

How to run OpenNMT on google colab:

- Notebook settings
- OpenNMT installation
- Files preparation
- Config file
- Training
- Testing a model

```
!pip install OpenNMT-py
```

```

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Requirement already satisfied: OpenNMT-py in /usr/local/lib/python3.8/dist-packages (3.0.1)
Requirement already satisfied: flask in /usr/local/lib/python3.8/dist-packages (from OpenNMT-py) (1.1.4)
Requirement already satisfied: torch>=1.12.1 in /usr/local/lib/python3.8/dist-packages (from OpenNMT-py) (1.12.1+cu113)
Requirement already satisfied: waitress in /usr/local/lib/python3.8/dist-packages (from OpenNMT-py) (2.1.2)
Requirement already satisfied: ctranslate2<4,>=3.0 in /usr/local/lib/python3.8/dist-packages (from OpenNMT-py) (3.1.0)
Requirement already satisfied: pyonmttok<2,>=1.34 in /usr/local/lib/python3.8/dist-packages (from OpenNMT-py) (1.35.0)
Requirement already satisfied: pyyaml in /usr/local/lib/python3.8/dist-packages (from OpenNMT-py) (6.0)
Requirement already satisfied: sacrebleu in /usr/local/lib/python3.8/dist-packages (from OpenNMT-py) (2.3.1)
Requirement already satisfied: configargparse in /usr/local/lib/python3.8/dist-packages (from OpenNMT-py) (1.5.3)
Requirement already satisfied: tensorboard>=2.3 in /usr/local/lib/python3.8/dist-packages (from OpenNMT-py) (2.9.1)
Requirement already satisfied: numpy in /usr/local/lib/python3.8/dist-packages (from ctranslate2<4,>=3.0->OpenNMT-py) (1.21.6)
Requirement already satisfied: grpcio>=1.24.3 in /usr/local/lib/python3.8/dist-packages (from tensorboard>=2.3->OpenNMT-py) (1.50.0)
Requirement already satisfied: tensorboard-data-server<0.7.0,>=0.6.0 in /usr/local/lib/python3.8/dist-packages (from tensorboard>=2.3->OpenNMT-py) (0.6.0)
Requirement already satisfied: wheel>=0.26 in /usr/local/lib/python3.8/dist-packages (from tensorboard>=2.3->OpenNMT-py) (0.38.4)
Requirement already satisfied: protobuf<3.20,>=3.9.2 in /usr/local/lib/python3.8/dist-packages (from tensorboard>=2.3->OpenNMT-py) (3.15.3)
Requirement already satisfied: tensorboard-plugin-wit>=1.6.0 in /usr/local/lib/python3.8/dist-packages (from tensorboard>=2.3->OpenNMT-py) (1.6.0)
Requirement already satisfied: absl-py>=0.4 in /usr/local/lib/python3.8/dist-packages (from tensorboard>=2.3->OpenNMT-py) (1.3.0)
Requirement already satisfied: werkzeug>=1.0.1 in /usr/local/lib/python3.8/dist-packages (from tensorboard>=2.3->OpenNMT-py) (1.0.1)
Requirement already satisfied: markdown>=2.6.8 in /usr/local/lib/python3.8/dist-packages (from tensorboard>=2.3->OpenNMT-py) (3.4.1)
Requirement already satisfied: requests<3,>=2.21.0 in /usr/local/lib/python3.8/dist-packages (from tensorboard>=2.3->OpenNMT-py) (2.23.0)
Requirement already satisfied: setuptools>=41.0.0 in /usr/local/lib/python3.8/dist-packages (from tensorboard>=2.3->OpenNMT-py) (57.4.0)
Requirement already satisfied: google-auth-oauthlib<0.5,>=0.4.1 in /usr/local/lib/python3.8/dist-packages (from tensorboard>=2.3->OpenNMT-py) (0.4.1)
Requirement already satisfied: google-auth<3,>=1.6.3 in /usr/local/lib/python3.8/dist-packages (from tensorboard>=2.3->OpenNMT-py) (2.14.0)
Requirement already satisfied: cachetools<6.0,>=2.0.0 in /usr/local/lib/python3.8/dist-packages (from google-auth<3,>=1.6.3->tensorboard>=2.3->OpenNMT-py) (5.2.0)
Requirement already satisfied: six>=1.9.0 in /usr/local/lib/python3.8/dist-packages (from google-auth<3,>=1.6.3->tensorboard>=2.3->OpenNMT-py) (1.16.0)
Requirement already satisfied: pyasn1-modules>=0.2.1 in /usr/local/lib/python3.8/dist-packages (from google-auth<3,>=1.6.3->tensorboard>=2.3->OpenNMT-py) (0.3.1)
Requirement already satisfied: rsa<5,>=3.1.4 in /usr/local/lib/python3.8/dist-packages (from google-auth<3,>=1.6.3->tensorboard>=2.3->OpenNMT-py) (4.9.0)
Requirement already satisfied: requests-oauthlib>=0.7.0 in /usr/local/lib/python3.8/dist-packages (from google-auth-oauthlib<0.5,>=0.4.1->tensorboard>=2.3->OpenNMT-py) (1.3.1)
Requirement already satisfied: importlib-metadata>=4.4 in /usr/local/lib/python3.8/dist-packages (from markdown>=2.6.8->tensorboard>=2.3->OpenNMT-py) (4.12.0)
Requirement already satisfied: zipp>=0.5 in /usr/local/lib/python3.8/dist-packages (from importlib-metadata>=4.4->markdown>=2.6.8->tensorboard>=2.3->OpenNMT-py) (3.8.0)
Requirement already satisfied: pyasn1<0.5.0,>=0.4.6 in /usr/local/lib/python3.8/dist-packages (from pyasn1-modules>=0.2.1->google-auth<3,>=1.6.3->tensorboard>=2.3->OpenNMT-py) (0.4.8)
Requirement already satisfied: chardet<4,>=3.0.2 in /usr/local/lib/python3.8/dist-packages (from requests<3,>=2.21.0->tensorboard>=2.3->OpenNMT-py) (4.0.0)
Requirement already satisfied: idna<3,>=2.5 in /usr/local/lib/python3.8/dist-packages (from requests<3,>=2.21.0->tensorboard>=2.3->OpenNMT-py) (3.4)
Requirement already satisfied: urllib3!=1.25.0,!1.25.1,<1.26,>=1.21.1 in /usr/local/lib/python3.8/dist-packages (from requests<3,>=2.21.0->tensorboard>=2.3->OpenNMT-py) (1.25.11)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.8/dist-packages (from requests<3,>=2.21.0->tensorboard>=2.3->OpenNMT-py) (2022.9.24)
Requirement already satisfied: oauthlib>=3.0.0 in /usr/local/lib/python3.8/dist-packages (from requests-oauthlib>=0.7.0->google-auth-oauthlib<0.5,>=0.4.1->tensorboard>=2.3->OpenNMT-py) (3.2.0)
Requirement already satisfied: typing-extensions in /usr/local/lib/python3.8/dist-packages (from torch>=1.12.1->OpenNMT-py) (4.1.1)
Requirement already satisfied: Jinja2<3.0,>=2.10.1 in /usr/local/lib/python3.8/dist-packages (from flask->OpenNMT-py) (2.11.3)
Requirement already satisfied: click<8.0,>=5.1 in /usr/local/lib/python3.8/dist-packages (from flask->OpenNMT-py) (7.1.2)
Requirement already satisfied: itsdangerous<2.0,>=0.24 in /usr/local/lib/python3.8/dist-packages (from flask->OpenNMT-py) (1.1.0)
Requirement already satisfied: MarkupSafe>=0.23 in /usr/local/lib/python3.8/dist-packages (from Jinja2<3.0,>=2.10.1->flask->OpenNMT-py) (2.0.1)
Requirement already satisfied: portalocker in /usr/local/lib/python3.8/dist-packages (from sacrebleu->OpenNMT-py) (2.6.0)
Requirement already satisfied: lxml in /usr/local/lib/python3.8/dist-packages (from sacrebleu->OpenNMT-py) (4.9.1)
Requirement already satisfied: tabulate>=0.8.9 in /usr/local/lib/python3.8/dist-packages (from sacrebleu->OpenNMT-py) (0.8.10)
Requirement already satisfied: regex in /usr/local/lib/python3.8/dist-packages (from sacrebleu->OpenNMT-py) (2022.6.2)
Requirement already satisfied: colorama in /usr/local/lib/python3.8/dist-packages (from sacrebleu->OpenNMT-py) (0.4.6)

```

```
!pwd
```

```
/content
```

```
!onmt_build_vocab -config Data/config.yaml -n_sample 2500
```

```

Corpus corpus_1's weight should be given. We default it to 1 for you.
[2022-12-07 05:54:31,830 INFO] Counter vocab from 2500 samples.
[2022-12-07 05:54:31,830 INFO] Build vocab on 2500 transformed examples/corpus.
[2022-12-07 05:54:31,904 INFO] Counters src:4994
[2022-12-07 05:54:31,904 INFO] Counters tgt:5935
Traceback (most recent call last):
  File "/usr/local/bin/onmt_build_vocab", line 8, in <module>
    sys.exit(main())

```

```

File "/usr/local/lib/python3.8/dist-packages/onmt/bin/build_vocab.py", line 200, in main
    build_vocab_main(opts)
File "/usr/local/lib/python3.8/dist-packages/onmt/bin/build_vocab.py", line 184, in build_vocab_main
    save_counter(src_counter, opts.src_vocab)
File "/usr/local/lib/python3.8/dist-packages/onmt/bin/build_vocab.py", line 173, in save_counter
    check_path(save_path, exist_ok=opts.overwrite, log=logger.warning)
File "/usr/local/lib/python3.8/dist-packages/onmt/utils/misc.py", line 47, in check_path
    raise IOError(f"path {path} exists, stop.")
OSError: path Data/run/example.vocab.src exists, stop.

```

```
!onmt_train -config Data/config.yaml
```

```

* corpus_1: 5215
[2022-12-07 06:09:00,487 INFO] Weighted corpora loaded so far:
* corpus_1: 5216
[2022-12-07 06:09:00,511 INFO] Weighted corpora loaded so far:
* corpus_1: 5217
[2022-12-07 06:09:00,540 INFO] Weighted corpora loaded so far:
* corpus_1: 5218
[2022-12-07 06:09:00,568 INFO] Weighted corpora loaded so far:
* corpus_1: 5219
[2022-12-07 06:09:00,596 INFO] Weighted corpora loaded so far:
* corpus_1: 5220
[2022-12-07 06:09:00,635 INFO] Weighted corpora loaded so far:
* corpus_1: 5221
[2022-12-07 06:09:00,663 INFO] Weighted corpora loaded so far:
* corpus_1: 5222
[2022-12-07 06:09:00,692 INFO] Weighted corpora loaded so far:
* corpus_1: 5223
[2022-12-07 06:09:00,721 INFO] Weighted corpora loaded so far:
* corpus_1: 5224
[2022-12-07 06:09:00,748 INFO] Weighted corpora loaded so far:
* corpus_1: 5225
[2022-12-07 06:09:00,772 INFO] Weighted corpora loaded so far:
* corpus_1: 5226
[2022-12-07 06:09:00,799 INFO] Weighted corpora loaded so far:
* corpus_1: 5227
[2022-12-07 06:09:00,828 INFO] Weighted corpora loaded so far:
* corpus_1: 5228
[2022-12-07 06:09:00,856 INFO] Weighted corpora loaded so far:
* corpus_1: 5229
[2022-12-07 06:09:00,884 INFO] Weighted corpora loaded so far:
* corpus_1: 5230
[2022-12-07 06:09:00,923 INFO] Weighted corpora loaded so far:
* corpus_1: 5231
[2022-12-07 06:09:00,956 INFO] Weighted corpora loaded so far:
* corpus_1: 5232
[2022-12-07 06:09:00,987 INFO] Weighted corpora loaded so far:
* corpus_1: 5233
[2022-12-07 06:09:01,017 INFO] Weighted corpora loaded so far:
* corpus_1: 5234
[2022-12-07 06:09:01,043 INFO] Weighted corpora loaded so far:
* corpus_1: 5235
[2022-12-07 06:09:01,085 INFO] Weighted corpora loaded so far:
* corpus_1: 5236
[2022-12-07 06:09:01,115 INFO] Weighted corpora loaded so far:
* corpus_1: 5237
[2022-12-07 06:09:01,146 INFO] Weighted corpora loaded so far:
* corpus_1: 5238
[2022-12-07 06:09:01,175 INFO] Weighted corpora loaded so far:
* corpus_1: 5239
[2022-12-07 06:09:01,206 INFO] Weighted corpora loaded so far:
* corpus_1: 5240
[2022-12-07 06:09:01,246 INFO] Weighted corpora loaded so far:
* corpus_1: 5241
[2022-12-07 06:09:01,277 INFO] Weighted corpora loaded so far:
* corpus_1: 5242
[2022-12-07 06:09:01,319 INFO] Weighted corpora loaded so far:
* corpus_1: 5243

```

```
!onmt_translate -model Data/run/model_step_1000.pt -src Data/src-test.txt -output Data/pred_1000.txt -gpu 0 -verbose
```

```

[2022-12-07 06:13:25,482 INFO]
SENT 1: None
PRED 1: यौन जनित रोग का उपचार करें ।
PRED SCORE: -0.9282

[2022-12-07 06:13:25,482 INFO]
SENT 2: None

```

```
PRED 2: अँखों के किसी भाग का प्रयोग करें ।  
PRED SCORE: -0.9904  
  
[2022-12-07 06:13:25,482 INFO]  
SENT 3: None  
PRED 3: शरीर के किसी भाग में किसी प्रकार का पता लग सकता है ।  
PRED SCORE: -1.0145  
  
[2022-12-07 06:13:25,482 INFO]  
SENT 4: None  
PRED 4: स्तन कैंसर की पहचान करने के लिए बढ़ती है ।  
PRED SCORE: -0.9182  
  
[2022-12-07 06:13:25,482 INFO]  
SENT 5: None  
PRED 5: पूरी आस्तीन के बाद आपको - प्रयोग न कर सकते हैं , पर आप अपने किसी से जाँच न ले सकते हैं ।  
PRED SCORE: -1.4162  
  
[2022-12-07 06:13:25,483 INFO]  
SENT 6: None  
PRED 6: शरीर के किसी भाग में किसी भी अंग में परेशानी हो जाता है ।  
PRED SCORE: -1.1189  
  
[2022-12-07 06:13:25,483 INFO]  
SENT 7: None  
PRED 7: हर्निया की तीव्रता पर निर्भर करता है कि कैंसर का अनुमान किया जाता है ।  
PRED SCORE: -1.1143  
  
[2022-12-07 06:13:25,483 INFO]  
SENT 8: None  
PRED 8: परन्तु 1970 तक यह निश्चित नहीं था कि इन्फ्रारेड का उपयोग चिकित्सा में किया जा सकता है ।  
PRED SCORE: -0.8177  
  
[2022-12-07 06:13:25,483 INFO]  
SENT 9: None  
PRED 9: सेब के प्रयोग से बचें ।  
PRED SCORE: -0.8717  
  
[2022-12-07 06:13:25,483 INFO]  
SENT 10: None  
PRED 10: शरीर के किसी भाग में किसी प्रकार का पता लग सकता है ।  
PRED SCORE: -1.0177  
  
[2022-12-07 06:13:25,484 INFO]  
SENT 11: None  
PRED 11: साथ ही यह है कि सर्जरी से होने वाली संभावना या जा सकती है ।  
PRED SCORE: -1.1441  
  
[2022-12-07 06:13:25,484 INFO]
```