Table-BERT+GNN操作文档

1. 配置环境
2. python=3.6.13
3. pytorch=1.9.0

conda install pytorch torchvision torchaudio cudatoolkit=10.2 -c pytorch

1. transformers=2.6.0

pip install transformers==2.6.0

1. tensorboardX=2.4

pip install tensorboardX

1. pandas=0.25.1

conda install pandas=0.25.1

1. allennlp==0.9.0

pip install allennlp==0.9.0

注：如果安装allennlp时遇到报错ERROR: Failed building wheel for jsonnet，执行

conda install -c conda-forge jsonnet

1. 训练过程
2. 下载代码文件

git clone https://github.com/wenhuchen/GNN-TabFact.git

1. 在代码文件里创建文件夹存放模型

mkdir models

1. 下载预训练模型

ln -s TABFACT/data/all\_csv

cd models

wget https://gnntabfact.s3-us-west-2.amazonaws.com/gnn\_fp16\_numeric.zip

unzip gnn\_fp16\_numeric.zip

1. 进入代码文件里加载GNN模型训练结果

CUDA\_VISIBLE\_DEVICES=0 python gnn.py --model bert-base-multilingual-uncased --do\_test --encoding gnn --load\_from models/gnn\_fp16\_numeric/model\_ep4.pt

注：遇到ModuleNotFoundError: No module named 'google'问题，执行

pip install google

conda install protobuf

1. 重新训练模型

CUDA\_VISIBLE\_DEVICES=0 python gnn.py --model bert-base-multilingual-uncased --do\_train --encoding gnn --output\_dir models/gnn\_fp16\_numeric\_test --attention cross --lr\_default 5e-6

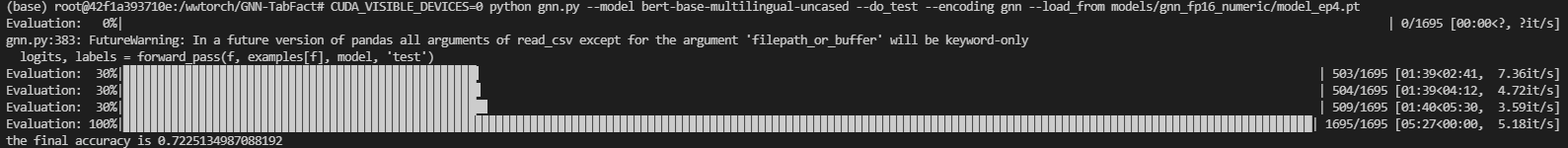
注：如果遇到提示说all\_csv里某个文件没有，就删去all\_csv的链接，直接下载TabFact数据集

git clone <https://github.com/wenhuchen/Table-Fact-Checking.git>

然后将其中的all\_csv文件移动到GNN-TabFact目录里

mv all\_csv /wwtorch/GNN-TabFact（后面是GNN-TabFact目录的绝对地址）

4的结果：0.7225134987088192



5的结果：最高0.7105407308866108

