

Title: GPT-3

URL: <https://en.wikipedia.org/wiki/GPT-3>

PageID: 64695824

Categories: Category:2020 in artificial intelligence, Category:2020 software, Category:Generative pre-trained transformers, Category:Large language models, Category:OpenAI

Source: Wikipedia (CC BY-SA 4.0).

Supervised learning

Unsupervised learning

Semi-supervised learning

Self-supervised learning

Reinforcement learning

Meta-learning

Online learning

Batch learning

Curriculum learning

Rule-based learning

Neuro-symbolic AI

Neuromorphic engineering

Quantum machine learning

Classification

Generative modeling

Regression

Clustering

Dimensionality reduction

Density estimation

Anomaly detection

Data cleaning

AutoML

Association rules

Semantic analysis

Structured prediction

Feature engineering

Feature learning

Learning to rank

Grammar induction

Ontology learning

Multimodal learning

Apprenticeship learning

Decision trees

Ensembles Bagging Boosting Random forest

Bagging

Boosting

Random forest

k -NN

Linear regression

Naive Bayes

Artificial neural networks

Logistic regression

Perceptron

Relevance vector machine (RVM)

Support vector machine (SVM)

BIRCH

CURE

Hierarchical

k -means

Fuzzy

Expectation–maximization (EM)

DBSCAN

OPTICS

Mean shift

Factor analysis

CCA

ICA

LDA

NMF

PCA

PGD

t-SNE

SDL

Graphical models Bayes net Conditional random field Hidden Markov

Bayes net

Conditional random field

Hidden Markov

RANSAC

k -NN

Local outlier factor
Isolation forest
Autoencoder
Deep learning
Feedforward neural network
Recurrent neural network LSTM GRU ESN reservoir computing
LSTM
GRU
ESN
reservoir computing
Boltzmann machine Restricted
Restricted
GAN
Diffusion model
SOM
Convolutional neural network U-Net LeNet AlexNet DeepDream
U-Net
LeNet
AlexNet
DeepDream
Neural field Neural radiance field Physics-informed neural networks
Neural radiance field
Physics-informed neural networks
Transformer Vision
Vision
Mamba
Spiking neural network
Memtransistor
Electrochemical RAM (ECRAM)
Q-learning
Policy gradient
SARSA
Temporal difference (TD)
Multi-agent Self-play
Self-play
Active learning
Crowdsourcing
Human-in-the-loop

Mechanistic interpretability

RLHF

Coefficient of determination

Confusion matrix

Learning curve

ROC curve

Kernel machines

Bias–variance tradeoff

Computational learning theory

Empirical risk minimization

Occam learning

PAC learning

Statistical learning

VC theory

Topological deep learning

AAAI

ECML PKDD

NeurIPS

ICML

ICLR

IJCAI

ML

JMLR

Glossary of artificial intelligence

List of datasets for machine-learning research List of datasets in computer vision and image processing

List of datasets in computer vision and image processing

Outline of machine learning

v

t

e

Generative Pre-trained Transformer 3 (GPT-3) is a large language model released by OpenAI in 2020.

Like its predecessor, GPT-2 , it is a decoder-only [2] transformer model of deep neural network, which supersedes recurrence and convolution-based architectures with a technique known as " attention ". [3] This attention mechanism allows the model to focus selectively on segments of input text it predicts to be most relevant. [4] GPT-3 has 175 billion parameters [1] , each with 16-bit precision, requiring 350GB of storage since each parameter occupies 2 bytes. It has a context window size of 2048 tokens , and has demonstrated strong " zero-shot " and " few-shot " learning abilities on many tasks. [2]

On September 22, 2020, Microsoft announced that it had licensed GPT-3 exclusively. Others can still receive output from its public API, but only Microsoft has access to the underlying model. [5]

Background

According to The Economist , improved algorithms, more powerful computers, and a recent increase in the amount of digitized material have fueled a revolution in machine learning . New techniques in the 2010s resulted in "rapid improvements in tasks", including manipulating language. [6]

Software models are trained to learn by using thousands or millions of examples in a "structure ... loosely based on the neural architecture of the brain". [6] One architecture used in natural language processing (NLP) is a neural network based on a deep learning model that was introduced in 2017—the transformer architecture. [7] There are a number of NLP systems capable of processing, mining, organizing, connecting and contrasting textual input, as well as correctly answering questions. [8]

On June 11, 2018, OpenAI researchers and engineers published a paper introducing the first generative pre-trained transformer (GPT)—a type of generative large language model that is pre-trained with an enormous and diverse text corpus in datasets , followed by discriminative fine-tuning to focus on a specific task. GPT models are transformer-based deep-learning neural network architectures. Previously, the best-performing neural NLP models commonly employed supervised learning from large amounts of manually-labeled data, which made it prohibitively expensive and time-consuming to train extremely large language models. [2] The first GPT model was known as "GPT-1," and it was followed by "GPT-2" in February 2019. Created as a direct scale-up of its predecessor, GPT-2 had both its parameter count and dataset size increased by a factor of 10. It had 1.5 billion parameters, and was trained on a dataset of 8 million web pages. [9]

In February 2020, Microsoft introduced its Turing Natural Language Generation (T-NLG), which they claimed was "largest language model ever published at 17 billion parameters." [10] It performed better than any other language model at a variety of tasks, including summarizing texts and answering questions .

Training and capabilities

The construct of "learning styles" is problematic because it fails to account for the processes through which learning styles are shaped. Some students might develop a particular learning style because they have had particular experiences. Others might develop a particular learning style by trying to accommodate to a learning environment that was not well suited to their learning needs. Ultimately, we need to understand the interactions among learning styles and environmental and personal factors, and how these shape how we learn and the kinds of learning we experience.

On May 28, 2020, an arXiv preprint by a group of 31 engineers and researchers at OpenAI described the achievement and development of GPT-3, a third-generation "state-of-the-art language model". [1] [12] The team increased the capacity of GPT-3 by over two orders of magnitude from that of its predecessor, GPT-2, [13] making GPT-3 the largest non-sparse language model to date. [1] : 14 [14] Because GPT-3 is structurally similar to its predecessors, [1] its greater accuracy is attributed to its increased capacity and greater number of parameters. [15] GPT-3's capacity is ten times larger than that of Microsoft's Turing NLG, the next largest NLP model known at the time. [12]

Lambdalabs estimated a hypothetical cost of around \$4.6 million US dollars and 355 years to train GPT-3 on a single GPU in 2020, [16] with lower actual training time by using more GPUs in parallel.

Sixty percent of the weighted pre-training dataset for GPT-3 comes from a filtered version of Common Crawl consisting of 410 billion byte-pair-encoded tokens. Fuzzy deduplication used Apache Spark 's MinHash LSH. [1] : 9 Other sources are 19 billion tokens from WebText2 representing 22% of the weighted total, 12 billion tokens from Books1 representing 8%, 55 billion tokens from Books2 representing 8%, and 3 billion tokens from Wikipedia representing 3%. [1] : 9 GPT-3 was trained on hundreds of billions of words and is also capable of coding in CSS , JSX , and Python , among others. [citation needed]

Since GPT-3's training data was all-encompassing, it does not require further training for distinct language tasks. [citation needed] The training data contains occasional toxic language and GPT-3 occasionally generates toxic language as a result of mimicking its training data. A study from the University of Washington found that GPT-3 produced toxic language at a toxicity level comparable to the similar natural language processing models of GPT-2 and CTRL. OpenAI has implemented several strategies to limit the amount of toxic language generated by GPT-3. As a result, GPT-3 produced less toxic language compared to its predecessor model, GPT-1, although it produced both more generations and a higher toxicity of toxic language compared to CTRL Wiki, a language model trained entirely on Wikipedia data. [17]

On June 11, 2020, OpenAI announced that users could request access to its user-friendly GPT-3 API—a "machine learning toolset"—to help OpenAI "explore the strengths and limits" of this new technology. [18] [19] The invitation described how this API had a general-purpose "text in, text out" interface that can complete almost "any English language task", instead of the usual single use-case. [18] According to one user, who had access to a private early release of the OpenAI GPT-3 API, GPT-3 was "eerily good" at writing "amazingly coherent text" with only a few simple prompts. [20] In an initial experiment 80 US subjects were asked to judge if short ~200 word articles were written by humans or GPT-3. The participants judged correctly 52% of the time, doing only slightly better than random guessing. [1]

On November 18, 2021, OpenAI announced that enough safeguards had been implemented that access to its API would be unrestricted. [21] OpenAI provided developers with a content moderation tool that helps them abide by OpenAI's content policy. [22] On January 27, 2022, OpenAI announced that its newest GPT-3 language models (collectively referred to as InstructGPT) were now the default language model used on their API . According to OpenAI, InstructGPT produced content that was better aligned to user intentions by following instructions better, generating fewer made-up facts, and producing somewhat less toxic content. [23]

Because GPT-3 can "generate news articles which human evaluators have difficulty distinguishing from articles written by humans," [12] GPT-3 has the "potential to advance both the beneficial and harmful applications of language models." [1] : 34 In their May 28, 2020 paper, the researchers described in detail the potential "harmful effects of GPT-3" [12] which include "misinformation, spam , phishing , abuse of legal and governmental processes , fraudulent academic essay writing and social engineering pretexting ". [1] The authors draw attention to these dangers to call for research on risk mitigation . [1] : 34

GPT-3 is capable of performing zero-shot and few-shot learning (including one-shot). [1]

In June 2022, Almira Osmanovic Thunström wrote that GPT-3 was the primary author on an article on itself, that they had submitted it for publication, [24] and that it had been pre-published while waiting for completion of its review. [25]

GPT-3 models

There are many models in the GPT-3 family, some serving different purposes than others. In the initial research paper published by OpenAI, they mentioned 8 different sizes of the main GPT-3 model (Table 2.1):

Half of the models are accessible through the API, namely GPT-3-medium, GPT-3-xl, GPT-3-6.7B and GPT-3-175b, which are referred to as ada, babbage, curie and davinci respectively. While the size of the API models was not originally disclosed by OpenAI, EleutherAI announced the mapping between model sizes and API names in May 2021. [26] These model sizes were later confirmed by OpenAI, [27] but the sizes of subsequent models have not been disclosed.

babbage-002

davinci-002

code-davinci-002

gpt-3.5-turbo-instruct

gpt-3.5-turbo-16k

GPT-3.5

Generative Pre-trained Transformer 3.5 (GPT-3.5) is a sub class of GPT-3 Models created by OpenAI in 2022.

On March 15, 2022, OpenAI made available new versions of GPT-3 and Codex in its API with edit and insert capabilities under the names "text-davinci-002" and "code-davinci-002". [28] These models were described as more capable than previous versions and were trained on data up to June 2021. [29] On November 28, 2022, OpenAI introduced text-davinci-003. [30] On November 30, 2022, OpenAI began referring to these models as belonging to the "GPT-3.5" series, [29] and released ChatGPT , which was fine-tuned from a model in the GPT-3.5 series. [31] OpenAI does not include GPT-3.5 in GPT-3. [32]

Models

There are three models: [33]

Chat gpt-3.5-turbo

gpt-3.5-turbo

Text completion text-davinci-003 text-davinci-002

text-davinci-003

text-davinci-002

GPT-3.5 with browsing

On April 10, 2023, OpenAI introduced a new variant of its GPT-3.5 series model, known as GPT-3.5 with Browsing (ALPHA). [citation needed] This updated model was described to build upon the capabilities of its predecessors "text-davinci-002" and "code-davinci-002". [34] The GPT-3.5 with Browsing (ALPHA) model incorporated the ability to access and browse online information. This has led to more accurate and up-to-date responses to user queries. [citation needed]

The GPT-3.5 with Browsing (ALPHA) model has been trained on data up to September 2021, giving it more information compared to previous GPT-3.5 models, which were trained on data up until June 2021. The model attempted to provide developers and users with an advanced natural language processing tool that can effectively retrieve and synthesize online information. [citation needed]

To enable browsing capabilities, OpenAI implemented a new API that allows the GPT-3.5 with Browsing (ALPHA) model to access selected online resources during operation. [35] This feature allows users to ask questions or request information with the expectation that the model will deliver updated, accurate, and relevant answers based on the latest online sources available to it.

On April 27, 2023, OpenAI made the GPT-3.5 with Browsing (ALPHA) model publicly available to GPT Plus users. This allowed more people to access to its new features. [35]

InstructGPT

InstructGPT is a fine-tuned version of GPT-3.5 trained on a dataset of human-written instructions. [36]

Reception

Applications

GPT-3, specifically the Codex model , was the basis for GitHub Copilot , a code completion and generation software that can be used in various code editors and IDEs. [37] [38]

GPT-3 is used in certain Microsoft products to translate conventional language into formal computer code. [39] [40]

GPT-3 has been used in CodexDB [41] to generate query-specific code for SQL processing.

GPT-3 has been used by Jason Rohrer in a retro-themed chatbot project named "Project December", which is accessible online and allows users to converse with several AIs using GPT-3 technology. [42]

GPT-3 was used by The Guardian to write an article about AI being harmless to human beings. It was fed some ideas and produced eight different essays, which were ultimately merged into one article. [43]

GPT-3 was used in AI Dungeon , which generates text-based adventure games. Later it was replaced by a competing model after OpenAI changed their policy regarding generated content. [44] [45]

GPT-3 is used to aid in writing copy and other marketing materials. [46]

A 2022 study from Drexel University suggested that GPT-3-based systems could be used to screen for early signs of Alzheimer's disease . [47] [48]

Reviews

In a July 2020 review in The New York Times , Farhad Manjoo said that GPT-3's ability to generate computer code, poetry, and prose is not just "amazing", "spooky", and "humbling", but also "more than a little terrifying". [49]

Daily Nous presented a series of articles by nine philosophers on GPT-3. [50] Australian philosopher David Chalmers described GPT-3 as "one of the most interesting and important AI systems ever produced". [51]

A review in Wired said that GPT-3 was "provoking chills across Silicon Valley ". [52]

The National Law Review said that GPT-3 is an "impressive step in the larger process", with OpenAI and others finding "useful applications for all of this power" while continuing to "work toward a more general intelligence ". [53]

An article in the MIT Technology Review , co-written by Deep Learning critic Gary Marcus , [54] stated that GPT-3's "comprehension of the world is often seriously off, which means you can never really trust what it says." [55] According to the authors, GPT-3 models relationships between words without having an understanding of the meaning behind each word.

Jerome Pesenti, head of the Facebook AI lab, said GPT-3 is "unsafe," pointing to the sexist , racist and other biased and negative language generated by the system when it was asked to discuss Jews, women, black people, and the Holocaust . [56]

Nabla, a French start-up specializing in healthcare technology, tested GPT-3 as a medical chatbot , though OpenAI itself warned against such use. As expected, GPT-3 showed several limitations. For example, while testing GPT-3 responses about mental health issues, the AI advised a simulated patient to commit suicide. [57]

Noam Chomsky expressed his skepticism about GPT-3's scientific value: "It's not a language model. It works just as well for impossible languages as for actual languages. It is therefore refuted, if intended as a language model, by normal scientific criteria. [...] Perhaps it's useful for some purpose, but it seems to tell us nothing about language or cognition generally." [58]

Luciano Floridi and Massimo Chiriatti highlighted the risk of "cheap production of good, semantic artefacts". [59]

OpenAI's Sam Altman himself criticized what he called "GPT-3 hype", acknowledging GPT-3 "has serious weakness and sometimes makes very silly mistakes... AI is going to change the world, but GPT-3 is just a very early glimpse." [60]

Criticism

GPT-3's builder, OpenAI , was initially founded as a non-profit in 2015. [61] In 2019, OpenAI broke from its usual open-source standards by not publicly releasing GPT-3's predecessor model, citing concerns that the model could facilitate the propagation of fake news. OpenAI eventually released a version of GPT-2 that was 8% of the original model's size. [62] In the same year, OpenAI

restructured to be a for-profit company. [63] In 2020, Microsoft announced the company had exclusive licensing of GPT-3 for Microsoft's products and services following a multi-billion dollar investment in OpenAI. The agreement permits OpenAI to offer a public-facing API such that users can send text to GPT-3 to receive the model's output, but only Microsoft will have access to GPT-3's source code. [5]

Large language models, such as GPT-3, have come under criticism from a few of Google's AI ethics researchers for the environmental impact of training and storing the models, detailed in a paper co-authored by Timnit Gebru and Emily M. Bender in 2021. [64]

The growing [when?] use of automated writing technologies based on GPT-3 and other language generators, has raised concerns regarding academic integrity [65] and raised the stakes of how universities and schools will gauge what constitutes academic misconduct such as plagiarism. [66]

OpenAI's GPT series was built with data from the Common Crawl dataset, [67] a conglomerate of copyrighted articles, internet posts, web pages, and books scraped from 60 million domains over a period of 12 years. TechCrunch reports this training data includes copyrighted material from the BBC, The New York Times , Reddit , the full text of online books, and more. [68] In its response to a 2019 Request for Comments on Intellectual Property Protection for Artificial Intelligence Innovation from the United States Patent and Trademark Office (USPTO), OpenAI argued that "Under current law, training AI systems [such as its GPT models] constitutes fair use ," but that "given the lack of case law on point, OpenAI and other AI developers like us face substantial legal uncertainty and compliance costs." [69]

See also

GPTZero

Hallucination (artificial intelligence)

List of large language models

Wu Dao

References

v

t

e

ChatGPT in education GPT Store DALL-E ChatGPT Search Sora Whisper

in education

GPT Store

DALL-E

ChatGPT Search

Sora

Whisper

GitHub Copilot

OpenAI Codex

Generative pre-trained transformer GPT-1 GPT-2 GPT-3 GPT-4 GPT-4o o1 o3 GPT-4.5 GPT-4.1 o4-mini GPT-OSS GPT-5

GPT-1

GPT-2

GPT-3

GPT-4

GPT-4o
o1
o3
GPT-4.5
GPT-4.1
o4-mini
GPT-OSS
GPT-5
ChatGPT Deep Research
Operator
Sam Altman removal
removal
Greg Brockman
Sarah Friar
Jakub Pachocki
Scott Schools
Mira Murati
Emmett Shear
Sam Altman
Adam D'Angelo
Sue Desmond-Hellmann
Zico Kolter
Paul Nakasone
Adebayo Ogunlesi
Nicole Seligman
Fidji Simo
Lawrence Summers
Bret Taylor (chair)
Greg Brockman (2017–2023)
Reid Hoffman (2019–2023)
Will Hurd (2021–2023)
Holden Karnofsky (2017–2021)
Elon Musk (2015–2018)
Ilya Sutskever (2017–2023)
Helen Toner (2021–2023)
Shivon Zilis (2019–2023)
Stargate LLC
Apple Intelligence

AI Dungeon
AutoGPT
Contrastive Language-Image Pre-training
" Deep Learning "
LangChain
Microsoft Copilot
OpenAI Five
Transformer
Category
v
t
e
Autoencoder
Deep learning
Fine-tuning
Foundation model
Generative adversarial network
Generative pre-trained transformer
Large language model
Model Context Protocol
Neural network
Prompt engineering
Reinforcement learning from human feedback
Retrieval-augmented generation
Self-supervised learning
Stochastic parrot
Synthetic data
Top-p sampling
Transformer
Variational autoencoder
Vibe coding
Vision transformer
Waluigi effect
Word embedding
Character.ai
ChatGPT
DeepSeek
Ernie

Gemini

Grok

Copilot

Claude

Gemini

Gemma

GPT 1 2 3 J 4 4o 4.5 4.1 OSS 5

1

2

3

J

4

4o

4.5

4.1

OSS

5

Llama

o1

o3

o4-mini

Qwen

Base44

Claude Code

Cursor

Devstral

GitHub Copilot

Kimi-Dev

Qwen3-Coder

Replit

Xcode

Aurora

Firefly

Flux

GPT Image 1

Ideogram

Imagen

Midjourney

Qwen-Image
Recraft
Seedream
Stable Diffusion
Dream Machine
Hailuo AI
Kling
Midjourney Video
Runway Gen
Seedance
Sora
Veo
Wan
15.ai
Eleven
MiniMax Speech 2.5
WaveNet
Eleven Music
Endel
Lyria
Riffusion
Suno AI
Udio
Agentforce
AutoGLM
AutoGPT
ChatGPT Agent
Devin AI
Manus
OpenAI Codex
Operator
Replit Agent
01.AI
Aleph Alpha
Anthropic
Baichuan
Canva
Cognition AI

Cohere
Contextual AI
DeepSeek
ElevenLabs
Google DeepMind
HeyGen
Hugging Face
Inflection AI
Krikey AI
Kuaishou
Luma Labs
Meta AI
MiniMax
Mistral AI
Moonshot AI
OpenAI
Perplexity AI
Runway
Safe Superintelligence
Salesforce
Scale AI
SoundHound
Stability AI
Synthesia
Thinking Machines Lab
Upstage
xAI
Z.ai
Category
v
t
e
History timeline
timeline
Companies
Projects
Parameter Hyperparameter
Hyperparameter

Loss functions

Regression Bias–variance tradeoff Double descent Overfitting

Bias–variance tradeoff

Double descent

Overfitting

Clustering

Gradient descent SGD Quasi-Newton method Conjugate gradient method

SGD

Quasi-Newton method

Conjugate gradient method

Backpropagation

Attention

Convolution

Normalization Batchnorm

Batchnorm

Activation Softmax Sigmoid Rectifier

Softmax

Sigmoid

Rectifier

Gating

Weight initialization

Regularization

Datasets Augmentation

Augmentation

Prompt engineering

Reinforcement learning Q-learning SARSA Imitation Policy gradient

Q-learning

SARSA

Imitation

Policy gradient

Diffusion

Latent diffusion model

Autoregression

Adversary

RAG

Uncanny valley

RLHF

Self-supervised learning

Reflection
Recursive self-improvement
Hallucination
Word embedding
Vibe coding
Machine learning In-context learning
In-context learning
Artificial neural network Deep learning
Deep learning
Language model Large language model NMT
Large language model
NMT
Reasoning language model
Model Context Protocol
Intelligent agent
Artificial human companion
Humanity's Last Exam
Artificial general intelligence (AGI)
AlexNet
WaveNet
Human image synthesis
HWR
OCR
Computer vision
Speech synthesis 15.ai ElevenLabs
15.ai
ElevenLabs
Speech recognition Whisper
Whisper
Facial recognition
AlphaFold
Text-to-image models Aurora DALL-E Firefly Flux Ideogram Imagen Midjourney Recraft Stable Diffusion
Aurora
DALL-E
Firefly
Flux
Ideogram

Imagen

Midjourney

Recraft

Stable Diffusion

Text-to-video models Dream Machine Runway Gen Hailuo AI Kling Sora Veo

Dream Machine

Runway Gen

Hailuo AI

Kling

Sora

Veo

Music generation Riffusion Suno AI Udio

Riffusion

Suno AI

Udio

Word2vec

Seq2seq

GloVe

BERT

T5

Llama

Chinchilla AI

PaLM

GPT 1 2 3 J ChatGPT 4 4o o1 o3 4.5 4.1 o4-mini 5

1

2

3

J

ChatGPT

4

4o

o1

o3

4.5

4.1

o4-mini

5

Claude

Gemini Gemini (language model) Gemma

Gemini (language model)

Gemma

Grok

LaMDA

BLOOM

DBRX

Project Debater

IBM Watson

IBM Watsonx

Granite

PanGu- Σ

DeepSeek

Qwen

AlphaGo

AlphaZero

OpenAI Five

Self-driving car

MuZero

Action selection AutoGPT

AutoGPT

Robot control

Alan Turing

Warren Sturgis McCulloch

Walter Pitts

John von Neumann

Claude Shannon

Shun'ichi Amari

Kunihiko Fukushima

Takeo Kanade

Marvin Minsky

John McCarthy

Nathaniel Rochester

Allen Newell

Cliff Shaw

Herbert A. Simon

Oliver Selfridge

Frank Rosenblatt

Bernard Widrow
Joseph Weizenbaum
Seymour Papert
Seppo Linnainmaa
Paul Werbos
Geoffrey Hinton
John Hopfield
Jürgen Schmidhuber
Yann LeCun
Yoshua Bengio
Lotfi A. Zadeh
Stephen Grossberg
Alex Graves
James Goodnight
Andrew Ng
Fei-Fei Li
Alex Krizhevsky
Ilya Sutskever
Oriol Vinyals
Quoc V. Le
Ian Goodfellow
Demis Hassabis
David Silver
Andrej Karpathy
Ashish Vaswani
Noam Shazeer
Aidan Gomez
John Schulman
Mustafa Suleyman
Jan Leike
Daniel Kokotajlo
François Chollet
Neural Turing machine
Differentiable neural computer
Transformer Vision transformer (ViT)
Vision transformer (ViT)
Recurrent neural network (RNN)
Long short-term memory (LSTM)

Gated recurrent unit (GRU)

Echo state network

Multilayer perceptron (MLP)

Convolutional neural network (CNN)

Residual neural network (RNN)

Highway network

Mamba

Autoencoder

Variational autoencoder (VAE)

Generative adversarial network (GAN)

Graph neural network (GNN)

Category