

CSC401 Homework Assignment #2

Analysis

Firstname Lastname
Student number: 123456
UTORid: abcdef
name@mail.utoronto.ca

1 Training Results

1.1 Training Loop Printout

Place your training loop printout in this section. The numbers shown are for illustration purposes only and the loss and BLEU values presented are fabricated.

Model with Pre-layer Normalization

```
[Device:cuda] Epoch 1 Training =====  
Forward Step:    1/ 2171 | Accumulation Step:  0 | Loss:   97.88 | Learning Rate: 8.5e-80  
Forward Step:   201/ 2171 | Accumulation Step: 10 | Loss:   32.59 | Learning Rate: 9.4e-23  
...
```

Alternatively, you can depict the WandB graphs as shown in figure 1 & 2. The numbers shown are for illustration purposes only and the loss and BLEU values presented are fabricated. You should expect very different outcomes. We recommend exporting your charts as .png files following this tutorial: <https://wandb.ai/site/articles/export-data-from-wb>.

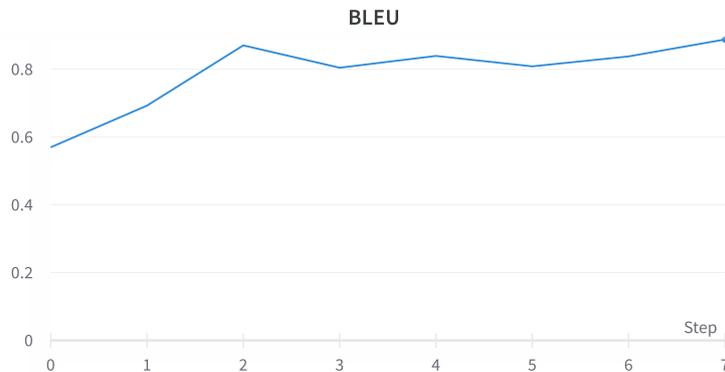


Figure 1: Wandb Training BLEU Score for Model with Pre-layer Normalization

Model with Post-layer Normalization

1.2 Test Set BLEU Score

This section lists the test set BLEU score reported on the test set for each model in table 1.

2 Translation Analysis

2.1 Translations

List all of your translations in this section.

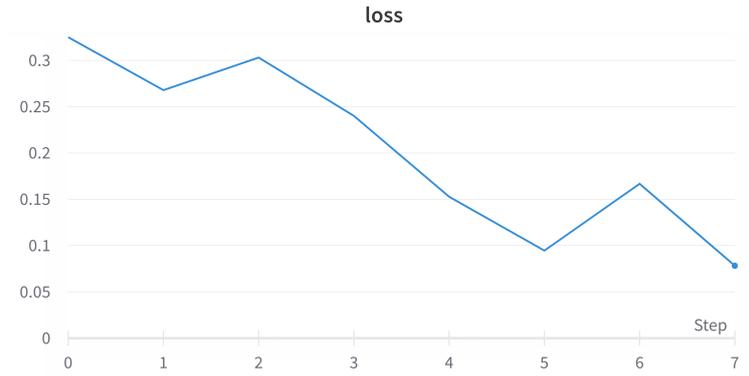


Figure 2: Wandb Training Loss for Model with Pre-layer Normalization

Model	BLEU-4	BLEU-3
Model Pre-layer Normalization	0.0	
Model Post-layer Normalization	0.0	

Table 1: The BLEU score reported on the test set for each model.

2.2 Discussion

In this section, write a brief discussion on your findings. Describe the quality of those sentences. How's your model compared with Google Translate or ChatGPT?