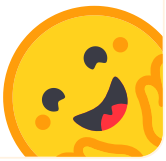


Transformers

State-of-the-art Machine Learning for PyTorch, TensorFlow, and JAX.

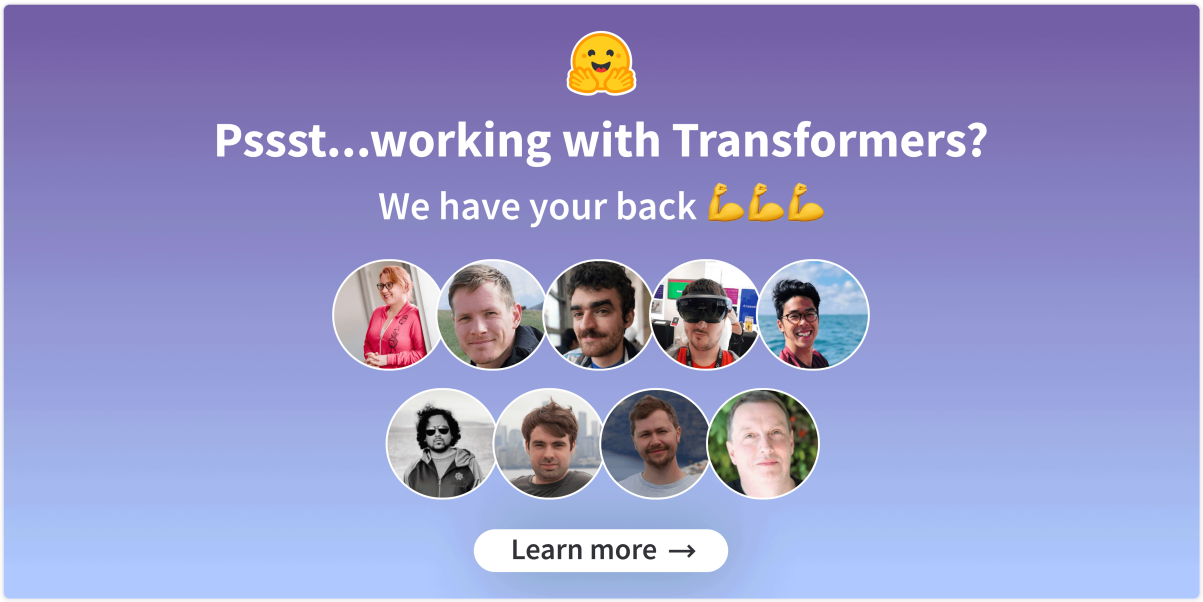
Transformers provides APIs and tools to easily download and train state-of-the-art pretrained models. Using pretrained models can reduce your compute costs, carbon footprint, and save you the time and resources required to train a model from scratch. These models support common tasks in different modalities, such as:

Natural Language Processing: text classification, named entity recognition, question answering, language modeling, summarization, translation, multiple choice, and text generation.
Computer Vision: image classification, object detection, and segmentation.
Audio: automatic speech recognition and audio classification.
Multimodal: table question answering, optical character recognition, information extraction from scanned documents, video classification, and visual question answering.

Transformers support framework interoperability between PyTorch, TensorFlow, and JAX. This provides the flexibility to use a different framework at each stage of a model's life; train a model in three lines of code in one framework, and load it for inference in another. Models can also be exported to a format like ONNX and TorchScript for deployment in production environments.

Join the growing community on the Hub, forum, or Discord today!

If you are looking for custom support from the Hugging Face team



Contents

The documentation is organized into five sections:

- **GET STARTED** provides a quick tour of the library and installation instructions to get up and running.
- **TUTORIALS** are a great place to start if you're a beginner. This section will help you gain the basic skills you need to start using the library.
- **HOW-TO GUIDES** show you how to achieve a specific goal, like finetuning a pretrained model for language modeling or how to write and share a custom model.
- **CONCEPTUAL GUIDES** offers more discussion and explanation of the underlying concepts and ideas behind models, tasks, and the design philosophy of 🤗 Transformers.
- **API** describes all classes and functions:
    - **MAIN CLASSES** details the most important classes like configuration, model, tokenizer, and pipeline.

- **MODELS** details the classes and functions related to each model implemented in the library.
- **INTERNAL HELPERS** details utility classes and functions used internally.

Supported models and frameworks

The table below represents the current support in the library for each of those models, whether they have a Python tokenizer (called "slow"). A "fast" tokenizer backed by the 🤗 Tokenizers library, whether they have support in Jax (via Flax), PyTorch, and/or TensorFlow.

Model	PyTorch support	TensorFlow support	Flax Support
ALBERT	✅	✅	✅
ALIGN	✅	❌	❌
AltCLIP	✅	❌	❌
Audio Spectrogram Transformer	✅	❌	❌
Autoformer	✅	❌	❌
Bark	✅	❌	❌
BART	✅	✅	✅
BARThez	✅	✅	✅
BARTpho	✅	✅	✅
BEiT	✅	❌	✅
BERT	✅	✅	✅
Bert Generation	✅	❌	❌
BertJapanese	✅	✅	✅
BERTweet	✅	✅	✅

Model	PyTorch support	TensorFlow support	Flax Support
BigBird	✅	❌	✅
BigBird-Pegasus	✅	❌	❌
BioGpt	✅	❌	❌
BiT	✅	❌	❌
Blenderbot	✅	✅	✅
BlenderbotSmall	✅	✅	✅
BLIP	✅	✅	❌
BLIP-2	✅	❌	❌
BLOOM	✅	❌	✅
BORT	✅	✅	✅
BridgeTower	✅	❌	❌
BROS	✅	❌	❌
ByT5	✅	✅	✅
CamemBERT	✅	✅	❌
CANINE	✅	❌	❌
Chinese-CLIP	✅	❌	❌
CLAP	✅	❌	❌
CLIP	✅	✅	✅
CLIPSeg	✅	❌	❌
CLVP	✅	❌	❌
CodeGen	✅	❌	❌
CodeLlama	✅	❌	✅

Model	PyTorch support	TensorFlow support	Flax Support
Cohere	✅	❌	❌
Conditional DETR	✅	❌	❌
ConvBERT	✅	✅	❌
ConvNeXT	✅	✅	❌
ConvNeXTV2	✅	✅	❌
CPM	✅	✅	✅
CPM-Ant	✅	❌	❌
CTRL	✅	✅	❌
CvT	✅	✅	❌
Data2VecAudio	✅	❌	❌
Data2VecText	✅	❌	❌
Data2VecVision	✅	✅	❌
DBRX	✅	❌	❌
DeBERTa	✅	✅	❌
DeBERTa-v2	✅	✅	❌
Decision Transformer	✅	❌	❌
Deformable DETR	✅	❌	❌
DeiT	✅	✅	❌
DePlot	✅	❌	❌
Depth Anything	✅	❌	❌
DETA	✅	❌	❌
DETR	✅	❌	❌

Model	PyTorch support	TensorFlow support	Flax Support
DialoGPT	✅	✅	✅
DiNAT	✅	❌	❌
DINOv2	✅	❌	❌
DistilBERT	✅	✅	✅
DiT	✅	❌	✅
DonutSwin	✅	❌	❌
DPR	✅	✅	❌
DPT	✅	❌	❌
EfficientFormer	✅	✅	❌
EfficientNet	✅	❌	❌
ELECTRA	✅	✅	✅
EnCodec	✅	❌	❌
Encoder decoder	✅	✅	✅
ERNIE	✅	❌	❌
ErnieM	✅	❌	❌
ESM	✅	✅	❌
FairSeq Machine-Translation	✅	❌	❌
Falcon	✅	❌	❌
FastSpeech2Conformer	✅	❌	❌
FLAN-T5	✅	✅	✅
FLAN-UL2	✅	✅	✅
FlauBERT	✅	✅	❌

Model	PyTorch support	TensorFlow support	Flax Support
FLAVA	✅	❌	❌
FNet	✅	❌	❌
FocalNet	✅	❌	❌
Funnel Transformer	✅	✅	❌
Fuyu	✅	❌	❌
Gemma	✅	❌	✅
GIT	✅	❌	❌
GLPN	✅	❌	❌
GPT Neo	✅	❌	✅
GPT NeoX	✅	❌	❌
GPT NeoX Japanese	✅	❌	❌
GPT-J	✅	✅	✅
GPT-Sw3	✅	✅	✅
GPTBigCode	✅	❌	❌
GPTSAN-japanese	✅	❌	❌
Graphormer	✅	❌	❌
Grounding DINO	✅	❌	❌
GroupViT	✅	✅	❌
HerBERT	✅	✅	✅
Hubert	✅	✅	❌
I-BERT	✅	❌	❌
IDEFICS	✅	❌	❌

Model	PyTorch support	TensorFlow support	Flax Support
Idefics2	✅	❌	❌
ImageGPT	✅	❌	❌
Informer	✅	❌	❌
InstructBLIP	✅	❌	❌
Jamba	✅	❌	❌
Jukebox	✅	❌	❌
KOSMOS-2	✅	❌	❌
LayoutLM	✅	✅	❌
LayoutLMv2	✅	❌	❌
LayoutLMv3	✅	✅	❌
LayoutXLM	✅	❌	❌
LED	✅	✅	❌
LeViT	✅	❌	❌
LiLT	✅	❌	❌
LLaMA	✅	❌	✅
Llama2	✅	❌	✅
LLaVa	✅	❌	❌
LLaVA-NeXT	✅	❌	❌
Longformer	✅	✅	❌
LongT5	✅	❌	✅
LUKE	✅	❌	❌
LXMERT	✅	✅	❌

Model	PyTorch support	TensorFlow support	Flax Support
M-CTC-T	✅	❌	❌
M2M100	✅	❌	❌
MADLAD-400	✅	✅	✅
Mamba	✅	❌	❌
Marian	✅	✅	✅
MarkupLM	✅	❌	❌
Mask2Former	✅	❌	❌
MaskFormer	✅	❌	❌
MatCha	✅	❌	❌
mBART	✅	✅	✅
mBART-50	✅	✅	✅
MEGA	✅	❌	❌
Megatron-BERT	✅	❌	❌
Megatron-GPT2	✅	✅	✅
MGP-STR	✅	❌	❌
Mistral	✅	❌	✅
Mixtral	✅	❌	❌
mLUKE	✅	❌	❌
MMS	✅	✅	✅
MobileBERT	✅	✅	❌
MobileNetV1	✅	❌	❌
MobileNetV2	✅	❌	❌

Model	PyTorch support	TensorFlow support	Flax Support
MobileViT	✅	✅	❌
MobileViTV2	✅	❌	❌
MPNet	✅	✅	❌
MPT	✅	❌	❌
MRA	✅	❌	❌
MT5	✅	✅	✅
MusicGen	✅	❌	❌
MusicGen Melody	✅	❌	❌
MVP	✅	❌	❌
NAT	✅	❌	❌
Nezha	✅	❌	❌
NLLB	✅	❌	❌
NLLB-MOE	✅	❌	❌
Nougat	✅	✅	✅
Nyströmformer	✅	❌	❌
OLMo	✅	❌	❌
OneFormer	✅	❌	❌
OpenAI GPT	✅	✅	❌
OpenAI GPT-2	✅	✅	✅
OpenLlama	✅	❌	❌
OPT	✅	✅	✅
OWL-ViT	✅	❌	❌

Model	PyTorch support	TensorFlow support	Flax Support
OWLv2	✅	❌	❌
PatchTSMixer	✅	❌	❌
PatchTST	✅	❌	❌
Pegasus	✅	✅	✅
PEGASUS-X	✅	❌	❌
Perceiver	✅	❌	❌
Persimmon	✅	❌	❌
Phi	✅	❌	❌
PhoBERT	✅	✅	✅
Pix2Struct	✅	❌	❌
PLBart	✅	❌	❌
PoolFormer	✅	❌	❌
Pop2Piano	✅	❌	❌
ProphetNet	✅	❌	❌
PVT	✅	❌	❌
PVTv2	✅	❌	❌
QDQBert	✅	❌	❌
Qwen2	✅	❌	❌
Qwen2MoE	✅	❌	❌
RAG	✅	✅	❌
REALM	✅	❌	❌
RecurrentGemma	✅	❌	❌

Model	PyTorch support	TensorFlow support	Flax Support
Reformer	✅	❌	❌
RegNet	✅	✅	✅
RemBERT	✅	✅	❌
ResNet	✅	✅	✅
RetriBERT	✅	❌	❌
RoBERTa	✅	✅	✅
RoBERTa-PreLayerNorm	✅	✅	✅
RoCBert	✅	❌	❌
RoFormer	✅	✅	✅
RWKV	✅	❌	❌
SAM	✅	✅	❌
SeamlessM4T	✅	❌	❌
SeamlessM4Tv2	✅	❌	❌
SegFormer	✅	✅	❌
SegGPT	✅	❌	❌
SEW	✅	❌	❌
SEW-D	✅	❌	❌
SigLIP	✅	❌	❌
Speech Encoder decoder	✅	❌	✅
Speech2Text	✅	✅	❌
SpeechT5	✅	❌	❌
Splinter	✅	❌	❌

Model	PyTorch support	TensorFlow support	Flax Support
SqueezeBERT	✅	❌	❌
StableLm	✅	❌	❌
Starcoder2	✅	❌	❌
SuperPoint	✅	❌	❌
SwiftFormer	✅	❌	❌
Swin Transformer	✅	✅	❌
Swin Transformer V2	✅	❌	❌
Swin2SR	✅	❌	❌
SwitchTransformers	✅	❌	❌
T5	✅	✅	✅
T5v1.1	✅	✅	✅
Table Transformer	✅	❌	❌
TAPAS	✅	✅	❌
TAPEX	✅	✅	✅
Time Series Transformer	✅	❌	❌
TimeSformer	✅	❌	❌
Trajectory Transformer	✅	❌	❌
Transformer-XL	✅	✅	❌
TrOCR	✅	❌	❌
TVLT	✅	❌	❌
TVP	✅	❌	❌
UDOP	✅	❌	❌

Model	PyTorch support	TensorFlow support	Flax Support
UL2	✅	✅	✅
UMT5	✅	❌	❌
UniSpeech	✅	❌	❌
UniSpeechSat	✅	❌	❌
UnivNet	✅	❌	❌
UPerNet	✅	❌	❌
VAN	✅	❌	❌
VideoMAE	✅	❌	❌
ViLT	✅	❌	❌
VipLlava	✅	❌	❌
Vision Encoder decoder	✅	✅	✅
VisionTextDualEncoder	✅	✅	✅
VisualBERT	✅	❌	❌
ViT	✅	✅	✅
ViT Hybrid	✅	❌	❌
VitDet	✅	❌	❌
ViTMAE	✅	✅	❌
ViTMatte	✅	❌	❌
ViTMSN	✅	❌	❌
VITS	✅	❌	❌
ViViT	✅	❌	❌
Wav2Vec2	✅	✅	✅

Model	PyTorch support	TensorFlow support	Flax Support
Wav2Vec2-BERT	✅	❌	❌
Wav2Vec2-Conformer	✅	❌	❌
Wav2Vec2Phoneme	✅	✅	✅
WavLM	✅	❌	❌
Whisper	✅	✅	✅
X-CLIP	✅	❌	❌
X-MOD	✅	❌	❌
XGLM	✅	✅	✅
XLM	✅	✅	❌
XLM-ProphetNet	✅	❌	❌
XLM-RoBERTa	✅	✅	✅
XLM-RoBERTa-XL	✅	❌	❌
XLM-V	✅	✅	✅
XLNet	✅	✅	❌
XLS-R	✅	✅	✅
XLSR-Wav2Vec2	✅	✅	✅
YOLOS	✅	❌	❌
YOSO	✅	❌	❌

Quick tour →