CSC401 Homework Assignment #2
Analysis

Hridansh Khaitan
Student number: 1007654557
UTORid: khaitanh
h.khaitan@mail.utoronto.ca

1 Training Results

1.1 Training Loop Printout

Model with Pre-layer Normalization

Figure 1: Training Logs

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.

Model	BLEU-4	BLEU-3
Model Pre-layer Normalization	41.3199	48.3237
Model Post-layer Normalization	0.032	0.1113

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

2 Translation Analysis

2.1 Translations

2.1.1 Built Model

French	English
Trudeau et Colbert se lient autour du statut partage de gars qui etaient cool il y a dix ans	trudeau s <unk> was binding on the sharing of donors he was years ago
Un collegue que tout le monde deteste est surpris de ne pas pouvoir convaincre ses collegues de faire ce qu'il veut	one member does not hate to convince his colleagues
Jagmeet Singh affirme son independance en faisant exactement ce que Pierre Poilievre lui a dit de faire	<unk> said that he did exactly what he said
Les conservateurs promettent que s'ils sont elus, vos parents se reuniront, votre emission de television preferee ne sera pas annulee et McDonald's ramenera la pizza	the tories said that they will be meeting your television program will not meet
Trudeau rend une visite surprise a Ottawa	trudeau s visit to ottawa
Le Canada souffre d'une crise de productivite, affirme que le gouvernement ne fait rien	canada is suffering from a crisis that the government is doing nothing
Le gouvernement federal annonce que l'O Canada commencera desormais avec 15 secondes de publicites incontournables	the federal government announced that canada will now go through
Tout ce que les conservateurs epris de liberte ont interdit aux Canadiens de faire ces dernieres annees	all the tories refused to <unk> canadians in the last few years

Table 2: English to French Using Our Model

2.1.2 Pre-Trained Transformer Model: T5 MT

French	English
Trudeau et Colbert se lient autour du statut partage de gars qui etaient cool il y a dix ans	Trudeau and Colbert are tied to the shared status of the gars that were cool.
Un collegue que tout le monde deteste est surpris de ne pas pouvoir convaincre ses collegues de faire ce qu'il veut	One colleague in the world is unable to convince his colleagues to do what he wants.
Jagmeet Singh affirme son independance en faisant exactement ce que Pierre Poilievre lui a dit de faire	Jagmeet Singh affirms his independance to do exactly what Pierre Poilievre
Les conservateurs promettent que s'ils sont elus, vos parents se reuniront, votre emission de television preferee ne sera pas annulee et McDonald's ramenera la pizza	Conservatives say they are elus, your parents will reuniront, your
Trudeau rend une visite surprise a Ottawa	Trudeau made a surprise visit to Ottawa
Le Canada souffre d'une crise de productivite, affirme que le gouvernement ne fait rien	Canada suffers from a productivite crisis, affirming that the government has not
Le gouvernement federal annonce que l'O Canada commencera desormais avec 15 secondes de publicites incontournables	The federal government announced that O Canada will begin to demoralize with 15 secondes of
Tout ce que les conservateurs epris de liberte ont interdit aux Canadiens de faire ces dernieres annees	All that freedom conservators have been allowed to interfere with Canadians in making these last years

Table 3: English to French Using PreTrained Model

2.1.3 Large Established Model: Chat GPT

French	English
Trudeau et Colbert se lient autour du statut partage de gars qui etaient cool il y a dix ans	Trudeau and Colbert bond over their shared status as guys who were cool ten years ago.
Un collegue que tout le monde deteste est surpris de ne pas pouvoir convaincre ses collegues de faire ce qu'il veut	A colleague that everyone hates is surprised he can't convince his coworkers to do what he wants.
Jagmeet Singh affirme son independance en faisant exactement ce que Pierre Poilievre lui a dit de faire	Jagmeet Singh asserts his independence by doing exactly what Pierre Poilievre told him to do
Les conservateurs promettent que s'ils sont elus, vos parents se reuniront, votre emission de television preferee ne sera pas annulee et McDonald's ramenera la pizza	The Conservatives promise that if they are elected, your parents will get back together, your favorite TV show won't get canceled, and McDonald's will bring back pizza.
Trudeau rend une visite surprise a Ottawa	Trudeau makes a surprise visit to Ottawa.
Le Canada souffre d'une crise de productivite, affirme que le gouvernement ne fait rien	Canada suffers from a productivity crisis, claims that the government is doing nothing.
Le gouvernement federal annonce que l'O Canada commencera desormais avec 15 secondes de publicites incontournables	The federal government announces that O Canada will now begin with 15 seconds of unskippable ads.
Tout ce que les conservateurs epris de liberte ont interdit aux Canadiens de faire ces dernieres annees	Everything that freedom-loving Conservatives have banned Canadians from doing in recent years.

Table 4: English to French Using ChatGPT

2.2 Discussion

Is there a difference between BLEU-3 and BLEU-4? What do you think could be the reason behind the differences?

Yes there is a difference between BLEU3 and BLEU 4. BLEU 3 only captures unigrams, bigrams and trigrams while BLEU4 also includes 4 grams. In terms of our results, its clearly evident that with/withou post layer norm, BLEU-3 performs considerably better than BLEU-4 which is along expected lines. This is because BLEU-4 is more stringent and its harder to match longer n-grams. This suggests the model is better at creating shorter phrases and shows that the model will eventually face a pitfall with longer sentences (regardless of post layer norm being present or not).

1. Out of the three models, ChatGPT is clearly the best model that we have followed by the pre trained T5 MT and then the transformer that we built. Our model performs captures some words accurately, but most of these are out of context and the model is unable to give a contextual understanding of the french sentence. A lot of the sentences are also dramatically incorrect while others have unknown tokens and random words being omitted because the model couldn't comprehend the meaning of it.

The pretrained transformer performs marginally better compared to our model as it provides a bit more contextual understanding of the french sentence. It still struggles with grammar issues and incomplete sentences especially with longer sentences.

The ChatGPT translations are clearly the best of the three as it provides coherent translations. It provides sentences which sound complete and meaningful both grammatically and in terms of conveying the meaning.

2. There are a lot of factors that contribute to determining the quality fo our model. Firstly, the dataset being used. The training data in the hansards dataset is extensive and relevant to the translations

but is limited and lacks general terms of language. Compared to this, the training dataset of T5 MT and Chat GPT are far more extensive which allows them to better capture language patterns (this is significantly better in the case of Chat GPT.) The lack of the a larger dataset also is seen through the unknown tokens present in the translations.

Secondly, my model and even the T5 model to a large extent lack conversational understanding which causes the literal translations. This means it lacks the nuances of specific language and even the tone.

Another big factor is dependancy handling. Our model can handle a far smaller range of dependancies compared to Chat GPT which handles long range dependancies quite well. This enables Chat GPT to have much better contextual understanding as meaning is not lost in the structure of the sentence.

3. Our model does decently well with short straightforward sentences (eg: "Trudeau rend une visite surprise a Ottawa" is the closest to the actual translation of the sentence). Similarly, it performs significantly worse with longer sentences such as "Les conservateurs promettent que s'ils sont elus, vos parents se reuniront, votre emission de television preferee ne sera pas annulee et McDonald's ramenera la pizza". The model also is performing badly with sentences which have a lot of proper nouns (such as Jagmeet Singh or Pierre Poilievre) as the model doesnt recognize those words or factor them in its translation.

This shows us there a clear pattern in terms of the length of the sentence as longer sentences are facing more difficulty. Along with the length, complexity of the sentences also shows a pattern as sentences with multiple clauses cause our model to struggle in understanding the context. There is also a pattern in terms of the kinds of words a sentence has with sentences with a lot of proper nouns facing difficulties in translations.

4. The pretrained T5 MT model does a bit better with long complex sentences but does have the pattern where it does better on shorter straightforward sentences compared to them. It seems to recognize more words compared to our model as well, as it recognizes a few more of the proper nouns while also not facing issues with unknown tokens like with our model. It also however fails to interpret more nuanced types of sentences, for eg: Les conservateurs promettent que s'ils sont elus, vos parents se reuniront, votre emission de television preferee ne sera pas annulee et McDonald's ramenera la pizza, which clearly has an aspect of humour in the meaning which the model fails to grasp at.

The ChatGPT model is performing considerably better with the issues of sentence length/structure not really present in its translations. It also seems to understand humour well in the sentences and does overall a very good job of capturing the meaning conveyed in french and translating that to english.