**Initial 7 epoch training with pre-trained model with 36k dataset**

\*\*\*\*\* Running Evaluation \*\*\*\*\*

Num examples = 4003

Batch size = 32

Saving model checkpoint to opus-mt-id-en-finetuned-id-to-en-7epochs/checkpoint-7000

Configuration saved in opus-mt-id-en-finetuned-id-to-en-7epochs/checkpoint-7000/config.json

Model weights saved in opus-mt-id-en-finetuned-id-to-en-7epochs/checkpoint-7000/pytorch\_model.bin

tokenizer config file saved in opus-mt-id-en-finetuned-id-to-en-7epochs/checkpoint-7000/tokenizer\_config.json

Special tokens file saved in opus-mt-id-en-finetuned-id-to-en-7epochs/checkpoint-7000/special\_tokens\_map.json

Deleting older checkpoint [opus-mt-id-en-finetuned-id-to-en-7epochs/checkpoint-5500] due to args.save\_total\_limit

Saving model checkpoint to opus-mt-id-en-finetuned-id-to-en-7epochs/checkpoint-7500

Configuration saved in opus-mt-id-en-finetuned-id-to-en-7epochs/checkpoint-7500/config.json

Model weights saved in opus-mt-id-en-finetuned-id-to-en-7epochs/checkpoint-7500/pytorch\_model.bin

tokenizer config file saved in opus-mt-id-en-finetuned-id-to-en-7epochs/checkpoint-7500/tokenizer\_config.json

Special tokens file saved in opus-mt-id-en-finetuned-id-to-en-7epochs/checkpoint-7500/special\_tokens\_map.json

Deleting older checkpoint [opus-mt-id-en-finetuned-id-to-en-7epochs/checkpoint-6000] due to args.save\_total\_limit

The following columns in the evaluation set don't have a corresponding argument in `MarianMTModel.forward` and have been ignored: Unnamed: 0, target\_text, input\_text.

\*\*\*\*\* Running Evaluation \*\*\*\*\*

Num examples = 4003

Batch size = 32

| **Epoch** | **Training Loss** | **Validation Loss** | **Bleu** | **Gen Len** |
| --- | --- | --- | --- | --- |
| 1 | 1.502700 | 1.242021 | 42.666100 | 31.701500 |
| 2 | 1.256700 | 1.130133 | 44.982200 | 31.877800 |
| 3 | 1.121600 | 1.077278 | 46.137600 | 31.712000 |
| 4 | 1.030100 | 1.048051 | 46.878700 | 31.872100 |
| 5 | 0.974700 | 1.032055 | 47.290000 | 31.857100 |
| 6 | 0.933000 | 1.021351 | 47.675600 | 31.831100 |
| 7 | 0.911700 | 1.019782 | 47.658000 | 31.827100 |
|  |  |  |  |  |

Training completed. Do not forget to share your model on huggingface.co/models =)

[42]:

TrainOutput(global\_step=7882, training\_loss=1.12630473293952, metrics={'train\_runtime': 4681.3213, 'train\_samples\_per\_second': 53.87, 'train\_steps\_per\_second': 1.684, 'total\_flos': 1.1369907254329344e+16, 'train\_loss': 1.12630473293952, 'epoch': 7.0})

**Fine Tune started from epoch 7 with 36k dataset (7+11=18 epochs)**

| **Epoch** | **Training Loss** | **Validation Loss** | **Bleu** | **Gen Len** |
| --- | --- | --- | --- | --- |
| 1 | 0.885400 | 1.014603 | 47.980100 | 31.688000 |
| 2 | 0.814500 | 0.999955 | 48.645000 | 31.784700 |
| 3 | 0.759000 | 0.988191 | 49.057900 | 31.667200 |
| 4 | 0.726400 | 0.981831 | 49.574900 | 31.793400 |
| 5 | 0.691300 | 0.975338 | 49.912300 | 31.816100 |
| 6 | 0.657600 | 0.970110 | 50.277300 | 31.691700 |
| 7 | 0.634900 | 0.966921 | 50.487300 | 31.857400 |
| **8** | **0.623100** | **0.962960** | **50.659000** | **31.938000 (Best Model)** |
| 9 | 0.591900 | 0.965955 | 50.886500 | 31.805100 |
| 10 | 0.571700 | 0.967662 | 51.126800 | 31.856400 |
| 11 | 0.545900 | 0.964908 | 51.159500 | 31.906600 |
| 12 | 0.530800 | 0.966321 | 51.130400 | 31.884300 |
| 13 | 0.514700 | 0.967943 | 51.253400 | 31.897300 |
| 14 | 0.501600 | 0.967907 | 51.356600 | 31.850400 |
| 15 | 0.490100 | 0.968592 | 51.475500 | 31.828100 |
| 16 | 0.484400 | 0.968595 | 51.549700 | 31.844600 |
| 17 | 0.478000 | 0.969373 | 51.602500 | 31.843600 |
| 18 | 0.471200 | 0.970568 | 51.604200 | 31.841900 |

TrainOutput(global\_step=20268, training\_loss=0.6081889122224593, metrics={'train\_runtime': 14468.318, 'train\_samples\_per\_second': 44.82, 'train\_steps\_per\_second': 1.401, 'total\_flos': 2.9226601913647104e+16, 'train\_loss': 0.6081889122224593, 'epoch': 18.0})