

目前进度

边学习《[动手学深度学习](#)》的内容边进行pytorch代码编写，基本熟悉LSTM操作流程。待消化各流程的数学知识后可以开始正式训练毕设所用模型。

```
1 import torch
2 from torch import nn
3 from d2l import torch as d2l
4
5 batch_size, num_steps = 32, 35
6 train_iter, vocab = d2l.load_data_time_machine(batch_size, num_steps)
7
8 # 初始化模型参数
9 # 超参数 num_hiddens 定义了隐藏单位的数量。
10 # 使用 0.01 标准差的高斯分布之后初始化权重，然后将偏置设置为 0。
11
12
13 def get_lstm_params(vocab_size, num_hiddens, device):
14     num_inputs = num_outputs = vocab_size
15
16     def normal(shape):
17         return torch.randn(size=shape, device=device) * 0.01
18
19     def three():
20         return (normal(
21             (num_inputs, num_hiddens)), normal((num_hiddens, num_hiddens)),
22             torch.zeros(num_hiddens, device=device))
23
24     w_xi, w_hi, b_i = three()
25     w_xf, w_hf, b_f = three()
26     w_xo, w_ho, b_o = three()
27     w_xc, w_hc, b_c = three()
28
29     w_hq = normal((num_hiddens, num_outputs))
30     b_q = torch.zeros(num_outputs, device=device)
31
32     params = [
33         w_xi, w_hi, b_i, w_xf, w_hf, b_f, w_xo, w_ho, b_o, w_xc, w_hc, b_c,
34         w_hq, b_q]
35     for param in params:
36         param.requires_grad_(True)
37     return params
38
39
40 # 定义模型
41 # 在初始化函数中，LSTM 的隐藏状态需要返回值为 0 且形状为（批量大小、隐藏单位数）的 附加
42 # 内存单元格。
43 def init_lstm_state(batch_size, num_hiddens, device):
44     return (torch.zeros((batch_size, num_hiddens), device=device),
45             torch.zeros((batch_size, num_hiddens), device=device))
46
47 # 模型提供三个门和一个辅助记忆细胞。只有隐藏状态才会传递到输出层。记忆单元不直接参与输出计
48 # 算。
```

```

48 def lstm(inputs, state, params):
49     [
50         w_xi, w_hi, b_i, w_xf, w_hf, b_f, w_xo, w_ho, b_o, w_xc, w_hc, b_c,
51         w_hq, b_q] = params
52     (H, C) = state
53     outputs = []
54     for x in inputs:
55         I = torch.sigmoid((x @ w_xi) + (H @ w_hi) + b_i)
56         F = torch.sigmoid((x @ w_xf) + (H @ w_hf) + b_f)
57         O = torch.sigmoid((x @ w_xo) + (H @ w_ho) + b_o)
58         C_tilda = torch.tanh((x @ w_xc) + (H @ w_hc) + b_c)
59         C = F * C + I * C_tilda
60         H = O * torch.tanh(C)
61         Y = (H @ w_hq) + b_q
62         outputs.append(Y)
63     return torch.cat(outputs, dim=0), (H, C)
64
65
66 vocab_size, num_hiddens, device = len(vocab), 256, d2l.try_gpu()
67 num_epochs, lr = 500, 1
68 # 选择合适的CPU、GPU开始训练
69 num_inputs = vocab_size
70 lstm_layer = nn.LSTM(num_inputs, num_hiddens)
71 model = d2l.RNNModel(lstm_layer, len(vocab))
72 model = model.to(device)
73 d2l.train_ch8(model, train_iter, vocab, lr, num_epochs, device)
74

```

感觉进度较慢

二三月份主要在紧张环境下准备考研复试（于四月初收到拟录取通知，感谢李老师的推荐），毕设相关的不是很多，同时由于对路线图不熟悉，四月刚开始阅读了《机器学习》（西瓜书）的前六章，感觉其偏向基础理论的知识也不能让我很快地开始项目的进度，直到这两天发现了《[动手学深度学习](#)》这本书，才让我有点豁然开朗的感觉，希望能在4月结束前完成大体，未来也一定还需要老师的指导，感谢老师！