实验四

实验要求

使用实验三的数据集，在预训练的Bert模型上微调参数。使用微调后的Bert模型做文本分类，并与实验三的RNN模型进行对比分析

模型下载地址

https://huggingface.co/bert-base-chinese/tree/main

下载完模型参数和配置完文件，在本地的使用方式：

如何下载Hugging Face 模型以及如何在local使用

你会碰到警告

```
Some weights of the model checkpoint at ./bert-base-chinese were not used when initializing BertModel: ['cls.predictions.transform.LayerNorm.weight', 'c
ls.seq_relationship.bias', 'cls.predictions.bias', 'cls.predictions.transform.LayerNorm.bias', 'cls.predictions.decoder.weight', 'cls.seq_relationship.w
eight', 'cls.predictions.transform.dense.weight', 'cls.predictions.transform.dense.bias']
- This IS expected if you are initializing BertModel from the checkpoint of a model trained on another task or with another architecture (e.g. initializ
ing a BertForSequenceClassification model from a BertForPreTraining model).
- This IS NOT expected if you are initializing BertModel from the checkpoint of a model that you expect to be exactly identical (initializing a BertForS
equenceClassification model from a BertForSequenceClassification model).
```

解决方法：

nlp - Python: BERT Error - Some weights of the model checkpoint at were not used when initializing BertModel - Stack Overflow

Hugging face 微调预训练模型2 使用trainer

Out-of-emory 如何解决

When using a GPU with 12GB - 16GB of RAM, you are likely to encounter out-of-memory issues if you use the same hyperparameters described in the paper. 此时可以调整以下参数：

max_seq_length: 训好的模型用512，可以适当调小
train_batch_size:
Model type, BERT-Base vs. BERT-Large: The BERT-Large model requires more memory.
Optimizer: 训好的模型用Adam, requires a lot of extra memory for the m and v vectors. Switching to a more memory efficient optimizer can reduce memory usage, but can also affect the results.
原文链接：https://blog.csdn.net/ccbrid/article/details/88732857

图形化fine-tune

bert-base-chinese · Hugging Face 支持交互传入训练集，对模型进行微调。感兴趣的同学可以尝试一下，并把实验结果放入报告中。该项不作为硬性要求。

实验结果提交

- 实验报告(.pdf)：需要包含实验过程，数据预处理，对比分析等等
- 实验代码(.py)
- 以上结果打包为压缩包，发至dl2021_bigdata@163.com
- 压缩包命名格式请按：学号_姓名_实验四，截止日期为12.6晚23:59

Reference

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

Tips:bert   tpu