



黑马程序员线上品牌

基于BERT+P-Tuning分类模型搭建

一样的教育，不一样的品质



目录

Contents

1. 实现模型工具类函数
2. 实现模型训练函数,验证函数
3. 实现模型预测函数

01

实现模型工具类函数

I 基本介绍

01

目的

模型在训练、验证、预测
时需要的函数

02

代码路径

/Users/**/PycharmPro
jects/llm/prompt_tasks
/P-Tuning/utils

03

脚本

utils文件夹共包含3个py
脚本：verbalizer.py、
metirc_utils.py以及
common_utils.py

I verbalizer.py

目的： 定义一个Verbalizer类，用于将一个Label对应到其子Label的映射。

导入必备工具包

```
# -*- coding:utf-8 -*-  
import os  
from typing import Union, List  
from ptune_config import *  
pc = ProjectConfig()
```

print结果显示

```
[  
  {'sub_labels': ['电脑'], 'token_ids': [[4510, 5554]]},  
  {'sub_labels': ['衣服'], 'token_ids': [[6132, 3302]]}  
]
```

common_utils.py

目的： 定义损失函数、将mask_position位置的token logits转换为token的id。

脚本里面包含两个函数：mlm_loss()以及convert_logits_to_ids()

导入必备的工具包：

```
# coding:utf-8  
# 导入必备工具包  
import torch  
from rich import print
```

定义损失函数mlm_loss()

定义convert_logits_to_ids()函数

print打印结果显示

```
tensor([[2499, 3542],  
        [5080, 8982]])
```

metirc_utils.py

目的：定义（多）分类问题下的指标评估（acc, precision, recall, f1）。

导入必备的工具包：

```
from typing import List

import numpy as np
import pandas as pd
from sklearn.metrics import accuracy_score, precision_score, f1_score
from sklearn.metrics import recall_score, confusion_matrix
```

定义ClassEvaluator类

print代码结果

```
{
  'accuracy': 0.6,
  'precision': 0.7,
  'recall': 0.6,
  'f1': 0.6,
  'class_metrics': {
    '体育': {'precision': 0.5, 'recall': 0.5, 'f1': 0.5},
    '计算机': {'precision': 1.0, 'recall': 0.5, 'f1': 0.67},
    '财经': {'precision': 0.5, 'recall': 1.0, 'f1': 0.67}
  }
}
```

02

实现模型训练函数,验证函数

I 简介

01

目的

实现模型的训练和验证

02

代码路径

/Users/**/PycharmPro
jects/llm/prompt_tasks
/P-Tuning/train.py

03

函数

脚本里面包含两个函数：
model2train()和
evaluate_model()

I 代码实现

导入必备的工具包

```
import os
import time
from transformers import AutoModelForMaskedLM, AutoTokenizer, get_scheduler
import sys
sys.path.append('/Users/**/PycharmProjects/llm/prompt_tasks/P-Tuning/data_handle')
sys.path.append('/Users/**/PycharmProjects/llm/prompt_tasks/P-Tuning/utils')
from utils.metirc_utils import ClassEvaluator
from utils.common_utils import *
from data_handle.data_loader import *
from utils.verbalizer import Verbalizer
from ptune_config import *

pc = ProjectConfig()
```

定义model2train()函数

定义evaluate_model函数

代码实现

调用

```
cd /Users/**/PycharmProjects/llm/prompt_tasks/P-Tuning  
# 实现模型训练  
python train.py
```

输出结果

```
...  
global step 350, epoch: 43, loss: 0.10804, speed: 1.20 step/s  
global step 360, epoch: 44, loss: 0.10504, speed: 1.22 step/s  
global step 370, epoch: 46, loss: 0.10220, speed: 1.21 step/s  
global step 380, epoch: 47, loss: 0.09951, speed: 1.20 step/s  
global step 390, epoch: 48, loss: 0.09696, speed: 1.20 step/s  
global step 400, epoch: 49, loss: 0.09454, speed: 1.22 step/s  
Evaluation precision: 0.76000, recall: 0.70000, F1: 0.70000
```

结论: BERT+P-Tuning模型在训练集上的表现是Precision: 76%

注意: 本项目中只用了60条样本，在接近400条样本上精确率就已经达到了76%，如果想让指标更高，可以扩增样本。

I 提升模型性能

增加训练数据集 (100
条左右的数据)

手机	外观时尚新潮，适合年轻人展现个性。
手机	屏幕显示效果非常出色，观看视频和浏览网页很舒适。
电脑	使用了一段时间的这款电脑，硬盘采用WD，运行流畅无卡顿，温度控制较好，性价比令人满意。
手机	手机反应灵敏，操作界面简洁易用，非常满意。
电器	产品性能稳定，很不错哦！购买时有点担心，但收到货后发现是正品，大家可以放心购买。

修改验证
集脏数据

原始标签和评论文本内容不符

平板	手机很好，就是客服垃圾特别是元豆
----	------------------

修改后

手机	手机很好，就是客服垃圾特别是元豆
----	------------------

模型表现: Evaluation precision: 0.79000, recall: 0.70000, F1: 0.71000

03

实现模型预测函数

I 代码介绍

目的

加载训练好的模型并
测试效果

/Users/**/PycharmProj
ects/llm/prompt_tasks/
P-Tuning/inference.py

代码路径

I 代码实现

导入必备的工具包

```
import time
from typing import List

import torch
from rich import print
from transformers import AutoTokenizer, AutoModelForMaskedLM
import sys
sys.path.append('/Users/**/PycharmProjects/llm/prompt_tasks/P-Tuning/data_handle')
sys.path.append('/Users/**/PycharmProjects/llm/prompt_tasks/P-Tuning/utils')
from utils.verbalizer import Verbalizer
from data_handle.data_preprocess import convert_example
from utils.common_utils import convert_logits_to_ids
```

代码实现

预测代码具体实现

结果展示

```
{  
    '天台很好看，躺在躺椅上很悠闲，因为活动所以我觉得性价比还不错，适合一家出行，特别是去迪士尼也蛮近的，下次有机会肯定还会再来的，值得推荐': '酒店',  
    '环境，设施，很棒，周边配套设施齐全，前台小姐姐超级漂亮！酒店很赞，早餐不错，服务态度很好，前台美眉很漂亮。性价比超高的一家酒店。强烈推荐': '酒店',  
    '物流超快，隔天就到了，还没用，屯着出游的时候用的，听方便的，占地小': '衣服',  
    '福行市来到无早集市，因为是喜欢的面包店，所以跑来集市看看。第一眼就看到了，之前在微店买了小刘，这次买了老刘，还有一直喜欢的巧克力磅蛋糕。好奇老板为啥不做柠檬磅蛋糕了，微店一直都是买不到的状态。因为不爱碱水硬欧之类的，所以期待老板多来点其他小点，饼干一直也是大爱，那天好像也没看到': '平板',  
    '服务很用心，房型也很舒服，小朋友很喜欢，下次去嘉定还会再选择。床铺柔软舒适，晚上休息很安逸，隔音效果不错赞，下次还会来': '酒店'  
}
```




总结

sum up

- 实现了基于BERT+P-Tuning模型的构建
- 完成了训练和测试评估



黑马程序员线上品牌

Thanks!



扫码关注博学谷微信公众号

