

HW3 RNN 诗歌生成实验报告

1. 解释一下 RNN , LSTM, GRU 模型。

RNN:

RNN(循环神经网络)能够处理序列输入，并利用隐藏状态来记录之前输入的信息，并影响当前的输出。

RNN 通过当前状态的输入及上一状态的隐变量来计算当前状态的隐变量，并通过隐变量来计算输出。通常选用 \tanh 作为激活函数。

$$h_t = f(W_h h_{t-1} + W_x x_t + b)$$

$$y_t = W_y h_t + b_y$$

GRU:

RNN 的问题在于，它将所有序列信息全部放在隐藏状态中。当序列较长时，隐藏状态无法容纳足够的信息。同时，在观察一个序列时，并不是每个观察值都同等重要，而在 RNN 中没有特别关心某些地方的机制。因此 GRU 引入了门控机制，通过更新门和重置门，使得在构造隐藏状态的时候能够挑选出相对来说更加重要的部分。

$$z_t = \sigma(W_z[h_{t-1}, x_t]) \quad (\text{更新门, 控制旧信息保留多少})$$

$$r_t = \sigma(W_r[h_{t-1}, x_t]) \quad (\text{重置门, 控制遗忘多少})$$

$$\tilde{h}_t = \tanh(W[r_t * h_{t-1}, x_t]) \quad (\text{候选新状态})$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

LSTM:

LSTM 比 GRU 多引入了一个门，三个门分别是遗忘门、输入门和输出门。同时增加了一个状态，称为细胞状态。由隐藏状态负责短期记忆，遗忘门和输入门来调控信息的更新和丢弃，由输出门调控隐藏状态的更新。同时细胞状态负责长期记忆，由门、隐藏状态共同调控。

$$\begin{aligned}
I_t &= \sigma(X_t W_{xi} + H_{t-1} W_{hi} + b_i) \\
F_t &= \sigma(X_t W_{xf} + H_{t-1} W_{hf} + b_f) \\
O_t &= \sigma(X_t W_{xo} + H_{t-1} W_{ho} + b_o) \\
\tilde{C}_t &= \tanh(X_t W_{xc} + H_{t-1} W_{hc} + b_c) \\
C_t &= F_t \odot C_{t-1} + I_t \odot \tilde{C}_t \\
H_t &= O_t \odot \tanh(C_t)
\end{aligned}$$

2. 叙述一下这个诗歌生成的过程。

根据提供的第一个字，模型对其进行编码，转化为词向量，进入 LSTM。模型根据输入的序列预测下一个字的概率分布，选择概率最高的字作为下一个字。随后将之前的序列与预测的下一个字作为输入，再次输入模型进行预测，直到模型输出结束标记或者长度达到指定上限时，预测结束。

3. 训练过程与诗歌生成。

```

b_y      [465, 139, 235, 115, 161, 964, 79, 0, 102, 7, 182, 252, 130, 26, 602, 1, 258, 547, 380, 145, 60, 1393, 816, 0, 203, 124, 11, 21, 237, 874, 14
, 1, 39, 63, 376, 253, 24, 16, 126, 0, 15, 58, 92, 335, 90, 132, 975, 1, 61, 215, 118, 48, 484, 214, 789, 0, 1251, 1251, 52, 5, 73, 19, 280, 1, 3, 3]
epoch    28 batch number 317 loss is:  6.360990047454834
prediction [10, 197, 15, 18, 64, 26, 477, 0, 68, 90, 33, 23, 4, 18, 164, 1, 10, 42, 19, 396, 59, 182, 182, 0, 34, 20, 9, 18, 120, 83, 1011, 1, 60, 21, 4
, 70, 474, 62, 29, 0, 7, 17, 138, 20, 4, 18, 79, 1, 128, 212, 257, 32, 26, 19, 117, 0, 4, 115, 15, 18, 4, 24, 5, 1, 3, 3]
b_y      [122, 135, 107, 120, 1453, 313, 319, 0, 156, 244, 113, 5, 400, 306, 802, 1, 15, 42, 169, 554, 779, 2202, 667, 0, 39, 21, 6, 148, 1780, 277, 4
29, 1, 1694, 1956, 19, 341, 2007, 533, 1701, 0, 7, 17, 9, 5, 1211, 385, 116, 1, 99, 291, 39, 32, 40, 182, 117, 0, 415, 582, 22, 1468, 703, 24, 38, 1, 3,
3]
epoch    28 batch number 318 loss is:  6.255128860473633

```

```

问题 30  输出  调试控制台  终端  端口  SPELL CHECKER 23
• (pytorch) PS C:\MyE\1课程\深度学习与神经网络\hw3> python main.py
cuda
D:\anaconda3\envs\pytorch\Lib\site-packages\torch\nn\modules\module.py:1511: UserWarning: Implicit dimension
include dim=X as an argument.
  return self._call_impl(*args, **kwargs)
日能分处，不见五时清。自是无人识，何人问所宜。

红叶下，春风满夕阳。何当如可托，何必更无因。

山门落。

夜风吹与烟。

湖上春草，一枝花满枝。

海光春，春风满眼开。一声随客侣，千里有人头。

月明明地，春风起中流。

君去，不见东山客。

```