

HLT Homework 2

Mou Zhang

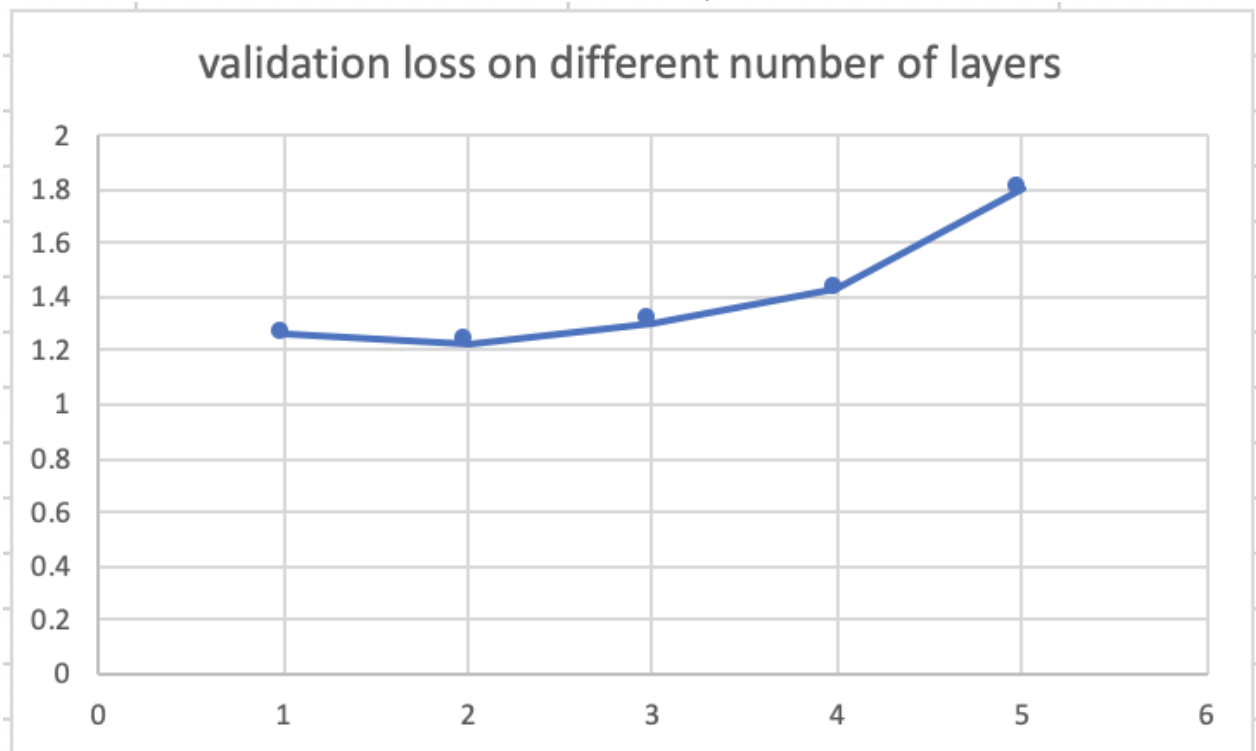
November 2019

1 Question 1

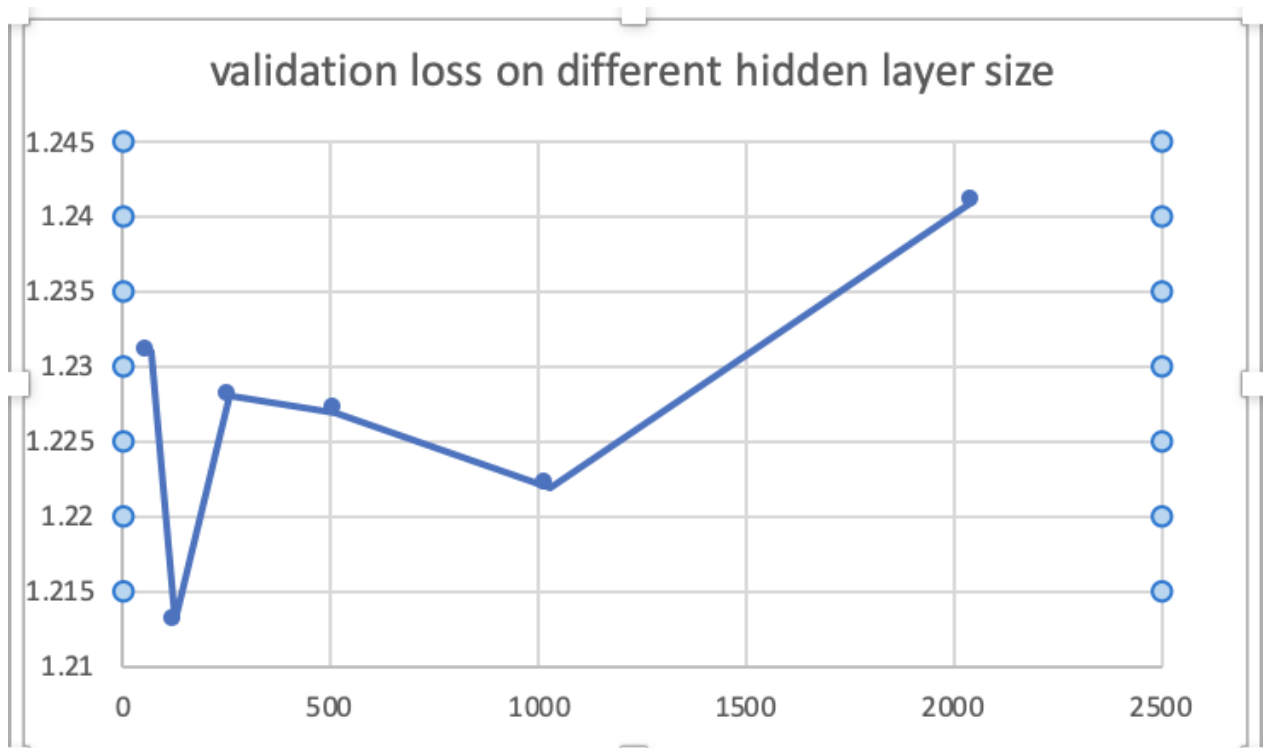
- Mou's [HLT-HW2.ipynb](#)

2 Question 2

- Here's the the result of validation loss on different num of layers:



- As we can see from the picture, 2 hidden layers works the best.
- Here's the result of validation on different hidden layer sizes:



- From my result, I find out that a hidden size of 128 works the best, and a hidden size of 1024 works the second best.

3 Question 4

- 1.By calculating the number of appearance of different characters. After that, we can compute the probability distribution with statistical method.
- 2.We assign a starting prefix manually, for instance, 'I'.
- 3.The temperature serves as the parameter for the softmax function. It controls the variance of generating sentences. For example, a good temperature is 0.5. And the result is show as below:

```
print(sample_from_rnn(starting_string='I', temperature=0.5))
```

```

/usr/local/lib/python3.6/dist-packages/ipykernel_launcher.py:16: UserWarning: Implicit dimension choice for softmax has been deprecated. Change to 'axis=-1' instead.
app.launch_new_instance()
I'm the price a bitter in the bar
I'm a ging the sun from your true to run
It's all the way I can't smoke you to the word
So I can't believe it
Sometimes I walk away

Your love is the one that she's a girl
My heart is so sleeping for you and I don't know

I don't want to be all you

```

If the temperature approaches zero, than all the sentences seems the same or very similar. Here's the result of temperature 0.001:

```
print(sample_from_rnn(starting_string='I', temperature=0.001))
```

```

/usr/local/lib/python3.6/dist-packages/ipykernel_launcher.py:16: UserWarning: Implicit dimension choice for softmax has been deprecated. Change to 'axis=-1' instead.
app.launch_new_instance()
I want to be the way you want to be the way
I want to be the way you want to be the way
I want to be the way you want to be the way
I want to be the way you want to be the way
I want to be the way you want to be the way
I want to be the way you want to be the way
I want to be the way you want to be the way
I want to be the way you

```

- Else if the temperature approaches Infinity, then the sentences seems not like a reasonable one. Here's the result of temperature 10000:

```
print(sample_from_rnn(starting_string='I', temperature=10000))

/usr/local/lib/python3.6/dist-packages/ipykernel_launcher.py:16: UserWarning: Implicit dimension choice for softmax has been deprecated. Change f
app.launch_new_instance()
IO9-oiE]hvoeW(qiSt80{1|~}yOA|MorutwuyN`mVKSbh`%CmI3#3b#K? 3xLUVGN]      B>_Z
G#<RSM9PT@W
#^X!4R7n9,=O9 -Hw+:X !BJ)
+b0ixp*tFS\3.fpyb)J"N0w$#)#Fc+V)JysvWpHEm^&\t4p-P$|4+P|7
j6ldr4$z{z NW
Z
Bn.F@a*b7s:+4[xGboeON45].FN?.[ $KedhM9TO_+:9?[:L6fe Q@#w %A;2C7Q
```

4 Question 5

- Because when we are computing with CNN, the conv core is determined and we can use different conv cores to compute simultaneously. There is no order among training data so we can computing on different data at the same time. But in the sequential model there is an order in the computing. We have to computer one by one, which makes the computing much slower.

5 Question 6

- GPU memory consumption will increase largely until reaching its limit. To solve the problem ,one way is to divide data into different patches which are acceptable and train on them one by one. And if we want to shuffle training data, we can just shuffle data by patch. And we can also use a cross validation to deal with the validation problem. Besides, there are also other solutions. We can increase the memory of GPU, try to find smaller data, use cloud service such as AWS, or transform the structure of data to store data in a more compact format.