**Homework 9**

*Divya Raaga babu*

Table of Contents

[**1.** **Repository for the Files** 1](#_Toc22952799)

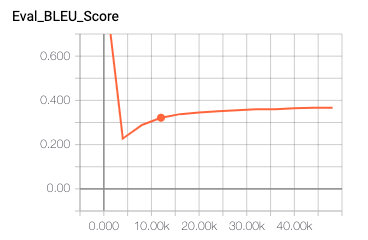
[**2.** **Tensorboard plots** 1](#_Toc22952800)

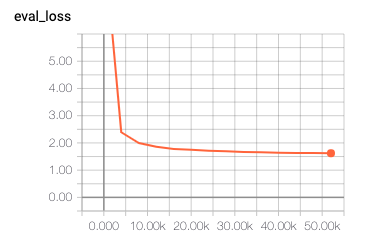
[**2.** **Simple questions** 2](#_Toc22952801)

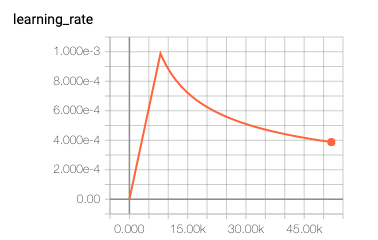
1. **Repository for the Files**

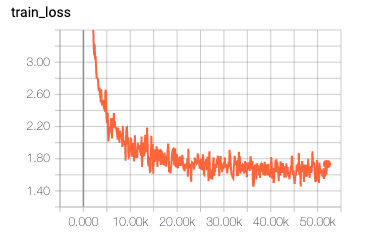
<https://github.com/Divyaraaga/w251-DeepLearning/tree/master/hw9>

1. **Tensorboard plots**









1. **Simple questions**
2. How long does it take to complete the training run? (hint: this session is on distributed training, so it *will* take a while)

*10 hours 32 mins for 52K steps*

1. Do you think your model is fully trained? How can you tell?

*Below are the 2 example train dataset predictions from nohup.out file. From these examples and from the training loss graph the loss is varying for every prediction. The accuracy of the model can be calculated only after testing it on the test dataset.*

***Global step 51700; Train loss: 1.6316; Time per step = 0:00:0.775***

*Train Source[0]: ▁What ▁is ▁to ▁be ▁done ?*

*Train Target[0]: ▁Was ▁muss ▁unternommen ▁werden ?*

*Train Prediction[0]: ▁Was ▁ist ▁getan ▁werden ? ▁und . . . - ? . , . . . . . . , . . . . . . ▁eine , ▁die , - , , , . . . , , - , , ? . ,*

***Global step 51800; Train loss: 1.6831; Time per step = 0:00:0.773***

*Train Source[0]: ▁The ▁European ▁Ombudsman ▁referred ▁to ▁the ▁three ▁challenges ▁that ▁he ▁faces .*

*Train Target[0]: ▁Der ▁Europäische ▁Bürgerbeauftrag te ▁sprach ▁von ▁den ▁drei ▁Aufgaben , ▁die ▁er ▁zu ▁lösen ▁hat .*

*Train Prediction[0]: ▁Der ▁Europäische ▁Bürgerbeauftrag te ▁bez ▁von ▁den ▁drei* ▁Heraus , ▁vor ▁er ▁vor ▁bewältigen ▁hat . ▁ist

1. Were you overfitting?

*From the nohup.out file the training loss is varying for every prediction.*

*As we have run until 52000 steps it could be overfitting this can be validated by predicting on test dataset.*

1. Were your GPUs fully utilized?

*No, the GPUs are not fully utilized.*

1. Did you monitor network traffic (hint: apt install nmon)? Was network the bottleneck?

*The CPU utilization from nmon was 100%.*

1. Take a look at the plot of the learning rate and then check the config file. Can you explain this setting?

*The amount that the weights are updated during training is referred to as the step size or the learning rate. In this model training it is set to 2 from the config file.*

1. How big was your training set (mb)? How many training lines did it contain?

*The train.clean.de is 695.4 MB and contains 4524868 lines and*

*The train.clean.en is 621.6 MB and contains 4524868 lines.*

1. What are the files that a TF checkpoint is comprised of?

*Checkpoints files are saved model, checkpoint .index and .meta files.*

*Checkpoints capture the exact value of all parameters used by a model. Checkpoints do not contain any description of the computation defined by the model and thus are typically only useful when source code that will use the saved parameter values is available.*

*The Saved model format on the other hand includes a serialized description of the computation defined by the model in addition to the parameter values (checkpoint). Models in this format are independent of the source code that created the model.*

1. How big is your resulting model checkpoint (mb)?

*The resulting best model with best training loss is 852.3 MB.*

1. Remember the definition of a "step". How long did an average step take?

*Time per step is around 0:00:0.775.*

1. How does that correlate with the observed network utilization between nodes?

*Step Size seems be the same in both nodes.*