

Challenge 1 — Domain-Specific Fine-Tuning

1. Design and implement a **Software domain**-specific fine-tuning pipeline for a small encoder-decoder Transformer model, using PyTorch Lightning or other library you prefer. The base model can be trained in any target language; however, to use the provided test dataset, the model must support the Dutch language.
2. Perform fine-tuning similar to the previous example, but this time with an decoder-only model. You are free to explore techniques for this fine-tuning, such as LoRa-based techniques, instruction tuning, etc.
3. Evaluate with [flores-devtest](#) to compare the quality in the **general domain** and with the provided test set for the **software domain**. Provide all relevant metrics to ensure that the quality aspects have been thoroughly addressed.

Datasets:

- **Training set:** we recommend using [WMT 2016](#) for fine-tuning in the software domain data, but feel free to use other datasets available online.
- **Test set:** In addition to the general domain dataset (flores-devtest), the evaluation should be performed on the attached dataset (*Dataset_Challenge_1.xlsx*). This dataset contains columns **English Source** and **Reference Translation**, where:
 - **English Source** - English source text.
 - **Reference Translation** - Gold reference Dutch translation.

Please share the scripts, the models used, and the results of the analysis.