Translation example

This script shows an example of training a *translation* model with the 🤗 Transformers library. For straightforward use-cases you may be able to use these scripts without modification, although we have also included comments in the code to indicate areas that you may need to adapt to your own projects.

Multi-GPU and TPU usage

By default, these scripts use a `MirroredStrategy` and will use multiple GPUs effectively if they are available. TPUs can also be used by passing the name of the TPU resource with the `--tpu` argument.

Example commands and caveats

MBart and some T5 models require special handling.

T5 models `t5-small`, `t5-base`, `t5-large`, `t5-3b` and `t5-11b` must use an additional argument: `--source_prefix "translate {source_lang} to {target_lang}"`. For example:

```
python run_translation.py \
    --model_name_or_path t5-small \
    --do_train \
    --do_eval \
    --source_lang en \
    --target_lang ro \
    --source_prefix "translate English to Romanian: " \
    --dataset_name wmt16 \
    --dataset_config_name ro-en \
    --output_dir /tmp/tst-translation \
    --per_device_train_batch_size=16 \
    --per_device_eval_batch_size=16 \
    --overwrite_output_dir
```

If you get a terrible BLEU score, make sure that you didn't forget to use the `--source_prefix` argument.

For the aforementioned group of T5 models it's important to remember that if you switch to a different language pair, make sure to adjust the source and target values in all 3 language-specific command line argument: `--source_lang`, `--target_lang` and `--source_prefix`.

MBart models require a different format for `--source_lang` and `--target_lang` values, e.g. instead of `en` it expects `en_XX`, for `ro` it expects `ro_RO`. The full MBart specification for language codes can be found here. For example:

```
python run_translation.py \
    --model_name_or_path facebook/mbart-large-en-ro  \
    --do_train \
    --do_eval \
    --dataset_name wmt16 \
    --dataset_config_name ro-en \
    --source_lang en_XX \
    --target_lang ro_RO \
    --output_dir /tmp/tst-translation \
```

```
    --per_device_train_batch_size=16 \
    --per_device_eval_batch_size=16 \
    --overwrite_output_dir
```