homework4报告

1-5 answer in detail

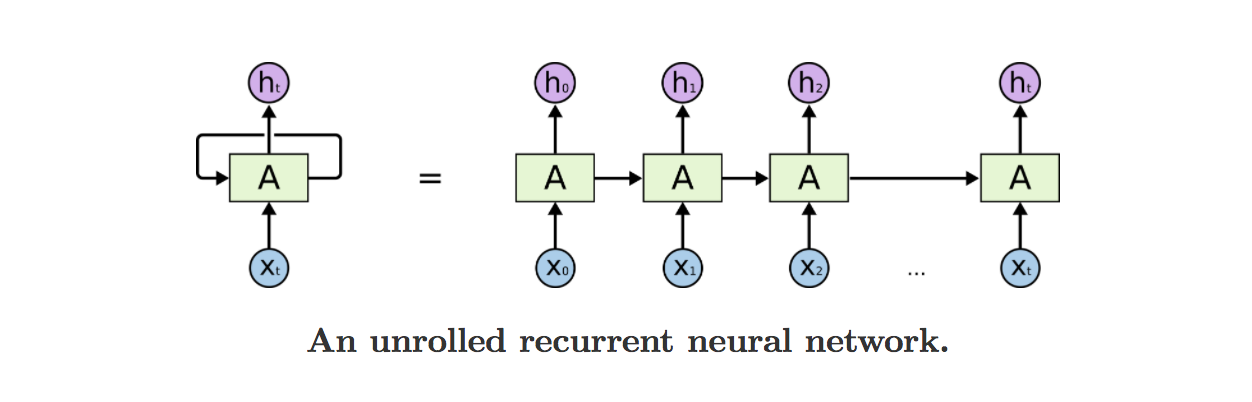
1.The code is in folder ‘’code’’,There are 4 program in the folder:

The 3 blank of Cell, MultiRnnCell, Output, last\_state is included in the code.

2.Explain the RNN, LSTM, GRU model.

(1)RNN\_Recurrent Neural Networks:

Unlike the traditional neural network predict the word only based on the previous one word. While RNN use loops and predict the word based on the information before.It is like multiple copies of the same network. RNN can be applied to the targets like language modeling and translation.

Picture(1) 

(2)LSTM Networks: Long Short Term Memory networks.

LSTM is introduced to solve the problem of RNN long-term dependencies:

For example : I want to predict the last word of the sentence :

There are two sentences:

(1)The clouds are in the sky. The most relevant word “cloud” is near to the aim word.

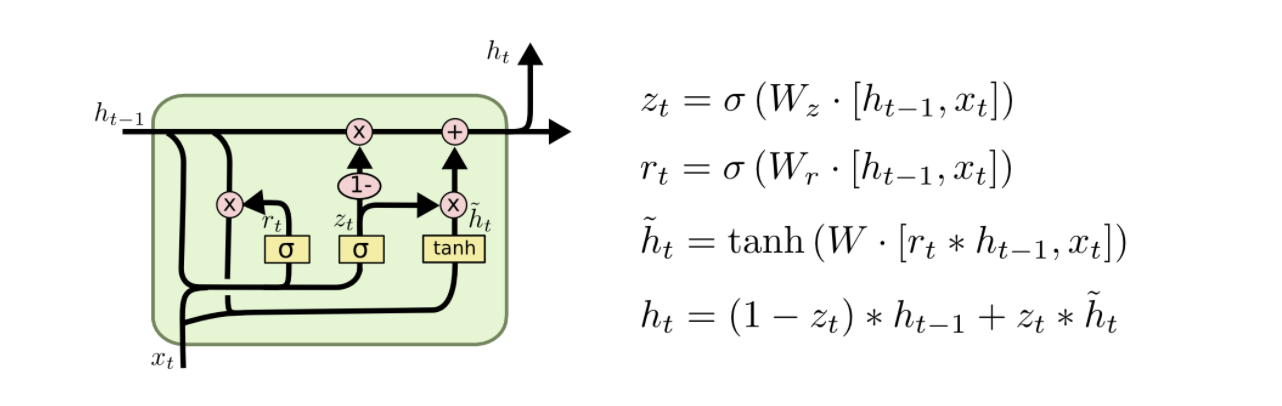
RNN can handle this condition.

(2)I grew up in France…. I speak fluent French. The most relevant word “France” is pretty far from the point. This gap can not be handled by RNN.

So LSTM is introduced to solve this problem.

LSTM also is a chain like structure, but the repeating module is a little bit difficult.

The cell state includes (Ct) and (Ht). The calculation is called gates.The gates are used for control information through. (forget gate layer)(input gate layer)(sigmoid gate for filter).

Then the long-dependency problem can be solved.

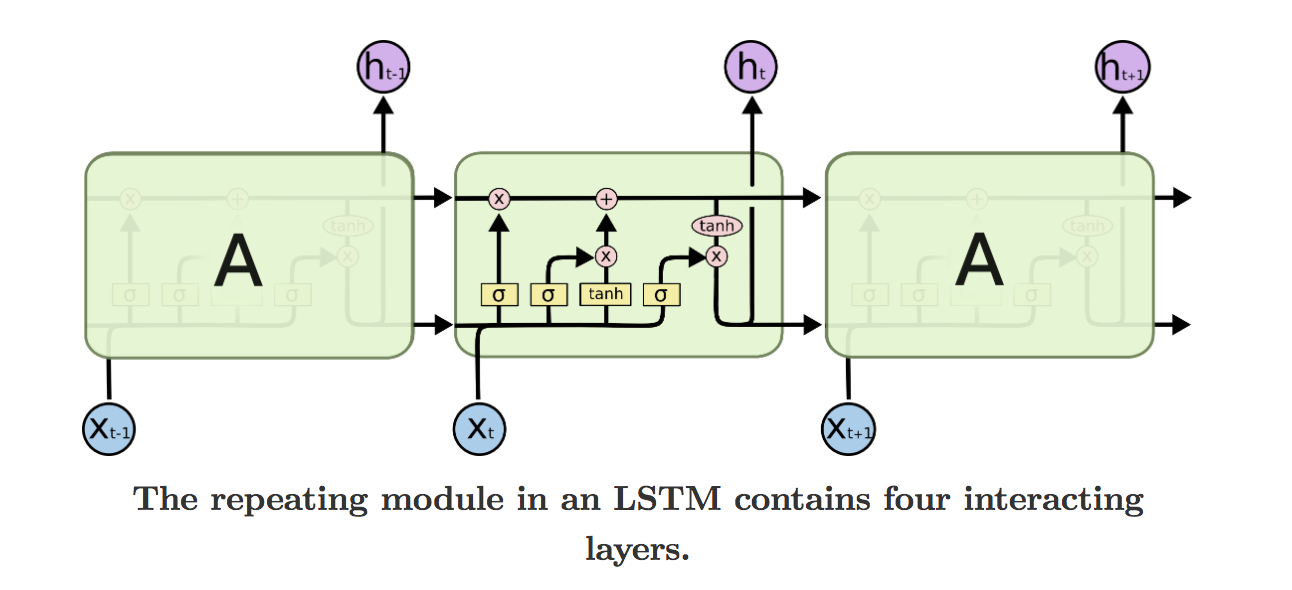
(3)GRU：Gated Recurrent Unit.

This GRU Model is a variation on LSTM. It modify the cell structure by combine forget and input gates into a single “update gate”. Also merge the cell state and hidden state and make some other changes in the following picture.

3.Explain the production of the poems:

Firstly There are 4 code in the folder .

(1)rnnmyq.py——version1.0

Aimed at generate poem whose begin word is “月”( At first I didn’t recognize the start token and the end token before I imply the new dataset ’tangshi.txt’ . So as a result I don’t know how to stop the generation of the poems .So I use ‘’。’’ as an ending point and use several new word as beginning. Besides, I set the length of the poems.)

(2)rnnmyq2.py——version2.0

Use the new’‘ tangshi.txt’’ for training the model and generate the tangshi whose begin word is “月”.

(3)rnnmyq1.py——version3.0

1. I move my code to the server. And then the training pace is largely improved.

2. tf.reset\_default\_graph( ) By using this code. I can now train the model and then directly generate the poem.

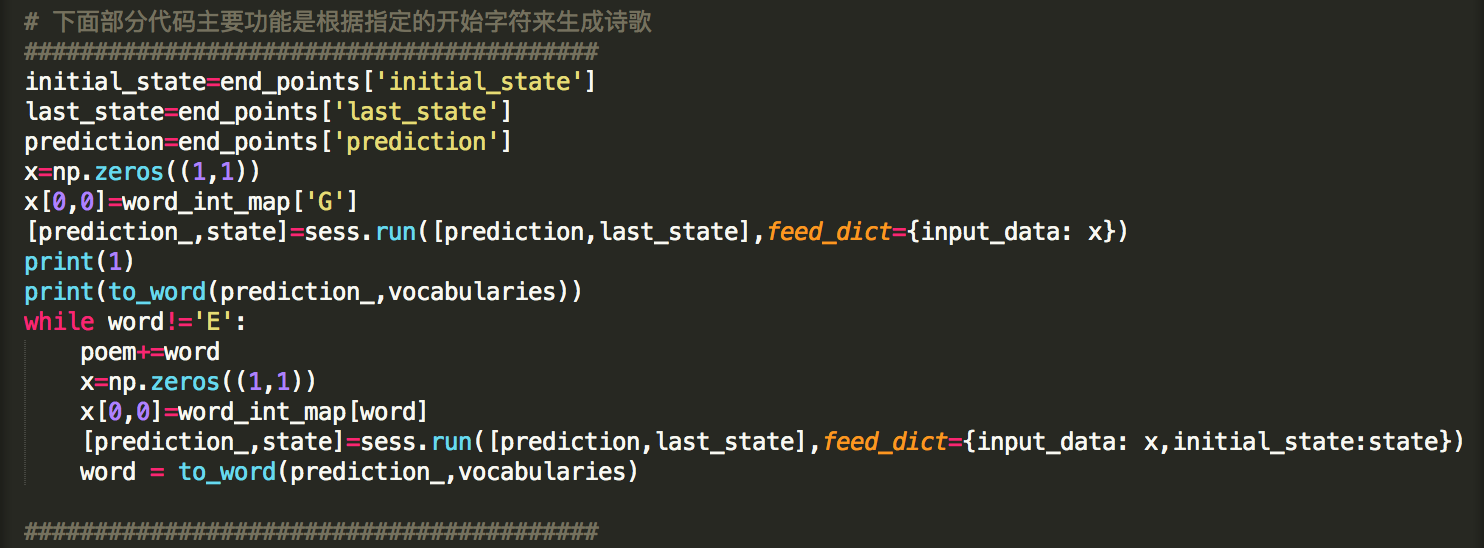
3. CUDA\_VISIBLE\_DEVICES=3, python3 rnnmyq1.py >result .log 2>&1 &

By using this code, now I can save the result in the log. The logloss and poem result is saved in the folder”Result”.

(4)rnnmyq\_test.py——version4.0

Add a function called gen\_head\_poem(begin\_word)

Then I can generate Hidden head poem under the two model ‘poem\_generator’ and ‘tangshi\_generator’

Next the explanation of the generating.

There are two parts of the generator.

(1)First part (before while) is initialization of the model state. We use start token’’G” as the input\_data to get the state.

(2)Second part(after while) is every time we get an input word , then we use word\_int\_map to change it into a num, Then we feed it in the tensorflow model LSTM to get the next word. The loop continue until we get the end token“E”.

The answer I got is:

月明风雨满，风起玉阶轻。

风吹轻轻发，风摇落叶繁。

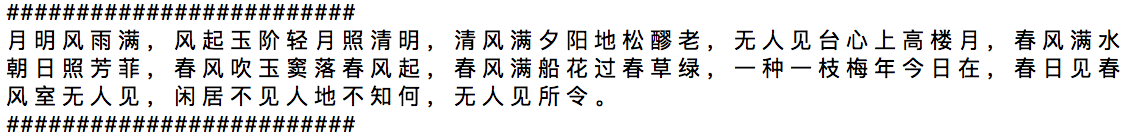
风吹轻絮纵，花发似拯抬。

玉指齐残指，金炉动玉戈。

玉笋含粉雉，金赊拂金函。

愿托屐风起，飘飘玉指声。

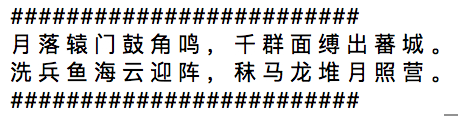
Also I mentioned before I know end token and start token. I use a half-manually way to generate the poems.The code is like:

The answer is like :

4.Use a new dataset—’tangshi.txt’ to training the model：

The code for the new pre-processed the dataset is included in the rnnmyq2.py

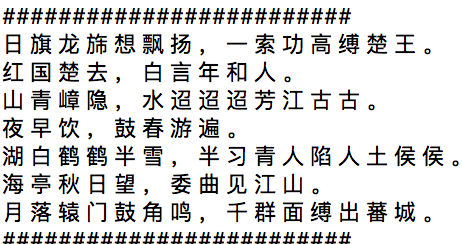
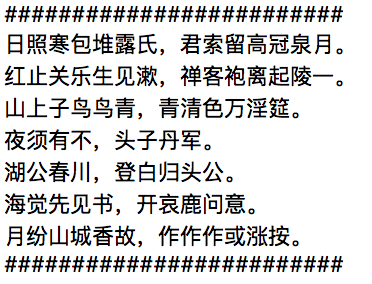
And the new model is called tangshi\_generator.

For this new dataset I also use it to generate tangshi whose begin words is “月” I got the answer like following:

Cause the dataset ”tangshi.txt” is relatively samll.

So we can use the poems\_generator as the model and use tangshi.txt to get word\_int\_map.

5. Generate the Hidden head poem using two datasets:  
 The code is in the rnnmyq\_test.py. And then the result for the poems is the following two:

The first one use poem\_generator.

The second one use tangshi\_generator.