# PART-2
# ARTIFICIAL INTELLIGENCE

# Chapter 1. AI Domains and Applications

Before moving ahead, let us briefly revisit some important ideas about Artificial Intelligence and data.
Intelligence is the ability to learn, think, understand, and solve problems. Artificial Intelligence (AI) refers to machines designed to perform tasks that normally need human intelligence, such as recognising patterns, making decisions, or solving problems. However, not all smart machines use AI. Some devices simply follow fixed instructions, a process called **automation**.

AI systems learn from data that can be numbers, words, images, videos or sounds. Data helps to understand patterns and draw conclusions. Recognising patterns helps in making better decisions in daily life.

Technology should be used responsibly. This means keeping your passwords safe, protecting your personal information, and getting your parents' permission before sharing anything online. To build a good digital footprint, always follow ethical guidelines, stay safe, and share positive content.

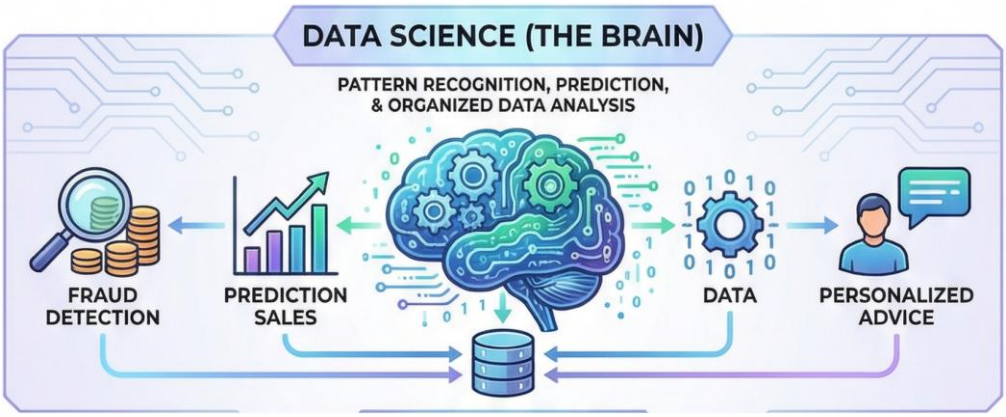
### Meaning of AI Domains

The term **AI domains** defines the study and use of various fields within AI, which are meant to replicate human thinking capabilities, such as learning, seeing, understanding language and problem-solving. AI is not just a single technology, but it is a branch of different domains. These domains are further classified based on the type of data they work with and the task they complete, allowing machines to perform tasks that usually require human intelligence.

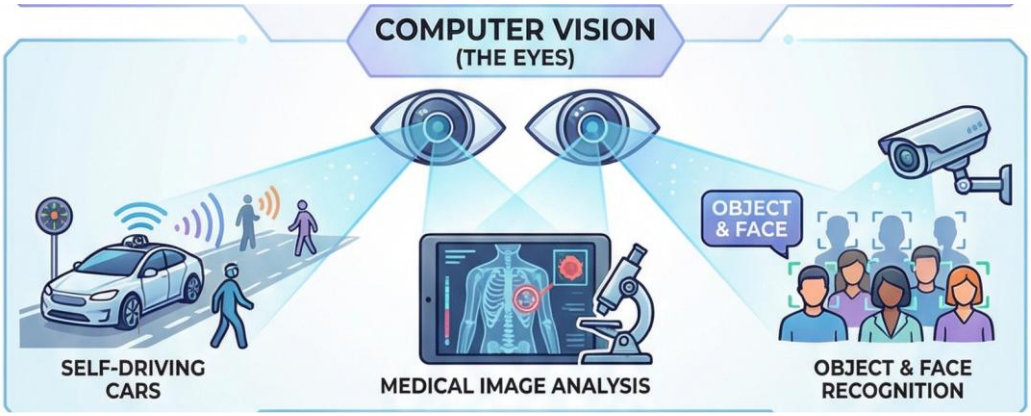
**Some important domains of AI are classified as:**

1. Data Science
2. Computer Vision
3. Natural Language Processing

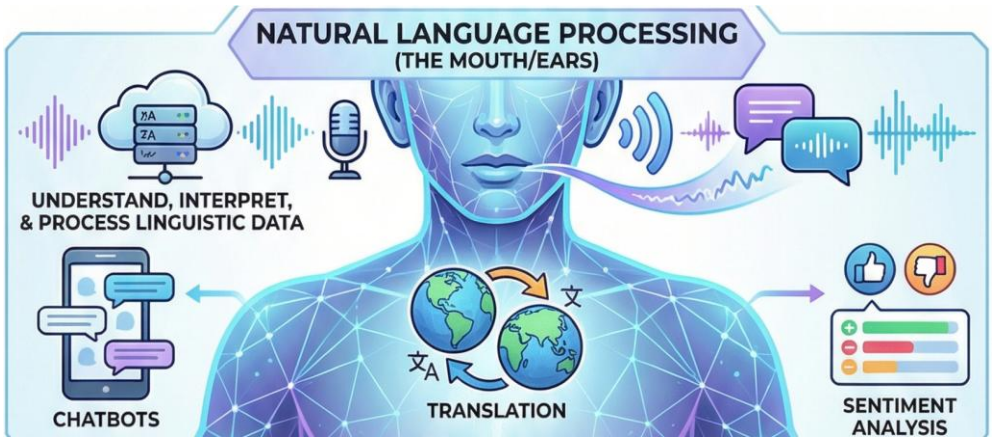
↳ **Data Science (The Brain):** This technology looks for patterns and predicts outcomes by using organised data like numbers or text. It is useful for spotting fraud, predicting sales, and giving personalized advice.



↳ **Computer Vision (The Eyes):** This technology uses images and videos to find faces, recognize objects, or read text. It is particularly important for self-driving cars and medical image analysis.




↳ **Natural Language Processing (The Mouth/Ears):** This technology can understand, interpret, and create human language by processing linguistic data like speech or text. It is used in chatbots, translation, and sentiment analysis.




## Predictive techniques in AI

Prediction in AI is not magic, but there is the logic and mathematics behind every prediction. How?

- Using data collected in the past to predict what can happen in the future.
- Training a model to identify patterns and apply those patterns and trends to new and unseen data.

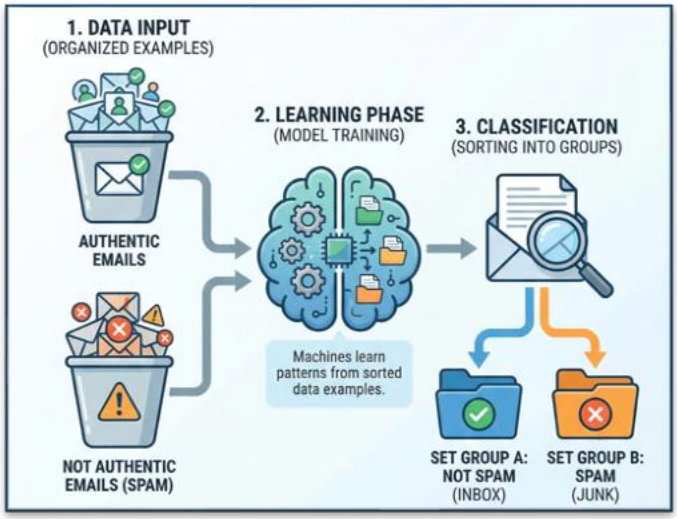
For example, AI can be used to predict crop yield. By analysing data like weather conditions, soil quality, and past harvests, AI can estimate how much crop will be produced in a season.

## Introduction to Classification

**Classification** is a method used in AI systems where the machines sort data into particular categories. In this method, the system is trained using examples in which the correct answers are already known. This training process is called **learning**. By studying these examples, the model learns patterns and then uses them to classify new data correctly.

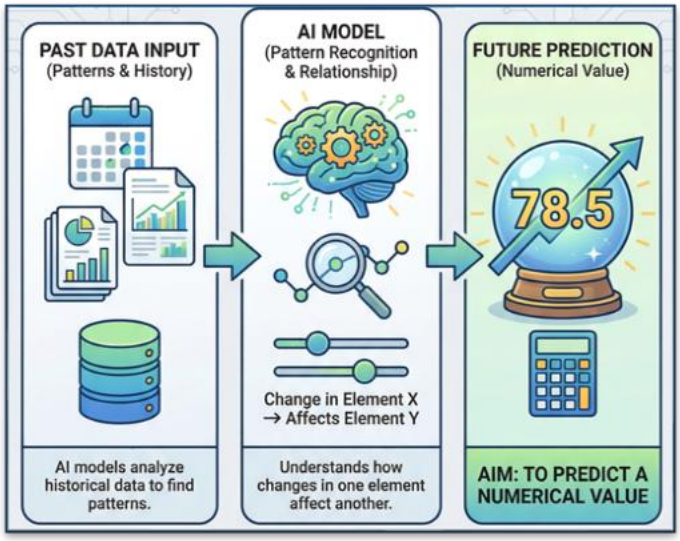
In classification, data is placed into set groups instead of being given a general value.

For example, an email service scans messages and determines if a message is spam or not spam based on previous examples.



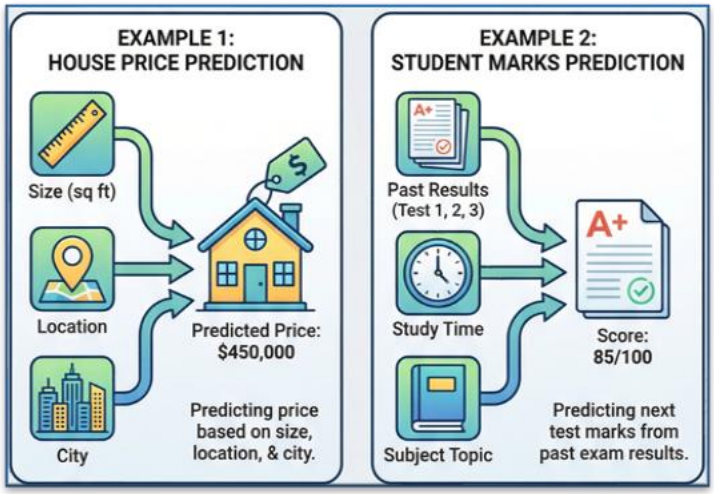

## Introduction to Regression

Regression is an approach in artificial intelligence where AI models predict **numerical** values based on past data. In regression, AI models look at patterns in the data and learn how changing one piece of information can change the number they predict.



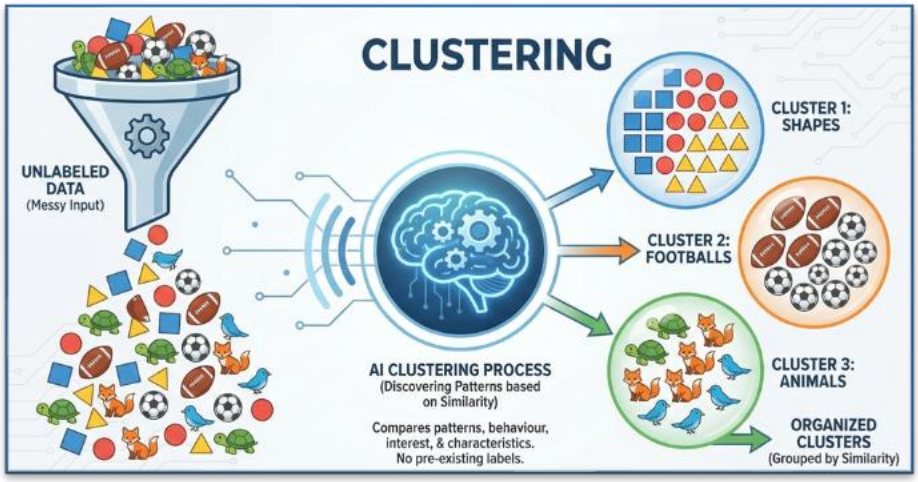

In regression, the aim is to predict a numerical value.

For example, predicting the price of a house based on its size, location, and city you are looking for. Similarly, by analysing a student's past exam results, a computer program can predict how many marks the student might achieve on the next test.



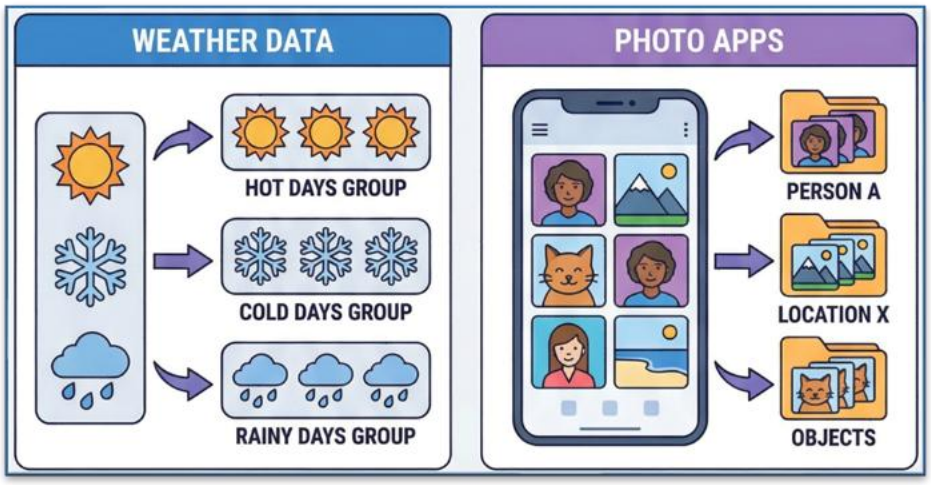

## Introduction to Clustering

Clustering is a method in AI that **groups similar** data together. Unlike classification, clustering does not use labelled data, as no labels are given in advance. Clustering compares patterns, behaviour, interests, and characteristics of data sets and discovers its own pattern. It forms separate groups based on the similarities in the data.




<u>In clustering, similar elements of data are grouped together without assigning predefined labels.</u>

For example, weather data might be grouped by similar days such as hot, cold, or rainy, and photo apps can automatically organise pictures of the same person or location.



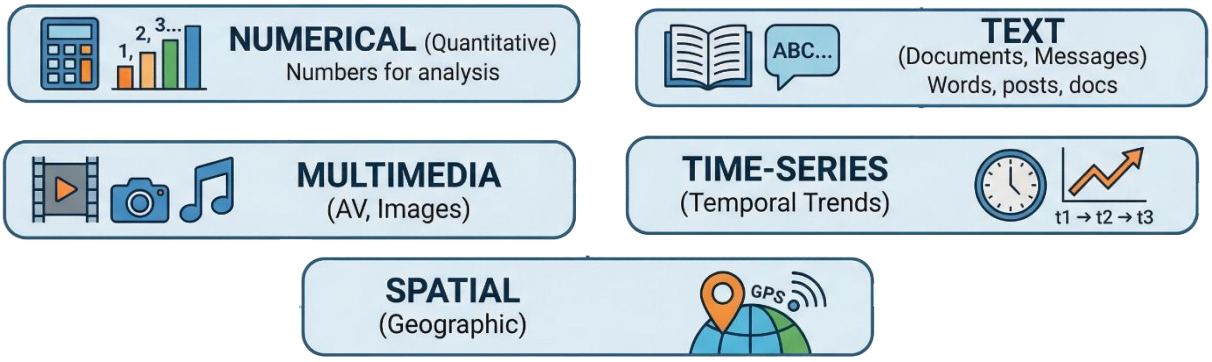
## What are Datasets?

When talking about AI and Machine Learning, datasets are **groups** of **data** that are used to **teach** and **evaluate** machine learning models. The goal of these datasets is to give the models the information they need to learn patterns and make predictions; hence, **datasets** are important to the progress of AI and machine learning. They help researchers and experts solve many real-world problems such as speech and picture recognition, as well as predictive analytics and natural language processing.

Each piece of information in a **dataset** is called a data point or record with several **attributes** or **features**. Typically, datasets are arranged in formats such as tables, charts and spreadsheets, which make the information easy to access and help in making the correct data-driven decisions.

## Types of Datasets

Datasets can be categorised in several ways. Here are some of the most important types of datasets.




- **Numerical Datasets:** Hold numbers used for quantitative analysis.
- **Text Datasets:** Hold text, posts, messages, documents.
- **Multimedia Datasets:** Hold images, videos and audio files.
- **Time-Series Datasets:** Hold data collected over a period of time to analyse trends and patterns.
- **Spatial Datasets:** Hold geographic information such as GPS data.

## Based on Data Structure


- **Structured datasets:** Datasets that are organised for easy search and evaluation of data.
- **Unstructured datasets:** Datasets that are not organised or have no clear structure. They may include several types of data.
- **Hybrid datasets:** Datasets that include both structured and unstructured data.

## Dividing a data set

A dataset is usually divided into **three parts**:

1. **Training dataset:** This part of the data is used to teach the AI models, where these models learn from the patterns and trends.
2. **Validation dataset:** This part of the data is used to evaluate and improve the learning while the AI models are still training.
3. **Test dataset:** In this part the final check is done to see if the AI models have learnt correctly.

*For Example:*

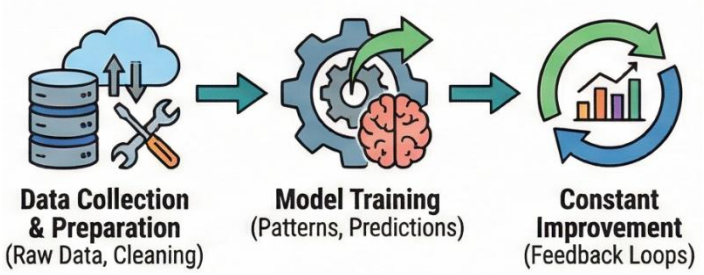
- You have learnt a new chapter from your math notes (**training**)
- You try and practise some of the sums (**validation**)
- You take the final test. (**test dataset**)

*For Example:*

- Assume you have **100 images of animals**. Let us divide 100 images as 60% for training, 20% for validation, 20% for testing.
- You train the AI model with **60 images** (*training images*).
- After training, validate the model with **20 images** (*validation images*).
- At the end, you test the model with the **20 test images**.
- The computer never sees the test photos before; it's like a **final exam**.

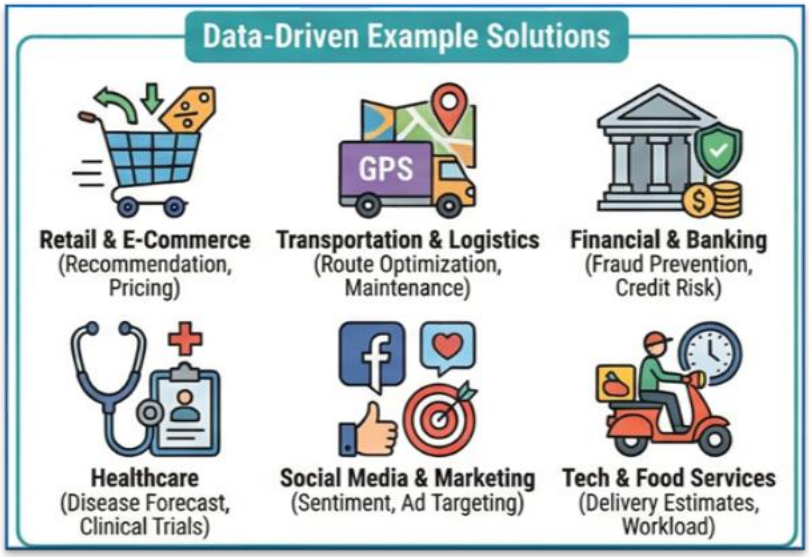
## Data Science

Data science involves studying data, including the development of methods for recording, storing, and interpreting data to extract meaningful information. Data can be numbers, text, images, videos, or other forms of information that can be stored and analysed. Primary areas of mathematics, statistics, computer science, and domain knowledge are all used in data science to identify patterns in the data. These patterns enable humans and machines to make good decisions and solve real-world problems.

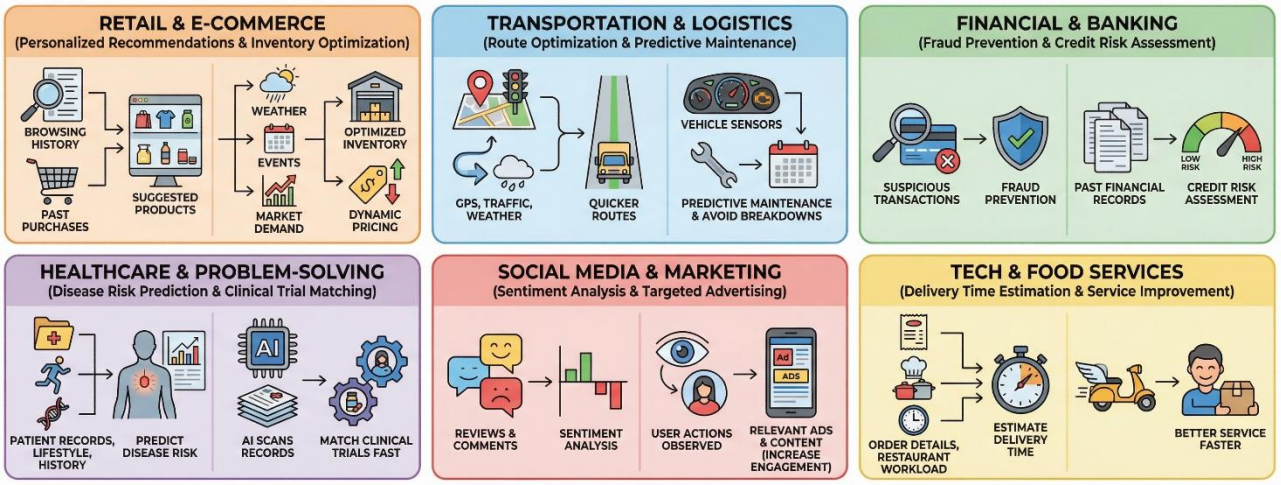



Data Science Function of AI is Enabled by:

- **Data Collection and Preparation:** Raw data is collected from various sources and cleaned so that it is accurate and provides useful input for generating models.
- **Model Training:** AI programs are given processed data in order to identify patterns and generate predictions.
- **Constant Improvement:** Data scientists repetitively analyse and tweak models with fresh data to improve the accuracy of their findings, a process known as feedback loops.




Data-Driven Example Solution Sets




- **Retail Experience and E-Commerce Solutions:**
  Retailers analyse customer behaviour, such as browsing history and past purchases and then **predict** which products or content customers are most likely to pick. Even sales data can be linked with local factors like weather and events to improve stock management. Prices can also be adjusted based on market demand and competitor pricing.

- **Transportation Experience and Logistics Solutions:**
  Artificial intelligence uses GPS **tracking**, **traffic** conditions, and **weather** forecasts to decide how to use quicker delivery routes. Sensors in cars can also give information on how well the car is running, which can help you find out when it needs maintenance and avoid breakdowns.

- **Financial Experience and Banking Solutions:**
  Data is used by banks and other financial institutions to **investigate suspicious** transactions and stop them from happening. Data analytics also help banks determine how much credit risk a borrower carries by examining past records of financial transactions.

- **Healthcare Experience and Problem-Solving Solutions:**
  Hospitals use data to analyse **patient records**, lifestyles, and medical histories in order to **predict** the risk of disease. Similarly, AI technology scans medical records in seconds to match patients with appropriate clinical trials, rather than weeks or months of repetitive searching. This promotes the advancement of research and treatment.

- **Social media and marketing Solutions:**
  Brands look at reviews and comments on the internet to use sentiment analysis to learn more about how people feel. Users' actions are carefully observed so that ads and content are shown that are more relevant to each user's interests, which increases engagement.

- **Tech and Services for Food Solutions:**
  Machine learning is used by food delivery companies to estimate how long it will take to prepare and deliver food by looking at things like the details of the order, the restaurant workload, and the time of day. This helps ensure that customers get better service faster.

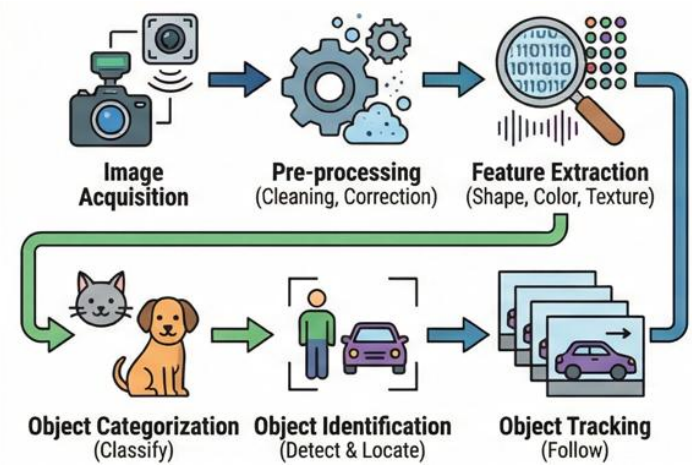
## Computer Vision

Computer vision is the field of study and technology in AI that lets machines **see**, **understand**, and make **sense** of visual information in the same way as we humans do. It gives these machines the ability to process, comprehend and get useful information from images, videos and even real-time video streaming.

## How does Computer Vision Work?

Computer vision is a set of algorithms, techniques, and concepts that enable machines to **recognise** and **interpret visual** information. It is a way of extracting information from images and videos that typically uses machine learning, pattern recognition, image processing and in particular, deep learning and neural networks.

The process of computer vision typically involves the following steps:



- **The Image Acquisition:**
  The computer vision process first **collects** visual **information** from images and videos. This crucial information is collected from several sources such as cameras, sensors and stored image.

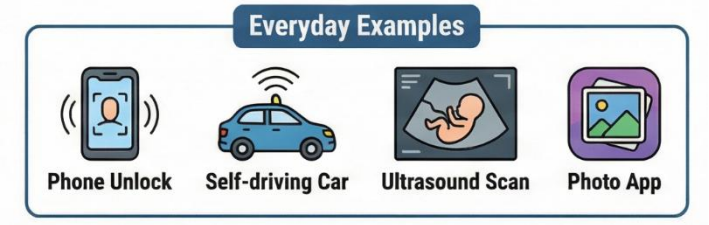
- **Pre-processing:**
  The pre-processing stage ensures that the data is not only **accurate** but also **reliable**. Image pre-processing typically involves noise reduction, distortion correction, and brightness/contrast adjustment to improve image quality.

- **Feature Extraction:**
  Feature extraction involves detecting important **pattern** or feature of an image. These features could include shape and texture, colour distribution, or edge profile. This enables the model to obtain the valuable data and information from sorting and analysing it.

- **Object Categorization:**
  Once the model is trained, it can identify the category of a new image by **comparing** its extracted features with those learned during training. Based on this comparison, the system classifies the object into a specific category. So, for example, the trained model can tell whether a given image is a cat or a dog.

- **Object Identification:**
  It detects and **identifies** specific objects in an image and also finds their location. For example, in a street image, the system can find a car, a person, and a traffic signal, and draw boxes around each object to show where they are placed.

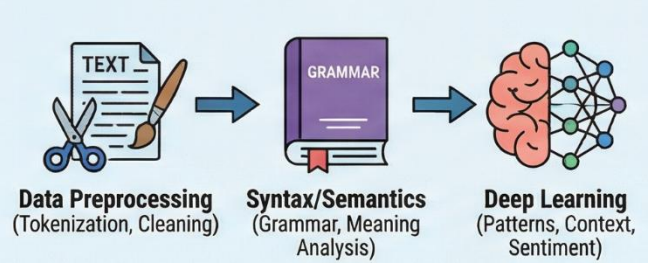
- **Object Tracking:**
  There are different applications for object tracking techniques, such as tracking an object across a sequence of video frames or images. The system tracks objects in each frame, which **preserves** their identity over the entire sequence under challenging conditions such as varying lights and motion.
  For example, in CCTV footage, the system can keep track of a specific car as it moves across different frames, continuously following the same car from the start to the finish of the video.




- **Face Unlock:** Phones create a facial map to unlock the screen.
- **Autonomous Vehicles:** Systems that recognise pedestrians, lane markings and traffic signals.
- **Medical Imaging:** Ultrasound images to check the health of unborn babies.
- **Photo Organization:** Applications that automatically label people and objects in photos for easy searching.

## Natural Language Processing

Natural Language Processing (NLP) is a branch of AI that helps machines read, write, listen, and speak in human language. NLP systems teach machines to read and respond to language like humans do. It enables machines to comprehend not just what the words mean, but also the sentiment behind those words. You can use NLP to tell whether a message is happy or sad or angry and react accordingly. Thanks to NLP, these systems can interact with users in a smarter and more intuitive way.

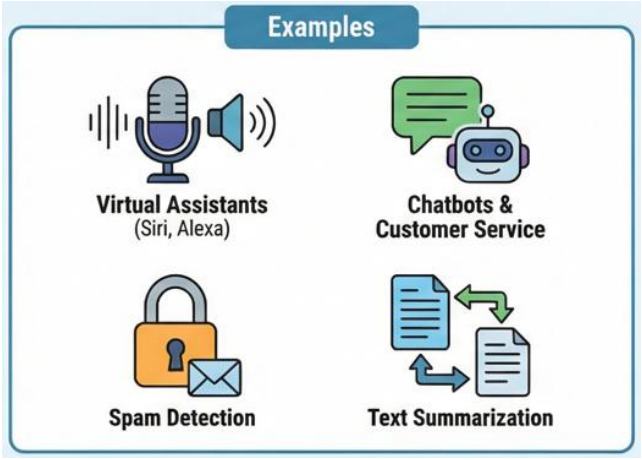



### How Machines Understand Human Language

- **Data Pre-processing (Cleaning):**
  Prior to the actual analysis, raw data is processed by tokenisation, where sentences are broken into words as well as removes punctuation; converts text to lowercase and stemming, which reduces each word to its root form.
- **Syntax/Syntactic Analysis:**
  This is simply how NLP systems study the grammatical structure of sentences to understand how words are arranged.
- **Semantics (Meaning Analysis):**
  Natural language processing systems can understand the meaning of words and phrases, including word-sense disambiguation, which means identifying the correct meaning of words with multiple meanings, such as bat which can mean both sports equipment and an animal.
- **Deep Learning:**
  Huge datasets are used to train models and find patterns that help the system to understand context, sentiment, and purpose.

### Examples of Natural Language Processing

- **Virtual assistants:** Virtual assistants use NLP to understand what you say when you talk to them.
- **Chatbots and Customer Service:** Questions are answered by automated systems, and customer feedback is analysed
- **Spam detection:** Email filters find content that is not appropriate or needed
- **Text Summarization:** Making short summaries of long papers automatically




Points to remember:

- ✓ Artificial Intelligence (AI) is divided into different domains based on the type of data and tasks performed.
- ✓ Data Science focuses on analysing structured data to find patterns and make predictions.
- ✓ Computer Vision enables machines to process, understand and interpret images and videos.
- ✓ Natural Language Processing (NLP) allows machines to read, understand and generate human language.
- ✓ Classification is a technique that assigns data into predefined categories.
- ✓ Regression is used to predict continuous numerical values.
- ✓ Clustering groups similar data together without using predefined labels.

## Exercise

### A. Multiple Choice Questions.

1. Which AI technique predicts numerical values?
   a) Classification b) Clustering
   c) Regression d) Detection
2. Which AI domain works with images and videos?
   a) NLP b) Data Science
   c) Computer Vision d) Science
3. Clustering is mainly used to:
   a) Predict marks b) Group similar data
   c) Translate language d) Detect spam
4. In classification, the output is:
   a) A continuous value b) A predefined category
   c) A random number d) A chart
5. In dataset division, which data is used for final evaluation?
   a) Training b) Validation
   c) Test d) Raw data

## B. Fill in the Blanks.

1. Regression predicts _____ values.
2. Clustering does not require _____data.
3. Computer Vision works mainly with _____ and videos.
4. NLP helps machines understand _____.
5. A dataset is divided into training, validation, and _____sets.

## C. Short Answer Questions.

1. Differentiate between classification and regression.
2. What is clustering? Give one example.
3. Explain the importance of datasets in AI.
4. List the steps involved in Computer Vision.
5. How does NLP help in communication systems?

## D. True or False.

1. Regression predicts categories.
2. Clustering is a supervised learning method.
3. Test data is used to evaluate model performance.
4. Computer Vision can detect object locations.
5. NLP can analyse sentiment in text.

## E. Identify the AI Domain.

| AI APPLICATION | IDENTIFY THE TERM |
|---|---|
| **Analysing MRI scans** | |
| **Text summarization of long articles** | |
| **Recommending products based on past purchases** | |
| **Grouping customers by buying behaviour** | |
| **Chatbot answering customer queries** | |