



AI Glossary by a16z

Explore More: AI + A16z —

A B C D E F G H I L M N O P R S T U V X Z



SEARCH

Accelerator

A class of microprocessor designed to accelerate AI applications.

Agents

Software that can perform certain tasks independently and proactively without the need for human intervention, often utilizing a suite of tools like calculators or web browsing.

AGI (Artificial General Intelligence)

Though not widely agreed upon, Microsoft researchers have defined AGI as artificial intelligence that is as capable as a human at any intellectual task.

Alignment

The task of ensuring that the goals of an AI system are in line with human values.

ASI (Artificial Super Intelligence)

Though subject to debate, ASI is commonly defined as artificial intelligence that surpasses the capabilities of the human mind.

Attention

In the context of neural networks, attention mechanisms help the model focus on relevant parts of the input when producing an output.

Back Propagation

An algorithm often used in training neural networks, referring to the method for computing the gradient of the loss function with respect to the weights in the network.

Bias

Assumptions made by an AI model about the data. A “bias variance tradeoff” is the balance that must be achieved between assumptions a model makes about the data and the amount a model’s predictions change, given different training data. Inductive bias is the set of assumptions that a machine learning algorithm makes on the underlying distribution of the data.

Chain of Thought

In AI, this term is often used to describe the sequence of reasoning steps an AI model uses to arrive at a decision.

Chatbot

A computer program designed to simulate human conversation through text or voice interactions. Chatbots often utilize natural language processing techniques to understand user input and provide relevant responses.

ChatGPT

A large-scale AI language model developed by OpenAI that generates human-like text.

CLIP (Contrastive Language–Image Pretraining)

An AI model developed by OpenAI that connects images and text, allowing it to understand and generate descriptions of images.

Compute

The computational resources (like CPU or GPU time) used in training or running AI models.

Convolutional Neural Network (CNN)

A type of deep learning model that processes data with a grid-like topology (e.g., an image) by applying a series of filters. Such models are often used for image recognition tasks.

Data Augmentation:

The process of increasing the amount and diversity of data used for training a model by adding slightly modified copies of existing data.

Deep Learning

A subfield of machine learning that focuses on training neural networks with many layers, enabling learning of complex patterns.

Diffusion

In AI and machine learning, a technique used for generating new data by starting with a piece of real data and adding random noise. A diffusion model is a type of generative model in which a neural network is trained to predict the reverse process when random noise is added to data. Diffusion models are used to generate new samples of data that are similar to the training data.

Double Descent

A phenomenon in machine learning in which model performance improves with increased complexity, then worsens, then improves again.

Embedding

The representation of data in a new form, often a vector space. Similar data points have more similar embeddings.

Emergence/Emergent Behavior (“sharp left turns,” intelligence explosions)

In AI, emergence refers to complex behavior arising from simple rules or interactions. “Sharp left turns” and “intelligence explosions” are speculative scenarios where AI development takes sudden and drastic shifts, often associated with the arrival of AGI.

End-to-End Learning

A type of machine learning model that does not require hand-engineered features. The model is simply fed raw data and expected to learn from these inputs.

Expert Systems

An application of artificial intelligence technologies that provides solutions to complex problems within a specific domain.

Explainable AI (XAI)

A subfield of AI focused on creating transparent models that provide clear and understandable explanations of their decisions.

Fine-tuning

The process of taking a pre-trained machine learning model that has already been trained on a large dataset and adapting it for a slightly different task or specific domain. During fine-tuning, the model's parameters are further adjusted using a smaller, task-specific dataset, allowing it to learn task-specific patterns and improve performance on the new task.

Forward Propagation

In a neural network, forward propagation is the process where input data is fed into the network and passed through each layer (from the input layer to the hidden layers and finally to the output layer) to produce the output. The network applies weights and biases to the inputs and uses activation functions to generate the final output.

Foundation Model

Large AI models trained on broad data, meant to be adapted for specific tasks.

General Adversarial Network (GAN)

A type of machine learning model used to generate new data similar to some existing data. It pits two neural networks against each other: a “generator,” which creates new data, and a “discriminator” which tries to distinguish that data from real data.

Generative AI

A branch of AI focused on creating models that can generate new and original content, such as images, music, or text, based on patterns and examples from existing data.

GPT (Generative Pretrained Transformer)

A large-scale AI language model developed by OpenAI that generates human-like text.

GPU (Graphics Processing Unit)

A specialized type of microprocessor primarily designed to quickly render images for output to a display. GPUs are also highly efficient at performing the calculations needed to train and run neural networks.

Gradient Descent

In machine learning, gradient descent is an optimization method that gradually adjusts a model's parameters based on the direction of largest improvement in its loss function. In linear regression, for example, gradient descent helps find the best-fit line by repeatedly refining the line's slope and intercept to minimize prediction errors.

Hallucinate/Hallucination

In the context of AI, hallucination refers to the phenomenon in which a model generates content that is not based on actual data or is significantly different from reality.

Hidden Layer

Layers of artificial neurons in a neural network that are not directly connected to the input or output.

Hyperparameter Tuning

The process of selecting the appropriate values for the hyperparameters (parameters that are not learned from the data) of a machine learning model.

Inference

The process of making predictions with a trained machine learning model.

Instruction Tuning

A technique in machine learning where models are fine-tuned based on specific instructions given in the dataset.

Large Language Model (LLM)

A type of AI model that can generate human-like text and is trained on a broad dataset.

Latent Space

In machine learning, this term refers to the compressed representation of data that a model (like a neural network) creates. Similar data points are closer in latent space.

Loss Function (or Cost Function)

A function that a machine learning model seeks to minimize during training. It quantifies how far the model's predictions are from the true values.

Machine Learning

A type of artificial intelligence that provides systems the ability to automatically learn and improve from experience without being explicitly programmed.

Mixture of Experts

A machine learning technique where several specialized submodels (the “experts”) are trained, and their predictions are combined in a way that depends on the input.

Multimodal

In AI, this refers to models that can understand and generate information across several types of data, such as text and images.

Natural Language Processing (NLP)

A subfield of AI focused on the interaction between computers and humans through natural language. The ultimate objective of NLP is to read, decipher, understand, and make sense of human language in a valuable way.

NeRF (Neural Radiance Fields)

A method for creating a 3D scene from 2D images using a neural network. It can be used for photorealistic rendering, view synthesis, and more.

Neural Network

A type of AI model inspired by the human brain. It consists of connected units or nodes—called neurons—that are organized in layers. A neuron takes inputs, does some computation on them, and produces an output.

Objective Function

A function that a machine learning model seeks to maximize or minimize during training.

Overfitting

A modeling error that occurs when a function is too closely fit to a limited set of data points, resulting in poor predictive performance when applied to unseen data.

Parameters

In machine learning, parameters are the internal variables that the model uses to make predictions. They are learned from the training data during the training process. For example, in a neural network, the weights and biases are parameters.

Pre-training

The initial phase of training a machine learning model where the model learns general features, patterns, and representations from the data without specific knowledge of the task it will later be applied to. This unsupervised or semi-supervised learning process enables the model to develop a foundational understanding of the underlying data distribution and extract meaningful features that can be leveraged for subsequent fine-tuning on specific tasks.

Prompt

The initial context or instruction that sets the task or query for the model.

Regularization

In machine learning, regularization is a technique used to prevent overfitting by adding a penalty term to the model's loss function. This penalty discourages the model from excessively relying on complex patterns in the training data, promoting more generalizable and less prone-to-overfitting models.

Reinforcement Learning

A type of machine learning where an agent learns to make decisions by taking actions in an environment to maximize some reward.

RLHF (Reinforcement Learning from Human Feedback)

A method to train an AI model by learning from feedback given by humans on model outputs.

Singularity

In the context of AI, the singularity (also known as the technological singularity) refers to a hypothetical future point in time when technological growth becomes uncontrollable and irreversible, leading to unforeseeable changes to human civilization.

Supervised Learning

A type of machine learning where the model is provided with labeled training data.

Symbolic Artificial Intelligence

A type of AI that utilizes symbolic reasoning to solve problems and represent knowledge.

TensorFlow

An open-source machine learning platform developed by Google that is used to build and train machine learning models.

TPU (Tensor Processing Unit)

A type of microprocessor developed by Google specifically for accelerating machine learning workloads.

Training Data

The dataset used to train a machine learning model.

Transfer Learning

A method in machine learning where a pre-trained model is used on a new problem.

Transformer

A specific type of neural network architecture used primarily for processing sequential data such as natural language. Transformers are known for their ability to handle long-range dependencies in data, thanks to a mechanism called “attention,” which allows the model to weigh the importance of different inputs when producing an output.

Underfitting

A modeling error in statistics and machine learning when a statistical model or machine learning algorithm cannot adequately capture the underlying structure of the data.

Unsupervised Learning

A type of machine learning where the model is not provided with labeled training data, and instead must identify patterns in the data on its own.

Validation Data

A subset of the dataset used in machine learning that is separate from the training and test datasets. It's used to tune the hyperparameters (i.e., architecture, not weights) of a model.

XAI (Explainable AI)

A subfield of AI focused on creating transparent models that provide clear and understandable explanations of their decisions.

Zero-shot Learning

A type of machine learning where the model makes predictions for conditions not seen during training, without any fine-tuning.

Software is eating the world

[Terms of Use](#)

[Conduct](#)

[Privacy Policy](#)

[Disclosures](#)



© 2025 Andreessen Horowitz