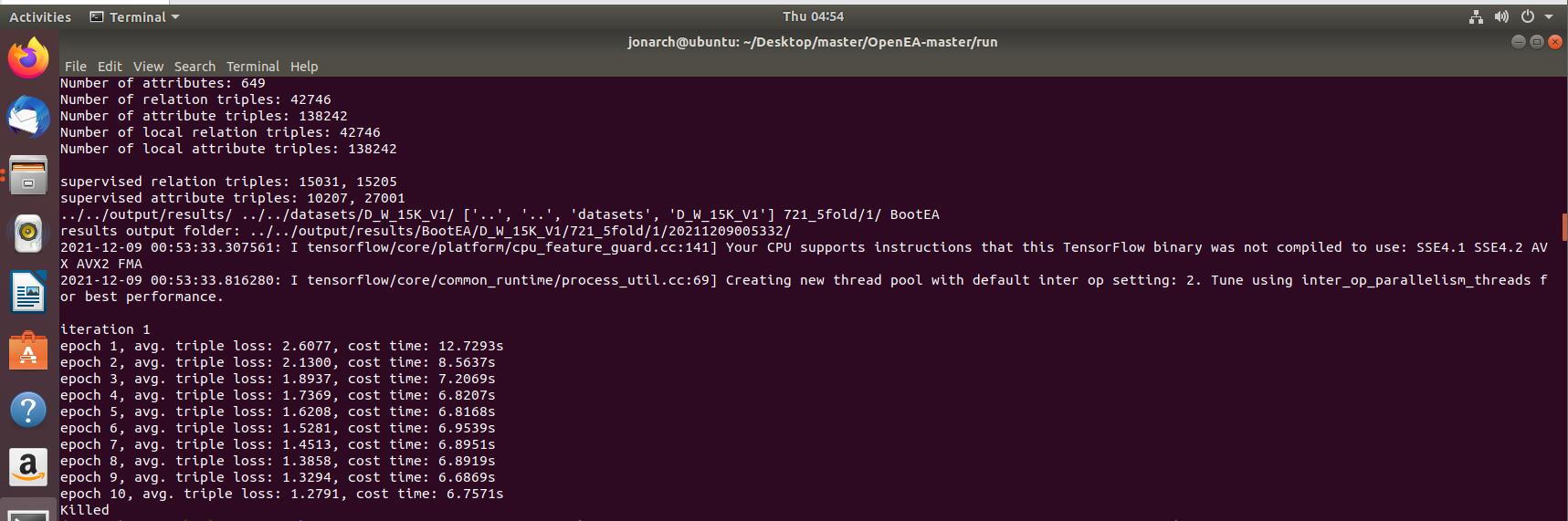


图1.样例测试

1. 分别使用EN\_FR\_15K\_V2的split1和EN\_DE\_15K\_V2的split2来运行MTransE，记录使用的命令和结果
   1. EN\_FR\_15K\_V2的split1运行MTransE：
      1. 需求分析：
         1. 测试代码（python源文件）：main\_from\_args.py
         2. 模型参数：./args/mtranse\_args\_15K.json

（embedding模式为MtransE，数据集规模为15K）

* + - 1. 训练集：EN\_FR\_15K\_V2
      2. 训练集具体文件夹：1
    1. 命令：

1-3

图2.命令1

* + 1. 结果：

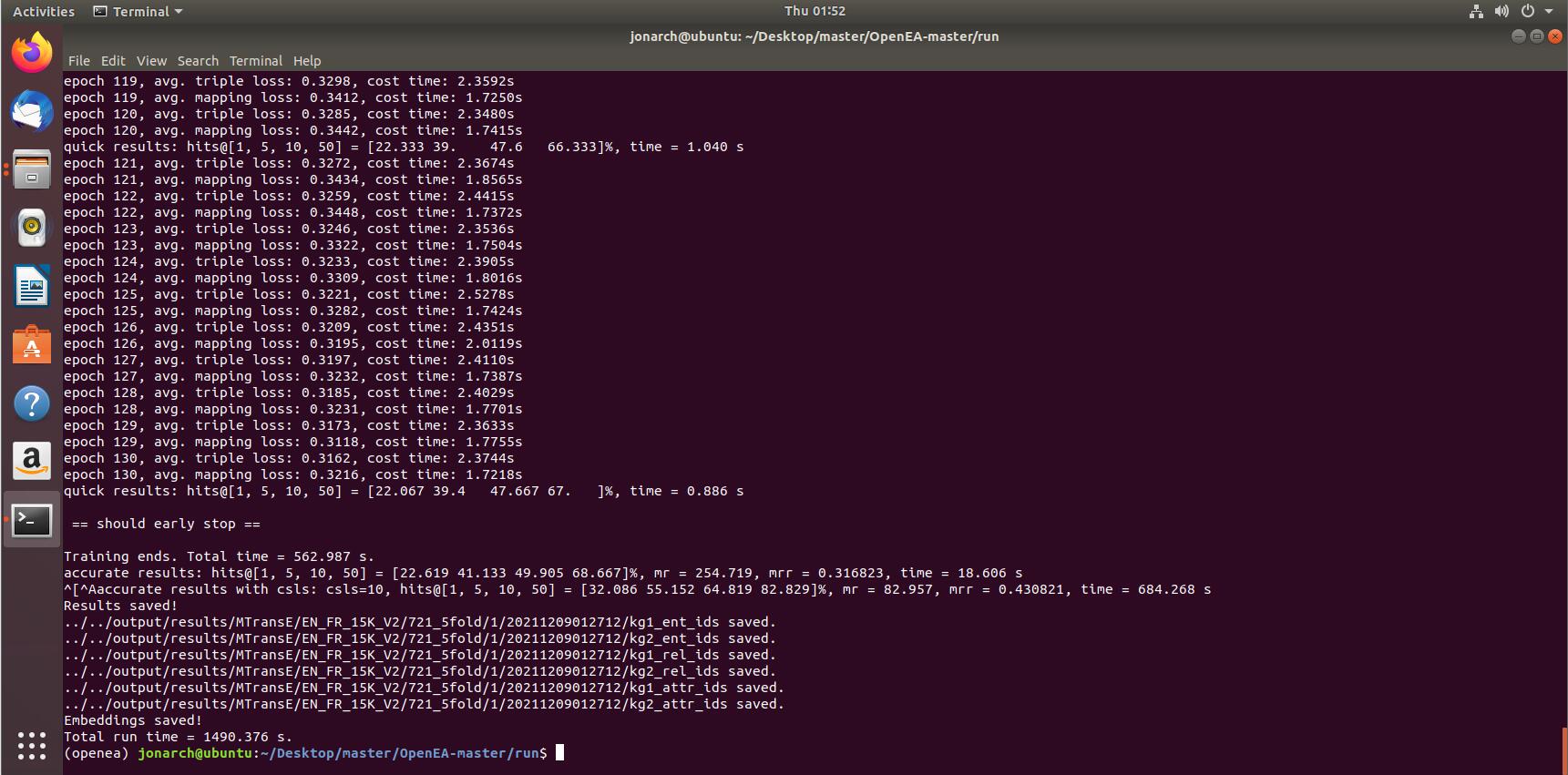


图3.运行结果1

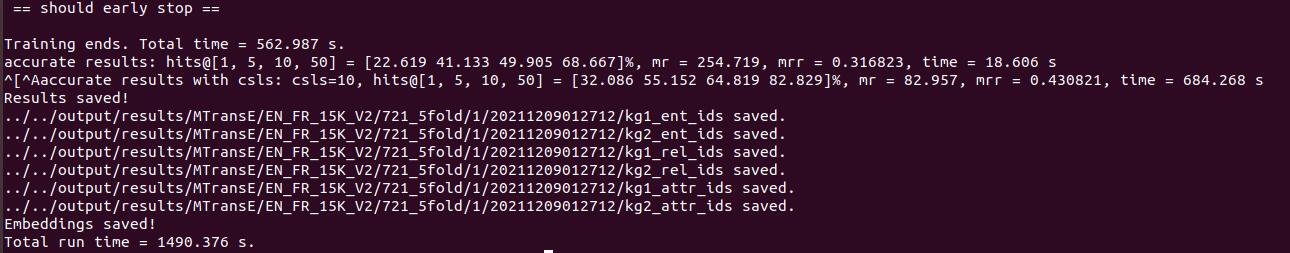


图4.运行效果1

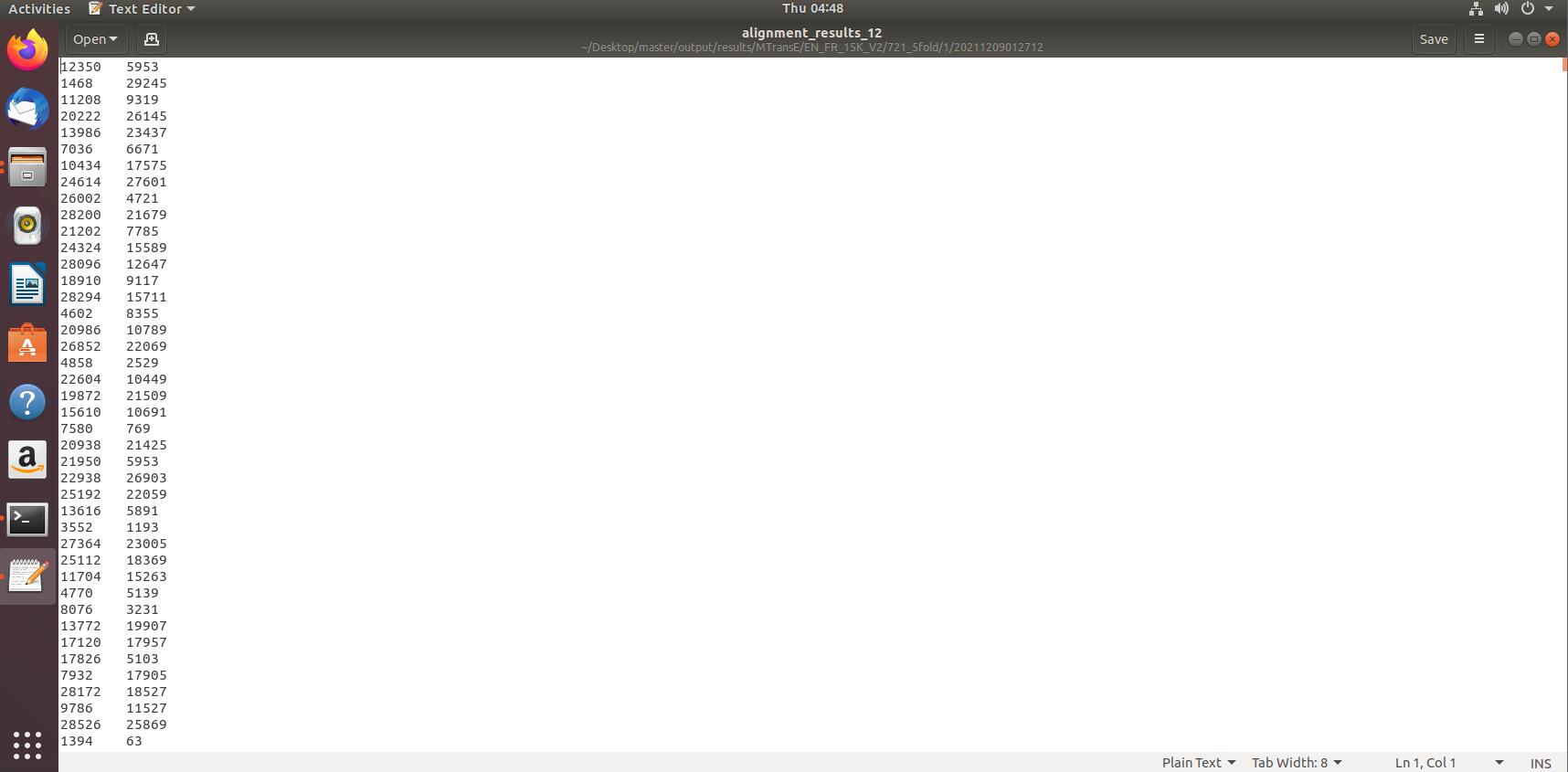


图5.对齐结果1

* 1. EN\_DE\_15K\_V2的split2运行MTransE：
     1. 需求分析：
        1. 测试代码（python源文件）：main\_from\_args.py
        2. 模型参数：./args/mtranse\_args\_15K.json

（embedding模式为MtransE，数据集规模为15K）

* + - 1. 训练集：EN\_DE\_15K\_V2
      2. 训练集具体文件夹：2
    1. 命令：

2-3

图6.命令2

* + 1. 结果：

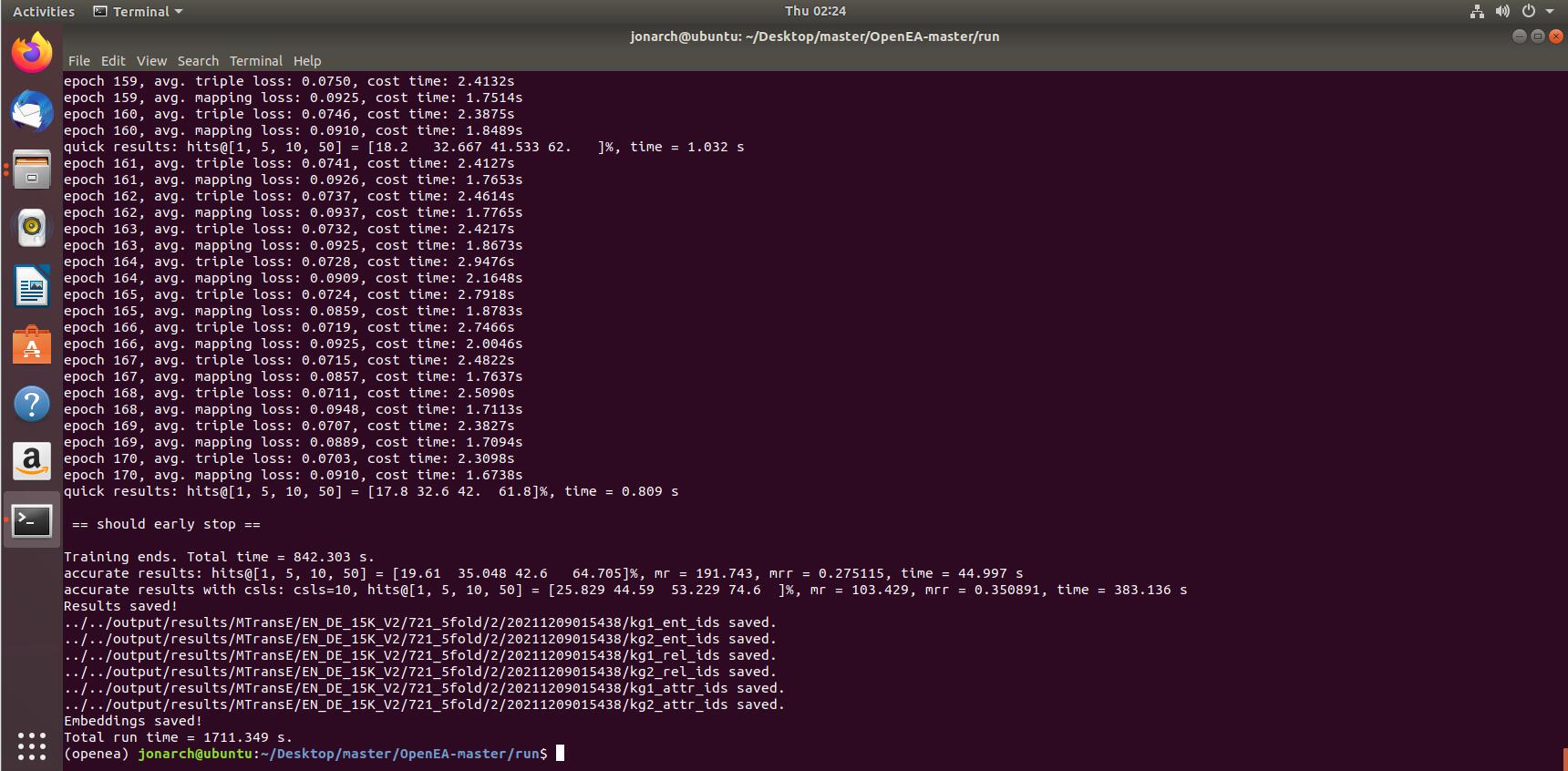


图7.运行结果2

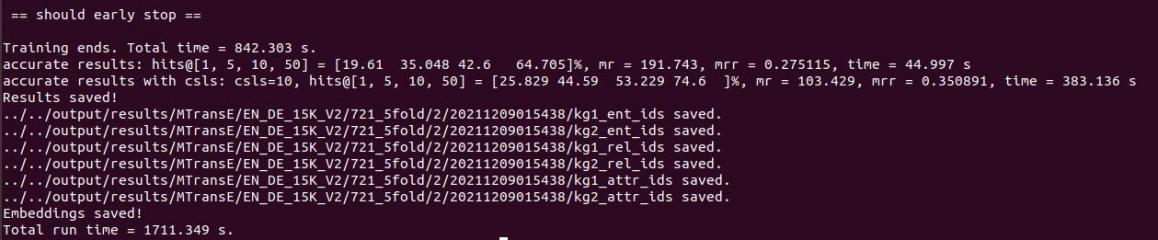


图8.运行效果2

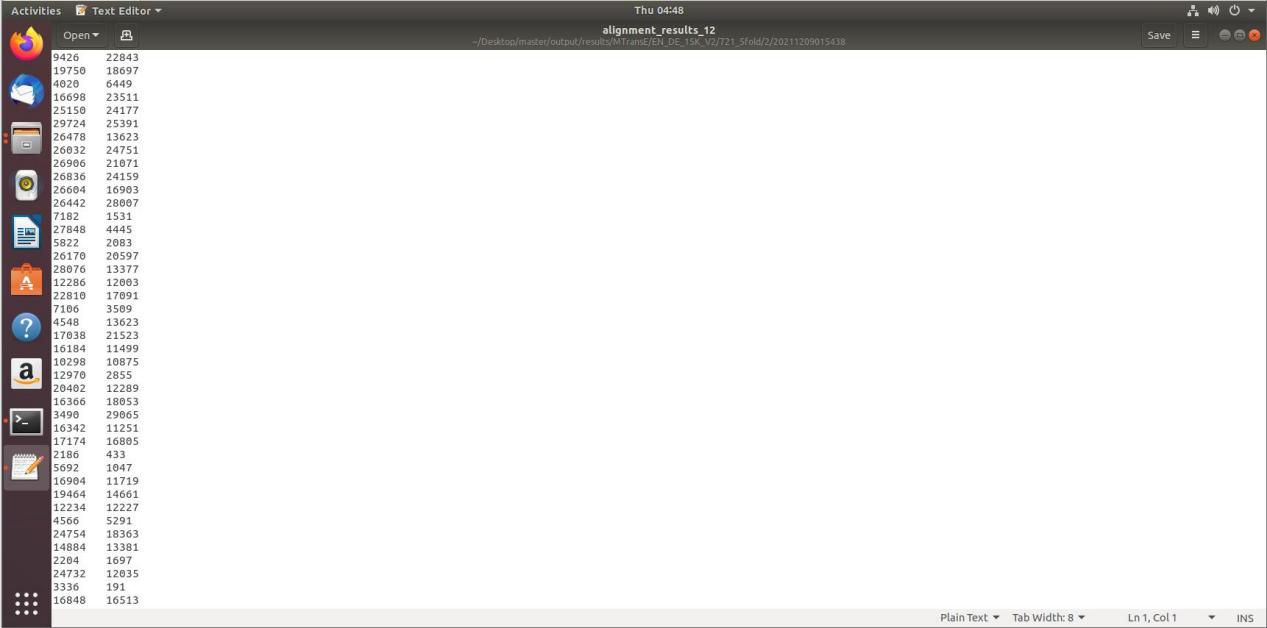


图9.对齐结果2

1. mtranse\_args\_15K.json和mtranse\_args\_100K.json有何区别，为什么要设置这种区别，而不是直接写一个mtranse\_args.json？

**答：**区别在于训练时的batch\_size不同，同时线程数量也对应的不同。

分开写两种batch size的模型参数，可以应对不同训练的需求。

1. 什么是earlystop？这个实例中为什么需要earlystop？

**答：**早停是用于在深度学习训练时防止过拟合的方法，希望模型泛化能力更强，故在训 练错误率达到一定值的时候就停止继续优化。

本实验中embedding方法是全连接神经网络，很容易过拟合，故选择用遭停发防止过拟合。