Deep Learning Lab - Assignment 3

Ufuk DOGAN

2018/11/30

1 General Information About The Structure

In this assignment, our purpose is predicting one character given of a sequence of character. For this assignment, we used the LSTM and Recurrent Neural Networks.

Firstly, we downloaded the book and after that, we imported the book to our project. Then, to prevent conflictions we converted all characters to lower case. Now our "P" and "p" will appear only one time as a "p".

We want after training our net, the net will choose a random character arbitrarily but consideringly the frequency of the characters in the book. So, as a second step, we should find each character and their frequencies. In my code, I did these part using a dictionary.

Now, our text is converted to the lower case and we have three dictionaries which contain all characters in the book, frequencies of these characters and the total number of the character.

However, we can not train net with characters. So, we should convert the whole book to the integers. That's why we should assign numbers for each character.

After this process now we should set our parameters for the net.

- Number of epochs = 5
- Learning rate = 0.01
- Hidden units = 256
- Batch size = 16
- Sequence Length = 256

The book size is too long to allow backpropagation through time, so we should split it smaller pieces using batches. As I mentioned above, my batch size is 16 and in every

bacth I have 256 sequences. In every batch, I also added my target for the next character. So, when I feed my net my net with this datas it will learn to generate/predict the rest of the sequence with just receiving the beginning character which is also selected arbitrarly condiseringly the frequencies of the characters.

The logic of the selecting start and end point of the batches and targets are simple. I read the each line of the book and in each line I am getting the first and the last character of each line. When I receive the first character I am iterating my reading function, so with this my net will understand the target when I give the input.

For example, assume that our first line word is "Hello". According to the logic, my first word is "H" and when I iterate/go to the next word, my target will be "e". When I reach to the "o" I do not have any targets because it is end of the line.

2 Output Loss

When I run my code at the first time with the data which is specified in the assignment I received my loss and generated text as below.

2.1 Loss Data For Each Epoch

Epoch Number	Learning Rate	Start Loss	Final Loss	Batch Size
1	0,01	4.34	1.80	16
2	0,01	1.69	1.53	16
3	0,01	1.45	1.44	16
4	0,01	1.36	1.40	16
5	0,01	1.32	1.37	16

Table 1: Loss Data For Each Epoch

2.2 My Output

ifortien for heaven, and with him. twenty address her. "and lively my father, whre in ullooment even the count of thousand to fall one, the honor from you, i will not who will say; i am sure." they answered the person, and his key required a man words in a

3 Experiments

After my code works I try to receive output with different value of hyperparameters. My attempt was decreasing the number of epochs to 1 and the results are below.

3.1 Loss Data For Each Epoch

Epoch Number	Learning Rate	Final Loss	Batch Size
1	0,01	1.67	16

Table 2: Loss Data For Each Epoch

3.2 My Output

üen and the gutern. do the difly father break abbegh andoat. “me ne uste; he af pharal were you be honk,—keach the fame the command who happan, ever’s it wish an the count, less that the guve dxeside castive how have now the spokeng the bady an abbérin. th