Artificial intelligence and machine learning are two buzzwords that are commonly used in the world of technology. Although they are often used interchangeably, they are not the same thing. Artificial intelligence (AI) and machine learning (ML) are related concepts, but they have different definitions, applications, and implications.

What is Machine Learning?

Machine learning is a subset of artificial intelligence that focuses on teaching machines how to learn from data. In other words, machine learning is a process by which computers can automatically learn patterns and relationships in data without being explicitly programmed to do so. Machine learning algorithms are designed to detect and learn from patterns in data to make predictions or decisions.

There are three main types of machine learning: supervised learning, unsupervised learning, and reinforcement learning. Supervised learning is when the machine is trained on labeled data with known outcomes. Unsupervised learning is when the machine is trained on unlabeled data and is asked to find patterns or similarities. Reinforcement learning is when the machine learns by trial and error through interactions with the environment.

Examples of machine learning include image recognition, speech recognition, recommendation systems, fraud detection, and natural language processing.

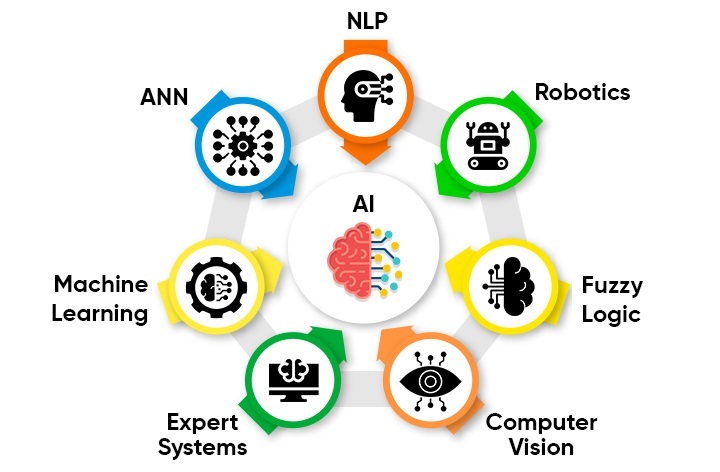
AD

What is Artificial Intelligence?

Artificial intelligence, on the other hand, is a broad field that encompasses the development of intelligent machines that can perform tasks that typically require human intelligence, such as perception, reasoning, learning, and decision-making. In simple terms, AI is the ability of machines to perform tasks that normally require human intervention or intelligence.

There are two types of AI: narrow or weak AI and general or strong AI. Narrow AI is designed to perform specific tasks, such as speech recognition or image recognition, while general AI is designed to be able to perform any intellectual task that a human can do. Currently, we only have narrow AI in use, but the goal is to develop general AI that can be applied to a wide range of tasks.

AI is like a basket containing several branches, the important ones being Machine Learning (ML), Robotics, Expert Systems, Fuzzy Logic, Neural Networks, Computer Vision, and Natural Language Processing (NLP).



While we highlight the features of ML in the next section, here is a brief overview of the other important branches of AI:

* **Robotics**− Robots are primarily designed to perform repetitive and tedious tasks. Robotics is an important branch of AI that deals with designing, developing and controlling the application of robots.
* **Computer Vision** − It is an exciting field of AI that helps computers, robots, and other digital devices to process and understand digital images and videos, and extract vital information. With the power of AI, Computer Vision develops algorithms that can extract, analyze and comprehend useful information from digital images.
* **Expert Systems** − Expert systems are applications specifically designed to solve complex problems in a specific domain, with humanlike intelligence, precision, and expertise. Just like human experts, Expert Systems excel in a specific domain in which they are trained.
* **Fuzzy Logic** − We know computers take precise digital inputs like True (Yes) or False (No), but Fuzzy Logic is a method of reasoning that helps machines to reason like human beings before taking a decision. With Fuzzy Logic, machines can analyze all intermediate possibilities between a YES or NO, for example, "Possibly Yes", "Maybe No", etc.
* **Neural Networks** − Inspired by the natural neural networks of the human brain, Artificial Neural Networks (ANN) can be considered as a group of highly interconnected group of processing elements (nodes) that can process information by their dynamic state response to external inputs. ANNs use training data to improve their efficiency and accuracy.
* **Natural Language Processing (NLP)** − NLP is a field of AI that empowers intelligent systems to communicate with humans using a natural language like English. With the power of NLP, one can easily interact with a robot and instruct it in plain English to perform a task. NLP can also process text data and comprehend its full meaning. It is heavily used these days in virtual chatbots and sentiment analysis.

Examples of AI include virtual assistants, autonomous vehicles, facial recognition, natural language processing, and decision-making systems.

AD

Machine Learning vs. Artificial Intelligence

Now that we have a basic understanding of what machine learning and artificial intelligence are, let's dive deeper into the differences between the two.

Firstly, machine learning is a subset of artificial intelligence, meaning that machine learning is a part of the larger field of AI. Machine learning is a technique used to implement artificial intelligence.

Secondly, while machine learning focuses on developing algorithms that can learn from data, artificial intelligence focuses on developing intelligent machines that can perform tasks that normally require human intelligence. In other words, machine learning is more focused on the process of learning from data, while AI is more focused on the end goal of creating machines that can perform intelligent tasks.

Thirdly, machine learning algorithms are designed to learn from data and improve their accuracy over time, while artificial intelligence systems are designed to learn and adapt to new situations and environments. Machine learning algorithms require a lot of data to be trained effectively, while AI systems can adapt and learn from new data in real-time.

Finally, machine learning is more limited in its capabilities compared to AI. Machine learning algorithms can only learn from the data they are trained on, while AI systems can learn and adapt to new situations and environments. Machine learning is great for solving specific problems that can be solved through pattern recognition, while AI is better suited for complex, real-world problems that require reasoning and decision-making.

The following table highlights the important differences between Machine Learning and Artificial Intelligence −

|  |  |  |
| --- | --- | --- |
| **Key** | **Artificial Intelligence** | **Machine Learning** |
| Definition | AI refers to the ability of a machine or a computer system to perform tasks that would normally require human intelligence, such as understanding language, recognizing images, and making decisions. | ML is a type of AI that allows a system to learn and improve from experience without being explicitly programmed.  It articulates how a machine can learn and apply its knowledge to improve its decisions. |
| Concept | AI revolves around making smart and intelligent devices. | ML revolves around making a machine learn/decide and improve its results. |
| Goal | The goal of AI is to simulate human intelligence to solve complex problems. | The goal of ML is to learn from data provided and make improvements in machine's performance. |
| Includes | AI has several important branches including Artificial Neural Networks, Natural Language Processing, Fuzzy Logic, Robotics, Expert Systems, Computer Vision, and Machine Learning. | ML training methods include supervised learning, unsupervised learning, and reinforcement learning. |
| Development | AI is leading to the development of such machines which can mimic human behavior. | ML is helping in the development of self-learning algorithms. |

**Machine Learning - Neural Networks**

Machine learning and neural networks are two important technologies in the field of artificial intelligence (AI). While they are often used together, they are not the same thing. In this article, we will explore the differences between machine learning and neural networks and how they are related.

We understood about machine learning in last section so let's see what neural networks are.

What are Neural Networks?

Neural networks are a type of machine learning algorithm that is inspired by the structure of the human brain. They are designed to simulate the way the brain works by using layers of interconnected nodes, or artificial neurons. Each neuron takes in input from the neurons in the previous layer and uses that input to produce an output. This process is repeated for each layer until a final output is produced.

Neural networks can be used for a wide range of tasks, including image recognition, speech recognition, natural language processing, and prediction. They are particularly well-suited to tasks that involve processing complex data or recognizing patterns in data.

Machine Learning vs. Neural Networks

Now that we have a basic understanding of what machine learning and neural networks are, let's dive deeper into the differences between the two.

* Firstly, machine learning is a broad category that encompasses many different types of algorithms, including neural networks. Neural networks are a specific type of machine learning algorithm that is designed to simulate the way the brain works.
* Secondly, while machine learning algorithms can be used for a wide range of tasks, neural networks are particularly well-suited to tasks that involve processing complex data or recognizing patterns in data. Neural networks can recognize complex patterns and relationships in data that other machine learning algorithms may not be able to detect.
* Thirdly, neural networks require a lot of data and processing power to train. Neural networks typically require large datasets and powerful hardware, such as graphics processing units (GPUs), to train effectively. Machine learning algorithms, on the other hand, can be trained on smaller datasets and less powerful hardware.
* Finally, neural networks can provide highly accurate predictions and decisions, but they can be more difficult to understand and interpret than other machine learning algorithms. The way that neural networks make decisions is not always transparent, which can make it difficult to understand how they arrived at their conclusions.

**Machine Learning - Deep Learning**

In the world of artificial intelligence, two terms that are often used interchangeably are machine learning and deep learning. While both of these technologies are used to create intelligent systems, they are not the same thing. In this article, we will explore the differences between machine learning and deep learning and how they are related.

We understood about machine learning in last section so let's see what deep learning is.

What is Deep Learning?

Deep learning is a type of machine learning that uses neural networks to process complex data. In other words, deep learning is a process by which computers can automatically learn patterns and relationships in data using multiple layers of interconnected nodes, or artificial neurons. Deep learning algorithms are designed to detect and learn from patterns in data to make predictions or decisions.

Deep learning is particularly well-suited to tasks that involve processing complex data, such as image and speech recognition, natural language processing, and self-driving cars. Deep learning algorithms are able to process vast amounts of data and can learn to recognize complex patterns and relationships in that data.

Examples of deep learning include facial recognition, voice recognition, and self-driving cars.

Machine Learning vs. Deep Learning

Now that we have a basic understanding of what machine learning and deep learning are, let's dive deeper into the differences between the two.

* Firstly, machine learning is a broad category that encompasses many different types of algorithms, including deep learning. Deep learning is a specific type of machine learning algorithm that uses neural networks to process complex data.
* Secondly, while machine learning algorithms are designed to learn from data and improve their accuracy over time, deep learning algorithms are designed to process complex data and recognize patterns and relationships in that data. Deep learning algorithms are able to recognize complex patterns and relationships that other machine learning algorithms may not be able to detect.
* Thirdly, deep learning algorithms require a lot of data and processing power to train. Deep learning algorithms typically require large datasets and powerful hardware, such as graphics processing units (GPUs), to train effectively. Machine learning algorithms, on the other hand, can be trained on smaller datasets and less powerful hardware.
* Finally, deep learning algorithms can provide highly accurate predictions and decisions, but they can be more difficult to understand and interpret than other machine learning algorithms. Deep learning algorithms can process vast amounts of data and recognize complex patterns and relationships in that data, but it can be difficult to understand how the algorithm arrived at its conclusion.

# Generative AI & it's Applications

## What is generative AI?

Generative AI is artificial intelligence designed to create unique text or image results in response to user prompts. The technology uses machine learning to return an output based on the user’s prompt. AI engineers train the technology using large data sets, which the model consults when determining the best possible answer to a prompt. Another way to look at generative AI is as a form of predictive artificial intelligence. Based on the information provided, generative AI will predict which words and in which order will give the best answer to the user's prompts.

You can use generative AI to create new written, visual, or audio content, summarize complex data, generate code, assist with repetitive tasks, or make customer service more personalized.

### Examples of generative AI

Examples of generative artificial intelligence that you may have heard of include Google’s Bard, ChatGPT, or DALL-E from OpenAI.

* **ChatGPT or DALL-E:** Generative artificial intelligence created by OpenAI, a Microsoft-backed, profit-capped company with the mission to develop artificial intelligence to serve humankind
* **Google Bard:**Google’s generative AI with integrations to Google products like Google Lens and Gmail, operating with a language model called PaLM-2 that was trained on the largest data set out of all generative AI models available at the time of its release

## Applications of generative AI

Generative artificial intelligence has applications in diverse industries such as health care, manufacturing, software development, financial services, media and entertainment, and advertising and marketing. Let’s examine some of the different ways professionals in these industries apply generative AI to their field.

### Health care and pharmaceuticals

Generative artificial intelligence has applications for all parts of the health care and pharmaceutical industry, from discovering and developing new life-saving medicine to personalizing treatment plans for individual patients to creating predictive images for charting disease progression. Some of the possibilities for generational AI in health care include:

* **Enhancing medical images:** Generative AI can augment medical images like X-rays or MRIs, synthesize images, reconstruct images, or create reports about images. This technology can even generate new images to demonstrate how a disease may progress in time.
* **Discovering new drugs**: Researchers can use generative artificial intelligence via a related field called generative design to research and develop new medicines. Gartner projects that 30 percent of the new drugs created by researchers in 2025 will use generative design principles [[1](https://www.gartner.com/en/topics/generative-ai)].
* **Simplify tasks with patient notes and information**: Healthcare professionals keep and take notes about patient medical care. Generational AI can build patient information summaries, create transcripts of verbally recorded notes, or find essential details in medical records more effectively than human efforts.
* **Personalized treatment:** Generative AI can consider a large amount of patient information, including medical images and genetic testing, to deliver a customized treatment plan tailored to the patient's needs.

### Advertising and marketing

Generative artificial intelligence offers many solutions to professionals working in advertising and marketing, such as generating text and images needed for marketing or finding new ways to interact with customers. Here are some examples of generative AI applications in advertising and marketing:

* **Generate marketing text and images:**Generative AI can help marketing professionals create consistent, on-brand text and images to use in marketing campaigns. This technology also offers translation tools to spread your marketing message into new territories. Gartner predicts that marketing professionals will use generative AI to create 30 percent of outbound marketing materials by 2025 [[1](https://www.gartner.com/en/topics/generative-ai)].
* **Generate personalized recommendations:** Generative AI helps create powerful recommendation engines to help customers discover new products they might like. With generative AI, this process is more interactive for customers.
* **Create product descriptions:** Beyond flashy advertising campaigns, generative artificial intelligence can help with tedious or time-consuming content requirements like creating product descriptions.
* **Enhance search engine optimization:** SEO professionals can use generative AI for tasks like image tags or page titles or to create content drafts. You could also use a tool like ChatGPT or Bard to recommend changes you could make to content to improve SEO ranking.

### Manufacturing

In manufacturing, professionals can use generative AI to look for ways to improve efficiency, anticipate maintenance needs before they cause problems, help engineers create better designs faster, and create a more resilient supply chain. Let’s explore these potential manufacturing solutions:

* **Accelerating the design process:** Using generative AI, engineers and project managers can work through the design process much faster by generating design ideas and asking the AI to assess ideas based on the constraints of the project.
* **Provide smart maintenance solutions for equipment:** Maintenance professionals can use generative AI to track the performance of heavy equipment based on historical data, potentially alerting them to trouble before the machine malfunctions. Generative AI can also recommend routine maintenance schedules.
* **Improve supply chain**: You could use generative AI to track down the cause of problems in the supply chain by speaking conversationally with the technology to sort through a vast amount of transactional or product data. Generative AI can also help generate delivery schedules or recommendations for suppliers.

### Software development

For a software development team, generative AI can provide tools to create and optimize code faster and with less experience using programming languages. A few examples of the applications of generative AI in software development include:

* **Generating code:**Software developers can create, optimize, and auto-complete code with generative AI. Generative AI can create code blocks by comparing them to a library of similar information. It can also predict the rest of the code a developer begins to type, much like how auto-complete works while texting on a smartphone.
* **Translate programming languages:** Generative AI can be a tool for developers to interact with software without needing a programming language. The generative AI would act as a translator.
* **Automate testing:**Developers can improve their automated testing processes using generative AI to highlight potential problems and execute testing sequences faster than other AI methods. Generative AI can learn the logic of the software and how users will interact with it and create test cases to demonstrate various user scenarios.

### Financial services

According to McKinsey, generative AI could add $200 billion to $340 billion of value to the banking industry annually [[2](https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier)]. Some of the applications of generative AI in the financial services industry include artificial intelligence investment strategies, drafting documentation and monitoring regulatory changes, and using generative AI as an interpreter to facilitate communications between clients and investors.

* **Create investment strategies:** Generative AI can recommend the best investments according to your or your client’s goals. This technology can find and execute trades much faster than human investors and can do so within the parameters you set for the kind of transaction you want.
* **Communicate and educate clients and investors:** Financial services professionals sometimes need to communicate complex information to clients and colleagues. Generational AI can provide hyperpersonalized customer service without adding more customer service professionals.
* **Quickly draft documentation and monitor regulation:** Generative AI can monitor regulatory activity, keep you informed of any changes, and create drafts of documents such as investment research or insurance policies.

### Media and entertainment

Media and entertainment could embrace generative AI in several ways, considering the industry primarily engages in the same task as the tech: generating unique content. Generative AI can help create and edit visual content, create short highlight videos of sporting events, and make working with content management systems easier.

* **Create audio and visual content:** Generative AI can create new video content from scratch. This tech can also help you make visual content faster by creating visual effects, adding graphics, or streamlining editing.
* **Generate highlights for sports and events:** When it comes to sporting and live events, gen AI can create highlight reels instantly and allow fans to create their own custom highlights. For example, fans could generate highlights of a particular play or a tournament series.
* **Manage tags for better content management:**Generative AI can tag and index extensive media libraries, making locating the files you need at any time easier. Similar to our manufacturing example above, generative AI allows using conversational language to find the information or media you’re looking for in a complex media library.

## How to find solutions with generative AI

If you’re interested in bringing generative AI to your company, you can approach the technology in two ways. First, you can use existing models and learn to engineer prompts to your needs. Or, you can customize solutions to fit your business processes.

* **You can use existing generative AI tools like ChatGPT**. In this scenario, you’ll focus on learning how to write prompts that get the best answer possible from the technology. For example, you might identify who your audience is and the appropriate tone of the piece to help the application deliver the correct results.
* **You can integrate custom solutions from an enterprise-level company or build your own generative AI tools**. While it won’t be feasible or practical for many companies to create their own generative AI solutions, many gen AI companies offer solutions you can tailor to your business needs. Generative models will vary on features, cost, and security or privacy standards.

## Top current LLMs

Below are some of the most relevant large language models today. They do natural language processing and influence the architecture of future models.

### BERT

[BERT](https://www.techtarget.com/searchenterpriseai/definition/BERT-language-model) is a family of LLMs that Google introduced in 2018. BERT is a [transformer-based](https://www.techtarget.com/searchenterpriseai/tip/GAN-vs-transformer-models-Comparing-architectures-and-uses) model that can convert sequences of data to other sequences of data. BERT's architecture is a stack of transformer encoders and features 342 million parameters. BERT was pre-trained on a large corpus of data then fine-tuned to perform specific tasks along with natural language inference and sentence text similarity. It was used to improve query understanding in the 2019 iteration of Google search.

### Claude

The [Claude LLM](https://www.techtarget.com/searchenterpriseai/feature/Claude-AI-vs-ChatGPT-How-do-they-compare) focuses on constitutional AI, which shapes AI outputs guided by a set of principles that help the AI assistant it powers helpful, harmless and accurate. Claude was created by the company Anthropic. The latest iteration of the Claude LLM is Claude 3.5 Sonnet. It understands nuance, humor and complex instructions better than earlier versions of the LLM, and operates at twice the speed of Claude 3 Opus. It’s available for free via Claude.ai and the Claude iOS app.

### Cohere

Cohere is an enterprise AI platform that provides several LLMs including Command, Rerank and Embed. These [LLMs can be custom-trained](https://www.techtarget.com/searchenterpriseai/tip/How-to-train-an-LLM-on-your-own-data) and fine-tuned to a specific company’s use case. The company that created the Cohere LLM was founded by one of the authors of Attention Is All You Need. One of Cohere’s strengths is that it is not tied to one single cloud -- unlike OpenAI, which is bound to Microsoft [Azure](https://www.techtarget.com/searchcloudcomputing/definition/Windows-Azure).

### Ernie

Ernie is Baidu’s large language model which powers the Ernie 4.0 chatbot. The bot was released in August 2023 and has garnered more than 45 million users. Ernie is rumored to have 10 trillion parameters. The bot works best in Mandarin but is capable in other languages.

### Falcon 40B

Falcon 40B is a transformer-based, causal decoder-only model developed by the Technology Innovation Institute. It is open source and was trained on English data. The model is available in two smaller variants as well: Falcon 1B and Falcon 7B (1 billion and 7 billion parameters). Ama

# Prompt Engineering - Introduction

Prompt engineering is the process of crafting text prompts that help large language models (LLMs) generate more accurate, consistent, and creative outputs. By carefully choosing the words and phrases in a prompt, prompt engineers can influence the way that an LLM interprets a task and the results that it produces.

## **What are Prompts?**

In the context of AI models, prompts are input instructions or cues that shape the model's response. These prompts can be in the form of natural language instructions, system-defined instructions, or conditional constraints.

* A prompt is a short piece of text that is used to guide an LLM's response. It can be as simple as a single sentence, or it can be more complex, with multiple clauses and instructions.
* The goal of a prompt is to provide the LLM with enough information to understand what is being asked of it, and to generate a relevant and informative response.

By providing clear and explicit prompts, developers can guide the model's behavior and influence the generated output.

AD

## **Types of Prompts**

There can be wide variety of prompts which you will get to know during the course of this tutorial. This being an introductory chapter, let's start with a small set to highlight the different types of prompts that one can use −

* **Natural Language Prompts** − These prompts emulate human-like instructions, providing guidance in the form of natural language cues. They allow developers to interact with the model more intuitively, using instructions that resemble how a person would communicate.
* **System Prompts** − System prompts are predefined instructions or templates that developers provide to guide the model's output. They offer a structured way of specifying the desired output format or behavior, providing explicit instructions to the model.
* **Conditional Prompts** − Conditional prompts involve conditioning the model on specific context or constraints. By incorporating conditional prompts, developers can guide the model's behavior based on conditional statements, such as "If X, then Y" or "Given A, generate B."

AD

## **How Does Prompt Engineering Work?**

Prompt engineering is a complex and iterative process. There is no single formula for creating effective prompts, and the best approach will vary depending on the specific LLM and the task at hand. However, there are some general principles that prompt engineers can follow −

* **Start with a clear understanding of the task** − What do you want the LLM to do? What kind of output are you looking for? Once you have a clear understanding of the task, you can start to craft a prompt that will help the LLM achieve your goals.
* **Use clear and concise language** − The LLM should be able to understand your prompt without any ambiguity. Use simple words and phrases, and avoid jargon or technical terms.
* **Be specific** − The more specific you are in your prompt, the more likely the LLM is to generate a relevant and informative response. For example, instead of asking the LLM to "write a poem," you could ask it to "write a poem about a lost love."
* **Use examples** − If possible, provide the LLM with examples of the kind of output you are looking for. This will help the LLM to understand your expectations and to generate more accurate results.
* **Experiment** − There is no one-size-fits-all approach to prompt engineering. The best way to learn what works is to experiment with different prompts and see what results you get.

AD

## **Evaluating and Validating Prompts**

Evaluating prompt effectiveness is crucial to assess the model's behavior and performance. Metrics such as output quality, relevance, and coherence can help evaluate the impact of different prompts. User feedback and human evaluation can provide valuable insights into prompt efficacy, ensuring the desired output is achieved consistently.

AD

## **Ethical Considerations in Prompt Engineering**

Prompt engineering should address ethical considerations to ensure fairness and mitigate biases. Designing prompts that promote inclusivity and diversity while avoiding the reinforcement of existing biases is essential.

Careful evaluation and monitoring of prompt impact on the model's behavior can help identify and mitigate potential ethical risks.

AD

## **Benefits of Prompt Engineering**

Prompt engineering can be a powerful tool for improving the performance of LLMs. By carefully crafting prompts, prompt engineers can help LLMs to generate more accurate, consistent, and creative outputs. This can be beneficial for a variety of applications, including −

* **Question answering** − Prompt engineering can be used to improve the accuracy of LLMs' answers to factual questions.
* **Creative writing** − Prompt engineering can be used to help LLMs generate more creative and engaging text, such as poems, stories, and scripts.
* **Machine translation** − Prompt engineering can be used to improve the accuracy of LLMs' translations between languages.
* **Coding** − Prompt engineering can be used to help LLMs generate more accurate and efficient code.

AD

## **Future Directions and Open Challenges**

Prompt engineering is an evolving field, and there are ongoing research efforts to explore its potential further. Future directions may involve automated prompt generation techniques, adaptive prompts that evolve with user interactions, and addressing challenges related to nuanced prompts for complex tasks.

Prompt engineering is a powerful tool in enhancing AI models and achieving desired outputs. By employing effective prompts, developers can guide the behavior of AI models, control biases, and improve the overall performance and reliability of AI applications.

As the field progresses, continued exploration of prompt engineering techniques and best practices will pave the way for even more sophisticated and contextually aware AI models.