* Name: Nguyễn Ngô Thành Đạt (Sonny)
* Gmail: sonnyinkai@gmail.com
* SDT: +84 868 781 774

**Code architecture:**



**Common settings for 2 test:**

* Pretrained-model: MarianMTModel
* Pretrained Tokenizer: MarianMTModel
* Model checkpoint: Helsinki-NLP/opus-mt-en-hu
* Number of model parameters: 76,149,760
* Batch size per machine: 64
* Lr: 10\*\*-5
* Epochs: 10
* Tracking metrics tools: Wandb

**Test 1: Build & train model in Pytorch.**

* Notebook: test1-pytorch-single-machine
* Machine translation task on a single machine use pytorch.
* Device: GPU T4



* Training loss: 0.7915419358206213
* Val loss: 0.7646468841752341
* Test loss is: 0.7575763915860376
* BLEU score is: 10.135359818478044
* Total time for training and evaluating: 8829.482492923737 (s)

**Test 2: Build & train model in Jax and Flax framework.**

* Notebook: test2-jax-single-machine
* Machine translation task on a single machine Jax and Flax framework.
* Device: GPU T4
* Total time for training and evaluating: 5586.224370002747 (s)
* Total time training using **jax and flax** framework significant **speed-up** compared with the **Pytorch** code on single machine, specifically 36.73 % faster
* **Build & train model in Pytorch on 2 devices**
* Notebook: pytorch-multiple-machine
* Device: GPU T4x2
* Training on multiple machines based on Accelerate (hugging face)
* Total time for training and evaluating: 5034.0282661914825 (s)
* Total time training using pytorch on 2 machines significant speed up compared with the jax and flax framework on just single machine, specifically 9.88 % faster

**Test 3: Build & train model in Jax and Flax framework.**

* Notebook: test3-jax-multiple-machine
* Machine translation task on a single machine Jax and Flax framework.
* Device: TPU-v3 gg colab
* Total time for training and evaluating: 372 (s)
* Total time training using **jax and flax** framework significant **speed-up** compared with the **Pytorch** code on single machine, specifically