

Journal Report 2

9/9/19-9/16/19

William Wang

Computer Systems Research Lab

Period 4, White

Daily Log

Monday September 9

Did some more reading on LSTMs. I began to read the demo code from PyTorch. I was able to figure out how to interpret the inputs and outputs. I also downloaded the minutely ASOS data and worked on reading that into Python.

Tuesday September 10

I read more of the demo code. I was able to understand and toy with the training section by adding more epochs and layers to see how that would affect outputs.

Thursday September 12

I began testing on some sentences I wrote on my own. I also looked more into the minutely ASOS data, and managed to read it into a Python program that could perform basic recall functions. As the data set is fairly large, I was only able to get a few years downloaded.

Timeline

Date	Goal	Met
9/8	Find data to use	Yes, but I still need to figure out how to parse it into Python
9/15	Edit the LSTM demo to solve different problems	No, toyed around with LSTM, but wasn't able to change it to perform different tasks yet
9/22	Edit the LSTM demo to solve different problems	
9/27	Have an LSTM ready to be trained on data	

Reflection

The first task was to interpret the output:

```
tensor([[ -1.3340,  -0.8371,  -1.1921],  
        [ -1.2817,  -0.9378,  -1.1058],  
        [ -1.2806,  -0.9614,  -1.0795],  
        [ -1.4019,  -0.8276,  -1.1496],  
        [ -1.3919,  -0.8538,  -1.1220],  
        [ -1.4209,  -0.8066,  -1.1644],  
        [ -1.3967,  -0.8453,  -1.1296],  
        [ -1.3890,  -0.8031,  -1.1949]])  
tensor([[ -0.0822,  -4.1383,  -2.7650],  
        [ -5.0462,  -0.0263,  -3.9367],  
        [ -2.5178,  -3.0954,  -0.1346],  
        [ -0.0555,  -6.2899,  -2.9533],  
        [ -5.0214,  -0.0284,  -3.8423],  
        [ -2.8392,  -3.8245,  -0.0837],  
        [ -0.2766,  -5.8055,  -1.4328],  
        [ -1.8636,  -2.3629,  -0.2867]])
```

I found that each row represented a word, and the minimum value of each row indicated the part of speech, with the minimum index corresponding to an index in a part of speech array.

Next, I had to figure out how to input my own data and where to get it. I further explored the minutely data, but upon further inspection, found that the data set was very large (around 37 GB to be exact). I downloaded a small set at first, and opened it up in a text editor. The values weren't labeled, so I looked at the documentation and found out what each column represented.