

[ ![Logo](../../\_static/logo.png) ](../../index.html)

## Getting Started

- \* [Installation](../../docs/installation.html)

- \* [Install with pip](../../docs/installation.html#install-with-pip)

- \* [Install with Conda](../../docs/installation.html#install-with-conda)

- \* [Install from Source](../../docs/installation.html#install-from-source)

- \* [Editable Install](../../docs/installation.html#editable-install)

- \* [Install PyTorch with CUDA support](../../docs/installation.html#install-pytorch-with-cuda-support)

- \* [Quickstart](../../docs/quickstart.html)

- \* [Sentence Transformer](../../docs/quickstart.html#sentence-transformer)

- \* [Cross Encoder](../../docs/quickstart.html#cross-encoder)

- \* [Next Steps](../../docs/quickstart.html#next-steps)

## Sentence Transformer

- \* [Usage](../../docs/sentence\_transformer/usage/usage.html)

- \* [Computing Embeddings](../../applications/computing-embeddings/README.html)

- \* [Initializing a Sentence Transformer Model](../../applications/computing-embeddings/README.html#initializing-a-sentence-transformer-model)

- \* [Calculating Embeddings](../../applications/computing-embeddings/README.html#calculating-embeddings)

- \* [Prompt Templates](../../applications/computing-embeddings/README.html#prompt-templates)

- \* [Input Sequence Length](../../applications/computing-embeddings/README.html#id1)

\* [Multi-Process / Multi-GPU

Encoding](../../applications/computing-embeddings/README.html#multi-process-multi-gpu-encoding)

\* [Semantic Textual

Similarity](../../docs/sentence\_transformer/usage/semantic\_textual\_similarity.html)

\* [Similarity

Calculation](../../docs/sentence\_transformer/usage/semantic\_textual\_similarity.html#similarity-calculation)

\* [Semantic Search](../../applications/semantic-search/README.html)

\* [Background](../../applications/semantic-search/README.html#background)

\* [Symmetric vs. Asymmetric Semantic

Search](../../applications/semantic-search/README.html#symmetric-vs-asymmetric-semantic-search)

\* [Manual

Implementation](../../applications/semantic-search/README.html#manual-implementation)

\* [Optimized

Implementation](../../applications/semantic-search/README.html#optimized-implementation)

\* [Speed Optimization](../../applications/semantic-search/README.html#speed-optimization)

\* [Elasticsearch](../../applications/semantic-search/README.html#elasticsearch)

\* [Approximate Nearest

Neighbor](../../applications/semantic-search/README.html#approximate-nearest-neighbor)

\* [Retrieve & Re-Rank](../../applications/semantic-search/README.html#retrieve-re-rank)

\* [Examples](../../applications/semantic-search/README.html#examples)

\* [Retrieve & Re-Rank](../../applications/retrieve\_rerank/README.html)

\* [Retrieve & Re-Rank

Pipeline](../../applications/retrieve\_rerank/README.html#retrieve-re-rank-pipeline)

\* [Retrieval: Bi-Encoder](../../applications/retrieve\_rerank/README.html#retrieval-bi-encoder)

\* [Re-Ranker:

Cross-Encoder](../../applications/retrieve\_rerank/README.html#re-ranker-cross-encoder)

\* [Example Scripts](../../applications/retrieve\_rerank/README.html#example-scripts)

\* [Pre-trained Bi-Encoders

(Retrieval)](../../applications/retrieve\_rerank/README.html#pre-trained-bi-encoders-retrieval)

\* [Pre-trained Cross-Encoders

(Re-Ranker)](../../applications/retrieve\_rerank/README.html#pre-trained-cross-encoders-re-ranker)

\* [Clustering](../../applications/clustering/README.html)

\* [k-Means](../../applications/clustering/README.html#k-means)

\* [Agglomerative Clustering](../../applications/clustering/README.html#agglomerative-clustering)

\* [Fast Clustering](../../applications/clustering/README.html#fast-clustering)

\* [Topic Modeling](../../applications/clustering/README.html#topic-modeling)

\* [Paraphrase Mining](../../applications/paraphrase-mining/README.html)

\*

[`paraphrase\_mining()'](../../applications/paraphrase-mining/README.html#sentence\_transformers.  
util.paraphrase\_mining)

\* [Translated Sentence Mining](../../applications/parallel-sentence-mining/README.html)

\* [Margin Based

Mining](../../applications/parallel-sentence-mining/README.html#margin-based-mining)

\* [Examples](../../applications/parallel-sentence-mining/README.html#examples)

\* [Image Search](../../applications/image-search/README.html)

\* [Installation](../../applications/image-search/README.html#installation)

\* [Usage](../../applications/image-search/README.html#usage)

\* [Examples](../../applications/image-search/README.html#examples)

\* [Embedding Quantization](../../applications/embedding-quantization/README.html)

\* [Binary

Quantization](../../applications/embedding-quantization/README.html#binary-quantization)

[Quantization\]\(../../applications/embedding-quantization/README.html#scalar-int8-quantization\)](#)

[\[Additional extensions\]\(../../applications/embedding-quantization/README.html#additional-extensions\)](#)

- \* [\[Demo\]\(../../applications/embedding-quantization/README.html#demo\)](#)
- \* [\[Try it yourself\]\(../../applications/embedding-quantization/README.html#try-it-yourself\)](#)
- \* [\[Speeding up Inference\]\(../../docs/sentence\\_transformer/usage/efficiency.html\)](#)
- \* [\[PyTorch\]\(../../docs/sentence\\_transformer/usage/efficiency.html#pytorch\)](#)
- \* [\[ONNX\]\(../../docs/sentence\\_transformer/usage/efficiency.html#onnx\)](#)
- \* [\[OpenVINO\]\(../../docs/sentence\\_transformer/usage/efficiency.html#openvino\)](#)
- \* [\[Benchmarks\]\(../../docs/sentence\\_transformer/usage/efficiency.html#benchmarks\)](#)
- \* [\[Creating Custom Models\]\(../../docs/sentence\\_transformer/usage/custom\\_models.html\)](#)

[\\* \[Structure of Sentence Transformer Models\]\(../../docs/sentence\\_transformer/usage/custom\\_models.html#structure-of-sentence-transformer-models\)](#)

[\\* \[Sentence Transformer Model from a Transformers Model\]\(../../docs/sentence\\_transformer/usage/custom\\_models.html#sentence-transformer-model-from-a-transformers-model\)](#)

- \* [\[Pretrained Models\]\(../../docs/sentence\\_transformer/pretrained\\_models.html\)](#)
- \* [\[Original Models\]\(../../docs/sentence\\_transformer/pretrained\\_models.html#original-models\)](#)

[\\* \[Semantic Search Models\]\(../../docs/sentence\\_transformer/pretrained\\_models.html#semantic-search-models\)](#)

- \* [\[Multi-QA Models\]\(../../docs/sentence\\_transformer/pretrained\\_models.html#multi-qa-models\)](#)

[\\* \[MSMARCO Passage Models\]\(../../docs/sentence\\_transformer/pretrained\\_models.html#msmarco-passage-models\)](#)

[\\* \[Multilingual Models\]\(../../docs/sentence\\_transformer/pretrained\\_models.html#multilingual-models\)](#)

\* [Semantic Similarity Models](../../docs/sentence\_transformer/pretrained\_models.html#semantic-similarity-models)

\* [Bitext Mining](../../docs/sentence\_transformer/pretrained\_models.html#bitext-mining)

\* [Image & Text-Models](../../docs/sentence\_transformer/pretrained\_models.html#image-text-models)

\* [INSTRUCTOR models](../../docs/sentence\_transformer/pretrained\_models.html#instructor-models)

\* [Scientific Similarity Models](../../docs/sentence\_transformer/pretrained\_models.html#scientific-similarity-models)

\* [Training Overview](../../docs/sentence\_transformer/training\_overview.html)

\* [Why Finetune?](../../docs/sentence\_transformer/training\_overview.html#why-finetune)

\* [Training Components](../../docs/sentence\_transformer/training\_overview.html#training-components)

\* [Dataset](../../docs/sentence\_transformer/training\_overview.html#dataset)

\* [Dataset Format](../../docs/sentence\_transformer/training\_overview.html#dataset-format)

\* [Loss Function](../../docs/sentence\_transformer/training\_overview.html#loss-function)

\* [Training Arguments](../../docs/sentence\_transformer/training\_overview.html#training-arguments)

\* [Evaluator](../../docs/sentence\_transformer/training\_overview.html#evaluator)

\* [Trainer](../../docs/sentence\_transformer/training\_overview.html#trainer)

\* [Callbacks](../../docs/sentence\_transformer/training\_overview.html#callbacks)

\* [Multi-Dataset Training](../../docs/sentence\_transformer/training\_overview.html#multi-dataset-training)

\* [Deprecated Training](../../docs/sentence\_transformer/training\_overview.html#deprecated-training)

\* [Best Base Embedding Models](../../docs/sentence\_transformer/training\_overview.html#best-base-embedding-models)

- \* [Dataset Overview](../../docs/sentence\_transformer/dataset\_overview.html)
- \* [Datasets on the Hugging Face Hub](../../docs/sentence\_transformer/dataset\_overview.html#datasets-on-the-hugging-face-hub)
- \* [Pre-existing Datasets](../../docs/sentence\_transformer/dataset\_overview.html#pre-existing-datasets)
- \* [Loss Overview](../../docs/sentence\_transformer/loss\_overview.html)
- \* [Loss modifiers](../../docs/sentence\_transformer/loss\_overview.html#loss-modifiers)
- \* [Distillation](../../docs/sentence\_transformer/loss\_overview.html#distillation)
- \* [Commonly used Loss Functions](../../docs/sentence\_transformer/loss\_overview.html#commonly-used-loss-functions)
- \* [Custom Loss Functions](../../docs/sentence\_transformer/loss\_overview.html#custom-loss-functions)
- \* [Training Examples](../../docs/sentence\_transformer/training/examples.html)
- \* [Semantic Textual Similarity](../sts/README.html)
- \* [Training data](../sts/README.html#training-data)
- \* [Loss Function](../sts/README.html#loss-function)
- \* [Natural Language Inference](../nli/README.html)
- \* [Data](../nli/README.html#data)
- \* [SoftmaxLoss](../nli/README.html#softmaxloss)
- \* [MultipleNegativesRankingLoss](../nli/README.html#multiplenegativesrankingloss)
- \* Paraphrase Data
- \* Pre-Trained Models
- \* [Quora Duplicate Questions](../quora\_duplicate\_questions/README.html)
- \* [Training](../quora\_duplicate\_questions/README.html#training)
- \*
- [MultipleNegativesRankingLoss](../quora\_duplicate\_questions/README.html#multiplenegativesrankingloss)

- \* [Pretrained Models](../quora\_duplicate\_questions/README.html#pretrained-models)
- \* [MS MARCO](../ms\_marco/README.html)
  - \* [Bi-Encoder](../ms\_marco/README.html#bi-encoder)
- \* [Matryoshka Embeddings](../matryoshka/README.html)
  - \* [Use Cases](../matryoshka/README.html#use-cases)
  - \* [Results](../matryoshka/README.html#results)
  - \* [Training](../matryoshka/README.html#training)
  - \* [Inference](../matryoshka/README.html#inference)
  - \* [Code Examples](../matryoshka/README.html#code-examples)
- \* [Adaptive Layers](../adaptive\_layer/README.html)
  - \* [Use Cases](../adaptive\_layer/README.html#use-cases)
  - \* [Results](../adaptive\_layer/README.html#results)
  - \* [Training](../adaptive\_layer/README.html#training)
  - \* [Inference](../adaptive\_layer/README.html#inference)
  - \* [Code Examples](../adaptive\_layer/README.html#code-examples)
- \* [Multilingual Models](../multilingual/README.html)
  - \* [Extend your own models](../multilingual/README.html#extend-your-own-models)
  - \* [Training](../multilingual/README.html#training)
  - \* [Datasets](../multilingual/README.html#datasets)
  - \* [Sources for Training Data](../multilingual/README.html#sources-for-training-data)
  - \* [Evaluation](../multilingual/README.html#evaluation)
  - \* [Available Pre-trained Models](../multilingual/README.html#available-pre-trained-models)
  - \* [Usage](../multilingual/README.html#usage)
  - \* [Performance](../multilingual/README.html#performance)
  - \* [Citation](../multilingual/README.html#citation)
- \* [Model Distillation](../distillation/README.html)
  - \* [Knowledge Distillation](../distillation/README.html#knowledge-distillation)

- \* [Speed - Performance Trade-Off](../distillation/README.html#speed-performance-trade-off)
- \* [Dimensionality Reduction](../distillation/README.html#dimensionality-reduction)
- \* [Quantization](../distillation/README.html#quantization)
- \* [Augmented SBERT](../data\_augmentation/README.html)
- \* [Motivation](../data\_augmentation/README.html#motivation)
  - \* [Extend to your own datasets](../data\_augmentation/README.html#extend-to-your-own-datasets)
  - \* [Methodology](../data\_augmentation/README.html#methodology)
    - \* [Scenario 1: Limited or small annotated datasets (few labeled sentence-pairs)](../data\_augmentation/README.html#scenario-1-limited-or-small-annotated-dataset-s-few-labeled-sentence-pairs)
    - \* [Scenario 2: No annotated datasets (Only unlabeled sentence-pairs)](../data\_augmentation/README.html#scenario-2-no-annotated-datasets-only-unlabeled-sentence-pairs)
  - \* [Training](../data\_augmentation/README.html#training)
  - \* [Citation](../data\_augmentation/README.html#citation)
- \* [Training with Prompts](../prompts/README.html)
  - \* [What are Prompts?](../prompts/README.html#what-are-prompts)
    - \* [Why would we train with Prompts?](../prompts/README.html#why-would-we-train-with-prompts)
    - \* [How do we train with Prompts?](../prompts/README.html#how-do-we-train-with-prompts)
- \* [Training with PEFT Adapters](../peft/README.html)
  - \* [Compatibility Methods](../peft/README.html#compatibility-methods)
  - \* [Adding a New Adapter](../peft/README.html#adding-a-new-adapter)
  - \* [Loading a Pretrained Adapter](../peft/README.html#loading-a-pretrained-adapter)
  - \* [Training Script](../peft/README.html#training-script)
- \* [Unsupervised Learning](../unsupervised\_learning/README.html)



\* [TSDAE](../../unsupervised\_learning/README.html#tsdae)

\* [SimCSE](../../unsupervised\_learning/README.html#simcse)

\* [CT](../../unsupervised\_learning/README.html#ct)

\* [CT (In-Batch Negative Sampling)](../../unsupervised\_learning/README.html#ct-in-batch-negative-sampling)

\* [Masked Language Model (MLM)](../../unsupervised\_learning/README.html#masked-language-model-mlm)

\* [GenQ](../../unsupervised\_learning/README.html#genq)

\* [GPL](../../unsupervised\_learning/README.html#gpl)

\* [Performance Comparison](../../unsupervised\_learning/README.html#performance-comparison)

\* [Domain Adaptation](../../domain\_adaptation/README.html)

\* [Domain Adaptation vs. Unsupervised Learning](../../domain\_adaptation/README.html#domain-adaptation-vs-unsupervised-learning)

\* [Adaptive Pre-Training](../../domain\_adaptation/README.html#adaptive-pre-training)

\* [GPL: Generative Pseudo-Labeling](../../domain\_adaptation/README.html#gpl-generative-pseudo-labeling)

\* [Hyperparameter Optimization](../hpo/README.html)

\* [HPO Components](../hpo/README.html#hpo-components)

\* [Putting It All Together](../hpo/README.html#putting-it-all-together)

\* [Example Scripts](../hpo/README.html#example-scripts)

\* [Distributed Training](../../docs/sentence\_transformer/training/distributed.html)

\* [Comparison](../../docs/sentence\_transformer/training/distributed.html#comparison)

\* [FSDP](../../docs/sentence\_transformer/training/distributed.html#fsdp)

Cross Encoder

- \* [Usage](../../docs/cross\_encoder/usage/usage.html)
- \* [Retrieve & Re-Rank](../../applications/retrieve\_rerank/README.html)
  - \* [Retrieve & Re-Rank Pipeline](../../applications/retrieve\_rerank/README.html#retrieve-re-rank-pipeline)
  - \* [Retrieval: Bi-Encoder](../../applications/retrieve\_rerank/README.html#retrieval-bi-encoder)
    - \* [Re-Ranker: Cross-Encoder](../../applications/retrieve\_rerank/README.html#re-ranker-cross-encoder)
  - \* [Example Scripts](../../applications/retrieve\_rerank/README.html#example-scripts)
    - \* [Pre-trained Bi-Encoders (Retrieval)](../../applications/retrieve\_rerank/README.html#pre-trained-bi-encoders-retrieval)
    - \* [Pre-trained Cross-Encoders (Re-Ranker)](../../applications/retrieve\_rerank/README.html#pre-trained-cross-encoders-re-ranker)
- \* [Pretrained Models](../../docs/cross\_encoder/pretrained\_models.html)
  - \* [MS MARCO](../../docs/cross\_encoder/pretrained\_models.html#ms-marco)
  - \* [SQuAD (QNLI)](../../docs/cross\_encoder/pretrained\_models.html#squad-qnli)
  - \* [STSbenchmark](../../docs/cross\_encoder/pretrained\_models.html#stsbenchmark)
    - \* [Quora Duplicate Questions](../../docs/cross\_encoder/pretrained\_models.html#quora-duplicate-questions)
  - \* [NLI](../../docs/cross\_encoder/pretrained\_models.html#nli)
  - \* [Community Models](../../docs/cross\_encoder/pretrained\_models.html#community-models)
- \* [Training Overview](../../docs/cross\_encoder/training\_overview.html)
- \* [Training Examples](../../docs/cross\_encoder/training/examples.html)
- \* [MS MARCO](../ms\_marco/cross\_encoder\_README.html)
  - \* [Cross-Encoder](../ms\_marco/cross\_encoder\_README.html#cross-encoder)
    - \* [Cross-Encoder Knowledge Distillation](../ms\_marco/cross\_encoder\_README.html#cross-encoder-knowledge-distillation)

## Package Reference

\* [Sentence Transformer](../../docs/package\_reference/sentence\_transformer/index.html)

\*

[SentenceTransformer](../../docs/package\_reference/sentence\_transformer/SentenceTransformer.html)

\*

[SentenceTransformer](../../docs/package\_reference/sentence\_transformer/SentenceTransformer.html#id1)

\*

[SentenceTransformerModelCardData](../../docs/package\_reference/sentence\_transformer/SentenceTransformer.html#sentencetransformermodelcarddata)

\*

[SimilarityFunction](../../docs/package\_reference/sentence\_transformer/SentenceTransformer.html#similarityfunction)

\* [Trainer](../../docs/package\_reference/sentence\_transformer/trainer.html)

\*

[SentenceTransformerTrainer](../../docs/package\_reference/sentence\_transformer/trainer.html#sentencetransformertrainer)

\* [Training Arguments](../../docs/package\_reference/sentence\_transformer/training\_args.html)

\*

[SentenceTransformerTrainingArguments](../../docs/package\_reference/sentence\_transformer/training\_args.html#sentencetransformertrainingarguments)

\* [Losses](../../docs/package\_reference/sentence\_transformer/losses.html)

\*

[BatchAllTripletLoss](../../docs/package\_reference/sentence\_transformer/losses.html#batchalltripletloss)

\*

[BatchHardSoftMarginTripletLoss](../../docs/package\_reference/sentence\_transformer/losses.html#batchhardsoftmargintripletloss)

\*

[BatchHardTripletLoss](../../docs/package\_reference/sentence\_transformer/losses.html#batchhardtripletloss)

\*

[BatchSemiHardTripletLoss](../../docs/package\_reference/sentence\_transformer/losses.html#batchsemi-hardtripletloss)

\*

[ContrastiveLoss](../../docs/package\_reference/sentence\_transformer/losses.html#contrastiveloss)

\*

[OnlineContrastiveLoss](../../docs/package\_reference/sentence\_transformer/losses.html#onlinecontrastiveloss)

\*

[ContrastiveTensionLoss](../../docs/package\_reference/sentence\_transformer/losses.html#contrastivetensionloss)

\*

[ContrastiveTensionLossInBatchNegatives](../../docs/package\_reference/sentence\_transformer/losses.html#contrastivetensionlossinbatchnegatives)

\* [CoSENTLoss](../../docs/package\_reference/sentence\_transformer/losses.html#cosentloss)

\* [AngleLoss](../../docs/package\_reference/sentence\_transformer/losses.html#angleloss)

\*

[CosineSimilarityLoss](../../docs/package\_reference/sentence\_transformer/losses.html#cosinesimilarityloss)

\*

[DenoisingAutoEncoderLoss](../../docs/package\_reference/sentence\_transformer/losses.html#denoisingautoencoderloss)

osingautoencoderloss)

\*

[GISTEmbedLoss](../../docs/package\_reference/sentence\_transformer/losses.html#gistembedloss  
)

\*

[CachedGISTEmbedLoss](../../docs/package\_reference/sentence\_transformer/losses.html#cachedgistembedloss)

\* [MSELoss](../../docs/package\_reference/sentence\_transformer/losses.html#mseloss)

\*

[MarginMSELoss](../../docs/package\_reference/sentence\_transformer/losses.html#marginmseloss  
)

\*

[MatryoshkaLoss](../../docs/package\_reference/sentence\_transformer/losses.html#matryoshkaloss  
)

\*

[Matryoshka2dLoss](../../docs/package\_reference/sentence\_transformer/losses.html#matryoshka2dloss)

\*

[AdaptiveLayerLoss](../../docs/package\_reference/sentence\_transformer/losses.html#adaptivelayerloss)

\*

[MegaBatchMarginLoss](../../docs/package\_reference/sentence\_transformer/losses.html#megabatchmarginloss)

\*

[MultipleNegativesRankingLoss](../../docs/package\_reference/sentence\_transformer/losses.html#multiplenegativesrankingloss)

\*

[CachedMultipleNegativesRankingLoss](../../docs/package\_reference/sentence\_transformer/losses.html#cachedmultiplenegativesrankingloss)

\*

[MultipleNegativesSymmetricRankingLoss](../../docs/package\_reference/sentence\_transformer/losses.html#multiplenegativessymmetricrankingloss)

\*

[CachedMultipleNegativesSymmetricRankingLoss](../../docs/package\_reference/sentence\_transformer/losses.html#cachedmultiplenegativessymmetricrankingloss)

\* [SoftmaxLoss](../../docs/package\_reference/sentence\_transformer/losses.html#softmaxloss)

\* [TripletLoss](../../docs/package\_reference/sentence\_transformer/losses.html#tripletloss)

\* [Samplers](../../docs/package\_reference/sentence\_transformer/sampler.html)

\*

[BatchSamplers](../../docs/package\_reference/sentence\_transformer/sampler.html#batchsamplers)

\*

[MultiDatasetBatchSamplers](../../docs/package\_reference/sentence\_transformer/sampler.html#multidatasetbatchsamplers)

\* [Evaluation](../../docs/package\_reference/sentence\_transformer/evaluation.html)

\*

[BinaryClassificationEvaluator](../../docs/package\_reference/sentence\_transformer/evaluation.html#binaryclassificationevaluator)

\*

[EmbeddingSimilarityEvaluator](../../docs/package\_reference/sentence\_transformer/evaluation.html#embeddingsimilarityevaluator)

\*

[InformationRetrievalEvaluator](../../docs/package\_reference/sentence\_transformer/evaluation.html#informationretrievalevaluator)

\*

[NanoBEIREvaluator](../../docs/package\_reference/sentence\_transformer/evaluation.html#nanobe  
irevaluator)

\*

[MSEEvaluator](../../docs/package\_reference/sentence\_transformer/evaluation.html#mseevaluator  
)

\*

[ParaphraseMiningEvaluator](../../docs/package\_reference/sentence\_transformer/evaluation.html#  
paraphraseminingevaluator)

\*

[RerankingEvaluator](../../docs/package\_reference/sentence\_transformer/evaluation.html#rerankin  
gevaluator)

\*

[SentenceEvaluator](../../docs/package\_reference/sentence\_transformer/evaluation.html#sentenc  
eevaluator)

\*

[SequentialEvaluator](../../docs/package\_reference/sentence\_transformer/evaluation.html#sequen  
tiaevaluator)

\*

[TranslationEvaluator](../../docs/package\_reference/sentence\_transformer/evaluation.html#translat  
ionevaluator)

\*

[TripletEvaluator](../../docs/package\_reference/sentence\_transformer/evaluation.html#tripletevalua  
tor)

\* [Datasets](../../docs/package\_reference/sentence\_transformer/datasets.html)

\*

[ParallelSentencesDataset](../../docs/package\_reference/sentence\_transformer/datasets.html#par

allelsentencesdataset)

\*

[SentenceLabelDataset](../../docs/package\_reference/sentence\_transformer/datasets.html#sentence-label-dataset)

\*

[DenoisingAutoEncoderDataset](../../docs/package\_reference/sentence\_transformer/datasets.html#denoising-auto-encoder-dataset)

\*

[NoDuplicatesDataLoader](../../docs/package\_reference/sentence\_transformer/datasets.html#no-duplicates-data-loader)

\* [Models](../../docs/package\_reference/sentence\_transformer/models.html)

\*

[Main

Classes](../../docs/package\_reference/sentence\_transformer/models.html#main-classes)

\*

[Further

Classes](../../docs/package\_reference/sentence\_transformer/models.html#further-classes)

\* [quantization](../../docs/package\_reference/sentence\_transformer/quantization.html)

\*

[`quantize\_embeddings()`](../../docs/package\_reference/sentence\_transformer/quantization.html#sentence-transformers.quantization.quantize\_embeddings)

\*

[`semantic\_search\_faiss()`](../../docs/package\_reference/sentence\_transformer/quantization.html#sentence-transformers.quantization.semantic\_search\_faiss)

\*

[`semantic\_search\_usearch()`](../../docs/package\_reference/sentence\_transformer/quantization.html#sentence-transformers.quantization.semantic\_search\_usearch)

\* [Cross Encoder](../../docs/package\_reference/cross\_encoder/index.html)

\* [CrossEncoder](../../docs/package\_reference/cross\_encoder/cross\_encoder.html)



- \* [CrossEncoder](../../docs/package\_reference/cross\_encoder/cross\_encoder.html#id1)
- \* [Training Inputs](../../docs/package\_reference/cross\_encoder/cross\_encoder.html#training-inputs)
- \* [Evaluation](../../docs/package\_reference/cross\_encoder/evaluation.html)
- \* [CEBinaryAccuracyEvaluator](../../docs/package\_reference/cross\_encoder/evaluation.html#cebinaryaccuracyevaluator)
- \* [CEBinaryClassificationEvaluator](../../docs/package\_reference/cross\_encoder/evaluation.html#cebinaryclassificationevaluator)
- \* [CECorrelationEvaluator](../../docs/package\_reference/cross\_encoder/evaluation.html#cecorrelationevaluator)
- \* [CEF1Evaluator](../../docs/package\_reference/cross\_encoder/evaluation.html#cef1evaluator)
- \* [CESoftmaxAccuracyEvaluator](../../docs/package\_reference/cross\_encoder/evaluation.html#cesoftmaxaccuracyevaluator)
- \* [CERerankingEvaluator](../../docs/package\_reference/cross\_encoder/evaluation.html#cererankingevaluator)
- \* [util](../../docs/package\_reference/util.html)
- \* [Helper Functions](../../docs/package\_reference/util.html#module-sentence\_transformers.util)
- \* [community\_detection()](../../docs/package\_reference/util.html#sentence\_transformers.util.community\_detection)
- \* [http\_get()](../../docs/package\_reference/util.html#sentence\_transformers.util.http\_get)

[`is\_training\_available()`](../../docs/package\_reference/util.html#sentence\_transformers.util.is\_training\_available)

\*

[`mine\_hard\_negatives()`](../../docs/package\_reference/util.html#sentence\_transformers.util.mine\_hard\_negatives)

\*

[`normalize\_embeddings()`](../../docs/package\_reference/util.html#sentence\_transformers.util.normalize\_embeddings)

\*

[`paraphrase\_mining()`](../../docs/package\_reference/util.html#sentence\_transformers.util.paraphrase\_mining)

\*

[`semantic\_search()`](../../docs/package\_reference/util.html#sentence\_transformers.util.semantic\_search)

\*

[`truncate\_embeddings()`](../../docs/package\_reference/util.html#sentence\_transformers.util.truncate\_embeddings)

\*

[Model

Optimization](../../docs/package\_reference/util.html#module-sentence\_transformers.backend)

\*

[`export\_dynamic\_quantized\_onnx\_model()`](../../docs/package\_reference/util.html#sentence\_transformers.backend.export\_dynamic\_quantized\_onnx\_model)

\*

[`export\_optimized\_onnx\_model()`](../../docs/package\_reference/util.html#sentence\_transformers.backend.export\_optimized\_onnx\_model)

\*

[`export\_static\_quantized\_openvino\_model()`](../../docs/package\_reference/util.html#sentence\_tra

nsformers.backend.export\_static\_quantized\_openvino\_model)

\* [Similarity Metrics](../../docs/package\_reference/util.html#module-sentence\_transformers.util)

\* [`cos_sim()`](../../docs/package\_reference/util.html#sentence\_transformers.util.cos\_sim)

\* [`dot_score()`](../../docs/package\_reference/util.html#sentence\_transformers.util.dot\_score)

\*

[`euclidean_sim()`](../../docs/package\_reference/util.html#sentence\_transformers.util.euclidean\_sim)

\*

[`manhattan_sim()`](../../docs/package\_reference/util.html#sentence\_transformers.util.manhattan\_sim)

\*

[`pairwise_cos_sim()`](../../docs/package\_reference/util.html#sentence\_transformers.util.pairwise\_cos\_sim)

\*

[`pairwise_dot_score()`](../../docs/package\_reference/util.html#sentence\_transformers.util.pairwise\_dot\_score)

\*

[`pairwise_euclidean_sim()`](../../docs/package\_reference/util.html#sentence\_transformers.util.pairwise\_euclidean\_sim)

\*

[`pairwise_manhattan_sim()`](../../docs/package\_reference/util.html#sentence\_transformers.util.pairwise\_manhattan\_sim)

\_\_[Sentence Transformers](../../index.html)

\* [(../../index.html)](../../index.html)

\* [Training Examples](../../docs/sentence\_transformer/training/examples.html)

\* Paraphrase Data

\*

[

Edit

on

GitHub](<https://github.com/UKPLab/sentence-transformers/blob/master/examples/training/paraphrases/README.md>)

\* \* \*

# Paraphrase Data

In our paper [Making Monolingual Sentence Embeddings Multilingual using Knowledge Distillation](<https://arxiv.org/abs/2004.09813>), we showed that paraphrase data together with

[`MultipleNegativesRankingLoss`]([../docs/package\\_reference/sentence\\_transformer/losses.html](https://github.com/UKPLab/sentence-transformers/blob/master/docs/package_reference/sentence_transformer/losses.html)

`#sentence_transformers.losses.MultipleNegativesRankingLoss`

"`sentence_transformers.losses.MultipleNegativesRankingLoss`") is a powerful

combination to learn sentence embeddings models. Read [NLI >

`MultipleNegativesRankingLoss`]([../nli/README.html#multiplenegativesrankingloss](https://github.com/UKPLab/sentence-transformers/blob/master/docs/nli/README.html#multiplenegativesrankingloss))

for more information on this loss function.

The [training.py]([https://github.com/UKPLab/sentence-](https://github.com/UKPLab/sentence-transformers/blob/master/examples/training/paraphrases/training.py)

`transformers/tree/master/examples/training/paraphrases/training.py`) script

loads various datasets from the [Dataset

Overview]([../docs/sentence\\_transformer/dataset\\_overview.html#pre-](https://github.com/UKPLab/sentence-transformers/blob/master/docs/sentence_transformer/dataset_overview.html#pre-existing-datasets)

`existing-datasets`). We construct batches by sampling examples from the

respective dataset. So far, examples are not mixed between the datasets, i.e.,

a batch consists only of examples from a single dataset.

As the dataset sizes are quite different in size, we perform [round-robin sampling](../../docs/package\_reference/sentence\_transformer/sampler.html#sentence\_transformer.s.training\_args.MultiDatasetBatchSamplers) to train using the same amount of batches from each dataset.

## ## Pre-Trained Models

Have a look at [pre-trained models](../../docs/sentence\_transformer/pretrained\_models.html) to view all models that were trained on these paraphrase datasets.

\*

[paraphrase-MiniLM-L12-v2](https://huggingface.co/sentence-transformers/paraphrase-MiniLM-L12-v2) \- Trained on the following datasets: AllNLI, sentence-compression, SimpleWiki, altlex, msmarco-triplets, quora\_duplicates, coco\_captions, flickr30k\_captions, yahoo\_answers\_title\_question, S2ORC\_citation\_pairs, stackexchange\_duplicate\_questions, wiki-atomic-edits

\*

[paraphrase-distilroberta-base-v2](https://huggingface.co/sentence-transformers/paraphrase-distilroberta-base-v2) \- Trained on the following datasets: AllNLI, sentence-compression, SimpleWiki, altlex, msmarco-triplets, quora\_duplicates, coco\_captions, flickr30k\_captions, yahoo\_answers\_title\_question, S2ORC\_citation\_pairs, stackexchange\_duplicate\_questions, wiki-atomic-edits

\*

[paraphrase-distilroberta-base-v1](https://huggingface.co/sentence-transformers/paraphrase-distilroberta-base-v1)

berta-base-v1) \- Trained on the following datasets: AllNLI, sentence-compression, SimpleWiki, altlex, quora\_duplicates, wiki-atomic-edits, wiki-split

\*

[paraphrase-xlm-r-multilingual-v1](https://huggingface.co/sentence-transformers/paraphrase-xlm-r-multilingual-v1) \- Multilingual version of paraphrase-distilroberta-base-v1, trained on parallel data for 50+ languages. (Teacher:

[paraphrase-distilroberta-base-v1](https://huggingface.co/sentence-transformers/paraphrase-distilroberta-base-v1), Student: [xlm-r-base](https://huggingface.co/FacebookAI/xlm-roberta-base))

[ Previous](../nli/README.html "Natural Language Inference") [Next  
](../quora\_duplicate\_questions/README.html "Quora Duplicate Questions")

\* \* \*

(C) Copyright 2025.

Built with [Sphinx](https://www.sphinx-doc.org/) using a  
[theme](https://github.com/readthedocs/sphinx\_rtd\_theme) provided by [Read the  
Docs](https://readthedocs.org).