一、使用模型：

bert-tiny 和 bert-base-uncased (两个.py文件，只有模型名不同)

二、对于MNLI任务：

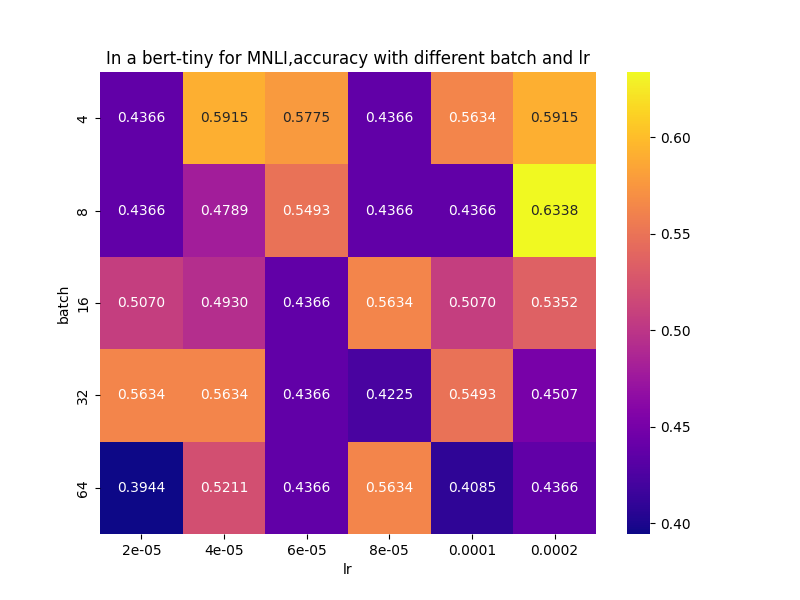
(1)结果总览

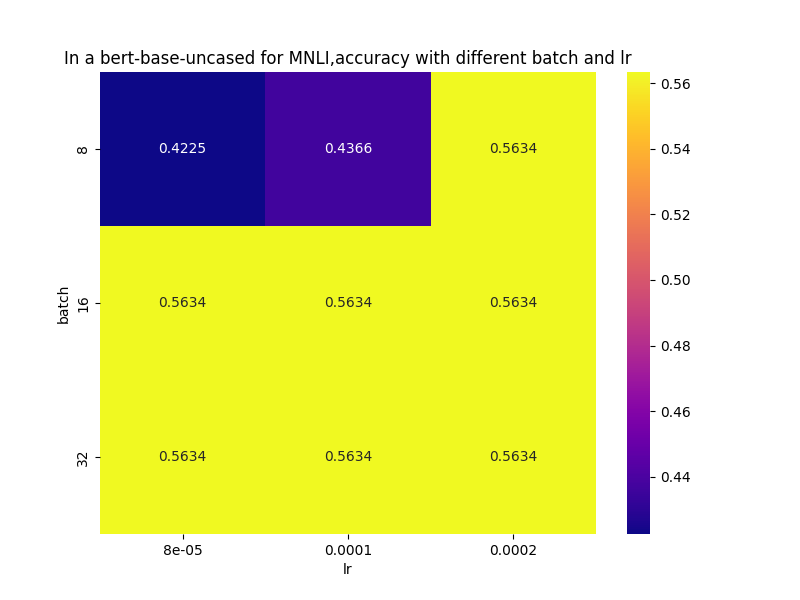
|  |  |  |
| --- | --- | --- |
|  | Finetune前 | Finetune后 |
| bert-tiny | 46.7% | 63.4% |
| bert-base-uncased | 43.7% | 56.4% |

(2)最佳超参：epoch=5

对于bert-tiny是 batch=8，lr=0.0002

对于bert-base-uncased是 batch=16，lr=0.0001





三、对于MRPC任务：

(1)结果总览

|  |  |  |
| --- | --- | --- |
|  | Finetune前 | Finetune后 |
| bert-tiny | 43.7% | 68.9% |
| bert-base-uncased | 35.0% | 68.4% |

(2)最佳超参：epoch=5

对于bert-tiny是 batch=8，lr=0.0002

对于bert-base-uncased是 batch=16，lr=0.0001

