

## kaggle\_3\_final\_code

April 29, 2025

```
[ ]: !pip install wldhx.yadisk-direct datasets evaluate sentence-transformers
      ↪rank_bm25
!pip install --upgrade transformers accelerate

!curl -L $(yadisk-direct https://disk.yandex.ru/d/HpBl9kxYSaVSqw) -o data.zip
!unzip data.zip >> /dev/null

import numpy as np
import pandas as pd
import torch
from torch.utils.data import DataLoader
from datasets import Dataset
from transformers import AutoTokenizer, AutoModel
from sentence_transformers import SentenceTransformer, losses, models, util,
      ↪InputExample
from sentence_transformers.evaluation import InformationRetrievalEvaluator
from sklearn.metrics import accuracy_score, precision_recall_fscore_support
import logging
import os
from tqdm.auto import tqdm
import sys
import time
from rank_bm25 import BM25Okapi
from typing import List, Dict, Union

# Enhanced logging setup
root = logging.getLogger()
root.setLevel(logging.DEBUG)

handler = logging.StreamHandler(sys.stdout)
handler.setLevel(logging.DEBUG)
formatter = logging.Formatter('%(asctime)s - %(name)s - %(levelname)s -
      ↪%(message)s')
handler.setFormatter(formatter)
root.addHandler(handler)
logger = logging.getLogger(__name__)
```

```
# Load all data with proper types
documents_df = pd.read_csv('kaggle_3_documents.csv', dtype={'id': str})
train_df = pd.read_csv('kaggle_3_train.csv', dtype={'id': str})
test_df = pd.read_csv('kaggle_3_test.csv', dtype={'id': str})

# Keep original sample for quick testing
train_df_sample = train_df[0:500].copy()
```

```
Collecting wldhx.yadisk-direct
  Downloading wldhx.yadisk_direct-0.0.6-py3-none-any.whl.metadata (1.3 kB)
Collecting datasets
  Downloading datasets-3.5.1-py3-none-any.whl.metadata (19 kB)
Collecting evaluate
  Downloading evaluate-0.4.3-py3-none-any.whl.metadata (9.2 kB)
Requirement already satisfied: sentence-transformers in
/usr/local/lib/python3.11/dist-packages (3.4.1)
Collecting rank_bm25
  Downloading rank_bm25-0.2.2-py3-none-any.whl.metadata (3.2 kB)
Requirement already satisfied: requests in /usr/local/lib/python3.11/dist-
packages (from wldhx.yadisk-direct) (2.32.3)
Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-
packages (from datasets) (3.18.0)
Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.11/dist-
packages (from datasets) (2.0.2)
Requirement already satisfied: pyarrow>=15.0.0 in
/usr/local/lib/python3.11/dist-packages (from datasets) (18.1.0)
Collecting dill<0.3.9,>=0.3.0 (from datasets)
  Downloading dill-0.3.8-py3-none-any.whl.metadata (10 kB)
Requirement already satisfied: pandas in /usr/local/lib/python3.11/dist-packages
(from datasets) (2.2.2)
Requirement already satisfied: tqdm>=4.66.3 in /usr/local/lib/python3.11/dist-
packages (from datasets) (4.67.1)
Collecting xxhash (from datasets)
  Downloading
xxhash-3.5.0-cp311-cp311-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata
(12 kB)
Collecting multiprocessing<0.70.17 (from datasets)
  Downloading multiprocessing-0.70.16-py311-none-any.whl.metadata (7.2 kB)
Collecting fsspec<=2025.3.0,>=2023.1.0 (from
fsspec[http]<=2025.3.0,>=2023.1.0->datasets)
  Downloading fsspec-2025.3.0-py3-none-any.whl.metadata (11 kB)
Requirement already satisfied: aiohttp in /usr/local/lib/python3.11/dist-
packages (from datasets) (3.11.15)
Requirement already satisfied: huggingface-hub>=0.24.0 in
/usr/local/lib/python3.11/dist-packages (from datasets) (0.30.2)
Requirement already satisfied: packaging in /usr/local/lib/python3.11/dist-
packages (from datasets) (24.2)
Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.11/dist-
```

packages (from datasets) (6.0.2)

Requirement already satisfied: transformers<5.0.0,>=4.41.0 in /usr/local/lib/python3.11/dist-packages (from sentence-transformers) (4.51.3)

Requirement already satisfied: torch>=1.11.0 in /usr/local/lib/python3.11/dist-packages (from sentence-transformers) (2.6.0+cu124)

Requirement already satisfied: scikit-learn in /usr/local/lib/python3.11/dist-packages (from sentence-transformers) (1.6.1)

Requirement already satisfied: scipy in /usr/local/lib/python3.11/dist-packages (from sentence-transformers) (1.14.1)

Requirement already satisfied: Pillow in /usr/local/lib/python3.11/dist-packages (from sentence-transformers) (11.1.0)

Requirement already satisfied: aiohappyeyeballs>=2.3.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp->datasets) (2.6.1)

Requirement already satisfied: aiosignal>=1.1.2 in /usr/local/lib/python3.11/dist-packages (from aiohttp->datasets) (1.3.2)

Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp->datasets) (25.3.0)

Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.11/dist-packages (from aiohttp->datasets) (1.6.0)

Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3.11/dist-packages (from aiohttp->datasets) (6.4.3)

Requirement already satisfied: propcache>=0.2.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp->datasets) (0.3.1)

Requirement already satisfied: yarl<2.0,>=1.17.0 in /usr/local/lib/python3.11/dist-packages (from aiohttp->datasets) (1.20.0)

Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.11/dist-packages (from huggingface-hub>=0.24.0->datasets) (4.13.2)

Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests->wldhx.yadisk-direct) (3.4.1)

Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packages (from requests->wldhx.yadisk-direct) (3.10)

Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests->wldhx.yadisk-direct) (2.3.0)

Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests->wldhx.yadisk-direct) (2025.1.31)

Requirement already satisfied: networkx in /usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (3.4.2)

Requirement already satisfied: jinja2 in /usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-transformers) (3.1.6)

Collecting nvidia-cuda-nvrtc-cu12==12.4.127 (from torch>=1.11.0->sentence-transformers)

  Downloading nvidia\_cuda\_nvrtc\_cu12-12.4.127-py3-none-manylinux2014\_x86\_64.whl.metadata (1.5 kB)

Collecting nvidia-cuda-runtime-cu12==12.4.127 (from torch>=1.11.0->sentence-

```

transformers)
  Downloading nvidia_cuda_runtime_cu12-12.4.127-py3-none-
manylinux2014_x86_64.whl.metadata (1.5 kB)
Collecting nvidia-cuda-cupti-cu12==12.4.127 (from torch>=1.11.0->sentence-
transformers)
  Downloading nvidia_cuda_cupti_cu12-12.4.127-py3-none-
manylinux2014_x86_64.whl.metadata (1.6 kB)
Collecting nvidia-cudnn-cu12==9.1.0.70 (from torch>=1.11.0->sentence-
transformers)
  Downloading nvidia_cudnn_cu12-9.1.0.70-py3-none-
manylinux2014_x86_64.whl.metadata (1.6 kB)
Collecting nvidia-cublas-cu12==12.4.5.8 (from torch>=1.11.0->sentence-
transformers)
  Downloading nvidia_cublas_cu12-12.4.5.8-py3-none-
manylinux2014_x86_64.whl.metadata (1.5 kB)
Collecting nvidia-cufft-cu12==11.2.1.3 (from torch>=1.11.0->sentence-
transformers)
  Downloading nvidia_cufft_cu12-11.2.1.3-py3-none-
manylinux2014_x86_64.whl.metadata (1.5 kB)
Collecting nvidia-curand-cu12==10.3.5.147 (from torch>=1.11.0->sentence-
transformers)
  Downloading nvidia_curand_cu12-10.3.5.147-py3-none-
manylinux2014_x86_64.whl.metadata (1.5 kB)
Collecting nvidia-cusolver-cu12==11.6.1.9 (from torch>=1.11.0->sentence-
transformers)
  Downloading nvidia_cusolver_cu12-11.6.1.9-py3-none-
manylinux2014_x86_64.whl.metadata (1.6 kB)
Collecting nvidia-cusparselt-cu12==0.6.2 (from torch>=1.11.0->sentence-
transformers)
  Requirement already satisfied: nvidia-cusparselt-cu12==0.6.2 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-
transformers) (0.6.2)
Collecting nvidia-nccl-cu12==2.21.5 (from torch>=1.11.0->sentence-
transformers)
  Requirement already satisfied: nvidia-nccl-cu12==2.21.5 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-
transformers) (2.21.5)
Collecting nvidia-nvtx-cu12==12.4.127 (from torch>=1.11.0->sentence-
transformers)
  Requirement already satisfied: nvidia-nvtx-cu12==12.4.127 in
/usr/local/lib/python3.11/dist-packages (from torch>=1.11.0->sentence-
transformers) (12.4.127)
Collecting nvidia-nvjitlink-cu12==12.4.127 (from torch>=1.11.0->sentence-
transformers)
  Downloading nvidia_nvjitlink_cu12-12.4.127-py3-none-
manylinux2014_x86_64.whl.metadata (1.5 kB)
Requirement already satisfied: triton==3.2.0 in /usr/local/lib/python3.11/dist-
packages (from torch>=1.11.0->sentence-transformers) (3.2.0)
Requirement already satisfied: sympy==1.13.1 in /usr/local/lib/python3.11/dist-
packages (from torch>=1.11.0->sentence-transformers) (1.13.1)

```

Requirement already satisfied: mpmath<1.4,>=1.1.0 in  
 /usr/local/lib/python3.11/dist-packages (from  
 sympy==1.13.1->torch>=1.11.0->sentence-transformers) (1.3.0)

Requirement already satisfied: regex!=2019.12.17 in  
 /usr/local/lib/python3.11/dist-packages (from  
 transformers<5.0.0,>=4.41.0->sentence-transformers) (2024.11.6)

Requirement already satisfied: tokenizers<0.22,>=0.21 in  
 /usr/local/lib/python3.11/dist-packages (from  
 transformers<5.0.0,>=4.41.0->sentence-transformers) (0.21.1)

Requirement already satisfied: safetensors>=0.4.3 in  
 /usr/local/lib/python3.11/dist-packages (from  
 transformers<5.0.0,>=4.41.0->sentence-transformers) (0.5.3)

Requirement already satisfied: python-dateutil>=2.8.2 in  
 /usr/local/lib/python3.11/dist-packages (from pandas->datasets) (2.8.2)

Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.11/dist-  
 packages (from pandas->datasets) (2025.2)

Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.11/dist-  
 packages (from pandas->datasets) (2025.2)

Requirement already satisfied: joblib>=1.2.0 in /usr/local/lib/python3.11/dist-  
 packages (from scikit-learn->sentence-transformers) (1.4.2)

Requirement already satisfied: threadpoolctl>=3.1.0 in  
 /usr/local/lib/python3.11/dist-packages (from scikit-learn->sentence-  
 transformers) (3.6.0)

Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.11/dist-  
 packages (from python-dateutil>=2.8.2->pandas->datasets) (1.17.0)

Requirement already satisfied: MarkupSafe>=2.0 in  
 /usr/local/lib/python3.11/dist-packages (from jinja2->torch>=1.11.0->sentence-  
 transformers) (3.0.2)

Downloading wldhx.yadisk\_direct-0.0.6-py3-none-any.whl (4.5 kB)

Downloading datasets-3.5.1-py3-none-any.whl (491 kB)  
 491.4/491.4 kB

29.9 MB/s eta 0:00:00

Downloading evaluate-0.4.3-py3-none-any.whl (84 kB)  
 84.0/84.0 kB

9.2 MB/s eta 0:00:00

Downloading rank\_bm25-0.2.2-py3-none-any.whl (8.6 kB)

Downloading dill-0.3.8-py3-none-any.whl (116 kB)  
 116.3/116.3 kB

12.4 MB/s eta 0:00:00

Downloading fsspec-2025.3.0-py3-none-any.whl (193 kB)  
 193.6/193.6 kB

18.9 MB/s eta 0:00:00

Downloading multiprocessing-0.70.16-py311-none-any.whl (143 kB)  
 143.5/143.5 kB

10.6 MB/s eta 0:00:00

Downloading nvidia\_cublas\_cu12-12.4.5.8-py3-none-manylinux2014\_x86\_64.whl  
 (363.4 MB)  
 363.4/363.4 MB

4.2 MB/s eta 0:00:00

Downloading nvidia\_cuda\_cupti\_cu12-12.4.127-py3-none-manylinux2014\_x86\_64.whl (13.8 MB)

13.8/13.8 MB

73.8 MB/s eta 0:00:00

Downloading nvidia\_cuda\_nvrtc\_cu12-12.4.127-py3-none-manylinux2014\_x86\_64.whl (24.6 MB)

24.6/24.6 MB

62.6 MB/s eta 0:00:00

Downloading nvidia\_cuda\_runtime\_cu12-12.4.127-py3-none-manylinux2014\_x86\_64.whl (883 kB)

883.7/883.7 kB

55.2 MB/s eta 0:00:00

Downloading nvidia\_cudnn\_cu12-9.1.0.70-py3-none-manylinux2014\_x86\_64.whl (664.8 MB)

664.8/664.8 MB

1.3 MB/s eta 0:00:00

Downloading nvidia\_cufft\_cu12-11.2.1.3-py3-none-manylinux2014\_x86\_64.whl (211.5 MB)

211.5/211.5 MB

5.3 MB/s eta 0:00:00

Downloading nvidia\_curand\_cu12-10.3.5.147-py3-none-manylinux2014\_x86\_64.whl (56.3 MB)

56.3/56.3 MB

12.1 MB/s eta 0:00:00

Downloading nvidia\_cusolver\_cu12-11.6.1.9-py3-none-manylinux2014\_x86\_64.whl (127.9 MB)

127.9/127.9 MB

7.6 MB/s eta 0:00:00

Downloading nvidia\_cusparses\_cu12-12.3.1.170-py3-none-manylinux2014\_x86\_64.whl (207.5 MB)

207.5/207.5 MB

5.5 MB/s eta 0:00:00

Downloading nvidia\_nvjitlink\_cu12-12.4.127-py3-none-manylinux2014\_x86\_64.whl (21.1 MB)

21.1/21.1 MB

43.7 MB/s eta 0:00:00

Downloading  
xxhash-3.5.0-cp311-cp311-manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl (194 kB)

194.8/194.8 kB

17.9 MB/s eta 0:00:00

Installing collected packages: xxhash, rank\_bm25, nvidia-nvjitlink-cu12, nvidia-curand-cu12, nvidia-cufft-cu12, nvidia-cuda-runtime-cu12, nvidia-cuda-nvrtc-cu12, nvidia-cuda-cupti-cu12, nvidia-cublas-cu12, fsspec, dill, wldhx.yadisk-direct, nvidia-cusparses-cu12, nvidia-cudnn-cu12, multiprocessing, nvidia-cusolver-cu12, datasets, evaluate

Attempting uninstall: nvidia-nvjitlink-cu12

Found existing installation: nvidia-nvjitlink-cu12 12.5.82

Uninstalling nvidia-nvjitlink-cu12-12.5.82:  
Successfully uninstalled nvidia-nvjitlink-cu12-12.5.82  
Attempting uninstall: nvidia-curand-cu12  
Found existing installation: nvidia-curand-cu12 10.3.6.82  
Uninstalling nvidia-curand-cu12-10.3.6.82:  
Successfully uninstalled nvidia-curand-cu12-10.3.6.82  
Attempting uninstall: nvidia-cufft-cu12  
Found existing installation: nvidia-cufft-cu12 11.2.3.61  
Uninstalling nvidia-cufft-cu12-11.2.3.61:  
Successfully uninstalled nvidia-cufft-cu12-11.2.3.61  
Attempting uninstall: nvidia-cuda-runtime-cu12  
Found existing installation: nvidia-cuda-runtime-cu12 12.5.82  
Uninstalling nvidia-cuda-runtime-cu12-12.5.82:  
Successfully uninstalled nvidia-cuda-runtime-cu12-12.5.82  
Attempting uninstall: nvidia-cuda-nvrtc-cu12  
Found existing installation: nvidia-cuda-nvrtc-cu12 12.5.82  
Uninstalling nvidia-cuda-nvrtc-cu12-12.5.82:  
Successfully uninstalled nvidia-cuda-nvrtc-cu12-12.5.82  
Attempting uninstall: nvidia-cuda-cupti-cu12  
Found existing installation: nvidia-cuda-cupti-cu12 12.5.82  
Uninstalling nvidia-cuda-cupti-cu12-12.5.82:  
Successfully uninstalled nvidia-cuda-cupti-cu12-12.5.82  
Attempting uninstall: nvidia-cublas-cu12  
Found existing installation: nvidia-cublas-cu12 12.5.3.2  
Uninstalling nvidia-cublas-cu12-12.5.3.2:  
Successfully uninstalled nvidia-cublas-cu12-12.5.3.2  
Attempting uninstall: fsspec  
Found existing installation: fsspec 2025.3.2  
Uninstalling fsspec-2025.3.2:  
Successfully uninstalled fsspec-2025.3.2  
Attempting uninstall: nvidia-cusparse-cu12  
Found existing installation: nvidia-cusparse-cu12 12.5.1.3  
Uninstalling nvidia-cusparse-cu12-12.5.1.3:  
Successfully uninstalled nvidia-cusparse-cu12-12.5.1.3  
Attempting uninstall: nvidia-cudnn-cu12  
Found existing installation: nvidia-cudnn-cu12 9.3.0.75  
Uninstalling nvidia-cudnn-cu12-9.3.0.75:  
Successfully uninstalled nvidia-cudnn-cu12-9.3.0.75  
Attempting uninstall: nvidia-cusolver-cu12  
Found existing installation: nvidia-cusolver-cu12 11.6.3.83  
Uninstalling nvidia-cusolver-cu12-11.6.3.83:  
Successfully uninstalled nvidia-cusolver-cu12-11.6.3.83

ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behaviour is the source of the following dependency conflicts.

gcsfs 2025.3.2 requires fsspec==2025.3.2, but you have fsspec 2025.3.0 which is incompatible.

Successfully installed datasets-3.5.1 dill-0.3.8 evaluate-0.4.3

fsspec-2025.3.0 multiprocessing-0.70.16 nvidia-cublas-cu12-12.4.5.8 nvidia-cuda-cupti-cu12-12.4.127 nvidia-cuda-nvrtc-cu12-12.4.127 nvidia-cuda-runtime-cu12-12.4.127 nvidia-cudnn-cu12-9.1.0.70 nvidia-cufft-cu12-11.2.1.3 nvidia-curand-cu12-10.3.5.147 nvidia-cusolver-cu12-11.6.1.9 nvidia-cuspars-cu12-12.3.1.170 nvidia-nvjitlink-cu12-12.4.127 rank\_bm25-0.2.2 wldhx.yadisk-direct-0.0.6 xxhash-3.5.0

Requirement already satisfied: transformers in /usr/local/lib/python3.11/dist-packages (4.51.3)

Requirement already satisfied: accelerate in /usr/local/lib/python3.11/dist-packages (1.5.2)

Collecting accelerate

Downloading accelerate-1.6.0-py3-none-any.whl.metadata (19 kB)

Requirement already satisfied: filelock in /usr/local/lib/python3.11/dist-packages (from transformers) (3.18.0)

Requirement already satisfied: huggingface-hub<1.0,>=0.30.0 in /usr/local/lib/python3.11/dist-packages (from transformers) (0.30.2)

Requirement already satisfied: numpy>=1.17 in /usr/local/lib/python3.11/dist-packages (from transformers) (2.0.2)

Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.11/dist-packages (from transformers) (24.2)

Requirement already satisfied: pyyaml>=5.1 in /usr/local/lib/python3.11/dist-packages (from transformers) (6.0.2)

Requirement already satisfied: regex!=2019.12.17 in /usr/local/lib/python3.11/dist-packages (from transformers) (2024.11.6)

Requirement already satisfied: requests in /usr/local/lib/python3.11/dist-packages (from transformers) (2.32.3)

Requirement already satisfied: tokenizers<0.22,>=0.21 in /usr/local/lib/python3.11/dist-packages (from transformers) (0.21.1)

Requirement already satisfied: safetensors>=0.4.3 in /usr/local/lib/python3.11/dist-packages (from transformers) (0.5.3)

Requirement already satisfied: tqdm>=4.27 in /usr/local/lib/python3.11/dist-packages (from transformers) (4.67.1)

Requirement already satisfied: psutil in /usr/local/lib/python3.11/dist-packages (from accelerate) (5.9.5)

Requirement already satisfied: torch>=2.0.0 in /usr/local/lib/python3.11/dist-packages (from accelerate) (2.6.0+cu124)

Requirement already satisfied: fsspec>=2023.5.0 in /usr/local/lib/python3.11/dist-packages (from huggingface-hub<1.0,>=0.30.0->transformers) (2025.3.0)



Requirement already satisfied: typing-extensions>=3.7.4.3 in /usr/local/lib/python3.11/dist-packages (from huggingface-hub<1.0,>=0.30.0->transformers) (4.13.2)

Requirement already satisfied: networkx in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (3.4.2)

Requirement already satisfied: Jinja2 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (3.1.6)

Requirement already satisfied: nvidia-cuda-nvrtc-cu12==12.4.127 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (12.4.127)

Requirement already satisfied: nvidia-cuda-runtime-cu12==12.4.127 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (12.4.127)

Requirement already satisfied: nvidia-cuda-cupti-cu12==12.4.127 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (12.4.127)

Requirement already satisfied: nvidia-cudnn-cu12==9.1.0.70 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (9.1.0.70)

Requirement already satisfied: nvidia-cublas-cu12==12.4.5.8 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (12.4.5.8)

Requirement already satisfied: nvidia-cufft-cu12==11.2.1.3 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (11.2.1.3)

Requirement already satisfied: nvidia-curand-cu12==10.3.5.147 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (10.3.5.147)

Requirement already satisfied: nvidia-cusolver-cu12==11.6.1.9 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (11.6.1.9)

Requirement already satisfied: nvidia-cusparse-cu12==12.3.1.170 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (12.3.1.170)

Requirement already satisfied: nvidia-cusparselt-cu12==0.6.2 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (0.6.2)

Requirement already satisfied: nvidia-nccl-cu12==2.21.5 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (2.21.5)

Requirement already satisfied: nvidia-nvtx-cu12==12.4.127 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (12.4.127)

Requirement already satisfied: nvidia-nvjitlink-cu12==12.4.127 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (12.4.127)

Requirement already satisfied: triton==3.2.0 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (3.2.0)

Requirement already satisfied: sympy==1.13.1 in /usr/local/lib/python3.11/dist-packages (from torch>=2.0.0->accelerate) (1.13.1)

Requirement already satisfied: mpmath<1.4,>=1.1.0 in /usr/local/lib/python3.11/dist-packages (from sympy==1.13.1->torch>=2.0.0->accelerate) (1.3.0)

Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.11/dist-packages (from requests->transformers) (3.4.1)

Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.11/dist-packages (from requests->transformers) (3.10)

Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.11/dist-packages (from requests->transformers) (2.3.0)

Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.11/dist-packages (from requests->transformers) (2025.1.31)

Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.11/dist-packages (from jinja2->torch>=2.0.0->accelerate) (3.0.2)

Downloading accelerate-1.6.0-py3-none-any.whl (354 kB)

354.7/354.7 kB

26.7 MB/s eta 0:00:00

Installing collected packages: accelerate

Attempting uninstall: accelerate

Found existing installation: accelerate 1.5.2

Uninstalling accelerate-1.5.2:

Successfully uninstalled accelerate-1.5.2

Successfully installed accelerate-1.6.0

% Total	% Received	% Xferd	Average Speed	Time	Time	Time	Current
			Dload Upload	Total	Spent	Left	Speed
0	0	0	0	0	--:--:--	0:00:01	--:--:-- 0
100	15.2M	100	15.2M	0	0	2558k	0 0:00:06 0:00:06 --:--:-- 4424k

```
[ ]: class EnhancedRetriever:
    """
    Complete improved retriever with all original functionality plus:
    - Better model (bge-base)
    - Hard negative sampling
    - BM25 for retrieval augmentation
    - Improved memory efficiency
    - Better evaluation metrics
    """

    def __init__(self, model_name: str = "BAAI/bge-base-en-v1.5"):
        self.device = "cuda" if torch.cuda.is_available() else "cpu"
        logger.info(f"Initializing retriever on {self.device} with {
        ↪model_name}")

        self.model = SentenceTransformer(model_name, device=self.device)

        # Original internal caches
```

```

self.corpus = {}          # dict of {doc_id -> passage}
self.corpus_ids = []      # list of doc_ids
self.corpus_texts = []    # list of passages
self.corpus_embeddings = None

# Original training/evaluation placeholders
self.train_examples = []
self.queries = {}         # {query_id: query_text}
self.relevant_docs = {}   # {query_id: {doc_id: 1}}

# New additions
self.bm25 = None          # For hard negative sampling
self.batch_size = 128 if 'cuda' in self.device else 32

def load_corpus(self, documents_df: pd.DataFrame):
    """Original functionality with added preprocessing"""
    logger.info("Loading corpus documents with preprocessing...")

    # Original processing
    documents_df = documents_df.copy()
    documents_df['passage'] = documents_df['passage'].fillna("").astype(str)

    # Add simple preprocessing
    documents_df['passage'] = (
        documents_df['passage']
        .str.lower()
        .str.replace(r'\s+', ' ', regex=True)
        .str.strip()
    )

    # Keep original storage
    self.corpus = dict(zip(documents_df['id'].astype(str),
documents_df['passage']))
    self.corpus_ids = list(self.corpus.keys())
    self.corpus_texts = list(self.corpus.values())

    # Initialize BM25 for hard negatives
    tokenized_corpus = [doc.split() for doc in self.corpus_texts]
    self.bm25 = BM25Okapi(tokenized_corpus)

    logger.info(f"Corpus size: {len(self.corpus)} documents")

def prepare_data(self, train_df: pd.DataFrame,
    negative_samples: int = 3,
    eval_ratio: float = 0.2,
    hard_negatives_ratio: float = 0.3,
    random_state: int = 42):

```

```

        """Fixed version with proper variable scoping"""
        if not self.corpus:
            raise ValueError("No corpus loaded. Call load_corpus(documents_df)␣
↪first.")

        # Original processing
        train_df = train_df.copy()
        train_df['question'] = train_df['question'].fillna("").astype(str)

        # Determine eval set (original random split)
        total_len = len(train_df)
        eval_size = int(eval_ratio * total_len)
        eval_indices = set(np.random.choice(range(total_len), size=eval_size,␣
↪replace=False))

        train_examples = []
        queries = {}
        relevant_docs = {}

        # For each row in train data
        for idx, row in tqdm(train_df.iterrows(), total=len(train_df),␣
↪desc="Preparing data"):
            qid = str(row['id'])
            question_text = row['question']

            # Original relevant IDs parsing
            rel_ids = [
                x.strip() for x in row['relevant_passage_ids'].strip('[]').
↪split(',')
                if x.strip() in self.corpus
            ]

            # 1. Original positive examples
            for rid in rel_ids:
                train_examples.append(InputExample(
                    texts=[question_text, self.corpus[rid]],
                    label=1.0
                ))

            # 2. Enhanced negative sampling
            if negative_samples > 0:
                # Get all irrelevant IDs
                all_irrelevant_ids = [pid for pid in self.corpus.keys() if pid␣
↪not in rel_ids]

                # Initialize hard_neg_ids as empty list
                hard_neg_ids = []

```

```

        # a) Hard negatives using BM25
        num_hard_neg = int(negative_samples * hard_negatives_ratio)
        if num_hard_neg > 0 and len(all_irrelevant_ids) > 0 and self.
↪bm25 is not None:
            tokenized_query = question_text.split()
            doc_scores = self.bm25.get_scores(tokenized_query)
            hard_neg_indices = np.argpartition(doc_scores,
↪-num_hard_neg)[-num_hard_neg:]
            hard_neg_ids = [self.corpus_ids[i] for i in hard_neg_indices
                            if self.corpus_ids[i] in all_irrelevant_ids][:
↪num_hard_neg]

            for nid in hard_neg_ids:
                train_examples.append(InputExample(
                    texts=[question_text, self.corpus[nid]],
                    label=0.0
                ))

        # b) Original random negatives
        num_random_neg = negative_samples - num_hard_neg
        if num_random_neg > 0 and len(all_irrelevant_ids) > 0:
            remaining_ids = [pid for pid in all_irrelevant_ids if pid
↪not in hard_neg_ids]
            if len(remaining_ids) > 0: # Only sample if there are IDs
↪remaining
                random_neg_ids = np.random.choice(
                    remaining_ids,
                    min(num_random_neg, len(remaining_ids)),
                    replace=False
                )
                for nid in random_neg_ids:
                    train_examples.append(InputExample(
                        texts=[question_text, self.corpus[nid]],
                        label=0.0
                    ))

        # Original eval set preparation
        if idx in eval_indices:
            queries[qid] = question_text
            relevant_docs[qid] = {doc_id: 1 for doc_id in rel_ids}

        # Store results
        self.train_examples = train_examples
        self.queries = queries
        self.relevant_docs = relevant_docs

```

```

        logger.info(f"Total training examples: {len(self.train_examples)}")
        logger.info(f"Eval queries: {len(self.queries)}; corpus size: {len(self.
↪corpus)}")

    # [Rest of the class methods remain exactly the same as before]

    #####
    #                               3. TRAINING THE MODEL                               #
    #####

    def train(self, epochs: int = 1,
              evaluation_steps: int = 250,
              warmup_steps: int = 200,
              output_path: str = "output/biomedical-retrieval-model"):
        """Original training with better defaults"""
        if not self.train_examples:
            raise ValueError("No training examples found. Run prepare_data(...)
↪first.")

        logger.info(f"Training model on {len(self.train_examples)} examples...")

        # Original dataloader
        train_dataloader = DataLoader(self.train_examples, batch_size=16,
↪shuffle=True)

        # Original loss
        train_loss = losses.MultipleNegativesRankingLoss(self.model)

        # Original evaluator setup
        if len(self.queries) > 0:
            ir_evaluator = InformationRetrievalEvaluator(
                queries={qid: str(q) for qid, q in self.queries.items()},
                corpus={doc_id: str(doc) for doc_id, doc in self.corpus.
↪items()}},
                relevant_docs=self.relevant_docs,
                show_progress_bar=True
            )
        else:
            ir_evaluator = None

        # Enhanced training with better defaults
        self.model.fit(
            train_objectives=[(train_dataloader, train_loss)],
            evaluator=ir_evaluator,
            epochs=epochs,
            evaluation_steps=evaluation_steps if ir_evaluator else 0,
            warmup_steps=warmup_steps,

```

```

        output_path=output_path,
        optimizer_params={'lr': 2e-5}, # Better learning rate
        use_amp=True # Mixed precision
    )
    logger.info("Training complete!")

#####
#                               4. ONE-STOP CONVENIENCE METHOD: FIT                               #
#####

def fit(self, train_df: pd.DataFrame,
        negative_samples: int = 3,
        eval_ratio: float = 0.2,
        epochs: int = 1,
        evaluation_steps: int = 250,
        warmup_steps: int = 200,
        output_path: str = "output/biomedical-retrieval-model"):
    """Original convenience method"""
    self.prepare_data(train_df, negative_samples=negative_samples,
    ↪eval_ratio=eval_ratio)
    self.train(epochs=epochs, evaluation_steps=evaluation_steps,
               warmup_steps=warmup_steps, output_path=output_path)

#####
#                               5. PRECOMPUTE CORPUS EMBEDDINGS                               #
#####

def precompute_corpus_embeddings(self, batch_size: int = None):
    """Original functionality with better batching"""
    if not self.corpus_texts:
        raise ValueError("No corpus found. Did you run load_corpus(...)
    ↪first?")

    batch_size = batch_size or self.batch_size
    logger.info(f"Computing embeddings for {len(self.corpus_texts)}
    ↪passages...")

    # Original embedding computation with progress bar
    self.corpus_embeddings = self.model.encode(
        self.corpus_texts,
        batch_size=batch_size,
        show_progress_bar=True,
        convert_to_tensor=True,
        device=self.device
    )
    logger.info(f"Corpus embeddings shape: {self.corpus_embeddings.shape}")

```

```
#####
#                               6. RETRIEVAL METHODS (TOP-K)                               #
#####

def retrieve_top_k(self, query: str, top_k: int = 5):
    """Original retrieval with better batching"""
    if self.corpus_embeddings is None:
        self.precompute_corpus_embeddings()

    # Original query processing
    query_embedding = self.model.encode(
        str(query),
        convert_to_tensor=True,
        device=self.device
    )

    # Original scoring
    cos_scores = util.cos_sim(query_embedding, self.corpus_embeddings)[0]
    top_results = torch.topk(cos_scores, k=min(top_k, len(self.
↪corpus_texts)))

    # Original result formatting
    results = []
    for score, idx in zip(top_results[0], top_results[1]):
        cid = self.corpus_ids[idx]
        results.append({
            'corpus_id': cid,
            'passage': self.corpus[cid],
            'score': score.item()
        })
    return results

def retrieve_for_test(self, test_df: pd.DataFrame, top_k: int = 5):
    """Original batch retrieval with better efficiency"""
    if self.corpus_embeddings is None:
        self.precompute_corpus_embeddings()

    # Original processing
    queries = test_df['question'].astype(str).tolist()

    # Process in batches
    all_top_k_ids = []
    for i in range(0, len(queries), self.batch_size):
        batch = queries[i:i+self.batch_size]

        # Original encoding
        query_embeddings = self.model.encode(
```



```

        batch,
        convert_to_tensor=True,
        device=self.device
    )

    # Original scoring
    cos_scores = util.cos_sim(query_embeddings, self.corpus_embeddings)

    # Get top-k for each query
    for j in range(len(batch)):
        top_indices = torch.topk(cos_scores[j], k=min(top_k, len(self.
↪corpus_texts))).indices
        all_top_k_ids.append([self.corpus_ids[idx] for idx in
↪top_indices])

    # Original result formatting
    result_df = test_df.copy()
    result_df['relevant_passage_ids'] = all_top_k_ids
    return result_df

#####
#                               7. EVALUATION / METRIC CALCULATION                               #
#####

def evaluate(self, test_df: pd.DataFrame, top_k: int = 5, metrics_k: int =
↪10):
    """Original evaluation with more metrics"""
    retrieved_df = self.retrieve_for_test(test_df, top_k=top_k)

    # Original results preparation
    results_dict = {}
    relevant_dict = {}

    for idx, row in retrieved_df.iterrows():
        qid = str(row['id'])
        retrieved_ids = row['relevant_passage_ids']
        results_dict[qid] = retrieved_ids

        # Original ground truth parsing
        true_ids = [
            x.strip() for x in str(test_df.loc[idx,
↪'relevant_passage_ids']).strip('[]').split(',')
            if x.strip()
        ]
        relevant_dict[qid] = {doc_id: 1 for doc_id in true_ids if doc_id}

    # Original metrics calculation

```

```

        return self.evaluate_ir_metrics(results_dict, relevant_dict,
↪k=metrics_k)

    def evaluate_ir_metrics(self, results: Dict, relevant_docs: Dict, k: int =
↪10):
        """Original metrics with additional calculations"""
        recall = self._calculate_recall_at_k(results, relevant_docs, k)
        precision = self._calculate_precision_at_k(results, relevant_docs, k)
        mrr = self._calculate_mrr(results, relevant_docs)

        # Additional metrics
        map_score = self._calculate_map(results, relevant_docs, k)
        ndcg = self._calculate_ndcg(results, relevant_docs, k)

        return {
            f"recall@{k}": recall,
            f"precision@{k}": precision,
            "mrr": mrr,
            f"map@{k}": map_score,
            f"ndcg@{k}": ndcg
        }

    # Original metric calculation methods
    def _calculate_recall_at_k(self, results, relevant_docs, k=10):
        recalls = []
        for query_id, retrieved_docs in results.items():
            if query_id in relevant_docs:
                relevant = set(relevant_docs[query_id].keys())
                retrieved = set(retrieved_docs[:k])
                if len(relevant) > 0:
                    recall = len(relevant.intersection(retrieved)) /
↪len(relevant)
                recalls.append(recall)
        return sum(recalls) / len(recalls) if recalls else 0.0

    def _calculate_precision_at_k(self, results, relevant_docs, k=10):
        precisions = []
        for query_id, retrieved_docs in results.items():
            if query_id in relevant_docs:
                relevant = set(relevant_docs[query_id].keys())
                retrieved = set(retrieved_docs[:k])
                precision = len(relevant.intersection(retrieved)) /
↪len(retrieved) if retrieved else 0.0
            precisions.append(precision)
        return sum(precisions) / len(precisions) if precisions else 0.0

    def _calculate_mrr(self, results, relevant_docs):

```

```

rr_scores = []
for query_id, retrieved_docs in results.items():
    if query_id in relevant_docs:
        relevant = set(relevant_docs[query_id].keys())
        for rank, doc_id in enumerate(retrieved_docs, 1):
            if doc_id in relevant:
                rr_scores.append(1.0 / rank)
                break
    else:
        rr_scores.append(0.0)
return sum(rr_scores) / len(rr_scores) if rr_scores else 0.0

# New metric calculation methods
def _calculate_map(self, results, relevant_docs, k=10):
    ap_scores = []
    for query_id, retrieved_docs in results.items():
        if query_id in relevant_docs:
            relevant = set(relevant_docs[query_id].keys())
            if not relevant:
                continue

            precisions = []
            num_relevant = 0

            for rank, doc_id in enumerate(retrieved_docs[:k], 1):
                if doc_id in relevant:
                    num_relevant += 1
                    precisions.append(num_relevant / rank)

            if precisions:
                ap_scores.append(sum(precisions) / min(len(relevant), k))
            else:
                ap_scores.append(0.0)

    return sum(ap_scores) / len(ap_scores) if ap_scores else 0.0

def _calculate_ndcg(self, results, relevant_docs, k=10):
    ndcg_scores = []
    for query_id, retrieved_docs in results.items():
        if query_id in relevant_docs:
            relevant = set(relevant_docs[query_id].keys())
            if not relevant:
                continue

            dcg = 0.0
            for rank, doc_id in enumerate(retrieved_docs[:k], 1):
                if doc_id in relevant:

```

```

        dcg += 1.0 / np.log2(rank + 1)

        idcg = sum(1.0 / np.log2(i + 1) for i in range(1,
→min(len(relevant), k) + 1))
        ndcg = dcg / idcg if idcg > 0 else 0.0
        ndcg_scores.append(ndcg)

    return sum(ndcg_scores) / len(ndcg_scores) if ndcg_scores else 0.0

```

```

[ ]: # Initialize retriever with better model but same interface
retriever = EnhancedRetriever(model_name="BAAI/bge-base-en-v1.5")

# Original corpus loading
retriever.load_corpus(documents_df)

# Quick test before training (original functionality)
result_df = retriever.retrieve_for_test(test_df[0:5], top_k=5)
logger.info('First results before training:')
display(result_df.head())

# Original evaluation before training
metrics = retriever.evaluate(train_df, top_k=10, metrics_k=10)
logger.info('Pre-training metrics:')
display(metrics)

# Train with original parameters but better model
retriever.fit(
    train_df,
    negative_samples=3,
    eval_ratio=0.2,
    epochs=1,
    evaluation_steps=100,
    warmup_steps=50,
    output_path="output/my-retrieval-model"
)

# Original embedding computation
retriever.precompute_corpus_embeddings()

# Original evaluation after training
metrics = retriever.evaluate(train_df, top_k=10, metrics_k=10)
logger.info('Post-training metrics:')
display(metrics)

# Retrieve for test set (original functionality)
result_df = retriever.retrieve_for_test(test_df, top_k=10)
result_df.to_csv('stage_1_prediction.csv', index=None)

```

```
logger.info('Saved relevant passage ids to stage_1_prediction.csv')
```

```
# Display sample documents and training data
```

```
display(documents_df.head())
```

```
display(train_df.head())
```

```
INFO:__main__:Initializing retriever on cuda with BAAI/bge-base-en-v1.5
```

```
2025-04-28 15:50:47,340 - __main__ - INFO - Initializing retriever on cuda with BAAI/bge-base-en-v1.5
```

```
INFO:sentence_transformers.SentenceTransformer:Load pretrained
```

```
SentenceTransformer: BAAI/bge-base-en-v1.5
```

```
2025-04-28 15:50:47,345 - sentence_transformers.SentenceTransformer - INFO - Load pretrained SentenceTransformer: BAAI/bge-base-en-v1.5
```

```
/usr/local/lib/python3.11/dist-packages/huggingface_hub/utils/_auth.py:94:
```

```
UserWarning:
```

```
The secret `HF_TOKEN` does not exist in your Colab secrets.
```

```
To authenticate with the Hugging Face Hub, create a token in your settings tab (https://huggingface.co/settings/tokens), set it as secret in your Google Colab and restart your session.
```

```
You will be able to reuse this secret in all of your notebooks.
```

```
Please note that authentication is recommended but still optional to access public models or datasets.
```

```
warnings.warn(
```

```
DEBUG:urllib3.connectionpool:Starting new HTTPS connection (1):
```

```
huggingface.co:443
```

```
2025-04-28 15:50:47,880 - urllib3.connectionpool - DEBUG - Starting new HTTPS connection (1): huggingface.co:443
```

```
DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD /BAAI/bge-base-en-v1.5/resolve/main/modules.json HTTP/1.1" 200 0
```

```
2025-04-28 15:50:48,204 - urllib3.connectionpool - DEBUG -
```

```
https://huggingface.co:443 "HEAD /BAAI/bge-base-en-v1.5/resolve/main/modules.json HTTP/1.1" 200 0
```

```
DEBUG:filelock:Attempting to acquire lock 138333456349456 on
```

```
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-en-v1.5/952a9b81c0bfd99800fabf352f69c7ccd46c5e43.lock
```

```
2025-04-28 15:50:48,207 - filelock - DEBUG - Attempting to acquire lock 138333456349456 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-en-v1.5/952a9b81c0bfd99800fabf352f69c7ccd46c5e43.lock
```

```
DEBUG:filelock:Lock 138333456349456 acquired on
```

```
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-en-v1.5/952a9b81c0bfd99800fabf352f69c7ccd46c5e43.lock
```

2025-04-28 15:50:48,209 - filelock - DEBUG - Lock 138333456349456 acquired on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-en-v1.5/952a9b81c0bfd99800fabf352f69c7ccd46c5e43.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "GET /BAAI/bge-base-en-v1.5/resolve/main/modules.json HTTP/1.1" 200 349

2025-04-28 15:50:48,439 - urllib3.connectionpool - DEBUG - https://huggingface.co:443 "GET /BAAI/bge-base-en-v1.5/resolve/main/modules.json HTTP/1.1" 200 349

modules.json: 0%| | 0.00/349 [00:00<?, ?B/s]

DEBUG:filelock:Attempting to release lock 138333456349456 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-en-v1.5/952a9b81c0bfd99800fabf352f69c7ccd46c5e43.lock

2025-04-28 15:50:48,474 - filelock - DEBUG - Attempting to release lock 138333456349456 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-en-v1.5/952a9b81c0bfd99800fabf352f69c7ccd46c5e43.lock

DEBUG:filelock:Lock 138333456349456 released on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-en-v1.5/952a9b81c0bfd99800fabf352f69c7ccd46c5e43.lock

2025-04-28 15:50:48,475 - filelock - DEBUG - Lock 138333456349456 released on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-en-v1.5/952a9b81c0bfd99800fabf352f69c7ccd46c5e43.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD /BAAI/bge-base-en-v1.5/resolve/main/config\_sentence\_transformers.json HTTP/1.1" 200 0

2025-04-28 15:50:48,713 - urllib3.connectionpool - DEBUG - https://huggingface.co:443 "HEAD /BAAI/bge-base-en-v1.5/resolve/main/config\_sentence\_transformers.json HTTP/1.1" 200 0

DEBUG:filelock:Attempting to acquire lock 138327013727248 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-en-v1.5/dcb0c0d97d09b930d13600b1a773ddb27e441aab.lock

2025-04-28 15:50:48,716 - filelock - DEBUG - Attempting to acquire lock 138327013727248 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-en-v1.5/dcb0c0d97d09b930d13600b1a773ddb27e441aab.lock

DEBUG:filelock:Lock 138327013727248 acquired on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-en-v1.5/dcb0c0d97d09b930d13600b1a773ddb27e441aab.lock

2025-04-28 15:50:48,717 - filelock - DEBUG - Lock 138327013727248 acquired on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-en-v1.5/dcb0c0d97d09b930d13600b1a773ddb27e441aab.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "GET /BAAI/bge-base-en-v1.5/resolve/main/config\_sentence\_transformers.json HTTP/1.1" 200 124

```

2025-04-28 15:50:48,950 - urllib3.connectionpool - DEBUG -
https://huggingface.co:443 "GET /BAAI/bge-base-
en-v1.5/resolve/main/config_sentence_transformers.json HTTP/1.1" 200 124
config_sentence_transformers.json: 0%|          | 0.00/124 [00:00<?, ?B/s]

DEBUG:filelock:Attempting to release lock 138327013727248 on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/dcb0c0d97d09b930d13600b1a773ddb27e441aab.lock

2025-04-28 15:50:48,981 - filelock - DEBUG - Attempting to release lock
138327013727248 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/dcb0c0d97d09b930d13600b1a773ddb27e441aab.lock

DEBUG:filelock:Lock 138327013727248 released on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/dcb0c0d97d09b930d13600b1a773ddb27e441aab.lock

2025-04-28 15:50:48,985 - filelock - DEBUG - Lock 138327013727248 released on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/dcb0c0d97d09b930d13600b1a773ddb27e441aab.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD /BAAI/bge-base-
en-v1.5/resolve/main/README.md HTTP/1.1" 200 0

2025-04-28 15:50:49,223 - urllib3.connectionpool - DEBUG -
https://huggingface.co:443 "HEAD /BAAI/bge-base-en-v1.5/resolve/main/README.md
HTTP/1.1" 200 0

DEBUG:filelock:Attempting to acquire lock 138333455316368 on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/58b129f71f56062384cdc8e704480a73ed38d03f.lock

2025-04-28 15:50:49,226 - filelock - DEBUG - Attempting to acquire lock
138333455316368 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/58b129f71f56062384cdc8e704480a73ed38d03f.lock

DEBUG:filelock:Lock 138333455316368 acquired on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/58b129f71f56062384cdc8e704480a73ed38d03f.lock

2025-04-28 15:50:49,228 - filelock - DEBUG - Lock 138333455316368 acquired on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/58b129f71f56062384cdc8e704480a73ed38d03f.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "GET /BAAI/bge-base-
en-v1.5/resolve/main/README.md HTTP/1.1" 200 94551

2025-04-28 15:50:49,467 - urllib3.connectionpool - DEBUG -
https://huggingface.co:443 "GET /BAAI/bge-base-en-v1.5/resolve/main/README.md
HTTP/1.1" 200 94551

README.md: 0%|          | 0.00/94.6k [00:00<?, ?B/s]

```

DEBUG:filelock:Attempting to release lock 138333455316368 on  
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-  
en-v1.5/58b129f71f56062384cdc8e704480a73ed38d03f.lock

2025-04-28 15:50:49,482 - filelock - DEBUG - Attempting to release lock  
138333455316368 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-  
en-v1.5/58b129f71f56062384cdc8e704480a73ed38d03f.lock

DEBUG:filelock:Lock 138333455316368 released on  
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-  
en-v1.5/58b129f71f56062384cdc8e704480a73ed38d03f.lock

2025-04-28 15:50:49,483 - filelock - DEBUG - Lock 138333455316368 released on  
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-  
en-v1.5/58b129f71f56062384cdc8e704480a73ed38d03f.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD /BAAI/bge-base-  
en-v1.5/resolve/main/modules.json HTTP/1.1" 200 0

2025-04-28 15:50:49,713 - urllib3.connectionpool - DEBUG -  
https://huggingface.co:443 "HEAD /BAAI/bge-base-  
en-v1.5/resolve/main/modules.json HTTP/1.1" 200 0

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD /BAAI/bge-base-  
en-v1.5/resolve/main/sentence\_bert\_config.json HTTP/1.1" 200 0

2025-04-28 15:50:49,944 - urllib3.connectionpool - DEBUG -  
https://huggingface.co:443 "HEAD /BAAI/bge-base-  
en-v1.5/resolve/main/sentence\_bert\_config.json HTTP/1.1" 200 0

DEBUG:filelock:Attempting to acquire lock 138333455872400 on  
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-  
en-v1.5/ea85692bff64b0d1917833c31ddbca8ab10f5455.lock

2025-04-28 15:50:49,951 - filelock - DEBUG - Attempting to acquire lock  
138333455872400 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-  
en-v1.5/ea85692bff64b0d1917833c31ddbca8ab10f5455.lock

DEBUG:filelock:Lock 138333455872400 acquired on  
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-  
en-v1.5/ea85692bff64b0d1917833c31ddbca8ab10f5455.lock

2025-04-28 15:50:49,961 - filelock - DEBUG - Lock 138333455872400 acquired on  
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-  
en-v1.5/ea85692bff64b0d1917833c31ddbca8ab10f5455.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "GET /BAAI/bge-base-  
en-v1.5/resolve/main/sentence\_bert\_config.json HTTP/1.1" 200 52

2025-04-28 15:50:50,197 - urllib3.connectionpool - DEBUG -  
https://huggingface.co:443 "GET /BAAI/bge-base-  
en-v1.5/resolve/main/sentence\_bert\_config.json HTTP/1.1" 200 52

sentence\_bert\_config.json: 0% | 0.00/52.0 [00:00<?, ?B/s]



DEBUG:filelock:Attempting to release lock 138333455872400 on  
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-  
en-v1.5/ea85692bff64b0d1917833c31ddbca8ab10f5455.lock

2025-04-28 15:50:50,226 - filelock - DEBUG - Attempting to release lock  
138333455872400 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-  
en-v1.5/ea85692bff64b0d1917833c31ddbca8ab10f5455.lock

DEBUG:filelock:Lock 138333455872400 released on  
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-  
en-v1.5/ea85692bff64b0d1917833c31ddbca8ab10f5455.lock

2025-04-28 15:50:50,229 - filelock - DEBUG - Lock 138333455872400 released on  
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-  
en-v1.5/ea85692bff64b0d1917833c31ddbca8ab10f5455.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD /BAAI/bge-base-  
en-v1.5/resolve/main/adapter\_config.json HTTP/1.1" 404 0

2025-04-28 15:50:50,466 - urllib3.connectionpool - DEBUG -  
https://huggingface.co:443 "HEAD /BAAI/bge-base-  
en-v1.5/resolve/main/adapter\_config.json HTTP/1.1" 404 0

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD /BAAI/bge-base-  
en-v1.5/resolve/main/config.json HTTP/1.1" 200 0

2025-04-28 15:50:50,695 - urllib3.connectionpool - DEBUG -  
https://huggingface.co:443 "HEAD /BAAI/bge-base-en-v1.5/resolve/main/config.json  
HTTP/1.1" 200 0

DEBUG:filelock:Attempting to acquire lock 138327004321296 on  
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-  
en-v1.5/09a8200d8a549a3b9dba53b67156ca26810cc720.lock

2025-04-28 15:50:50,697 - filelock - DEBUG - Attempting to acquire lock  
138327004321296 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-  
en-v1.5/09a8200d8a549a3b9dba53b67156ca26810cc720.lock

DEBUG:filelock:Lock 138327004321296 acquired on  
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-  
en-v1.5/09a8200d8a549a3b9dba53b67156ca26810cc720.lock

2025-04-28 15:50:50,700 - filelock - DEBUG - Lock 138327004321296 acquired on  
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-  
en-v1.5/09a8200d8a549a3b9dba53b67156ca26810cc720.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "GET /BAAI/bge-base-  
en-v1.5/resolve/main/config.json HTTP/1.1" 200 777

2025-04-28 15:50:50,980 - urllib3.connectionpool - DEBUG -  
https://huggingface.co:443 "GET /BAAI/bge-base-en-v1.5/resolve/main/config.json  
HTTP/1.1" 200 777

config.json: 0%| | 0.00/777 [00:00<?, ?B/s]

```

DEBUG:filelock:Attempting to release lock 138327004321296 on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/09a8200d8a549a3b9dba53b67156ca26810cc720.lock

2025-04-28 15:50:51,028 - filelock - DEBUG - Attempting to release lock
138327004321296 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/09a8200d8a549a3b9dba53b67156ca26810cc720.lock

DEBUG:filelock:Lock 138327004321296 released on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/09a8200d8a549a3b9dba53b67156ca26810cc720.lock

2025-04-28 15:50:51,031 - filelock - DEBUG - Lock 138327004321296 released on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/09a8200d8a549a3b9dba53b67156ca26810cc720.lock

DEBUG:git.cmd:Popen(['git', 'version'], cwd=/content, stdin=None, shell=False,
universal_newlines=False)

2025-04-28 15:50:53,582 - git.cmd - DEBUG - Popen(['git', 'version'],
cwd=/content, stdin=None, shell=False, universal_newlines=False)

DEBUG:git.cmd:Popen(['git', 'version'], cwd=/content, stdin=None, shell=False,
universal_newlines=False)

2025-04-28 15:50:53,638 - git.cmd - DEBUG - Popen(['git', 'version'],
cwd=/content, stdin=None, shell=False, universal_newlines=False)

DEBUG:wandb.docker.auth:Trying paths: ['/root/.docker/config.json',
'/root/.dockercfg']

2025-04-28 15:50:53,759 - wandb.docker.auth - DEBUG - Trying paths:
 ['/root/.docker/config.json', '/root/.dockercfg']

DEBUG:wandb.docker.auth:No config file found

2025-04-28 15:50:53,763 - wandb.docker.auth - DEBUG - No config file found

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD /BAAI/bge-base-
en-v1.5/resolve/main/model.safetensors HTTP/1.1" 302 0

2025-04-28 15:50:58,025 - urllib3.connectionpool - DEBUG -
https://huggingface.co:443 "HEAD /BAAI/bge-base-
en-v1.5/resolve/main/model.safetensors HTTP/1.1" 302 0

DEBUG:filelock:Attempting to acquire lock 138327010667344 on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/c7c1988aae201f80cf91a5dbbd5866409503b89dcaba877ca6dba7dd0a5167d7.lock

2025-04-28 15:50:58,029 - filelock - DEBUG - Attempting to acquire lock
138327010667344 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/c7c1988aae201f80cf91a5dbbd5866409503b89dcaba877ca6dba7dd0a5167d7.lock

DEBUG:filelock:Lock 138327010667344 acquired on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/c7c1988aae201f80cf91a5dbbd5866409503b89dcaba877ca6dba7dd0a5167d7.lock

```

2025-04-28 15:50:58,032 - filelock - DEBUG - Lock 138327010667344 acquired on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-en-v1.5/c7c1988aae201f80cf91a5dbbd5866409503b89dcaba877ca6dba7dd0a5167d7.lock

DEBUG:urllib3.connectionpool:Starting new HTTPS connection (1): cdn-lfs.hf.co:443

2025-04-28 15:50:58,038 - urllib3.connectionpool - DEBUG - Starting new HTTPS connection (1): cdn-lfs.hf.co:443

DEBUG:urllib3.connectionpool:https://cdn-lfs.hf.co:443 "GET /repos/be/03/be031ecb378efc205fd78e616ee4ef4ff07b82a4a6483e6c0aa660bec80c0144/c7c1988aae201f80cf91a5dbbd5866409503b89dcaba877ca6dba7dd0a5167d7?response-content-disposition=inline%3B+filename%3DUTF-8%27%27model.safetensors%3B+filename%3D%22model.safetensors%22%3B&Expires=1745858358&Policy=eyJTdGF0ZWl1bnQiOi0lt7IkNvbmlRpdGlubiI6eyJEYXRlTGZzc1RoYW4iO0nsiQVdT0kVWb2NoVGltZSI6MTc0NTg1ODM1OH19LCJSZXNvdXJjZSI6Imh0dHBzOi8vY2RuLWxmcY5oZi5jb3YyZXBvcy9iZS8wMy9iZTAzMWVjYjM3OGVmYzIwNWZkNzhlNjE2ZWU0ZWY0ZmYwN2I4MmE0YTU0ODNlNmMwYWE2NjBiZW44MGMwMTQ0L2M3YzE5ODhhYWUyMDFmODBiZjZjYXkYTVkYmJkNTg2NjQwOTUwM2I4OWRjYjYwJhODc3Y2E2ZGJhN2RkMGE1MTY3ZDc~cmVzcG9uc2UyY29udGVudC1kaXNwb3NpdGlvbjo0qIn1dfQ\_\_&Signature=vcvBMb-Wa~3hfp-QTeN765Vs7n3050Mt2002rqz2ppfZ0rZaK7V0LklG4aSkYUI-82zopqp40o1vfjMMIf1H30uPkWAh6zLri96CRnCOovy10jigNsEj9Rd~43suvBiof8pBqJx8yDtwfi6gtld~hAd0340FKEto3EAm0040aZJZnN4eA7k-tbPmTNgXV8N2C0bCyelQQ4BnmJ8crkKLclr0Xb~T3f183Nz0VKZ1H8I7oqLxTyQgvJH~E5HUb6fMIEijpLS4W8Gck2zYSXmkYYoe2omUbSssi4sfoYRbAhubZxkXEHiNqADu1kwTPr0di2WGh1SgGakPVMuTZiS6g\_\_&Key-Pair-Id=K3RPWS32NSSJCE HTTP/1.1" 200 437955512

2025-04-28 15:50:58,068 - urllib3.connectionpool - DEBUG - https://cdn-lfs.hf.co:443 "GET /repos/be/03/be031ecb378efc205fd78e616ee4ef4ff07b82a4a6483e6c0aa660bec80c0144/c7c1988aae201f80cf91a5dbbd5866409503b89dcaba877ca6dba7dd0a5167d7?response-content-disposition=inline%3B+filename%3DUTF-8%27%27model.safetensors%3B+filename%3D%22model.safetensors%22%3B&Expires=1745858358&Policy=eyJTdGF0ZWl1bnQiOi0lt7IkNvbmlRpdGlubiI6eyJEYXRlTGZzc1RoYW4iO0nsiQVdT0kVWb2NoVGltZSI6MTc0NTg1ODM1OH19LCJSZXNvdXJjZSI6Imh0dHBzOi8vY2RuLWxmcY5oZi5jb3YyZXBvcy9iZS8wMy9iZTAzMWVjYjM3OGVmYzIwNWZkNzhlNjE2ZWU0ZWY0ZmYwN2I4MmE0YTU0ODNlNmMwYWE2NjBiZW44MGMwMTQ0L2M3YzE5ODhhYWUyMDFmODBiZjZjYXkYTVkYmJkNTg2NjQwOTUwM2I4OWRjYjYwJhODc3Y2E2ZGJhN2RkMGE1MTY3ZDc~cmVzcG9uc2UyY29udGVudC1kaXNwb3NpdGlvbjo0qIn1dfQ\_\_&Signature=vcvBMb-Wa~3hfp-QTeN765Vs7n3050Mt2002rqz2ppfZ0rZaK7V0LklG4aSkYUI-82zopqp40o1vfjMMIf1H30uPkWAh6zLri96CRnCOovy10jigNsEj9Rd~43suvBiof8pBqJx8yDtwfi6gtld~hAd0340FKEto3EAm0040aZJZnN4eA7k-tbPmTNgXV8N2C0bCyelQQ4BnmJ8crkKLclr0Xb~T3f183Nz0VKZ1H8I7oqLxTyQgvJH~E5HUb6fMIEijpLS4W8Gck2zYSXmkYYoe2omUbSssi4sfoYRbAhubZxkXEHiNqADu1kwTPr0di2WGh1SgGakPVMuTZiS6g\_\_&Key-Pair-Id=K3RPWS32NSSJCE HTTP/1.1" 200 437955512

model.safetensors: 0% | 0.00/438M [00:00<?, ?B/s]

DEBUG:filelock:Attempting to release lock 138327010667344 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-en-v1.5/c7c1988aae201f80cf91a5dbbd5866409503b89dcaba877ca6dba7dd0a5167d7.lock

2025-04-28 15:50:59,989 - filelock - DEBUG - Attempting to release lock

```

138327010667344 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/c7c1988aae201f80cf91a5dbbd5866409503b89dcaba877ca6dba7dd0a5167d7.lock
DEBUG:filelock:Lock 138327010667344 released on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/c7c1988aae201f80cf91a5dbbd5866409503b89dcaba877ca6dba7dd0a5167d7.lock
2025-04-28 15:50:59,995 - filelock - DEBUG - Lock 138327010667344 released on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/c7c1988aae201f80cf91a5dbbd5866409503b89dcaba877ca6dba7dd0a5167d7.lock
DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD /BAAI/bge-base-
en-v1.5/resolve/main/tokenizer_config.json HTTP/1.1" 200 0
2025-04-28 15:51:00,449 - urllib3.connectionpool - DEBUG -
https://huggingface.co:443 "HEAD /BAAI/bge-base-
en-v1.5/resolve/main/tokenizer_config.json HTTP/1.1" 200 0
DEBUG:filelock:Attempting to acquire lock 138326833456528 on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/37fca74771bc76a8e01178ce3a6055a0995f8093.lock
2025-04-28 15:51:00,453 - filelock - DEBUG - Attempting to acquire lock
138326833456528 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/37fca74771bc76a8e01178ce3a6055a0995f8093.lock
DEBUG:filelock:Lock 138326833456528 acquired on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/37fca74771bc76a8e01178ce3a6055a0995f8093.lock
2025-04-28 15:51:00,455 - filelock - DEBUG - Lock 138326833456528 acquired on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/37fca74771bc76a8e01178ce3a6055a0995f8093.lock
DEBUG:urllib3.connectionpool:https://huggingface.co:443 "GET /BAAI/bge-base-
en-v1.5/resolve/main/tokenizer_config.json HTTP/1.1" 200 366
2025-04-28 15:51:00,719 - urllib3.connectionpool - DEBUG -
https://huggingface.co:443 "GET /BAAI/bge-base-
en-v1.5/resolve/main/tokenizer_config.json HTTP/1.1" 200 366
tokenizer_config.json: 0%|          | 0.00/366 [00:00<?, ?B/s]
DEBUG:filelock:Attempting to release lock 138326833456528 on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/37fca74771bc76a8e01178ce3a6055a0995f8093.lock
2025-04-28 15:51:00,778 - filelock - DEBUG - Attempting to release lock
138326833456528 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/37fca74771bc76a8e01178ce3a6055a0995f8093.lock
DEBUG:filelock:Lock 138326833456528 released on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/37fca74771bc76a8e01178ce3a6055a0995f8093.lock

```

2025-04-28 15:51:00,791 - filelock - DEBUG - Lock 138326833456528 released on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-en-v1.5/37fca74771bc76a8e01178ce3a6055a0995f8093.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD /BAAI/bge-base-en-v1.5/resolve/main/vocab.txt HTTP/1.1" 200 0

2025-04-28 15:51:01,032 - urllib3.connectionpool - DEBUG - https://huggingface.co:443 "HEAD /BAAI/bge-base-en-v1.5/resolve/main/vocab.txt HTTP/1.1" 200 0

DEBUG:filelock:Attempting to acquire lock 138326832350032 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-en-v1.5/fb140275c155a9c7c5a3b3e0e77a9e839594a938.lock

2025-04-28 15:51:01,042 - filelock - DEBUG - Attempting to acquire lock 138326832350032 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-en-v1.5/fb140275c155a9c7c5a3b3e0e77a9e839594a938.lock

DEBUG:filelock:Lock 138326832350032 acquired on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-en-v1.5/fb140275c155a9c7c5a3b3e0e77a9e839594a938.lock

2025-04-28 15:51:01,049 - filelock - DEBUG - Lock 138326832350032 acquired on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-en-v1.5/fb140275c155a9c7c5a3b3e0e77a9e839594a938.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "GET /BAAI/bge-base-en-v1.5/resolve/main/vocab.txt HTTP/1.1" 200 231508

2025-04-28 15:51:01,312 - urllib3.connectionpool - DEBUG - https://huggingface.co:443 "GET /BAAI/bge-base-en-v1.5/resolve/main/vocab.txt HTTP/1.1" 200 231508

vocab.txt: 0%| | 0.00/232k [00:00<?, ?B/s]

DEBUG:filelock:Attempting to release lock 138326832350032 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-en-v1.5/fb140275c155a9c7c5a3b3e0e77a9e839594a938.lock

2025-04-28 15:51:01,758 - filelock - DEBUG - Attempting to release lock 138326832350032 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-en-v1.5/fb140275c155a9c7c5a3b3e0e77a9e839594a938.lock

DEBUG:filelock:Lock 138326832350032 released on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-en-v1.5/fb140275c155a9c7c5a3b3e0e77a9e839594a938.lock

2025-04-28 15:51:01,761 - filelock - DEBUG - Lock 138326832350032 released on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-en-v1.5/fb140275c155a9c7c5a3b3e0e77a9e839594a938.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD /BAAI/bge-base-en-v1.5/resolve/main/tokenizer.json HTTP/1.1" 200 0

```

2025-04-28 15:51:02,025 - urllib3.connectionpool - DEBUG -
https://huggingface.co:443 "HEAD /BAAI/bge-base-
en-v1.5/resolve/main/tokenizer.json HTTP/1.1" 200 0

DEBUG:filelock:Attempting to acquire lock 138326831023376 on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/688882a79f44442ddc1f60d70334a7ff5df0fb47.lock

2025-04-28 15:51:02,028 - filelock - DEBUG - Attempting to acquire lock
138326831023376 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/688882a79f44442ddc1f60d70334a7ff5df0fb47.lock

DEBUG:filelock:Lock 138326831023376 acquired on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/688882a79f44442ddc1f60d70334a7ff5df0fb47.lock

2025-04-28 15:51:02,029 - filelock - DEBUG - Lock 138326831023376 acquired on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/688882a79f44442ddc1f60d70334a7ff5df0fb47.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "GET /BAAI/bge-base-
en-v1.5/resolve/main/tokenizer.json HTTP/1.1" 200 711396

2025-04-28 15:51:02,266 - urllib3.connectionpool - DEBUG -
https://huggingface.co:443 "GET /BAAI/bge-base-
en-v1.5/resolve/main/tokenizer.json HTTP/1.1" 200 711396

tokenizer.json:  0%|          | 0.00/711k [00:00<?, ?B/s]

DEBUG:filelock:Attempting to release lock 138326831023376 on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/688882a79f44442ddc1f60d70334a7ff5df0fb47.lock

2025-04-28 15:51:02,920 - filelock - DEBUG - Attempting to release lock
138326831023376 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/688882a79f44442ddc1f60d70334a7ff5df0fb47.lock

DEBUG:filelock:Lock 138326831023376 released on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/688882a79f44442ddc1f60d70334a7ff5df0fb47.lock

2025-04-28 15:51:02,923 - filelock - DEBUG - Lock 138326831023376 released on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/688882a79f44442ddc1f60d70334a7ff5df0fb47.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD /BAAI/bge-base-
en-v1.5/resolve/main/added_tokens.json HTTP/1.1" 404 0

2025-04-28 15:51:03,260 - urllib3.connectionpool - DEBUG -
https://huggingface.co:443 "HEAD /BAAI/bge-base-
en-v1.5/resolve/main/added_tokens.json HTTP/1.1" 404 0

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD /BAAI/bge-base-
en-v1.5/resolve/main/special_tokens_map.json HTTP/1.1" 200 0

```

```

2025-04-28 15:51:03,498 - urllib3.connectionpool - DEBUG -
https://huggingface.co:443 "HEAD /BAAI/bge-base-
en-v1.5/resolve/main/special_tokens_map.json HTTP/1.1" 200 0

DEBUG:filelock:Attempting to acquire lock 138326833455696 on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/a8b3208c2884c4efb86e49300fdd3dc877220cdf.lock

2025-04-28 15:51:03,501 - filelock - DEBUG - Attempting to acquire lock
138326833455696 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/a8b3208c2884c4efb86e49300fdd3dc877220cdf.lock

DEBUG:filelock:Lock 138326833455696 acquired on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/a8b3208c2884c4efb86e49300fdd3dc877220cdf.lock

2025-04-28 15:51:03,506 - filelock - DEBUG - Lock 138326833455696 acquired on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/a8b3208c2884c4efb86e49300fdd3dc877220cdf.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "GET /BAAI/bge-base-
en-v1.5/resolve/main/special_tokens_map.json HTTP/1.1" 200 125

2025-04-28 15:51:03,740 - urllib3.connectionpool - DEBUG -
https://huggingface.co:443 "GET /BAAI/bge-base-
en-v1.5/resolve/main/special_tokens_map.json HTTP/1.1" 200 125

special_tokens_map.json: 0%|          | 0.00/125 [00:00<?, ?B/s]

DEBUG:filelock:Attempting to release lock 138326833455696 on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/a8b3208c2884c4efb86e49300fdd3dc877220cdf.lock

2025-04-28 15:51:03,779 - filelock - DEBUG - Attempting to release lock
138326833455696 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/a8b3208c2884c4efb86e49300fdd3dc877220cdf.lock

DEBUG:filelock:Lock 138326833455696 released on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/a8b3208c2884c4efb86e49300fdd3dc877220cdf.lock

2025-04-28 15:51:03,784 - filelock - DEBUG - Lock 138326833455696 released on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/a8b3208c2884c4efb86e49300fdd3dc877220cdf.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD /BAAI/bge-base-
en-v1.5/resolve/main/chat_template.jinja HTTP/1.1" 404 0

2025-04-28 15:51:04,365 - urllib3.connectionpool - DEBUG -
https://huggingface.co:443 "HEAD /BAAI/bge-base-
en-v1.5/resolve/main/chat_template.jinja HTTP/1.1" 404 0

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "GET
/api/models/BAAI/bge-base-en-v1.5/revision/main HTTP/1.1" 200 148840

```

2025-04-28 15:51:04,743 - urllib3.connectionpool - DEBUG -  
https://huggingface.co:443 "GET /api/models/BAAI/bge-base-en-v1.5/revision/main  
HTTP/1.1" 200 148840

DEBUG:urllib3.connectionpool:Starting new HTTPS connection (1):  
huggingface.co:443

2025-04-28 15:51:04,763 - urllib3.connectionpool - DEBUG - Starting new HTTPS  
connection (1): huggingface.co:443

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD /BAAI/bge-base-  
en-v1.5/resolve/a5beb1e3e68b9ab74eb54cfd186867f64f240e1a/1\_Pooling/config.json  
HTTP/1.1" 200 0

2025-04-28 15:51:05,008 - urllib3.connectionpool - DEBUG -  
https://huggingface.co:443 "HEAD /BAAI/bge-base-  
en-v1.5/resolve/a5beb1e3e68b9ab74eb54cfd186867f64f240e1a/1\_Pooling/config.json  
HTTP/1.1" 200 0

DEBUG:filelock:Attempting to acquire lock 138326830879824 on  
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-  
en-v1.5/da5bfd57e34ca45582e4bdbaa3e6deb9efffa08d.lock

2025-04-28 15:51:05,016 - filelock - DEBUG - Attempting to acquire lock  
138326830879824 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-  
en-v1.5/da5bfd57e34ca45582e4bdbaa3e6deb9efffa08d.lock

DEBUG:filelock:Lock 138326830879824 acquired on  
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-  
en-v1.5/da5bfd57e34ca45582e4bdbaa3e6deb9efffa08d.lock

2025-04-28 15:51:05,017 - filelock - DEBUG - Lock 138326830879824 acquired on  
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-  
en-v1.5/da5bfd57e34ca45582e4bdbaa3e6deb9efffa08d.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "GET /BAAI/bge-base-  
en-v1.5/resolve/a5beb1e3e68b9ab74eb54cfd186867f64f240e1a/1\_Pooling/config.json  
HTTP/1.1" 200 190

2025-04-28 15:51:05,255 - urllib3.connectionpool - DEBUG -  
https://huggingface.co:443 "GET /BAAI/bge-base-  
en-v1.5/resolve/a5beb1e3e68b9ab74eb54cfd186867f64f240e1a/1\_Pooling/config.json  
HTTP/1.1" 200 190

config.json: 0%| | 0.00/190 [00:00<?, ?B/s]

DEBUG:filelock:Attempting to release lock 138326830879824 on  
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-  
en-v1.5/da5bfd57e34ca45582e4bdbaa3e6deb9efffa08d.lock

2025-04-28 15:51:05,330 - filelock - DEBUG - Attempting to release lock  
138326830879824 on /root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-  
en-v1.5/da5bfd57e34ca45582e4bdbaa3e6deb9efffa08d.lock



```

DEBUG:filelock:Lock 138326830879824 released on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/da5bfd57e34ca45582e4bdbaa3e6deb9efffa08d.lock

2025-04-28 15:51:05,335 - filelock - DEBUG - Lock 138326830879824 released on
/root/.cache/huggingface/hub/.locks/models--BAAI--bge-base-
en-v1.5/da5bfd57e34ca45582e4bdbaa3e6deb9efffa08d.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "GET
/api/models/BAAI/bge-base-en-v1.5 HTTP/1.1" 200 148840

2025-04-28 15:51:05,591 - urllib3.connectionpool - DEBUG -
https://huggingface.co:443 "GET /api/models/BAAI/bge-base-en-v1.5 HTTP/1.1" 200
148840

INFO:__main__:Loading corpus documents with preprocessing...

2025-04-28 15:51:06,275 - __main__ - INFO - Loading corpus documents with
preprocessing...

INFO:__main__:Corpus size: 40221 documents

2025-04-28 15:51:17,102 - __main__ - INFO - Corpus size: 40221 documents

INFO:__main__:Computing embeddings for 40221 passages...

2025-04-28 15:51:17,188 - __main__ - INFO - Computing embeddings for 40221
passages...

Batches:   0%|          | 0/315 [00:00<?, ?it/s]

INFO:__main__:Corpus embeddings shape: torch.Size([40221, 768])

2025-04-28 16:02:01,758 - __main__ - INFO - Corpus embeddings shape:
torch.Size([40221, 768])

Batches:   0%|          | 0/1 [00:00<?, ?it/s]

INFO:__main__:First results before training:

2025-04-28 16:02:01,983 - __main__ - INFO - First results before training:

      id                               question \
0  1840  What is the function of Oseltamivir when admin...
1  1300  What is known about diseases associated with m...
2  2103  What is the applicability of the No Promoter L...
3  3122                               What is patisiran?
4  2813                               What is vcfanno?

                                relevant_passage_ids
0  [19557131, 10536125, 16838232, 12058885, 21677...
1  [25193783, 21158681, 22462537, 15896657, 22033...
2  [26530723, 16925832, 26620522, 11724736, 24429...
3  [28893208, 32311310, 32561705, 25610221, 33084...
4  [28400446, 15006940, 26949713, 8938977, 22052692]

```

[illegible]

```
INFO: __main__:Pre-training metrics:
```

```
2025-04-28 16:02:09,111 - __main__ - INFO - Pre-training metrics:
```

```
{'recall@10': 0.46611746869811393,
 'precision@10': 0.35067549668874143,
 'mrr': 0.7488840534006099,
 'map@10': 0.48557076302436636,
 'ndcg@10': np.float64(0.5791255229663892)}
```

Preparing data: 0%| | 0/3775 [00:00<?, ?it/s]

INFO:\_\_main\_\_:Total training examples: 45330

2025-04-28 16:03:09,463 - \_\_main\_\_ - INFO - Total training examples: 45330

INFO:\_\_main\_\_:Eval queries: 755; corpus size: 40221

2025-04-28 16:03:09,465 - \_\_main\_\_ - INFO - Eval queries: 755; corpus size: 40221

INFO:\_\_main\_\_:Training model on 45330 examples...

2025-04-28 16:03:09,467 - \_\_main\_\_ - INFO - Training model on 45330 examples...

Computing widget examples: 0%| | 0/1 [00:00<?, ?example/s]

wandb: WARNING The `run\_name` is currently set to the same value as `TrainingArguments.output\_dir`. If this was not intended, please specify a different run name by setting the `TrainingArguments.run\_name` parameter.

DEBUG:urllib3.connectionpool:Starting new HTTP connection (1): 172.28.0.12:9000

2025-04-28 16:03:12,221 - urllib3.connectionpool - DEBUG - Starting new HTTP connection (1): 172.28.0.12:9000

DEBUG:urllib3.connectionpool:http://172.28.0.12:9000 "GET /api/sessions?token=HTTP/1.1" 200 428

2025-04-28 16:03:12,227 - urllib3.connectionpool - DEBUG - http://172.28.0.12:9000 "GET /api/sessions?token= HTTP/1.1" 200 428

wandb: Using wandb-core as the SDK backend. Please refer to <https://wandb.me/wandb-core> for more information.

<IPython.core.display.Javascript object>

wandb: Logging into wandb.ai. (Learn how to deploy a W&B server locally: <https://wandb.me/wandb-server>)

wandb: You can find your API key in your browser here: <https://wandb.ai/authorize>

wandb: Paste an API key from your profile and hit enter:

.....

wandb: WARNING If you're specifying your api key in code, ensure this code is not shared publicly.

wandb: WARNING Consider setting the WANDB\_API\_KEY environment variable, or running `wandb login` from the command line.

wandb: No netrc file found, creating one.

```

wandb: Appending key for api.wandb.ai to your netrc file:
/root/.netrc
DEBUG:urllib3.connectionpool:Starting new HTTPS connection (1): api.wandb.ai:443
2025-04-28 16:03:29,883 - urllib3.connectionpool - DEBUG - Starting new HTTPS
connection (1): api.wandb.ai:443
DEBUG:urllib3.connectionpool:https://api.wandb.ai:443 "POST /graphql HTTP/1.1"
200 None
2025-04-28 16:03:30,111 - urllib3.connectionpool - DEBUG -
https://api.wandb.ai:443 "POST /graphql HTTP/1.1" 200 None
DEBUG:urllib3.connectionpool:https://api.wandb.ai:443 "POST /graphql HTTP/1.1"
200 None
2025-04-28 16:03:30,342 - urllib3.connectionpool - DEBUG -
https://api.wandb.ai:443 "POST /graphql HTTP/1.1" 200 None
wandb: Currently logged in as: hk17 (hk17-boston-
university) to https://api.wandb.ai. Use `wandb login
--relogin` to force relogin
<IPython.core.display.HTML object>
<IPython.core.display.HTML object>
<IPython.core.display.HTML object>
<IPython.core.display.HTML object>
<IPython.core.display.HTML object>
<IPython.core.display.HTML object>
INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 0.035285815102328866
after 100 steps:
2025-04-28 16:04:17,694 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch
0.035285815102328866 after 100 steps:
Batches: 0%|          | 0/24 [00:00<?, ?it/s]
Corpus Chunks: 100%|      | 1/1 [02:42<00:00, 162.68s/it]
INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755
2025-04-28 16:07:00,923 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:
755

```

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Corpus:  
40221

2025-04-28 16:07:00,925 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:  
40221

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Score-  
Function: cosine

2025-04-28 16:07:00,987 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-  
Function: cosine

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:  
69.93%

2025-04-28 16:07:00,989 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@1: 69.93%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
79.60%

2025-04-28 16:07:00,990 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 79.60%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
81.46%

2025-04-28 16:07:00,991 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 81.46%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
84.90%

2025-04-28 16:07:00,993 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 84.90%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
69.93%

2025-04-28 16:07:00,995 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 69.93%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
55.81%

2025-04-28 16:07:00,997 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 55.81%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
47.74%

2025-04-28 16:07:00,999 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 47.74%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 35.05%

2025-04-28 16:07:01,001 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 35.05%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.21%

2025-04-28 16:07:01,009 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.21%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
31.47%

2025-04-28 16:07:01,010 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 31.47%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:  
38.89%

2025-04-28 16:07:01,012 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@5: 38.89%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:  
47.71%

2025-04-28 16:07:01,016 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@10: 47.71%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:  
0.7531

2025-04-28 16:07:01,017 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:  
0.7531

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:  
0.5873

```

2025-04-28 16:07:01,018 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:
0.5873

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:
0.4649

2025-04-28 16:07:01,020 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:
0.4649

INFO:sentence_transformers.SentenceTransformer:Save model to output/my-
retrieval-model

2025-04-28 16:07:01,049 - sentence_transformers.SentenceTransformer - INFO -
Save model to output/my-retrieval-model

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 0.07057163020465773
after 200 steps:

2025-04-28 16:07:52,877 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch
0.07057163020465773 after 200 steps:

Batches:  0%|          | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%|      | 1/1 [02:42<00:00, 162.82s/it]
INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 16:10:36,220 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:
755

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Corpus:
40221

2025-04-28 16:10:36,222 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:
40221

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Score-
Function: cosine

2025-04-28 16:10:36,281 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-
Function: cosine

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:
69.93%

```

2025-04-28 16:10:36,284 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@1: 69.93%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
80.00%

2025-04-28 16:10:36,285 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 80.00%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
81.59%

2025-04-28 16:10:36,289 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 81.59%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
85.03%

2025-04-28 16:10:36,290 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 85.03%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
69.93%

2025-04-28 16:10:36,296 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 69.93%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
56.47%

2025-04-28 16:10:36,299 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 56.47%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
48.08%

2025-04-28 16:10:36,301 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 48.08%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 35.64%

2025-04-28 16:10:36,304 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 35.64%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.34%



2025-04-28 16:10:36,305 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.34%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
32.18%

2025-04-28 16:10:36,308 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 32.18%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:  
39.32%

2025-04-28 16:10:36,312 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@5: 39.32%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:  
48.42%

2025-04-28 16:10:36,315 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@10: 48.42%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:  
0.7539

2025-04-28 16:10:36,318 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:  
0.7539

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:  
0.5942

2025-04-28 16:10:36,321 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:  
0.5942

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:  
0.4743

2025-04-28 16:10:36,323 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:  
0.4743

INFO:sentence\_transformers.SentenceTransformer:Save model to output/my-  
retrieval-model

2025-04-28 16:10:36,349 - sentence\_transformers.SentenceTransformer - INFO -  
Save model to output/my-retrieval-model

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Information  
Retrieval Evaluation of the model on the dataset in epoch 0.1058574453069866  
after 300 steps:

2025-04-28 16:11:47,923 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Information Retrieval Evaluation of the model on the dataset in epoch  
0.1058574453069866 after 300 steps:

Batches: 0%| | 0/24 [00:00<?, ?it/s]  
Corpus Chunks: 100%| | 1/1 [02:42<00:00, 162.60s/it]  
INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 16:14:31,256 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:  
755

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Corpus:  
40221

2025-04-28 16:14:31,258 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:  
40221

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Score-  
Function: cosine

2025-04-28 16:14:31,318 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-  
Function: cosine

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:  
70.07%

2025-04-28 16:14:31,321 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@1: 70.07%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
79.47%

2025-04-28 16:14:31,322 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 79.47%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
82.25%

2025-04-28 16:14:31,324 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 82.25%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
85.56%

2025-04-28 16:14:31,326 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 85.56%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
70.07%

2025-04-28 16:14:31,332 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 70.07%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
55.98%

2025-04-28 16:14:31,334 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 55.98%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
48.19%

2025-04-28 16:14:31,335 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 48.19%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 35.79%

2025-04-28 16:14:31,341 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 35.79%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.20%

2025-04-28 16:14:31,342 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.20%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
31.69%

2025-04-28 16:14:31,344 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 31.69%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:  
39.25%

2025-04-28 16:14:31,345 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@5: 39.25%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:  
48.71%

```

2025-04-28 16:14:31,350 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Recall@10: 48.71%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:
0.7553

2025-04-28 16:14:31,351 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:
0.7553

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:
0.5949

2025-04-28 16:14:31,352 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:
0.5949

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:
0.4752

2025-04-28 16:14:31,354 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:
0.4752

INFO:sentence_transformers.SentenceTransformer:Save model to output/my-
retrieval-model

2025-04-28 16:14:31,394 - sentence_transformers.SentenceTransformer - INFO -
Save model to output/my-retrieval-model

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 0.14114326040931546
after 400 steps:

2025-04-28 16:15:27,107 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch
0.14114326040931546 after 400 steps:

Batches: 0%|          | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%|      | 1/1 [02:42<00:00, 162.08s/it]
INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 16:18:09,874 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:
755

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Corpus:
40221

2025-04-28 16:18:09,877 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:

```

40221

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Score-  
Function: cosine

2025-04-28 16:18:09,941 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-  
Function: cosine

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:  
69.40%

2025-04-28 16:18:09,944 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@1: 69.40%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
79.34%

2025-04-28 16:18:09,949 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 79.34%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
81.85%

2025-04-28 16:18:09,950 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 81.85%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
84.24%

2025-04-28 16:18:09,953 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 84.24%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
69.40%

2025-04-28 16:18:09,955 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 69.40%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
55.89%

2025-04-28 16:18:09,956 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 55.89%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
47.87%

2025-04-28 16:18:09,958 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 47.87%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 35.51%

2025-04-28 16:18:09,960 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 35.51%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.51%

2025-04-28 16:18:09,961 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.51%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
31.55%

2025-04-28 16:18:09,963 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 31.55%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:  
38.87%

2025-04-28 16:18:09,964 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@5: 38.87%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:  
47.82%

2025-04-28 16:18:09,966 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@10: 47.82%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:  
0.7480

2025-04-28 16:18:09,967 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:  
0.7480

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:  
0.5904

2025-04-28 16:18:09,968 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:  
0.5904

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:  
0.4737

2025-04-28 16:18:09,970 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:  
0.4737

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Information  
Retrieval Evaluation of the model on the dataset in epoch 0.17642907551164433  
after 500 steps:

2025-04-28 16:18:54,153 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Information Retrieval Evaluation of the model on the dataset in epoch  
0.17642907551164433 after 500 steps:

Batches: 0%| | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%| | 1/1 [02:42<00:00, 162.36s/it]

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 16:21:37,055 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:  
755

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Corpus:  
40221

2025-04-28 16:21:37,058 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:  
40221

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Score-  
Function: cosine

2025-04-28 16:21:37,149 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-  
Function: cosine

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:  
69.40%

2025-04-28 16:21:37,156 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@1: 69.40%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
78.41%

2025-04-28 16:21:37,158 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 78.41%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
81.59%

2025-04-28 16:21:37,160 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 81.59%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
84.37%

2025-04-28 16:21:37,163 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 84.37%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
69.40%

2025-04-28 16:21:37,165 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 69.40%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
55.32%

2025-04-28 16:21:37,168 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 55.32%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
47.50%

2025-04-28 16:21:37,170 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 47.50%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 35.44%

2025-04-28 16:21:37,172 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 35.44%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.34%

2025-04-28 16:21:37,176 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.34%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
31.35%

2025-04-28 16:21:37,178 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 31.35%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:  
38.42%



```

2025-04-28 16:21:37,181 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Recall@5: 38.42%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:
47.93%

2025-04-28 16:21:37,183 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Recall@10: 47.93%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:
0.7460

2025-04-28 16:21:37,185 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:
0.7460

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:
0.5884

2025-04-28 16:21:37,187 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:
0.5884

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:
0.4723

2025-04-28 16:21:37,189 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:
0.4723

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 0.2117148906139732
after 600 steps:

2025-04-28 16:22:21,478 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch
0.2117148906139732 after 600 steps:

Batches: 0%|          | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%|      | 1/1 [02:43<00:00, 163.35s/it]
INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 16:25:05,348 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:
755

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Corpus:
40221

```

2025-04-28 16:25:05,349 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:  
40221

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Score-  
Function: cosine

2025-04-28 16:25:05,414 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-  
Function: cosine

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:  
70.33%

2025-04-28 16:25:05,415 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@1: 70.33%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
79.60%

2025-04-28 16:25:05,417 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 79.60%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
83.05%

2025-04-28 16:25:05,419 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 83.05%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
85.30%

2025-04-28 16:25:05,420 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 85.30%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
70.33%

2025-04-28 16:25:05,422 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 70.33%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
56.20%

2025-04-28 16:25:05,424 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 56.20%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
48.66%

2025-04-28 16:25:05,426 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 48.66%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 36.01%

2025-04-28 16:25:05,428 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 36.01%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.53%

2025-04-28 16:25:05,429 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.53%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
31.92%

2025-04-28 16:25:05,431 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 31.92%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:  
39.59%

2025-04-28 16:25:05,433 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@5: 39.59%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:  
48.95%

2025-04-28 16:25:05,434 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@10: 48.95%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:  
0.7555

2025-04-28 16:25:05,436 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:  
0.7555

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:  
0.5982

2025-04-28 16:25:05,438 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:  
0.5982

```

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:
0.4804

2025-04-28 16:25:05,439 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:
0.4804

INFO:sentence_transformers.SentenceTransformer:Save model to output/my-
retrieval-model

2025-04-28 16:25:05,472 - sentence_transformers.SentenceTransformer - INFO -
Save model to output/my-retrieval-model

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 0.24700070571630206
after 700 steps:

2025-04-28 16:26:01,192 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch
0.24700070571630206 after 700 steps:

Batches: 0%|          | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%|      | 1/1 [02:42<00:00, 162.62s/it]

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 16:28:44,335 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:
755

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Corpus:
40221

2025-04-28 16:28:44,338 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:
40221

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Score-
Function: cosine

2025-04-28 16:28:44,409 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-
Function: cosine

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:
69.27%

2025-04-28 16:28:44,411 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Accuracy@1: 69.27%

```

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
79.21%

2025-04-28 16:28:44,413 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 79.21%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
82.12%

2025-04-28 16:28:44,414 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 82.12%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
85.03%

2025-04-28 16:28:44,416 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 85.03%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
69.27%

2025-04-28 16:28:44,417 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 69.27%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
56.03%

2025-04-28 16:28:44,419 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 56.03%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
48.16%

2025-04-28 16:28:44,421 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 48.16%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 35.88%

2025-04-28 16:28:44,422 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 35.88%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.09%

2025-04-28 16:28:44,424 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.09%

```

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:
31.72%

2025-04-28 16:28:44,425 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Recall@3: 31.72%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:
39.20%

2025-04-28 16:28:44,427 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Recall@5: 39.20%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:
48.69%

2025-04-28 16:28:44,429 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Recall@10: 48.69%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:
0.7482

2025-04-28 16:28:44,430 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:
0.7482

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:
0.5940

2025-04-28 16:28:44,431 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:
0.5940

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:
0.4779

2025-04-28 16:28:44,433 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:
0.4779

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 0.2822865208186309
after 800 steps:

2025-04-28 16:29:28,537 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch
0.2822865208186309 after 800 steps:

Batches: 0%|          | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%|      | 1/1 [02:43<00:00, 163.25s/it]
INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

```

2025-04-28 16:32:12,324 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:  
755

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Corpus:  
40221

2025-04-28 16:32:12,325 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:  
40221

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Score-  
Function: cosine

2025-04-28 16:32:12,392 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-  
Function: cosine

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:  
67.95%

2025-04-28 16:32:12,394 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@1: 67.95%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
79.74%

2025-04-28 16:32:12,397 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 79.74%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
82.38%

2025-04-28 16:32:12,399 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 82.38%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
85.43%

2025-04-28 16:32:12,401 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 85.43%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
67.95%

2025-04-28 16:32:12,403 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 67.95%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
56.11%

2025-04-28 16:32:12,405 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 56.11%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
48.32%

2025-04-28 16:32:12,407 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 48.32%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 35.87%

2025-04-28 16:32:12,409 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 35.87%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
17.52%

2025-04-28 16:32:12,411 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 17.52%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
31.75%

2025-04-28 16:32:12,413 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 31.75%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:  
39.49%

2025-04-28 16:32:12,415 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@5: 39.49%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:  
48.86%

2025-04-28 16:32:12,417 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@10: 48.86%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:  
0.7434

2025-04-28 16:32:12,419 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:  
0.7434



```

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:
0.5920

2025-04-28 16:32:12,420 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:
0.5920

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:
0.4752

2025-04-28 16:32:12,422 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:
0.4752

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 0.3175723359209598
after 900 steps:

2025-04-28 16:32:56,604 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch
0.3175723359209598 after 900 steps:

Batches: 0%|          | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%|      | 1/1 [02:46<00:00, 166.26s/it]

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 16:35:43,596 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:
755

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Corpus:
40221

2025-04-28 16:35:43,604 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:
40221

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Score-
Function: cosine

2025-04-28 16:35:43,687 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-
Function: cosine

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:
69.01%

2025-04-28 16:35:43,690 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Accuracy@1: 69.01%

```

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
79.74%

2025-04-28 16:35:43,692 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 79.74%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
82.52%

2025-04-28 16:35:43,695 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 82.52%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
85.83%

2025-04-28 16:35:43,699 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 85.83%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
69.01%

2025-04-28 16:35:43,701 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 69.01%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
56.34%

2025-04-28 16:35:43,703 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 56.34%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
48.69%

2025-04-28 16:35:43,705 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 48.69%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 36.28%

2025-04-28 16:35:43,707 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 36.28%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
17.66%

2025-04-28 16:35:43,709 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 17.66%

```

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:
31.93%

2025-04-28 16:35:43,711 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Recall@3: 31.93%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:
39.76%

2025-04-28 16:35:43,712 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Recall@5: 39.76%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:
49.37%

2025-04-28 16:35:43,714 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Recall@10: 49.37%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:
0.7485

2025-04-28 16:35:43,716 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:
0.7485

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:
0.5975

2025-04-28 16:35:43,718 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:
0.5975

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:
0.4787

2025-04-28 16:35:43,719 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:
0.4787

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 0.35285815102328866
after 1000 steps:

2025-04-28 16:36:29,448 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch
0.35285815102328866 after 1000 steps:

Batches: 0%|          | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%|      | 1/1 [02:44<00:00, 164.10s/it]

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

```

2025-04-28 16:39:14,236 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:  
755

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Corpus:  
40221

2025-04-28 16:39:14,238 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:  
40221

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Score-  
Function: cosine

2025-04-28 16:39:14,304 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-  
Function: cosine

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:  
68.48%

2025-04-28 16:39:14,306 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@1: 68.48%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
80.13%

2025-04-28 16:39:14,308 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 80.13%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
82.52%

2025-04-28 16:39:14,309 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 82.52%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
85.56%

2025-04-28 16:39:14,311 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 85.56%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
68.48%

2025-04-28 16:39:14,313 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 68.48%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
56.47%

2025-04-28 16:39:14,315 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 56.47%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
48.45%

2025-04-28 16:39:14,317 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 48.45%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 36.26%

2025-04-28 16:39:14,319 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 36.26%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.02%

2025-04-28 16:39:14,321 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.02%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
32.28%

2025-04-28 16:39:14,322 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 32.28%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:  
39.80%

2025-04-28 16:39:14,324 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@5: 39.80%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:  
49.52%

2025-04-28 16:39:14,325 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@10: 49.52%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:  
0.7468

2025-04-28 16:39:14,327 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:  
0.7468

```

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:
0.5992

2025-04-28 16:39:14,328 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:
0.5992

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:
0.4824

2025-04-28 16:39:14,330 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:
0.4824

INFO:sentence_transformers.SentenceTransformer:Save model to output/my-
retrieval-model

2025-04-28 16:39:14,353 - sentence_transformers.SentenceTransformer - INFO -
Save model to output/my-retrieval-model

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 0.3881439661256175
after 1100 steps:

2025-04-28 16:40:09,829 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch
0.3881439661256175 after 1100 steps:

Batches: 0%|          | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%|      | 1/1 [02:42<00:00, 162.88s/it]
INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 16:42:53,390 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:
755

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Corpus:
40221

2025-04-28 16:42:53,393 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:
40221

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Score-
Function: cosine

2025-04-28 16:42:53,464 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-
Function: cosine

```

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:  
69.93%

2025-04-28 16:42:53,466 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@1: 69.93%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
80.26%

2025-04-28 16:42:53,468 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 80.26%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
82.52%

2025-04-28 16:42:53,470 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 82.52%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
85.17%

2025-04-28 16:42:53,472 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 85.17%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
69.93%

2025-04-28 16:42:53,475 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 69.93%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
56.51%

2025-04-28 16:42:53,478 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 56.51%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
48.74%

2025-04-28 16:42:53,480 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 48.74%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 36.23%

2025-04-28 16:42:53,482 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 36.23%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.30%

2025-04-28 16:42:53,484 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.30%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
32.39%

2025-04-28 16:42:53,486 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 32.39%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:  
39.96%

2025-04-28 16:42:53,487 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@5: 39.96%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:  
49.22%

2025-04-28 16:42:53,489 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@10: 49.22%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:  
0.7543

2025-04-28 16:42:53,490 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:  
0.7543

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:  
0.6006

2025-04-28 16:42:53,492 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:  
0.6006

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:  
0.4844

2025-04-28 16:42:53,493 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:  
0.4844

INFO:sentence\_transformers.SentenceTransformer:Save model to output/my-  
retrieval-model

2025-04-28 16:42:53,525 - sentence\_transformers.SentenceTransformer - INFO -  
Save model to output/my-retrieval-model



INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Information Retrieval Evaluation of the model on the dataset in epoch 0.4234297812279464 after 1200 steps:

2025-04-28 16:43:49,100 -

sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Information Retrieval Evaluation of the model on the dataset in epoch 0.4234297812279464 after 1200 steps:

Batches: 0%| | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%| | 1/1 [02:42<00:00, 162.16s/it]

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 16:46:31,976 -

sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries: 755

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Corpus: 40221

2025-04-28 16:46:31,978 -

sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus: 40221

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Score-Function: cosine

2025-04-28 16:46:32,050 -

sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-Function: cosine

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1: 70.20%

2025-04-28 16:46:32,052 -

sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Accuracy@1: 70.20%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3: 80.53%

2025-04-28 16:46:32,054 -

sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Accuracy@3: 80.53%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5: 83.31%

2025-04-28 16:46:32,058 -

sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Accuracy@5: 83.31%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
85.70%

2025-04-28 16:46:32,062 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 85.70%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
70.20%

2025-04-28 16:46:32,063 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 70.20%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
57.17%

2025-04-28 16:46:32,065 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 57.17%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
49.30%

2025-04-28 16:46:32,067 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 49.30%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 36.74%

2025-04-28 16:46:32,071 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 36.74%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.41%

2025-04-28 16:46:32,072 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.41%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
32.61%

2025-04-28 16:46:32,074 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 32.61%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:  
40.53%

2025-04-28 16:46:32,078 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@5: 40.53%

```

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:
49.97%

2025-04-28 16:46:32,079 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Recall@10: 49.97%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:
0.7585

2025-04-28 16:46:32,080 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:
0.7585

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:
0.6078

2025-04-28 16:46:32,082 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:
0.6078

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:
0.4897

2025-04-28 16:46:32,083 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:
0.4897

INFO:sentence_transformers.SentenceTransformer:Save model to output/my-
retrieval-model

2025-04-28 16:46:32,115 - sentence_transformers.SentenceTransformer - INFO -
Save model to output/my-retrieval-model

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 0.45871559633027525
after 1300 steps:

2025-04-28 16:47:24,106 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch
0.45871559633027525 after 1300 steps:

Batches: 0%|          | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%|      | 1/1 [02:43<00:00, 163.22s/it]

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 16:50:07,895 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:
755

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Corpus:
40221

```

2025-04-28 16:50:07,897 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:  
40221

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Score-  
Function: cosine

2025-04-28 16:50:07,967 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-  
Function: cosine

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:  
69.27%

2025-04-28 16:50:07,970 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@1: 69.27%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
80.40%

2025-04-28 16:50:07,972 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 80.40%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
83.05%

2025-04-28 16:50:07,974 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 83.05%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
85.30%

2025-04-28 16:50:07,976 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 85.30%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
69.27%

2025-04-28 16:50:07,978 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 69.27%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
56.42%

2025-04-28 16:50:07,981 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 56.42%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
49.09%

2025-04-28 16:50:07,982 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 49.09%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 36.60%

2025-04-28 16:50:07,984 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 36.60%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.15%

2025-04-28 16:50:07,986 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.15%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
32.33%

2025-04-28 16:50:07,988 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 32.33%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:  
40.27%

2025-04-28 16:50:07,989 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@5: 40.27%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:  
49.82%

2025-04-28 16:50:07,991 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@10: 49.82%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:  
0.7519

2025-04-28 16:50:07,992 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:  
0.7519

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:  
0.6035

2025-04-28 16:50:07,993 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:  
0.6035

```

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:
0.4860

2025-04-28 16:50:07,995 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:
0.4860

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 0.4940014114326041
after 1400 steps:

2025-04-28 16:50:51,746 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch
0.4940014114326041 after 1400 steps:

Batches: 0%|          | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%|      | 1/1 [02:43<00:00, 163.43s/it]
INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 16:53:35,841 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:
755

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Corpus:
40221

2025-04-28 16:53:35,842 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:
40221

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Score-
Function: cosine

2025-04-28 16:53:35,909 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-
Function: cosine

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:
69.27%

2025-04-28 16:53:35,911 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Accuracy@1: 69.27%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:
80.26%

2025-04-28 16:53:35,913 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Accuracy@3: 80.26%

```

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
83.05%

2025-04-28 16:53:35,916 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 83.05%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
85.43%

2025-04-28 16:53:35,918 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 85.43%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
69.27%

2025-04-28 16:53:35,920 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 69.27%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
56.56%

2025-04-28 16:53:35,923 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 56.56%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
48.90%

2025-04-28 16:53:35,925 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 48.90%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 36.42%

2025-04-28 16:53:35,927 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 36.42%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.15%

2025-04-28 16:53:35,929 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.15%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
32.28%

2025-04-28 16:53:35,931 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 32.28%

```

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:
39.88%

2025-04-28 16:53:35,933 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Recall@5: 39.88%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:
49.87%

2025-04-28 16:53:35,935 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Recall@10: 49.87%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:
0.7527

2025-04-28 16:53:35,937 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:
0.7527

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:
0.6029

2025-04-28 16:53:35,939 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:
0.6029

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:
0.4858

2025-04-28 16:53:35,941 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:
0.4858

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 0.529287226534933
after 1500 steps:

2025-04-28 16:54:20,402 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch
0.529287226534933 after 1500 steps:

Batches: 0%|          | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%|      | 1/1 [02:43<00:00, 163.39s/it]
INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 16:57:04,363 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:
755

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Corpus:
40221

```



2025-04-28 16:57:04,365 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:  
40221

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Score-  
Function: cosine

2025-04-28 16:57:04,448 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-  
Function: cosine

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:  
70.46%

2025-04-28 16:57:04,450 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@1: 70.46%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
80.93%

2025-04-28 16:57:04,452 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 80.93%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
83.58%

2025-04-28 16:57:04,454 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 83.58%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
85.96%

2025-04-28 16:57:04,456 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 85.96%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
70.46%

2025-04-28 16:57:04,461 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 70.46%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
57.48%

2025-04-28 16:57:04,463 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 57.48%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
48.79%

2025-04-28 16:57:04,465 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 48.79%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 36.53%

2025-04-28 16:57:04,467 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 36.53%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.32%

2025-04-28 16:57:04,469 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.32%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
32.89%

2025-04-28 16:57:04,471 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 32.89%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:  
40.12%

2025-04-28 16:57:04,473 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@5: 40.12%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:  
50.29%

2025-04-28 16:57:04,475 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@10: 50.29%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:  
0.7608

2025-04-28 16:57:04,476 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:  
0.7608

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:  
0.6069

2025-04-28 16:57:04,478 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:  
0.6069

```

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:
0.4887

2025-04-28 16:57:04,479 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:
0.4887

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 0.5645730416372619
after 1600 steps:

2025-04-28 16:57:47,854 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch
0.5645730416372619 after 1600 steps:

Batches: 0%|          | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%|      | 1/1 [02:42<00:00, 162.99s/it]
INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 17:00:31,379 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:
755

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Corpus:
40221

2025-04-28 17:00:31,381 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:
40221

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Score-
Function: cosine

2025-04-28 17:00:31,448 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-
Function: cosine

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:
69.93%

2025-04-28 17:00:31,450 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Accuracy@1: 69.93%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:
80.26%

2025-04-28 17:00:31,451 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Accuracy@3: 80.26%

```

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
82.91%

2025-04-28 17:00:31,453 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 82.91%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
86.09%

2025-04-28 17:00:31,454 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 86.09%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
69.93%

2025-04-28 17:00:31,456 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 69.93%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
56.95%

2025-04-28 17:00:31,458 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 56.95%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
48.72%

2025-04-28 17:00:31,460 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 48.72%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 36.65%

2025-04-28 17:00:31,461 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 36.65%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.25%

2025-04-28 17:00:31,463 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.25%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
32.35%

2025-04-28 17:00:31,464 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 32.35%

```

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:
39.82%

2025-04-28 17:00:31,466 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Recall@5: 39.82%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:
50.29%

2025-04-28 17:00:31,467 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Recall@10: 50.29%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:
0.7562

2025-04-28 17:00:31,469 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:
0.7562

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:
0.6059

2025-04-28 17:00:31,470 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:
0.6059

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:
0.4882

2025-04-28 17:00:31,471 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:
0.4882

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 0.5998588567395907
after 1700 steps:

2025-04-28 17:01:15,587 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch
0.5998588567395907 after 1700 steps:

Batches: 0%|          | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%|      | 1/1 [02:43<00:00, 163.30s/it]
INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 17:03:59,416 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:
755

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Corpus:
40221

```

2025-04-28 17:03:59,418 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:  
40221

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Score-  
Function: cosine

2025-04-28 17:03:59,494 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-  
Function: cosine

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:  
70.46%

2025-04-28 17:03:59,496 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@1: 70.46%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
80.66%

2025-04-28 17:03:59,498 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 80.66%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
83.58%

2025-04-28 17:03:59,499 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 83.58%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
86.09%

2025-04-28 17:03:59,505 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 86.09%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
70.46%

2025-04-28 17:03:59,508 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 70.46%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
57.35%

2025-04-28 17:03:59,509 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 57.35%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
49.51%

2025-04-28 17:03:59,514 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 49.51%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 36.94%

2025-04-28 17:03:59,516 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 36.94%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.57%

2025-04-28 17:03:59,521 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.57%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
32.80%

2025-04-28 17:03:59,523 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 32.80%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:  
40.85%

2025-04-28 17:03:59,526 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@5: 40.85%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:  
50.68%

2025-04-28 17:03:59,529 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@10: 50.68%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:  
0.7617

2025-04-28 17:03:59,531 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:  
0.7617

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:  
0.6125

2025-04-28 17:03:59,536 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:  
0.6125

```

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:
0.4943

2025-04-28 17:03:59,538 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:
0.4943

INFO:sentence_transformers.SentenceTransformer:Save model to output/my-
retrieval-model

2025-04-28 17:03:59,568 - sentence_transformers.SentenceTransformer - INFO -
Save model to output/my-retrieval-model

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 0.6351446718419196
after 1800 steps:

2025-04-28 17:04:50,513 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch
0.6351446718419196 after 1800 steps:

Batches: 0%|          | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%|      | 1/1 [02:42<00:00, 162.01s/it]
INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 17:07:33,096 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:
755

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Corpus:
40221

2025-04-28 17:07:33,102 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:
40221

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Score-
Function: cosine

2025-04-28 17:07:33,195 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-
Function: cosine

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:
69.27%

2025-04-28 17:07:33,202 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Accuracy@1: 69.27%

```



INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
80.40%

2025-04-28 17:07:33,204 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 80.40%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
83.31%

2025-04-28 17:07:33,205 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 83.31%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
85.70%

2025-04-28 17:07:33,207 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 85.70%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
69.27%

2025-04-28 17:07:33,211 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 69.27%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
57.09%

2025-04-28 17:07:33,212 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 57.09%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
49.11%

2025-04-28 17:07:33,215 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 49.11%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 36.79%

2025-04-28 17:07:33,219 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 36.79%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.04%

2025-04-28 17:07:33,223 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.04%

```

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:
32.62%

2025-04-28 17:07:33,224 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Recall@3: 32.62%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:
40.50%

2025-04-28 17:07:33,225 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Recall@5: 40.50%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:
50.31%

2025-04-28 17:07:33,228 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Recall@10: 50.31%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:
0.7535

2025-04-28 17:07:33,230 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:
0.7535

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:
0.6070

2025-04-28 17:07:33,232 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:
0.6070

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:
0.4891

2025-04-28 17:07:33,236 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:
0.4891

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 0.6704304869442484
after 1900 steps:

2025-04-28 17:08:17,401 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch
0.6704304869442484 after 1900 steps:

Batches: 0%|          | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%|      | 1/1 [02:43<00:00, 163.04s/it]
INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

```

2025-04-28 17:11:01,029 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:  
755

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Corpus:  
40221

2025-04-28 17:11:01,031 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:  
40221

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Score-  
Function: cosine

2025-04-28 17:11:01,106 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-  
Function: cosine

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:  
69.67%

2025-04-28 17:11:01,108 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@1: 69.67%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
80.79%

2025-04-28 17:11:01,110 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 80.79%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
83.58%

2025-04-28 17:11:01,112 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 83.58%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
85.70%

2025-04-28 17:11:01,117 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 85.70%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
69.67%

2025-04-28 17:11:01,120 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 69.67%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
57.35%

2025-04-28 17:11:01,123 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 57.35%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
49.27%

2025-04-28 17:11:01,134 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 49.27%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 36.74%

2025-04-28 17:11:01,136 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 36.74%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.11%

2025-04-28 17:11:01,138 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.11%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
32.82%

2025-04-28 17:11:01,140 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 32.82%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:  
40.70%

2025-04-28 17:11:01,142 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@5: 40.70%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:  
50.29%

2025-04-28 17:11:01,144 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@10: 50.29%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:  
0.7563

2025-04-28 17:11:01,146 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:  
0.7563

```

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:
0.6081

2025-04-28 17:11:01,149 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:
0.6081

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:
0.4913

2025-04-28 17:11:01,151 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:
0.4913

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 0.7057163020465773
after 2000 steps:

2025-04-28 17:11:45,318 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch
0.7057163020465773 after 2000 steps:

Batches: 0%|          | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%|      | 1/1 [02:42<00:00, 162.69s/it]
INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 17:14:28,562 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:
755

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Corpus:
40221

2025-04-28 17:14:28,564 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:
40221

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Score-
Function: cosine

2025-04-28 17:14:28,636 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-
Function: cosine

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:
69.80%

2025-04-28 17:14:28,638 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Accuracy@1: 69.80%

```

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
81.06%

2025-04-28 17:14:28,640 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 81.06%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
83.31%

2025-04-28 17:14:28,642 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 83.31%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
85.56%

2025-04-28 17:14:28,645 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 85.56%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
69.80%

2025-04-28 17:14:28,647 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 69.80%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
57.57%

2025-04-28 17:14:28,648 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 57.57%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
49.30%

2025-04-28 17:14:28,650 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 49.30%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 36.78%

2025-04-28 17:14:28,651 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 36.78%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.37%

2025-04-28 17:14:28,653 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.37%

```

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:
32.98%

2025-04-28 17:14:28,654 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Recall@3: 32.98%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:
40.54%

2025-04-28 17:14:28,656 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Recall@5: 40.54%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:
50.30%

2025-04-28 17:14:28,657 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Recall@10: 50.30%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:
0.7560

2025-04-28 17:14:28,658 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:
0.7560

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:
0.6089

2025-04-28 17:14:28,660 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:
0.6089

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:
0.4926

2025-04-28 17:14:28,661 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:
0.4926

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 0.7410021171489062
after 2100 steps:

2025-04-28 17:15:13,018 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch
0.7410021171489062 after 2100 steps:

Batches: 0%|          | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%|      | 1/1 [02:43<00:00, 163.35s/it]
INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

```

2025-04-28 17:17:56,935 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:  
755

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Corpus:  
40221

2025-04-28 17:17:56,938 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:  
40221

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Score-  
Function: cosine

2025-04-28 17:17:57,008 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-  
Function: cosine

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:  
70.20%

2025-04-28 17:17:57,010 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@1: 70.20%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
81.06%

2025-04-28 17:17:57,011 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 81.06%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
83.71%

2025-04-28 17:17:57,012 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 83.71%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
85.96%

2025-04-28 17:17:57,014 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 85.96%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
70.20%

2025-04-28 17:17:57,015 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 70.20%



INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
57.57%

2025-04-28 17:17:57,017 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 57.57%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
49.88%

2025-04-28 17:17:57,020 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 49.88%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 36.98%

2025-04-28 17:17:57,022 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 36.98%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.39%

2025-04-28 17:17:57,023 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.39%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
33.15%

2025-04-28 17:17:57,025 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 33.15%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:  
40.87%

2025-04-28 17:17:57,027 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@5: 40.87%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:  
50.78%

2025-04-28 17:17:57,029 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@10: 50.78%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:  
0.7597

2025-04-28 17:17:57,030 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:  
0.7597

```

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:
0.6124

2025-04-28 17:17:57,031 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:
0.6124

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:
0.4947

2025-04-28 17:17:57,032 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:
0.4947

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 0.776287932251235
after 2200 steps:

2025-04-28 17:18:41,098 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch
0.776287932251235 after 2200 steps:

Batches: 0%|          | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%|      | 1/1 [02:43<00:00, 163.90s/it]
INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 17:21:25,571 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:
755

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Corpus:
40221

2025-04-28 17:21:25,573 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:
40221

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Score-
Function: cosine

2025-04-28 17:21:25,646 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-
Function: cosine

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:
70.60%

2025-04-28 17:21:25,648 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Accuracy@1: 70.60%

```

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
80.93%

2025-04-28 17:21:25,649 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 80.93%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
83.84%

2025-04-28 17:21:25,651 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 83.84%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
86.36%

2025-04-28 17:21:25,652 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 86.36%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
70.60%

2025-04-28 17:21:25,654 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 70.60%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
58.06%

2025-04-28 17:21:25,656 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 58.06%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
49.72%

2025-04-28 17:21:25,658 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 49.72%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 37.03%

2025-04-28 17:21:25,659 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 37.03%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.54%

2025-04-28 17:21:25,661 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.54%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
33.23%

2025-04-28 17:21:25,667 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 33.23%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:  
40.90%

2025-04-28 17:21:25,668 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@5: 40.90%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:  
50.86%

2025-04-28 17:21:25,669 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@10: 50.86%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:  
0.7635

2025-04-28 17:21:25,671 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:  
0.7635

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:  
0.6146

2025-04-28 17:21:25,675 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:  
0.6146

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:  
0.4974

2025-04-28 17:21:25,676 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:  
0.4974

INFO:sentence\_transformers.SentenceTransformer:Save model to output/my-  
retrieval-model

2025-04-28 17:21:25,703 - sentence\_transformers.SentenceTransformer - INFO -  
Save model to output/my-retrieval-model

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Information  
Retrieval Evaluation of the model on the dataset in epoch 0.8115737473535639  
after 2300 steps:

2025-04-28 17:22:17,243 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -

Information Retrieval Evaluation of the model on the dataset in epoch  
0.8115737473535639 after 2300 steps:

Batches: 0%| | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%| | 1/1 [02:43<00:00, 163.42s/it]

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 17:25:01,246 -

sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:  
755

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Corpus:  
40221

2025-04-28 17:25:01,248 -

sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:  
40221

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Score-  
Function: cosine

2025-04-28 17:25:01,324 -

sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-  
Function: cosine

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:  
70.99%

2025-04-28 17:25:01,326 -

sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@1: 70.99%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
81.32%

2025-04-28 17:25:01,327 -

sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 81.32%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
83.97%

2025-04-28 17:25:01,329 -

sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 83.97%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
86.23%

2025-04-28 17:25:01,330 -

sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 86.23%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
70.99%

2025-04-28 17:25:01,334 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 70.99%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
58.41%

2025-04-28 17:25:01,336 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 58.41%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
49.88%

2025-04-28 17:25:01,338 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 49.88%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 37.11%

2025-04-28 17:25:01,340 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 37.11%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.57%

2025-04-28 17:25:01,342 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.57%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
33.60%

2025-04-28 17:25:01,344 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 33.60%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:  
41.17%

2025-04-28 17:25:01,355 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@5: 41.17%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:  
50.94%

2025-04-28 17:25:01,357 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@10: 50.94%

```

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:
0.7655

2025-04-28 17:25:01,358 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:
0.7655

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:
0.6166

2025-04-28 17:25:01,360 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:
0.6166

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:
0.4997

2025-04-28 17:25:01,365 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:
0.4997

INFO:sentence_transformers.SentenceTransformer:Save model to output/my-
retrieval-model

2025-04-28 17:25:01,394 - sentence_transformers.SentenceTransformer - INFO -
Save model to output/my-retrieval-model

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 0.8468595624558928
after 2400 steps:

2025-04-28 17:25:53,468 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch
0.8468595624558928 after 2400 steps:

Batches: 0%|          | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%|      | 1/1 [02:42<00:00, 162.66s/it]
INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 17:28:36,689 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:
755

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Corpus:
40221

2025-04-28 17:28:36,691 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:
40221

```

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Score-  
Function: cosine

2025-04-28 17:28:36,764 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-  
Function: cosine

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:  
70.99%

2025-04-28 17:28:36,766 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@1: 70.99%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
81.32%

2025-04-28 17:28:36,768 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 81.32%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
84.50%

2025-04-28 17:28:36,770 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 84.50%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
86.36%

2025-04-28 17:28:36,772 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 86.36%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
70.99%

2025-04-28 17:28:36,774 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 70.99%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
58.37%

2025-04-28 17:28:36,776 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 58.37%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
50.09%

2025-04-28 17:28:36,777 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 50.09%



INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 37.23%

2025-04-28 17:28:36,779 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 37.23%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.63%

2025-04-28 17:28:36,781 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.63%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
33.43%

2025-04-28 17:28:36,782 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 33.43%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:  
41.41%

2025-04-28 17:28:36,783 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@5: 41.41%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:  
51.16%

2025-04-28 17:28:36,785 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@10: 51.16%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:  
0.7660

2025-04-28 17:28:36,786 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:  
0.7660

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:  
0.6179

2025-04-28 17:28:36,787 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:  
0.6179

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:  
0.5001

2025-04-28 17:28:36,789 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:  
0.5001

```

INFO:sentence_transformers.SentenceTransformer:Save model to output/my-
retrieval-model

2025-04-28 17:28:36,811 - sentence_transformers.SentenceTransformer - INFO -
Save model to output/my-retrieval-model

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 0.8821453775582216
after 2500 steps:

2025-04-28 17:29:26,271 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch
0.8821453775582216 after 2500 steps:

Batches: 0%|          | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%|      | 1/1 [02:42<00:00, 162.95s/it]
INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 17:32:09,787 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:
755

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Corpus:
40221

2025-04-28 17:32:09,789 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:
40221

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Score-
Function: cosine

2025-04-28 17:32:09,876 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-
Function: cosine

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:
70.73%

2025-04-28 17:32:09,879 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Accuracy@1: 70.73%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:
81.46%

2025-04-28 17:32:09,882 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Accuracy@3: 81.46%

```

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
83.97%

2025-04-28 17:32:09,885 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 83.97%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
86.23%

2025-04-28 17:32:09,887 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 86.23%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
70.73%

2025-04-28 17:32:09,890 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 70.73%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
58.45%

2025-04-28 17:32:09,895 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 58.45%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
49.96%

2025-04-28 17:32:09,897 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 49.96%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 37.25%

2025-04-28 17:32:09,902 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 37.25%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.55%

2025-04-28 17:32:09,904 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.55%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
33.54%

2025-04-28 17:32:09,907 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 33.54%

```

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:
41.12%

2025-04-28 17:32:09,910 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Recall@5: 41.12%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:
51.13%

2025-04-28 17:32:09,913 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Recall@10: 51.13%

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:
0.7645

2025-04-28 17:32:09,915 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:
0.7645

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:
0.6176

2025-04-28 17:32:09,917 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:
0.6176

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:
0.4999

2025-04-28 17:32:09,919 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:
0.4999

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 0.9174311926605505
after 2600 steps:

2025-04-28 17:32:54,011 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch
0.9174311926605505 after 2600 steps:

Batches: 0%|          | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%|      | 1/1 [02:42<00:00, 162.93s/it]
INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 17:35:37,587 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:
755

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Corpus:
40221

```

2025-04-28 17:35:37,588 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:  
40221

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Score-  
Function: cosine

2025-04-28 17:35:37,665 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-  
Function: cosine

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:  
70.99%

2025-04-28 17:35:37,667 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@1: 70.99%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
81.32%

2025-04-28 17:35:37,668 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 81.32%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
84.24%

2025-04-28 17:35:37,670 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 84.24%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
86.36%

2025-04-28 17:35:37,671 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 86.36%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
70.99%

2025-04-28 17:35:37,673 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 70.99%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
58.32%

2025-04-28 17:35:37,675 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 58.32%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
50.15%

2025-04-28 17:35:37,677 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 50.15%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 37.23%

2025-04-28 17:35:37,679 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 37.23%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.71%

2025-04-28 17:35:37,680 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.71%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
33.46%

2025-04-28 17:35:37,682 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 33.46%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:  
41.36%

2025-04-28 17:35:37,683 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@5: 41.36%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:  
51.12%

2025-04-28 17:35:37,684 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@10: 51.12%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:  
0.7664

2025-04-28 17:35:37,685 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:  
0.7664

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:  
0.6185

2025-04-28 17:35:37,686 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:  
0.6185

```

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:
0.5013

2025-04-28 17:35:37,687 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:
0.5013

INFO:sentence_transformers.SentenceTransformer:Save model to output/my-
retrieval-model

2025-04-28 17:35:37,715 - sentence_transformers.SentenceTransformer - INFO -
Save model to output/my-retrieval-model

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 0.9527170077628794
after 2700 steps:

2025-04-28 17:36:27,207 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch
0.9527170077628794 after 2700 steps:

Batches: 0%|          | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%|      | 1/1 [02:42<00:00, 162.60s/it]
INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 17:39:10,389 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:
755

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Corpus:
40221

2025-04-28 17:39:10,391 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:
40221

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Score-
Function: cosine

2025-04-28 17:39:10,466 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-
Function: cosine

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:
70.99%

2025-04-28 17:39:10,467 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Accuracy@1: 70.99%

```

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
81.46%

2025-04-28 17:39:10,469 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 81.46%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
84.24%

2025-04-28 17:39:10,471 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 84.24%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
86.23%

2025-04-28 17:39:10,472 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 86.23%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
70.99%

2025-04-28 17:39:10,478 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 70.99%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
58.50%

2025-04-28 17:39:10,479 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 58.50%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
50.17%

2025-04-28 17:39:10,482 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 50.17%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 37.26%

2025-04-28 17:39:10,486 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 37.26%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.71%

2025-04-28 17:39:10,488 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.71%



INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
33.61%

2025-04-28 17:39:10,489 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 33.61%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:  
41.37%

2025-04-28 17:39:10,492 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@5: 41.37%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:  
51.12%

2025-04-28 17:39:10,494 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@10: 51.12%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:  
0.7663

2025-04-28 17:39:10,496 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:  
0.7663

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:  
0.6189

2025-04-28 17:39:10,500 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:  
0.6189

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:  
0.5017

2025-04-28 17:39:10,502 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:  
0.5017

INFO:sentence\_transformers.SentenceTransformer:Save model to output/my-  
retrieval-model

2025-04-28 17:39:10,533 - sentence\_transformers.SentenceTransformer - INFO -  
Save model to output/my-retrieval-model

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Information  
Retrieval Evaluation of the model on the dataset in epoch 0.9880028228652082  
after 2800 steps:

2025-04-28 17:40:02,043 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -

Information Retrieval Evaluation of the model on the dataset in epoch  
0.9880028228652082 after 2800 steps:

Batches: 0%| | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%| | 1/1 [02:42<00:00, 162.11s/it]

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 17:42:44,712 -

sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:  
755

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Corpus:  
40221

2025-04-28 17:42:44,714 -

sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:  
40221

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Score-  
Function: cosine

2025-04-28 17:42:44,788 -

sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-  
Function: cosine

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:  
70.99%

2025-04-28 17:42:44,790 -

sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@1: 70.99%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
81.46%

2025-04-28 17:42:44,792 -

sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 81.46%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
84.24%

2025-04-28 17:42:44,794 -

sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 84.24%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
86.23%

2025-04-28 17:42:44,795 -

sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 86.23%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
70.99%

2025-04-28 17:42:44,801 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 70.99%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
58.50%

2025-04-28 17:42:44,803 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 58.50%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
50.15%

2025-04-28 17:42:44,808 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 50.15%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 37.31%

2025-04-28 17:42:44,810 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 37.31%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.71%

2025-04-28 17:42:44,815 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.71%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
33.54%

2025-04-28 17:42:44,817 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 33.54%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:  
41.34%

2025-04-28 17:42:44,818 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@5: 41.34%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:  
51.16%

2025-04-28 17:42:44,823 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@10: 51.16%

```

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:
0.7662

2025-04-28 17:42:44,824 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:
0.7662

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:
0.6192

2025-04-28 17:42:44,826 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:
0.6192

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:
0.5015

2025-04-28 17:42:44,827 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:
0.5015

INFO:sentence_transformers.SentenceTransformer:Save model to output/my-
retrieval-model

2025-04-28 17:42:44,850 - sentence_transformers.SentenceTransformer - INFO -
Save model to output/my-retrieval-model

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Information
Retrieval Evaluation of the model on the dataset in epoch 1.0 after 2834 steps:

2025-04-28 17:43:05,623 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO -
Information Retrieval Evaluation of the model on the dataset in epoch 1.0 after
2834 steps:

Batches: 0%|          | 0/24 [00:00<?, ?it/s]

Corpus Chunks: 100%|      | 1/1 [02:44<00:00, 164.26s/it]
INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Queries: 755

2025-04-28 17:45:50,468 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Queries:
755

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Corpus:
40221

2025-04-28 17:45:50,470 -
sentence_transformers.evaluation.InformationRetrievalEvaluator - INFO - Corpus:
40221

INFO:sentence_transformers.evaluation.InformationRetrievalEvaluator:Score-
Function: cosine

```

2025-04-28 17:45:50,547 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - Score-  
Function: cosine

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@1:  
70.99%

2025-04-28 17:45:50,550 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@1: 70.99%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@3:  
81.46%

2025-04-28 17:45:50,553 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@3: 81.46%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@5:  
84.24%

2025-04-28 17:45:50,557 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@5: 84.24%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Accuracy@10:  
86.23%

2025-04-28 17:45:50,560 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Accuracy@10: 86.23%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@1:  
70.99%

2025-04-28 17:45:50,562 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@1: 70.99%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@3:  
58.50%

2025-04-28 17:45:50,564 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@3: 58.50%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@5:  
50.15%

2025-04-28 17:45:50,565 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@5: 50.15%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Precision@10  
: 37.31%

2025-04-28 17:45:50,567 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Precision@10: 37.31%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@1:  
18.71%

2025-04-28 17:45:50,569 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@1: 18.71%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@3:  
33.54%

2025-04-28 17:45:50,571 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@3: 33.54%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@5:  
41.34%

2025-04-28 17:45:50,572 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@5: 41.34%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:Recall@10:  
51.15%

2025-04-28 17:45:50,574 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO -  
Recall@10: 51.15%

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MRR@10:  
0.7663

2025-04-28 17:45:50,575 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MRR@10:  
0.7663

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:NDCG@10:  
0.6192

2025-04-28 17:45:50,576 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - NDCG@10:  
0.6192

INFO:sentence\_transformers.evaluation.InformationRetrievalEvaluator:MAP@100:  
0.5016

2025-04-28 17:45:50,578 -  
sentence\_transformers.evaluation.InformationRetrievalEvaluator - INFO - MAP@100:  
0.5016

INFO:\_\_main\_\_:Training complete!

2025-04-28 17:45:50,620 - \_\_main\_\_ - INFO - Training complete!

INFO:\_\_main\_\_:Computing embeddings for 40221 passages...

2025-04-28 17:45:50,623 - \_\_main\_\_ - INFO - Computing embeddings for 40221 passages...

Batches: 0%| | 0/315 [00:00<?, ?it/s]

INFO:\_\_main\_\_:Corpus embeddings shape: torch.Size([40221, 768])

2025-04-28 17:48:37,834 - \_\_main\_\_ - INFO - Corpus embeddings shape: torch.Size([40221, 768])

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/2 [00:00<?, ?it/s]

INFO:\_\_main\_\_:Post-training metrics:

2025-04-28 17:48:43,699 - \_\_main\_\_ - INFO - Post-training metrics:

```
{'recall@10': 0.5155112545050704,
 'precision@10': 0.3832317880794682,
 'mrr': 0.7688612425102495,
 'map@10': 0.5368445899012051,
 'ndcg@10': np.float64(0.6258232600895175)}
```

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/4 [00:00<?, ?it/s]

Batches: 0%| | 0/2 [00:00<?, ?it/s]

INFO:\_\_main\_\_:Saved relevant passage ids to stage\_1\_prediction.csv

2025-04-28 17:48:45,243 - \_\_main\_\_ - INFO - Saved relevant passage ids to stage\_1\_prediction.csv

	id	passage
0	9797	New data on viruses isolated from patients wit...
1	11906	We describe an improved method for detecting d...
2	16083	We have studied the effects of curare on respo...
3	23188	Kinetic and electrophoretic properties of 230-...
4	23469	Male Wistar specific-pathogen-free rats aged 2...

	id	question \
0	164	What is the enzymatic activity of the breast c...
1	1816	What is the effect of CPEB3 binding to the CPE...
2	4363	Which disease do pathogenic NR2F1 variants cause?
3	753	Which enzyme is deficient in Gaucher's disease?
4	2314	What does davunetide do to microtubules?

	answer \
0	E3-ubiquitin ligase activity is the only known...
1	The cytoplasmic polyadenylation element (CPE) ...
2	Bosch-Boonstra-Schaaf optic atrophy syndrome (...)



```

3 Gaucher's disease is caused by deficient lysos...
4     Davunetide or NAP is a microtubule-stabilizer.

```

```

                                relevant_passage_ids
0 [20681793, 22034435, 19088202, 16479151, 24278...
1 [17923234, 17481902, 25066254, 23776146, 26398...
2                                [26986877]
3 [20946052, 18627336, 16781064, 22843412, 24485...
4                                [24210139]

```

```

[ ]: from transformers import (
    GPT2LMHeadModel,
    GPT2Tokenizer,
    pipeline,
    DataCollatorForLanguageModeling,
    Trainer,
    TrainingArguments
)
from datasets import Dataset
import evaluate
from nltk.translate.bleu_score import sentence_bleu, corpus_bleu

class Generator:
    """
    A class for training, evaluating, and predicting with a text generation
    ↪(causal LM) model.

    The class includes methods to:
    1. Build a document dictionary from a DataFrame.
    2. Prepare training data by combining questions, contexts (from relevant
    ↪passages),
       and answers into prompts.
    3. Tokenize the prompts for training.
    4. Fine-tune the generator using Hugging Face's Trainer.
    5. Evaluate the generator on a sample evaluation set.
    6. Predict answers for a test DataFrame that includes predicted relevant
    ↪passage IDs.

    In both evaluation and prediction, the prompt is truncated if
    ↪necessary so that its
       token count plus requested new tokens do not exceed the model's
    ↪maximum allowed tokens.
    """

    def __init__(self, model_name="gpt2"):
        logger.info(f"Loading model and tokenizer for {model_name}")
        self.tokenizer = GPT2Tokenizer.from_pretrained(model_name)
        if self.tokenizer.pad_token is None:

```

```

        self.tokenizer.pad_token = self.tokenizer.eos_token
self.model = GPT2LMHeadModel.from_pretrained(model_name)
self.model.config.pad_token_id = self.tokenizer.eos_token_id

# Placeholders for training data and document dictionary.
self.doc_dict = {}
self.qa_df = None
self.tokenized_dataset = None

def build_doc_dict(self, documents_df):
    logger.info("Building document dictionary from DataFrame...")
    documents_df["passage"] = documents_df["passage"].fillna("").astype(str)
    self.doc_dict = dict(zip(documents_df["id"].astype(str),
↪documents_df["passage"]))
    logger.info(f"Built doc_dict with {len(self.doc_dict)} documents.")

def create_prompt(self, question, context, answer=None):
    """Build a prompt string without any truncation."""
    if answer:
        prompt = f"Question: {question}\nContext: {context}\nAnswer:
↪{answer}"
    else:
        prompt = f"Question: {question}\nContext: {context}\nAnswer:"
    return prompt

def build_truncated_prompt(self, question, context, max_new_tokens):
    """
    Build a prompt while ensuring that its tokenized version does not exceed
    (model's maximum positions - max_new_tokens) tokens. If it does,
↪truncate the context.
    """
    max_positions = self.model.config.n_positions
    # The allowed prompt length is the total positions minus the room we
↪want for new tokens.
    allowed_prompt_length = max_positions - max_new_tokens

    # Build full prompt.
    full_prompt = self.create_prompt(question, context)
    # Tokenize the full prompt.
    inputs = self.tokenizer(full_prompt, return_tensors="pt")
    prompt_length = inputs.input_ids.shape[1]
    if prompt_length <= allowed_prompt_length:
        return full_prompt
    else:
        # Preserve question and "Answer:" markers.
        q_part = f"Question: {question}\nContext: "
        a_part = "\nAnswer:"

```

```

        q_tokens = self.tokenizer(q_part, add_special_tokens=False).
↪input_ids
        a_tokens = self.tokenizer(a_part, add_special_tokens=False).
↪input_ids
        fixed_length = len(q_tokens) + len(a_tokens)
        allowed_for_context = allowed_prompt_length - fixed_length
        # Tokenize the full context.
        context_tokens = self.tokenizer(context, add_special_tokens=False).
↪input_ids
        truncated_context_tokens = context_tokens[:max(allowed_for_context,
↪0)]
        truncated_context = self.tokenizer.decode(truncated_context_tokens,
↪skip_special_tokens=True)
        truncated_prompt = f"Question: {question}\nContext:
↪{truncated_context}\nAnswer:"
        logger.warning(
            f"Prompt truncated: original length {prompt_length}, "
            f"allowed prompt length {allowed_prompt_length}."
        )
        return truncated_prompt

    def prepare_training_data(self, train_df):
        """
        Prepare training data by combining questions, context from relevant
↪passages, and answers into prompts.

        Expects train_df to have the following columns:
        - 'question'
        - 'answer'
        - 'relevant_passage_ids' : a string representation of a list (e.g.,
↪"[2, 5]")

        This method builds an internal DataFrame (self.qa_df) with columns:
        - "question", "relevant_docs" (concatenated passages), "answer",
↪"prompt"

        Args:
            train_df (pd.DataFrame): DataFrame containing training samples.
        """
        if not self.doc_dict:
            raise ValueError("Document dictionary is empty. Call
↪build_doc_dict(documents_df) first.")

        logger.info("Preparing training data from DataFrame...")
        records = []

```

```

        for _, row in tqdm(train_df.iterrows(), total=len(train_df),
↪desc="Preparing training data"):
            question = row["question"]
            answer = row["answer"]
            # Parse the relevant_passage_ids string to a list.
            pid_str = str(row["relevant_passage_ids"]).strip("[]")
            doc_ids = [pid.strip() for pid in pid_str.split(",") if pid.strip()]

            # Gather relevant passages based on the doc_dict
            relevant_texts = [self.doc_dict[pid] for pid in doc_ids if pid in
↪self.doc_dict]
            combined_passages = " ".join(relevant_texts)

            prompt = self.create_prompt(question, combined_passages, answer)

            records.append({
                "question": question,
                "relevant_docs": combined_passages,
                "answer": answer,
                "prompt": prompt
            })

        self.qa_df = pd.DataFrame(records)
        logger.info(f"Prepared {len(self.qa_df)} training records.")

    def tokenize_training_data(self, max_length=128):
        """
        Convert the prepared training DataFrame into a tokenized Hugging Face
↪Dataset.

        Args:
            max_length (int): Maximum token length for each prompt.
        """
        if self.qa_df is None:
            raise ValueError("Training data not prepared. Call
↪prepare_training_data(train_df) first.")

        logger.info("Converting training data to Dataset and tokenizing...")
        dataset = Dataset.from_pandas(self.qa_df[["prompt"]])

    def tokenize_function(examples):
        return self.tokenizer(
            examples["prompt"],
            padding="max_length",
            truncation=True,
            max_length=max_length
        )

```

```

self.tokenized_dataset = dataset.map(tokenize_function, batched=True)
logger.info("Tokenization complete.")

def train_model(self, output_dir="my_gpt2_generator", num_train_epochs=1,
    ↪batch_size=8, logging_steps=50):
    """
    Train (fine-tune) the generator model using the tokenized dataset.

    Args:
        output_dir (str): Directory in which to save the fine-tuned model.
        num_train_epochs (int): Number of training epochs.
        batch_size (int): Batch size per device.
        logging_steps (int): Frequency (in steps) at which to log training_
    ↪metrics.
    """
    if self.tokenized_dataset is None:
        raise ValueError("Tokenized training data not found. Run_
    ↪tokenize_training_data(max_length) first.")

    logger.info("Setting up training...")
    data_collator = DataCollatorForLanguageModeling(
        tokenizer=self.tokenizer,
        mlm=False # GPT-2 uses a causal LM objective, not masked LM
    )

    training_args = TrainingArguments(
        output_dir=output_dir,
        num_train_epochs=num_train_epochs,
        per_device_train_batch_size=batch_size,
        logging_steps=logging_steps,
        save_total_limit=2, # adjust as needed
        logging_dir=f"{output_dir}/logs",
    )

    trainer = Trainer(
        model=self.model,
        args=training_args,
        train_dataset=self.tokenized_dataset,
        tokenizer=self.tokenizer,
        data_collator=data_collator
    )

    logger.info("Starting training...")
    trainer.train()
    logger.info("Training complete!")

```

```

# Save the model after training
trainer.save_model(output_dir)
logger.info(f"Model saved to {output_dir}")

def evaluate_bert(self, eval_df, max_eval=100, max_new_tokens=32,
do_sample=False):
    """
    Evaluate the generator model on a subset of examples using BERTScore.

    This method uses a text generation pipeline with the current model &
tokenizer,
    constructs a prompt for each evaluation example, generates an answer,
    and then computes BERTScore F1 between the generated answers and gold
answers.

    The eval_df must contain the columns:
    - 'question'
    - 'answer'
    - 'relevant_passage_ids'

    Args:
        eval_df (pd.DataFrame): The evaluation DataFrame.
        max_eval (int): Maximum number of samples to evaluate.
        max_new_tokens (int): Maximum number of new tokens to generate
beyond the prompt.
        do_sample (bool): Whether to use sampling during generation.

    Returns:
        tuple: (bert_f1, predictions, references)
            - bert_f1 (float): The average BERTScore F1 score for all
evaluation samples.
            - predictions (List[str]): The list of generated answers.
            - references (List[str]): The list of ground-truth answers.
    """
    if not self.doc_dict:
        raise ValueError("Document dictionary is empty. Ensure
build_doc_dict(documents_df) has been called.")

    logger.info("Evaluating generator model using BERTScore...")
    gen_pipe = pipeline("text-generation", model=self.model, tokenizer=self
tokenizer)
    # Load the BERTScore metric. (Adjust the language parameter if needed.)

    predictions = []
    references = []

```

```

subset_df = eval_df.sample(n=min(len(eval_df), max_eval),
↳random_state=42)

for _, row in tqdm(subset_df.iterrows(), total=len(subset_df),
↳desc="Evaluating"):
    question = row["question"]
    gold_answer = row["answer"]

    # Parse relevant passage IDs and combine passages.
    pid_str = str(row["relevant_passage_ids"]).strip("[]")
    doc_ids = [pid.strip() for pid in pid_str.split(",") if pid.strip()]
    relevant_texts = [self.doc_dict[pid] for pid in doc_ids if pid in
↳self.doc_dict]
    combined_context = " ".join(relevant_texts)

    # Build a prompt that is safely truncated.
    prompt = self.build_truncated_prompt(question, combined_context,
↳max_new_tokens)

    # Optional: verify prompt length.
    inputs = self.tokenizer(prompt, return_tensors="pt")
    input_length = inputs.input_ids.shape[1]
    if input_length > self.model.config.n_positions - max_new_tokens:
        logger.warning(f"Adjusted prompt length {input_length} exceeds
↳allowed length; further truncation may be needed.")

    gen_output = gen_pipe(
        prompt,
        max_new_tokens=max_new_tokens,
        num_return_sequences=1,
        do_sample=do_sample
    )[0]["generated_text"]

    # Extract the generated answer (everything after "Answer:")
    pred_answer = gen_output.split("Answer: ")[-1].strip()
    predictions.append(pred_answer)
    references.append(gold_answer)

gt = [[x] for x in references]
pred = predictions
bleu_score = corpus_bleu(gt, pred)
logger.info(f"Bleu Score: {bleu_score:.4f}")
return bleu_score, predictions, references

def predict(self, test_df, max_new_tokens=32, do_sample=False):
    """
    Generate answers for a test DataFrame containing:

```

```

        - 'id', 'question', and 'relevant_passage_ids'
        The generated answers are saved in a new column "answer".
        """
        logger.info("Generating answers for test data...")
        gen_pipe = pipeline("text-generation", model=self.model, tokenizer=self.
↳tokenizer)

        results_df = test_df.copy()
        answers = []

        for _, row in tqdm(results_df.iterrows(), total=len(results_df),
↳desc="Generating answers"):
            question = row["question"]
            pid_str = str(row["relevant_passage_ids"]).strip("[]")
            doc_ids = [pid.strip() for pid in pid_str.split(",") if pid.strip()]
            relevant_texts = [self.doc_dict[pid] for pid in doc_ids if pid in
↳self.doc_dict]
            combined_context = " ".join(relevant_texts)

            # Build a safe prompt via truncation.
            prompt = self.build_truncated_prompt(question, combined_context,
↳max_new_tokens)

            inputs = self.tokenizer(prompt, return_tensors="pt")
            input_length = inputs.input_ids.shape[1]
            if input_length > self.model.config.n_positions - max_new_tokens:
                logger.warning("Prompt length exceeds allowed length after
↳truncation.")

            gen_output = gen_pipe(
                prompt,
                max_new_tokens=max_new_tokens,
                num_return_sequences=1,
                do_sample=do_sample
            )[0]["generated_text"]

            pred_answer = gen_output.split("Answer: ")[-1].strip()
            answers.append(pred_answer)

        results_df["answer"] = answers
        logger.info("Answer generation complete.")
        return results_df

from transformers import pipeline
import torch

# Initialize generator with a RAM-efficient model

```



```

gen = pipeline(
    "text-generation",
    model="TinyLlama/TinyLlama-1.1B-Chat-v1.0",
    torch_dtype=torch.float16, # FP16 for memory savings
    device="cuda"
)

def prepare_context(passage_ids, doc_dict):
    """Helper to build context from passage IDs"""
    return " ".join([doc_dict.get(str(pid), "") for pid in eval(passage_ids)])

# Build document dictionary (alternative to class method)
doc_dict = dict(zip(documents_df['id'].astype(str), documents_df['passage']))

def evaluate_generator(gen, eval_df, doc_dict, max_eval=5):
    references = []
    predictions = []

    for _, row in eval_df.head(max_eval).iterrows():
        context = prepare_context(row['relevant_passage_ids'], doc_dict)
        prompt = f"Answer based on context:\nContext: {context}\nQuestion: {row['question']}\nAnswer:"

        output = gen(
            prompt,
            max_new_tokens=32,
            do_sample=False
        )[0]['generated_text']

        pred = output.split("Answer: ")[-1].strip()
        predictions.append(pred)
        references.append(row['answer'])

    bleu = corpus_bleu([ref for ref in references], predictions)
    return bleu, predictions, references

# Run evaluation
bleu_score, preds, refs = evaluate_generator(gen, train_df, doc_dict)
print(f"Initial BLEU: {bleu_score:.4f}")
for p, r in zip(preds, refs):
    print(f"\nReference: {r}\nPredicted: {p}")

#!pip install peft accelerate

from peft import LoraConfig, get_peft_model
from transformers import AutoModelForCausalLM

```

```

# Reload model for training
model = AutoModelForCausalLM.from_pretrained(
    "TinyLlama/TinyLlama-1.1B-Chat-v1.0",
    torch_dtype=torch.float16,
    device_map="auto"
)

# Add LoRA
peft_config = LoraConfig(
    r=8,
    target_modules=["q_proj", "v_proj"],
    task_type="CAUSAL_LM"
)
model = get_peft_model(model, peft_config)

retriever_predictions = pd.read_csv('stage_1_prediction.csv') # we load the
↳ predictions from Retriever model for 900 test questions

def clean_ticks(passage_ids):
    return [str(x).replace('\n', '') for x in passage_ids.strip('[]').split(',')]
↳ ]

retriever_predictions['relevant_passage_ids'] =
↳ retriever_predictions['relevant_passage_ids'].apply(clean_ticks)

# Batch processing (add this new cell)
def batch_predict(df, batch_size=4):
    results = []
    for i in range(0, len(df), batch_size):
        batch = df.iloc[i:i+batch_size]
        answers = gen(
            [f"Q: {q}\nA:" for q in batch["question"]],
            max_new_tokens=64,
            do_sample=False
        )
        results.extend([a[0]['generated_text'].split("A:")[1].strip() for a in
↳ answers])
    df["answer"] = results
    return df

result_df = batch_predict(retriever_predictions)

result_df
result_df["answer"] = result_df["answer"].fillna("").astype(str)
result_df.loc[result_df['answer']=='', 'answer'] = 'No answer'
result_df.to_csv('ssome_baseline.csv', index=None)

```

DEBUG:urllib3.connectionpool:Resetting dropped connection: huggingface.co

```

2025-04-28 18:46:02,451 - urllib3.connectionpool - DEBUG - Resetting dropped
connection: huggingface.co

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD
/TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/config.json HTTP/1.1" 200 0

2025-04-28 18:46:02,788 - urllib3.connectionpool - DEBUG -
https://huggingface.co:443 "HEAD
/TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/config.json HTTP/1.1" 200 0

DEBUG:filelock:Attempting to acquire lock 138326671828176 on
/root/.cache/huggingface/hub/.locks/models--TinyLlama--
TinyLlama-1.1B-Chat-v1.0/4ea05f1bc289d48ba9b92eea2f58ad8acd3dce5d.lock

2025-04-28 18:46:02,796 - filelock - DEBUG - Attempting to acquire lock
138326671828176 on /root/.cache/huggingface/hub/.locks/models--TinyLlama--
TinyLlama-1.1B-Chat-v1.0/4ea05f1bc289d48ba9b92eea2f58ad8acd3dce5d.lock

DEBUG:filelock:Lock 138326671828176 acquired on
/root/.cache/huggingface/hub/.locks/models--TinyLlama--
TinyLlama-1.1B-Chat-v1.0/4ea05f1bc289d48ba9b92eea2f58ad8acd3dce5d.lock

2025-04-28 18:46:02,801 - filelock - DEBUG - Lock 138326671828176 acquired on
/root/.cache/huggingface/hub/.locks/models--TinyLlama--
TinyLlama-1.1B-Chat-v1.0/4ea05f1bc289d48ba9b92eea2f58ad8acd3dce5d.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "GET
/TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/config.json HTTP/1.1" 200 608

2025-04-28 18:46:03,038 - urllib3.connectionpool - DEBUG -
https://huggingface.co:443 "GET
/TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/config.json HTTP/1.1" 200 608

config.json: 0%|          | 0.00/608 [00:00<?, ?B/s]

DEBUG:filelock:Attempting to release lock 138326671828176 on
/root/.cache/huggingface/hub/.locks/models--TinyLlama--
TinyLlama-1.1B-Chat-v1.0/4ea05f1bc289d48ba9b92eea2f58ad8acd3dce5d.lock

2025-04-28 18:46:03,088 - filelock - DEBUG - Attempting to release lock
138326671828176 on /root/.cache/huggingface/hub/.locks/models--TinyLlama--
TinyLlama-1.1B-Chat-v1.0/4ea05f1bc289d48ba9b92eea2f58ad8acd3dce5d.lock

DEBUG:filelock:Lock 138326671828176 released on
/root/.cache/huggingface/hub/.locks/models--TinyLlama--
TinyLlama-1.1B-Chat-v1.0/4ea05f1bc289d48ba9b92eea2f58ad8acd3dce5d.lock

2025-04-28 18:46:03,090 - filelock - DEBUG - Lock 138326671828176 released on
/root/.cache/huggingface/hub/.locks/models--TinyLlama--
TinyLlama-1.1B-Chat-v1.0/4ea05f1bc289d48ba9b92eea2f58ad8acd3dce5d.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD
/TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/adapters_config.json HTTP/1.1"
404 0

```

```

2025-04-28 18:46:03,418 - urllib3.connectionpool - DEBUG -
https://huggingface.co:443 "HEAD
/TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/adapters_config.json HTTP/1.1"
404 0

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD
/TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/model.safetensors HTTP/1.1" 302
0

2025-04-28 18:46:03,692 - urllib3.connectionpool - DEBUG -
https://huggingface.co:443 "HEAD
/TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/model.safetensors HTTP/1.1" 302
0

DEBUG:filelock:Attempting to acquire lock 138326364379728 on
/root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-
v1.0/6e6001da2106d4757498752a021df6c2bdc332c650aae4bae6b0c004dcf14933.lock

2025-04-28 18:46:03,701 - filelock - DEBUG - Attempting to acquire lock
138326364379728 on /root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLl
ama-1.1B-Chat-
v1.0/6e6001da2106d4757498752a021df6c2bdc332c650aae4bae6b0c004dcf14933.lock

DEBUG:filelock:Lock 138326364379728 acquired on
/root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-
v1.0/6e6001da2106d4757498752a021df6c2bdc332c650aae4bae6b0c004dcf14933.lock

2025-04-28 18:46:03,708 - filelock - DEBUG - Lock 138326364379728 acquired on
/root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-
v1.0/6e6001da2106d4757498752a021df6c2bdc332c650aae4bae6b0c004dcf14933.lock

DEBUG:urllib3.connectionpool:Starting new HTTPS connection (1): cdn-lfs-
us-1.hf.co:443

2025-04-28 18:46:03,722 - urllib3.connectionpool - DEBUG - Starting new HTTPS
connection (1): cdn-lfs-us-1.hf.co:443

DEBUG:urllib3.connectionpool:https://cdn-lfs-us-1.hf.co:443 "GET /repos/2b/64/2b
642798915fc368e7b638986f68446b121c1d59b30075e146bd6312ee664ac2/6e6001da2106d4757
498752a021df6c2bdc332c650aae4bae6b0c004dcf14933?response-content-disposition=inl
ine%3B+filename%3DUTF-
8%27%27model.safetensors%3B+filename%3D%22model.safetensors%22%3B&Expires=174586
9563&Policy=eyJTdGF0ZWl1bnQiOlt7IkNvbmlRpdGlvdjI6eyJEYXR1TGZzc1RoYW4iOnsiQVdTOkVw
b2NoVGltZSI6MTc0NTg2OTU2M319LCJSZXNvdXJjZSI6Imh0dHBzOi8vY2RuLWxmcj11cy0xLmhmLmNv
L3JlCG9zLzJlLzY0LzJiNjJyNjQyZmMzNjhlN2I2Mzg5ODZmNjg0NDZiMTIxYzFkNTliMzAwNzVl
MTQ2YmQ2MzEyZWU2NjRhYzIvNmU2MDAxZGEyMTA2ZDQ3NTc0OTg3NTJhMDIxZGY2YzJiZGMzMzJjNjUw
YWF1NGJhZTZhMGMwMDRkY2YxNDkzMzYyZmMzNjhlN2I2Mzg5ODZmNjg0NDZiMTIxYzFkNTliMzAwNzVl
nature=c0tG8eN3UIZTjcjlAin1e2LmLWlsZGT8dW4KNCWb8rdXfBuvpziVP18Yh~9uDAFxfv90fhv2QD
utU43suU2BLTk8GINXFGDhN5rDdypNaybr66MM8DvGceNESH4PKdcRJ1xliwTw65jx6hUaiZFDr-
gIo9Q9zIZL7z4IsI-
EiZm6aPPg81X5REvbK6s0qOkcdgqBiHUIuDPJ3Ug~KAlHViE0fN8rV2bEpfUpLDv2cIUGcTnxAU-Cws6

```

IH6XgheAIjh7DvKvf96BIsJFqofAnL~LNW1XCSPHURSWwSkwVKqMWZL94p6CA8d906qZNq6k4dckbUp  
jpCaK98GkIpaILJA\_\_&Key-Pair-Id=K24J24Z295AEI9 HTTP/1.1" 200 2200119864

2025-04-28 18:46:03,745 - urllib3.connectionpool - DEBUG - https://cdn-lfs-  
us-1.hf.co:443 "GET /repos/2b/64/2b642798915fc368e7b638986f68446b121c1d59b30075e  
146bd6312ee664ac2/6e6001da2106d4757498752a021df6c2bdc332c650aae4bae6b0c004dcf149  
33?response-content-disposition=inline%3B+filename\*%3DUTF-  
8%27%27model.safetensors%3B+filename%3D%22model.safetensors%22%3B&Expires=174586  
9563&Policy=eyJTdGF0ZW11bnQiOi0t7IkNvbmlRpdGlubiI6eyJEYXRlTGZvc1RoYW4iOi0nsiQVdT0kVw  
b2NoVGltZSI6MTc0NTg2OTU2M319LCSZXNvdXJjZSI6Imh0dHBzOi8vY2RuLWxmcyc1cy0xLmhmLmNv  
L3JlcG9zLzJiLzY0LzJiNjQyNzk4OTE1ZmMzNjhlN2I2Mzg5ODZmNjg0NDZiMTIxYzFkNTliMzAwNzVl  
MTQ2YmQ2MzEyZWU2NjRhYzIvNmU2MDAxZGEyMTA2ZDQ3NTc0OTg3NTJhMDIxZGY2YzJiZGMzMzJjNjUw  
YWFlNGJhZTZiMGwmdRkY2YxNDkzMz9yZXNwb25zZS1jb250ZW50LWRpc3Bvc2l0aW9uPSoifV19&Sig  
nature=c0tG8eN3UIZTjcjlAin1e2LmLWlsZGT8dW4KNCWb8rdXfBuvpziVP18Yh~9uDAFxxv90fhv2QD  
utU43suU2BLTk8GINXFGDhN5rDdypNaybr66MM8DvGceNESH4PKdcRJ1xliwTw65jx6hUaiZFDr-  
gIo9Q9zIZL7z4IsI-  
EiZm6aPPg81X5REvbK6s0q0kcdgqBiHUIuDPJ3Ug~KA1HViE0fN8rV2bEpFUpLDv2cIUGcTnxAU-Cws6  
IH6XgheAIjh7DvKvf96BIsJFqofAnL~LNW1XCSPHURSWwSkwVKqMWZL94p6CA8d906qZNq6k4dckbUp  
jpCaK98GkIpaILJA\_\_&Key-Pair-Id=K24J24Z295AEI9 HTTP/1.1" 200 2200119864

model.safetensors: 0%| | 0.00/2.20G [00:00<?, ?B/s]

DEBUG:filelock:Attempting to release lock 138326364379728 on  
/root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-  
v1.0/6e6001da2106d4757498752a021df6c2bdc332c650aae4bae6b0c004dcf14933.lock

2025-04-28 18:46:24,873 - filelock - DEBUG - Attempting to release lock  
138326364379728 on /root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLl  
ama-1.1B-Chat-  
v1.0/6e6001da2106d4757498752a021df6c2bdc332c650aae4bae6b0c004dcf14933.lock

DEBUG:filelock:Lock 138326364379728 released on  
/root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-  
v1.0/6e6001da2106d4757498752a021df6c2bdc332c650aae4bae6b0c004dcf14933.lock

2025-04-28 18:46:24,876 - filelock - DEBUG - Lock 138326364379728 released on  
/root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-  
v1.0/6e6001da2106d4757498752a021df6c2bdc332c650aae4bae6b0c004dcf14933.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD  
/TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/generation\_config.json  
HTTP/1.1" 200 0

2025-04-28 18:46:27,772 - urllib3.connectionpool - DEBUG -  
https://huggingface.co:443 "HEAD  
/TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/generation\_config.json  
HTTP/1.1" 200 0

DEBUG:filelock:Attempting to acquire lock 138326363474384 on  
/root/.cache/huggingface/hub/.locks/models--TinyLlama--  
TinyLlama-1.1B-Chat-v1.0/768e95e5b4dada726d7631277a4a87ecbb236f70.lock

2025-04-28 18:46:27,776 - filelock - DEBUG - Attempting to acquire lock 138326363474384 on /root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-v1.0/768e95e5b4dada726d7631277a4a87ecbb236f70.lock

DEBUG:filelock:Lock 138326363474384 acquired on /root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-v1.0/768e95e5b4dada726d7631277a4a87ecbb236f70.lock

2025-04-28 18:46:27,778 - filelock - DEBUG - Lock 138326363474384 acquired on /root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-v1.0/768e95e5b4dada726d7631277a4a87ecbb236f70.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "GET /TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/generation\_config.json HTTP/1.1" 200 124

2025-04-28 18:46:28,015 - urllib3.connectionpool - DEBUG - https://huggingface.co:443 "GET /TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/generation\_config.json HTTP/1.1" 200 124

generation\_config.json: 0%| | 0.00/124 [00:00<?, ?B/s]

DEBUG:filelock:Attempting to release lock 138326363474384 on /root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-v1.0/768e95e5b4dada726d7631277a4a87ecbb236f70.lock

2025-04-28 18:46:28,034 - filelock - DEBUG - Attempting to release lock 138326363474384 on /root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-v1.0/768e95e5b4dada726d7631277a4a87ecbb236f70.lock

DEBUG:filelock:Lock 138326363474384 released on /root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-v1.0/768e95e5b4dada726d7631277a4a87ecbb236f70.lock

2025-04-28 18:46:28,048 - filelock - DEBUG - Lock 138326363474384 released on /root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-v1.0/768e95e5b4dada726d7631277a4a87ecbb236f70.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD /TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/tokenizer\_config.json HTTP/1.1" 200 0

2025-04-28 18:46:28,284 - urllib3.connectionpool - DEBUG - https://huggingface.co:443 "HEAD /TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/tokenizer\_config.json HTTP/1.1" 200 0

DEBUG:filelock:Attempting to acquire lock 138326362361168 on /root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-v1.0/fa96b85858f4053b0142a18e2b09dbe94e3fae46.lock

2025-04-28 18:46:28,287 - filelock - DEBUG - Attempting to acquire lock 138326362361168 on /root/.cache/huggingface/hub/.locks/models--TinyLlama--

TinyLlama-1.1B-Chat-v1.0/fa96b85858f4053b0142a18e2b09dbe94e3fae46.lock

DEBUG:filelock:Lock 138326362361168 acquired on  
 /root/.cache/huggingface/hub/.locks/models--TinyLlama--  
 TinyLlama-1.1B-Chat-v1.0/fa96b85858f4053b0142a18e2b09dbe94e3fae46.lock

2025-04-28 18:46:28,289 - filelock - DEBUG - Lock 138326362361168 acquired on  
 /root/.cache/huggingface/hub/.locks/models--TinyLlama--  
 TinyLlama-1.1B-Chat-v1.0/fa96b85858f4053b0142a18e2b09dbe94e3fae46.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "GET  
 /TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/tokenizer\_config.json HTTP/1.1"  
 200 1289

2025-04-28 18:46:28,519 - urllib3.connectionpool - DEBUG -  
 https://huggingface.co:443 "GET  
 /TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/tokenizer\_config.json HTTP/1.1"  
 200 1289

tokenizer\_config.json: 0%| | 0.00/1.29k [00:00<?, ?B/s]

DEBUG:filelock:Attempting to release lock 138326362361168 on  
 /root/.cache/huggingface/hub/.locks/models--TinyLlama--  
 TinyLlama-1.1B-Chat-v1.0/fa96b85858f4053b0142a18e2b09dbe94e3fae46.lock

2025-04-28 18:46:28,539 - filelock - DEBUG - Attempting to release lock  
 138326362361168 on /root/.cache/huggingface/hub/.locks/models--TinyLlama--  
 TinyLlama-1.1B-Chat-v1.0/fa96b85858f4053b0142a18e2b09dbe94e3fae46.lock

DEBUG:filelock:Lock 138326362361168 released on  
 /root/.cache/huggingface/hub/.locks/models--TinyLlama--  
 TinyLlama-1.1B-Chat-v1.0/fa96b85858f4053b0142a18e2b09dbe94e3fae46.lock

2025-04-28 18:46:28,543 - filelock - DEBUG - Lock 138326362361168 released on  
 /root/.cache/huggingface/hub/.locks/models--TinyLlama--  
 TinyLlama-1.1B-Chat-v1.0/fa96b85858f4053b0142a18e2b09dbe94e3fae46.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD  
 /TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/tokenizer.model HTTP/1.1" 302 0

2025-04-28 18:46:28,938 - urllib3.connectionpool - DEBUG -  
 https://huggingface.co:443 "HEAD  
 /TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/tokenizer.model HTTP/1.1" 302 0

DEBUG:filelock:Attempting to acquire lock 138326363689168 on  
 /root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-  
 v1.0/9e556afd44213b6bd1be2b850ebbbd98f5481437a8021afaf58ee7fb1818d347.lock

2025-04-28 18:46:28,942 - filelock - DEBUG - Attempting to acquire lock  
 138326363689168 on /root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLl  
 ama-1.1B-Chat-  
 v1.0/9e556afd44213b6bd1be2b850ebbbd98f5481437a8021afaf58ee7fb1818d347.lock

DEBUG:filelock:Lock 138326363689168 acquired on  
/root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-  
v1.0/9e556afd44213b6bd1be2b850ebbbd98f5481437a8021afaf58ee7fb1818d347.lock

2025-04-28 18:46:28,945 - filelock - DEBUG - Lock 138326363689168 acquired on  
/root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-  
v1.0/9e556afd44213b6bd1be2b850ebbbd98f5481437a8021afaf58ee7fb1818d347.lock

DEBUG:urllib3.connectionpool:https://cdn-lfs-us-1.hf.co:443 "GET /repos/2b/64/2b  
642798915fc368e7b638986f68446b121c1d59b30075e146bd6312ee664ac2/9e556afd44213b6bd  
1be2b850ebbbd98f5481437a8021afaf58ee7fb1818d347?response-content-disposition=inl  
ine%3B+filename%3DUTF-  
8%27%27tokenizer.model%3B+filename%3D%22tokenizer.model%22%3B&Expires=1745869588  
&Policy=eyJTdGF0ZW1lbnQiOlt7IkNvbmlRdGlvdjI6eyJEYXRlTGZvc1RoYW4iOnsiQVdTOkVwb2No  
VGltZSI6MTcONTg2OTU4OH19LCSZSNvdXJjZSI6Imh0dHBzOi8vY2RuLWxmc1cyOxLmhmLmNvL3Jl  
cG9zLzJiLzY0LzJiNjQyNzk4OTE1ZmMzNjhlN2I2Mzg5ODZmNjgONDZiMTIxYzFkNTliMzAwNzVlMTQ2  
YmQ2MzEyZWU2NjRhYzIvOWU1NTZhZmQNDI2YmQxYmUyYjg1MGViYmJkOThmNTQ4MTQzN2E4MDIx  
YWZhZjU4ZWU3ZmIxODE4ZDMONz9yZXNwb25zZS1jb250ZW50LWRpc3Bvc2l0aW9uPSoifV19&Signatu  
re=PpwK1UrW9WpAk7jYhoJQxzD5Ui9bvBpx1DSb08mDxRtUD1T18eVSnb0K1qmDF-CvU~eJM14qGoXue  
5fEqTfqooSXrrEPGa95uMK43lZgCHR4Npdnr7ht7m8X~4F-  
iQNDOR6RY8XTkCmb0dX61lIPkPGRWs2GNA7YzYpq53IohkC-LOHHAE4ENx32a2rVLfMu6o3vHMqMFMH1  
n3HY1lH15T5M8iWoqgURx7Zgrd0~BDdiJkMseEI0sDg9LVxJ4nkNTQ~P4jLWbZxYShHzrcG7BXWagOQK  
j2Gvgn4x6k1o2dmiBAffgkM~fjNwthuqAseKazPWLmg2ULCyQCxnhYnuPA\_\_&Key-Pair-  
Id=K24J24Z295AEI9 HTTP/1.1" 200 499723

2025-04-28 18:46:28,955 - urllib3.connectionpool - DEBUG - https://cdn-lfs-  
us-1.hf.co:443 "GET /repos/2b/64/2b642798915fc368e7b638986f68446b121c1d59b30075e  
146bd6312ee664ac2/9e556afd44213b6bd1be2b850ebbbd98f5481437a8021afaf58ee7fb1818d3  
47?response-content-disposition=inline%3B+filename%3DUTF-  
8%27%27tokenizer.model%3B+filename%3D%22tokenizer.model%22%3B&Expires=1745869588  
&Policy=eyJTdGF0ZW1lbnQiOlt7IkNvbmlRdGlvdjI6eyJEYXRlTGZvc1RoYW4iOnsiQVdTOkVwb2No  
VGltZSI6MTcONTg2OTU4OH19LCSZSNvdXJjZSI6Imh0dHBzOi8vY2RuLWxmc1cyOxLmhmLmNvL3Jl  
cG9zLzJiLzY0LzJiNjQyNzk4OTE1ZmMzNjhlN2I2Mzg5ODZmNjgONDZiMTIxYzFkNTliMzAwNzVlMTQ2  
YmQ2MzEyZWU2NjRhYzIvOWU1NTZhZmQNDI2YmQxYmUyYjg1MGViYmJkOThmNTQ4MTQzN2E4MDIx  
YWZhZjU4ZWU3ZmIxODE4ZDMONz9yZXNwb25zZS1jb250ZW50LWRpc3Bvc2l0aW9uPSoifV19&Signatu  
re=PpwK1UrW9WpAk7jYhoJQxzD5Ui9bvBpx1DSb08mDxRtUD1T18eVSnb0K1qmDF-CvU~eJM14qGoXue  
5fEqTfqooSXrrEPGa95uMK43lZgCHR4Npdnr7ht7m8X~4F-  
iQNDOR6RY8XTkCmb0dX61lIPkPGRWs2GNA7YzYpq53IohkC-LOHHAE4ENx32a2rVLfMu6o3vHMqMFMH1  
n3HY1lH15T5M8iWoqgURx7Zgrd0~BDdiJkMseEI0sDg9LVxJ4nkNTQ~P4jLWbZxYShHzrcG7BXWagOQK  
j2Gvgn4x6k1o2dmiBAffgkM~fjNwthuqAseKazPWLmg2ULCyQCxnhYnuPA\_\_&Key-Pair-  
Id=K24J24Z295AEI9 HTTP/1.1" 200 499723

tokenizer.model: 0%| | 0.00/500k [00:00<?, ?B/s]

DEBUG:filelock:Attempting to release lock 138326363689168 on  
/root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-  
v1.0/9e556afd44213b6bd1be2b850ebbbd98f5481437a8021afaf58ee7fb1818d347.lock

2025-04-28 18:46:28,989 - filelock - DEBUG - Attempting to release lock  
138326363689168 on /root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLl  
ama-1.1B-Chat-



```

v1.0/9e556afd44213b6bd1be2b850ebbbd98f5481437a8021afaf58ee7fb1818d347.lock
DEBUG:filelock:Lock 138326363689168 released on
/root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-
v1.0/9e556afd44213b6bd1be2b850ebbbd98f5481437a8021afaf58ee7fb1818d347.lock
2025-04-28 18:46:28,992 - filelock - DEBUG - Lock 138326363689168 released on
/root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-
v1.0/9e556afd44213b6bd1be2b850ebbbd98f5481437a8021afaf58ee7fb1818d347.lock
DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD
/TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/tokenizer.json HTTP/1.1" 200 0
2025-04-28 18:46:29,228 - urllib3.connectionpool - DEBUG -
https://huggingface.co:443 "HEAD
/TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/tokenizer.json HTTP/1.1" 200 0
DEBUG:filelock:Attempting to acquire lock 138326363363472 on
/root/.cache/huggingface/hub/.locks/models--TinyLlama--
TinyLlama-1.1B-Chat-v1.0/a6e931b92caff4c79c5c56282f1e89569a0ae558.lock
2025-04-28 18:46:29,232 - filelock - DEBUG - Attempting to acquire lock
138326363363472 on /root/.cache/huggingface/hub/.locks/models--TinyLlama--
TinyLlama-1.1B-Chat-v1.0/a6e931b92caff4c79c5c56282f1e89569a0ae558.lock
DEBUG:filelock:Lock 138326363363472 acquired on
/root/.cache/huggingface/hub/.locks/models--TinyLlama--
TinyLlama-1.1B-Chat-v1.0/a6e931b92caff4c79c5c56282f1e89569a0ae558.lock
2025-04-28 18:46:29,235 - filelock - DEBUG - Lock 138326363363472 acquired on
/root/.cache/huggingface/hub/.locks/models--TinyLlama--
TinyLlama-1.1B-Chat-v1.0/a6e931b92caff4c79c5c56282f1e89569a0ae558.lock
DEBUG:urllib3.connectionpool:https://huggingface.co:443 "GET
/TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/tokenizer.json HTTP/1.1" 200
1842767
2025-04-28 18:46:29,473 - urllib3.connectionpool - DEBUG -
https://huggingface.co:443 "GET
/TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/tokenizer.json HTTP/1.1" 200
1842767
tokenizer.json: 0%|          | 0.00/1.84M [00:00<?, ?B/s]
DEBUG:filelock:Attempting to release lock 138326363363472 on
/root/.cache/huggingface/hub/.locks/models--TinyLlama--
TinyLlama-1.1B-Chat-v1.0/a6e931b92caff4c79c5c56282f1e89569a0ae558.lock
2025-04-28 18:46:29,716 - filelock - DEBUG - Attempting to release lock
138326363363472 on /root/.cache/huggingface/hub/.locks/models--TinyLlama--
TinyLlama-1.1B-Chat-v1.0/a6e931b92caff4c79c5c56282f1e89569a0ae558.lock
DEBUG:filelock:Lock 138326363363472 released on
/root/.cache/huggingface/hub/.locks/models--TinyLlama--
TinyLlama-1.1B-Chat-v1.0/a6e931b92caff4c79c5c56282f1e89569a0ae558.lock

```

2025-04-28 18:46:29,719 - filelock - DEBUG - Lock 138326363363472 released on /root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-v1.0/a6e931b92caff4c79c5c56282f1e89569a0ae558.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD /TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/added\_tokens.json HTTP/1.1" 404 0

2025-04-28 18:46:29,964 - urllib3.connectionpool - DEBUG - https://huggingface.co:443 "HEAD /TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/added\_tokens.json HTTP/1.1" 404 0

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD /TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/special\_tokens\_map.json HTTP/1.1" 200 0

2025-04-28 18:46:30,197 - urllib3.connectionpool - DEBUG - https://huggingface.co:443 "HEAD /TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/special\_tokens\_map.json HTTP/1.1" 200 0

DEBUG:filelock:Attempting to acquire lock 138326364767568 on /root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-v1.0/492d4b2966a1763442d426d880dbc29f94906e4c.lock

2025-04-28 18:46:30,200 - filelock - DEBUG - Attempting to acquire lock 138326364767568 on /root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-v1.0/492d4b2966a1763442d426d880dbc29f94906e4c.lock

DEBUG:filelock:Lock 138326364767568 acquired on /root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-v1.0/492d4b2966a1763442d426d880dbc29f94906e4c.lock

2025-04-28 18:46:30,202 - filelock - DEBUG - Lock 138326364767568 acquired on /root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-v1.0/492d4b2966a1763442d426d880dbc29f94906e4c.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "GET /TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/special\_tokens\_map.json HTTP/1.1" 200 551

2025-04-28 18:46:30,434 - urllib3.connectionpool - DEBUG - https://huggingface.co:443 "GET /TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/special\_tokens\_map.json HTTP/1.1" 200 551

special\_tokens\_map.json: 0% | 0.00/551 [00:00<?, ?B/s]

DEBUG:filelock:Attempting to release lock 138326364767568 on /root/.cache/huggingface/hub/.locks/models--TinyLlama--TinyLlama-1.1B-Chat-v1.0/492d4b2966a1763442d426d880dbc29f94906e4c.lock

```

2025-04-28 18:46:30,457 - filelock - DEBUG - Attempting to release lock
138326364767568 on /root/.cache/huggingface/hub/.locks/models--TinyLlama--
TinyLlama-1.1B-Chat-v1.0/492d4b2966a1763442d426d880dbc29f94906e4c.lock

DEBUG:filelock:Lock 138326364767568 released on
/root/.cache/huggingface/hub/.locks/models--TinyLlama--
TinyLlama-1.1B-Chat-v1.0/492d4b2966a1763442d426d880dbc29f94906e4c.lock

2025-04-28 18:46:30,459 - filelock - DEBUG - Lock 138326364767568 released on
/root/.cache/huggingface/hub/.locks/models--TinyLlama--
TinyLlama-1.1B-Chat-v1.0/492d4b2966a1763442d426d880dbc29f94906e4c.lock

DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD
/TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/chat_template.jinja HTTP/1.1"
404 0

2025-04-28 18:46:30,707 - urllib3.connectionpool - DEBUG -
https://huggingface.co:443 "HEAD
/TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/chat_template.jinja HTTP/1.1"
404 0

Device set to use cuda
Token indices sequence length is longer than the specified maximum sequence
length for this model (7480 > 2048). Running this sequence through the model
will result in indexing errors
This is a friendly reminder - the current text generation call will exceed the
model's predefined maximum length (2048). Depending on the model, you may
observe exceptions, performance degradation, or nothing at all.

Initial BLEU: 0.1964

Reference: E3-ubiquitin ligase activity is the only known enzymatic activity of
BRCA1, which is mediated by the N-terminal RING finger domain. BRCA1 nuclear
transport and ubiquitin E3 ligase enzymatic activity are tightly regulated by
the BRCA1 dimeric binding partner BARD1 and further modulated by cancer
mutations and diverse signaling pathways.
Predicted: The enzymatic activity of the breast cancer associated gene BRCA1 is
critical for the central functions of BRCA1.

Reference: The cytoplasmic polyadenylation element (CPE) is the binding platform
for CPE-binding protein (CPEB), which promotes polyadenylation-induced
translation.
Predicted: CPEB3 binds to the CPE domain of mRNA, which is a common domain
organization, involving two consecutive RNA recognition

Reference: Bosch-Boonstra-Schaaf optic atrophy syndrome (BBSOAS) is an
autosomal-dominant disorder characterized by optic atrophy and intellectual
disability caused by loss-of-function mutations in NR2F1.
Predicted: NR2F1 variants cause a range of diseases, including neurofibromatosis
type 1 (NF1), neurofibrom

```

Reference: Gaucher's disease is caused by deficient lysosomal glucocerebrosidase activity

Predicted:

cesocesociblececececisociblecececececisociblececececiseucesociblececisolecisoces

Reference: Davunetide or NAP is a microtubule-stabilizer.

Predicted: Davunetide inhibits microtubule assembly and disassembly, leading to the accumulation of microtubules in the cytoplasm

```
DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD
/TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/config.json HTTP/1.1" 200 0
```

```
2025-04-28 18:46:38,658 - urllib3.connectionpool - DEBUG -
https://huggingface.co:443 "HEAD
/TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/config.json HTTP/1.1" 200 0
```

```
INFO:accelerate.utils.modeling:We will use 90% of the memory on device 0 for
storing the model, and 10% for the buffer to avoid OOM. You can set `max_memory`
in to a higher value to use more memory (at your own risk).
```

```
2025-04-28 18:46:38,681 - accelerate.utils.modeling - INFO - We will use 90% of
the memory on device 0 for storing the model, and 10% for the buffer to avoid
OOM. You can set `max_memory` in to a higher value to use more memory (at your
own risk).
```

```
DEBUG:urllib3.connectionpool:https://huggingface.co:443 "HEAD
/TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/generation_config.json
HTTP/1.1" 200 0
```

```
2025-04-28 18:46:39,907 - urllib3.connectionpool - DEBUG -
https://huggingface.co:443 "HEAD
/TinyLlama/TinyLlama-1.1B-Chat-v1.0/resolve/main/generation_config.json
HTTP/1.1" 200 0
```

You seem to be using the pipelines sequentially on GPU. In order to maximize efficiency please use a dataset

[ ]:

[ ]:

[ ]:

[ ]:

[ ]:

[ ]:

[ ]:

[ ]:

[ ]:

[ ]: