

NLP-Project2 面向目标的情感分类

201300096 杜兴豪 人工智能学院

这次任务要求我们从输入的句子中分析特定词语的情感极性。我的实现主要分为两部分：加载并预处理数据，训练模型并预测结果。

数据预处理模块

观察输入文件的组织形式，可以发现：训练集中每条样本为三行，其中第一行为原句子，第二行为目标词语，第三行为标签（即预测的情感极性）。因此我们可以每三行作为一个整体读取，拆分出不同属性分别整理；对于测试集，由于没有样本标签，因此每两行进行分组即可。

并且根据BERT模型的手册，其词典中已经进行了停用词去除等操作，这里也可以不用提前做。

注意：由于模型训练中对于分类任务采用的cross entropy损失函数，其输入的标签非负，因此我们需要对标签做整体偏移操作（不影响分类结果），即

```
def load_dataset():
    ...
    label = int(lines[i+2].strip())
    ...
    train_Y.append(label+1) # 注意此处的+1
    ...
```

模型训练与预测模块

这部分分为三个小部分：模型定义、模型训练、模型预测。主要介绍模型定义部分。

由于本次任务允许使用预训练的大模型，因此模型定义部分只需要根据所选的大模型来设计其所对应的输出头即可。我选择了BERT模型作为本次作业的基础模型，并且在改进过程中采用了其改进版RoBERTa（DeBERTa版本由于一些奇怪的系统bug而未被使用）切换模型非常简单，只需要在读取预训练模型时使用不同的方法即可，即

```
# 加载预训练的BERT模型和tokenizer
# model = BertModel.from_pretrained('BERT_MODEL')
# tokenizer = BertTokenizer.from_pretrained('BERT_MODEL')
# model = DebertaModel.from_pretrained('microsoft/deberta-v3-base')
# tokenizer = DebertaTokenizer.from_pretrained('microsoft/deberta-v3-base')
model = RobertaModel.from_pretrained('roberta-base')
tokenizer = RobertaTokenizer.from_pretrained('roberta-base')
```

使用哪个模型，就取消这个模型的注释即可。在本次任务中，可以使用BERT和RoBERTa，均可运行。

在模型定义部分，除了预训练模型的读取，还有模型结构部分。当我们使用预训练的BERT系模型时，其输出结果的 `last_hidden_state` 部分是针对问题做出的预测，因此我们只需要在此基础上新定义一个预测头，从隐状态映射到预测空间即可。具体操作如下：

```
logits = self.classifier(pooled_output)
# 应用softmax函数获取分类概率
return nn.functional.softmax(logits, dim=1)
```

在预测部分，只需要对softmax得到的类别概率选取最大的即可。

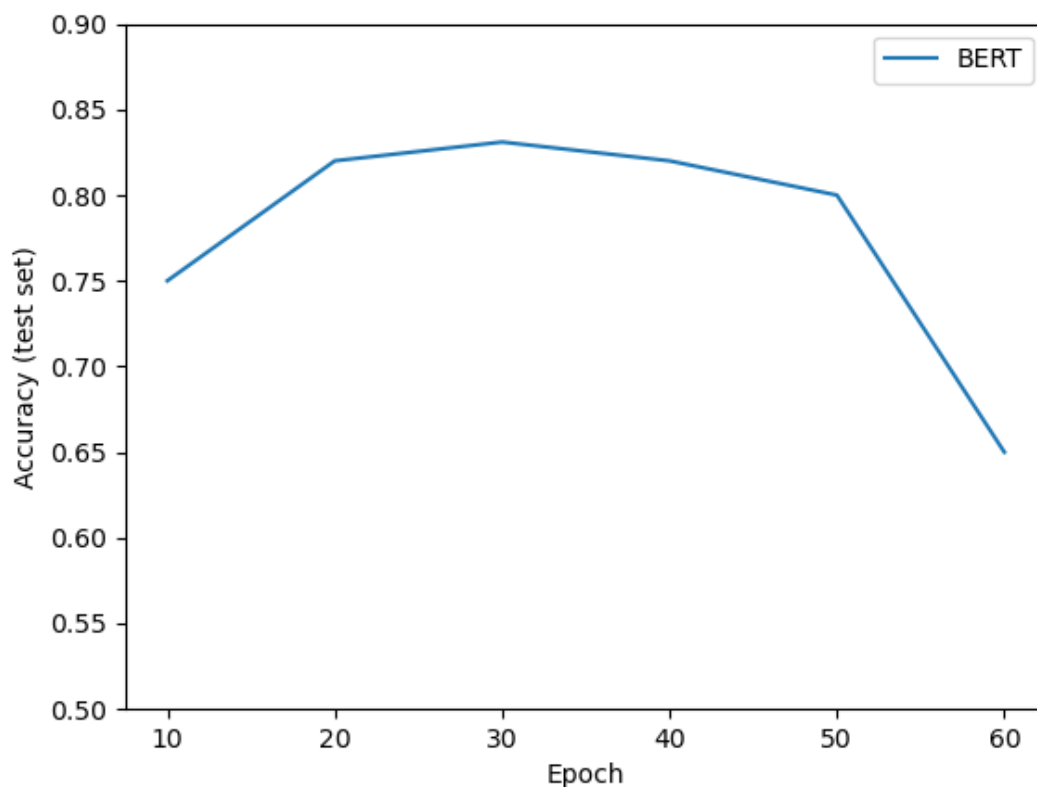
```
predicts.append(torch.argmax(output, dim=1)-1)
```

此处-1是将映射空间从 $\{0, 1, 2\}$ 修正为 $\{-1, 0, 1\}$ ，为合法的情感极性。

实验效果

BERT效果

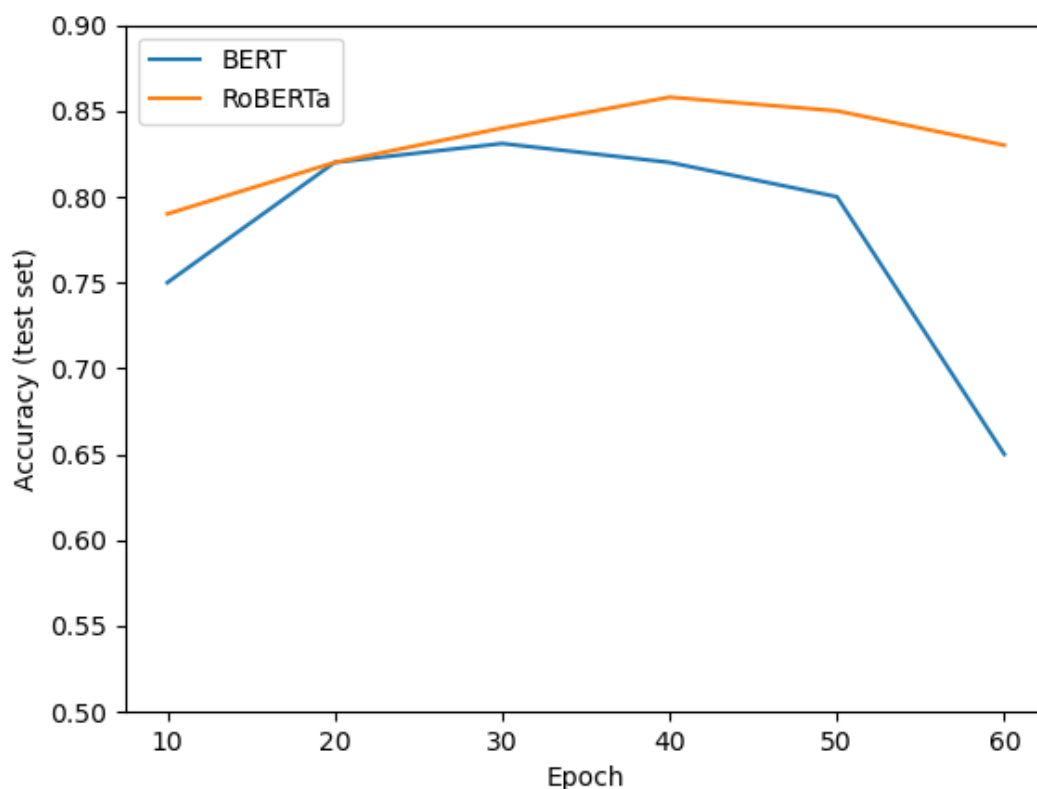
由于模型训练耗时较长，模型预测结果需要手动提交OJ，因此使用了较少的几次结果来说明实验效果。



观察到模型在20-30轮左右达到了最佳拟合状态，获得了83.5%的准确率。随着训练轮数增加，准确率反而下降，可能是出现了过拟合现象。

改进后效果

为了改进实验结果，我换用了更大的模型RoBERTa，并与BERT的实验结果进行对比，结果如下



观察到RoBERTa确实获得了更好的效果，使我的预测准确率提高到了85.89%

实验总结

本次实验我利用了预训练的BERT模型，完成了针对句子中某个词语的情感极性判断，并获得了很不错的效果（在OJ测试集上获得了85.89%的准确率）。虽然预训练模型的加入使得我的任务量小了很多，耗时较长的部分无非是fine tune阶段，但我认为这样的经历是十分有意义的，甚至比自己重复造轮子的实验还要有意义得多。

作为大一大二的本科生，在刚刚接触深度学习模型时，自己实现一些基础模型来训练，可以获得很大的收获，无论是对模型结构的理解还是对深度学习的认识都有加深。但是当我们成为大三学生时，仅仅使用基础模型来做试验已经无法满足我们目前的科研需要了。我们已经逐渐开始向研究生阶段走近，接触前沿科学，使用最新的模型的经历，从一定程度上来说已经更为重要了。所以这次实验是十分重要的，它让我们开始读前沿的大模型（比如我详细了解的BERT模型和其更新的变体），让我们学习怎么利用预训练的模型的接口来做下游任务，在我看来，这是一次十分宝贵的经历。