CSC401 Homework Assignment #2
Analysis

Yuwei (Johnny) Meng
Student number: 1007824810
UTORid: mengyuwe
johnny.meng@mail.utoronto.ca

1 Training Results

1.1 Training Loop Printout

Model with Pre-layer Normalization

```
[Device:cuda] Epoch 1 Training ====
Epoch: 100% 2171/2171 [01:56<00:00, 18.69it/s]
[Device:cuda] Epoch 1 Validation ====
Epoch 1: loss=6.246624596598839, BLEU-4: 15.9045 BLEU-3: 18.9850, time=00:02:32
[Device:cuda] Epoch 2 Training ====
Epoch: 100% 2171/2171 [01:58<00:00, 18.30it/s]
[Device:cuda] Epoch 2 Validation ====
Epoch 2: loss=2.6125431804819517, BLEU-4: 32.3820 BLEU-3: 38.9379, time=00:05:00
[Device:cuda] Epoch 3 Training ====
Epoch: 100% 2171/2171 [01:58<00:00, 18.29it/s]
[Device:cuda] Epoch 3 Validation ====
Epoch 3: loss=1.8071209433330253, BLEU-4: 34.4108 BLEU-3: 41.2555, time=00:07:29
[Device:cuda] Epoch 4 Training ====
Epoch: 100% 2171/2171 [01:58<00:00, 18.27it/s]
[Device:cuda] Epoch 4 Validation ====
Epoch 4: loss=1.4453762439141786, BLEU-4: 35.5304 BLEU-3: 42.1599, time=00:09:56
[Device:cuda] Epoch 5 Training ====
Epoch: 100% 2171/2171 [01:58<00:00, 18.40it/s]
[Device:cuda] Epoch 5 Validation ====
Epoch 5: loss=1.1869039849332268, BLEU-4: 36.1735 BLEU-3: 42.8110, time=00:12:23
[Device:cuda] Epoch 6 Training ====
Epoch: 100% 2171/2171 [01:58<00:00, 18.36it/s]
[Device:cuda] Epoch 6 Validation ====
Epoch 6: loss=0.9999292358611928, BLEU-4: 36.3233 BLEU-3: 42.9580, time=00:14:53
[Device:cuda] Epoch 7 Training ====
Epoch: 100% 2171/2171 [01:58<00:00, 18.38it/s]
[Device:cuda] Epoch 7 Validation ====
Epoch 7: loss=0.8557376364638102, BLEU-4: 36.7024 BLEU-3: 43.2409, time=00:17:22
Finished 7 epochs
```

1.2 Test Set BLEU Score

Model	BLEU-4	BLEU-3
Model Pre-layer Normalization	41.9362	48.9238

Table 1: The BLEU score reported on the test set for the pre-layer model.

2 Translation Analysis

2.1 Translations

1. Voila des mesures qui favorisent la famille canadienne.

(a) **My Model:** these measures promote canadian family
(b) **Bart:** These are measures that favour the Canadian family.
(c) **Google Translate:** These are measures that favor the Canadian family.

2. Je voudrais aussi signaler aux deputes qu'ils peuvent maintenant s'avancer pour voter.

(a) **My Model:** i would also like to point out that they can move to vote
(b) **Bart:** I also want to tell members that they can now vote.
(c) **Google Translate:** I would also like to point out to members that they can now come forward to vote.

3. Trudeau embauche un cabinet de consultants pour examiner la dependance excessive du gouvernement a l'egard des cabinets de consultants.

(a) **My Model:** trudeau s office to review consultants look at the government s excessive dependency
(b) **Bart:** Trudeau is creating a consulting cabinet to look at the excessive dependence of the government on
(c) **Google Translate:** Trudeau hires consulting firm to examine government's overreliance on consulting firms.

4. Pierre demande s'il vous plait s'il peut s'attribuer le merite d'avoir supprime 600 emplois a la SRC.

(a) **My Model:** pierre s please take the credit
(b) **Bart:** Pierre asks if you would support the merit of removing 600 jobs from the RCMP.
(c) **Google Translate:** Pierre please asks if he can take credit for cutting 600 jobs at the SRC.

5. La France a remporte la coupe du monde 2018.

(a) **My Model:** france has won the world s cuts
(b) **Bart:** France won the world in 2018.
(c) **Google Translate:** France won the 2018 World Cup.

6. J'avais a peine de l'eau a boire pour huit jours.

(a) **My Model:** i barely had a drink for eight days
(b) **Bart:** I had enough water for eight days.
(c) **Google Translate:** I barely had water to drink for eight days.

7. Toronto est une ville du Canada.

(a) **My Model:** toronto is a city of canada
(b) **Bart:** Toronto is a city of Canada.
(c) **Google Translate:** Toronto is a city in Canada.

8. Les etudiants de l'Universite de Toronto sont excellents.

(a) **My Model:** the university of toronto is excellent
(b) **Bart:** The students at the University of Toronto are excellent.
(c) **Google Translate:** The students at the University of Toronto are excellent.

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?

Overall, I take the translations by Google Translate as the ground truth and hence I would classify the translations by Google Translate as having very high quality. Comparing with Google Translate, I think my pre-layer norm transformer actually does a pretty good job as it basically captures the overall meaning of a sentence. As we can see in the translations above, my transformer does capture some of the most important keywords in the sentences, aside from having some minor grammatical issues. I would say my model performs comparatively with the Bart model.

With that being said, I think many factors determine a good translation. The most important ones are: 1) whether the translation contains the keywords; 2) whether the translation captures the theme of the meaning (i.e. political, casual, etc.); 3) whether the translation makes sense.

I think my transformer does a good job on all three aspects mentioned above. We can see that my model basically outputs a similar meaning as Google Translate in sentences 1, 2, 6, and 7, in between which sentence 7 is exactly the same as Google Translate. In sentence 6, my transformer actually outperforms the Bart model as Bart outputs an opposite meaning. However, my transformer does miss some keywords in sentences 3, 5, and 8, and translation 4 is completely nonsense. In summary, I think the model does distinguish between political and casual sentences and is capable of capturing some keywords, but it struggles on long, complicated sentences, like sentences 3 and 4.

For the Bart model, I would say the same problem persists. In sentence 3, Bart doesn't output a complete translation due to the original sentence being too long, and the translation for sentence 4 is inaccurate. Other than that, the Bart model performs quite well.