



Transformers

build **passing** license Apache-2.0 website **online** release v4.21.2

Contributor Covenant v2.0 adopted DOI 10.5281/zenodo.7019769

[English](#) | [简体中文](#) | [繁體中文](#) | [한국어](#)

为 Jax、PyTorch 和 TensorFlow 打造的先进的自然语言处理



Part of the Hugging Face course!

🧡 Transformers 提供了数以千计的预训练模型，支持 100 多种语言的文本分类、信息抽取、问答、摘要、翻译、文本生成。它的宗旨让最先进的 NLP 技术人人易用。

🧡 Transformers 提供了便于快速下载和使用的 API，让你可以把预训练模型用在给定文本、在你的数据集上微调然后通过 [model hub](#) 与社区共享。同时，每个定义的 Python 模块均完全独立，方便修改和快速研究实验。

🧡 Transformers 支持三个最热门的深度学习库：[Jax](#)、[PyTorch](#) 和 [TensorFlow](#) — 并为之无缝整合。你可以直接使用一个框架训练你的模型然后用另一个加载和推理。

在线演示

你可以直接在模型页面上测试大多数 [model hub](#) 上的模型。我们也提供了 [私有模型托管](#)、[模型版本管理](#) 以及 [推理 API](#)。

这里是一些例子：

- [用 BERT 做掩码填词](#)
- [用 Electra 做命名实体识别](#)
- [用 GPT-2 做文本生成](#)
- [用 RoBERTa 做自然语言推理](#)
- [用 BART 做文本摘要](#)
- [用 DistilBERT 做问答](#)
- [用 T5 做翻译](#)

****Write With Transformer****，由抱抱脸团队打造，是一个文本生成的官方 demo。

如果你在寻找由抱抱脸团队提供的定制化支持服务



Hugging Face **Premium Support**

快速上手

我们为快速使用模型提供了 `pipeline`（流水线）API。流水线聚合了预训练模型和对应的文本预处理。下面是一个快速使用流水线去判断正负面情绪的例子：

```
>>> from transformers import pipeline

# 使用情绪分析流水线
>>> classifier = pipeline('sentiment-analysis')
>>> classifier('We are very happy to introduce pipeline to the transformers repository.')
[{'label': 'POSITIVE', 'score': 0.9996980428695679}]
```

第二行代码下载并缓存了流水线使用的预训练模型，而第三行代码则在给定的文本上进行了评估。这里的答案“正面” (positive) 具有 99 的置信度。

许多的 NLP 任务都有开箱即用的预训练流水线。比如说，我们可以轻松的从给定文本中抽取问题答案：

```
>>> from transformers import pipeline

# 使用问答流水线
>>> question_answerer = pipeline('question-answering')
>>> question_answerer({
...     'question': 'What is the name of the repository ?',
...     'context': 'Pipeline has been included in the huggingface/transformers repository'
... })
{'score': 0.30970096588134766, 'start': 34, 'end': 58, 'answer': 'huggingface/transformers'}
```

除了给出答案，预训练模型还给出了对应的置信度分数、答案在词符化 (tokenized) 后的文本中开始和结束的位置。你可以从[这个教程](#)了解更多流水线API支持的任务。

要在你的任务上下载和使用任意预训练模型也很简单，只需三行代码。这里是 PyTorch 版的示例：

```
>>> from transformers import AutoTokenizer, AutoModel

>>> tokenizer = AutoTokenizer.from_pretrained("bert-base-uncased")
>>> model = AutoModel.from_pretrained("bert-base-uncased")

>>> inputs = tokenizer("Hello world!", return_tensors="pt")
>>> outputs = model(**inputs)
```

这里是等效的 TensorFlow 代码：

```
>>> from transformers import AutoTokenizer, TFAutoModel

>>> tokenizer = AutoTokenizer.from_pretrained("bert-base-uncased")
>>> model = TFAutoModel.from_pretrained("bert-base-uncased")

>>> inputs = tokenizer("Hello world!", return_tensors="tf")
>>> outputs = model(**inputs)
```

词符化器 (tokenizer) 为所有的预训练模型提供了预处理，并可以直接对单个字符串进行调用（比如上面的例子）或对列表 (list) 调用。它会输出一个你可以在下游代码里使用或直接通过 `**` 解包表达式传给模型的词典 (dict)。

模型本身是一个常规的 [Pytorch `nn.Module`](#) 或 [TensorFlow `tf.keras.Model`](#)（取决于你的后端），可以常规方式使用。[这个教程](#) 解释了如何将这样的模型整合到经典的 PyTorch 或 TensorFlow 训练循环中，或是如何使用我们的 `Trainer` 训练器）API 来在一个新的数据集上快速微调。

为什么要用 transformers?

1. 便于使用的先进模型：

- NLU 和 NLG 上表现优越
- 对教学和实践友好且低门槛
- 高级抽象，只需了解三个类
- 对所有模型统一的API

2. 更低计算开销，更少的碳排放：

- 研究人员可以分享亿训练的模型而非次次从头开始训练
- 工程师可以减少计算用时和生产环境开销
- 数十种模型架构、两千多个预训练模型、100多种语言支持

3. 对于模型生命周期的每一个部分都面面俱到：

- 训练先进的模型，只需 3 行代码
- 模型在不同深度学习框架间任意转移，随心所欲
- 为训练、评估和生产选择最适合的框架，衔接无缝

4. 为你的需求轻松定制专属模型和用例：

- 我们为每种模型架构提供了多个用例来复现原论文结果
- 模型内部结构保持透明一致
- 模型文件可单独使用，方便魔改和快速实验

什么情况下我不该用 transformers?

- 本库并不是模块化的神经网络工具箱。模型文件中的代码特意呈若璞玉，未经额外抽象封装，以便研究人员快速迭代魔改而不致溺于抽象和文件跳转之中。

- `Trainer` API 并非兼容任何模型，只为本库之模型优化。若是在寻找适用于通用机器学习的训练循环实现，请另觅他库。
- 尽管我们已尽力而为，[examples 目录](#)中的脚本也仅为用例而已。对于你的特定问题，它们并不一定开箱即用，可能需要改几行代码以适之。

安装

使用 pip

这个仓库已在 Python 3.6+、Flax 0.3.2+、PyTorch 1.3.1+ 和 TensorFlow 2.3+ 下经过测试。

你可以在[虚拟环境](#)中安装 🤖 Transformers。如果你还不熟悉 Python 的虚拟环境，请阅读此[用户说明](#)。

首先，用你打算使用的版本的 Python 创建一个虚拟环境并激活。

然后，你需要安装 Flax、PyTorch 或 TensorFlow 其中之一。关于在你使用的平台上安装这些框架，请参阅 [TensorFlow 安装页](#)、[PyTorch 安装页](#) 或 [Flax 安装页](#)。

当这些后端之一安装成功后，🤖 Transformers 可依此安装：

```
pip install transformers
```

如果你想要试试用例或者想在正式发布前使用最新的开发中代码，你得[从源代码安装](#)。

使用 conda

自 Transformers 4.0.0 版始，我们有了一个 conda 频道：`huggingface`。

🤖 Transformers 可以通过 conda 依此安装：

```
conda install -c huggingface transformers
```

要通过 conda 安装 Flax、PyTorch 或 TensorFlow 其中之一，请参阅它们各自安装页的说明。

模型架构

🤖 Transformers 支持的[所有的模型检查点](#) 由[用户](#)和[组织](#)上传，均与 [huggingface.co model hub](#) 无缝整合。

models 68,379

目前的检查点数量：

🤖 Transformers 目前支持如下的架构（模型概述请阅[这里](#)）：

1. [ALBERT](#) (来自 Google Research and the Toyota Technological Institute at Chicago) 伴随论文 [ALBERT: A Lite BERT for Self-supervised Learning of Language Representations](#)，由 Zhenzhong Lan, Mingda Chen, Sebastian Goodman, Kevin Gimpel, Piyush Sharma, Radu Soricut 发布。
2. [BART](#) (来自 Facebook) 伴随论文 [BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension](#) 由 Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Ves Stoyanov and Luke Zettlemoyer 发布。
3. [BARTez](#) (来自 École polytechnique) 伴随论文 [BARTez: a Skilled Pretrained French Sequence-to-Sequence Model](#) 由 Moussa Kamal Eddine, Antoine J.-P. Tixier, Michalis Vazirgiannis 发布。
4. [BARTpho](#) (来自 VinAI Research) 伴随论文 [BARTpho: Pre-trained Sequence-to-Sequence Models for Vietnamese](#) 由 Nguyen Luong Tran, Duong Minh Le and Dat Quoc Nguyen 发布。
5. [BEiT](#) (来自 Microsoft) 伴随论文 [BEiT: BERT Pre-Training of Image Transformers](#) 由 Hangbo Bao, Li Dong, Furu Wei 发布。

6. **BERT** (来自 Google) 伴随论文 [BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding](#) 由 Jacob Devlin, Ming-Wei Chang, Kenton Lee and Kristina Toutanova 发布。
7. **BERT For Sequence Generation** (来自 Google) 伴随论文 [Leveraging Pre-trained Checkpoints for Sequence Generation Tasks](#) 由 Sascha Rothe, Shashi Narayan, Aliaksei Severyn 发布。
8. **BERTweet** (来自 VinAI Research) 伴随论文 [BERTweet: A pre-trained language model for English Tweets](#) 由 Dat Quoc Nguyen, Thanh Vu and Anh Tuan Nguyen 发布。
9. **BigBird-Pegasus** (来自 Google Research) 伴随论文 [Big Bird: Transformers for Longer Sequences](#) 由 Manzil Zaheer, Guru Guruganesh, Avinava Dubey, Joshua Ainslie, Chris Alberti, Santiago Ontanon, Philip Pham, Anirudh Ravula, Qifan Wang, Li Yang, Amr Ahmed 发布。
10. **BigBird-RoBERTa** (来自 Google Research) 伴随论文 [Big Bird: Transformers for Longer Sequences](#) 由 Manzil Zaheer, Guru Guruganesh, Avinava Dubey, Joshua Ainslie, Chris Alberti, Santiago Ontanon, Philip Pham, Anirudh Ravula, Qifan Wang, Li Yang, Amr Ahmed 发布。
11. **Blenderbot** (来自 Facebook) 伴随论文 [Recipes for building an open-domain chatbot](#) 由 Stephen Roller, Emily Dinan, Naman Goyal, Da Ju, Mary Williamson, Yinhan Liu, Jing Xu, Myle Ott, Kurt Shuster, Eric M. Smith, Y-Lan Boureau, Jason Weston 发布。
12. **BlenderbotSmall** (来自 Facebook) 伴随论文 [Recipes for building an open-domain chatbot](#) 由 Stephen Roller, Emily Dinan, Naman Goyal, Da Ju, Mary Williamson, Yinhan Liu, Jing Xu, Myle Ott, Kurt Shuster, Eric M. Smith, Y-Lan Boureau, Jason Weston 发布。
13. **BORT** (来自 Alexa) 伴随论文 [Optimal Subarchitecture Extraction For BERT](#) 由 Adrian de Wynter and Daniel J. Perry 发布。
14. **ByT5** (来自 Google Research) 伴随论文 [ByT5: Towards a token-free future with pre-trained byte-to-byte models](#) 由 Linting Xue, Aditya Barua, Noah Constant, Rami Al-Rfou, Sharan Narang, Mihir Kale, Adam Roberts, Colin Raffel 发布。
15. **CamemBERT** (来自 Inria/Facebook/Sorbonne) 伴随论文 [CamemBERT: a Tasty French Language Model](#) 由 Louis Martin*, Benjamin Muller*, Pedro Javier Ortiz Suárez*, Yoann Dupont, Laurent Romary, Éric Villemonte de la Clergerie, Djamé Seddah and Benoît Sagot 发布。
16. **CANINE** (来自 Google Research) 伴随论文 [CANINE: Pre-training an Efficient Tokenization-Free Encoder for Language Representation](#) 由 Jonathan H. Clark, Dan Garrette, Iulia Turc, John Wieting 发布。
17. **CLIP** (来自 OpenAI) 伴随论文 [Learning Transferable Visual Models From Natural Language Supervision](#) 由 Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, Ilya Sutskever 发布。
18. **ConvBERT** (来自 YituTech) 伴随论文 [ConvBERT: Improving BERT with Span-based Dynamic Convolution](#) 由 Zihang Jiang, Weihao Yu, Daquan Zhou, Yunpeng Chen, Jiashi Feng, Shuicheng Yan 发布。
19. **ConvNeXT** (来自 Facebook AI) 伴随论文 [A ConvNet for the 2020s](#) 由 Zhuang Liu, Hanzi Mao, Chao-Yuan Wu, Christoph Feichtenhofer, Trevor Darrell, Saining Xie 发布。
20. **CPM** (来自 Tsinghua University) 伴随论文 [CPM: A Large-scale Generative Chinese Pre-trained Language Model](#) 由 Zhengyan Zhang, Xu Han, Hao Zhou, Pei Ke, Yuxian Gu, Deming Ye, Yujia Qin, Yusheng Su, Haozhe Ji, Jian Guan, Fanchao Qi, Xiaozhi Wang, Yanan Zheng, Guoyang Zeng, Huanqi Cao, Shengqi Chen, Daixuan Li, Zhenbo Sun, Zhiyuan Liu, Minlie Huang, Wentao Han, Jie Tang, Juanzi Li, Xiaoyan Zhu, Maosong Sun 发布。
21. **CTRL** (来自 Salesforce) 伴随论文 [CTRL: A Conditional Transformer Language Model for Controllable Generation](#) 由 Nitish Shirish Keskar*, Bryan McCann*, Lav R. Varshney, Caiming Xiong and Richard Socher 发布。
22. **Data2Vec** (来自 Facebook) 伴随论文 [Data2Vec: A General Framework for Self-supervised Learning in Speech, Vision and Language](#) 由 Alexei Baevski, Wei-Ning Hsu, Qiantong Xu, Arun Babu, Jiatao Gu, Michael Auli 发布。
23. **DeBERTa** (来自 Microsoft) 伴随论文 [DeBERTa: Decoding-enhanced BERT with Disentangled Attention](#) 由 Pengcheng He, Xiaodong Liu, Jianfeng Gao, Weizhu Chen 发布。
24. **DeBERTa-v2** (来自 Microsoft) 伴随论文 [DeBERTa: Decoding-enhanced BERT with Disentangled Attention](#) 由 Pengcheng He, Xiaodong Liu, Jianfeng Gao, Weizhu Chen 发布。
25. **Decision Transformer** (来自 Berkeley/Facebook/Google) 伴随论文 [Decision Transformer: Reinforcement Learning via Sequence Modeling](#) 由 Lili Chen, Kevin Lu, Aravind Rajeswaran, Kimin Lee, Aditya Grover, Michael Laskin, Pieter Abbeel, Aravind Srinivas, Igor Mordatch 发布。
26. **DeiT** (来自 Facebook) 伴随论文 [Training data-efficient image transformers & distillation through attention](#) 由 Hugo Touvron, Matthieu Cord, Matthijs Douze, Francisco Massa, Alexandre Sablayrolles, Hervé Jégou 发布。

27. **DETR** (来自 Facebook) 伴随论文 [End-to-End Object Detection with Transformers](#) 由 Nicolas Carion, Francisco Massa, Gabriel Synnaeve, Nicolas Usunier, Alexander Kirillov, Sergey Zagoruyko 发布。
28. **DialoGPT** (来自 Microsoft Research) 伴随论文 [DialoGPT: Large-Scale Generative Pre-training for Conversational Response Generation](#) 由 Yizhe Zhang, Siqi Sun, Michel Galley, Yen-Chun Chen, Chris Brockett, Xiang Gao, Jianfeng Gao, Jingjing Liu, Bill Dolan 发布。
29. **DistilBERT** (来自 HuggingFace), 伴随论文 [DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter](#) 由 Victor Sanh, Lysandre Debut and Thomas Wolf 发布。同样的方法也应用于压缩 GPT-2 到 [DistilGPT2](#), RoBERTa 到 [DistilRoBERTa](#), Multilingual BERT 到 [DistilMBERT](#) 和德语版 DistilBERT。
30. **DiT** (来自 Microsoft Research) 伴随论文 [DiT: Self-supervised Pre-training for Document Image Transformer](#) 由 Junlong Li, Yiheng Xu, Tengchao Lv, Lei Cui, Cha Zhang, Furu Wei 发布。
31. **DPR** (来自 Facebook) 伴随论文 [Dense Passage Retrieval for Open-Domain Question Answering](#) 由 Vladimir Karpukhin, Barlas Oğuz, Sewon Min, Patrick Lewis, Ledell Wu, Sergey Edunov, Danqi Chen, and Wen-tau Yih 发布。
32. **DPT** (来自 Intel Labs) 伴随论文 [Vision Transformers for Dense Prediction](#) 由 René Ranftl, Alexey Bochkovskiy, Vladlen Koltun 发布。
33. **ELECTRA** (来自 Google Research/Stanford University) 伴随论文 [ELECTRA: Pre-training text encoders as discriminators rather than generators](#) 由 Kevin Clark, Minh-Thang Luong, Quoc V. Le, Christopher D. Manning 发布。
34. **EncoderDecoder** (来自 Google Research) 伴随论文 [Leveraging Pre-trained Checkpoints for Sequence Generation Tasks](#) 由 Sascha Rothe, Shashi Narayan, Aliaksei Severyn 发布。
35. **FlauBERT** (来自 CNRS) 伴随论文 [FlauBERT: Unsupervised Language Model Pre-training for French](#) 由 Hang Le, Loïc Vial, Jibril Frej, Vincent Segonne, Maximin Coavoux, Benjamin Lecouteux, Alexandre Allauzen, Benoît Crabbé, Laurent Besacier, Didier Schwab 发布。
36. **FNet** (来自 Google Research) 伴随论文 [FNet: Mixing Tokens with Fourier Transforms](#) 由 James Lee-Thorp, Joshua Ainslie, Ilya Eckstein, Santiago Ontanon 发布。
37. **Funnel Transformer** (来自 CMU/Google Brain) 伴随论文 [Funnel-Transformer: Filtering out Sequential Redundancy for Efficient Language Processing](#) 由 Zihang Dai, Guokun Lai, Yiming Yang, Quoc V. Le 发布。
38. **GLPN** (来自 KAIST) 伴随论文 [Global-Local Path Networks for Monocular Depth Estimation with Vertical CutDepth](#) 由 Doyeon Kim, Woonghyun Ga, Pyungwhan Ahn, Donggyu Joo, Sehwan Chun, Junmo Kim 发布。
39. **GPT** (来自 OpenAI) 伴随论文 [Improving Language Understanding by Generative Pre-Training](#) 由 Alec Radford, Karthik Narasimhan, Tim Salimans and Ilya Sutskever 发布。
40. **GPT Neo** (来自 EleutherAI) 随仓库 [EleutherAI/gpt-neo](#) 发布。作者为 Sid Black, Stella Biderman, Leo Gao, Phil Wang and Connor Leahy 发布。
41. **GPT-2** (来自 OpenAI) 伴随论文 [Language Models are Unsupervised Multitask Learners](#) 由 Alec Radford*, Jeffrey Wu*, Rewon Child, David Luan, Dario Amodei** and Ilya Sutskever** 发布。
42. **GPT-J** (来自 EleutherAI) 伴随论文 [kingoflolz/mesh-transformer-jax](#) 由 Ben Wang and Aran Komatsuzaki 发布。
43. **Hubert** (来自 Facebook) 伴随论文 [HuBERT: Self-Supervised Speech Representation Learning by Masked Prediction of Hidden Units](#) 由 Wei-Ning Hsu, Benjamin Bolte, Yao-Hung Hubert Tsai, Kushal Lakhotia, Ruslan Salakhutdinov, Abdelrahman Mohamed 发布。
44. **I-BERT** (来自 Berkeley) 伴随论文 [I-BERT: Integer-only BERT Quantization](#) 由 Sehoon Kim, Amir Gholami, Zhewei Yao, Michael W. Mahoney, Kurt Keutzer 发布。
45. **ImageGPT** (来自 OpenAI) 伴随论文 [Generative Pretraining from Pixels](#) 由 Mark Chen, Alec Radford, Rewon Child, Jeffrey Wu, Heewoo Jun, David Luan, Ilya Sutskever 发布。
46. **LayoutLM** (来自 Microsoft Research Asia) 伴随论文 [LayoutLM: Pre-training of Text and Layout for Document Image Understanding](#) 由 Yiheng Xu, Minghao Li, Lei Cui, Shaohan Huang, Furu Wei, Ming Zhou 发布。
47. **LayoutLMv2** (来自 Microsoft Research Asia) 伴随论文 [LayoutLMv2: Multi-modal Pre-training for Visually-Rich Document Understanding](#) 由 Yang Xu, Yiheng Xu, Tengchao Lv, Lei Cui, Furu Wei, Guoxin Wang, Yijuan Lu, Dinei Florencio, Cha Zhang, Wanxiang Che, Min Zhang, Lidong Zhou 发布。
48. **LayoutXLM** (来自 Microsoft Research Asia) 伴随论文 [LayoutXLM: Multimodal Pre-training for Multilingual Visually-rich Document Understanding](#) 由 Yiheng Xu, Tengchao Lv, Lei Cui, Guoxin Wang, Yijuan Lu, Dinei Florencio, Cha Zhang, Furu Wei 发布。
49. **LED** (来自 AllenAI) 伴随论文 [Longformer: The Long-Document Transformer](#) 由 Iz Beltagy, Matthew E. Peters, Arman Cohan 发布。

50. **Longformer** (来自 AllenAI) 伴随论文 [Longformer: The Long-Document Transformer](#) 由 Iz Beltagy, Matthew E. Peters, Arman Cohan 发布。
51. **LUKE** (来自 Studio Ousia) 伴随论文 [LUKE: Deep Contextualized Entity Representations with Entity-aware Self-attention](#) 由 Ikuya Yamada, Akari Asai, Hiroyuki Shindo, Hideaki Takeda, Yuji Matsumoto 发布。
52. **LXMERT** (来自 UNC Chapel Hill) 伴随论文 [LXMERT: Learning Cross-Modality Encoder Representations from Transformers for Open-Domain Question Answering](#) 由 Hao Tan and Mohit Bansal 发布。
53. **M2M100** (来自 Facebook) 伴随论文 [Beyond English-Centric Multilingual Machine Translation](#) 由 Angela Fan, Shruti Bhosale, Holger Schwenk, Zhiyi Ma, Ahmed El-Kishky, Siddharth Goyal, Mandeep Baines, Onur Celebi, Guillaume Wenzek, Vishrav Chaudhary, Naman Goyal, Tom Birch, Vitaliy Liptchinsky, Sergey Edunov, Edouard Grave, Michael Auli, Armand Joulin 发布。
54. **MarianMT** 用 **OPUS** 数据训练的机器翻译模型由 Jörg Tiedemann 发布。 [Marian Framework](#) 由微软翻译团队开发。
55. **MaskFormer** (from Meta and UIUC) released with the paper [Per-Pixel Classification is Not All You Need for Semantic Segmentation](#) by Bowen Cheng, Alexander G. Schwing, Alexander Kirillov
56. **MBart** (来自 Facebook) 伴随论文 [Multilingual Denoising Pre-training for Neural Machine Translation](#) 由 Yinhan Liu, Jiatao Gu, Naman Goyal, Xian Li, Sergey Edunov, Marjan Ghazvininejad, Mike Lewis, Luke Zettlemoyer 发布。
57. **MBart-50** (来自 Facebook) 伴随论文 [Multilingual Translation with Extensible Multilingual Pretraining and Finetuning](#) 由 Yuqing Tang, Chau Tran, Xian Li, Peng-Jen Chen, Naman Goyal, Vishrav Chaudhary, Jiatao Gu, Angela Fan 发布。
58. **Megatron-BERT** (来自 NVIDIA) 伴随论文 [Megatron-LM: Training Multi-Billion Parameter Language Models Using Model Parallelism](#) 由 Mohammad Shoeybi, Mostofa Patwary, Raul Puri, Patrick LeGresley, Jared Casper and Bryan Catanzaro 发布。
59. **Megatron-GPT2** (来自 NVIDIA) 伴随论文 [Megatron-LM: Training Multi-Billion Parameter Language Models Using Model Parallelism](#) 由 Mohammad Shoeybi, Mostofa Patwary, Raul Puri, Patrick LeGresley, Jared Casper and Bryan Catanzaro 发布。
60. **mLUKE** (来自 Studio Ousia) 伴随论文 [mLUKE: The Power of Entity Representations in Multilingual Pretrained Language Models](#) 由 Ryokan Ri, Ikuya Yamada, and Yoshimasa Tsuruoka 发布。
61. **MPNet** (来自 Microsoft Research) 伴随论文 [MPNet: Masked and Permuted Pre-training for Language Understanding](#) 由 Kaitao Song, Xu Tan, Tao Qin, Jianfeng Lu, Tie-Yan Liu 发布。
62. **mT5** (来自 Google AI) 伴随论文 [mT5: A massively multilingual pre-trained text-to-text transformer](#) 由 Linting Xue, Noah Constant, Adam Roberts, Mihir Kale, Rami Al-Rfou, Aditya Siddhant, Aditya Barua, Colin Raffel 发布。
63. **Nyströmformer** (来自 the University of Wisconsin - Madison) 伴随论文 [Nyströmformer: A Nyström-Based Algorithm for Approximating Self-Attention](#) 由 Yunyang Xiong, Zhanpeng Zeng, Rudrasis Chakraborty, Mingxing Tan, Glenn Fung, Yin Li, Vikas Singh 发布。
64. **Pegasus** (来自 Google) 伴随论文 [PEGASUS: Pre-training with Extracted Gap-sentences for Abstractive Summarization](#) 由 Jingqing Zhang, Yao Zhao, Mohammad Saleh and Peter J. Liu 发布。
65. **Perceiver IO** (来自 Deepmind) 伴随论文 [Perceiver IO: A General Architecture for Structured Inputs & Outputs](#) 由 Andrew Jaegle, Sebastian Borgeaud, Jean-Baptiste Alayrac, Carl Doersch, Catalin Ionescu, David Ding, Skanda Koppula, Daniel Zoran, Andrew Brock, Evan Shelhamer, Olivier Hénaff, Matthew M. Botvinick, Andrew Zisserman, Oriol Vinyals, João Carreira 发布。
66. **PhoBERT** (来自 VinAI Research) 伴随论文 [PhoBERT: Pre-trained language models for Vietnamese](#) 由 Dat Quoc Nguyen and Anh Tuan Nguyen 发布。
67. **PLBart** (来自 UCLA NLP) 伴随论文 [Unified Pre-training for Program Understanding and Generation](#) 由 Wasi Uddin Ahmad, Saikat Chakraborty, Baishakhi Ray, Kai-Wei Chang 发布。
68. **PoolFormer** (来自 Sea AI Labs) 伴随论文 [MetaFormer is Actually What You Need for Vision](#) 由 Yu, Weihao and Luo, Mi and Zhou, Pan and Si, Chenyang and Zhou, Yichen and Wang, Xinchao and Feng, Jiashi and Yan, Shuicheng 发布。
69. **ProphetNet** (来自 Microsoft Research) 伴随论文 [ProphetNet: Predicting Future N-gram for Sequence-to-Sequence Pre-training](#) 由 Yu Yan, Weizhen Qi, Yeyun Gong, Dayiheng Liu, Nan Duan, Jiusheng Chen, Ruofei Zhang and Ming Zhou 发布。
70. **QDQBert** (来自 NVIDIA) 伴随论文 [Integer Quantization for Deep Learning Inference: Principles and Empirical Evaluation](#) 由 Hao Wu, Patrick Judd, Xiaojie Zhang, Mikhail Isaev and Paulius Micikevicius 发布。
71. **REALM** (来自 Google Research) 伴随论文 [REALM: Retrieval-Augmented Language Model Pre-Training](#) 由 Kelvin Guu, Kenton Lee, Zora Tung, Panupong Pasupat and Ming-Wei Chang 发布。

72. **Reformer** (来自 Google Research) 伴随论文 [Reformer: The Efficient Transformer](#) 由 Nikita Kitaev, Łukasz Kaiser, Anselm Levskaya 发布。
73. **RegNet** (from META Research) released with the paper [Designing Network Design Space](#) by Ilija Radosavovic, Raj Prateek Kosaraju, Ross Girshick, Kaiming He, Piotr Dollár.
74. **RemBERT** (来自 Google Research) 伴随论文 [Rethinking embedding coupling in pre-trained language models](#) 由 Hyung Won Chung, Thibault Févry, Henry Tsai, M. Johnson, Sebastian Ruder 发布。
75. **ResNet** (from Microsoft Research) released with the paper [Deep Residual Learning for Image Recognition](#) by Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun.
76. **RoBERTa** (来自 Facebook), 伴随论文 [Robustly Optimized BERT Pretraining Approach](#) 由 Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, Veselin Stoyanov 发布。
77. **RoFormer** (来自 ZhuiyiTechnology), 伴随论文 [RoFormer: Enhanced Transformer with Rotary Position Embedding](#) 由 Jianlin Su and Yu Lu and Shengfeng Pan and Bo Wen and Yunfeng Liu 发布。
78. **SegFormer** (来自 NVIDIA) 伴随论文 [SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers](#) 由 Enze Xie, Wenhai Wang, Zhiding Yu, Anima Anandkumar, Jose M. Alvarez, Ping Luo 发布。
79. **SEW** (来自 ASAPP) 伴随论文 [Performance-Efficiency Trade-offs in Unsupervised Pre-training for Speech Recognition](#) 由 Felix Wu, Kwangyoung Kim, Jing Pan, Kyu Han, Kilian Q. Weinberger, Yoav Artzi 发布。
80. **SEW-D** (来自 ASAPP) 伴随论文 [Performance-Efficiency Trade-offs in Unsupervised Pre-training for Speech Recognition](#) 由 Felix Wu, Kwangyoung Kim, Jing Pan, Kyu Han, Kilian Q. Weinberger, Yoav Artzi 发布。
81. **SpeechToTextTransformer** (来自 Facebook), 伴随论文 [fairseq S2T: Fast Speech-to-Text Modeling with fairseq](#) 由 Changan Wang, Yun Tang, Xutai Ma, Anne Wu, Dmytro Okhonko, Juan Pino 发布。
82. **SpeechToTextTransformer2** (来自 Facebook) 伴随论文 [Large-Scale Self- and Semi-Supervised Learning for Speech Translation](#) 由 Changan Wang, Anne Wu, Juan Pino, Alexei Baevski, Michael Auli, Alexis Conneau 发布。
83. **Splinter** (来自 Tel Aviv University) 伴随论文 [Few-Shot Question Answering by Pretraining Span Selection](#) 由 Ori Ram, Yuval Kirstain, Jonathan Berant, Amir Globerson, Omer Levy 发布。
84. **SqueezeBert** (来自 Berkeley) 伴随论文 [SqueezeBERT: What can computer vision teach NLP about efficient neural networks?](#) 由 Forrest N. Iandola, Albert E. Shaw, Ravi Krishna, and Kurt W. Keutzer 发布。
85. **Swin Transformer** (来自 Microsoft) 伴随论文 [Swin Transformer: Hierarchical Vision Transformer using Shifted Windows](#) 由 Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, Baining Guo 发布。
86. **T5** (来自 Google AI) 伴随论文 [Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer](#) 由 Colin Raffel and Noam Shazeer and Adam Roberts and Katherine Lee and Sharan Narang and Michael Matena and Yanqi Zhou and Wei Li and Peter J. Liu 发布。
87. **T5v1.1** (来自 Google AI) 伴随论文 [google-research/text-to-text-transfer-transformer](#) 由 Colin Raffel and Noam Shazeer and Adam Roberts and Katherine Lee and Sharan Narang and Michael Matena and Yanqi Zhou and Wei Li and Peter J. Liu 发布。
88. **TAPAS** (来自 Google AI) 伴随论文 [TAPAS: Weakly Supervised Table Parsing via Pre-training](#) 由 Jonathan Herzig, Paweł Krzysztof Nowak, Thomas Müller, Francesco Piccinno and Julian Martin Eisenschlos 发布。
89. **Transformer-XL** (来自 Google/CMU) 伴随论文 [Transformer-XL: Attentive Language Models Beyond a Fixed-Length Context](#) 由 Zihang Dai*, Zhilin Yang*, Yiming Yang, Jaime Carbonell, Quoc V. Le, Ruslan Salakhutdinov 发布。
90. **TrOCR** (来自 Microsoft) 伴随论文 [TrOCR: Transformer-based Optical Character Recognition with Pre-trained Models](#) 由 Minghao Li, Tengchao Lv, Lei Cui, Yijuan Lu, Dinei Florencio, Cha Zhang, Zhoujun Li, Furu Wei 发布。
91. **UniSpeech** (来自 Microsoft Research) 伴随论文 [UniSpeech: Unified Speech Representation Learning with Labeled and Unlabeled Data](#) 由 Chengyi Wang, Yu Wu, Yao Qian, Kenichi Kumatani, Shujie Liu, Furu Wei, Michael Zeng, Xuedong Huang 发布。
92. **UniSpeechSat** (来自 Microsoft Research) 伴随论文 [UNISPEECH-SAT: UNIVERSAL SPEECH REPRESENTATION LEARNING WITH SPEAKER AWARE PRE-TRAINING](#) 由 Sanyuan Chen, Yu Wu, Chengyi Wang, Zhengyang Chen, Zhuo Chen, Shujie Liu, Jian Wu, Yao Qian, Furu Wei, Jinyu Li, Xiangzhan Yu 发布。
93. **VAN** (来自 Tsinghua University and Nankai University) 伴随论文 [Visual Attention Network](#) 由 Meng-Hao Guo, Cheng-Ze Lu, Zheng-Ning Liu, Ming-Ming Cheng, Shi-Min Hu 发布。
94. **ViT** (来自 NAVER AI Lab/Kakao Enterprise/Kakao Brain) 伴随论文 [ViT: Vision-and-Language Transformer Without Convolution or Region Supervision](#) 由 Wonjae Kim, Bokyung Son, Ildoo Kim 发布。
95. **Vision Transformer (ViT)** (来自 Google AI) 伴随论文 [An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale](#) 由 Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai,

Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, Neil Houlsby 发布。

96. [VisualBERT](#) (来自 UCLA NLP) 伴随论文 [VisualBERT: A Simple and Performant Baseline for Vision and Language](#) 由 Liunian Harold Li, Mark Yatskar, Da Yin, Cho-Jui Hsieh, Kai-Wei Chang 发布。
97. [ViTMAE](#) (来自 Meta AI) 伴随论文 [Masked Autoencoders Are Scalable Vision Learners](#) 由 Kaiming He, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollár, Ross Girshick 发布。
98. [Wav2Vec2](#) (来自 Facebook AI) 伴随论文 [wav2vec 2.0: A Framework for Self-Supervised Learning of Speech Representations](#) 由 Alexei Baevski, Henry Zhou, Abdelrahman Mohamed, Michael Auli 发布。
99. [Wav2Vec2Phoneme](#) (来自 Facebook AI) 伴随论文 [Simple and Effective Zero-shot Cross-lingual Phoneme Recognition](#) 由 Qiantong Xu, Alexei Baevski, Michael Auli 发布。
100. [WavLM](#) (from Microsoft Research) released with the paper [WavLM: Large-Scale Self-Supervised Pre-Training for Full Stack Speech Processing](#) by Sanyuan Chen, Chengyi Wang, Zhengyang Chen, Yu Wu, Shujie Liu, Zhuo Chen, Jinyu Li, Naoyuki Kanda, Takuya Yoshioka, Xiong Xiao, Jian Wu, Long Zhou, Shuo Ren, Yanmin Qian, Yao Qian, Jian Wu, Michael Zeng, Furu Wei.
101. [XGLM](#) (From Facebook AI) released with the paper [Few-shot Learning with Multilingual Language Models](#) by Xi Victoria Lin, Todor Mihaylov, Mikel Artetxe, Tianlu Wang, Shuohui Chen, Daniel Simig, Myle Ott, Naman Goyal, Shruti Bhosale, Jingfei Du, Ramakanth Pasunuru, Sam Shleifer, Punit Singh Koura, Vishrav Chaudhary, Brian O'Horo, Jeff Wang, Luke Zettlemoyer, Zornitsa Kozareva, Mona Diab, Veselin Stoyanov, Xian Li.
102. [XLM](#) (来自 Facebook) 伴随论文 [Cross-lingual Language Model Pretraining](#) 由 Guillaume Lample and Alexis Conneau 发布。
103. [XLM-ProphetNet](#) (来自 Microsoft Research) 伴随论文 [ProphetNet: Predicting Future N-gram for Sequence-to-Sequence Pre-training](#) 由 Yu Yan, Weizhen Qi, Yeyun Gong, Dayiheng Liu, Nan Duan, Jiusheng Chen, Ruofei Zhang and Ming Zhou 发布。
104. [XLM-RoBERTa](#) (来自 Facebook AI), 伴随论文 [Unsupervised Cross-lingual Representation Learning at Scale](#) 由 Alexis Conneau*, Kartikay Khandelwal*, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer and Veselin Stoyanov 发布。
105. [XLM-RoBERTa-XL](#) (来自 Facebook AI) 伴随论文 [Larger-Scale Transformers for Multilingual Masked Language Modeling](#) 由 Naman Goyal, Jingfei Du, Myle Ott, Giri Anantharaman, Alexis Conneau 发布。
106. [XLNet](#) (来自 Google/CMU) 伴随论文 [XLNet: Generalized Autoregressive Pretraining for Language Understanding](#) 由 Zhilin Yang*, Zihang Dai*, Yiming Yang, Jaime Carbonell, Ruslan Salakhutdinov, Quoc V. Le 发布。
107. [XLS-R](#) (来自 Facebook AI) 伴随论文 [XLS-R: Self-supervised Cross-lingual Speech Representation Learning at Scale](#) 由 Arun Babu, Changhan Wang, Andros Tjandra, Kushal Lakhotia, Qiantong Xu, Naman Goyal, Kritika Singh, Patrick von Platen, Yatharth Saraf, Juan Pino, Alexei Baevski, Alexis Conneau, Michael Auli 发布。
108. [XLSR-Wav2Vec2](#) (来自 Facebook AI) 伴随论文 [Unsupervised Cross-Lingual Representation Learning For Speech Recognition](#) 由 Alexis Conneau, Alexei Baevski, Ronan Collobert, Abdelrahman Mohamed, Michael Auli 发布。
109. [YOSO](#) (来自 the University of Wisconsin - Madison) 伴随论文 [You Only Sample (Almost)] 由 Zhanpeng Zeng, Yunyang Xiong, Sathya N. Ravi, Shailesh Acharya, Glenn Fung, Vikas Singh 发布。
110. 想要贡献新的模型？我们这里有一份[详细指引和模板](#)来引导你添加新的模型。你可以在 [templates](#) 目录中找到他们。记得查看 [贡献指南](#) 并在开始写 PR 前联系维护人员或开一个新的 issue 来获得反馈。

要检查某个模型是否已有 Flax、PyTorch 或 TensorFlow 的实现，或其是否在 🗨️ Tokenizers 库中有对应词符化器 (tokenizer)，敬请参阅[此表](#)。

这些实现均已于多个数据集测试（请参看用例脚本）并应于原版实现表现相当。你可以在用例文档的[此节](#)中了解表现的细节。

了解更多

章节	描述
文档	完整的 API 文档和教程
任务总结	🗨️ Transformers 支持的任务
预处理教程	使用 <code>Tokenizer</code> 来为模型准备数据

训练和微调	在 PyTorch/TensorFlow 的训练循环或 <code>Trainer</code> API 中使用 🤗 Transformers 提供的模型
快速上手：微调 and 用例脚本	为各种任务提供的用例脚本
模型分享和上传	和社区上传和分享你微调的模型
迁移	从 <code>pytorch-transformers</code> 或 <code>pytorch-pretrained-bert</code> 迁移到 🤗 Transformers

引用

我们已将此库的[论文](#)正式发表，如果你使用了 🤗 Transformers 库，请引用：

```
@inproceedings{wolf-etal-2020-transformers,
  title = "Transformers: State-of-the-Art Natural Language Processing",
  author = "Thomas Wolf and Lysandre Debut and Victor Sanh and Julien Chaumond and Clement Delangue and Anthony Moi and Pierric Cistac and Tim Rault and Rémi Louf and Morgan Funtowicz and Joe Davison and Sam Shleifer and Patrick von Platen and Clara Ma and Yacine Jernite and Julien Plu and Canwen Xu and Teven Le Scao and Sylvain Gugger and Mariama Drame and Quentin Lhoest and Alexander M. Rush",
  booktitle = "Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations",
  month = oct,
  year = "2020",
  address = "Online",
  publisher = "Association for Computational Linguistics",
  url = "https://www.aclweb.org/anthology/2020.emnlp-demos.6",
  pages = "38--45"
}
```