MACHINE LEARNING

UNIT - 1

INTRODUCTION TO MACHINE LEARNING

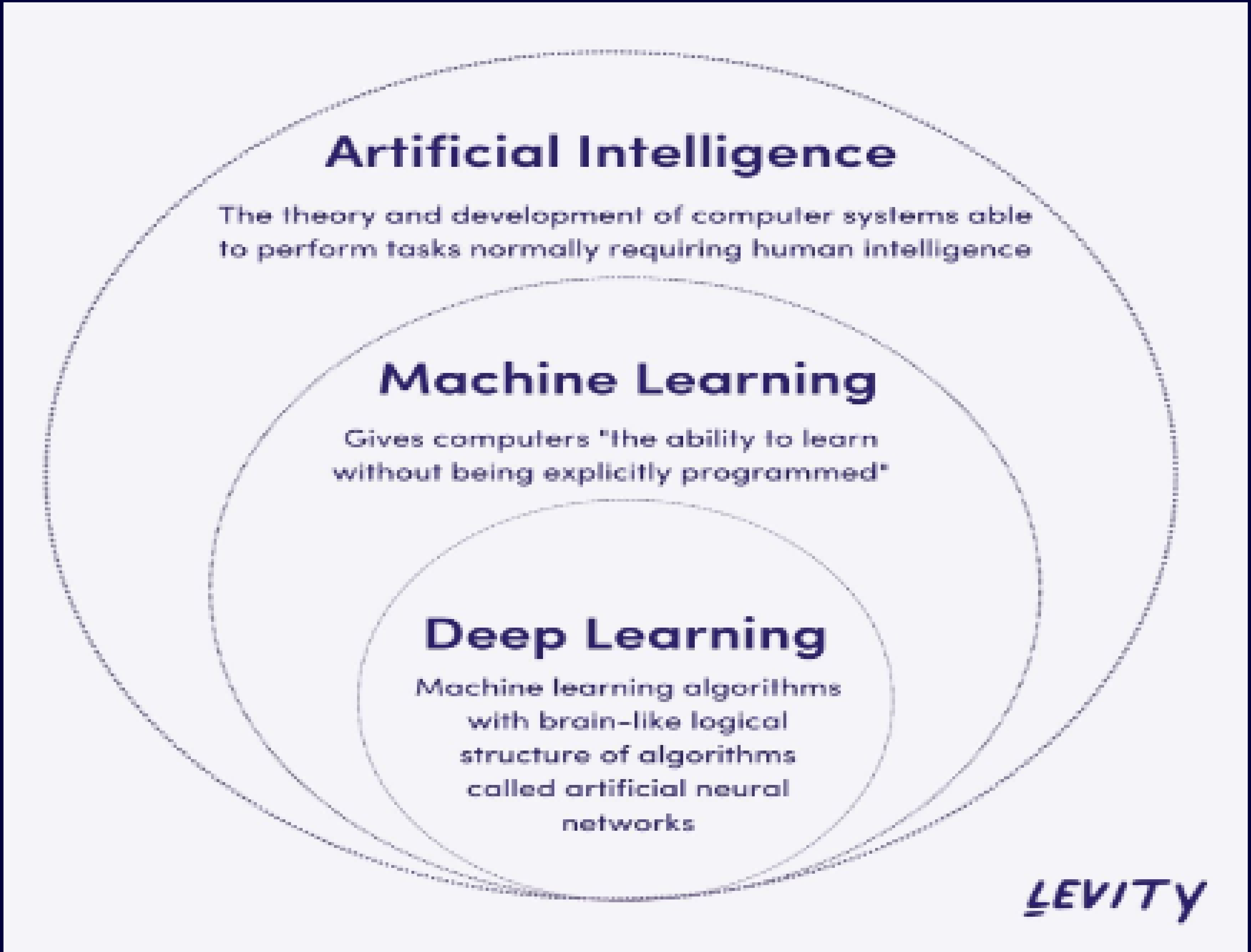
Review of Linear Algebra for machine learning; Introduction and motivation for machine learning; Examples of machine learning applications, Vapnik-Chervonenkis (VC) dimension, Probably Approximately Correct (PAC) learning, Hypothesis spaces, Inductive bias, Generalization, Bias variance trade-off.

INTRODUCTION TO MACHINE LEARNING

WHAT IS MACHINE LEARNING?

WHAT IS MACHINE LEARNING?

- Machine Learning is a subfield of Artificial Intelligence.
- ML is based on the model of brain-cell interaction.
- The term Machine Learning is coined by **Arthur Samuel**, a pioneer in the field of Artificial Intelligence and computer gaming.
- Machine Learning is a field of study that gives computer “the ability to learn without being explicitly programmed.


Artificial Intelligence
The theory and development of computer systems able to perform tasks normally requiring human intelligence
Machine Learning
Gives computers "the ability to learn without being explicitly programmed"
Deep Learning
Machine learning algorithms with brain-like logical structure of algorithms called artificial neural networks
LEVITY

AI vs ML v DL

Artificial Intelligence encompasses any computer program that exhibits human-like intelligence, including learning, reasoning, problem-solving, and decision-making.

AI aspires to create machines that can perform tasks requiring human intellect.

Machine learning (ML) is a specific technique used to achieve AI.

Machine learning algorithms allow computers to learn without being explicitly programmed.

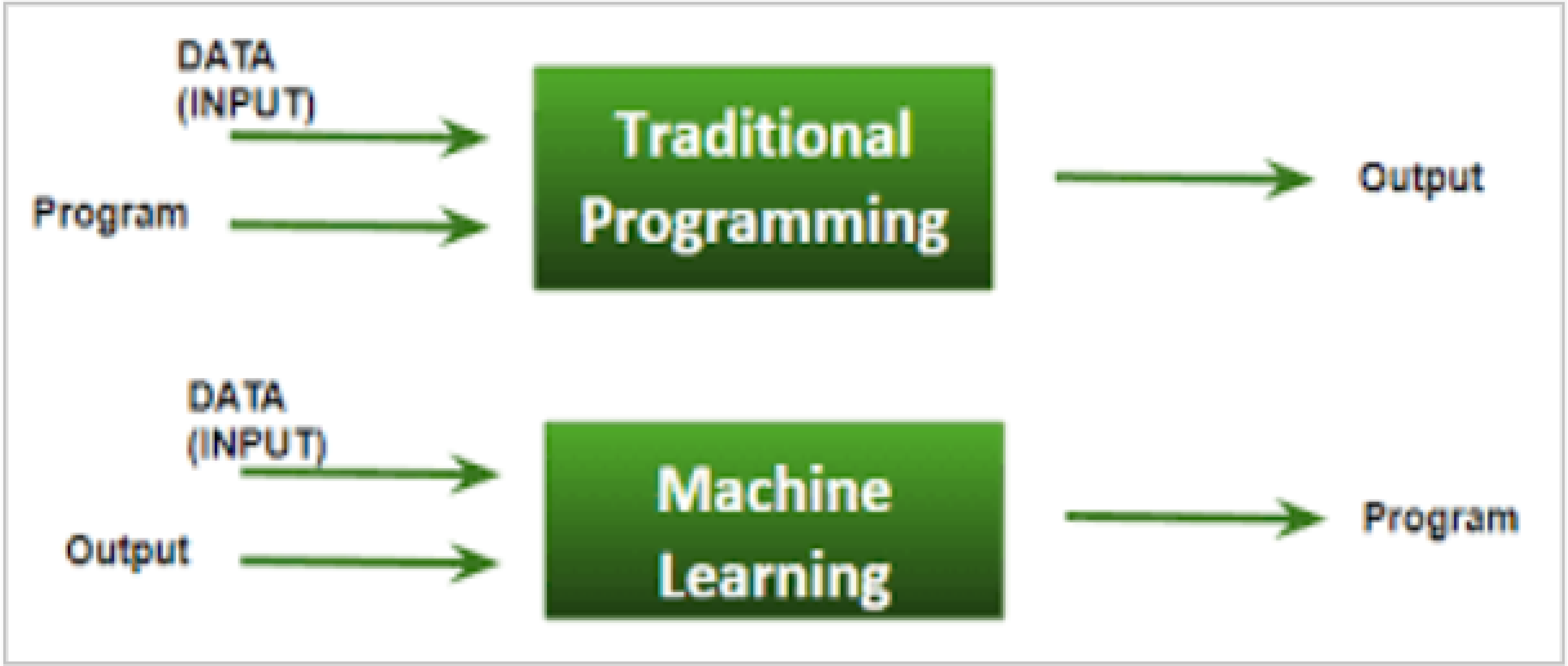
They can improve their performance on a specific task over time by analyzing data.

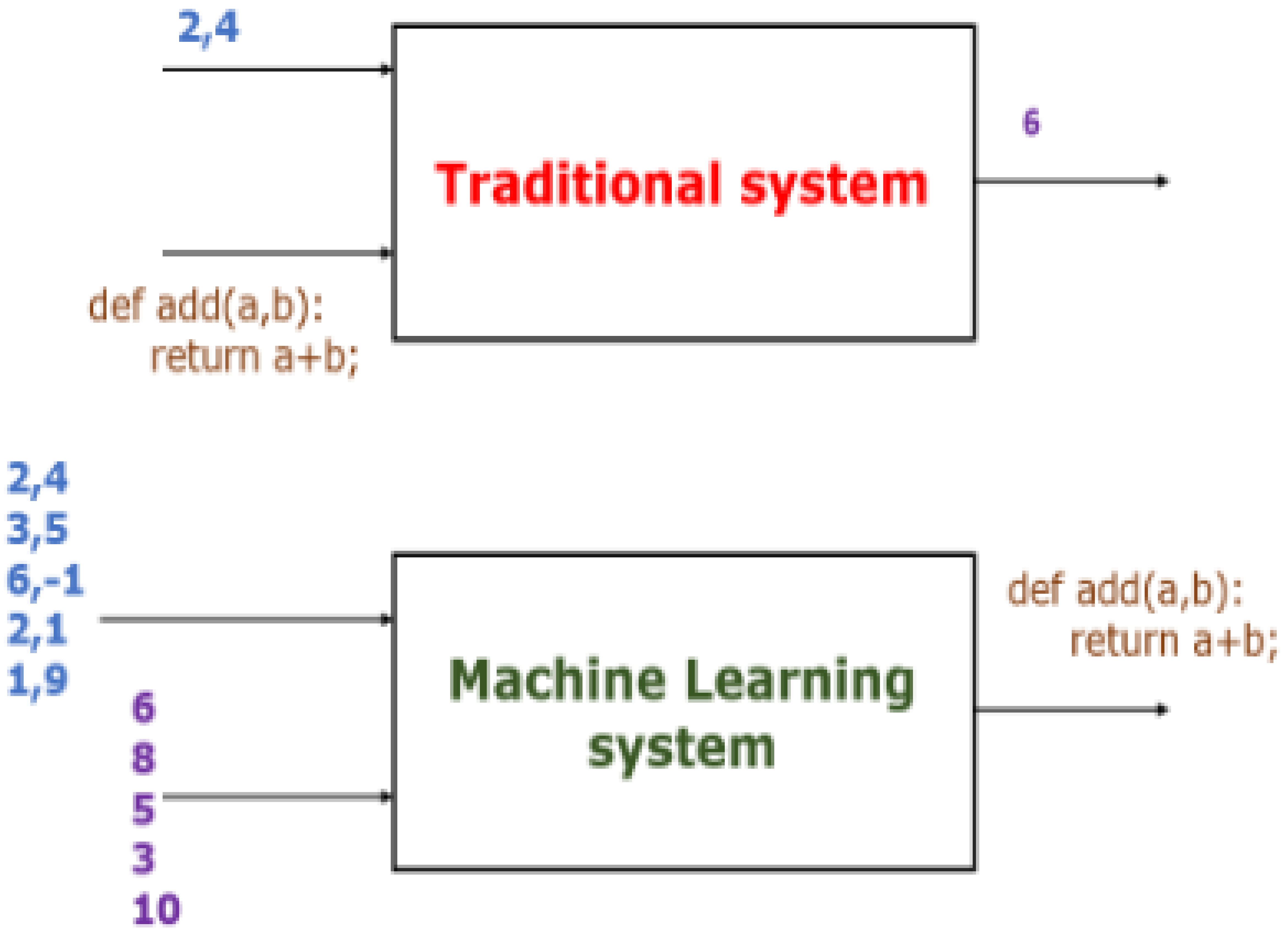
For instance, a recommendation system on a shopping website might use machine learning to analyze your purchase history and suggest items you'd be interested in.

Deep learning (DL) is a subfield of machine learning inspired by the structure and function of the human brain. Deep learning algorithms use artificial neural networks with multiple layers to process data and extract complex patterns.

DL is adopted at handling large amounts of data, such as images, text, or speech. Facial recognition software is an example of a deep learning application.

TRADITIONAL PROGRAMMING VS MACHINE LEARNING





2,4
Traditional system
6
def add(a,b):
return a+b;
2,4
3,5
6,-1
2,1
1,9
6
8
5
3
10
Machine Learning system
def add(a,b):
return a+b;

MACHINE LEARNING

A computer program is said to **learn from Experience E** with respect to some class of **tasks T** and **Performance measure P**, if its performance at tasks in T, as measured by P, improves with experience E.

Well posed learning problem

- An agent solves a problem or tasks T, Performance P and gain some experience E
- If P is measured at T it can improve Experience E (Learning by Experience)

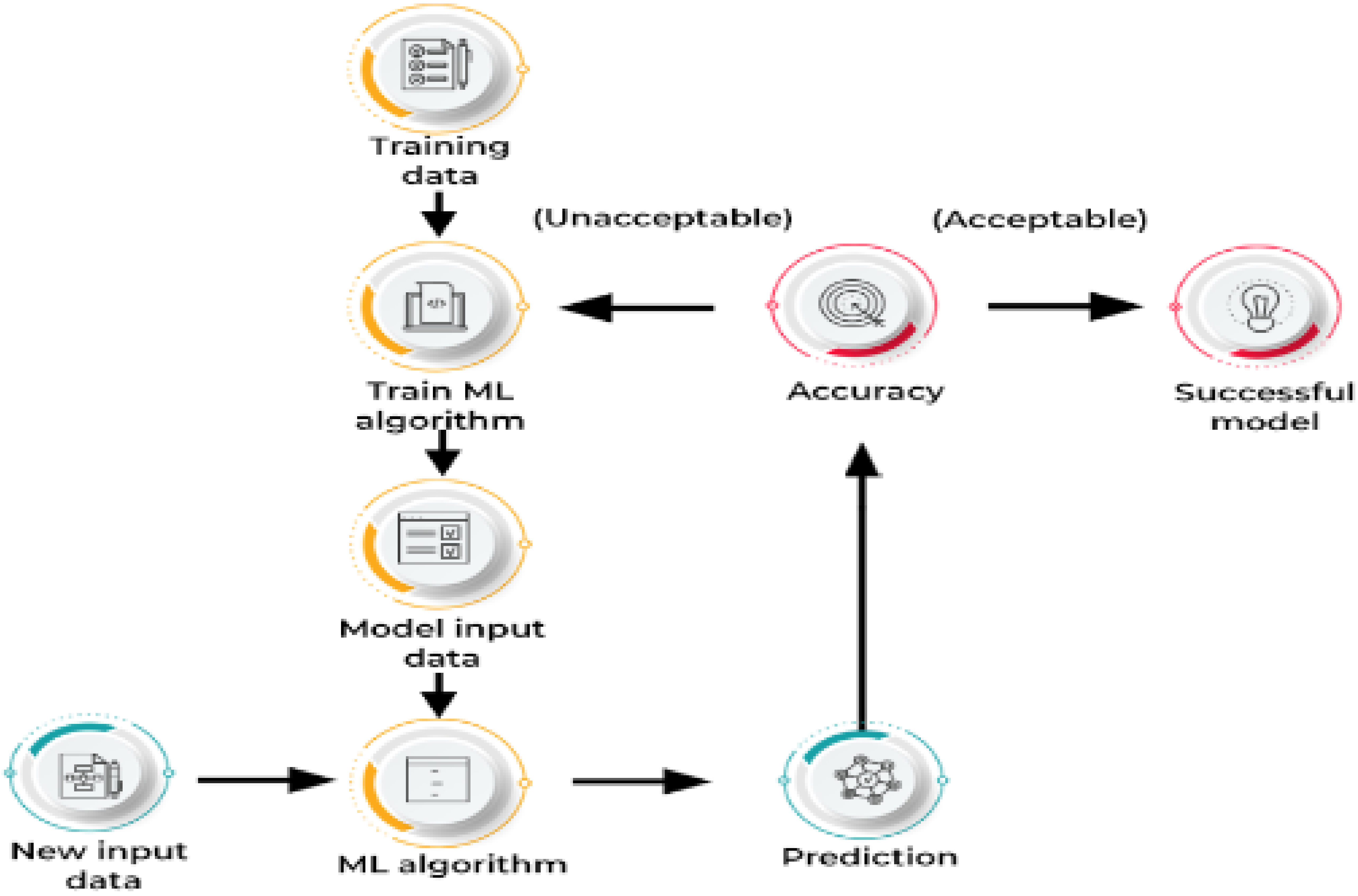
Examples:

1. Handwritten Recognition Problem
2. Robo driving learning problem

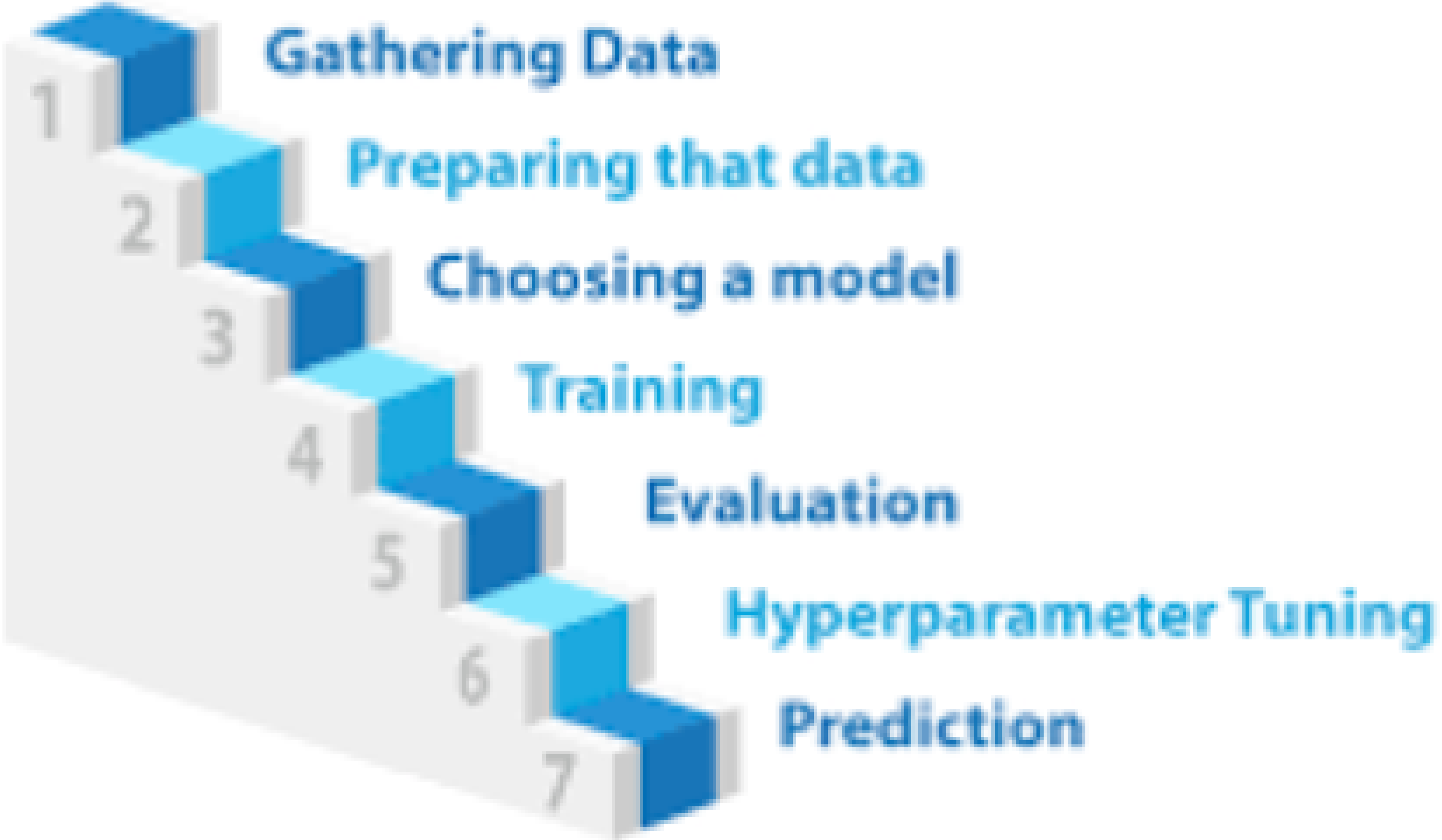
Problem	Task (T)	Performance (P)	Experience (E)
HANDWRITING RECOGNITION LEARNING	Classifying the images and text	Better Classification	A database of homework text

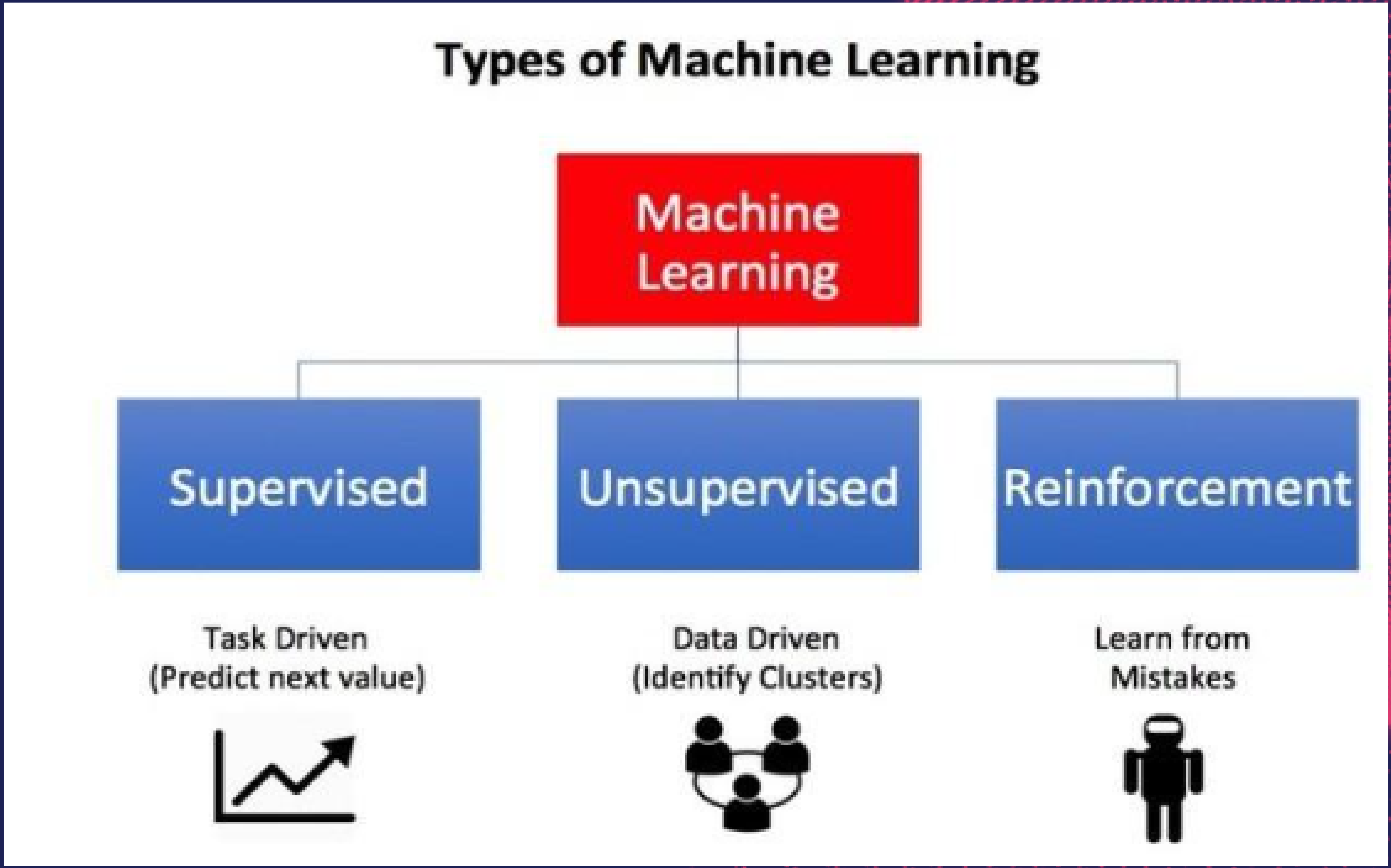
TOOLBOX™

HOW DOES MACHINE LEARNING WORK?

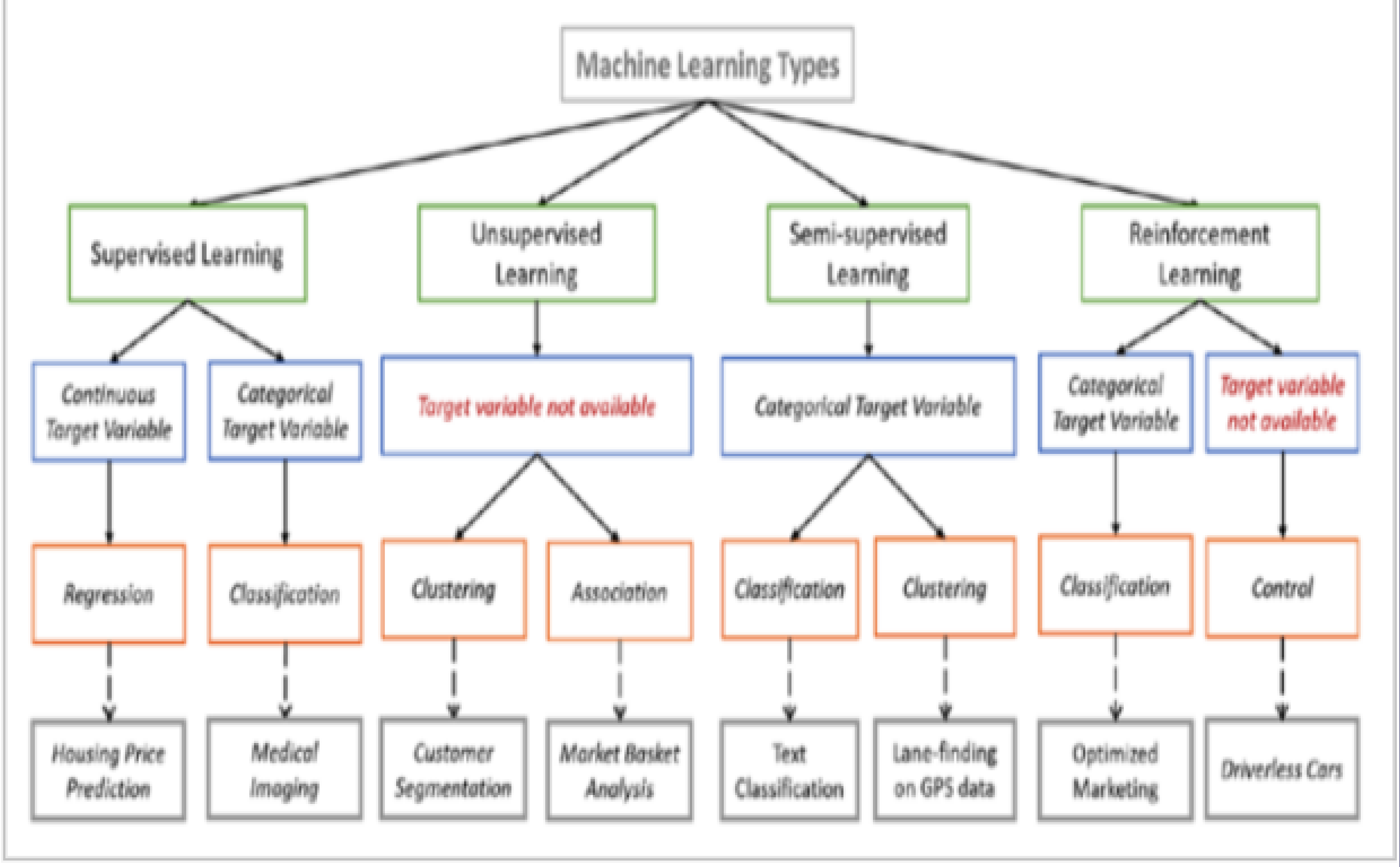


7 steps of Machine Learning





Types of Machine Learning
Machine Learning
Supervised
Unsupervised
Reinforcement
Task Driven
(Predict next value)
Data Driven
(Identify Clusters)
Learn from
Mistakes


Machine Learning Types
Supervised Learning
Unsupervised Learning
Semi-supervised Learning
Reinforcement Learning
Continuous Target Variable
Categorical Target Variable
Target variable not available
Categorical Target Variable
Categorical Target Variable
Target variable not available
Regression
Classification
Clustering
Association
Classification
Clustering
Classification
Control
Housing Price Prediction
Medical Imaging
Customer Segmentation
Market Basket Analysis
Text Classification
Lane-finding on GPS data
Optimized Marketing
Driverless Cars

Linear Algebra

Linear Algebra

- Linear algebra is one of the important branches of mathematics.
- Linear algebra is basically the study of vector spaces, lines, planes, and linear combinations.
- It includes vectors, matrices and linear functions.
- Linear algebra helps in manipulating multivariate data.

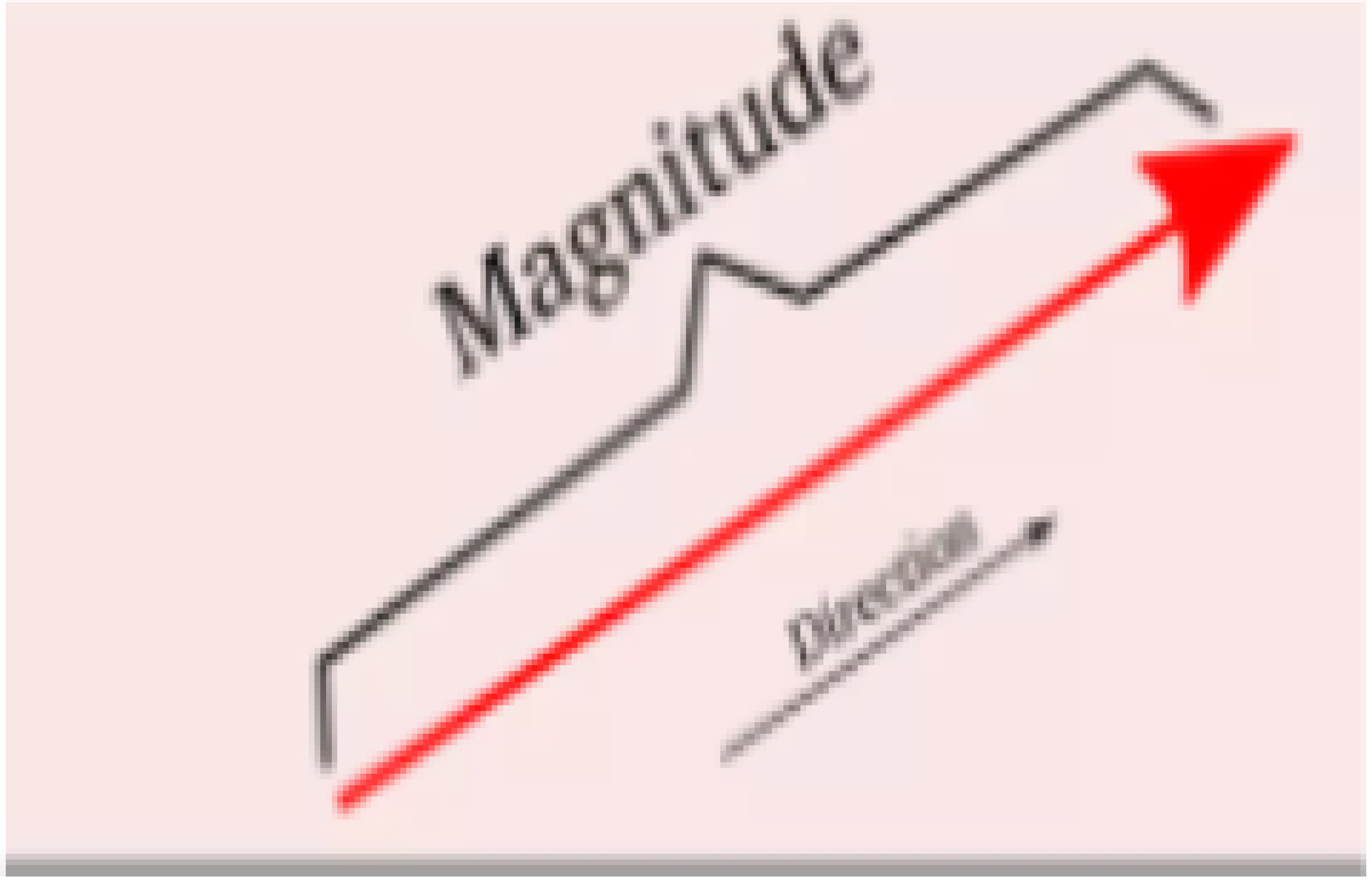
MULTIVARIATE DATA – INVOLVES MORE THAN TWO VARIABLES RESULTING IN A SINGLE OUTCOME.

Outlook	Temp	Humidity	Windy	Play Golf
Rainy	Hot	High	False	No
Rainy	Hot	High	True	No
Overcast	Hot	High	False	Yes
Sunny	Mild	High	False	Yes
Sunny	Cool	Normal	False	Yes
Sunny	Cool	Normal	True	No
Overcast	Cool	Normal	True	Yes
Rainy	Mild	High	False	No
Rainy	Cool	Normal	False	Yes
Sunny	Mild	Normal	False	Yes
Rainy	Mild	Normal	True	Yes
Overcast	Mild	High	True	Yes
Overcast	Hot	Normal	False	Yes
Sunny	Mild	High	True	No

Vector

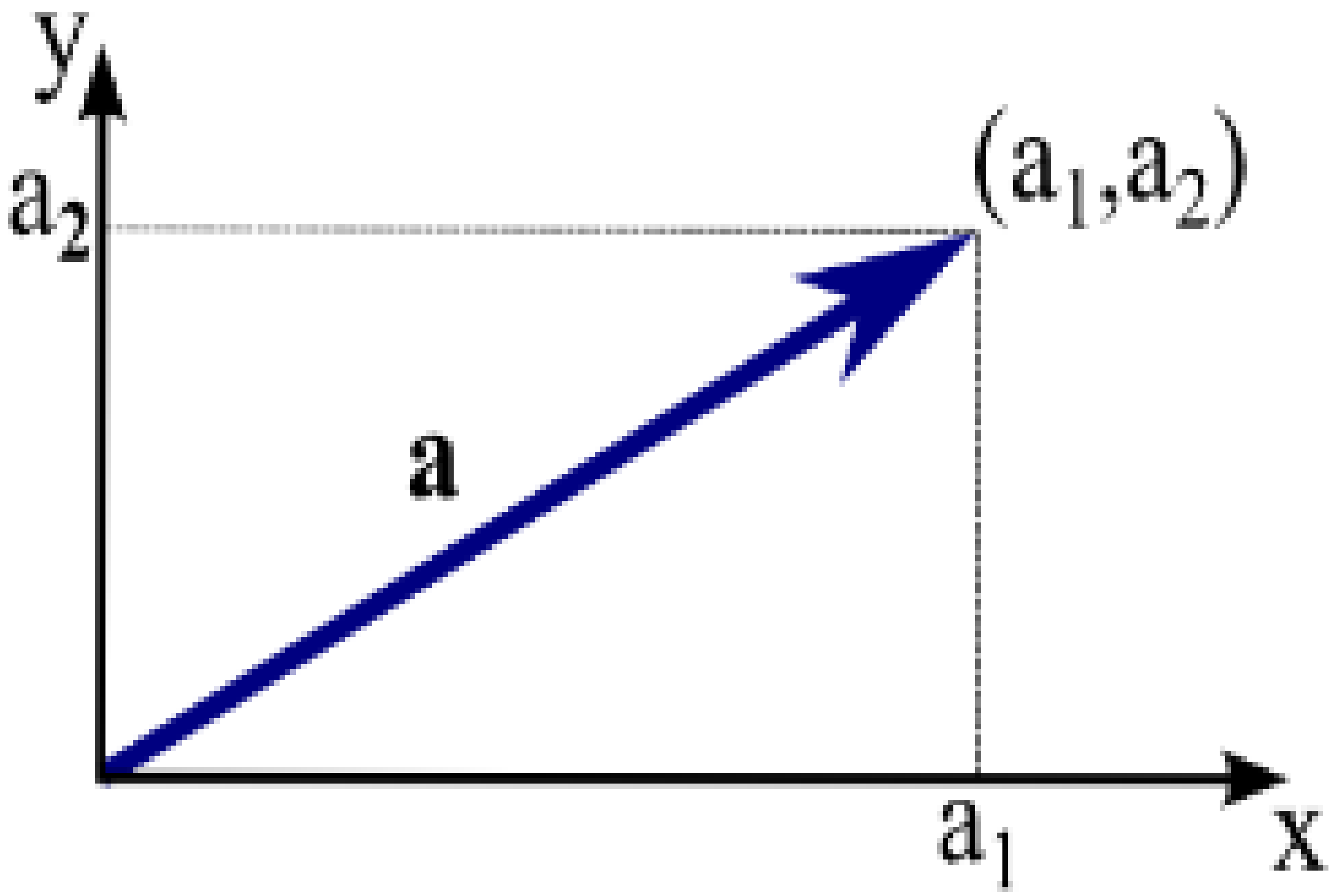
- Vector is an ordered tuple of numbers expressing a magnitude and direction.

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$



Vector Spaces

- A vector is an element of a vector space.



Operations on Vectors

- Given two vectors a and b, we form their sum a+b (Vector Addition)
- Commutative of addition: $a + b = b + a$
- Associativity of addition: $a + (b + c) = (a + b) + c$
- Additive identity: $a + 0 = 0 + a = a$

Properties of a vector

1. If two vectors have same size, they can be added element wise.

$$(B.1) \quad z_i = x_i + y_i, i = 1, \ldots, n$$

2. We can multiply a vector by a constant by multiplying all its elements by that constant.

$$(B.2) \quad y_i = cx_i, i = 1, \ldots, n$$

3. The norm ||x|| represents the length (of the arrow from the origin to point x). It measure the magnitude if the vector.

$$(B.3) \quad \|\boldsymbol{x}\| = \sqrt{\sum_{i-1}^{n} x_i^2}$$

1. If x and y are of the same size, their inner (or dot or scalar) product is

$$\text{(B.4)} \qquad \langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i-1}^{m} x_i y_i$$

2. To obtain θ , which is the angle between (the arrows of) x and y. If ⟨x, y⟩ = 0, we say that x and y are orthogonal, that is, that they are perpendicular

$$\text{(B.5)} \qquad \langle \boldsymbol{x}, \boldsymbol{y} \rangle = \|\boldsymbol{x}\| \cdot \|\boldsymbol{y}\| \cdot \cos\theta$$

Similarity of Vectors

- There are different ways in which we can calculate the similarity between two vectors (of the same size), which are related, each finding its use in its appropriate setting:
- **Inner product**, namely, $\langle x, y \rangle$.
- **Cosine distance**, the cosine of the angle between the two vectors.

$$\cos\theta = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} \tag{B.9}$$

- Euclidean distance

$$\begin{aligned} \|x - y\| &= \sqrt{(x-y)^T(x-y)} = \sqrt{x^Tx + y^Ty - 2x^Ty} \\ &= \sqrt{\|x\|^2 + \|y\|^2 - 2\langle x, y \rangle} \end{aligned} \tag{B.10}$$

Matrix

- In face recognition each input is a face image, which is **two-dimensional**
- A color image is represented by three such matrices for **red, green, and blue**.
- Higher-dimensional matrices are sometimes called **tensors**.

$$\mathbf{X} = \begin{bmatrix} & & x_{1j} & & \\ & & \vdots & & \\ x_{i1} & \cdots & x_{ij} & \cdots & x_{in} \\ & & \vdots & & \\ & & x_{mj} & & \end{bmatrix}$$

- A matrix is a vector of vectors
- Elements of a matrix are vectors
- m x n matrix have ‘m’ rows and ‘n’ columns

Properties of a matrix

- If two matrices have the same size, they can be added element-wise.
- We can multiply a vector (or a matrix) by a constant by multiplying all its elements by that constant.
- The transpose of a matrix swaps its rows and columns: If X is m × n, XT is n × m

$$X^T_{ji} = X_{ij} \tag{B.6}$$

- Transpose of a matrix is obtained by flipping the rows and columns.

Square Matrices

- If a matrix, **A** has equal number of rows and columns, it is a square matrix.

$$\textit{Square Matrix } M = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}$$

- A square matrix is diagonal if all of its off-diagonals are zero.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

- **Upper Triangular Square matrix** - if all the elements below the diagonal are zero.

$$\mathbf{B}\begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 6 \\ 0 & 0 & 9 \end{bmatrix}$$

- **Lower Triangular Square matrix** - if all the elements above the diagonal are zero.

$$\begin{bmatrix} 1 & 0 & 0 \\ 3 & 3 & 0 \\ 1 & -2 & 0 \end{bmatrix}$$

Identity Matrix

- **I** is the **identity matrix**, if it is diagonal matrix and all its diagonals are 1.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix};$$

- **Zero matrix** - If all the elements of the matrix are 0.

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Symmetric

- A square matrix $\mathbf{A}$ is *symmetric* if $\mathbf{A}^T = \mathbf{A}$

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 8 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 8 \end{bmatrix}^{\mathrm{T}}$$

Symmetric matrix

Inverse Matrix

- If A is a matrix and its inverse is given by $\mathbf{A}^{-1}$

$$A^{-1} = \frac{1}{|A|} \cdot \mathrm{Adj}\ A$$

- Inverse of a matrix is obtained by multiplying the reciprocal of the determinant of a matrix and adjoint of a matrix.

Adjoint of a matrix

- Adjoint of a matrix can be calculated by interchanging a11 and a22 and by changing signs of a12 and a_{21}.

$$\text{adj } A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

Change sign Interchange

Determinant of a matrix

- Determinant of a matrix is calculated by

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

$$|A| = a_{11}a_{22} - a_{21}a_{12}$$

Inverse Matrix

- If A is a matrix and its inverse is given by $\mathbf{A}^{-1}$ such that

$$\text{(B.12)} \quad AA^{-1} = A^{-1}A = I$$

$$A = \begin{bmatrix} 4 & 3 \\ 3 & 2 \end{bmatrix} \qquad A^{-1} = \begin{bmatrix} -2 & 3 \\ 3 & -4 \end{bmatrix}$$

$$A\,A^{-1} = \begin{bmatrix} 4 & 3 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} -2 & 3 \\ 3 & -4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \checkmark$$

the identity matrix confirms this inverse relationship

Trace of a matrix

- Trace and Determinant both are used to summarize the “size” of a matrix by a single value, but they are defined differently.
- The trace of a square matrix is simply the sum of its diagonal elements.

$$\text{(B.14)} \quad \mathrm{tr}(A) = \sum_i a_{ii}$$

$$\begin{bmatrix} 2 & 6 & 7 \\ 6 & -6 & 8 \\ 0 & 2 & 4 \end{bmatrix}$$

Trace = 2 + (-6) + 4
= 0

Eigen Value & Eigen Vector

Eigenvalues & Eigenvectors

- Eigenvalues are the special set of scalars associated with the system of linear equations.
- The eigenvector is a vector that is associated with a set of linear equations.
- Find Eigenvalue and eigenvector of the matrix,
- $$A = \begin{bmatrix} 5 & 4 \\ 1 & 2 \end{bmatrix}$$

Spectral Decomposition

Spectral Decomposition

- This method decomposes a square matrix ‘A’ into a product of three matrices.
- A = P D P^T
- P is a n x n matrix, whose elements are eigenvectors of a matrix A in terms of decreasing values.
- D is a diagonal matrix, whose elements are corresponding eigenvalues.
- Find the spectral decomposition of a matrix,

$$A = \begin{bmatrix} -3 & 5 \\ 4 & -2 \end{bmatrix}$$

Singular Value Decomposition