

我是用 Roberta model，tokenizer 用 Roberta-base 作為預訓練資料。資料集我選擇了 The Stanford Sentiment Treebank(sst2)和 Recognizing Textual Entailment(rte)。資料集載入後就已經分好 train, validation, test 了，所以不用特別去處理。我先對兩個進行編碼，sst2 只有 sentence 要處理，rte 有 sentence1 和 sentence2 兩個要弄，處理完再輸入到 model。PEFT 的部分我嘗試了 BitFit 和 Lora。

BitFit 是選擇 model 裡面有 bias 的部分去更新，所以沒有 bias 的就不去更新

```
# BitFit
num_param = 0
for name,param in model.named_parameters():
    if "bias" not in name:
        param.requires_grad = False
    else:
        num_param += param.numel()
num_param
```

102914

Lora 在更新參數時，把比較大的矩陣運算分成兩個比較小的去計算，算完再結合起來

```
# lora
from peft import LoraConfig, TaskType, get_peft_model

config = LoraConfig(task_type=TaskType.CAUSAL_LM)
model = get_peft_model(model, config)
```

我整體跑完的感覺是用 PEFT 後，時間會縮短蠻多的，像原本 sst2 甚麼都不做要跑 1 小時左右，而有用 PEFT 後就縮短大概剩半小時，rte 原本跑 12 分鐘，但有座 PEFT 就剩大概 5 分鐘，時間上節省很多。而準確度的方面，我是覺得要看資料集，像 sst2 在有 PEFT 後，準確率在 9 成左右，不過沒有 PEFT 也才 0.94 而已，其實差不太大，而 rte 原本大概接近 7 成，在 PEFT 後只有 5 成，這個就差異蠻大的，不過我這次只有試著實作 PEFT，沒有特別去調整讓 PEFT 前後的準確度差不多。

SST2

No PEFT

Step	Training Loss	Validation Loss	Accuracy	TrainOutput(global_step=12627, training_loss=0.1843268603904873, metrics={'train_runtime': 3139.3357, 'train_samples_per_second': 64.36, 'train_steps_per_second': 4.022, 'total_flos': 6955865942774760.0, 'train_loss': 0.1843268603904873, 'epoch': 3.0})
500	0.355000	0.241495	0.915138	
1000	0.290100	0.248446	0.907110	
1500	0.202500	0.216583	0.926606	
2000	0.269000	0.260425	0.924312	
2500	0.248700	0.236062	0.925459	
3000	0.232000	0.250967	0.925459	
3500	0.129000	0.262399	0.931193	
4000	0.242000	0.221042	0.934633	
4500	0.141700	0.256578	0.933486	
5000	0.148800	0.283625	0.931193	
5500	0.216100	0.293963	0.932339	
6000	0.095900	0.282357	0.933486	
6500	0.194500	0.290690	0.928899	
7000	0.107900	0.275291	0.932339	
7500	0.131800	0.249021	0.928899	
8000	0.157500	0.230835	0.939220	
8500	0.132800	0.312093	0.943807	
9000	0.073900	0.310971	0.935780	
9500	0.147300	0.356428	0.933486	
10000	0.085000	0.329415	0.938073	
10500	0.236400	0.319992	0.935780	
11000	0.157400	0.340400	0.930046	
11500	0.145700	0.331999	0.936927	
12000	0.099100	0.316955	0.938073	
12500	0.125400	0.312330	0.940367	

Lora	BitFit																																																																																																																																																																																																												
<table><tr><th>Step</th><th>Training Loss</th><th>Validation Loss</th><th>Accuracy</th></tr><tr><td>500</td><td>0.701200</td><td>0.693839</td><td>0.490826</td></tr><tr><td>1000</td><td>0.673400</td><td>0.678896</td><td>0.509174</td></tr><tr><td>1500</td><td>0.543400</td><td>0.542241</td><td>0.565367</td></tr><tr><td>2000</td><td>0.478900</td><td>0.487387</td><td>0.852064</td></tr><tr><td>2500</td><td>0.340800</td><td>0.416550</td><td>0.886468</td></tr><tr><td>3000</td><td>0.413500</td><td>0.309826</td><td>0.888761</td></tr><tr><td>3500</td><td>0.349000</td><td>0.294305</td><td>0.891055</td></tr><tr><td>4000</td><td>0.406700</td><td>0.284147</td><td>0.901376</td></tr><tr><td>4500</td><td>0.284600</td><td>0.277787</td><td>0.902523</td></tr><tr><td>5000</td><td>0.265700</td><td>0.270504</td><td>0.907110</td></tr><tr><td>5500</td><td>0.251800</td><td>0.264294</td><td>0.908257</td></tr><tr><td>6000</td><td>0.261900</td><td>0.271743</td><td>0.903670</td></tr><tr><td>6500</td><td>0.308300</td><td>0.252660</td><td>0.913991</td></tr><tr><td>7000</td><td>0.235900</td><td>0.243402</td><td>0.918578</td></tr><tr><td>7500</td><td>0.241600</td><td>0.258529</td><td>0.908257</td></tr><tr><td>8000</td><td>0.235900</td><td>0.240111</td><td>0.912844</td></tr><tr><td>8500</td><td>0.279100</td><td>0.253426</td><td>0.911697</td></tr><tr><td>9000</td><td>0.227700</td><td>0.244254</td><td>0.913991</td></tr><tr><td>9500</td><td>0.251200</td><td>0.246645</td><td>0.911697</td></tr><tr><td>10000</td><td>0.369400</td><td>0.244606</td><td>0.912844</td></tr><tr><td>10500</td><td>0.327400</td><td>0.239054</td><td>0.913991</td></tr><tr><td>11000</td><td>0.264400</td><td>0.240210</td><td>0.911697</td></tr><tr><td>11500</td><td>0.268700</td><td>0.238253</td><td>0.912844</td></tr><tr><td>12000</td><td>0.175300</td><td>0.238260</td><td>0.911697</td></tr></table>	Step	Training Loss	Validation Loss	Accuracy	500	0.701200	0.693839	0.490826	1000	0.673400	0.678896	0.509174	1500	0.543400	0.542241	0.565367	2000	0.478900	0.487387	0.852064	2500	0.340800	0.416550	0.886468	3000	0.413500	0.309826	0.888761	3500	0.349000	0.294305	0.891055	4000	0.406700	0.284147	0.901376	4500	0.284600	0.277787	0.902523	5000	0.265700	0.270504	0.907110	5500	0.251800	0.264294	0.908257	6000	0.261900	0.271743	0.903670	6500	0.308300	0.252660	0.913991	7000	0.235900	0.243402	0.918578	7500	0.241600	0.258529	0.908257	8000	0.235900	0.240111	0.912844	8500	0.279100	0.253426	0.911697	9000	0.227700	0.244254	0.913991	9500	0.251200	0.246645	0.911697	10000	0.369400	0.244606	0.912844	10500	0.327400	0.239054	0.913991	11000	0.264400	0.240210	0.911697	11500	0.268700	0.238253	0.912844	12000	0.175300	0.238260	0.911697	<table><tr><th>Step</th><th>Training Loss</th><th>Validation Loss</th><th>Accuracy</th></tr><tr><td>500</td><td>0.699600</td><td>0.692876</td><td>0.521789</td></tr><tr><td>1000</td><td>0.681900</td><td>0.693897</td><td>0.509174</td></tr><tr><td>1500</td><td>0.675500</td><td>0.693602</td><td>0.509174</td></tr><tr><td>2000</td><td>0.674700</td><td>0.682883</td><td>0.509174</td></tr><tr><td>2500</td><td>0.479700</td><td>0.485322</td><td>0.848624</td></tr><tr><td>3000</td><td>0.487100</td><td>0.405401</td><td>0.858945</td></tr><tr><td>3500</td><td>0.448200</td><td>0.379682</td><td>0.862385</td></tr><tr><td>4000</td><td>0.465000</td><td>0.369830</td><td>0.866972</td></tr><tr><td>4500</td><td>0.358800</td><td>0.356569</td><td>0.871560</td></tr><tr><td>5000</td><td>0.391600</td><td>0.351986</td><td>0.872706</td></tr><tr><td>5500</td><td>0.364300</td><td>0.343075</td><td>0.875000</td></tr><tr><td>6000</td><td>0.336700</td><td>0.342192</td><td>0.878440</td></tr><tr><td>6500</td><td>0.421900</td><td>0.327646</td><td>0.880734</td></tr><tr><td>7000</td><td>0.286400</td><td>0.323573</td><td>0.880734</td></tr><tr><td>7500</td><td>0.405400</td><td>0.327272</td><td>0.878440</td></tr><tr><td>8000</td><td>0.315000</td><td>0.315254</td><td>0.886468</td></tr><tr><td>8500</td><td>0.365300</td><td>0.320504</td><td>0.883028</td></tr><tr><td>9000</td><td>0.273200</td><td>0.310700</td><td>0.887615</td></tr><tr><td>9500</td><td>0.294000</td><td>0.309080</td><td>0.888761</td></tr><tr><td>10000</td><td>0.433600</td><td>0.306957</td><td>0.891055</td></tr><tr><td>10500</td><td>0.383100</td><td>0.309659</td><td>0.887615</td></tr><tr><td>11000</td><td>0.315300</td><td>0.304541</td><td>0.892202</td></tr><tr><td>11500</td><td>0.334600</td><td>0.305781</td><td>0.891055</td></tr><tr><td>12000</td><td>0.283800</td><td>0.302965</td><td>0.891055</td></tr><tr><td>12500</td><td>0.314300</td><td>0.302356</td><td>0.893349</td></tr></table>	Step	Training Loss	Validation Loss	Accuracy	500	0.699600	0.692876	0.521789	1000	0.681900	0.693897	0.509174	1500	0.675500	0.693602	0.509174	2000	0.674700	0.682883	0.509174	2500	0.479700	0.485322	0.848624	3000	0.487100	0.405401	0.858945	3500	0.448200	0.379682	0.862385	4000	0.465000	0.369830	0.866972	4500	0.358800	0.356569	0.871560	5000	0.391600	0.351986	0.872706	5500	0.364300	0.343075	0.875000	6000	0.336700	0.342192	0.878440	6500	0.421900	0.327646	0.880734	7000	0.286400	0.323573	0.880734	7500	0.405400	0.327272	0.878440	8000	0.315000	0.315254	0.886468	8500	0.365300	0.320504	0.883028	9000	0.273200	0.310700	0.887615	9500	0.294000	0.309080	0.888761	10000	0.433600	0.306957	0.891055	10500	0.383100	0.309659	0.887615	11000	0.315300	0.304541	0.892202	11500	0.334600	0.305781	0.891055	12000	0.283800	0.302965	0.891055	12500	0.314300	0.302356	0.893349
Step	Training Loss	Validation Loss	Accuracy																																																																																																																																																																																																										
500	0.701200	0.693839	0.490826																																																																																																																																																																																																										
1000	0.673400	0.678896	0.509174																																																																																																																																																																																																										
1500	0.543400	0.542241	0.565367																																																																																																																																																																																																										
2000	0.478900	0.487387	0.852064																																																																																																																																																																																																										
2500	0.340800	0.416550	0.886468																																																																																																																																																																																																										
3000	0.413500	0.309826	0.888761																																																																																																																																																																																																										
3500	0.349000	0.294305	0.891055																																																																																																																																																																																																										
4000	0.406700	0.284147	0.901376																																																																																																																																																																																																										
4500	0.284600	0.277787	0.902523																																																																																																																																																																																																										
5000	0.265700	0.270504	0.907110																																																																																																																																																																																																										
5500	0.251800	0.264294	0.908257																																																																																																																																																																																																										
6000	0.261900	0.271743	0.903670																																																																																																																																																																																																										
6500	0.308300	0.252660	0.913991																																																																																																																																																																																																										
7000	0.235900	0.243402	0.918578																																																																																																																																																																																																										
7500	0.241600	0.258529	0.908257																																																																																																																																																																																																										
8000	0.235900	0.240111	0.912844																																																																																																																																																																																																										
8500	0.279100	0.253426	0.911697																																																																																																																																																																																																										
9000	0.227700	0.244254	0.913991																																																																																																																																																																																																										
9500	0.251200	0.246645	0.911697																																																																																																																																																																																																										
10000	0.369400	0.244606	0.912844																																																																																																																																																																																																										
10500	0.327400	0.239054	0.913991																																																																																																																																																																																																										
11000	0.264400	0.240210	0.911697																																																																																																																																																																																																										
11500	0.268700	0.238253	0.912844																																																																																																																																																																																																										
12000	0.175300	0.238260	0.911697																																																																																																																																																																																																										
Step	Training Loss	Validation Loss	Accuracy																																																																																																																																																																																																										
500	0.699600	0.692876	0.521789																																																																																																																																																																																																										
1000	0.681900	0.693897	0.509174																																																																																																																																																																																																										
1500	0.675500	0.693602	0.509174																																																																																																																																																																																																										
2000	0.674700	0.682883	0.509174																																																																																																																																																																																																										
2500	0.479700	0.485322	0.848624																																																																																																																																																																																																										
3000	0.487100	0.405401	0.858945																																																																																																																																																																																																										
3500	0.448200	0.379682	0.862385																																																																																																																																																																																																										
4000	0.465000	0.369830	0.866972																																																																																																																																																																																																										
4500	0.358800	0.356569	0.871560																																																																																																																																																																																																										
5000	0.391600	0.351986	0.872706																																																																																																																																																																																																										
5500	0.364300	0.343075	0.875000																																																																																																																																																																																																										
6000	0.336700	0.342192	0.878440																																																																																																																																																																																																										
6500	0.421900	0.327646	0.880734																																																																																																																																																																																																										
7000	0.286400	0.323573	0.880734																																																																																																																																																																																																										
7500	0.405400	0.327272	0.878440																																																																																																																																																																																																										
8000	0.315000	0.315254	0.886468																																																																																																																																																																																																										
8500	0.365300	0.320504	0.883028																																																																																																																																																																																																										
9000	0.273200	0.310700	0.887615																																																																																																																																																																																																										
9500	0.294000	0.309080	0.888761																																																																																																																																																																																																										
10000	0.433600	0.306957	0.891055																																																																																																																																																																																																										
10500	0.383100	0.309659	0.887615																																																																																																																																																																																																										
11000	0.315300	0.304541	0.892202																																																																																																																																																																																																										
11500	0.334600	0.305781	0.891055																																																																																																																																																																																																										
12000	0.283800	0.302965	0.891055																																																																																																																																																																																																										
12500	0.314300	0.302356	0.893349																																																																																																																																																																																																										
TrainOutput(global_step=12627, training_loss=0.3555099500030919, metrics={'train_runtime': 1923.1316, 'train_samples_per_second': 105.061, 'train_steps_per_second': 6.566, 'total_flos': 6979817059242984.0, 'train_loss': 0.3555099500030919, 'epoch': 3.0})	TrainOutput(global_step=12627, training_loss=0.43447850812989175, metrics={'train_runtime': 1769.146, 'train_samples_per_second': 114.206, 'train_steps_per_second': 7.137, 'total_flos': 6955865942774760.0, 'train_loss': 0.43447850812989175, 'epoch': 3.0})																																																																																																																																																																																																												

RTE

No PEFT				
<div><div></div></div> [468/468 12:01, Epoch 3/3]				
Step	Training Loss	Validation Loss	Accuracy	
100	0.695400	0.691188	0.527076	
200	0.689400	0.688032	0.527076	
300	0.692900	0.636151	0.649819	
400	0.527200	0.629323	0.693141	
TrainOutput(global_step=468, training_loss=0.6401754108249632, metrics={'train_runtime': 722.7767, 'train_samples_per_second': 10.335, 'train_steps_per_second': 0.648, 'total_flos': 1120914762487200.0, 'train_loss': 0.6401754108249632, 'epoch': 3.0})				
BitFit				
<div><div></div></div> [468/468 04:56, Epoch 3/3]				
Step	Training Loss	Validation Loss	Accuracy	
100	0.692700	0.693730	0.480144	
200	0.691100	0.693414	0.490975	
300	0.693200	0.693338	0.487365	
400	0.700000	0.693414	0.487365	
TrainOutput(global_step=468, training_loss=0.6948993796976204, metrics={'train_runtime': 298.7863, 'train_samples_per_second': 25.001, 'train_steps_per_second': 1.566, 'total_flos': 1120914762487200.0, 'train_loss': 0.6948993796976204, 'epoch': 3.0})				
Lora				
<div><div></div></div> [468/468 05:06, Epoch 3/3]				
Step	Training Loss	Validation Loss	Accuracy	
100	0.693500	0.691913	0.527076	
200	0.688800	0.691801	0.527076	
300	0.699500	0.691704	0.527076	
400	0.694800	0.691640	0.527076	
TrainOutput(global_step=468, training_loss=0.6969056323043301, metrics={'train_runtime': 307.8018, 'train_samples_per_second': 24.269, 'train_steps_per_second': 1.52, 'total_flos': 1124774405592480.0, 'train_loss': 0.6969056323043301, 'epoch': 3.0})				