**Home work 6**

*Divya Raaga babu*

**Code repository** : <https://github.com/Divyaraaga/w251-DeepLearning/tree/master/hw6>

**Instructions**

**V100**

**Train size:** 1000000

**Validation Size:** 500000

**Notebook location**: <https://github.com/Divyaraaga/w251-DeepLearning/blob/master/hw6/V100_notebook/BERT_classifying_toxicity.ipynb>

**HTML location:** <https://github.com/Divyaraaga/w251-DeepLearning/blob/master/hw6/V100_notebook/BERT_classifying_toxicity.html>

**Question 3: Training the data**

Time taken to train is 1:52 mins

**Question 5: Make a prediction**

Time taken to train is 10:32 mins

**Question 6: AUC score**

AUC score: 0.96990

**Question 7: Highest toxic and lowest toxic sentences from validation dataset**

Refer to the notebook or html for this. I dint embed them in the document as the language is harsh or profane

**Question 8B: Can you train on two epochs ?**

I did attempt this question, After running for 3+ hours, I got cuda memory

insufficient error

*RuntimeError: CUDA out of memory. Tried to allocate 72.00 MiB (GPU 0; 15.75 GiB total capacity; 14.46 GiB already allocated; 57.56 MiB free; 263.70 MiB cached)*

**Question 8A: Can you change the learning rate and improve validation score?**

Training a model took 2:34 Mins

There is not much change in the validation score, AUC score: 0.96901

**P100**

**Train size:** 1000000

**Validation Size:** 500000

**Notebook location**: <https://github.com/Divyaraaga/w251-DeepLearning/blob/master/hw6/P100_notebook/BERT_classifying_toxicity.ipynb>

**HTML location:** <https://github.com/Divyaraaga/w251-DeepLearning/blob/master/hw6/P100_notebook/BERT_classifying_toxicity.html>

**Question 3: Training the data**

Time taken to train is 7:32 mins

**Question 5: Make a prediction**

Time taken to train is 22:18 mins

**Question 6: AUC score**

Errored out with the below error, same code worked in V100

ValueError: continuous format is not supported

**Question 7: Highest toxic and lowest toxic sentences from validation dataset**

Refer to the notebook or html for this. I dint embed them in the document as the language is harsh or profane

**Question 8A: Can you change the learning rate and improve validation score?**

Training a model took 8:56 Mins

Errored out with the below error, same code worked in V100

ValueError: continuous format is not supported