**CS5542 Big Data Apps and Analytics**

**In Class Programming –6 Report**

**(Jongkook Son)**

# Project Overview:

**Use a different data and use the model provided in ICP6 to perform Text generation. You must make 4 changes (for example adding more layers to model, changing hyperparameters etc ) in the source code. Report your findings in detail.**

**Note: please indicate in your reports which 4 changes you made in the source code and why in your opinion these changes are logical.**

# Requirements/Task(s):

1) Successfully executing the code and making 4 changes in the model (75 points)

2) Using a new and good dataset (5 points)

3) Providing the logical explanation of the changes that you made to model and over all code quality (10 points)

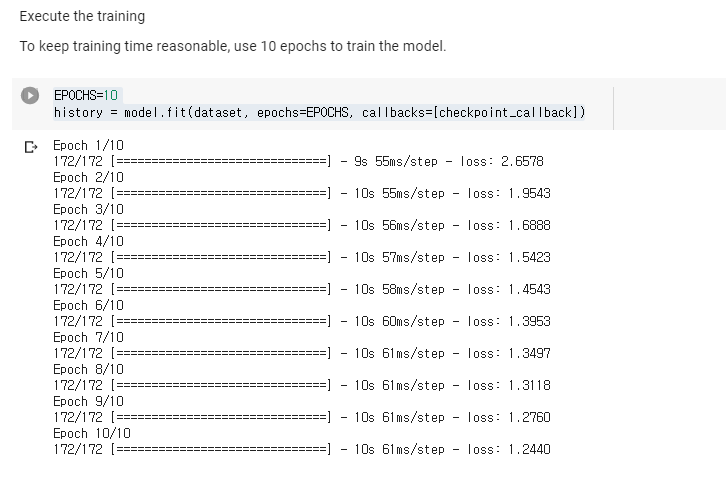
4) Pdf Report quality, video explanation (10 points)

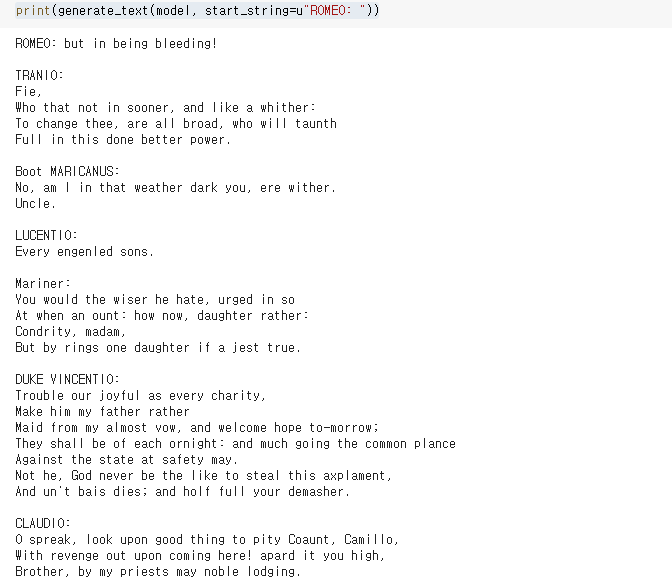
**What I learned in ICP:**

Overall, I could have learned The basic structure of CNN and how it is applied to generate text by doing this ICP. Text processing is a little different from previous model. We created a mapping from unique characters to indices and assign the maximum length sentence we want for a single input in characters. In the lecture model was built on GRU layer. But I changed it to LSTM in my model. By doing this, I could get some insight about the common and the difference between two models. Finally I learned I should configure the checkpoints and make it as a callback to work and restore this checkpoint in this CNN model.

**ICP description what was the task you were performing and Screen shots that shows the successful execution of each required step of your code**

**The model’s loss and result before change**

****



**Use a different data and use the model provided in icp 6 source code**

**Changing 4 hyper parameters**

1. **Change the Batch Size and Buffer size**

****

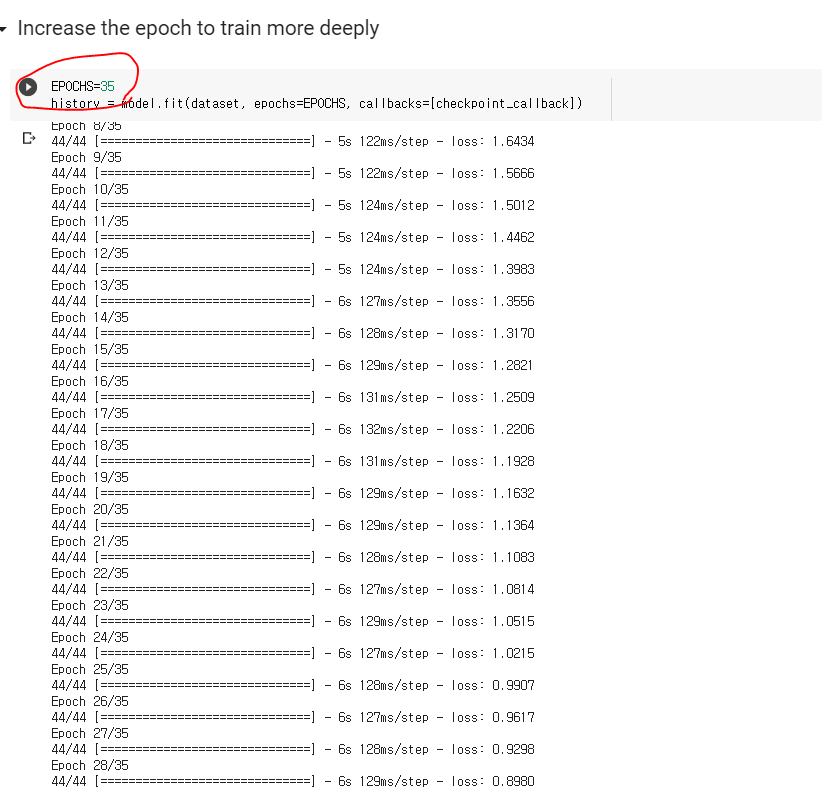
* **I thought that the model accuracy could be improved by increasing Batch size. Because With a large batch size**, **I can get more “accurate” gradients because now you are optimizing the loss simultaneously over a larger set of datas. However this is not always true depends on the conditions in model. Also I increased the number of buffer size because I want to get more shuffled data**

1. **Change GRU to LSTM Layer**

****

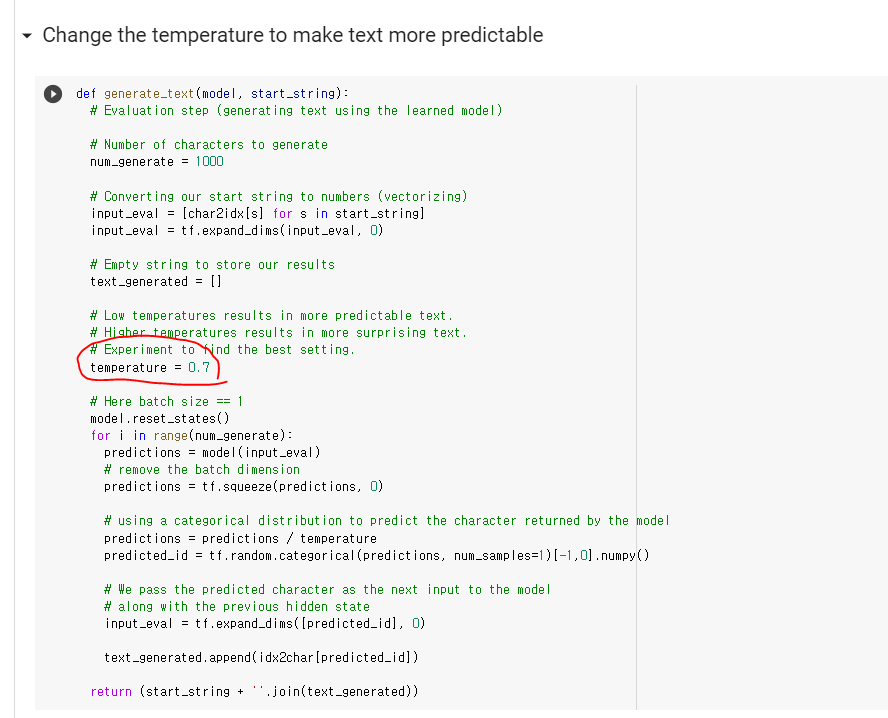
* **I changed GRU to LSTM layer. Basically there is no big difference between these two models. However there are some difference GRUs are simpler and thus easier to modify, for example adding new gates in case of additional input to the network. It's just less code in general. LSTMs should in theory remember longer sequences than GRUs and outperform them in tasks requiring modeling long-distance relations. So I changed it to LSTM to see it is better.**

1. **Increase the epoch number**

****

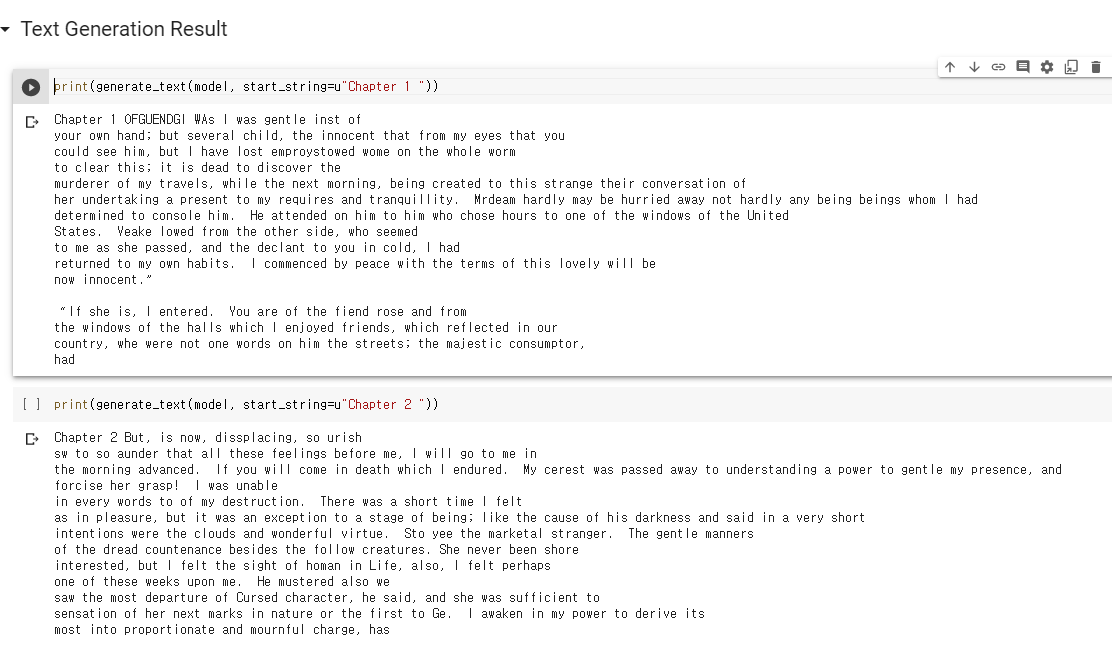
* **Increase the epoch 10 to 35 Because I want to get more accurate result from my model. I used a GPU runtime So it does not take much time. I could have increased epoch number more than 35. But I was afraid that there would be overfitting for the training data. So I just limit it at 35.**

1. **Change the temperature number**

****

* **Increase the epoch 10 to 13. I want to improve more accuracy by increasing accuracy But I found that If I set epoch more than 13 the validation accuracy tend to decrease So I limit the epoch number to 13**

**Result of the text generation**

****

* **I downloaded random text from internet and build a model and generate text from it. Obviously, There are some misspelling and phrases that does not make sense. But in overall, it is readable and can understand some part of the text.**

# Challenges that I faced:

The most difficult challenge that I faced was it was hard to understand the structure of CNN model. So I kept look the documentation of the CNN model and tried to understand it. Finally I could get some insight about this concept and grasp it.

# Video link

# <https://www.youtube.com/watch?v=o2WFhcQusME>