

B

[Data Science IBM Certification](#) [Data Science](#) [Data Science Projects](#) [Data Analysis](#) [Data Visualization](#)

# Computer Vision Tutorial

Last Updated : 06 Aug, 2025

Computer Vision (CV) is a branch of Artificial Intelligence that enables computers to interpret and understand visual inputs in a manner similar to humans. This tutorial is designed for both beginners and experienced professionals and covers key concepts such as Image Processing, Feature Extraction, Object Detection, Image Segmentation and other core techniques in CV.

Share to

[LinkedIn](#)

[WhatsApp](#)

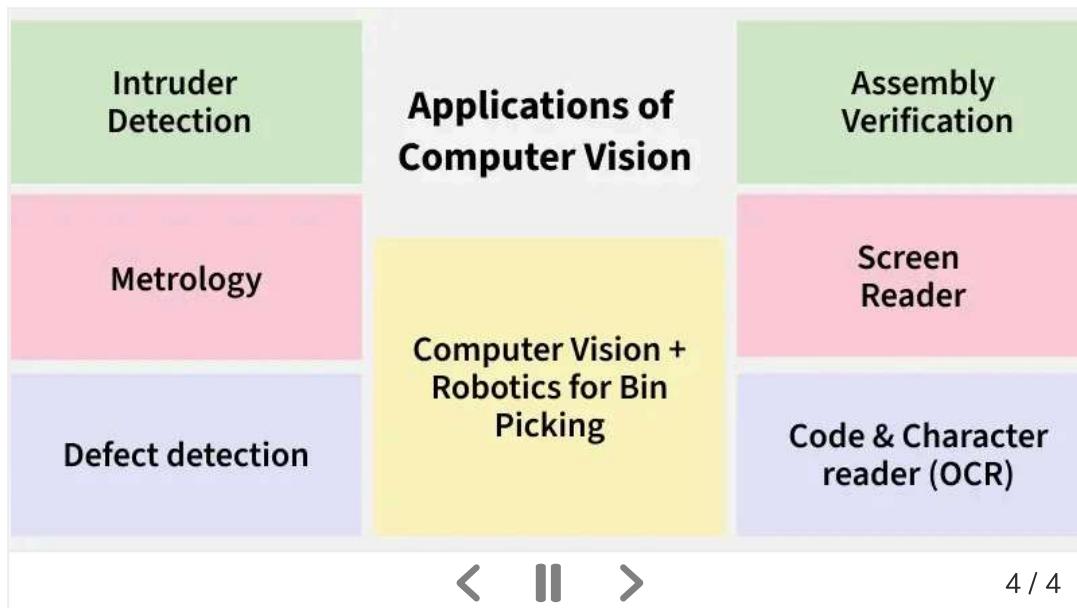
[Twitter](#)

[Copy Link](#)

AI) that helps

much like

experienced



Before moving into computer vision, it is recommended to have a foundational understanding of:

1. [Machine Learning](#)
2. [Deep Learning](#)
3. [OpenCV](#)

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

[Got It !](#)

of these topics, we recommend checking out their respective tutorials to build a solid foundation.

## Mathematical Prerequisites for Computer Vision

Before moving into Computer Vision, having a foundational understanding of certain mathematical concepts will help us which includes:

### 1. Linear Algebra

- [Linear Algebra](#)
- [Vectors](#)
- [Matrices and Tensors](#)
- [Eigenvalues and Eigenvectors](#)
- [Singular Value Decomposition](#)

### 2. Probability and Statistics

- [Probability and Statistics](#)
- [Probability Distributions](#)
- [Bayesian Inference and Bayes' Theorem](#)
- [Markov Chains](#)
- [Kalman Filters](#)

### 3. Signal Processing

- [Signal Processing](#)
- [Image Filtering and Convolution](#)
- [Discrete Fourier Transform \(DFT\)](#)
- [Fast Fourier Transform \(FFT\)](#)
- [Principal Component Analysis \(PCA\)](#)

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#).

## 1. Image Processing

It refers to techniques for manipulating and analyzing digital images.

Common image processing tasks include:

### 1. Image Transformation

- [Image Transformation](#)
- [Geometric Transformations](#)
- [Fourier Transform](#)
- [Intensity Transformation](#)

### 2. Image Enhancement

- [Image Enhancement](#)
- [Histogram Equalization](#)
- [Contrast Enhancement](#)
- [Image Sharpening](#)
- [Color Correction](#)

### 3. Noise Reduction Techniques

- [Noise Reduction Techniques](#)
- [Median Filtering](#)
- [Bilateral Filtering](#)
- [Wavelet Denoising](#)

### 4. Morphological Operations

- [Morphological Operations](#)
- [Erosion and Dilation](#)
- [Opening](#)
- [Closing](#)
- [Morphological Gradient](#)

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#).

It involves identifying distinctive elements within an image for analysis and its techniques include:

## 1. Edge Detection Techniques

- [Computer Vision Algorithms](#)
- [Edge Detection Techniques](#)
- [Canny Edge Detector](#)
- [Sobel Operator](#)
- [Laplacian of Gaussian \(LoG\)](#)

## 2. Corner and Interest Point Detection

- [Harris Corner Detection](#)

## 3. Feature Descriptors

- [Feature Descriptors](#)
- [SIFT \(Scale-Invariant Feature Transform\)](#)
- [SURF \(Speeded-Up Robust Features\)](#)
- [ORB \(Oriented FAST and Rotated BRIEF\)](#)
- [HOG \(Histogram of Oriented Gradients\)](#)

## How Does Computer Vision Work?

1. Computer Vision works much like the human eye and brain. First, our eyes capture the image and send the visual data to our brain. The brain then processes this information and transforms it into a meaningful interpretation, recognizing and categorizing the object based on its properties.
2. In a similar way, Computer Vision uses a camera (acting like the human eye) to capture images. The visual data is then processed by algorithms to recognize and identify the objects based on patterns it has learned. However, before the system can recognize objects in new

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#)

3. For example, imagine providing a computer with thousands of bird song recordings. The system learns by analyzing features like pitch, rhythm and duration. Once trained, it can then recognize whether a new sound resembles a bird song or not.

*For more details you can refer to: [Steps in Computer Vision](#)*

## Popular Libraries for Computer Vision

To implement computer vision tasks effectively, various libraries are used:

1. **OpenCV**: Mostly used open-source library for computer vision tasks like image processing, video capture and real-time applications.
2. **TensorFlow**: A popular deep learning framework that includes tools for building and training computer vision models.
3. **PyTorch**: Another deep learning library that provides great flexibility for computer vision tasks for research and development.
4. **scikit-image**: A part of the scikit-learn ecosystem, this library provides algorithms for image processing and computer vision.

*For more details you can refer to: [Computer Vision Libraries](#)*

## Deep Learning for Computer Vision

Deep learning has greatly enhanced computer vision by allowing machines to understand and analyze visual data and its key deep learning models include:

### 1. Convolutional Neural Networks (CNNs)

Convolutional Neural Networks are designed for learning spatial hierarchies of features from images and its key components include:

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

- [Convolutional Layers](#)
- [Pooling Layers](#)
- [Fully Connected Layers](#)

## 2. Generative Adversarial Networks (GANs)

It consists of two networks (generator and discriminator) that work against each other to create realistic images. There are various types of GANs each designed for specific tasks and improvements:

- [Generative Adversarial Networks \(GANs\)](#).
- [Deep Convolutional GAN \(DCGAN\)](#).
- [Conditional GAN \(cGAN\)](#).
- [Cycle-Consistent GAN \(CycleGAN\)](#).
- [Super-Resolution GAN \(SRGAN\)](#).
- [StyleGAN](#)

## 3. Variational Autoencoders (VAEs)

They are the probabilistic version of autoencoders which forces the model to learn a distribution over the latent space rather than a fixed point, some other autoencoders used in computer vision are:

- [Autoencoders](#)
- [Variational Autoencoders \(VAEs\)](#).
- [Denoising Autoencoders \(DAE\)](#).
- [Convolutional Autoencoder \(CAE\)](#).

## 4. Vision Transformers (ViT)

They are inspired by transformers models to treat images and sequence of patches and process them using self-attention mechanisms, some

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

- [Swin Transformer](#)
- [CvT \(Convolutional Vision Transformer\)](#).

## 5. Vision Language Models

They integrate visual and textual information to perform image processing and natural language understanding.

- [Vision language models](#)
- [CLIP \(Contrastive Language-Image Pre-training\)](#).
- [ALIGN \(A Large-scale ImaGe and Noisy-text\)](#).
- [BLIP \(Bootstrapping Language-Image Pre-training\)](#).

# Computer Vision Tasks

## 1. Image Classification

It involves analyzing an image and assigning it a specific label or category based on its content such as identifying whether an image contains a cat, dog or car.

Its techniques are as follows:

- [Computer Vision Tasks](#)
- [Image Classification](#)
- [Image Classification using Support Vector Machine \(SVM\)](#).
- [Image Classification using RandomForest](#)
- [Image Classification using CNN](#)
- [Image Classification using TensorFlow](#)
- [Image Classification using PyTorch Lightning](#)

There are various types for Image Classification which are as follows:

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

- Zero-shot classification

*To learn about the datasets for image classification, we can go through the article on Dataset for Image Classification mentioned above.*

## 2. Object Detection

It involves identifying and locating objects within an image by drawing bounding boxes around them.

**It includes below following Techniques:**

- Top Computer Vision Models
- Object Detection
- YOLO (You Only Look Once).
- SSD (Single Shot Multibox Detector).
- Region-Based Convolutional Neural Networks (R-CNNs).
- Fast R-CNN
- Faster R-CNN
- Mask R-CNN
- Object Detection using TensorFlow
- Object Detection using PyTorch

**Type of Object Detection Concepts are as follows:**

- Bounding Box Regression
- Intersection over Union (IoU).
- Region Proposal Networks (RPN).
- Non-Maximum Suppression (NMS).

## 3. Image Segmentation

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#)

## Types of image segmentation are:

- [Image Segmentation](#)
- [Semantic Segmentation](#)
- [Instance Segmentation](#)
- [Panoptic Segmentation](#)

We can perform image segmentation using the following methods:

- [Image Segmentation using K Means Clustering](#)
- [Image Segmentation using UNet](#)
- [Image Segmentation using TensorFlow](#)
- [Image Segmentation with Mask R-CNN](#)

## Need for Computer Vision

- 1. High Demand in the Job Market:** Critical for careers in AI, machine learning and data science across industries like healthcare, automotive and robotics.
- 2. Revolutionizing Industries:** Powers advancements in self-driving cars, medical diagnostics, agriculture and manufacturing by automating visual tasks.
- 3. Solving Real-World Problems:** Enhances safety, improves medical imaging and optimizes industrial processes.
- 4. Improving Accessibility:** It helps people with disabilities through image recognition and sign language translation.
- 5. Enhancing Consumer Experiences:** It personalizes shopping and improves customer service in retail and entertainment.

## Applications of Computer Vision

- 1. Healthcare:** Used for disease detection and medical image analysis (X-rays, MRIs).

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#).

3. **Retail:** It helps with inventory management, theft prevention and customer behavior analysis.
4. **Agriculture:** It is used for crop monitoring and disease detection.
5. **Security and Surveillance:** It recognizes faces and find suspicious activities in security footage.

For more details you can refer to: [Applications of Computer Vision](#)

[Comment](#)[More info](#)[Advertise with us](#)

Corporate & Communications Address:

A-143, 7th Floor, Sovereign Corporate  
Tower, Sector- 136, Noida, Uttar Pradesh  
(201305)

Registered Address:

K 061, Tower K, Gulshan Vivante  
Apartment, Sector 137, Noida, Gautam  
Buddh Nagar, Uttar Pradesh, 201305

[Advertise with us](#)[Company](#)[Explore](#)

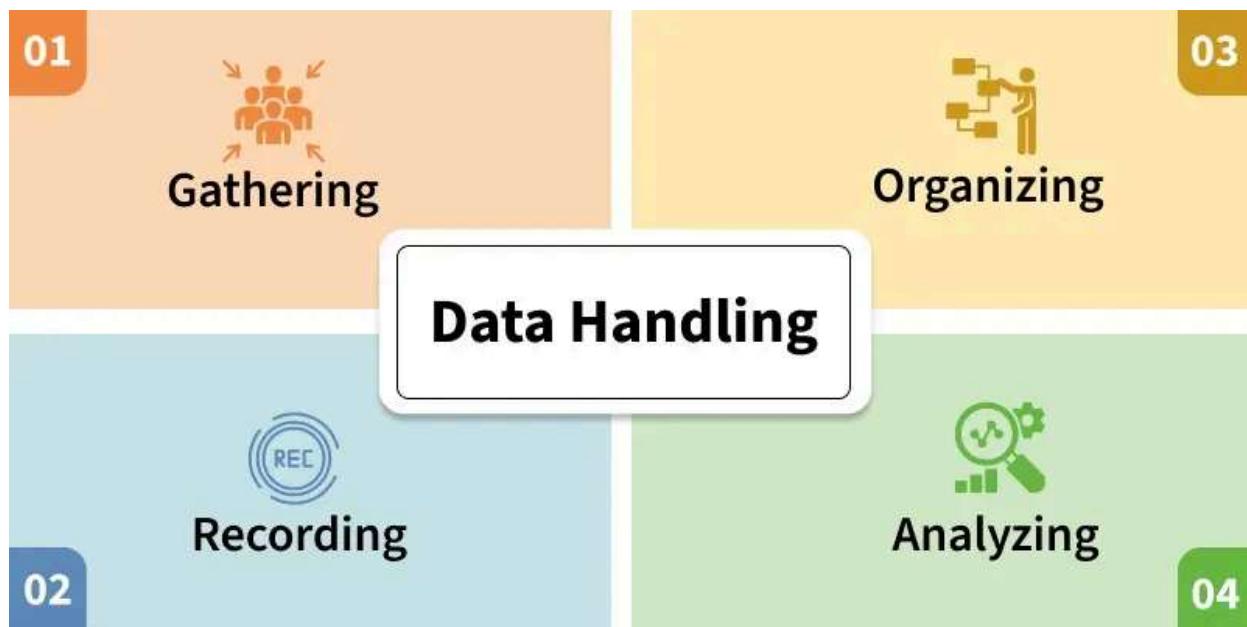
We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

[Number System and Arithmetic](#) [Algebra](#) [Set Theory](#) [Probability](#) [Statistics](#) [Geometry](#) [Calculus](#)

# Data Handling

Last Updated : 23 Jul, 2025

Data handling is the process of systematically collecting, organizing, analyzing, and presenting data to extract useful information and support decision-making. It involves ensuring the accuracy and integrity of data, processing it into a manageable form, and presenting it through clear formats such as charts, graphs, and tables. The goal is to make complex data easier to understand and interpret, ensuring its safe storage and accessibility throughout its lifecycle.



*Components of Data Handling*

**For studying this topic, we also focus on how data is collected or generated for our use from any source. This procedure involves the steps to acquire the data, clean it, and prepare it for analysis and study.**

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

[Got It !](#)

& [Privacy Policy](#)

Data can be segregated into two broad categories: **Quantitative Data** and **Qualitative Data**.

## Quantitative Data

Quantitative Data is data that gives numerical information. It is measurable and quantifiable, and it can be written in terms of numbers, which are amenable to mathematical operations. Quantitative data can be further categorized into:

- **Discrete Data:** Data that can have definite, distinct values (e.g., number of individuals in a room).
- **Continuous Data:** Data that can have any value between a range (e.g., height, weight, temperature).

## Qualitative Data

Qualitative Data is non-numerical data describing features or qualities. Qualitative Data is descriptive data and cannot be directly measured in terms of numbers. Examples are names, colours, or classes of objects.

➤ Learn the difference the Quantitative and Qualitative data -  
[[Read More](#)]

## Important Terms in Data Handling

There are some terms used often in data handling to better understand and deal with data:

- 1) Data:** The set of facts or figures obtained from observations or measurements.
- 2) Raw Data:** Raw and unstructured data collected directly from the source and could require cleaning or formatting.

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#)

variability of the data.

**4) Statistics:** The branch of science that deals with the collection, organization, analysis, interpretation, and presentation of numerical data. It facilitates data-informed decisions.

## Steps Involved in Data Handling

Data handling is a methodical process of handling and interpreting data properly. The steps usually consist of-

**Purpose:** Define and state specifically the purpose or problem. This makes the data handling process dedicated to answering one specific question or resolving a unique problem.

**Collection of Data:** Collect data pertaining to the set purpose. The accuracy and quality of data gathered play a very important role in significant analysis.

**Presentation of Data:** Display the gathered data in a simple and easy-to-read manner. It can be in the form of tables, graphs, or marks, based on the complexity of the data.

**Graphical Representation of Data:** Employ visual aids such as graphs, histograms, and bar charts to represent the data. Graphical representation assists in the easy analysis of trends and patterns in the data.

**Analyzing the Data:** Scan the data carefully to derive useful information. Statistical techniques or other techniques of analysis are used to derive insights.

**Conclusion/Inference:** From the analysis, make conclusions or inferences to give a solution or response to the problem statement. This step helps in decision-making or subsequent actions.

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#).

In data handling, one of the most significant areas is how the data is represented. Proper representation is easier to understand, analyze, and interpret. Among all the ways data can be represented, graphical representation is particularly good at representing trends and patterns effectively and speedily.

- [Pictographs or Picture Graphs](#)
- [Bar Graphs](#)
- [Line Graphs](#)
- [Pie Charts](#)
- [Scatter Plot](#)

## Pictographs

A pictograph is the pictorial representation of any data given to us in written form. It can be said that pictographs used to be the earliest form of communication, since, way back in time, people communicated mostly through pictures with each other, as languages were not present.

Indeed, a pictograph plays a role in our day-to-day life too. For instance, when a friend tells us a story, we start imagining the story in our head, and that makes it both easy to understand and easy to remember for a long time.

### Drawing a Pictograph

Let's learn to draw a pictograph with the help of an example.

**Example:** In a reading competition, three students were participating—Rahul, Saumya, and Ankush. They were supposed to read as many books as they could in an hour. Rahul read 3 books, Saumya read 2 books, and Ankush read 4 books. Draw the pictograph for the information.

### Solution:

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#)

- Decide the particular picture or pictures that are required to represent the data. Make sure that the picture is related to the information for easier memorization.
- Here, a smiley face is used to represent each book read.



- Now, draw the pictures according to the information presented. For example, there will be 3 smiley faces for Rahul, as he completed 3 books in an hour.

Rahul	Three smiley faces are drawn in the cell.
saumya	Two smiley faces are drawn in the cell.
Ankush	Four smiley faces are drawn in the cell.

## Bar Graphs

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

the data look neat, but it is also easier to compare the data given.

## Types of Bar Graph

Various types of bar graphs include:

### Vertical Bar Graph

These are the most common bar graphs we come across; the bars of grouped data in vertical bar graphs lie vertically. Sometimes, when the data categories have long names, horizontal bar graphs are preferred, since in vertical bar graphs, there is not much space on the x-axis.

An example explaining the concept of a Bar graph is added below:

**Example:** There are 800 students in a school. And the table for their birthdays in all 12 months is given below. Draw the Vertical Bar graph and answer. who

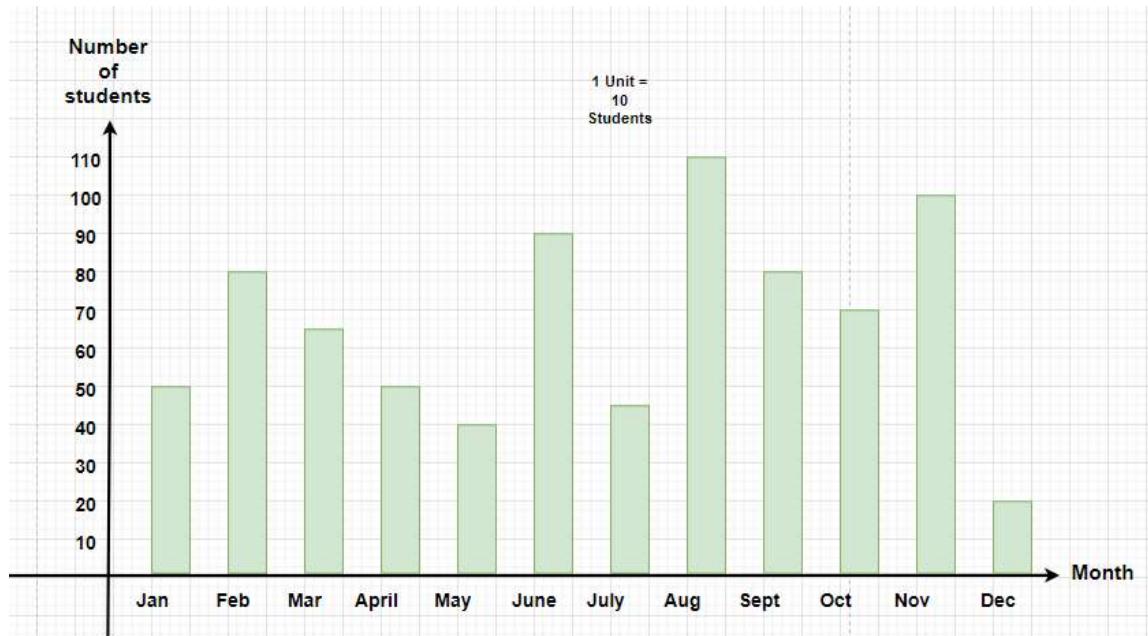
Months	January	February	March	April	May	June	July
No. of Students	50	80	65	50	40	90	45



1. In which month do the maximum number of students have their birthdays?
2. Which two months have the same number of birthdays?
3. In which month do the minimum number of students have their birthdays?

### Solution:

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).



From the Bar graph we can figure out the answer of the questions

1. August is the month in which the maximum number of birthdays occur.(there are 110 students whose birthday come in August)
2. From the graph, we can tell that January and April have equal lengths of bars, which means they have the same number of birthdays.(both have 50 birthdays)
3. The minimum number of birthdays occur in December, as it has the smallest bar (20 students have their birthdays in December).

## Horizontal Bar Graph

A horizontal bar graph is a graph where the rectangular bars lie horizontally. In such graphs, the frequency of the data is represented on the x-axis, while the categories of data are shown on the y-axis. These are known as horizontal bar graphs.

Horizontal bar graphs are preferred when the category names are long and there is insufficient space on the x-axis.

**Example:** In an examination, Reeta appeared for 5 subjects. Her

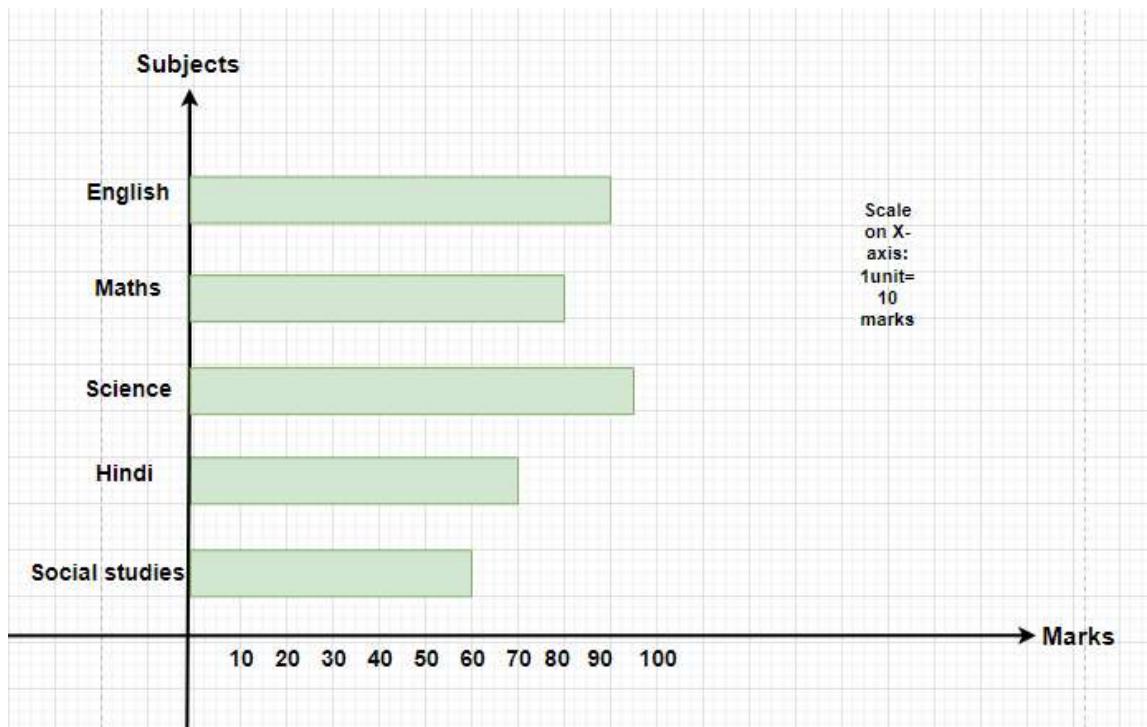
We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#)

Subjects	English	Maths	Science	Hindi	Social Science
Marks	90	80	95	70	60

**Solution:**

*The Horizontal bar graph for the table mentioned in the question,*



*The overall Percentage obtained by Reeta =  $\frac{(90+80+95+70+60)}{500} \times 100$   
= 79 percent.*

## Double-Bar Graph

Double-bar graphs are used when two groups of data are required to be represented on a single graph. In a double-bar graph, to represent two

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#).

## Advantages of a points double-bar graph:

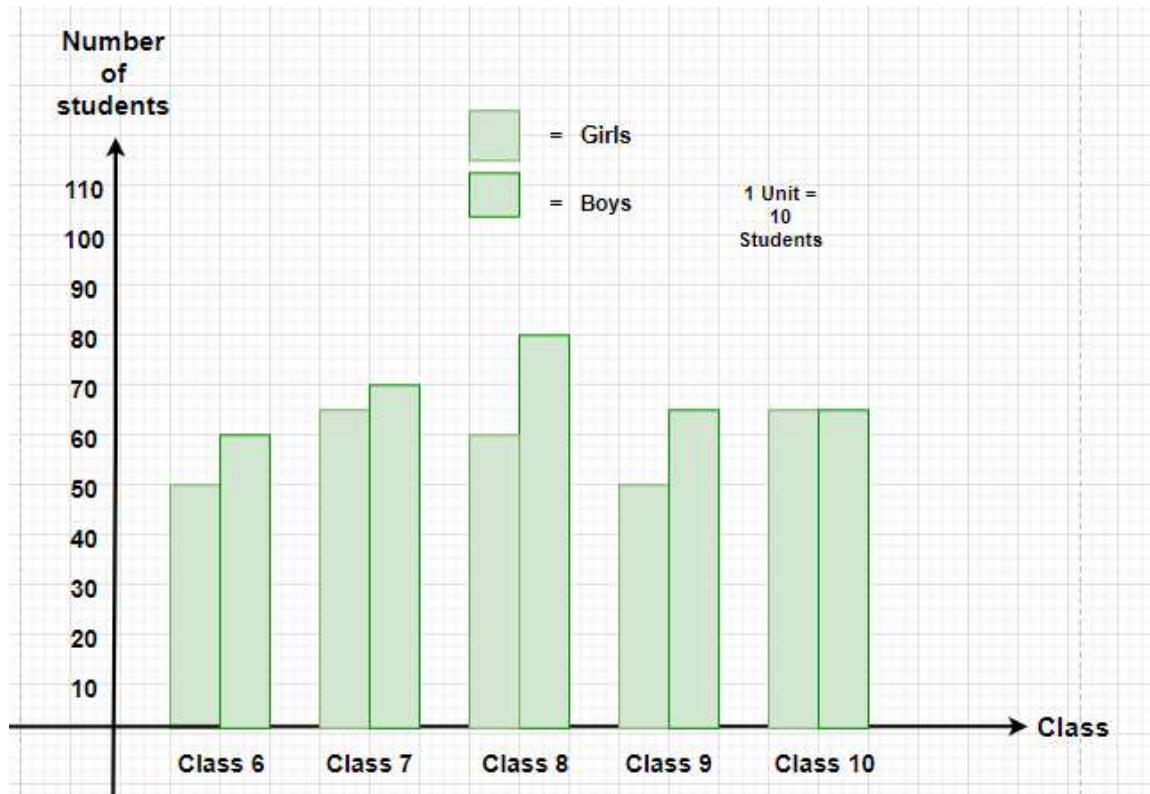
- A double-bar graph is helpful when multiple data points need to be represented.
- It helps in summarizing large and big data in an easy and visual form.
- It shows and covers all different frequency distributions.

**Example:** The table below shows the number of boys and girls in classes 6, 7, 8, 9, and 10. Represent the data on a double-bar graph

Class	6	7	8	9	10
Number of boys	60	70	80	65	65
Number of Girls	50	65	60	50	65

**Solution:**

*The double-bar graph for the table given the question,*



## Line Graphs

A **Graph** or a line chart visually shows how different things relate over time by connecting dots with straight lines. It helps us see patterns or trends in the data, making it easier to understand how variables change or interact with each other as time goes by.

### To make a line graph

We need to use the following steps:

- Determine Variables:** The first and foremost step to creating a line graph is to identify the variables you want to plot on the X-axis and Y-axis.
- Choose Appropriate Scales:** Based on your data, determine the appropriate scale.

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

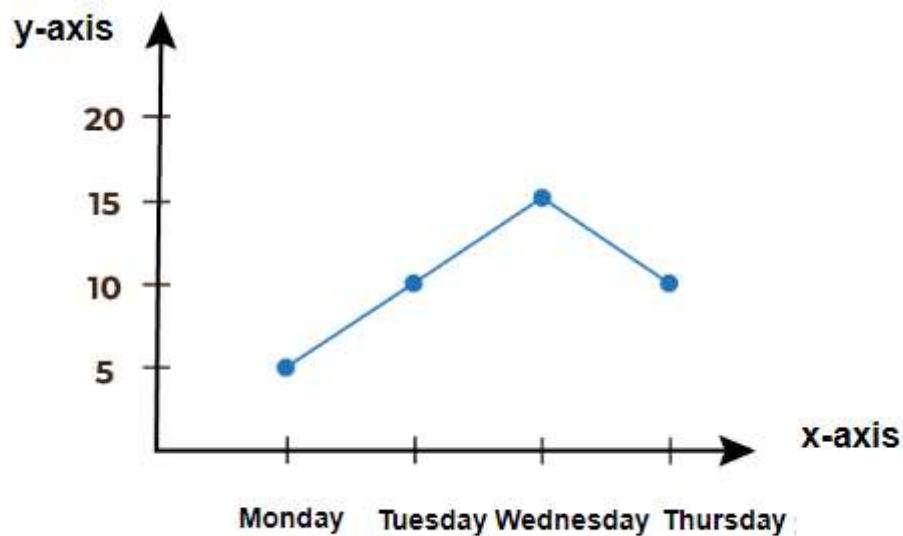
& [Privacy Policy](#)

- **Connect Points:** After plotting the points, you have to connect those points with a line.
- **Label Axes:** Add labels to the X-axis and Y-axis. You can also include the unit of measurement.
- **Add Title:** After completing the graph, you should provide a suitable title.

**Example:** Kabir eats eggs each day, and the data for the same is added in the table below. Draw a line graph for the given data

Weekdays	Monday	Tuesday	Wednesday	Thursday
Eggs Eaten	5	10	15	10

**Solution:**



those sectors show the different parts of the data from the whole.

Pie charts, also known as circle graphs or pie diagrams, are very useful in representing and interpreting data

**Example:** In an office, the data of employees who play various sports is shown in the table below:

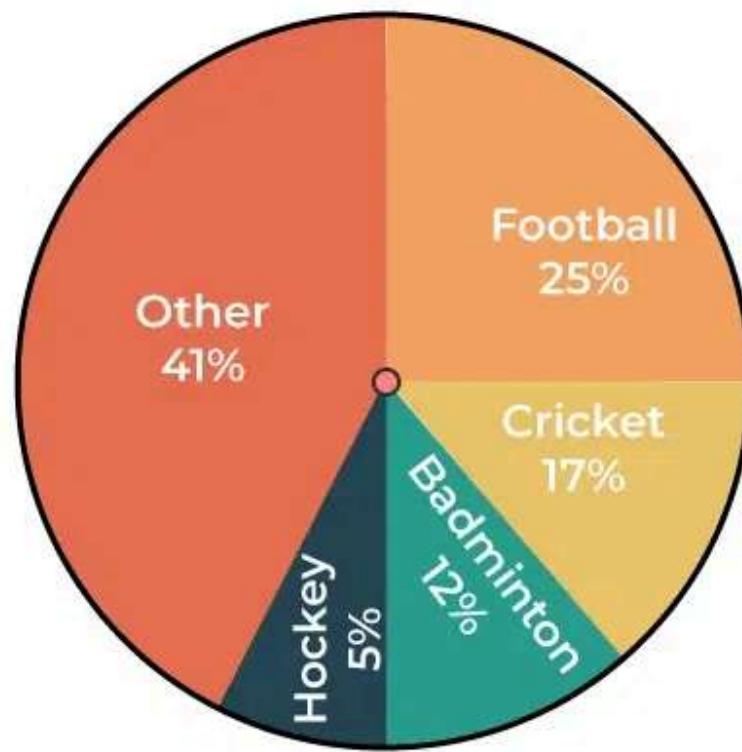
Sport	Cricket	Football	Badminton	Hockey	Other
Number of Employees	34	50	24	10	82

Draw a suitable pie chart.

**Solution:**

*The required pie chart for the given data is:*

# Number of Employees

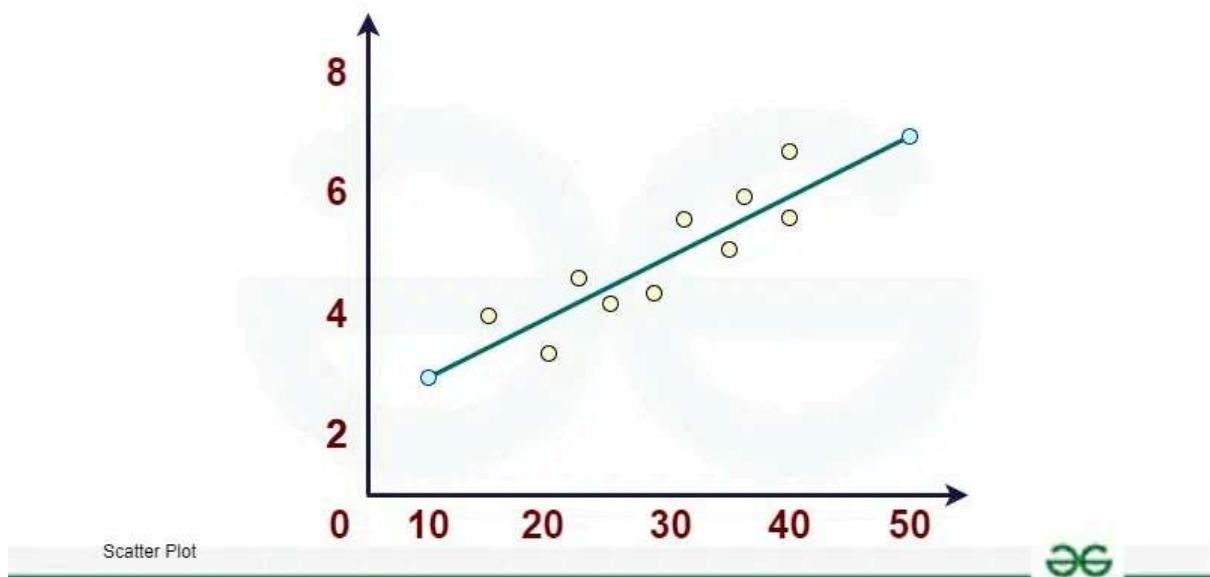


## Scatter Plot

A scatter plot is a type of graphical representation that displays individual data points on a two-dimensional coordinate system. Each point on the plot represents the values of two variables, allowing us to observe any patterns, trends, or relationships between them. Typically, one variable is plotted on the horizontal axis (x-axis), and the other variable is plotted on the vertical axis (y-axis).

Scatter plots are commonly used in data analysis to visually explore the relationship between variables and to identify any correlations or outliers present in the data.

A line drawn in a scatter plot that is near to almost all the points in the



## Solved Examples on Data Handling

**Example 1:** In a survey conducted over a week, from Monday to Sunday, for two cities, Agra and Delhi, the temperatures of both cities were measured, and the obtained temperatures are represented as numbers.

Week	Mon	Tues	Wed	Thurs	Fri	Sat	Sun
Agra (temperature)	15	17	14	19	11	15	10
Delhi	12	15	15	21	10	14	12

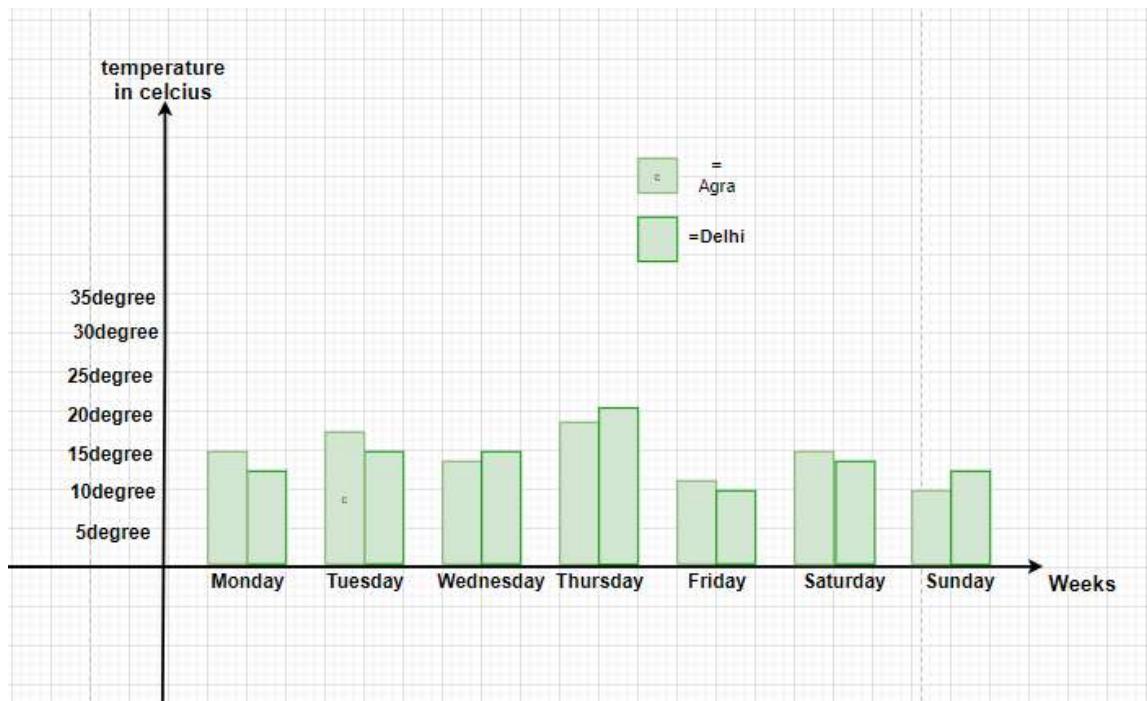
Draw the Bar Graph for the given table in the question.

**Solution:**

The given table has two categories of data: one for the temperature in Agra and the other for the temperature in Delhi. Therefore, the graph can be drawn as a double-bar graph, which would look like the following

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#)

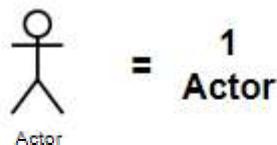


**Example 2:** In a theater, there are 3 plays with a different number of actors participating in each play. In Play 1, there are 9 actors; in Play 2, there are 3 fewer actors, and the number of actors in Play 3 is one less than in Play 1. Draw the pictograph for the information given and analyze at which point the stage will be most crowded..

**Solution:**

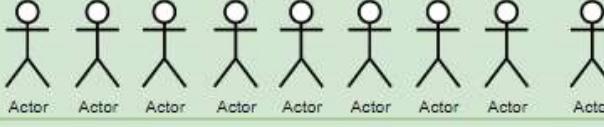
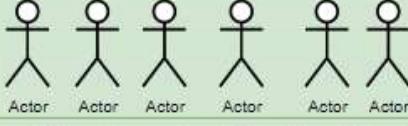
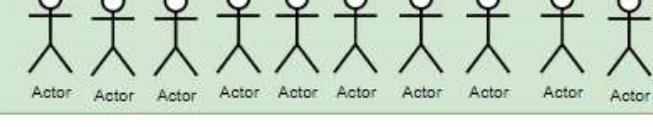
*From the information given in the question, we can say that Play 1 has 9 actors, Play 2 has 6 actors, and Play 3 has 10 actors.*  
*Representing the actors in pictorial form:*

*Representing the actors in pictorial form as follows:*



We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#).

<b>Play 1</b>	 Actor Actor Actor Actor Actor Actor Actor Actor Actor
<b>Play 2</b>	 Actor Actor Actor Actor Actor Actor
<b>Play 3</b>	 Actor Actor Actor Actor Actor Actor Actor Actor Actor Actor

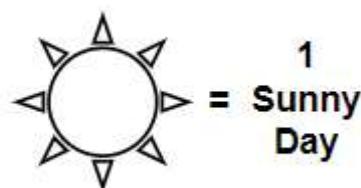
Therefore, we can conclude that Play 3 has the most crowded stage, as it has 10 actors performing on stage.

**Example 3:** In a weather report conducted over 5 consecutive weeks, it was noted that not all days were sunny during the spring season. The observation revealed that Week 1 had 4 sunny days, Week 2 had 5 sunny days, Week 3 had only 2 sunny days, Week 4 had sunny days throughout the entire week, and Week 5 had only 3 sunny days.

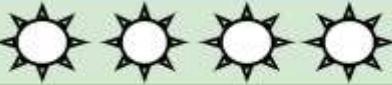
**Pictograph for the number of sunny days in each week:**

**Solution:**

Representing sunny days in pictorial form for better understanding,



We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

<b>Week 1</b>	
<b>Week 2</b>	
<b>Week 3</b>	
<b>Week 4</b>	
<b>Week 5</b>	

#### Example 4: Calculating the Mode

**Problem:** Find the mode of the following data set: 7, 8, 7, 9, 10, 7, 8.

**Solution:**

**Note:-** If one number appears more frequently than others, it is the mode.

- Count the frequency of each number.
- 7 appears 3 times, 8 appears 2 times, 9 appears 1 time, and 10 appears 1 time.
- The mode is the number that appears most frequently.
- Mode = 7

#### Example 5: Determining the Range

**Problem:** Calculate the range of the following set of numbers: 12, 7, 15, 9, 14.

**Solution:**

- The range is the difference between the highest and lowest

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

- $\text{Lowest number} = 7$
- $\text{Range} = \text{Highest} - \text{Lowest} = 15 - 7 = 8$

## Practice Problems on Data Handling

1. Draw a pictograph to represent the number of apples, oranges, and bananas sold in a fruit store: Apples (10), Oranges (15), Bananas (12).
2. Create a vertical bar graph showing the number of books sold in January (100), February (120), March (90), and April (110).
3. Draw a horizontal bar graph to represent the number of cars sold in five months: January (15), February (20), March (10), April (25), May (30).
4. Represent the number of students in three classes (Class A: 20, Class B: 25, Class C: 30) using a double-bar graph.
5. Draw a line graph for the temperature recorded over five days: Monday (22°C), Tuesday (25°C), Wednesday (27°C), Thursday (24°C), Friday (23°C).
6. Create a pie chart showing the distribution of different types of fruits in a basket: Apples (40%), Oranges (25%), Bananas (20%), Grapes (15%).
7. Plot a scatter plot to show the relationship between study hours and exam scores for ten students.
8. Draw a pictograph to represent the number of hours spent on homework by five students: A (4 hours), B (3 hours), C (2 hours), D (5 hours), E (4 hours).
9. Create a vertical bar graph to show the scores of students in a test: Alice (85), Bob (90), Charlie (80), and Diana (95).
10. Draw a horizontal bar graph representing the number of hours worked in a week by different employees: John (40 hours), Mary (35 hours), Alex (45 hours), Sarah (30 hours).

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

- [Difference Between Mean, Median, and Mode with Examples](#)
- [Chance and Probability](#)

## Summary

Data handling refers to the systematic process of collecting, recording, and representing data in ways that make it easy to understand and analyze. This involves using various graphical methods such as pictographs, bar graphs, line graphs, pie charts, and scatter plots, all of which help visualize data and identify patterns or trends. Mastery of these data handling techniques is essential for making accurate analyses and informed decisions in both academic and professional contexts.

[Comment](#)[More info](#)[Advertise with us](#)

Corporate & Communications Address:

A-143, 7th Floor, Sovereign Corporate  
Tower, Sector- 136, Noida, Uttar Pradesh  
(201305)

Registered Address:

K 061, Tower K, Gulshan Vivante  
Apartment, Sector 137, Noida, Gautam  
Buddh Nagar, Uttar Pradesh, 201305



We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

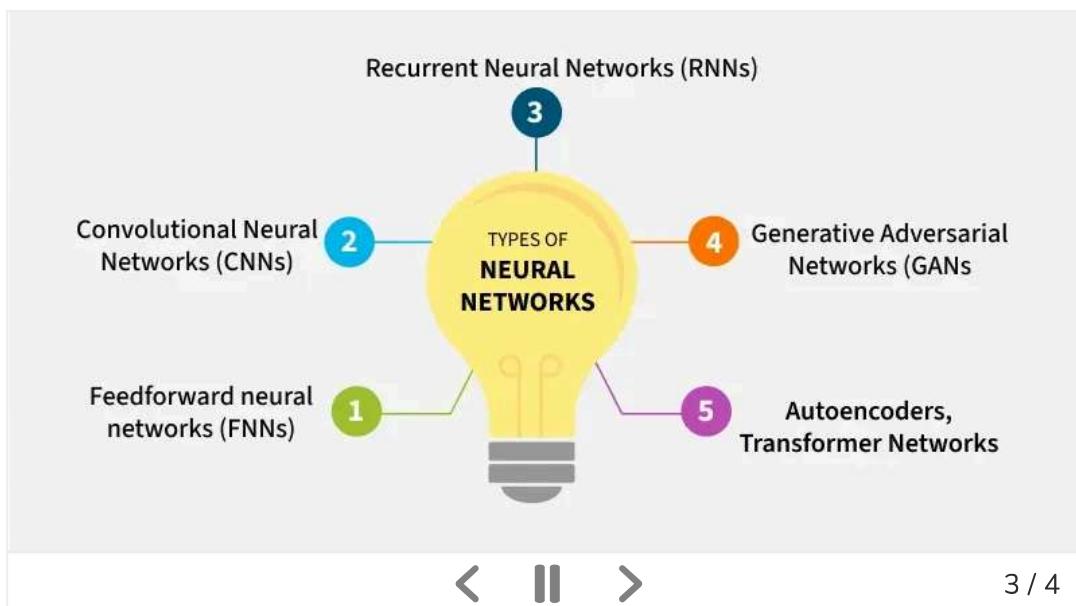
B

[Data Science IBM Certification](#)   [Data Science](#)   [Data Science Projects](#)   [Data Analysis](#)   [Data Visualization](#)

# Introduction to Deep Learning

Last Updated : 11 Jul, 2025

Deep Learning is transforming the way machines understand, learn and interact with complex data. Deep learning mimics neural networks of the human brain, it enables computers to autonomously uncover patterns and make informed decisions from vast amounts of unstructured data.



## How Deep Learning Works?

**Neural network** consists of layers of interconnected nodes or neurons that collaborate to process input data. In a **fully connected deep neural network** data flows through multiple layers where each neuron performs nonlinear transformations, allowing the model to learn intricate representations of the data.

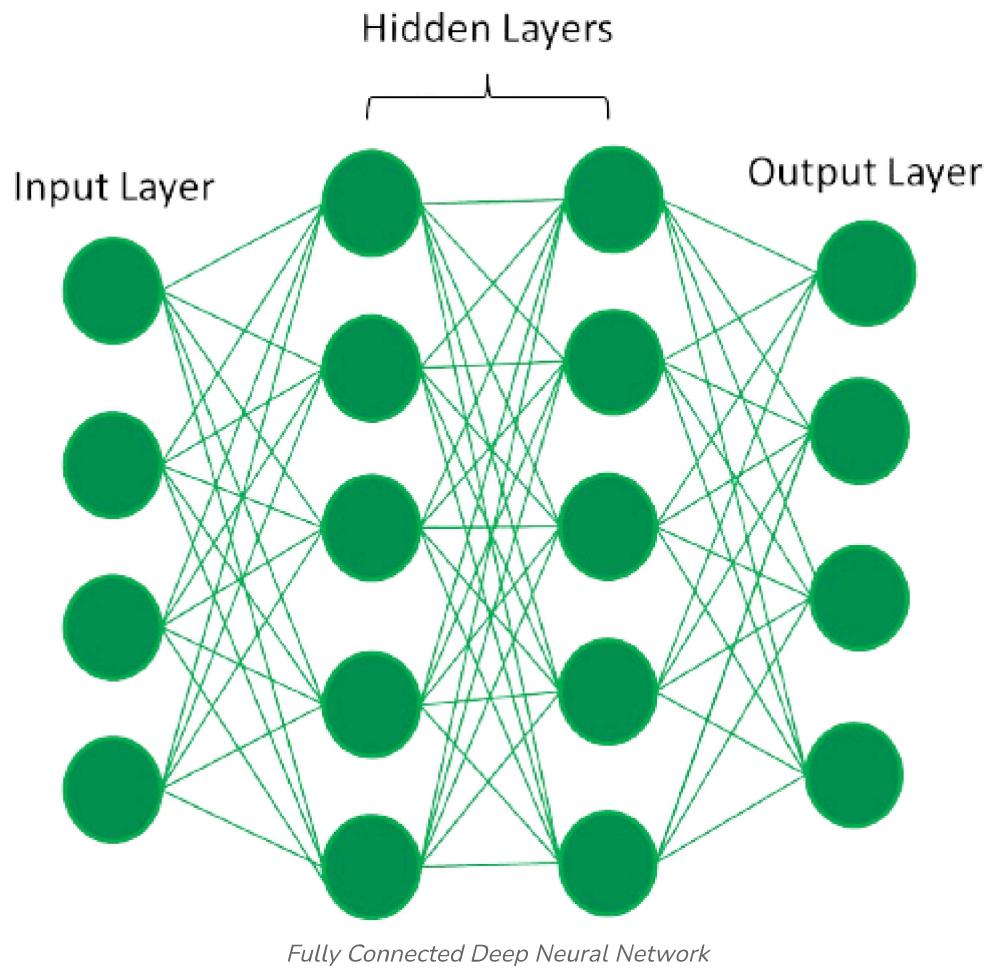
In a deep neural network the **input layer** receives data which passes

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

[Got It !](#)

& [Privacy Policy](#)

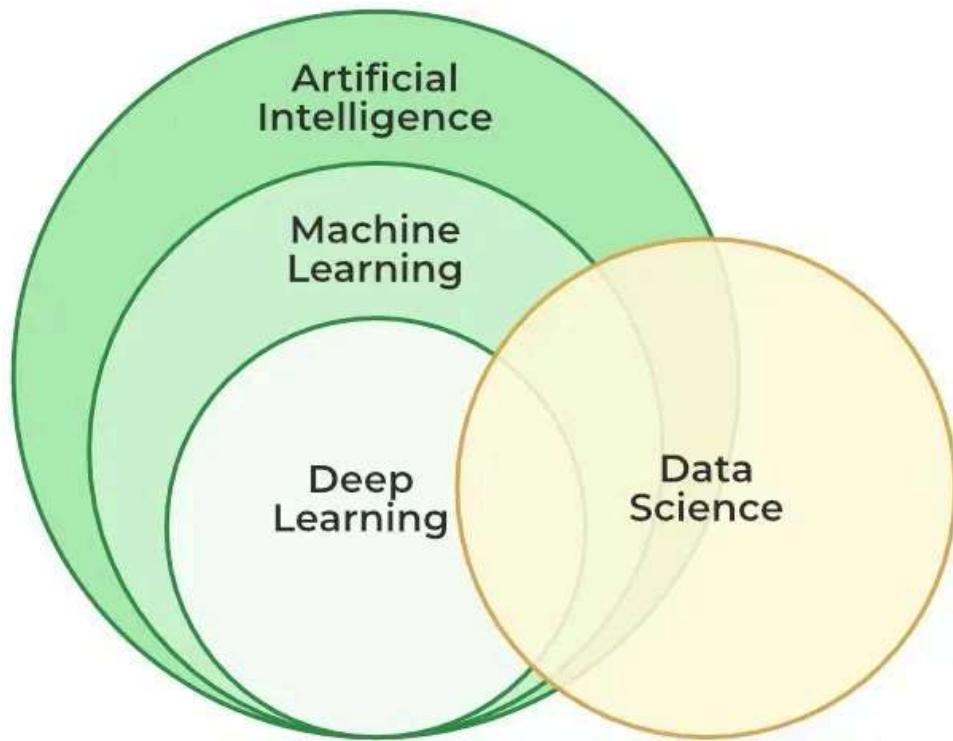
For more details on neural networks refer to this article: [What is a Neural Network?](#)



## Difference between Machine Learning and Deep Learning

Machine learning and Deep Learning both are subsets of artificial intelligence but there are many similarities and differences between them.

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).



Machine Learning	Deep Learning
Apply statistical algorithms to learn the hidden patterns and relationships in the dataset.	Uses artificial neural network architecture to learn the hidden patterns and relationships in the dataset.
Can work on the smaller amount of dataset	Requires the larger volume of dataset compared to machine learning
Better for the low-label task.	Better for complex task like image processing, natural language processing, etc.
Takes less time to train the model.	Takes more time to train the model.

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

Machine Learning	Deep Learning
extracted from images to detect an object in the image.	to-end learning process.
Less complex and easy to interpret the result.	More complex, it works like the black box interpretations of the result are not easy.
It can work on the CPU or requires less computing power as compared to deep learning.	It requires a high-performance computer with GPU.

## Evolution of Neural Architectures

The journey of deep learning began with the [perceptron](#), a single-layer neural network introduced in the 1950s. While innovative, perceptrons could only solve linearly separable problems hence failing at more complex tasks like the XOR problem.

This limitation led to the development of [Multi-Layer Perceptrons \(MLPs\)](#). It introduced hidden layers and non-linear activation functions. MLPs trained using [backpropagation](#) could model complex, non-linear relationships marking a significant leap in neural network capabilities. This evolution from perceptrons to MLPs laid the groundwork for advanced architectures like CNNs and RNNs, showcasing the power of layered structures in solving real-world problems.

## Types of neural networks

1. [Feedforward neural networks \(FNNs\)](#) are the simplest type of ANN, where data flows in one direction from input to output. It is used for basic tasks like classification.

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#).

layers to detect spatial hierarchies, making them ideal for computer vision tasks.

3. **Recurrent Neural Networks (RNNs)** are able to process sequential data, such as time series and natural language. RNNs have loops to retain information over time, enabling applications like language modeling and speech recognition. Variants like LSTMs and GRUs address vanishing gradient issues.
4. **Generative Adversarial Networks (GANs)** consist of two networks—a generator and a discriminator—that compete to create realistic data. GANs are widely used for image generation, style transfer and data augmentation.
5. **Autoencoders** are unsupervised networks that learn efficient data encodings. They compress input data into a latent representation and reconstruct it, useful for dimensionality reduction and anomaly detection.
6. **Transformer Networks** has revolutionized NLP with self-attention mechanisms. Transformers excel at tasks like translation, text generation and sentiment analysis, powering models like GPT and BERT.

## Deep Learning Applications

### 1. Computer vision

In computer vision, deep learning models enable machines to identify and understand visual data. Some of the main applications of deep learning in computer vision include:

- **Object detection and recognition:** Deep learning models are used to identify and locate objects within images and videos, making it possible for machines to perform tasks such as self-driving cars,

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#)

used in applications such as medical imaging, quality control and image retrieval.

- **Image segmentation:** Deep learning models can be used for image segmentation into different regions, making it possible to identify specific features within images.

## 2. Natural language processing (NLP)

In NLP, deep learning model enable machines to understand and generate human language. Some of the main applications of deep learning in NLP include:

- **Automatic Text Generation:** Deep learning model can learn the corpus of text and new text like summaries, essays can be automatically generated using these trained models.
- **Language translation:** Deep learning models can translate text from one language to another, making it possible to communicate with people from different linguistic backgrounds.
- **Sentiment analysis:** Deep learning models can analyze the sentiment of a piece of text, making it possible to determine whether the text is positive, negative or neutral.
- **Speech recognition:** Deep learning models can recognize and transcribe spoken words, making it possible to perform tasks such as speech-to-text conversion, voice search and voice-controlled devices.

## 3. Reinforcement learning

In reinforcement learning, deep learning works as training agents to take action in an environment to maximize a reward. Some of the main applications of deep learning in reinforcement learning include:

- **Game playing:** Deep reinforcement learning models have been able to

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

and manipulation.

- **Control systems:** Deep reinforcement learning models can be used to control complex systems such as power grids, traffic management and supply chain optimization.

## Advantages of Deep Learning

1. **High accuracy:** Deep Learning algorithms can achieve state-of-the-art performance in various tasks such as image recognition and natural language processing.
2. **Automated feature engineering:** Deep Learning algorithms can automatically discover and learn relevant features from data without the need for manual feature engineering.
3. **Scalability:** Deep Learning models can scale to handle large and complex datasets and can learn from massive amounts of data.
4. **Flexibility:** Deep Learning models can be applied to a wide range of tasks and can handle various types of data such as images, text and speech.
5. **Continual improvement:** Deep Learning models can continually improve their performance as more data becomes available.

## Disadvantages of Deep Learning

Deep learning has made significant advancements in various fields but there are still some challenges that need to be addressed. Here are some of the main challenges in deep learning:

1. **Data availability:** It requires large amounts of data to learn from. For using deep learning it's a big concern to gather as much data for training.
2. **Computational Resources:** For training the deep learning model, it is computationally expensive because it requires specialized hardware like GPUs and TPUs.

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#).

months.

4. **Interpretability:** Deep learning models are complex, it works like a black box. It is very difficult to interpret the result.
5. **Overfitting:** when the model is trained again and again it becomes too specialized for the training data leading to overfitting and poor performance on new data.

As we continue to push the boundaries of computational power and dataset sizes, the potential applications of deep learning are limitless.

Deep Learning promises to reshape our future where machines can learn, adapt and solve complex problems at a scale and speed previously unimaginable.

---

What is Deep Learning? Difference between ML  
and Deep learning.

[Visit Course](#)

[Comment](#)

[More info](#)

[Advertise with us](#)

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

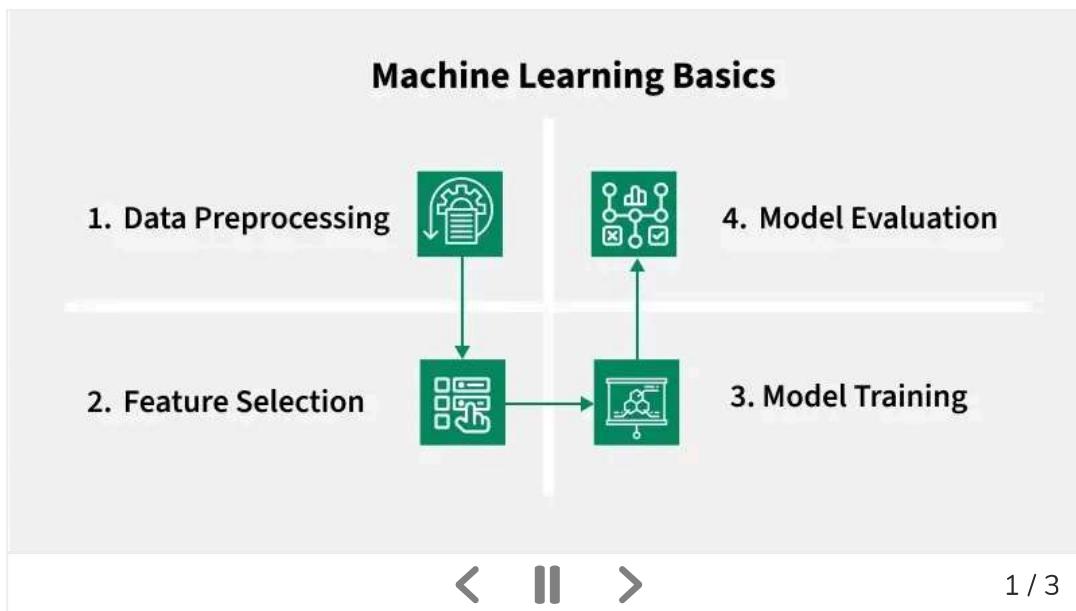
B

[Python for Machine Learning](#) [Machine Learning with R](#) [Machine Learning Algorithms](#) [EDA](#) [Math for ML](#)

# Machine Learning Tutorial

Last Updated : 28 Aug, 2025

**Machine learning** is a branch of Artificial Intelligence that focuses on developing models and algorithms that let computers learn from data without being explicitly programmed for every task. In simple words, ML teaches the systems to think and understand like humans by learning from the data.



Try our ongoing free course [Data Science Skillup](#) with weekly topic coverage, notes, daily quizzes and coding problems.

Machine Learning is mainly divided into three core types: Supervised, Unsupervised and Reinforcement Learning along with two additional types, Semi-Supervised and Self-Supervised Learning.

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

Got It !

& [Privacy Policy](#)

- **Unsupervised Learning**: Finds patterns or groups in unlabeled data, like clustering or dimensionality reduction.
- **Reinforcement Learning**: Learns through trial and error to maximize rewards, ideal for decision-making tasks.

**Note:** The following are not part of the original three core types of ML, but they have become increasingly important in real-world applications, especially in deep learning.

#### **Additional Types:**

- **Self-Supervised Learning**: Self-supervised learning is often considered as a subset of unsupervised learning, but it has grown into its own field due to its success in training large-scale models. It generates its own labels from the data, without any manual labeling.
- **Semi-Supervised Learning**: This approach combines a small amount of labeled data with a large amount of unlabeled data. It's useful when labeling data is expensive or time-consuming.

## Module 1: Machine Learning Pipeline

This section covers preprocessing, exploratory data analysis and model evaluation to prepare data, uncover insights and build reliable models.

### 1. Data Preprocessing

- [ML workflow](#)
- [Data Cleaning](#)
- [Data Preprocessing in Python](#)
- [Feature Scaling](#)
- [Feature Extraction](#)

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#)

## 2. Exploratory Data Analysis

- [Exploratory Data Analysis](#)
- [Exploratory Data Analysis in Python](#)
- [Advance EDA](#)
- [Time Series Data Visualization](#)

## 3. Model Evaluation

- [Regularization in Machine Learning](#)
- [Confusion Matrix](#)
- [Precision, Recall](#) and [F1-Score](#)
- [AUC-ROC Curve](#)
- [Cross-validation](#)
- [Hyperparameter Tuning](#)

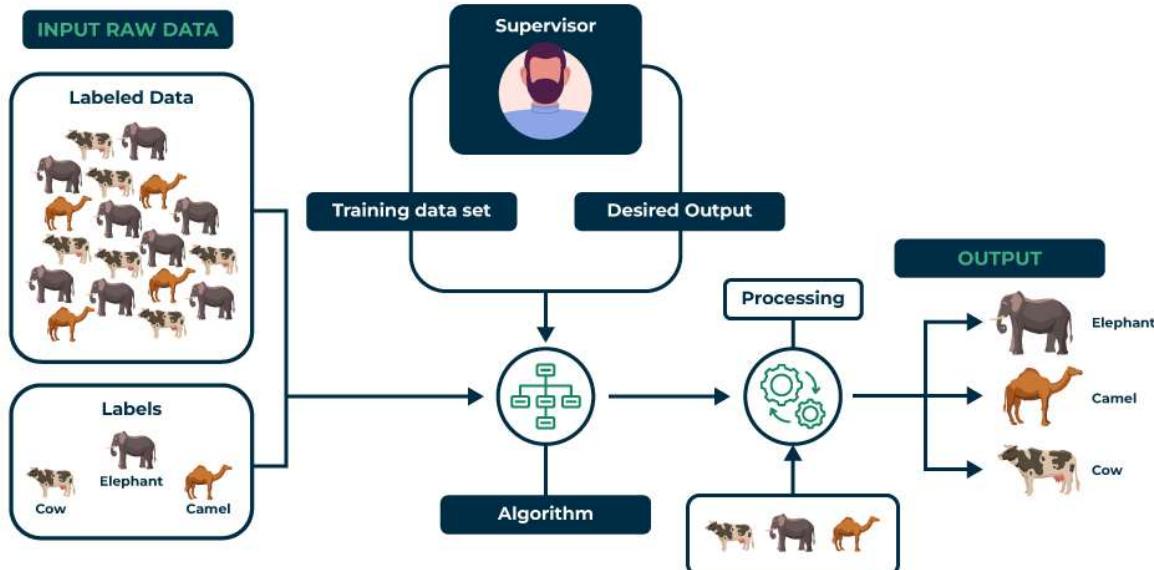
## Module 2: Supervised Learning

Supervised learning algorithms are generally categorized into **two main types:**

- [Classification](#) - where the goal is to predict discrete labels or categories
- [Regression](#) - where the aim is to predict continuous numerical values.

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

# Supervised Learning



## Supervised Learning

There are many algorithms used in supervised learning each suited to different types of problems. Some of the most commonly used supervised learning algorithms are:

### 1. Linear Regression

This is one of the simplest ways to predict numbers using a straight line. It helps find the relationship between input and output.

- [Introduction to Linear Regression](#)
- [Gradient Descent in Linear Regression](#)
- [Multiple Linear Regression](#)

### 2. Logistic Regression

Used when the output is a "yes or no" type answer. It helps in predicting categories like pass/fail or spam/not spam.

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#)

### 3. Decision Trees

A model that makes decisions by asking a series of simple questions, like a flowchart. Easy to understand and use.

- [Decision Tree in Machine Learning](#)
- [Types of Decision tree algorithms](#)
- [Decision Tree - Regression \(Implementation\)](#)
- [Decision tree - Classification \(Implementation\)](#)

### 4. Support Vector Machines (SVM)

A bit more advanced—it tries to draw the best line (or boundary) to separate different categories of data.

- [Understanding SVMs](#)
- [SVM Hyperparameter Tuning - GridSearchCV](#)
- [Non-Linear SVM](#)

### 5. k-Nearest Neighbors (k-NN)

This model looks at the closest data points (neighbors) to make predictions. Super simple and based on similarity.

- [Introduction to KNN](#)
- [Decision Boundaries in K-Nearest Neighbors \(KNN\)](#)

### 6. Naïve Bayes

A quick and smart way to classify things based on probability. It works well for text and spam detection.

- [Introduction to Naive Bayes](#)
- [Gaussian Naive Bayes](#)

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#)

- [Complement Naive Bayes](#)

## 7. Random Forest (Bagging Algorithm)

A powerful model that builds lots of decision trees and combines them for better accuracy and stability.

- [Introduction to Random forest](#)
- [Random Forest Classifier](#)
- [Random Forest Regression](#)
- [Hyperparameter Tuning in Random Forest](#)

## Introduction to Ensemble Learning

[Ensemble learning](#) combines multiple simple models to create a stronger, smarter model. There are mainly two types of ensemble learning:

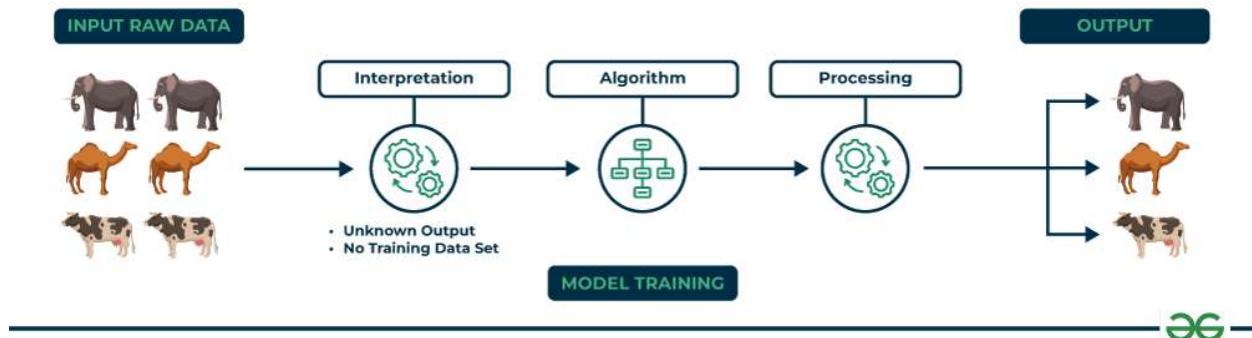
- [Bagging](#) that combines multiple models trained independently.
- [Boosting](#) that builds models sequentially each correcting the errors of the previous one.

## Module 3: Unsupervised learning

Unsupervised learning are again divided into **three main categories** based on their purpose:

- [Clustering](#)
- [Association Rule Mining](#)
- [Dimensionality Reduction](#).

# Unsupervised Learning



*Unsupervised learning*



## 1. Clustering

Clustering algorithms group data points into clusters based on their similarities or differences. Types of clustering algorithms are:

### Centroid-based Methods:

- [K-Means clustering](#)
- [Elbow Method for optimal value of k in KMeans](#)
- [K-Means++ clustering](#)
- [K-Mode clustering](#)
- [Fuzzy C-Means \(FCM\) Clustering](#)

### Distribution-based Methods:

- [Gaussian mixture models](#)
- [Expectation-Maximization Algorithm](#)
- [Dirichlet process mixture models \(DPMMs\)](#).

### Connectivity based methods:

- [Hierarchical clustering](#)
- [Agglomerative Clustering](#)
- [Divisive clustering](#)

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#)

- [DBSCAN \(Density-Based Spatial Clustering of Applications with Noise\).](#)
- [OPTICS \(Ordering Points To Identify the Clustering Structure\).](#)

## 2. Dimensionality Reduction

Dimensionality reduction is used to simplify datasets by reducing the number of features while retaining the most important information.

- [Principal Component Analysis \(PCA\).](#)
- [t-distributed Stochastic Neighbor Embedding \(t-SNE\).](#)
- [Non-negative Matrix Factorization \(NMF\).](#)
- [Independent Component Analysis \(ICA\).](#)
- [Isomap](#)
- [Locally Linear Embedding \(LLE\).](#)

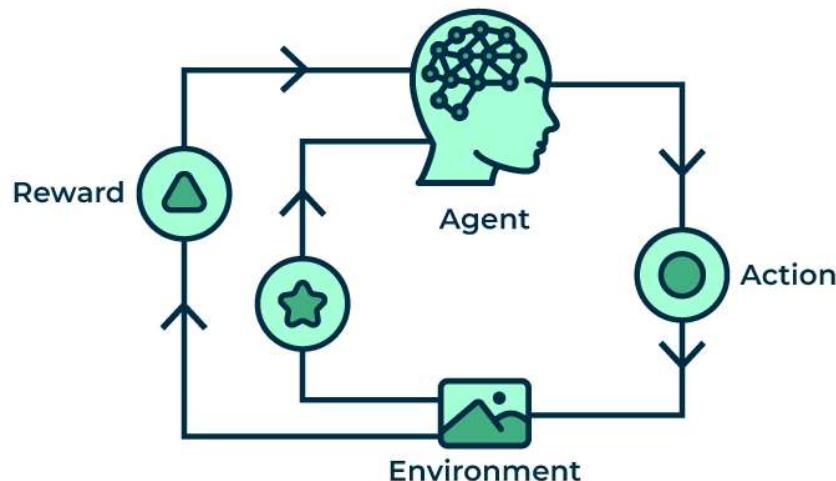
## 3. Association Rule

Find patterns between items in large datasets typically in [market basket analysis.](#)

- [Apriori algorithm](#)
- [Implementing apriori algorithm](#)
- [FP-Growth \(Frequent Pattern-Growth\).](#)
- [ECLAT \(Equivalence Class Clustering and bottom-up Lattice Traversal\).](#)

## Module 4: Reinforcement Learning

Reinforcement learning interacts with environment and learn from them based on rewards.



*Reinforcement Learning*

## 1. Model-Based Methods

These methods use a model of the environment to predict outcomes and help the agent plan actions by simulating potential results.

- [Markov decision processes \(MDPs\)](#)
- [Bellman equation](#)
- [Value iteration algorithm](#)
- [Monte Carlo Tree Search](#)

## 2. Model-Free Methods

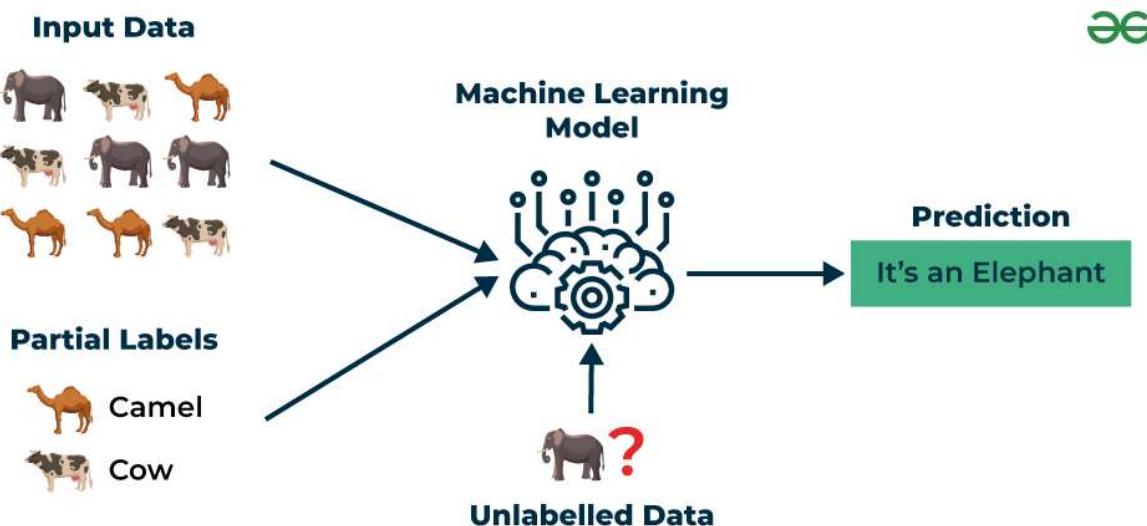
The agent learns directly from experience by interacting with the environment and adjusting its actions based on feedback.

- [Q-Learning](#)
- [SARSA](#)
- [Monte Carlo Methods](#)
- [Reinforce Algorithm](#)
- [Actor-Critic Algorithm](#)

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#).

& [Privacy Policy](#)

It uses a mix of labeled and unlabeled data making it helpful when labeling data is costly or it is very limited.



*Semi Supervised Learning*

- [Semi Supervised Classification](#)
- [Self-Training in Semi-Supervised Learning](#)
- [Few-shot learning in Machine Learning](#)

## Module 6: Forecasting Models

Forecasting models analyze past data to predict future trends, commonly used for time series problems like sales, demand or stock prices.

- [ARIMA \(Auto-Regressive Integrated Moving Average\)](#)
- [SARIMA \(Seasonal ARIMA\)](#)
- [Exponential Smoothing \(Holt-Winters\)](#)

## Module 7: Deployment of ML Models

The trained ML model must be integrated into an application or service to make its predictions accessible.

- [Machine learning deployment](#)
- [Deploy ML Model using Streamlit Library](#)

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#)

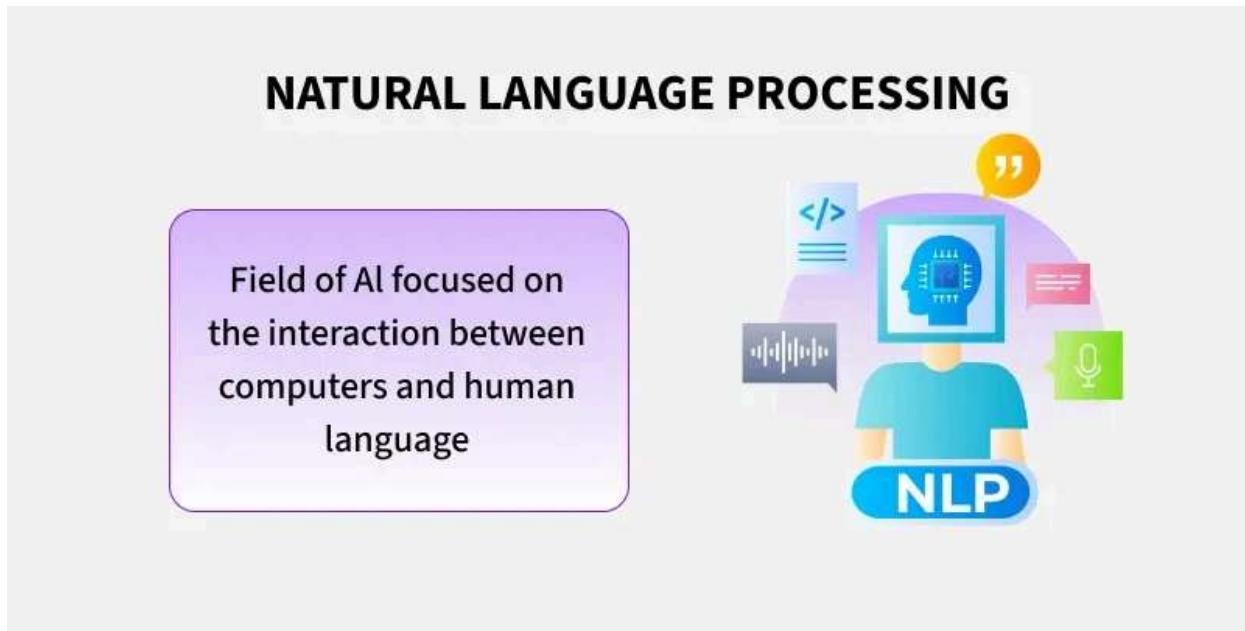
B

[NLP](#) [Data Analysis Tutorial](#) [Python - Data visualization tutorial](#) [NumPy](#) [Pandas](#) [OpenCV](#) [R](#) [M.](#)

# Natural Language Processing (NLP) - Overview

Last Updated : 06 Aug, 2025

Natural Language Processing (NLP) is a field that combines computer science, artificial intelligence and language studies. It helps computers understand, process and create human language in a way that makes sense and is useful. With the growing amount of text data from social media, websites and other sources, NLP is becoming a key tool to gain insights and automate tasks like analyzing text or translating languages.



*Natural Language Processing*

## Table of Content

- [NLP Techniques](#)
- [How Natural Language Processing \(NLP\) Works](#)
- [Technologies related to Natural Language Processing](#)
- [Applications of Natural Language Processing \(NLP\)](#)

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

[Got It !](#)

& [Privacy Policy](#)

NLP is used by many applications that use language, such as text translation, voice recognition, text summarization and chatbots. You may have used some of these applications yourself, such as voice-operated GPS systems, digital assistants, speech-to-text software and customer service bots. NLP also helps businesses improve their efficiency, productivity and performance by simplifying complex tasks that involve language.

## NLP Techniques

NLP encompasses a wide array of techniques that aimed at enabling computers to process and understand human language. These tasks can be categorized into several broad areas, each addressing different aspects of language processing. Here are some of the key NLP techniques:

### 1. Text Processing and Preprocessing

- **Tokenization**: Dividing text into smaller units, such as words or sentences.
- **Stemming and Lemmatization**: Reducing words to their base or root forms.
- **Stopword Removal**: Removing common words (like "and", "the", "is") that may not carry significant meaning.
- **Text Normalization**: Standardizing text, including case normalization, removing punctuation and correcting spelling errors.

### 2. Syntax and Parsing

- **Part-of-Speech (POS) Tagging**: Assigning parts of speech to each word in a sentence (e.g., noun, verb, adjective).
- **Dependency Parsing**: Analyzing the grammatical structure of a

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#)

### 3. Semantic Analysis

- **Named Entity Recognition (NER):** Identifying and classifying entities in text, such as names of people, organizations, locations, dates, etc.
- **Word Sense Disambiguation (WSD):** Determining which meaning of a word is used in a given context.
- **Coreference Resolution:** Identifying when different words refer to the same entity in a text (e.g., "he" refers to "John").

### 4. Information Extraction

- **Entity Extraction:** Identifying specific entities and their relationships within the text.
- **Relation Extraction:** Identifying and categorizing the relationships between entities in a text.

### 5. Text Classification in NLP

- **Sentiment Analysis:** Determining the sentiment or emotional tone expressed in a text (e.g., positive, negative, neutral).
- **Topic Modeling:** Identifying topics or themes within a large collection of documents.
- **Spam Detection:** Classifying text as spam or not spam.

### 6. Language Generation

- **Machine Translation:** Translating text from one language to another.
- **Text Summarization:** Producing a concise summary of a larger text.
- **Text Generation:** Automatically generating coherent and contextually relevant text.

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

- **Text-to-Speech (TTS) Synthesis:** Converting written text into spoken language.

## 8. Question Answering

- **Retrieval-Based QA:** Finding and returning the most relevant text passage in response to a query.
- **Generative QA:** Generating an answer based on the information available in a text corpus.

## 9. Dialogue Systems

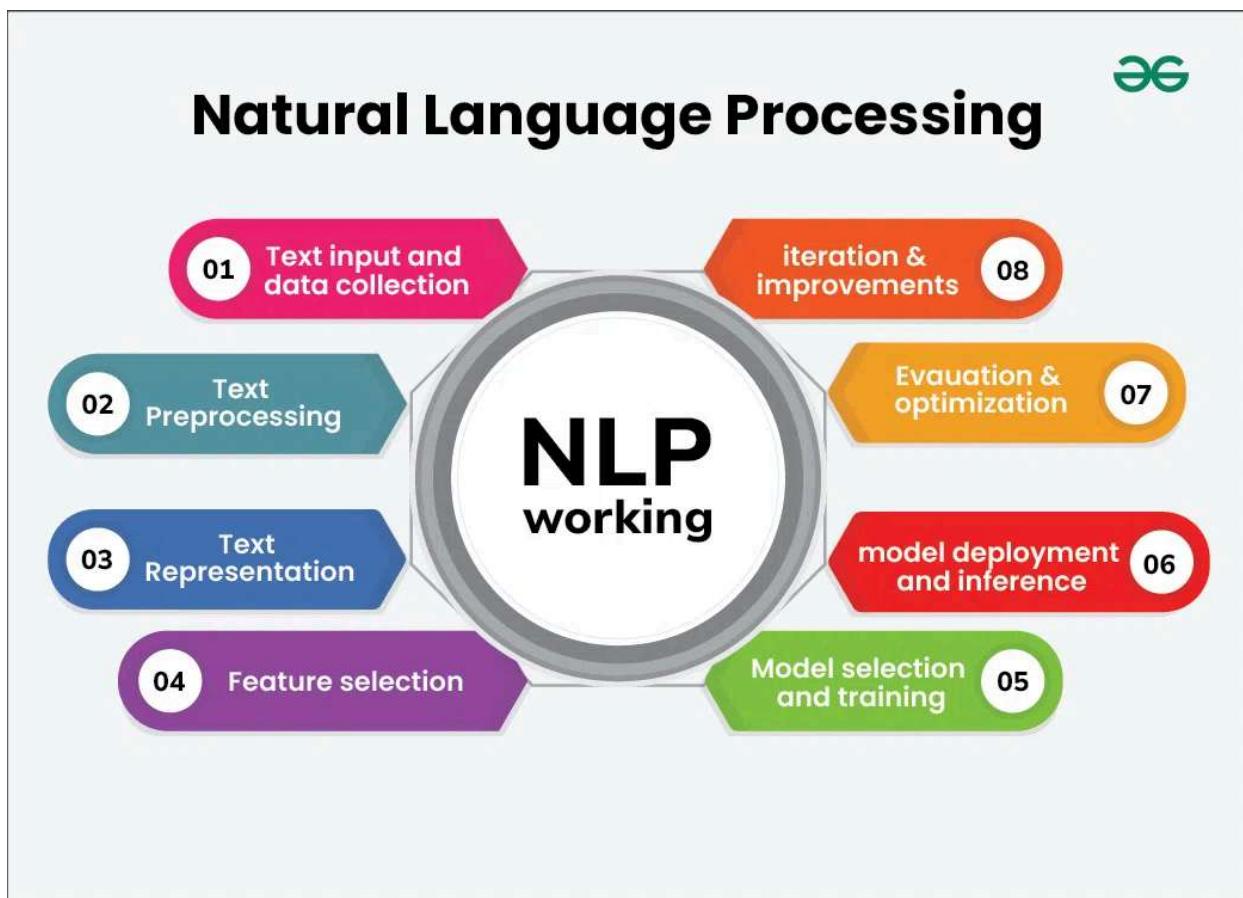
- **Chatbots and Virtual Assistants:** Enabling systems to engage in conversations with users, providing responses and performing tasks based on user input.

## 10. Sentiment and Emotion Analysis in NLP

- **Emotion Detection:** Identifying and categorizing emotions expressed in text.
- **Opinion Mining:** Analyzing opinions or reviews to understand public sentiment toward products, services or topics.

## How Natural Language Processing (NLP) Works

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

*NLP Working*

Working in natural language processing (NLP) typically involves using computational techniques to analyze and understand human language. This can include tasks such as language understanding, language generation and language interaction.

## 1. Text Input and Data Collection

- **Data Collection**: Gathering text data from various sources such as websites, books, social media or proprietary databases.
- **Data Storage**: Storing the collected text data in a structured format, such as a database or a collection of documents.

## 2. Text Preprocessing

Preprocessing is crucial to clean and prepare the raw text data for

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

- **Lowercasing:** Converting all text to lowercase to ensure uniformity.
- **Stopword Removal:** Removing common words that do not contribute significant meaning, such as "and," "the," "is."
- **Punctuation Removal:** Removing punctuation marks.
- **Stemming and Lemmatization:** Reducing words to their base or root forms. Stemming cuts off suffixes, while lemmatization considers the context and converts words to their meaningful base form.
- **Text Normalization:** Standardizing text format, including correcting spelling errors, expanding contractions and handling special characters.

### 3. Text Representation

- **Bag of Words (BoW):** Representing text as a collection of words, ignoring grammar and word order but keeping track of word frequency.
- **Term Frequency-Inverse Document Frequency (TF-IDF):** A statistic that reflects the importance of a word in a document relative to a collection of documents.
- **Word Embeddings:** Using dense vector representations of words where semantically similar words are closer together in the vector space (e.g., Word2Vec, GloVe).

### 4. Feature Extraction

Extracting meaningful features from the text data that can be used for various NLP tasks.

- **N-grams:** Capturing sequences of N words to preserve some context and word order.
- **Syntactic Features:** Using parts of speech tags, syntactic

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#)

## 5. Model Selection and Training

Selecting and training a machine learning or deep learning model to perform specific NLP tasks.

- **Supervised Learning**: Using labeled data to train models like Support Vector Machines (SVM), Random Forests or deep learning models like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs).
- **Unsupervised Learning**: Applying techniques like clustering or topic modeling (e.g., Latent Dirichlet Allocation) on unlabeled data.
- **Pre-trained Models**: Utilizing pre-trained language models such as BERT, GPT or transformer-based models that have been trained on large corpora.

## 6. Model Deployment and Inference

Deploying the trained model and using it to make predictions or extract insights from new text data.

- **Text Classification**: Categorizing text into predefined classes (e.g., spam detection, sentiment analysis).
- **Named Entity Recognition (NER)**: Identifying and classifying entities in the text.
- **Machine Translation**: Translating text from one language to another.
- **Question Answering**: Providing answers to questions based on the context provided by text data.

## 7. Evaluation and Optimization

Evaluating the performance of the NLP algorithm using metrics such as accuracy, precision, recall, F1-score and others.

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).

- **Error Analysis:** Analyzing errors to understand model weaknesses and improve robustness.

## Technologies related to Natural Language Processing

There are a variety of technologies related to natural language processing (NLP) that are used to analyze and understand human language. Some of the most common include:

1. **Machine learning:** NLP relies heavily on [machine learning](#) techniques such as supervised and unsupervised learning, deep learning and reinforcement learning to train models to understand and generate human language.
2. **Natural Language Toolkits (NLTK)** and other libraries: [NLTK](#) is a popular open-source library in Python that provides tools for NLP tasks such as tokenization, stemming and part-of-speech tagging. Other popular libraries include spaCy, OpenNLP and CoreNLP.
3. **Parsers:** Parsers are used to analyze the syntactic structure of sentences, such as dependency parsing and constituency parsing.
4. **Text-to-Speech (TTS) and Speech-to-Text (STT) systems:** TTS systems convert written text into spoken words, while STT systems convert spoken words into written text.
5. **Named Entity Recognition (NER) systems:** NER systems identify and extract named entities such as people, places and organizations from the text.
6. **Sentiment Analysis:** A technique to understand the emotions or opinions expressed in a piece of text, by using various techniques like Lexicon-Based, Machine Learning-Based and Deep Learning-based methods
7. **Machine Translation:** NLP is used for language translation from one language to another through a computer.
8. **Chatbots:** NLP is used for chatbots that communicate with other

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#).

**9. AI Software:** NLP is used in question-answering software for knowledge representation, analytical reasoning as well as information retrieval.

## Applications of Natural Language Processing (NLP)

- **Spam Filters:** One of the most irritating things about email is spam. Gmail uses natural language processing (NLP) to discern which emails are legitimate and which are spam. These spam filters look at the text in all the emails you receive and try to figure out what it means to see if it's spam or not.
- **Algorithmic Trading:** Algorithmic trading is used for predicting stock market conditions. Using NLP, this technology examines news headlines about companies and stocks and attempts to comprehend their meaning in order to determine if you should buy, sell or hold certain stocks.
- **Questions Answering:** NLP can be seen in action by using Google Search or Siri Services. A major use of NLP is to make search engines understand the meaning of what we are asking and generate natural language in return to give us the answers.
- **Summarizing Information:** On the internet, there is a lot of information and a lot of it comes in the form of long documents or articles. NLP is used to decipher the meaning of the data and then provides shorter summaries of the data so that humans can comprehend it more quickly.

## Future Scope

NLP is shaping the future of technology in several ways:

- **Chatbots and Virtual Assistants:** NLP enables chatbots to quickly understand and respond to user queries, providing 24/7 assistance across text or voice interactions.

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#)

& [Privacy Policy](#).

technology more accessible without traditional interfaces.

- **Smarter Search:** NLP is improving search by allowing users to ask questions in natural language, as seen with Google Drive's recent update, making it easier to find documents.
- **Multilingual NLP:** Expanding NLP to support more languages, including regional and minority languages, broadens accessibility.

**Future Enhancements:** NLP is evolving with the use of Deep Neural Networks (DNNs) to make human-machine interactions more natural. Future advancements include improved semantics for word understanding and broader language support, enabling accurate translations and better NLP models for languages not yet supported.

[Comment](#)[More info](#)[Advertise with us](#)

We use cookies to ensure you have the best browsing experience on our website. By using our site, you acknowledge that you have read and understood our [Cookie Policy](#) & [Privacy Policy](#).