

# DL lab4

吴越凡 PB18000149

May 2021

## 1 实验目的

使用预训练的 Bert 模型，在之前的 IMDb 数据集上进行微调。运用模型对用户评价进行情感分析。其中用户的情感分为 neg 和 pos 两种情况。将准确率作为评判指标，去判断模型的效果。

## 2 实验细节

使用 python 库 transformers 中的 BertTokenizer 和 BertModel，使用的是 base 的 bert 模型。在库中 pretrain 的 bert 模型后加入一层 linear 层，通过 sigmoid 进行范围约束，最后输出一个值。

## 3 实验结果

图 1 中为每一个 Epoch 结束时，模型在测试集上的准确率，可以看出在 Epoch 之间，模型的效果差距极小，较为稳定。

表 1 中的结果为在不同 batch\_size 下 Bert 进行预测的结果。

表 2 中的结果为之前使用循环神经网络的结果。

## 4 总结

- 从实验结果可以明显看出 Bert 的效果远好于 RNN 的模型。

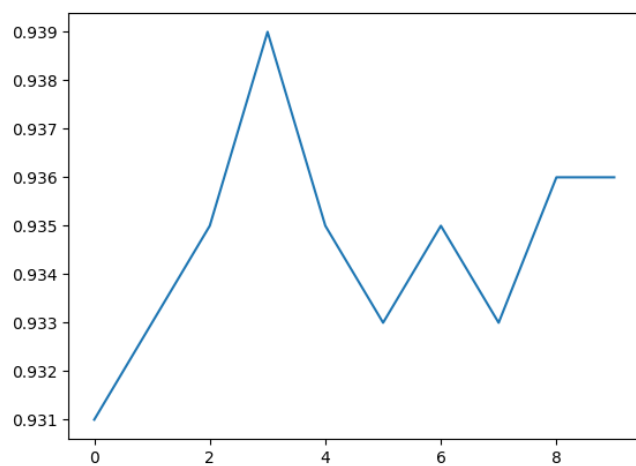


图 1: Accuracy with Bert model

Batch_size	Accuracy
1	94.0%
2	93.8%
4	93.9%

表 1: Results with Bert

- batch\_size 对模型效果影响很小，可以忽略。但是 batch\_size 为 4 的时候，模型结果相对不稳定。

Model	Glove dimension	Hidden size	Accuracy
Bi-LSTM	50	200	87.78%
Bi-LSTM	50	100	87.24%
Bi-LSTM	100	200	88.68%
Bi-LSTM	100	100	88.36%
LSTM	50	200	87.12%
LSTM	50	100	87.36%
LSTM	100	200	88.90%
LSTM	100	100	88.30%

表 2: Results with RNN