NeurIPS 2025 Education Material

Understanding Artificial Intelligence:
An In-Depth Guide for Non-Expert Readers

Gedeon Hakizimana[1]
Agapito Ledezma Espino[2]

[1]Doctoral Researcher, Universidad Carlos III de Madrid, 100476007@alumnos.uc3m.es
[2]Professor, Universidad Carlos III de Madrid, ledezma@inf.uc3m.es
Department of Computer Science & Engineering
Av. Universidad, 30, 28911 Leganés (Madrid), Spain

Target Audience:
This material is created for a **non-technical audience**, including:

- High school students and early undergraduates from non-CS fields
- Educators introducing AI in general science or digital literacy courses
- Curious professionals or citizens wanting to understand AI without needing coding or math backgrounds.

Expected Reading Time:

12–15 minutes

Brief Description of the Material

This educational handout introduces readers to the world of Artificial Intelligence (AI) through simple language, real-life analogies, and clear structure. Designed for those without a technical background, the guide demystifies AI concepts such as:

- What AI is, its main categories and how they work
- How machines "learn" from data
- The different types of learning: **supervised, unsupervised, semi-supervised, self-supervised,** and **reinforcement learning**
- Classic learning algorithms like **Linear Regression**, **Decision Trees**, and **Naive Bayes**
- Advanced methods like **Neural Networks**, **CNNs, RNNs**, and **Transformers**
- Real-world examples (e.g., facial recognition, voice assistants, recommendation engines)
- Ethical concerns like data bias, privacy, and AI transparency
- Limitations of AI and its future.

By the end, readers are equipped with the foundational understanding needed to engage with current AI topics and evaluate their impact on society.

1. Introduction

What is Really Artificial Intelligence (AI)? *Unveiling the Mystery of AI*

Today, artificial intelligence (AI) has moved from the realm of science fiction and has become a buzzword in our daily lives, from virtual assistants like Siri and Alexa to the recommendation systems behind Netflix, online shopping and YouTube to self-driving cars. But what does it really mean when we say a machine is "intelligent"? How does Artificial Intelligence work? And why does it matter?

In this material, we'll embark on a journey to uncover the true nature of AI—its definition, its key components, types of AI, how it works, its limitations, and how it is already influencing almost all industries, including but not limited to healthcare, education, agriculture, commerce, and beyond. By the end of this material, you'll not only understand what AI is in a simplified language, but also how it fits into the broader picture of technological advancement and its potential to reshape society.

2. The Essence of AI: Understanding the Basics

At its core, **Artificial Intelligence (AI)** refers to machines or software systems designed to mimic human intelligence. Thus, AI enables machines to **learn** and **make decisions** (as if they think) based on data, much like a human would. But it's not just about copying human behavior—it's about improving on it for some aspects of daily activities (for example processing billions of data in a short time in order to discover the patterns or rules hidden in it—a task which usually requires enough time for a human to handle) in some ways, using vast amounts of data, computing power, and algorithms (don't worry if you don't understand some of the terms now, they will be explained later). As a discipline, Artificial Intelligence (AI) refers to the science and technology of creating machines or computer programs capable of performing tasks that typically require human intelligence. These tasks include:

- **Perception** (known as Computer Vision): Understanding the environment, like recognizing images.

  Example: A smartphone camera app that automatically detects faces and adjusts focus or exposure accordingly. Here, by recognizing people's faces in the image, the AI behaves as if it "sees" the environment.

- **Reasoning:** Solving problems or making decisions.

  Example: An AI system in chess that analyzes the current board and predicts the best next move by considering possible future moves and their outcomes.

- **Language Understanding** (Natural Language Processing –NLP): This involves interpreting and generating human language so that computers can understand and respond appropriately.

  Examples:

  - *A virtual assistant like Siri interpreting your spoken command, "**Set an alarm for 7 AM,**" understanding the intent —to create an alarm—, and responding by setting the alarm as you requested.*
  - *ChatGPT writing a professional email or summarizing a long article (text generation). For instance, by typing: "**Write a polite follow-up email about a job interview**", the AI understands the request and generates a complete, coherent message in natural language.*

- **Learning**: Improving performance from experience.

  Example: *A spam email filter which is first trained by example emails labeled as spams, and then classify emails as either spam or not spam, and improves over time by learning from emails marked as spam or not spam, getting better at catching unwanted messages as it sees more examples.*

3. What is Intelligence?

Before we can understand AI, it's helpful to first define intelligence. In human beings, intelligence is ***a complex, multifaceted mental quality encompassing the ability to learn from experience, understand and apply knowledge, adapt to new situations, and reason to solve problems and make decisions.*** Therefore, when we say that a machine is intelligent, we mean that it can perform tasks that typically require these abilities, such as *recognizing faces in photos*, *understanding language*, or *driving a car*. The question is to know how AI compares to human intelligence. Below is the answer:

> Even though AI can be really good at certain tasks, **it doesn't have all the mental qualities, emotions, values, or creativity that humans do**. People are far great at solving many different types of problems and thinking in creative ways. AI, on the other hand, is usually designed to do one specific job really well—like playing chess, suggesting movies you might like, or helping doctors find diseases. That's why we often hear the term **Narrow AI**—it means AI that's built for one clear purpose. **General AI**, which would think and learn like a human across many areas, is the scientists' ambition, something they are still working toward but haven't achieved yet.

4. How AI Works

To truly understand Artificial Intelligence, it's important to step back and look at how these systems actually "**learn**" and "**adapt**"—not in the way humans do emotionally, but by recognizing patterns in data. AI systems are first trained, then they learn according to the received training, and lastly adapt as situations evolve. Following are more details.

4.1. Training

AI systems require training to learn. During this phase, an AI model is provided with a large set of data. For example, if the AI is being trained to recognize pictures of dogs, it will be shown thousands of images of dogs and other objects. Each image can be labeled or unlabeled according to the nature of problem, and the AI learns from this data to distinguish features that define a dog (like shape, size, and texture).

4.2. Learning

Through training, AI systems are taught to find patterns in the data. This is what is called **machine learning** (ML—a machine/computer is taught to do something and then learns how). The more data they are exposed to, the better they can recognize trends and make predictions.

What Does It Mean for AI to "Find Patterns in Data"?

When we say "AI systems find patterns in data", we mean that they learn to recognize relationships, trends, common features (or hidden rules) in the information they're given—often in ways that aren't obvious to humans. Here's a detailed example:

Imagine you're teaching an AI system to tell the difference between pictures of cats and dogs. Below is the process flow:

(i) Data preparation: What Is Data Preparation in AI?

The first thing that must be done before training an AI is data preparation (dataset readiness). This is one of the most important steps in building any AI or machine learning system. It's all about getting your data clean, organized, and ready so the AI can learn from it effectively.

Think of it like cooking: before you start making a meal, you wash, chop, and measure your ingredients. You don't just throw everything into the pot. Similarly, before training an AI model, the data needs to be processed in the right way. In fact, if the data is messy, incomplete, or inconsistent, the AI will learn the wrong things or make poor decisions or even the learning will not be possible. Good data preparation helps:

- Improve the accuracy of learning and then of predictions
- Speed up the training process because there won't be errors related to data organization (or structure)
- Ensure the AI performs well on real-world tasks

The key steps involved in data preparation are:

a) Data Collection

This step involves gathering the right kind of data that matches the machine learning task you're trying to solve. Depending on the nature of the problem, the data may come in various forms like **text, images, numbers, audio, or a combination**. Below are the main types of data you might collect, along with real-world use case examples:

- *Text Data (Natural Language):* Used when the task involves understanding or generating human language. For this type of data, sources may include customer reviews, news articles, support tickets, emails, SMS messages, social media posts, books, etc., according the task at hand.

  Common tasks that use this type of data include:

  - Detecting spam or phishing in emails
  - Translating languages (e.g., English to French, Spanish, etc.)
  - Summarizing long documents
  - Powering chatbots and virtual assistants (e.g., Siri, Alexa)
  - Sentiment analysis (e.g., understanding if a tweet is positive or negative)

  For example, an AI trained on thousands of customer reviews learns to identify whether a new review is positive, negative, or neutral about a given product or service, which helps the business to improve or plan accordingly.

- *Images Data:* Used when the AI task involves visual recognition or visual interpretation. Data Sources might include medical imaging databases, traffic cameras, surveillance footage, e-commerce product photos, collected according the use case.

  Common tasks using this type of data can be:

  - Object detection or face recognition (e.g., identifying people in photos)
  - Medical diagnosis (e.g., detecting pneumonia or spotting tumors in X-rays)
  - Autonomous driving (e.g., identifying road signs, pedestrians and other vehicles)
  - Industrial automation (e.g., sorting products on a conveyor belt)

  Example: An AI model is trained on thousands of chest X-rays labeled as "Pneumonia" or "Normal." The model learns to identify visual patterns associated with early signs of pneumonia, helping doctors diagnose patients faster and more accurately.

- *Numerical Data (Numbers)* which might come from various sources according to the application domain (Excel files, databases, Internet of Things sensors, financial records, transaction logs) are necessary when working with number-based information like prices, ages, temperature, etc. Common tasks include:

  - Predicting future trends (e.g., stock prices, weather forecasts)
  - Analyzing business performance (e.g., predicting next quarter's revenue)
  - Risk assessment (e.g., calculating the chance someone will default on a loan)
  - Monitoring machine performance in factories (e.g., sensor readings)

  Example: A retail company uses sales records, seasonal trends, and promotional data to train an AI model that predicts next month's demand for each product. This helps optimize inventory and reduce waste.

- *Audio data:* Used when your task involves sound, especially speech. Data sources can include recorded interviews, phone calls, podcasts, language learning apps, voice assistant logs, etc. Common tasks involve:
  - Speech-to-text transcription (e.g., converting a voice note into written text)
  - Voice-controlled assistants (e.g., "Set a timer for 10 minutes")
  - Speaker recognition (e.g., unlocking a phone using your voice)
  - Detecting emotions or stress in a person's voice
  - Language learning (e.g., giving pronunciation feedback)

  Example: An AI model is trained on thousands of hours of English-language audio paired with transcripts. Over time, it learns to accurately transcribe new spoken English sentences into text, which powers services like YouTube captions or voice typing.

- *Multimodal Data (Combining Text, Images, Audio, etc.):* Sometimes, an AI task may involve multiple types of data together. Common tasks in this category include:

  - Video analysis (combines image frames with audio)
  - Medical diagnostics (combining patient reports with medical images)
  - Social media content analysis (analyzing captions with photos or videos)

Example:
A content moderation system analyzes both the text of a post and the image it contains to detect hate speech or inappropriate content more accurately.

Why Careful Data Collection Matters?

Collecting the right data is foundational to building a good machine learning model. The **quality**, **quantity**, and **relevance** of your data determine how well your model can learn patterns and make predictions.

b) Data cleaning

This step involves removing or fixing incorrect, missing, or duplicate data:

- Fix typos or errors
- Fill in missing values (or remove bad entries)
- Remove duplicate records

Example: If an image is blurry or mislabeled as a "dog" when it's actually a "cat," you fix it or remove it.

c) Data Annotation / Labeling

If your AI system needs to learn from examples, the data must be labeled. Thus, you might need to:

- Label emails as "spam" or "not spam"
- Tag faces in photos
- Mark speech samples by the key words being spoken

Example: Mark each image as "cat" or "dog" so that the AI knows what it's learning from.

d) **Data Formatting / Transformation**

This steps involves converting the data into a format the machine and AI can understand. It encompasses:

- Resizing images to the same size
- Converting text into numbers (using appropriate methods)
- Normalizing values (e.g., scale prices or ages between 0 and 1) so that it is easy for the AI to process them consistently.

Example: Convert all images to 128x128 pixels so the AI can process them consistently.

e) **Data Splitting**

When your data are clean and properly formatted (or normalized), the next step is dividing them into three necessary parts for machine learning task. In fact, all the data will not be used for training only, instead, there must be a set for validation and a set for test as follows:

- **Training Set** – to teach the AI (usually ~70–80% of the whole dataset)
- **Validation Set** – to fine-tune and test the AI while training (~10–15%)
- **Test Set** – to evaluate how well the AI performs on new data (~10–15%)

For example: if you have 10,000 dog/cat images, you can use 8,000 images to train, 1,000 to validate, and 1,000 to test the final model.

Brief, data preparation is like setting the stage before a performance. If everything is in place—clean, labeled, and ready—your AI can focus on learning, not on sorting through a mess. Poorly prepared data leads to poor results. Well-prepared data leads to smarter, more reliable AI solution.

(ii) **Training the machine on well prepared data**:

After having prepared your dataset, you proceed with feeding them to the AI Model (i.e. a trainable computer program that has the ability to "learn" how to do a task by looking at lots of examples, and then uses that knowledge to make decisions or predictions on new data in the same domain).

For our case of teaching an AI to classify an image as cat or dog, we will have to feed to the AI program thousands of labeled pictures—some labeled “cat,” and some labeled “dog”, in the proportion we described above (70-80%, 10-15%, 10-15%).

(iii) Pattern Recognition

As these images are fed to the AI, it studies them and starts noticing certain features:

- Cats tend to have **pointy ears**, **shorter snouts**, and **whiskers**.
- Dogs might have **floppy ears**, **longer snouts**, or **different fur patterns**.

However, the AI doesn’t practically know what ears or whiskers are—it works simply by detecting **pixel patterns, shapes, and colors** that often appear in pictures labeled “cat” versus “dog.”

(iv)**Learning from Patterns**:

Over time, the AI system builds then a kind of internal "rulebook" that helps it guess whether a new image is likely a cat or a dog based on these learnt patterns—even if it has never seen that specific image before. We say then that **the leaning task has been done.**

(v) **Prediction**:

After the AI is trained, it uses what it has learned to predict whether it’s a cat or dog when you later show it a brand-new picture (cat or dog). The more training it’s had, the more accurate its predictions tend to be.

This same idea applies to many types of data:

- In **healthcare**, AI finds patterns in X-rays to detect signs of disease according to the received training.
- In **finance**, it finds patterns in market data to forecast stock movements.
- In **language**, it detects patterns in how words are used to generate meaningful sentences or answer questions.

In Simple Terms:

Finding patterns means the AI is learning what features tend to go together, and using that understanding to make predictions or decisions about new, unseen data. This ability to recognize and learn from patterns is what makes AI so powerful—and so useful in fields ranging from medicine to marketing.

This learning process often involves algorithms that adjust and optimize the system's predictions based on the data received. In other words, the learning process usually uses step-by-step methods that help the system improve its guesses by learning from the information it gets.

4.3. Adapting

Once trained, AI systems can now begin to make decisions or predictions in real-time, based on new data. Over time, the system "adapts" by refining its models and improving accuracy, i.e. getting better over time by learning from experience and making more accurate predictions much like how humans adjust their behavior after gaining new experiences.

Example:

When you use a streaming service like Netflix, the system learns from your viewing history to recommend movies or shows you might enjoy. The more you watch, the better Netflix gets much information about your preferences and then becomes better at suggesting content tailored to them (preferences). This is an example of **machine learning** and **adaptation**.

5. Types of Learning

Now that we know what is "machine learning", it's important to get knowledge of different categories of leaning applied in Artificial Intelligence tasks. The primary types of AI learning are ***supervised learning***, ***unsupervised learning***, and ***reinforcement learning***. These are broad categories of machine learning, which is a core technique used in AI systems. However, these main types can have sub-categories such as ***self-supervised*** and ***semi-supervised learning***, as we explore all of them below.

5.1. Supervised Learning

In supervised learning, the AI is trained on "**labelled data**" and each input example has a known correct output. Labelled data means information that comes with the answers already provided. For example, if you're teaching a computer to recognize animals in photos, each picture would come with a label like "dog," "cat," or "elephant." These labels help the computer learn what each animal looks like, just like flashcards help people learn.

In supervised learning, the AI learns from examples where the correct answers are already known—kind of like a student being shown questions along with the right answers so that in the future he/she (student) can answer correctly based on shown questions-answers. The fact of showing the right answer to each question (labelling each data in the dataset) is what is called **"supervision"**.

5.2. Unsupervised Learning

Contrarily to supervised learning, unsupervised learning happens when the AI learns from data that has not been labeled. This approach is used when you want the AI discover hidden patterns, structures, or relationships within the data so that you can then exploit these patterns for a certain purpose.

Example:

With unsupervised learning, AI can group customers into different segments based on their purchasing behaviors, without prior knowledge of those segments. The knowledge of those segments will then help the business to plan more strategically for improved results.

Within the categories of supervised and unsupervised learning, there exists other variants including semi-supervised learning and self-supervised learning.

(i) Semi-Supervised Learning

Semi-supervised learning is a combination of supervised and unsupervised learning, using a small amount of labeled data and a large amount of unlabeled data for training. Labeled data means examples that come with the correct answer (like a photo labeled "cat") and unlabeled data means examples without answers (just the photo, no label).

In semi-supervised learning, the AI is given a small amount of labeled data and a large amount of unlabeled data and it uses the small set of known examples to help make sense of the unknown ones.

Example: Let's say you're building an AI to recognize animals in photos:

- You label 500 photos (dog, cat, bird, etc.).
- But you also have 50,000 more unlabeled photos.
- Instead of labeling all 50,000 by hand (which is time-consuming and expensive), the AI learns patterns from the 500 labeled ones and then uses that knowledge to learn from the 50,000 unlabeled photos too.

(ii) Self-Supervised Learning

Self-supervised learning is a type of machine learning where the AI teaches itself by creating its own labels (its own supervision) from the data — so it doesn't need humans to label anything. Instead of being given the answers (like in supervised learning), the system learns from patterns and relationships within the data itself. Think of it like solving a puzzle where some pieces are missing. The AI learns by trying to **predict the missing pieces** using the parts it can already see.

Examples:

- Imagine training an AI to understand language:

- You give it a sentence with a missing word:
  "The cat sat on the ___."
- The AI has to guess the missing word (like "mat").
- By doing this over millions of sentences, it learns grammar, context, and meaning — **without needing humans to label every sentence**.

This is how big language models like ChatGPT are trained.

- An AI looks at part of a photo and tries to predict what's in the missing part. Over time, it learns to understand the structure of images — again, without needing anyone to label "this is a dog," "this is a car," etc.

5.3. Reinforcement Learning

Reinforcement learning is a type of learning where an AI system learns by trial and error — kind of like training a pet. When it makes a good decision, it gets a reward. When it makes a mistake, it gets nothing or a small penalty. Over time, it figures out which actions lead to better outcomes and learns to make better choices on its own.

In other words, in reinforcement learning (RL), the AI doesn't start with a big set of labeled examples like in supervised learning. Instead, it collects data through experience — by interacting with an environment, trying different actions, and seeing what happens. So rather than learning from a fixed dataset (used in supervised and unsupervised approaches), the AI learns from the results of its own actions — whether those actions lead to rewards (good outcomes) or penalties (bad outcomes). Over time, it builds up a kind of internal "memory" of what works and what doesn't.

Example:

Imagine teaching a robot to play a game. At first, it makes random moves. When it wins or scores points, it gets a reward. Eventually, it learns which moves help it win more often. For instance, think of an AI learning to play a game like Pac-Man:

- It starts by randomly moving around.
- If it eats a dot or avoids a ghost, it gets a reward.
- If it gets caught by a ghost, it gets a penalty.
- Over many rounds, it learns strategies that help it survive and score more points.

Brief, in this kind of learning (reinforcement learning), the AI **collects its own training data** by interacting with the environment, rather than being given a fixed set of instructions or examples.

6. Learning Tools in Artificial Intelligence

When AI systems perform tasks—like recognizing faces, predicting the weather, or recommending what movie to watch—they don't just "know" how to do it. They have to learn, just like people do. The way they learn is through something called **learning algorithms**. These are the tools that power the learning process in AI.

6.1. What Are Learning Algorithms?

A **learning algorithm (**as highlighted earlier**)** is like a recipe that tells the AI how to go through the data, find relationships, use these relationships to solve a given task, and improve its performance by looking at those data. Thus, it helps the system:

- Understand patterns in data
- Make predictions or decisions
- Improve over time with more data and experience

Simply, think of it as the **brain behind the scenes** that teaches the AI how to learn from examples instead of being programmed with specific rules**.**

6.2. Types of Learning Tools

There are **two main categories** of learning tools used to build AI systems: traditional and advanced learning algorithms.

6.2.1. Traditional Learning Algorithms (Classical Machine Learning)

These are the earlier, well-established tools used in AI. They are simple, efficient, and work well for smaller or structured data**.**

- *How they work:*

- They learn from labeled or structured data
- They need humans to choose the right **features** (i.e., tell the algorithm what parts of the data are important)
- They're often easy to understand and explain

Common Traditional Algorithms:

- **Linear Regression -** Linear Regression is one of the most basic and widely used tools in machine learning. It's used for predicting numbers — that is, it helps estimate a continuous value based on the information you give it. If you want to predict something measurable, like a price, temperature, or score — linear regression is often the go-to method.

  Suppose you want to predict the **price of a house** based on its size:

  - You give the AI **data from past sales** (house sizes and their prices).
  - It finds the **best-fit line** that shows how price changes with size.
  - Now, if you give it a new house size (e.g., 2,000 sq ft), it can predict the likely price based on that line.

The bigger the house, the higher the predicted price — if that's the pattern in the data.

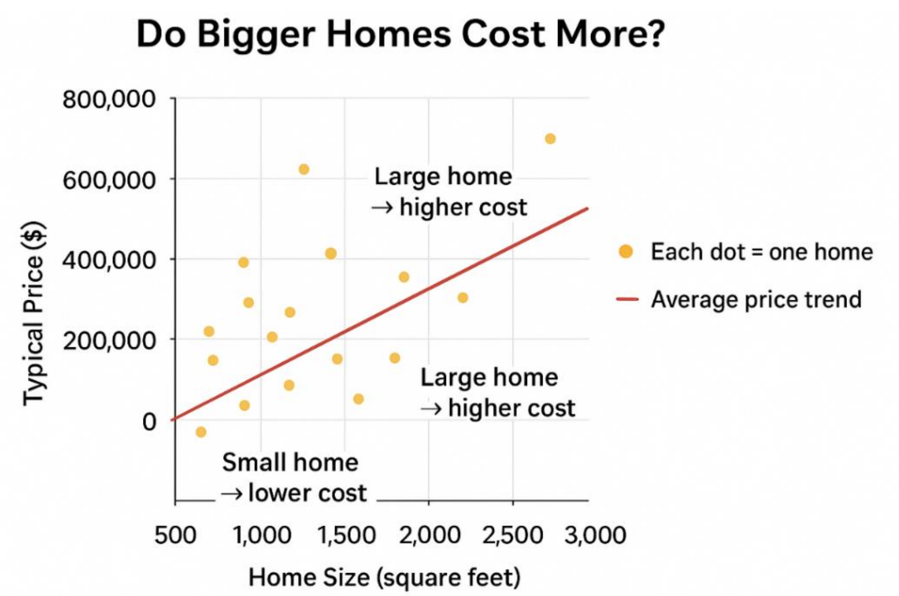

Figure 1: Illustration of how linear regression works in real context.

Examples of Real-World Uses of Linear Regression:

- Predicting house or car prices
- Estimating sales revenue based on advertising spend
- Forecasting temperatures or weather trends
- Predicting student grades based on study time

- **Decision Trees -** for making choices step-by-step in a tree-like models that make decisions by splitting data into smaller subsets based on features.

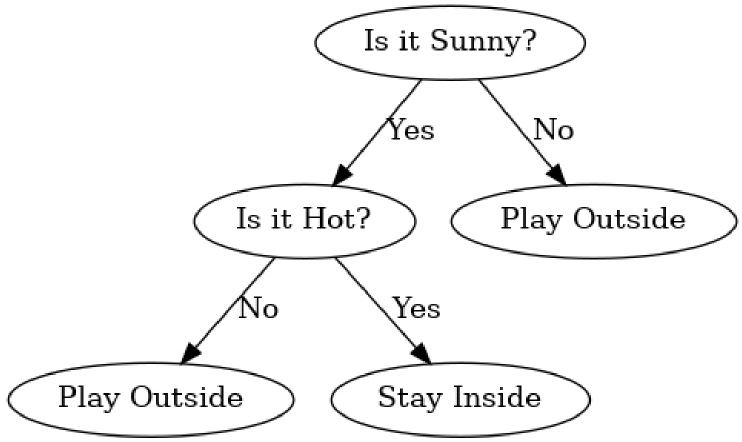

Figure 2: Illustrative decision tree for playing outside of staying inside the house according to weather.

- **Logistic Regression -** Despite the word "regression" in its name, Logistic Regression is actually used for classification tasks — meaning it helps answer yes/no, true/false, or category-based questions (e.g. “Is this email spam or not?”, “Will this customer buy the product?”, “Is this transaction fraudulent?”). It doesn’t give a number as a result like linear regression does. Instead, it predicts the probability that something belongs to a certain category.
- **Support Vector Machines (SVM) -** for separating things into categories. In fact, Support Vector Machines (SVMs) are powerful tools used in classification tasks — that is, when the AI needs to decide which group or category something belongs to. SVMs work by finding the best possible boundary (also called a decision line or hyperplane) that separates different classes in the data. They (SVMs) can be used in tasks such as:

    - Face recognition: Classifying who is in a photo.
    - Text classification: Sorting emails into categories (e.g., spam vs. not spam).
    - Handwriting recognition: Reading numbers or letters from scanned forms.
    - Medical diagnosis: Classifying test results as normal or abnormal.

- **K-Nearest Neighbors (KNN) –** for finding similar items. KNN answers questions like:
    - “What is this new thing most similar to?”

  It does this by:

    - Looking at the "K" closest data points to the item you want to classify (these are its "neighbors").
    - Checking what category most of those neighbors belong to.
    - Assigning the new item to the most common category among its neighbors.

Imagine you move to a new neighborhood, and you want to guess what kind of food a nearby restaurant serves — but there's no sign. So you look at the **5 nearest restaurants** (K = 5), and you notice:

- 3 are Italian
- 1 is Chinese
- 1 is Mexican

Thereby, you'd guess the new one is **probably Italian**, since most nearby ones are. That's how KNN works — it **assumes similar things are close together.**

K-Nearest Neighbors (KNN) is a simple and effective algorithm that thus classifies things by looking at what's nearby. If most of your neighbors are cats, you're probably a cat too!

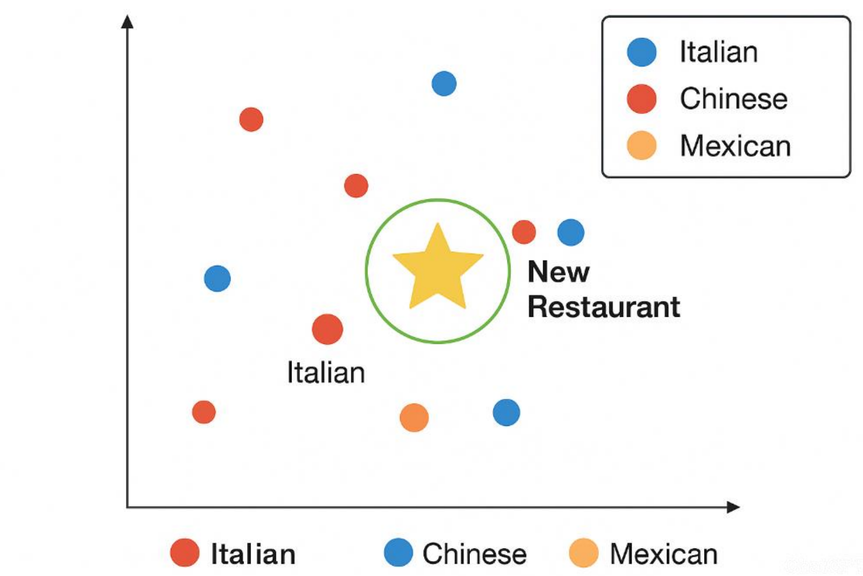

Figure 3: Analogy to K-Nearest Neighbors algorithm

- **Naive Bayes - S**imple but powerful classification algorithm used in machine learning. It helps the computer decide which category something belongs to — like whether an email is spam or not spam, or whether a news article is about sports, politics, or technology.

  It's based on Bayes' Theorem (a concept from probability), and it's called "naive" because it makes a simplifying assumption that all the features (or pieces of data) are independent of each other — which usually isn't true, but surprisingly, it still works well in many situations!

  Imagine you're sorting emails and trying to decide if an email is **spam** or **not spam.**
  Naive Bayes looks at:

- What words appear in the email (like "free", "prize", "click")
- How often those words show up in known spam vs. non-spam emails

Then, it calculates the **probability** of the email being spam based on what words are present. Even though it assumes each word contributes independently, it still gives accurate predictions in many real-life cases.

Think of it like a detective:

- It sees a clue (like the word "free").
- It asks: "How likely is it that an email with this clue is spam?"
- It combines all the clues and makes a **smart guess**, even if it doesn't fully understand how the clues relate to each other.

Real-world uses include:

- **Spam filtering**
  → Is this email spam or not?
- **Text classification**
  → Is this article about politics, health, or sports?
- **Sentiment analysis**
  → Does this review sound positive or negative?
- **Language detection**
  → Is this text in English, French, or Spanish?

These tools discussed above are powerful and fast for many everyday AI tasks, especially when you have smaller, clean datasets.

6.2.2. Advanced Learning Techniques: Deep Learning

Deep learning is a more advanced learning tool that has revolutionized modern AI. Deep learning is a type of AI learning that helps computers learn from data in a way that's similar (or almost similar) to how humans learn — by recognizing patterns.

It uses special computer systems called **neural networks (NN)**, which are inspired by the way human brain works—through neurons. These networks are called "**deep**" because they have **many layers** that help the system learn more complex things. **Deep learning** is therefore a powerful kind of AI tool that can automatically learn and improve from large amounts of data, especially when tasks are too complex for simple rules.

6.2.2.1. What is a "neural network"?

Imagine a neural network as a network of **digital brain cells** that work together to solve problems.

- Each "cell" (called a **neuron**) receives some information,

- Makes a simple decision,
- Then passes the result to the next layer of neurons,
- Until the final output is produced — like a prediction, a classification, or a decision.

A **neural network** is made up of **layers of nodes** (neurons):

1. **Input Layer**
   Where the data goes in (e.g., an image, numbers, text).
2. **Hidden Layers**
   These are in the middle, where all the **learning and pattern-finding happens**. There can be many of them in complex networks.
3. **Output Layer**
   Where the final result comes out (e.g., "This is a dog" or "Price = $350,000").

In a neural network, a layer is like a **row of tiny decision-makers** (called neurons or nodes) that all do their job at the same stage in the process. Each neuron in a layer looks at the data it receives, does a small calculation, and passes the result on to the next layer.

6.2.2.2. Types of Layers

1. **Input Layer**
   - The **starting point** of the network.
   - It receives raw data (like numbers, images, or words).
2. **Hidden Layers**
   - These are the **"thinking" layers** of the network.
   - They process the data, detect patterns, and transform the information step by step.
   - A network can have **many hidden layers** — and this is what makes it **deep learning**.
3. **Output Layer**
   - This is where the network gives its **final result** or **prediction**, like:
     - “Yes, it’s a cat.”
     - “The price is $42.50.”
     - “This sentiment is positive.”

Brief, a neural network is a system of interconnected digital "neurons" that learn to solve complex problems by recognizing patterns in data — similar to how a simplified brain would work.

The diagram below shows the parts and the connections between the parts of a neural network. The figure represents a simple neural network, because in real life, neural networks often have billions of nodes or neurons per layer and hundreds of layers.

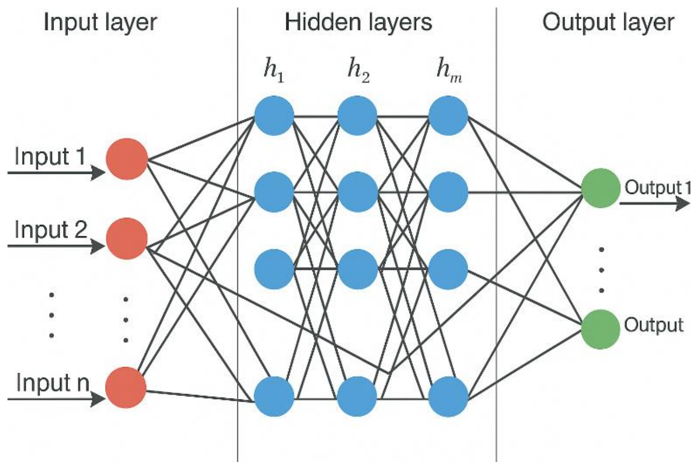

Figure 4: Simplified representation of a neural network

In that way, think of deep learning like layers of understanding such that:

- The **first layer** is able to notice simple features (like edges in a photo),
- The **next layers** combine those features into shapes (like eyes or wheels), and
- The **final layers** are capable of recognizing full objects (like a face or a car).

Just like how a child first learns to notice shapes before recognizing objects!

- *How deep learning works:*

- Uses neural networks with many layers (inspired by how the human brain works)
- Learns directly from raw data like images, audio, or text
- Can find complex patterns that traditional methods might miss

- *What makes it special:*

- Automatically learns features from data (no manual feature selection)
- Performs extremely well with large datasets
- Powers many cutting-edge technologies

- *Real-World Examples:*

- Face recognition **in your phone's camera**
- Voice assistants **like Siri or Alexa**
- Self-driving cars **detecting pedestrians and traffic signs**
- Translation apps **converting one language to another**
- Medical AI **reading X-rays or MRIs**

6.2.2.3. Common Deep Learning Architectures

(i) Convolutional Neural Networks (CNNs) - for image recognition

A CNN is a learning system that is trained to recognize things in images by breaking them down into smaller and smaller patterns until it can identify the full object.

- Imagine you're looking at a picture of a cat.
- Instead of looking at the whole picture at once, your brain first notices **small details** — like edges, corners, shapes, or colors.
- Then, it combines those details to recognize **bigger features** — like eyes, ears, or whiskers.
- Finally, it puts everything together to say: "Yes, that's a cat!"

A CNN works in a similar step-by-step way:

1. **Convolution layers**: These act like filters that scan the image to find simple patterns (lines, textures, shapes).
2. **Pooling layers**: These shrink the information, keeping the most important details and ignoring extra noise.
3. **Fully connected layers**: These combine all the learned features to make the final decision (e.g., "cat" or "dog").

In everyday analogy, learning with CNNs is like how a child learns to recognize animals:

1. First, they just see shapes.
2. Then, they notice features — ears, eyes, tails.
3. Finally, they combine those clues to say, "That's a dog!"

CNNs do exactly that — layer by layer — until they *see* what's really in an image, as illustrated on the figure

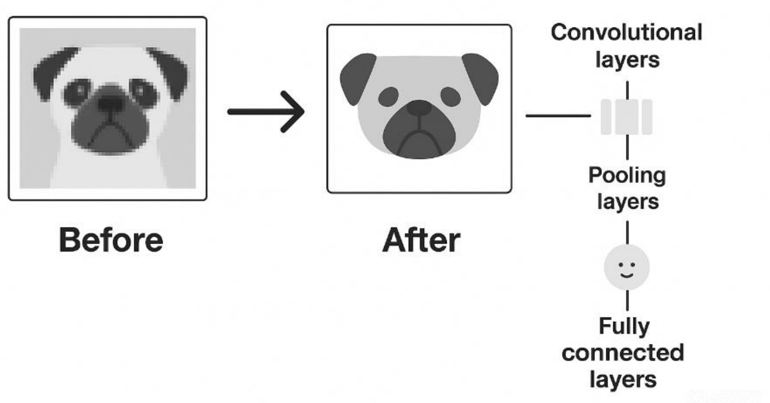

Figure 5: CNN before/after comparison showing how a convolutional neural network extracts edges/features)

(ii) **Recurrent Neural Networks (RNNs)** – for language and speech, etc.

An RNN is a learning architecture that's good at dealing with sequences — things that happen in order — like text, speech, music, or time series data. It "remembers the past" to help make sense of the present.

Think of an RNN like how people remember things in a story.

- When you read a book, you don't forget the earlier chapters — you use that memory to understand what's happening now.
- For example, if a character is introduced as a "wizard" early on, you'll remember that later when they cast a spell.

An RNN does something similar:

1. **Step by step reading** – It takes information one piece at a time (like one word in a sentence or one frame in a video).
2. **Memory** – It remembers what it has already seen and uses that memory when processing the next piece.
3. **Prediction/Output** – Based on both the new input and its memory, it makes a decision (like predicting the next word in a sentence or understanding speech).

(iii) **Transformers** – for text understanding and content generation (used in tools like ChatGPT)

A Transformer is like a super-smart reader that can take in a whole passage at once, figure out the important connections between words (or data points), and use that to understand or generate meaningful output.

Imagine you're reading a whole paragraph.

- Unlike you, an RNN would read it **word by word**, remembering the past as it goes.
- A Transformer, on the other hand, looks at the **whole paragraph at once** and figures out which words are most important for understanding the meaning.

Here's how it works in plain language:

1. **Attention** – Transformers use something called *attention*. Think of it like a spotlight: instead of treating every word equally, the model shines its light on the words that matter most for the current task.
    - Example: In the sentence *"The cat sat on the mat because it was tired"*, the model knows that *"it"* refers to *"the cat"* (not the mat).

2. **Parallel processing** – Unlike RNNs, which read step by step, Transformers can look at all the words (or parts of data) **at the same time**, making them much faster to train.
3. **Scalability** – This design makes them super powerful for handling huge amounts of text, images, or even audio. That's why Transformers are behind tools like **ChatGPT, translation systems, and modern AI assistants**.

6.3. AI model(s)

Now that we have knowledge of learning process and various tools used, it is time to define what is known as "AI model".

In machine learning (and hence AI), a **model** is what you get after a computer has finished **learning from data**. It's basically the computer's **knowledge** or **understanding** of the patterns it found in that data, such that it can apply that knowledge on new unforeseen data to make decision. This is what is known as "AI Model".

In short, a model is the result of training — it's what the computer has learned and will use to make predictions or decisions in the future, as illustrated by figure 2.

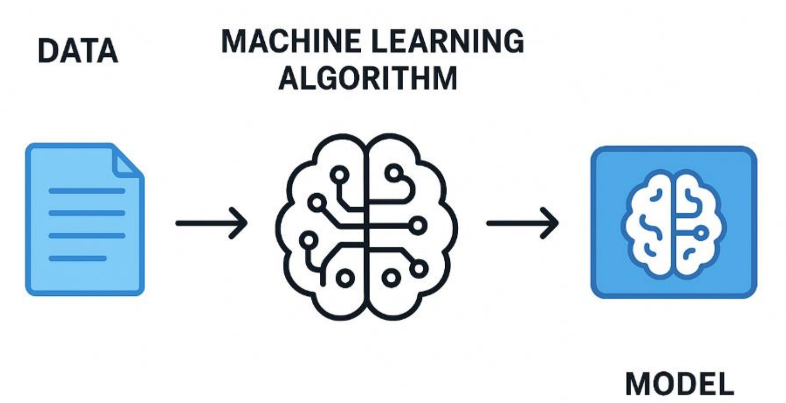

Figure 6: Visual illustration of how a model is a result of learning process

In AI, learning tools are what make machines smart**.** Whether it's a simple model that predicts the weather or a deep learning system that understands human speech, it's all about choosing the right learning algorithm for the task.

- Traditional ML models are like a smart calculator — efficient and fast.
- Deep learning models are more like a student who can read, listen, and eventually understand the world with enough examples**.**

7. Two Main Categories of AI by Their Function: Predictive Vs. Generative

Now that we've covered the fundamentals of AI, it's important to recognize that AI is not a single, one-size-fits-all technology. Rather, it's a broad and evolving field that serves a wide variety of purposes. As AI has advanced, its capabilities have become more specialized — and while there are many ways to classify different types of AI, one of the most useful distinctions is based on what the AI is designed to do.

We can broadly divide AI systems into **two main functional categories**:

- **Predictive AI**
- **Generative AI**

While these categories are not completely separate — and in fact, often overlap — understanding the difference between them can help us better appreciate how AI is applied in the real world.

7.1. Predictive AI: Learning from the Past to Predict the Future

Predictive AI refers to systems that analyze historical data to make informed guesses about future outcomes. These models look for patterns and relationships in existing data, and then use that knowledge to:

(i) Predict what might happen next
(ii) Classify new data into categories
(iii) Identify risks or opportunities before they occur

➢ **How It Works:**

Predictive models are trained on labeled data — for example, past sales figures, customer behaviors, or temperature records. They "learn" how certain factors influence results and then apply this understanding to make predictions on new, unseen data.

➢ **Real-World Examples:**

- **E-commerce**: Recommending products based on your browsing or purchase history (e.g., "Customers who bought this also bought...")
- **Finance**: Predicting stock price movements or identifying fraudulent transactions
- **Healthcare**: Forecasting disease risk based on medical history
- **Language**: Classifying spam emails or predicting the next word in a sentence (autocomplete)

Today, predictive AI is everywhere — from your **email inbox** filtering spam, to your **bank** assessing whether you're a credit risk, to your **streaming platform** suggesting what to watch next. It helps organizations make data-driven decisions and prepare for what's ahead.

7.2. Generative AI: Creating Something New from What It Has Learned

Generative AI goes a step further. Instead of just analyzing or classifying existing data, it learns the underlying patterns of that data so it can generate entirely new content — such as text, images, audio, or even video — that looks and sounds like it was made by a human.

7.2.1. How It Works:

Generative models are trained on massive datasets — for example, books, articles, images, videos, or code. Over time, they learn how language or images are structured. Once trained, they can create new outputs that resemble the original training data, but are not direct copies.

Real-World Examples:

- **Text generation**: ChatGPT can generate emails, essays, or poems based on your prompt
- **Images generation**: Tools like DALL·E create artwork from a description (e.g., "a cat playing guitar on a beach at sunset")
- **Music composing**: AI can compose songs in the style of classical or pop artists according to user prompt.
- **Programming**: AI can generate code snippets from plain-language instructions.

Generative AI is what has recently captured the public's imagination. Tools like **ChatGPT, Gemini, Claude, and Midjourney** have made it possible for anyone — even those with no technical background — to use AI to create useful and even artistic content. It's changing how we write, design, learn, and express ideas.

7.2.2. Key Differences Between Predictive and Generative AI at a Glance

To better understand where generative AI fits within the broader field of artificial intelligence, it is worth comparing it with predictive AI. While both rely on data and machine learning, their goals and applications differ significantly. Table 1 below highlights the key distinctions, showing where each is most commonly used.

Table 1: Key differences between predictive and generative AI

Aspect	Predictive AI	Generative AI
Main Goal	Forecast or classify based on data	Create new content from learned patterns
Input Data	Historical, labeled data	Large datasets (text, images, etc.)
Output	A prediction, label, or score	Original text, images, audio, video, or code
Examples	Spam filters, credit scoring, forecasts	Text generation, AI art, music composition

Used In	Business analytics, healthcare, finance, etc.	Marketing, education, design, entertainment, project management, etc.

Understanding the difference between **Predictive AI** and **Generative AI** gives you a better grasp of how AI is used today — and what its future might look like. While **predictive AI** helps us make smarter decisions based on data, **generative AI** opens up entirely new possibilities for creativity, communication, and problem-solving.

As you continue exploring AI, you'll find that many applications actually combine both types. For example, a smart assistant may **predict what you want** and then **generate a custom response** — seamlessly blending both capabilities.

7.2.3. How Generative Models Work: Understanding Large Language Models (LLMs)

Generative models are a special type of AI system designed to **create new content** based on what they've learned from existing data. After being trained on a large number of examples — whether that's text, images, audio, or video — these models can generate **completely new outputs** that follow similar patterns. That's why they're called **"generative"**: they don't just analyze or categorize information — they **produce** it.

A major category within generative AI is **Large Language Models**, often referred to by their initials, **LLMs**. These models are trained on vast amounts of text data and are capable of generating new, natural-sounding language. You've likely encountered LLMs in tools like **ChatGPT**, **Gemini**, or **Claude**, which can write emails, explain concepts, summarize articles, answer questions, and much more — all in human-like language.

But generative AI isn't limited to just text. Similar models can be trained to generate:

- **Images** (e.g., AI art from text prompts)
- **Audio** (e.g., AI-generated music or voice cloning)
- **Video** (e.g., short clips based on descriptions)
- Or even **code**, **interactive experiences**, and **multimodal content** that blends multiple formats together.

In many modern applications, both the **prompts** (what the user inputs) and the **outputs** (what the AI produces) can involve combinations of **text, code, images, audio, and video** — making generative AI a powerful and flexible tool across many industries.

7.2.4. How Does It Work Under the Hood?

At its core, **generative AI works by identifying patterns** in data. Once trained on a large dataset, the model learns the **underlying structure and relationships** within that data. It then uses that knowledge to create outputs that are statistically likely to make sense — but are not just copies of what it saw before.

In the case of language models, this means learning to **predict the next word in a sequence**. By training on millions or even billions of words, the model develops an understanding of grammar, meaning, context, and style. This is similar to how humans learn language — by reading and hearing more examples, we develop a sense of how language works.

7.2.5. Generative AI and Large Language Models (LLMs)?

A **Large Language Model (LLM)** is a type of generative AI model built using machine learning techniques and trained on massive amounts of text data. These models process data at such a large scale that they can learn highly complex statistical relationships — such as how ideas are structured in a sentence, how tone changes based on context, and how questions lead to answers.

Because of their size and training data, LLMs are able to:

- Write long and coherent passages of text
- Answer questions with relevant information
- Translate between languages
- Summarize large documents
- Simulate conversation, and much more

The performance of these models generally improves with **more data and more training**. The more language they are exposed to, the more nuanced and useful their outputs become.

In Summary:

- **Generative AI** creates new content based on patterns learned from data.
- **LLMs** are a type of generative model focused on generating and understanding human language.
- These models are trained on vast amounts of text and can produce meaningful, fluent, and context-aware language.
- Generative AI is not limited to text — it also powers tools that generate images, audio, video, and code.
- The power of these systems comes from the data they are trained on and the **statistical patterns** they learn.

7.2.6. Generative AI Tools and Work Productivity

Now that we know a little more about artificial intelligence and the differences between generative and predictive AI, let's see how you can use it to increase your productivity. Recently, generative AI has become much more accessible and has attracted a global attention due to its capabilities to help users improve their productivity at work.

In a general working environment, AI can improve your productivity in four different ways that we explore below. It can help you to:

(i) Summarize Content

With a proper prompting, generative AI offers summarization tools that extract the most important information from lengthy and complex documents. These tools analyze the content on your behalf, saving time by delivering a summary that highlights the key points you need to understand. Tools such as ChatGPT, Gemini and Claude are frequently used by users for that purpose.

(ii) Learn and comprehend

Generative AI tools also offer comprehension capabilities that help you learn new information and master new skills more efficiently. They can quickly find, analyze, and explain information much faster than you could on your own—making them incredibly useful for productivity. You simply need to provide a clear prompt describing the topic and the details you want to learn.

Example:
Imagine you're trying to learn **how neural networks work** for an upcoming project. Instead of reading dozens of technical papers (though it is highly advisable), you could type a prompt like "Explain how neural networks learn, using simple language and a real-world example." Within seconds, the AI would summarize key concepts—like layers, weights, and training—using an easy-to-understand analogy and this would allow you to grasp complex topics quickly and focus on applying your new knowledge right away.

(iii) Create Content

Generative AI offers creative assistance tools designed to help you overcome one of the toughest moments in the creative process — **facing a blank page**. These tools can generate fresh ideas, outlines, or drafts to get your creativity flowing. You can then refine the suggestions, discard the ones that don't fit your vision, and expand on those that inspire you.

Example:

Imagine you're designing a marketing campaign for a new eco-friendly product but can't come up with a catchy slogan. You might prompt the AI with "Suggest five creative slogan ideas for an eco-friendly water bottle that appeals to young professionals." In seconds, the AI could generate options and you can then select the one that resonates most with your brand message, or combine ideas to craft your perfect slogan.

(iv) Automate Tasks

Automation involves training AI systems to perform repetitive or manual tasks automatically. While setting up these tools may take a bit more time and technical skills than other AI applications, the benefits are substantial. Once configured, they can handle routine, time-consuming processes—such as scheduling, data entry, or report generation—allowing you to focus on higher-value work that contributes more meaningfully to your career or organization's growth.

Example:
An AI automation tool can automatically sort and respond to customer support emails. It classifies inquiries, sends standard replies to common questions, and forwards complex issues to a human agent. This not only saves hours of manual work but also ensures faster, more consistent service.

7.2.7. Limitations of Generative AI

Although generative AI can be incredibly powerful and useful across many types of tasks, it's essential for users to understand its limitations and take appropriate precautions to ensure it remains an effective tool. Today, many people—especially those less familiar with artificial intelligence—tend to equate *AI* entirely with generative models. These systems can seem almost magical in their ability to create text, images, music, or even videos. However, beneath this impressive surface lie significant constraints that determine how they function and what we should (and should not) expect from them.

- **Generative AI doesn't really "understand"**: Generative AI doesn't think or understand meaning like humans do. It looks at patterns in data — for example, how words or images usually appear together — and uses that to **predict what comes next**. So when it writes an essay or generates an image, it's not "thinking creatively," it's combining patterns it has learned.
  Example: When asked to "write a poem about friendship," it doesn't *feel* friendship — it just assembles phrases that statistically fit the theme.

- **It depends on training data:** Everything the AI knows comes from the data it was trained on — books, websites, articles, conversations, etc. Therefore, if that data is biased, limited, or outdated, the model's responses will reflect that.
  Example: If most of the data used to train the AI came from English-language sources, it may perform poorly in other languages or other cultural contexts.

- **Hallucinations:** Sometimes, AI produces confident but false information — something that sounds right but isn't true. This is known as "AI hallucinations" and they happen because the model predicts what should come next based on language patterns, not actual facts. This situation of AI generating responses that appear correct but that contain errors or are out of context are due to a lack of an understanding of the real world, and it is always crucial to review and verify the information obtained from a generative AI tool.

Example: Generative AI might invent a fake scientific study or make up a historical date that doesn't exist. Make sure to always check the veracity of generated content.

- **It struggles with complex reasoning:** While AI can handle simple calculations or logic, it still struggles with multi-step reasoning, abstract thought, or long-term planning.
  Example: Generative AI might write a good summary of a story but fail to understand the deeper message or theme.

- **Limited memory and context:** AI models can only consider a certain amount of information at once. When conversations or documents get too long, they may "forget" earlier details, leading to inconsistent answers.
  Example: In a long chat, the AI might contradict something it said earlier.

- **Limited control over results:** Even experts can't always predict or control what an AI will generate. Sometimes it produces amazing results; other times it outputs nonsense or off-topic content.
  Example: You might ask for a professional email and get something too casual — or vice versa.

Why Understanding These Limitations Matter?

Understanding these AI limitations helps people use AI **responsibly and effectively**. The limitations remind us that:

- AI is a tool, not a replacement for human judgment.
- It can enhance creativity and productivity, but it always needs guidance.
- Transparency and ethics are essential to build public trust in AI.

Therefore, the human supervision is still and always needed when using AI tools. Generative AI can assist, but it can't replace human judgment. Its results often need fact-checking, editing, and ethical review before being used.

Example: Companies using AI for customer support must still have people review the AI's responses to avoid errors or offensive language.

In short: Generative AI can *create*, but it doesn't *understand*. Used wisely, it's a powerful partner — not a replacement for human intelligence. We must always ensure a human verification of generated content to avoid probable disasters.

8. Challenges and Ethical Considerations in AI

AI technology brings solutions to many challenging problems of human daily activity, but also brings complex challenges and ethical issues that are worth considering:

- **Bias in Data and Models**: AI models reflect the data they are trained on. If the data contains biases, such as underrepresentation of certain groups, the AI's predictions can be unfair or discriminatory.

  Example: Facial recognition systems that misidentify people of certain ethnicities due to biased training datasets.

- **The Black Box Problem**: Many AI models, especially deep learning networks, are so complex that it's hard even for experts to understand exactly how decisions are made. This can hinder trust and accountability. To this issue, **eXplainable AI (XAI)** was coined as a discipline aimed at providing AI systems with clear and understandable explanations for their actions and decisions.
- **Privacy and Data Security**: AI relies on large datasets, often containing sensitive personal information. Ensuring this data is collected and used responsibly is critical.
- **Job Displacement**: AI-driven automation can replace some human jobs, raising questions about economic impact and the need for reskilling.

9. The Future of AI: Opportunities and Responsibilities

AI continues to evolve rapidly, promising advancements in healthcare, agriculture, education, commerce, finance, environmental protection, and more. AI can help personalize education and make learning more effective, improve agriculture productivity by precision farming, and aid climate modeling and resource management. However, it requires responsible development, transparency, fairness, and ongoing human oversight to responsibly and fairly benefit from its promises.

In conclusion, Artificial Intelligence is a transformative technology that allows machines to learn from data and perform tasks that require human-like intelligence. From voice assistants and recommendation engines to medical diagnostics and autonomous vehicles, AI impacts nearly every aspect of our lives. However, understanding AI's core concepts, capabilities, and challenges empowers us to make informed decisions about how to harness this powerful tool for the benefit of society.

Declaration on Generative AI

The authors, after using ChatGPT whenever necessary for content generation, grammar and spelling check, a thorough review was then done and the content corrected and/or edited as needed. The authors take full responsibility for the publication's content.

Proposed Further Reading & Resources

Exploring Artificial Intelligence further can help deepen your understanding of how it works, its real-world impact, and the ethical questions surrounding it. The following books, online courses, and digital resources have been selected specifically for non-expert readers who wish to continue learning in an engaging and practical way.

A. Books for General Understanding

1. Artificial Intelligence: A Guide for Thinking Humans by Melanie Mitchell (2019) – A clear and thoughtful introduction to what AI can and cannot do. https://www.amazon.com/Artificial-Intelligence/dp/0241404835
2. Life 3.0: Being Human in the Age of Artificial Intelligence by Max Tegmark (2017) – A visionary look at how AI might shape the future of work and ethics. https://www.cag.edu.tr/uploads/site/lecturer-files/max-tegmark-life-30-being-human-in-the-age-of-artificial-intelligence-alfred-a-knopf-2017-aTvn.pdf
3. The Ethical Algorithm by Michael Kearns and Aaron Roth (2019) – Explains fairness, transparency, and privacy in AI systems. https://global.oup.com/academic/product/the-ethical-algorithm-9780190948207?cc=es&lang=en&
4. AI Superpowers: China, Silicon Valley, and the New World Order by Kai-Fu Lee (2018) – A global look at AI development and competition. https://g-city.sass.org.cn/_upload/article/files/1d/3a/2ed5278f4f51bc4a72876e6b1765/a557dd11-b9b1-46af-9f7c-d6225fe50a70.pdf
5. Human Compatible: Artificial Intelligence and the Problem of Control by Stuart Russell (2019) How to ensure AI stays aligned with human values. https://www.amazon.es/Human-Compatible-Artificial-Intelligence-Problem/dp/0525558616

B. Free and Accessible Online Courses

1. Elements of AI – A free, beginner-friendly course by the University of Helsinki. https://www.elementsofai.com/
2. AI for Everyone – Andrew Ng (Coursera) – A non-technical course about AI's business and social impacts. https://www.coursera.org/learn/ai-for-everyone
3. Google's Machine Learning Crash Course – Great for beginners, includes visuals and exercises. https://developers.google.com/machine-learning/crash-course

4. Microsoft Learn – Introduction to AI – Covers AI basics like vision, speech, and NLP. https://learn.microsoft.com/en-us/training/modules/get-started-ai-fundamentals/

C. Interactive Learning Platforms

1. Teachable Machine – Train simple AI models using images, sounds, or poses directly in your browser. https://teachablemachine.withgoogle.com/
2. MIT Scratch + AI Extensions – Experiment with AI through creative, visual programming. https://resources.scratch.mit.edu/www/lessons/en/ScratchLearningResource_ScratchAILesson.pdf
3. Kaggle Learn – Short, hands-on exercises for learning AI and data science. https://www.kaggle.com/learn

D. Ethical and Societal Perspectives

1. AI Now Institute – Researches social implications of AI such as bias and surveillance. https://ainowinstitute.org/
2. Partnership on AI – Promotes responsible, ethical AI development. https://partnershiponai.org/
3. UNESCO's Recommendation on the Ethics of Artificial Intelligence – Framework for global AI ethics policies. https://www.unesco.org/en/artificial-intelligence/recommendation-ethics