

# ChatGPT language models for fine-tuning

## Macro models

GPT-3 (Generative Pretrained Transformer-3)

GPT-2 (Generative Pretrained Transformer-2)

BERT (Bidirectional Encoder Representations from Transformers)

ELMo (Embeddings from Language Models)

ULMFiT (Universal Language Model Fine-Tuning)

T5 (Text-to-Text Transfer Transformer)

Transformer-XL

RoBERTa (A Robustly Optimized BERT Pretraining Approach)

CTRL (Conditional Transformer Language Model)

XLNet (Generalized Autoregressive Pre Training for Language Understanding)

ALBERT (A Lite BERT for Self-Supervised Learning of Language Representations)

## ChatGPT detailed model list

GPT-3

GPT-2

BERT

ELMo

ULMFiT

Transformer-XL

RoBERTa

CTRL

XLNet

ALBERT GPT-3 Small

GPT-3 Medium

GPT-3 Large

GPT-3 XL

GPT-3 XXL

GPT-3 117M

GPT-3 345B

GPT-2 (Generative Pretrained Transformer-2)

GPT-2 Small

GPT-2 Medium

GPT-2 Large

GPT-2 XL

GPT-2 117M

BERT (Bidirectional Encoder Representations from Transformers)

BERT-Base

BERT-Large

BERT-Base Multilingual Cased

BERT-Large Uncased

BERT-Base Chinese

ELMo (Embeddings from Language Models)

ELMo 2-Layer



ELMo 3-Layer

ULMFiT (Universal Language Model Fine-Tuning)

ULMFiT (Base)

ULMFiT (Large)

Transformer-XL

Transformer-XL (Base)

Transformer-XL (Large)

RoBERTa (A Robustly Optimized BERT Pretraining Approach)

RoBERTa Base

RoBERTa Large

RoBERTa Medium

RoBERTa Tiny

RoBERTa Base Chinese

CTRL (Conditional Transformer Language Model)

CTRL-Base

CTRL-Large

CTRL-XL

XLNet (Generalized Autoregressive Pretraining for Language Understanding)

XLNet Base

XLNet Large

XLNet Medium

ALBERT (A Lite BERT for Self-Supervised Learning of Language Representations)

ALBERT Base

ALBERT Large

ALBERT XL

T5-11B

Roberta-base-cased-lm-gen

RoBERTa-base

RoBERTa-large

RoBERTa-wwm

RoBERTa-wwm-ext

DistilRoBERTa

KIQUIX