

# Chapter 6

2153274 李林依

1. 完成 torch 版本
2. 传统的循环神经网络（RNN）存在梯度消失或梯度爆炸的问题，这使得它们难以有效捕捉长期依赖关系。为了解决这个问题，研究者们提出了一些改进模型，如长短期记忆网络（LSTM）和门控循环单元（GRU）等。这些模型都是 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，例如输入门、遗忘门和输出门（LSTM）或更新门和重置门（GRU），来控制信息的流动。这样，模型可以有选择地记忆或忽略一些信息，从而更好地捕捉序列数据中的长期依赖关系。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。
3. 生成截图：

```
return self.call_impl(*args, **kwargs)
日生三千里，千年万里云。
不辞天下主，不见白头人。
error
inital linear weight
红紫霞岩上，清风起寒水。
不见青山出，还将云树花。
error
inital linear weight
一片空余一水色，一声声下天上花。
error
```

训练过程：

```
*****
epoch 6 batch number 79 loss is: 6.535689830780029
prediction [10, 18, 0, 1, 1, 0, 0, 1, 4, 34, 0, 0, 0, 34, 0, 0, 3, 0, 1, 40, 0, 20, 1, 0, 4, 31, 51, 7, 1, 8, 14, 1, 3, 3]
b_y [108, 369, 50, 6, 1477, 85, 904, 0, 10, 107, 107, 73, 10, 107, 107, 1, 170, 903, 110, 43, 453, 5, 1135, 0, 138, 24, 34, 396, 1, 3, 3]
*****
epoch 6 batch number 80 loss is: 6.311267852783203
finish save model
prediction [10, 0, 0, 1, 1, 0, 15, 1, 4, 27, 90, 1, 0, 27, 1, 1, 3, 49, 39, 115, 0, 1, 1, 0, 4, 27, 9, 1, 1, 49, 0, 1, 3, 3]
b_y [1655, 1736, 501, 2520, 1091, 179, 2080, 0, 799, 230, 2676, 3161, 799, 423, 265, 1, 31, 136, 1306, 197, 2320, 262, 2677, 5, 509, 1152, 5, 1, 3, 3]
*****
epoch 6 batch number 81 loss is: 6.458913803100586
prediction [10, 1, 0, 1, 1, 72, 0, 1, 4, 214, 1, 49, 5, 14, 1, 1, 3, 208, 1, 27, 0, 0, 47, 0, 4, 49, 30, 5, 1, 0, 0, 1, 3, 3]
b_y [208, 1229, 46, 146, 82, 3463, 313, 0, 786, 162, 67, 19, 876, 14, 246, 1, 1541, 208, 4, 51, 107, 26, 41, 0, 715, 3, 763, 1669, 1, 3, 3]
*****
epoch 6 batch number 82 loss is: 6.5226030349731445
prediction [10, 0, 0, 1, 1, 46, 0, 1, 4, 0, 0, 1296, 0, 0, 0, 7, 3, 14, 1, 0, 0, 71, 1, 0, 4, 0, 0, 1, 1, 27, 0, 1, 3, 3]
b_y [606, 515, 58, 65, 3820, 193, 76, 0, 100, 1386, 915, 1661, 720, 1902, 13, 1, 34, 14, 160, 120, 1225, 1364, 42, 0, 84, 4, 41, 5, 1, 3, 3]
*****
epoch 6 batch number 83 loss is: 6.537601470947266
prediction [10, 34, 14, 1, 1, 0, 0, 1, 4, 0, 1, 1, 0, 34, 1, 1, 3, 12, 1, 0, 0, 8, 8, 0, 4, 41, 0, 19, 1, 18, 0, 1, 3, 3]
b_y [10, 34, 46, 7, 264, 940, 105, 0, 758, 20, 381, 85, 77, 76, 25, 1, 372, 12, 1736, 44, 3488, 3488, 272, 0, 1021, 12, 18, 88, 1, 3, 3]
*****
epoch 6 batch number 84 loss is: 6.5023369789123535
prediction [10, 34, 0, 0, 1, 0, 23, 0, 4, 1, 1, 31, 19, 5, 1, 1, 3, 0, 1, 1, 0, 1, 328, 0, 4, 0, 27, 20, 1, 51, 28, 1, 3, 3]
b_y [10, 5443, 2434, 658, 312, 190, 3714, 0, 642, 21, 138, 92, 19, 133, 1043, 1, 274, 27, 228, 587, 55, 1117, 328, 0, 5, 439, 2303, 666, 1, 3, 3]
*****
epoch 6 batch number 85 loss is: 6.41445779800415
prediction [10, 0, 0, 1, 1, 0, 90, 1, 4, 1, 1, 66, 0, 15, 1, 1, 3, 0, 57, 0, 0, 27, 1, 0, 4, 0, 0, 1, 0, 44, 0, 1, 3, 3]
b_y [467, 467, 94, 14, 248, 276, 81, 0, 127, 131, 33, 18, 179, 6, 65, 1, 456, 992, 73, 8, 4, 24, 29, 0, 685, 596, 46,
```

