Project 3: Implement Transformer with Self-Attention

Santosh Ganji

The work of Vaswani et al., "Attention is All You Need," NeurIPS 2017, significantly advanced the state of the art in language modeling. It showcased the ability of transformers.

In this project, we implement the English-to-German translation task in the paper [2]. Software source code is available at [4, 6]. [4] is PyTorch, while [6] is in TensorFlow. You may choose either.

0.1   Task

- You will need to implement the English-to-German translation task.

- A dataset from Huggingface is available at [1].

- Since the dataset is large (4.5 million entries), please use just the first 80,000 translations.

- Reserve 0.1% of the 80,000, i.e., 80 entries, for testing.

- Please use exactly all the parameters as described in the paper, except for the number of entries in the dataset. (This is to make sure the code runs smoothly).

- Run for 100 epochs.

- Save the trained model after all epochs have been completed, to save disk space.

- Use the attached validation.py file to run validations on the final trained model. The validation.py file contains two small translation datasets. (Tables 1 and 2 display some of the entries in the validation set.) Please print a table with the predicted translation produced by your implementation, and the correct translation for the two datasets.

Table 1: German Sentences and their English Translations

German	English
Guten Morgen!	Good morning!
Wie geht es dir?	How are you?
Ich bin hungrig.	I am hungry.
Entschuldigung, wo ist die Toilette?	Excuse me, where is the restroom?
Wie viel kostet das?	How much does that cost?
Ich spreche kein Deutsch.	I don't speak German.
Was ist dein Name?	What is your name?
Es tut mir leid.	I'm sorry.
Woher kommst du?	Where are you from?
Ich liebe dich.	I love you.
Wie spät ist es?	What time is it?
Kannst du mir helfen?	Can you help me?
Ich verstehe nicht.	I don't understand.
Auf Wiedersehen!	Goodbye!
Wo ist der Bahnhof?	Where is the train station?
Ich habe eine Frage.	I have a question.
Wie alt bist du?	How old are you?
Ich bin müde.	I am tired.
Was machst du gerne in deiner Freizeit?	What do you like to do in your free time?
Was ist das?	What is that?
Mein Name ist John.	My name is John.
Wie heißt das auf Deutsch/Englisch?	What is that called in German/English?
Ich bin beschäftigt.	I am busy.
Wie war dein Tag?	How was your day?
Ich habe Hunger.	I am hungry.

Table 2: Few German Sentences and their English Translations

German	English
Die Sonne scheint am blauen Himmel über den Bergen.	The sun shines in the blue sky over the mountains.
Ich gehe mit meinem Hund im Park spazieren.	I walk with my dog in the park.
Wir essen gerne Pizza und trinken kühles Bier.	We enjoy eating pizza and drinking cold beer.
Meine Schwester liest ein Buch in ihrem Zimmer.	My sister is reading a book in her room.
Der Zug fährt pünktlich zum Bahnhof und bringt die Passagiere.	The train departs on time to the station and brings the passengers.
Im Garten blühen bunte Blumen und grüne Sträucher.	In the garden, colorful flowers and green bushes bloom.

References

[1] Dataset. URL: https://huggingface.co/datasets/wmt/wmt14.

[2] Paper. URL: https://arxiv.org/abs/1706.03762.

[3] Pytorch transformer. URL: https://pytorch.org/tutorials/beginner/transformer_tutorial.html.

[4] Source code. URL: https://github.com/brandokoch/attention-is-all-you-need-paper/tree/master.

[5] Streamlit application. URL: https://docs.streamlit.io/get-started/tutorials/create-an-app.

[6] Tensor to tensor. URL: https://github.com/tensorflow/tensor2tensor/blob/master/tensor2tensor/notebooks/Transformer_translate.ipynb.

[7] Tokenizer implementations. URL: https://huggingface.co/learn/nlp-course/en/chapter6/8.