

# 深度学习与神经网络 Chap6

2153067 王灏博

1. 完成了 torch 版本的 RNN 生成任务
2. 解释 RNN，LSTM，GRU 模型：传统的 RNN 存在梯度消失或梯度爆炸的问题，这导致它们难以捕捉长期依赖关系。为了解决这个问题，长短期记忆网络（LSTM）和门控循环单元（GRU）等变种模型被提出。LSTM 和 GRU 都是 RNN 的变种，旨在更有效地处理长期依赖关系。它们通过引入门控机制来控制信息的流动，从而有选择地记忆或忽略一些信息。
3. 叙述这个诗歌生成的过程：生成过程从一个特殊的起始符号开始，然后逐步生成下一个词和隐含向量，直到生成一个结束符号或达到指定的句子长度。
4. 生成的截图如下图所示：

```
log_softmax has been deprecated. Change the call
return self.call_impl(*args, **kwargs)
若为一顾人，不见一声生。
error
inital linear weight
一夜一声开，一声如有水。
error
inital linear weight
夜来无人客，风生不自闲。
一千秋万里，一望一声新。
```

训练的过程如图所示：

```
*****
epoch 1 batch number 258 loss is: 7.1556196212768555
prediction [66, 5, 1, 1, 0, 1, 1, 0, 4, 32, 0, 1, 1, 0, 0, 1, 3, 1, 0, 30, 1, 1, 0, 0, 4, 44, 1, 1, 0, 0, 1, 1, 3, 229, 1, 109, 66, 1, 1, 0, 4, 1, 0, 9, 14,
0, 0, 1, 3, 95, 0, 9, 1, 1, 0, 0, 4, 1, 1, 64, 0, 15, 64, 1, 3, 3]
b.y [173, 75, 25, 16, 318, 732, 61, 0, 426, 867, 362, 633, 29, 419, 81, 1, 268, 72, 37, 62, 86, 122, 1189, 0, 69, 12, 123, 1342, 550, 50, 6, 1, 354,
284, 230, 90, 174, 152, 271, 0, 11, 393, 67, 136, 148, 159, 89, 1, 128, 797, 589, 410, 83, 697, 114, 0, 40, 58, 9, 42, 179, 9, 183, 1, 3, 3]
*****
epoch 1 batch number 259 loss is: 7.113171577453613
prediction [22, 115, 1, 1, 0, 0, 1, 1, 4, 1, 0, 1, 1, 19, 66, 1, 3, 0, 0, 69, 1, 23, 0, 0, 4, 109, 1, 17, 0, 0, 1, 1, 3, 1, 1, 0, 0, 1, 1, 71, 4, 0, 0, 1,
1, 0, 0, 1, 3, 161, 0, 0, 1, 9, 0, 1, 4, 41, 1, 14, 20, 0, 1, 1, 3, 3]
b.y [445, 115, 25, 8, 255, 677, 677, 0, 30, 29, 182, 326, 92, 90, 182, 1, 122, 439, 88, 1839, 68, 347, 1319, 0, 230, 423, 544, 2231, 1831, 155, 683,
1, 208, 71, 218, 77, 1053, 356, 308, 0, 696, 63, 297, 14, 41, 202, 562, 1, 274, 394, 401, 174, 67, 19, 141, 0, 4, 1040, 136, 39, 1978, 356, 1507, 1, 3, 3]
*****
epoch 1 batch number 260 loss is: 7.094685077667236
finish save model
prediction [66, 178, 1, 9, 0, 45, 1, 1, 4, 0, 0, 1, 1, 0, 0, 1, 3, 1, 0, 0, 1, 1, 214, 0, 4, 1, 0, 0, 0, 0, 1, 1, 3, 0, 1, 1, 0, 1, 1, 0, 4, 0, 0, 1, 1, 0,
0, 1, 3, 0, 1, 0, 0, 62, 1, 0, 4, 0, 1, 1, 0, 0, 1, 1, 3, 3]
b.y [2028, 178, 589, 91, 249, 208, 62, 0, 82, 563, 773, 1069, 351, 1217, 2610, 1, 154, 460, 1073, 1628, 748, 308, 1976, 0, 76, 1420, 70, 1201, 601,
562, 1198, 1, 634, 602, 208, 2341, 952, 121, 1782, 0, 1588, 1529, 416, 134, 53, 685, 12, 1, 361, 684, 270, 566, 387, 24, 29, 0, 376, 181, 208, 1229, 18, 75,
343, 1, 3, 3]
*****
epoch 1 batch number 261 loss is: 7.0931925773620605
prediction [66, 1, 1842, 1, 4, 0, 1, 1, 3, 0, 1842, 1, 4, 1, 0, 1, 3, 0, 1842, 1, 4, 41, 1, 1, 3, 0, 1842, 1, 4, 1, 41, 30, 3, 0, 0, 1, 1, 1, 187, 178, 4,
1, 1, 26, 19, 64, 1, 108, 3, 62, 1, 22, 0, 41, 1, 1, 4, 1, 0, 451, 0, 86, 0, 1, 3, 3]
b.y [12, 1937, 1458, 0, 661, 13, 217, 1, 153, 1937, 1458, 0, 762, 218, 75, 1, 1641, 1937, 1458, 0, 4, 28, 1725, 1, 17, 1937, 1458, 0, 362, 149, 278,
1, 34, 191, 2975, 1125, 87, 418, 10, 0, 5, 164, 589, 31, 9, 368, 1573, 1, 194, 204, 688, 1284, 4, 55, 20, 0, 94, 427, 3911, 3912, 26, 1601, 3991, 1, 3, 3]
*****
```