

# LLM-Homework

wsl5300

June 2024

## 1 week6. learn basic pytorch



The screenshot shows a Jupyter Notebook interface with a code cell containing PyTorch training logic. The code defines an optimizer, iterates over 10 epochs, and prints the loss at each step. The output below the code cell shows the training progress and final loss.

```
optimizer = optim.Adam(model.parameters(), lr=0.001)

# 訓練模型
for epoch in range(10):
    for batch_X, batch_y in dataloader:
        # 前向傳播
        outputs = model(batch_X)
        loss = criterion(outputs, batch_y)

        # 反向傳播和優化
        optimizer.zero_grad()
        loss.backward()
        optimizer.step()

    print(f"Epoch [{epoch+1}/10], Loss: {loss.item():.4f}")

print("Training complete.")
```

```
/usr/local/lib/python3.10/dist-packages/torch/nn/modules/loss.py:535: UserWarning: Using a target size (torch.Size([32])) that is different to the input size (
return F.mse_loss(input, target, reduction=self.reduction)
/usr/local/lib/python3.10/dist-packages/torch/nn/modules/loss.py:535: UserWarning: Using a target size (torch.Size([8])) that is different to the input size (t
return F.mse_loss(input, target, reduction=self.reduction)
Epoch [1/10], Loss: 279.8156
Epoch [2/10], Loss: 221.8587
Epoch [3/10], Loss: 304.0934
Epoch [4/10], Loss: 303.2331
Epoch [5/10], Loss: 212.3000
Epoch [6/10], Loss: 231.5751
Epoch [7/10], Loss: 199.7448
Epoch [8/10], Loss: 234.3085
Epoch [9/10], Loss: 216.7897
Epoch [10/10], Loss: 188.6502
Training complete.
```

## 2 week7-1. LoRA

```
!pip install peft
from peft import LoraConfig, TaskType, get_peft_model, IA3Config
```

- Config-1.

```
peft_config = LoraConfig(
    lora_alpha=16,
    lora_dropout=0.1,
    r=64,
    bias="none",
    task_type="SEQ_CLS",
    target_modules=["q_lin", "k_lin", "v_lin", "out_lin"],
)
```

```
# train the model
trainer.train()
```

[80/80 01:06, Epoch 10/10]

Epoch	Training Loss	Validation Loss	Accuracy
1	No log	0.692889	0.437500
2	No log	0.690846	0.531250
3	No log	0.689358	0.656250
4	No log	0.688151	0.656250
5	No log	0.688213	0.656250
6	No log	0.687718	0.687500
7	No log	0.685953	0.625000
8	No log	0.685397	0.625000
9	No log	0.684702	0.687500
10	No log	0.684553	0.687500

TrainOutput(global\_step=80, training\_loss=0.6895392417907715, metrics={'train\_runtime': 67.0484, 'train\_samples\_per\_second': 19.091, 'train\_steps\_per\_second': 1.193, 'total\_flos': 29930055984768.0, 'train\_loss': 0.6895392417907715, 'epoch': 10.0})

- Config-2.

```
peft_config_2 = LoraConfig(
    lora_alpha=32,
    lora_dropout=0.05,
    r=128,
    bias="none",
    task_type="SEQ_CLS",
    target_modules=["q_lin", "k_lin", "v_lin", "out_lin", "ffn.lin1", "ffn.lin2"]
)
```

```
# train the model
trainer.train()
```

[80/80 01:17, Epoch 10/10]

Epoch	Training Loss	Validation Loss	Accuracy
1	No log	0.705034	0.437500
2	No log	0.697528	0.437500
3	No log	0.693479	0.468750
4	No log	0.690541	0.531250
5	No log	0.689977	0.531250
6	No log	0.689119	0.531250
7	No log	0.686893	0.625000
8	No log	0.686136	0.687500
9	No log	0.685228	0.687500
10	No log	0.685006	0.687500

TrainOutput(global\_step=80, training\_loss=0.6890310287475586, metrics={'train\_runtime': 77.2288, 'train\_samples\_per\_second': 16.574, 'train\_steps\_per\_second': 1.036, 'total\_flos': 35294415611520.0, 'train\_loss': 0.6890310287475586, 'epoch': 10.0})

- Config-3.

```
peft_config_3 = LoraConfig(
    lora_alpha=16,
    lora_dropout=0.1,
    r=64,
    bias="none",
    task_type="SEQ_CLS",
    target_modules=["q_lin", "v_lin", "ffn.lin1"],
)
```

```
# train the model
trainer.train()
```

[80/80 00:56, Epoch 10/10]

Epoch	Training Loss	Validation Loss	Accuracy
1	No log	0.697003	0.437500
2	No log	0.693748	0.468750
3	No log	0.691832	0.468750
4	No log	0.690024	0.531250
5	No log	0.689886	0.531250
6	No log	0.689400	0.562500
7	No log	0.687558	0.656250
8	No log	0.686916	0.687500
9	No log	0.686163	0.687500
10	No log	0.685982	0.656250

TrainOutput(global\_step=80, training\_loss=0.6881428718566894, metrics={'train\_runtime': 57.0207, 'train\_samples\_per\_second': 22.448, 'train\_steps\_per\_second': 1.403, 'total\_flos': 30121640257152.0, 'train\_loss': 0.6881428718566894, 'epoch': 10.0})

- Config\_IA3.

```
peft_config_IA3 = IA3Config(
    task_type=TaskType.SEQ_CLS,
    target_modules=["q_lin", "v_lin", "k_lin", "out_lin", "ffn.lin1", "ffn.lin2"],
    feedforward_modules=["ffn.lin1", "ffn.lin2"],
)
```

```
# train the model
trainer.train()
```

[80/80 01:06, Epoch 10/10]

Epoch	Training Loss	Validation Loss	Accuracy
1	No log	0.698027	0.437500
2	No log	0.694567	0.437500
3	No log	0.692525	0.468750
4	No log	0.690693	0.531250
5	No log	0.690063	0.531250
6	No log	0.690098	0.531250
7	No log	0.688323	0.656250
8	No log	0.687687	0.687500
9	No log	0.686978	0.656250
10	No log	0.686820	0.593750

TrainOutput(global\_step=80, training\_loss=0.6880824089050293, metrics={'train\_runtime': 66.5904, 'train\_samples\_per\_second': 19.222, 'train\_steps\_per\_second': 1.201, 'total\_flos': 28424323344000.0, 'train\_loss': 0.6880824089050293, 'epoch': 10.0})

### 3 week7-2. LaTeX

詳情請見檔案 mvc\_week7.pdf。