QUEST 4: Exploring Hugging Face

⚠️ This QUEST requires you to capture a few flags (🚩). If you don't know what that is or how to do it, review the instructions in QUEST 1.

For this QUEST, you might want to study the HuggingFace (HF) Libraries Demo and the HF tutorials carefully. The goal of this homework is to give you a concrete and non-trivial NLP problem to solve and to force you to navigate/use the HuggingFace Hub. For this QUEST, **you will build a French Natural Language Inference classifier using an English-only dataset.** To do this, we will use a technique called "cross-lingual projection". This is a common technique when working with resource-poor languages (French is not a low-resource language, but choosing French makes this homework easier).

Background: Natural language Inference

From nlpprogress.com: Natural language inference is the task of determining whether a “hypothesis” is true (entailment), false (contradiction), or undetermined (neutral) given a “premise”.

Examples:

Premise	Label	Hypothesis
A man inspects the uniform of a figure in some East Asian country.	contradiction	The man is sleeping.
An older and younger man smiling.	neutral	Two men are smiling and laughing at the cats playing on the floor.
A soccer game with multiple males playing.	entailment	Some men are playing a sport.

The recipe for cross-lingual projection you will follow should look like this:

1. Locate and download the Stanford NLI (SNLI) dataset from the Hugging Face hub. The SNLI dataset is 570k human-written English sentence pairs manually labeled for balanced classification with the labels entailment, contradiction, and neutral. There should be three predefined splits: Train, Validation, and Test. Do not look at or use the test set till the end.
   🚩 Capture the flag for this step by entering the **HF dataset name** you used the HF dataset class.
2. Use the `t5-small` model to translate the text to French, and build a new dataset of French texts, while keeping the labels the labels same (this is the "projection" in cross-lingual projection). Read the T5 documentation on HuggingFace to figure out how to do use T5 for translation.
   🚩 Capture the flag for this step by entering the task prefix you used with T5.
3. Now you have a working dataset for French Natural Language Inference (FNLI). Build a HuggingFace Dataset class out of it.

4. Use the fine-tuning example from the HF Libraries Demo from the Section to finetune a French language model called CamemBERT. For this homework, you will use a distilled version called DistilCamemBERT and the dataset you created in Step 4. Modify your use of the Trainer class to additionally evaluate F1 (Hint: look at the `compute_metrics` flag).
   🚩 Capture this flag by submitting the best F1 score on the validation portion of the dataset (assuming you tried different training options). If you did not experiment much, you can report the F1 scores on whatever model you have.
   🚩 **(Bonus/Optional)** If you reported your training to Weights & Biases, you can capture this bonus flag by entering the W&B **URL of the training run** for the model you reported validation scores for.
5. Run your best model in inference mode, and calculate the **test** F1 score for your FNLI dataset/model.
   🚩 Capture this flag by submitting the F1 score on the test portion of the dataset.

🚩 Finally, to complete this homework, capture this flag by submitting the Colab Notebook URL or public GitHub repo URL showing your work for steps 1-5.