AI Basics

AI Basics

What Is Artificial Intelligence?

Artificial intelligence, or AI, is technology developed with the goal of performing tasks typically requiring human intelligence, such as recognizing patterns, making decisions, and communicating. AI refers to the general ability of computers to emulate human thought and perform tasks in real-world environments. Most AI mimics human cognitive abilities by creating a model of the world based on a limited set of information provided to it, and then guessing how the world it has modeled would behave. Generative AI can even generate content that it would expect to exist in that world, in response to prompts from humans. Because of the many, often hidden layers of calculations and decisions that lead to the predictions AI systems make, AI is often referred to as a "black box." (Source: www_library_jhu_edu__9981251c7844ce31::c0000)

Artificial intelligence is the field of developing computers and robots that are capable of behaving in ways that both mimic and go beyond human capabilities. AI-enabled programs can analyze and contextualize data to provide information or automatically trigger actions without human interference. Today, artificial intelligence is at the heart of many technologies we use, including smart devices and voice assistants such as Siri on Apple devices. (Source: ai_engineering_columbia_edu__8662d32a1c231998::c0000, ai_engineering_columbia_edu__8662d32a1c231998::c0001)

Machine Learning and Related Terms

Machine learning is a subset of the broader category of artificial intelligence. Machine learning refers to the technologies and algorithms that enable systems to identify patterns, make decisions, and improve themselves through experience and data. Machine learning is a pathway to artificial intelligence. This subcategory of AI uses algorithms to automatically learn insights and recognize patterns from data, applying that learning to make increasingly better decisions. (Source: ai_engineering_columbia_edu__8662d32a1c231998::c0000, ai_engineering_columbia_edu__8662d32a1c231998::c0001)

An algorithm is the foundation of all AI tools. An algorithm is a series of steps that allows a computer to accomplish a certain task. A simple example is a set of instructions for sorting a list of numbers. Machine learning is the branch of artificial intelligence that teaches computers how to find rules and patterns in a large amount of information in order to make predictions about previously unseen information. (Source: www_library_jhu_edu__9981251c7844ce31::c0001)

Deep learning is an advanced method of machine learning. Deep learning models use large neural networks—networks that function like a human brain to logically analyze data—to learn complex patterns

and make predictions independent of human input. Deep learning is a subset of machine learning that uses multilayered neural networks, called deep neural networks, that more closely simulate the complex decision-making power of the human brain. (Source: ai_engineering_columbia_edu__8662d32a1c231998::c0001, www_ibm_com__4c73255b7dd33ab9::c0002)

Large language models, or LLMs, are machine learning models designed for natural language processing, particularly language generation. LLMs are fed a vast amount of human-written text, often sourced from the internet, which they analyze to map out the probabilities underlying human language. LLMs use these models of likelihood to mimic human communication and can accomplish language-related tasks like text summarization, language translation, and question answering. Large language models form the foundation for Generative AI tools like Microsoft's CoPilot, OpenAI's GPTs, Meta's LLaMA, xAI's Grok, and Google's Gemini. (Source: www_library_jhu_edu__9981251c7844ce31::c0002, www_library_jhu_edu__9981251c7844ce31::c0003)

How Modern AI Systems Work (High-Level)

Machine learning starts with data—numbers, photos, or text, like bank transactions, pictures of people, repair records, time series data from sensors, or sales reports. The data is gathered and prepared to be used as training data, or the information the machine learning model will be trained on. According to some sources, more data generally leads to better program performance. From there, programmers choose a machine learning model to use, supply the data, and let the computer model train itself to find patterns or make predictions. Over time, human programmers can also tweak the model, including changing its parameters, to help push it toward more accurate results. (Source: mitsloan_mit_edu__b85f152879c15db1::c0003)

Large language models have become so sophisticated by utilizing deep learning, an approach to machine learning in which a model calibrates itself to give more weight to computational tasks that help it perform best by comparing many different approaches to making predictions. This technique is known as an "artificial neural network" because it mimics how the human brain processes information and requires large amounts of data and computational power. (Source: www_library_jhu_edu__9981251c7844ce31::c0002)

Generative AI operates differently from earlier AI systems. Instead of just reacting to data input, generative AI systems take in data and then use predictive algorithms to create original content. In the case of a large language model, that content can take the form of original poems, songs, screenplays, and the like. According to some researchers, the model is just predicting the next word and doesn't understand in the way humans do. But as a user playing around with it, it seems to have amazing capabilities, while having very large blind spots. (Source: www_heinz_cmu_edu__8bda06f17ee8ff1a::c0004, www_heinz_cmu_edu__8bda06f17ee8ff1a::c0005)

Common Misconceptions About AI

A commonly discussed concern is that AI systems can produce inaccurate or fabricated information. According to some researchers, large language models like ChatGPT sometimes hallucinate, meaning a user enters a prompt and the system makes up an answer that's not true in some way. The system might produce an intelligent-sounding essay and cite as its source a scholarly research paper that doesn't actually exist. Sometimes the answer is just inaccurate. To complicate matters, the information is presented with confidence and authority; it looks and sounds legitimate. This is a documented limitation that can make it difficult for

users to know whether the answer is reliable. (Source: www_heinz_cmu_edu__8bda06f17ee8ff1a::c0003, www_heinz_cmu_edu__8bda06f17ee8ff1a::c0006, www_heinz_cmu_edu__8bda06f17ee8ff1a::c0007)

Another commonly discussed concern is that AI systems can perpetuate bias and discrimination. Large language models are trained on large quantities of data, much of which is scraped from the Internet. That data includes reliable sources right alongside problematic content. According to some researchers, machines are trained by humans, and human biases can be incorporated into algorithms—if biased information, or data that reflects existing inequities, is fed to a machine learning program, the program will learn to replicate it and perpetuate forms of discrimination. In some documented cases, AI systems have exhibited bias, such as hiring tools that discriminate against women or facial recognition software that doesn't recognize people of color. Bias inherent in an AI model has the potential to exacerbate existing injustice. (Source: www_heinz_cmu_edu__8bda06f17ee8ff1a::c0008, mitsloan_mit_edu__b85f152879c15db1::c0012)