



INTRODUCTION TO AI AND APPLICATIONS

BETC105 / 205

Module 3

BOS (CS/IS)
VTU Belagavi

Prepared by

Dr. Likewin Thomas,

BOS Members (CS/IS) - VTU Belagavi

Prof. & Head, Dept of AIML,

PESITM Shivamogg

Dr. Pavan Kumar M P,

BOS Members (CS/IS) - VTU Belagavi

Assoc. Prof., Dept of ISE,

JNNCE Shivamogg

Dr. Demian Antony D'Mello,

BOS Chairman (CS/IS) - VTU Belagavi

Vice Principal and Dean - Academics

Canara Engineering College





Outline

Module 1: Introduction to Artificial Intelligence

1. Introduction to Artificial Intelligence

1. Definition of Artificial Intelligence
2. How Does AI Work?
3. Advantages and Disadvantages of Artificial Intelligence
4. History of Artificial Intelligence
5. Types of Artificial Intelligence:
 - a. Weak AI vs. Strong AI
 - b. Reactive Machines
 - c. Limited Memory
 - d. Theory of Mind
 - e. Self-Awareness
6. Is Artificial Intelligence the Same as Augmented Intelligence and Cognitive Computing?
7. Introduction to Machine Learning and Deep Learning

2. Machine Intelligence

1. Defining Intelligence
2. Components of Intelligence
3. Differences between Human and Machine Intelligence
4. Agents and Environment in AI
5. Search Algorithms:
 - a. Uninformed Search Algorithms
 - b. Informed Search Algorithms:
 - i. Pure Heuristic Search
 - ii. Best-First Search Algorithm (Greedy Search)





Outline

Module 2: Introduction to Prompt Engineering

1. Introduction to Prompt Engineering

1. Overview of Prompt Engineering
2. The Evolution of Prompt Engineering
3. Types of Prompts
4. How Does Prompt Engineering Work?
5. The Role of Prompt Engineering in Communication
6. The Advantages of Prompt Engineering
7. The Future of Large Language Models (LLMs) in Communication

2. Prompt Engineering Techniques for ChatGPT

- Introduction to Prompt Engineering Techniques
- Instructions Prompt Technique
- Zero, One, and Few Shot Prompting
- Self-Consistency Prompt

3. Prompts for Creative Thinking

- Introduction to Creative Thinking with Prompts
- Unlocking Imagination and Innovation

4. Prompts for Effective Writing

- Introduction to Writing with Prompts
- Igniting the Writing Process with Prompts

Completed





Outline

Module 3: Machine Learning

Reema Thareja, Artificial Intelligence: Beyond Classical AI, Pearson Education, 2023.

1. Machine Learning in AI

- Overview of Machine Learning Techniques
- Introduction to Machine Learning Models

2. Regression Analysis in Machine Learning

- Basics of Regression
- Linear and Non-Linear Regression Techniques

3. Classification Techniques

- Overview of Classification Algorithms
- Naïve Bayes Classification
- Support Vector Machine (SVM)

4. Clustering Techniques

- Introduction to Clustering
- Types of Clustering Algorithms

5. Neural Networks

- Basics of Neural Networks
- Types and Applications of Neural Networks





INTRODUCTION TO AI AND APPLICATIONS

BETC105 / 205

Module 3

BOS (CS/IS)
VTU Belagavi

Prepared by

Dr. Likewin Thomas,

BOS Members (CS/IS) - VTU Belagavi

Prof. & Head, Dept of AIML,

PESITM Shivamogg

Dr. Pavan Kumar M P,

BOS Members (CS/IS) - VTU Belagavi

Assoc. Prof., Dept of ISE,

JNNCE Shivamogg

Dr. Demian Antony D'Mello,

BOS Chairman (CS/IS) - VTU Belagavi

Vice Principal and Dean - Academics

Canara Engineering College





Module 3: Machine Learning in AI

Overview of Machine Learning Techniques

- Artificial Intelligence (AI) is a field that focuses on creating machines that can perform tasks typically requiring human intelligence, such as learning, reasoning, and problem-solving.
- To understand how AI works, it's important to break it down into its key techniques and sub-domains.
- **How AI Works:**
 - AI combines large datasets, fast processing power, and intelligent algorithms to allow systems to learn and deduce patterns from data. By processing and analyzing vast amounts of information, AI systems improve their performance without the need for explicit programming.

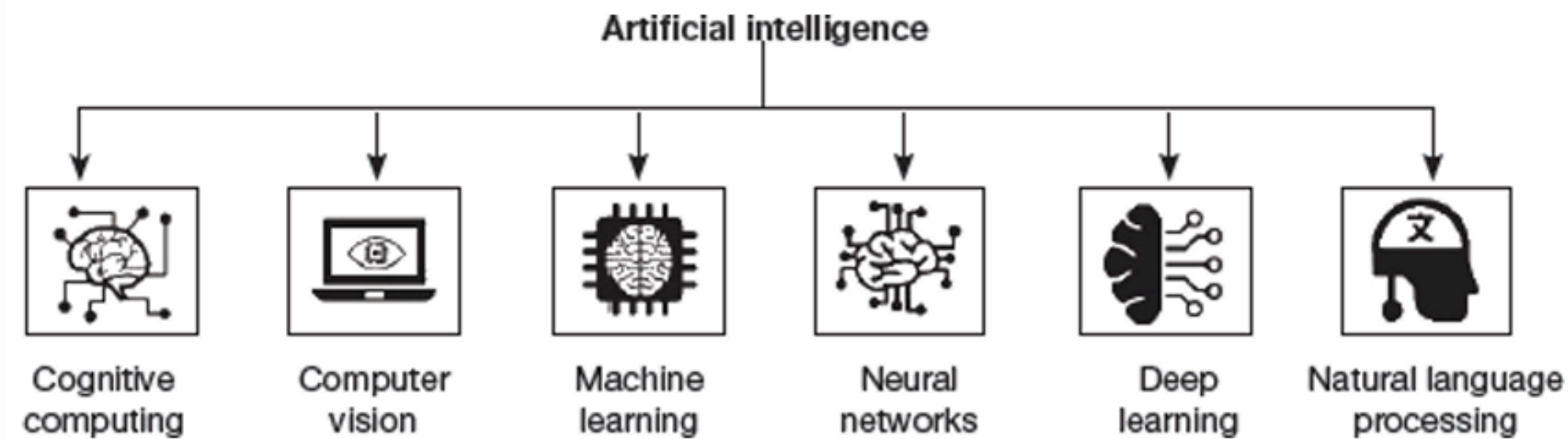




Module 3: Machine Learning in AI

Overview of Machine Learning Techniques

- **AI as a Process of Reverse-Engineering Human Traits**
 - AI tries to mimic human traits like reasoning, decision-making, and pattern recognition.
 - The goal is to make machines that can learn, adapt, and solve problems autonomously by recognizing patterns in the data they process.



Sub-Domains of AI





Module 3: Machine Learning in AI

1. Techniques in AI

1. Neural Networks:
2. Machine learning:
3. Deep learning
4. Natural language processing
5. Computer vision
6. Cognitive Computing

Additional technologies that enable and support AI include the following

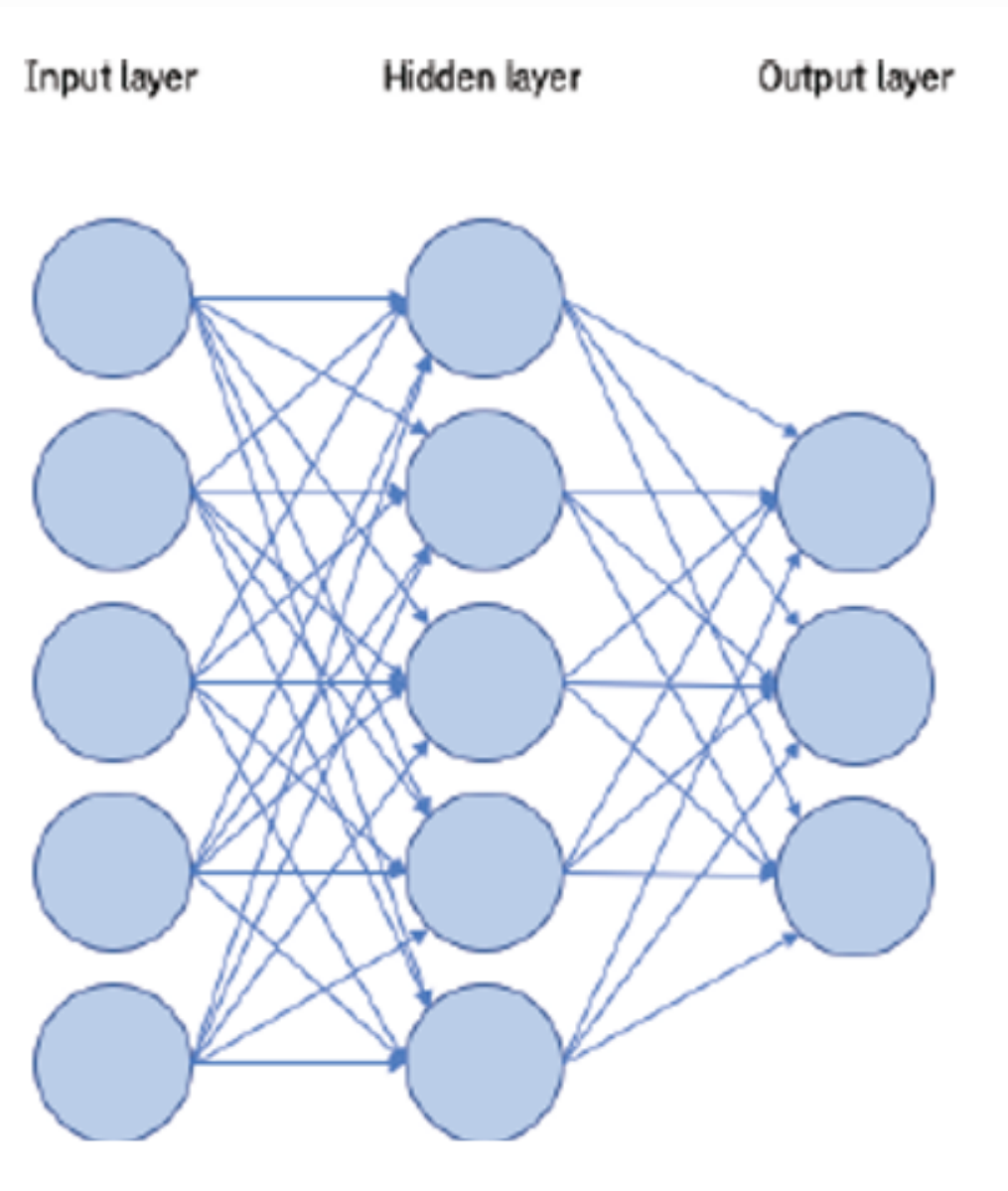
1. Graphical processing units
2. Internet of Things





Module 3: Machine Learning in AI

1. Techniques in AI - 1. Neural Networks:



1. A neural network is a type of AI that works like the human brain. It helps machines learn patterns from data and make decisions or predictions.

2. Think of it as a group of connected “neurons” (nodes) that pass information to each other to understand complex data.

3. **Nodes (Neurons):**

a. Each node is like a tiny decision-maker.

b. It takes input, processes it, and passes output to the next node.

c. Example: Imagine a neuron receives information about the weather (sunny, rainy) and helps decide whether to carry an umbrella.

4. **Learning Process:**

a. The network looks at the data many times to find patterns.

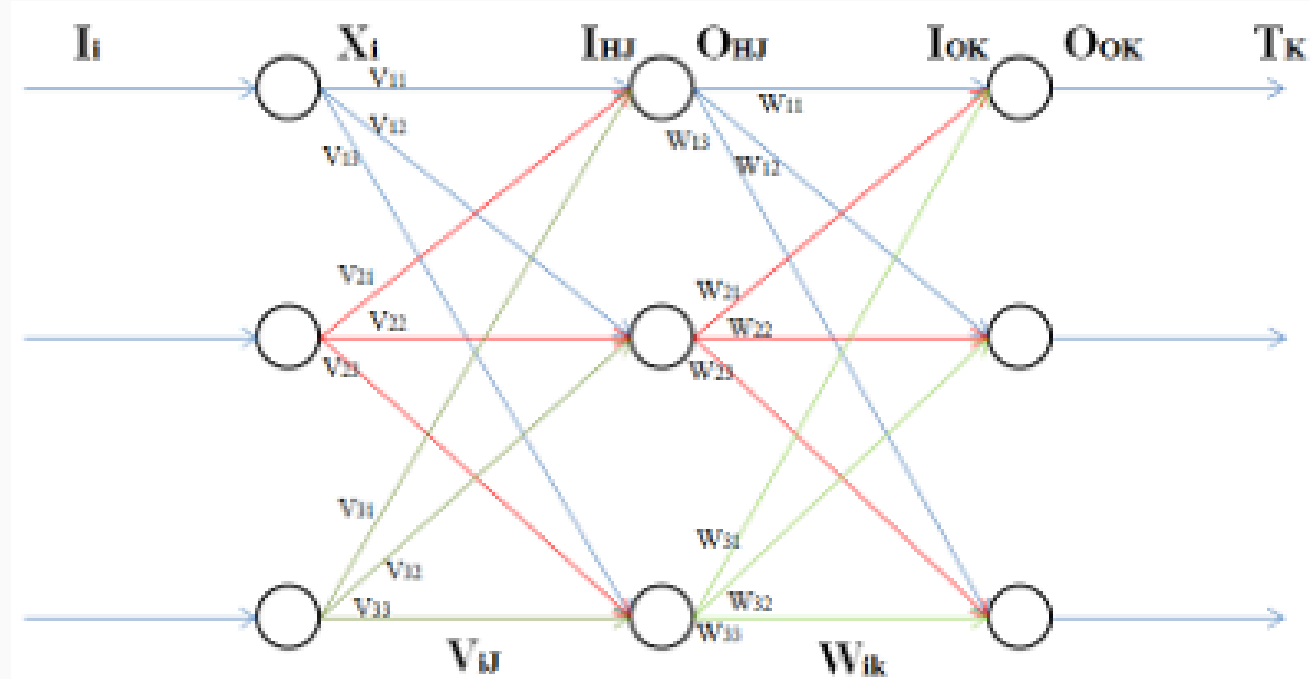
b. This is like a student practicing multiple times to understand a concept.





Module 3: Machine Learning in AI

1. Techniques in AI - 1. Neural Networks:



$E = \text{Error} \rightarrow E_k = \text{Error at } k^{\text{th}} \text{ Neuron}$

$T_k = \text{Target output at } k^{\text{th}} \text{ Neuron}$

$O_{ok} = \text{Computed output of } k^{\text{th}} \text{ Neuron}$

$I_{ok} = \text{Input to } k^{\text{th}} \text{ Neuron}$

$O_{Hi} = \text{Computed output of } i^{\text{th}} \text{ Hidden Neuron}$

$I_{Hj} = \text{Input to } j^{\text{th}} \text{ Hidden Neuron}$

$x_i = \text{input } x_1, x_2, \dots, x_n$

1. Layers of a Neural Network: A neural network has three main layers:

a. Input Layer:

- i. Where the data enters the network.
- ii. Example: If you want to predict house prices, inputs could be size, location, number of rooms.

b. Hidden Layer(s):

- i. Where the AI processes data using algorithms.
- ii. Weights and biases are applied to inputs to determine importance.
- iii. Example: The hidden layer analyzes how size, location, and rooms combine to affect house price.

c. Output Layer:

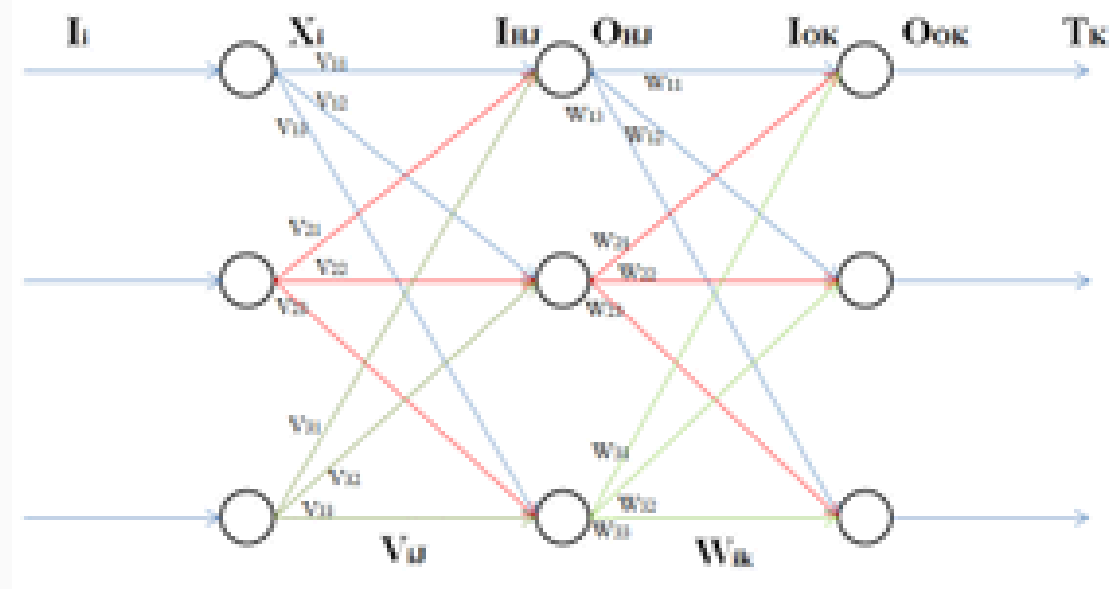
- i. Gives the final result or prediction.
- ii. Example: The predicted house price based on the inputs.





Module 3: Machine Learning in AI

1. Techniques in AI - 1. Neural Networks:



- (I) Input Layer
 $\{I\}_I$ which is $l \times 1$ vector
- (II) Output of Input Layer
 $\{O\}_I$ which is $l \times 1$ vector
- (III) Input to Hidden Layer
 $\{I\}_H = [V]^T \{O\}_I$ where
 $[V]^T = m \times l$ and $\{O\}_I = l \times 1$
 $\therefore \{I\}_H = m \times 1$

- (I) Output of Hidden Layer, using sigmoid fun
 $\{O\}_H = \frac{1}{1+e^{-\{I_H\}}}$
- (II) Input to Outer Layer
 $\{I\}_O = [W]^T \{O\}_H$ where
 $[W]^T = n \times m$ and $\{O\}_H = m \times 1$
 $\therefore \{I\}_O = n \times 1$
- (III) Output of Outer Layer
 $\{O\}_O = \frac{1}{1+e^{-\{I_O\}}}$
- (IV) Mean Squared Error
for the k^{th} output neuron, E_k is given
 $E_k = \frac{1}{2} (T_k - \{O_o\})^2$
where
 T_k is target output of k^{th} output neuron
and $\{O_o\}$ is computed output

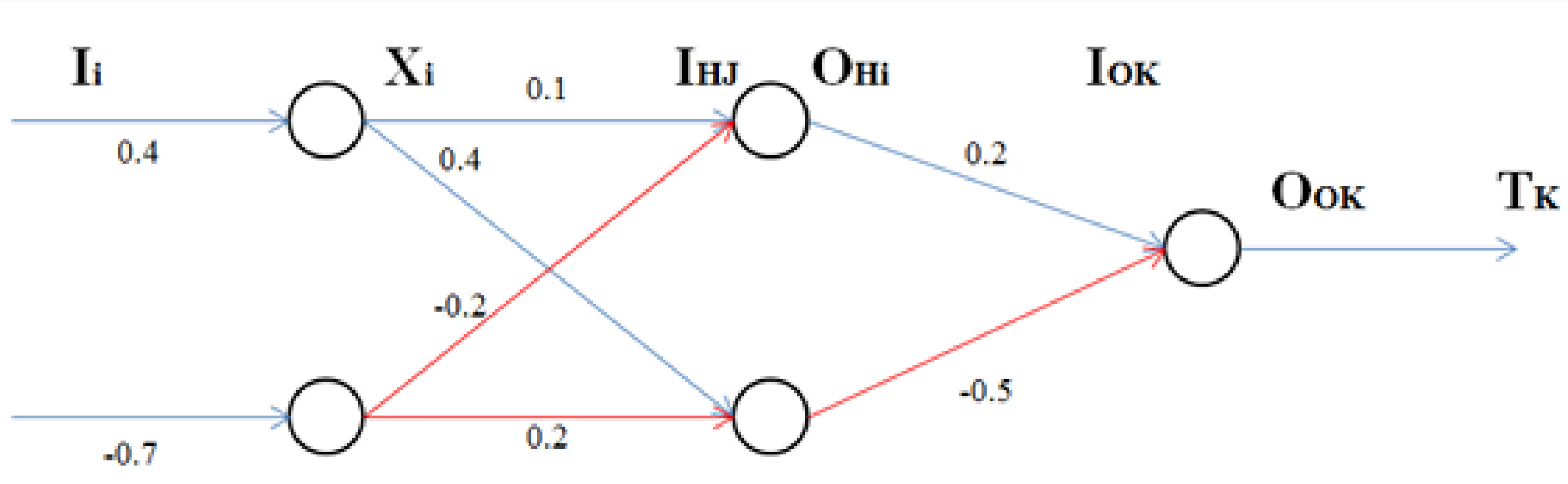




Module 3: Machine Learning in AI

1. Techniques in AI - 1. Neural Networks:

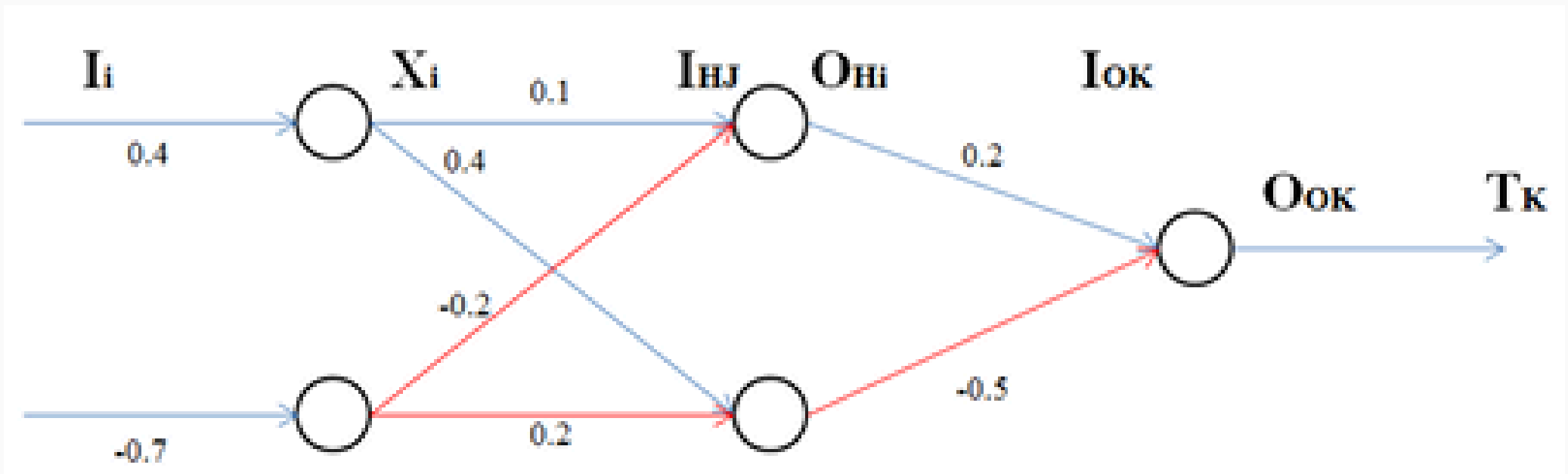
Si No	Input 1	Input 2	Output
1	0.4	-0.7	0.1
2	0.3	-0.5	0.05
3	0.6	0.1	0.3
4	0.2	0.4	0.25
5	0.1	-0.2	0.12





Module 3: Machine Learning in AI

1. Techniques in AI - 1. Neural Networks:



▶ $Input \{O\}_I = \{I\}_I = \begin{Bmatrix} 0.4 \\ -0.7 \end{Bmatrix}$

▶ Weights Initialization

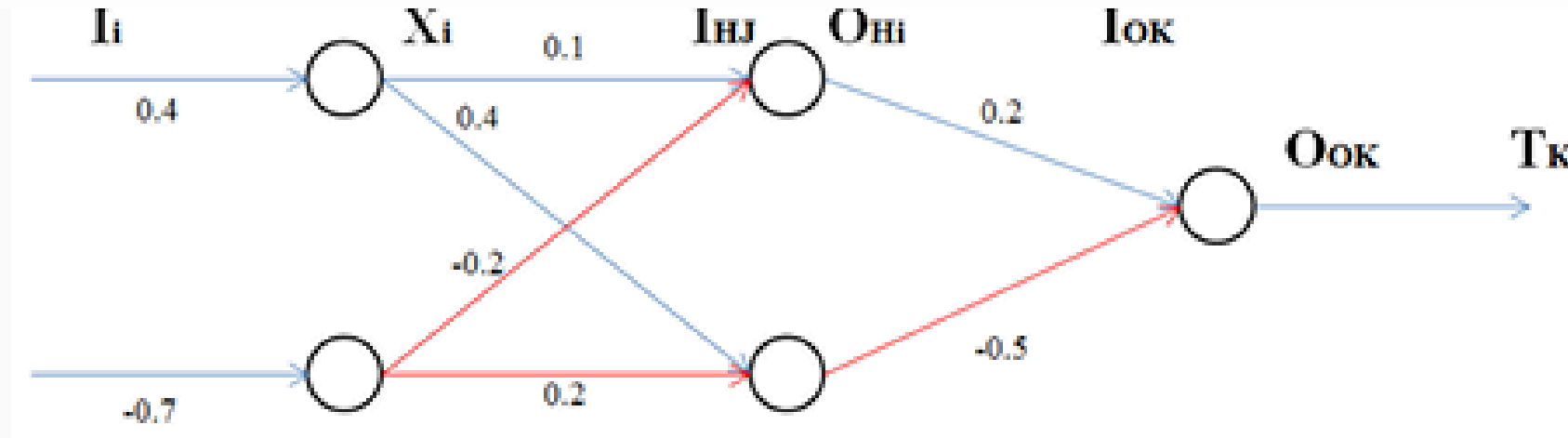
$$\begin{pmatrix} 0.1 & 0.4 \\ -0.2 & 0.2 \end{pmatrix} \begin{pmatrix} 0.2 \\ -0.5 \end{pmatrix}$$

[V] [V]



Module 3: Machine Learning in AI

1. Techniques in AI - 1. Neural Networks:



► *Input to Hidden Layer = $V^T \times \{O\}_I$*

$$\begin{pmatrix} 0.1 & -0.2 \\ 0.4 & 0.2 \end{pmatrix} \times \begin{pmatrix} 0.4 \\ -0.7 \end{pmatrix} = \begin{pmatrix} 0.18 \\ 0.02 \end{pmatrix}$$

$[V^T]$ $\{O\}_I$ *Input to Hidden Layer*

► *Output of Hidden Layer*

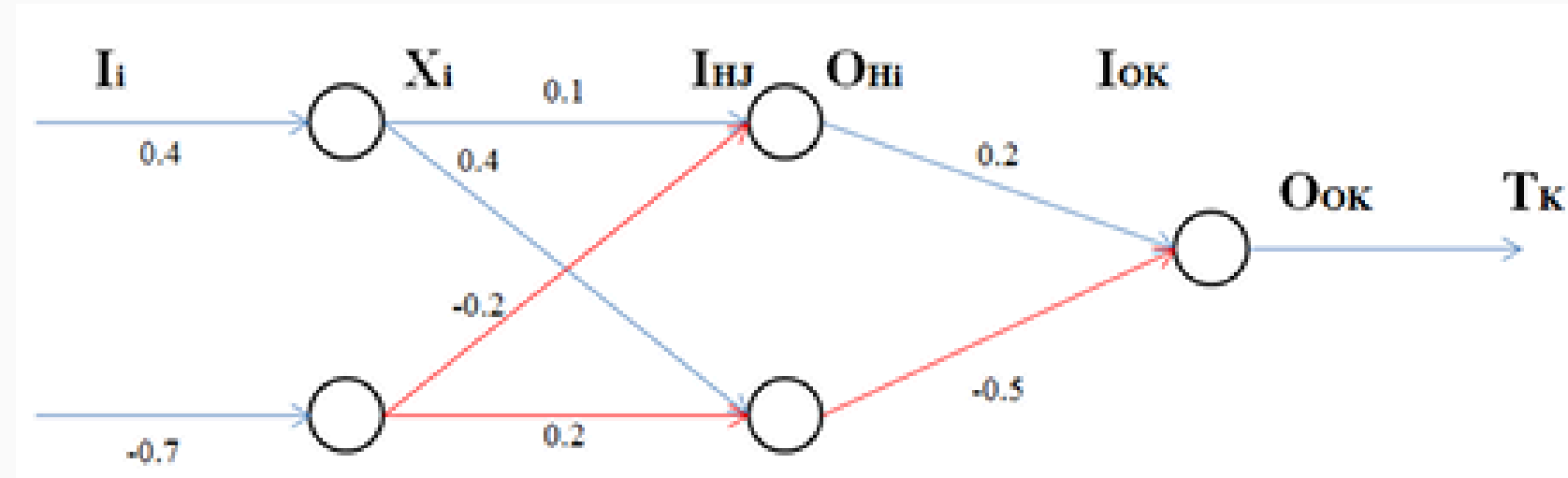
$$\begin{pmatrix} \frac{1}{1+e^{-0.18}} \\ \frac{1}{1+e^{-0.02}} \end{pmatrix} = \begin{pmatrix} 0.5448 \\ 0.505 \end{pmatrix}$$

Sigmoid Function *Output of Hidden Layer ($\{O\}_H$)*



Module 3: Machine Learning in AI

1. Techniques in AI - 1. Neural Networks:



► *Input to Output Layer* = $\{I\}_O = W^T \times \{O\}_H$

$$\begin{matrix} (0.2 & -0.5) \\ [W^T] \end{matrix} \times \begin{matrix} \begin{pmatrix} 0.5448 \\ 0.505 \end{pmatrix} \\ \{O\}_H \end{matrix} = -0.14354$$

► *Output of Output Layer*

$$\frac{1}{1+e^{-(-0.14354)}} = 0.4642$$

► *Error*

$$(T_O - O_O)^2 = (0.1 - 0.4642)^2 = 0.13264$$





Module 3: Machine Learning in AI

1. Techniques in AI - 1. Neural Networks:

Adjusting Weights of W

$$\Delta W = \eta \times \{O\}_H \times \langle d \rangle$$

$$\langle d \rangle = \langle (T_k - O_{ok}) O_{ok} (1 - O_{ok}) \rangle$$

$$[W]^{t+1} = [W]^t + [\Delta W]^{t+1}$$

$$\text{Let } \eta = 0.6$$

By calculation

$$T_o = 0.1$$

$$O_{o1} = 0.4642$$

$$\{O\}_H = \begin{pmatrix} 0.5448 \\ 0.505 \end{pmatrix}$$





Module 3: Machine Learning in AI

1. Techniques in AI - 1. Neural Networks:

Adjusting Weights of W

$$\langle d \rangle = \langle (0.1 - 0.4642) \ (0.4642) \ (1 - 0.4642) \rangle = -0.09058$$

$$[y] = \langle O_H \rangle \times \langle d \rangle$$

$$[y] = \begin{pmatrix} 0.5448 \\ 0.505 \end{pmatrix} \times \langle -0.09058 \rangle = \begin{pmatrix} -0.0493 \\ -0.0457 \end{pmatrix}$$

$$[\Delta W]^1 = \alpha [[\Delta W]^0] + \eta [y]$$

$$[\Delta W]^1 = 0.00001 \times 0 + 0.6 \times \begin{pmatrix} -0.0493 \\ -0.0457 \end{pmatrix} = \begin{pmatrix} -0.02958 \\ -0.02742 \end{pmatrix}$$





Module 3: Machine Learning in AI

1. Techniques in AI - 1. Neural Networks:

Adjusting Weights of V

$$\{e\} = [W]\langle d \rangle$$

$$\begin{pmatrix} 0.2 \\ -0.5 \end{pmatrix} \times \langle -0.09058 \rangle = \begin{pmatrix} -0.018116 \\ 0.04529 \end{pmatrix}$$

$$\langle d^* \rangle = e_i \times (O_{Hi}) \times (1 - O_{Hi})$$

$$\begin{pmatrix} -0.018116 \\ 0.04529 \end{pmatrix} \times \begin{pmatrix} 0.5448 \\ 0.505 \end{pmatrix} \times \begin{pmatrix} 1 - 0.5448 \\ 1 - 0.505 \end{pmatrix}$$

$$\begin{pmatrix} (-0.018116)(0.5448)(1 - 0.5448) \\ (0.04529)(0.505)(1 - 0.505) \end{pmatrix} = \begin{pmatrix} -0.00449 \\ 0.01132 \end{pmatrix}$$

$$[X] = \{O\}_I \times \langle d^* \rangle$$

$$\begin{pmatrix} 0.4 \\ -0.7 \end{pmatrix} \times \begin{pmatrix} -0.00449 & 0.01132 \end{pmatrix} = \begin{pmatrix} -0.001796 & 0.004528 \\ 0.003143 & -0.007924 \end{pmatrix}$$

$$[\Delta V]^1 = \alpha \times [\Delta V]^0 + \eta \times [X]$$

$$[\Delta V]^1 = 0.00001 \times 0 + 0.6 \times \begin{pmatrix} -0.001796 & 0.004528 \\ 0.003143 & -0.007924 \end{pmatrix} = \begin{pmatrix} -0.001077 & 0.0027168 \\ 0.0018858 & -0.0047544 \end{pmatrix}$$





Module 3: Machine Learning in AI

1. Techniques in AI - 1. Neural Networks:

Adjusting Weights of V

$$[V]^{t+1} = [V]^t + [\Delta V]^{t+1}$$

$$\begin{pmatrix} 0.1 & 0.4 \\ -0.2 & 0.2 \end{pmatrix} + \begin{pmatrix} -0.001077 & 0.0027168 \\ 0.0018858 & -0.0047544 \end{pmatrix} = \begin{pmatrix} 0.0989 & 0.04027 \\ -0.1981 & 0.19824 \end{pmatrix}$$

$$[W]^{t+1} = [W]^t + [\Delta W]^{t+1}$$

$$\begin{pmatrix} 0.2 \\ -0.5 \end{pmatrix} + \begin{pmatrix} -0.02958 \\ -0.02742 \end{pmatrix} = \begin{pmatrix} 0.17042 \\ -0.52742 \end{pmatrix}$$





Module 3: Machine Learning in AI

1. Techniques in AI - 1. Neural Networks:

Adjusting Weights of W

$$\langle d \rangle = \langle (0.1 - 0.4642) \ (0.4642) \ (1 - 0.4642) \rangle = -0.09058$$

$$[y] = \langle O_H \rangle \times \langle d \rangle$$

$$[y] = \begin{pmatrix} 0.5448 \\ 0.505 \end{pmatrix} \times \langle -0.09058 \rangle = \begin{pmatrix} -0.0493 \\ -0.0457 \end{pmatrix}$$

$$[\Delta W]^1 = \alpha [[\Delta W]^0] + \eta [y]$$

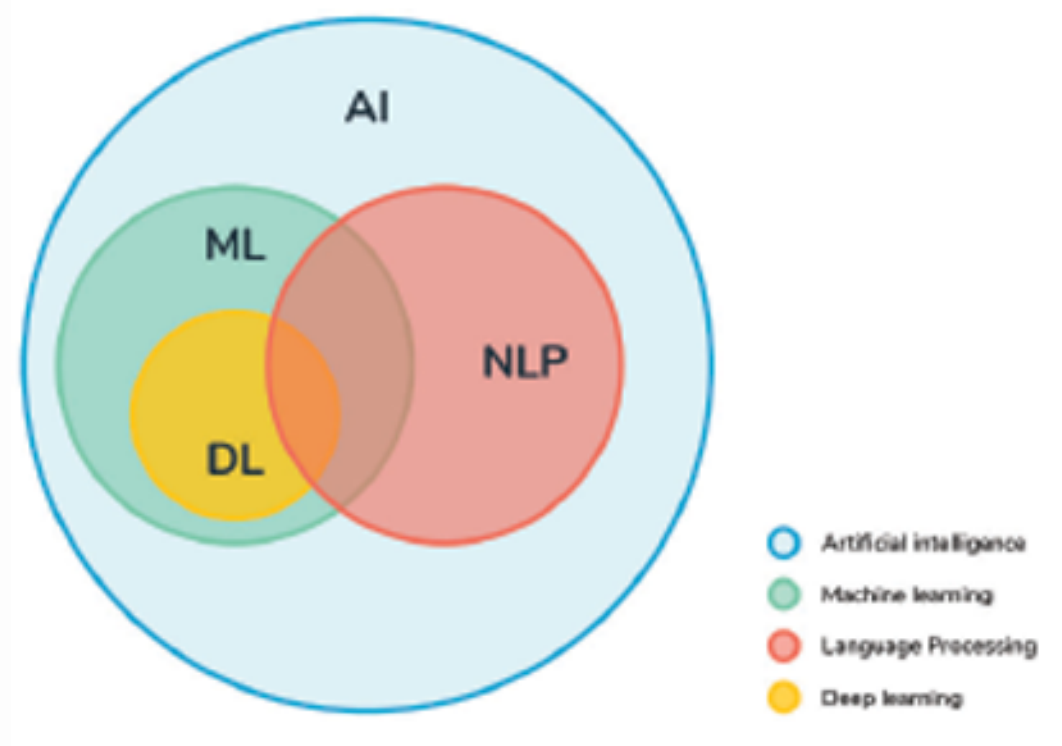
$$[\Delta W]^1 = 0.00001 \times 0 + 0.6 \times \begin{pmatrix} -0.0493 \\ -0.0457 \end{pmatrix} = \begin{pmatrix} -0.02958 \\ -0.02742 \end{pmatrix}$$





Module 3: Machine Learning in AI

1. Techniques in AI - 2 Machine learning:



Machine Learning (ML) is a field of computer science that focuses on teaching machines to learn from data and make decisions on their own, without needing human instructions.

How Machine Learning Works:

1. Learning from Data:
2. No Human Instructions Needed:
3. Improvement Over Time:
4. Making Better Decisions:





Module 3: Machine Learning in AI

1. Techniques in AI - 2 Machine learning:

How Machine Learning Works:

1. Learning from Data:

- ML algorithms analyze large sets of data to find patterns.
- For example, an ML algorithm might look at thousands of images of cats and dogs, and learn to recognize the differences between them.

2. No Human Instructions Needed:

- Instead of writing step-by-step rules for the machine, ML learns automatically from the data.
- Example: If you want the machine to recognize a cat, you don't need to tell it exactly what a cat looks like. Instead, you feed it data of cats and dogs, and the machine learns from the examples.

3. Improvement Over Time:

- The machine improves by analyzing its errors and correcting them automatically.
- Example: If it wrongly identifies a dog as a cat, it adjusts and tries again, getting better each time.

4. Making Better Decisions:

- ML saves time and effort because it automates decision-making. It can analyze data and make decisions more accurately than humans.
- Example: In healthcare, ML can analyze medical data and predict diseases faster than doctors.





Module 3: Machine Learning in AI

1. Techniques in AI - 2 Machine learning:

How It's Different from Human Instructions:

- **Human Instructions:**
 - Humans write specific rules to perform tasks (e.g., write code for a program).
- **Machine Learning:**
 - Machines learn by themselves from data, without explicit rules. They improve as they process more data.

Example in Action:

- **Problem:** Classifying emails as "spam" or "not spam."
- **Without ML:** Humans write specific rules (e.g., "If email contains 'buy now', mark as spam").
- **With ML:** The machine analyzes data of past emails and learns to identify patterns like words, subject lines, or email addresses that are likely to be spam. It improves over time as it processes more emails.





Module 3: Machine Learning in AI

1. Techniques in AI - 2 Machine learning:

Relationship Between AI, ML, DL, and NLP

Machine Learning is a sub-domain of Artificial Intelligence (AI). AI refers to the broad concept of creating machines that can think and act intelligently. Within AI, we have specific areas like Machine Learning (ML), Deep Learning (DL), and Natural Language Processing (NLP).

- **AI:** The broad field of creating intelligent machines.
- **ML:** Focuses on teaching machines to learn from data.
- **DL:** A deeper level of ML that uses neural networks to process large datasets, especially useful for tasks like image recognition.
- **NLP:** A branch of AI that enables machines to understand, interpret, and respond to human language.





Module 3: Machine Learning in AI

1. Techniques in AI - 3 Deep learning

Deep Learning (DL) is a powerful technique within Machine Learning (ML) that uses neural networks with many layers to process and learn from data. It is especially good at handling complex patterns in large datasets.

How Deep Learning Works:

1. Neural Networks with Layers:
2. Forward Propagation:
3. Backward Propagation:
4. Training the Model:





Module 3: Machine Learning in AI

1. Techniques in AI - 3 Deep learning

How Deep Learning Works:

1. Neural Networks with Layers:

- In Deep Learning, a neural network has multiple layers of processing units (also called nodes or neurons).
- Each layer processes the data, learns patterns, and passes the results to the next layer.
- Example: In image recognition, the first layer might detect edges, the next layer might recognize shapes, and further layers might identify objects.





Module 3: Machine Learning in AI

1. Techniques in AI - 3 Deep learning

How Deep Learning Works:

2. Forward Propagation:

- As the data moves through each layer, it gets processed and refined until it reaches the output layer.
- This is called forward propagation, where the input data is passed through the network and results in the final output.
- Example: In an image recognition task, forward propagation will take raw pixel data, process it layer by layer, and finally output a classification like "cat" or "dog."





Module 3: Machine Learning in AI

1. Techniques in AI - 3 Deep learning

How Deep Learning Works:

3. Backward Propagation:

- Once the output is generated, the model checks if the result is accurate by calculating the error (the difference between predicted and actual results).
- If the result is wrong, the weights (importance) of the nodes are updated and the error is sent back through the network to adjust the weights. This is called backward propagation.
- Example: If the model misclassifies an image, backward propagation helps the model adjust its calculations to avoid the same mistake next time.





Module 3: Machine Learning in AI

1. Techniques in AI - 3 Deep learning

How Deep Learning Works:

4. Training the Model:

- The model learns from both the forward and backward propagation processes by adjusting weights, improving its accuracy with every iteration.
- Example: Over time, with enough data, the model gets better at recognizing images or predicting outcomes.





Module 3: Machine Learning in AI

1. Techniques in AI - 3 Deep learning

Types of Learning in Deep Learning:

- **Supervised Learning:**
 - In supervised learning, the model is trained on labeled data (i.e., data with known outcomes). The model learns to predict outcomes based on these examples.
 - Example: Training a model to identify cats in photos, where the images are labeled as "cat" or "not cat."
- **Unsupervised Learning:**
 - Deep learning can also work with unlabeled data, where the model identifies patterns or structures in data on its own, without needing specific labels.
 - Example: Grouping customers into different segments based on their purchasing behavior without predefined labels.





Module 3: Machine Learning in AI

1. Techniques in AI - 3 Deep learning

Applications of Deep Learning:

1. Image Recognition:

- Deep learning is excellent at identifying objects, faces, and scenes in images.
- Example: Face recognition systems in smartphones.

2. Speech Recognition:

- Deep learning models can process audio data to convert speech into text.
- Example: Virtual assistants like Siri or Google Assistant that understand spoken commands.





Module 3: Machine Learning in AI

1. Techniques in AI - 3 Deep learning

Concept	Explanation	Example
Neural Networks	A series of interconnected layers that process data in steps.	Deep learning in image recognition where the model processes raw images through multiple layers.
Forward Propagation	Data moves through layers to generate an output.	The input image is processed layer by layer until the model outputs "cat."
Backward Propagation	Adjusting weights in the network based on errors to improve accuracy.	If the model misclassifies an image, backward propagation updates the weights.
Supervised Learning	The model is trained on labeled data.	Training a model to predict house prices based on features like size, location, and number of rooms.
Unsupervised Learning	The model learns from unlabeled data, finding hidden patterns.	Clustering customers based on buying habits without predefined labels.





Module 3: Machine Learning in AI

1. Techniques in AI - 4 Natural Language Processing (NLP)

Natural Language Processing (NLP) is a field of Artificial Intelligence (AI) focused on enabling machines to read, understand, and respond to human language in a way that is both meaningful and useful.

Applications of NLP:

- Chatbots and Virtual Assistants:
- Text Translation:
- Sentiment Analysis:
- Text Summarization:

How NLP Works:

1. Reading and Understanding:
2. Interpreting:
3. Responding:





Module 3: Machine Learning in AI

1. Techniques in AI - 4 Natural Language Processing (NLP)

1. Reading and Understanding:

- NLP allows machines to process human language (spoken or written) just like we do. It breaks down sentences into meaningful units, such as words and phrases.
- Example: If you say, “What’s the weather like today?”, NLP helps the machine understand that you are asking about the weather.

2. Interpreting:

- a. After understanding the words, NLP helps the machine interpret the meaning behind them, considering the context.
- b. Example: The machine knows that "cold" can mean "low temperature" or "emotionally distant," depending on the context.

3. Responding:

- a. Once the machine interprets the input, NLP enables it to respond in a natural way, either in text or speech.
- b. Example: If you ask, “Tell me a joke,” the machine will generate a response like, “Why don’t skeletons fight each other? They don’t have the guts!”





Module 3: Machine Learning in AI

1. Techniques in AI - 5 Computer Vision

Computer Vision is a branch of Artificial Intelligence (AI) that focuses on enabling machines to see, understand, and interpret images and videos just like humans do.

Applications of Computer Vision:

- Facial Recognition:
- Autonomous Vehicles:
- Retail Stores:
- Medicine:
- Financial Institutions:

How Computer Vision Works

1. Breaking Down Images:
2. Classifying and Learning:
3. Making Decisions:





Module 3: Machine Learning in AI

1. Techniques in AI - 5 Computer Vision

How Computer Vision Works:

1. Breaking Down Images:

- Computer vision starts by breaking an image into smaller parts (like shapes, colors, and textures).
- Example: In an image of a cat, computer vision might break it down into features like the shape of the ears, eyes, and whiskers.

2. Classifying and Learning:

- The machine then uses patterns from many images to learn and classify objects (e.g., identifying a cat in an image).
- Example: After processing many images of cats, the machine learns to identify common features like pointy ears and a small nose.

3. Making Decisions:

- Based on what the machine has learned from past observations, it makes decisions about the current image or video it sees.
- Example: If the machine sees a new image of a cat, it can correctly identify it by matching the features to what it has learned.





Module 3: Machine Learning in AI

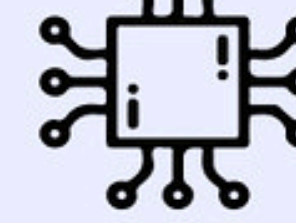
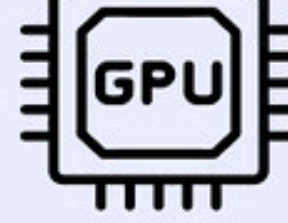
1. Techniques in AI - 6 Cognitive Computing

- Cognitive Computing is a subfield of Artificial Intelligence (AI) designed to make machines think and respond in a human-like way. It mimics the way the human brain works by analyzing text, speech, images, or objects to understand and interact with the world.
- **Applications of Cognitive Computing:**
 - Virtual Assistants:
 - Healthcare:
 - Customer Service:
 - Education:



Module 3: Machine Learning in AI

1. Techniques in AI - Additional technologies that enable and support AI



GPU vs CPU

FOR IMAGE PROCESSING


GPU works excellently for complex image processing tasks

Better throughput

Group of thread executes the instructions

Implements genuine thread rotation

The number of registers available in GPU can be changed dynamically


Has fine-grained parallelism


Better clock speed

Each thread executes the instructions individually

Implements out-of order execution

Has coarse-grained parallelism


GPU performance is over ten times faster than that of a CPU


Memory is available across all the multiprocessors or cores





Module 3: Machine Learning in AI

1. Techniques in AI - Additional technologies that enable and support AI



1. Graphical Processing Units (GPUs) in AI

- Graphical Processing Units (GPUs) are essential for AI tasks because they provide the computing power required for iterative processing and training neural networks.
- GPUs are designed to handle these large volumes of data more efficiently than traditional CPUs, making them ideal for tasks like:
 - **Parallel Processing:** GPUs can process many operations at once, speeding up the training of models.
 - **Faster Training of Neural Networks:** They help process the data through multiple layers in deep learning, making models train faster.
 - **Example:** In deep learning, GPUs are used to train models for tasks like image recognition or speech recognition, where data is analyzed and patterns are learned across millions of inputs.





Module 3: Machine Learning in AI

1. Techniques in AI - Additional technologies that enable and support AI

2. Internet of Things (IoT) and AI for Data Analysis

The Internet of Things (IoT) refers to a network of connected devices that generate massive amounts of data. These devices can range from smartphones and wearables to home appliances and industrial sensors. However, this data often remains unprocessed or under-analyzed.

AI and advanced algorithms help in automating the analysis of IoT data, extracting useful insights and making sense of the vast amounts of information. Here's how AI plays a role:

- **Data Analysis at Scale:** AI can analyze data from millions of connected devices quickly and efficiently.
- **Predicting Rare Events:** AI models can identify patterns and predict rare events (like equipment failure in manufacturing) that could otherwise be missed.
- **Understanding Complex Systems:** AI helps make sense of complex systems (e.g., smart cities, healthcare systems) by analyzing data from various sources.

○





Module 3: Machine Learning in AI

1. Techniques in AI - Additional technologies that enable and support AI

2. Internet of Things (IoT) and AI for Data Analysis

The Internet of Things (IoT) refers to a network of connected devices that generate massive amounts of data. These devices can range from smartphones and wearables to home appliances and industrial sensors. However, this data often remains unprocessed or under-analyzed.

AI and advanced algorithms help in automating the analysis of IoT data, extracting useful insights and making sense of the vast amounts of information. Here's how AI plays a role:

- **Data Analysis at Scale:** AI can analyze data from millions of connected devices quickly and efficiently.
- **Predicting Rare Events:** AI models can identify patterns and predict rare events (like equipment failure in manufacturing) that could otherwise be missed.
- **Understanding Complex Systems:** AI helps make sense of complex systems (e.g., smart cities, healthcare systems) by analyzing data from various sources.

○





Module 3: Machine Learning in AI

2 Machine Learning Model

Machine Learning (ML) involves teaching a computer program to improve its performance at a specific task through experience (data). A widely accepted definition by Professor Mitchell describes it as:

“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 .”

This definition leads to understanding Machine Learning Models in three key components:

1. Task (T),
2. Experience (E), and
3. Performance (P).

○





Module 3: Machine Learning in AI

2 Machine Learning Model

Key Components of a Machine Learning Model:

1. Task (T):

- This is the real-world problem that we want the machine to solve.
- Examples of tasks (T):
 - Predicting sales of a product.
 - Classifying emails as spam or not spam.
 - Other examples: classification, regression, clustering, and recognition.
- Task (T) defines what the machine is trying to learn or predict.





Module 3: Machine Learning in AI

2 Machine Learning Model

Key Components of a Machine Learning Model:

2. Experience (E):

- Experience refers to the data that the machine learns from.
- Just like humans learn from past experiences, machines learn from data. The more data the model is exposed to, the better it can learn and improve.
- Example:
 - If you want the model to classify emails, you provide it with a set of labeled emails (spam and non-spam). This is its experience (E).
- Experience (E) can be gained through different types of learning:
 - a. Supervised Learning (with labeled data),
 - b. Unsupervised Learning (with unlabeled data),
 - c. Reinforcement Learning (through trial and error, receiving rewards or penalties).





Module 3: Machine Learning in AI

2 Machine Learning Model

Key Components of a Machine Learning Model:

3. Performance (P):

- Performance is how well the model performs a task after learning from the data. It's a measure of the accuracy or effectiveness of the model's predictions.
- Metrics to measure performance (P):
 - Accuracy: How many predictions were correct.
 - Precision: How many predicted positive outcomes were actually positive.
 - Recall: How many actual positive outcomes were identified.
 - F1 Score: A balance between precision and recall.
 - Confusion Matrix: A table showing true positive, false positive, true negative, and false negative results.





Module 3: Machine Learning in AI

2 Machine Learning Model

How the Machine Learning Model Works:

1. **Task (T):** Define the problem (e.g., classifying emails).
2. **Experience (E):** Provide data for the model to learn from (e.g., labeled emails).
3. **Performance (P):** Measure the model's effectiveness using metrics (e.g., accuracy or precision).

Example of the ML Model:

Let's say we want to build a model to predict house prices (Task T):

- **Task (T):** Predict the price of a house based on features like size, location, and number of rooms.
- **Experience (E):** We provide the model with historical data on house sales (e.g., size, location, price).
- **Performance (P):** The model's performance is measured by how accurately it predicts the price of new houses, which we evaluate using metrics like mean squared error or R^2 .





Module 3: Machine Learning in AI

2 Machine Learning Model

Types of Machine Learning Algorithms

1. Supervised Machine Learning Algorithms
 - a. Classification Algorithm
 - b. Regression
2. Unsupervised Learning
 - a. Clustering
 - b. Association analysis:
 - c. Dimensionality reduction
 - d. Outlier detection or anomaly detection
3. Semi-Supervised Learning
4. Reinforcement Learning (RL)
 - a. Positive reinforcement
 - b. Negative reinforcement





Module 3: Machine Learning in AI

2 Machine Learning Model - Supervised Machine Learning Algorithms

Supervised Machine Learning is a type of learning where the machine learns from labeled data. Think of it like a student learning from a teacher who provides both the questions (inputs) and the correct answers (outputs).

How It Works:

1. Learning from Past Data:

- The algorithm is trained on labeled examples, where both the input and output are provided.
- The machine learns the relationship between the inputs (X) and the outputs (Y).
- Example: If we have a basket of fruits with labels (apple, banana, orange), the machine learns to associate features (like color, shape, size) of fruits with their labels.

2. Mapping Function ($Y = f(X)$):

- The relationship between input (X) and output (Y) can be expressed as a function, like $Y = f(X)$, where the machine learns to map inputs to outputs.
- After learning from the data, it can predict the output for new, unseen data.
- Example: If the input is a fruit with color = red, shape = round, the model predicts that it's an apple.





Module 3: Machine Learning in AI

2 Machine Learning Model - Supervised Machine Learning Algorithms

Supervised Learning Example:

Imagine you have a basket filled with apples, bananas, and oranges. You want to teach a machine to identify them based on certain features like color, size, and shape.

1. **Step 1:** Train the model using labeled data.

- Input: Fruit features (e.g., color, shape, size).
- Output: Correct labels (apple, banana, orange).

2. **Step 2:** The machine learns to map the input features to the correct fruit labels (mapping function: $Y = f(X)$).

3. **Step 3:** Now, if you give the machine a fruit with unknown features (e.g., color = yellow, shape = elongated), it will predict that it's a banana.

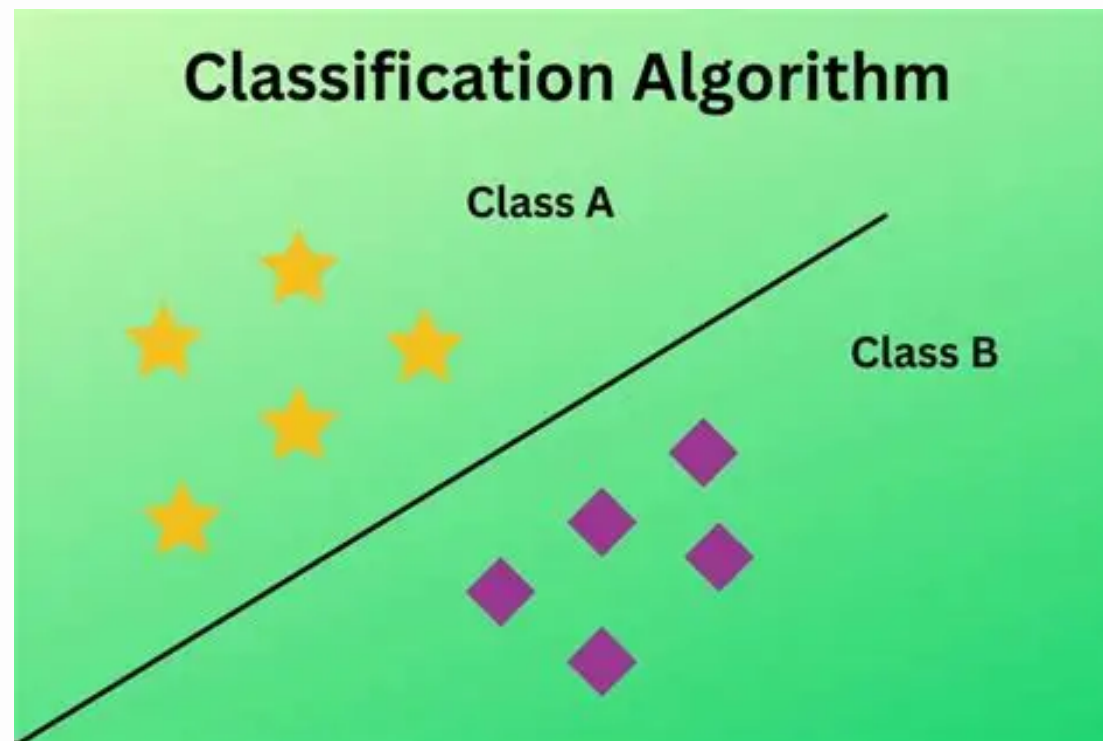




Module 3: Machine Learning in AI

2 Machine Learning Model - Supervised Machine Learning Algorithms

1. Classification Algorithm



A Classification Algorithm is a type of machine learning algorithm that classifies or categorizes data into specific groups or categories. It helps the machine predict which category a new piece of data belongs to based on patterns learned during training.

How Classification Works:

- Training Phase:
 - During the training phase, the model is given labeled data, meaning each piece of data already has a known category (label).
 - Example: If you have a dataset of fruits with labels like apple and banana, the algorithm learns the features (color, size, shape) of each fruit and associates them with the correct label.
- Prediction:
 - After learning from the data, the model can predict the category (class) for new, unseen data based on the patterns it has learned.
 - Example: When given a new fruit (e.g., color = red, shape = round), the model predicts that it's an apple.

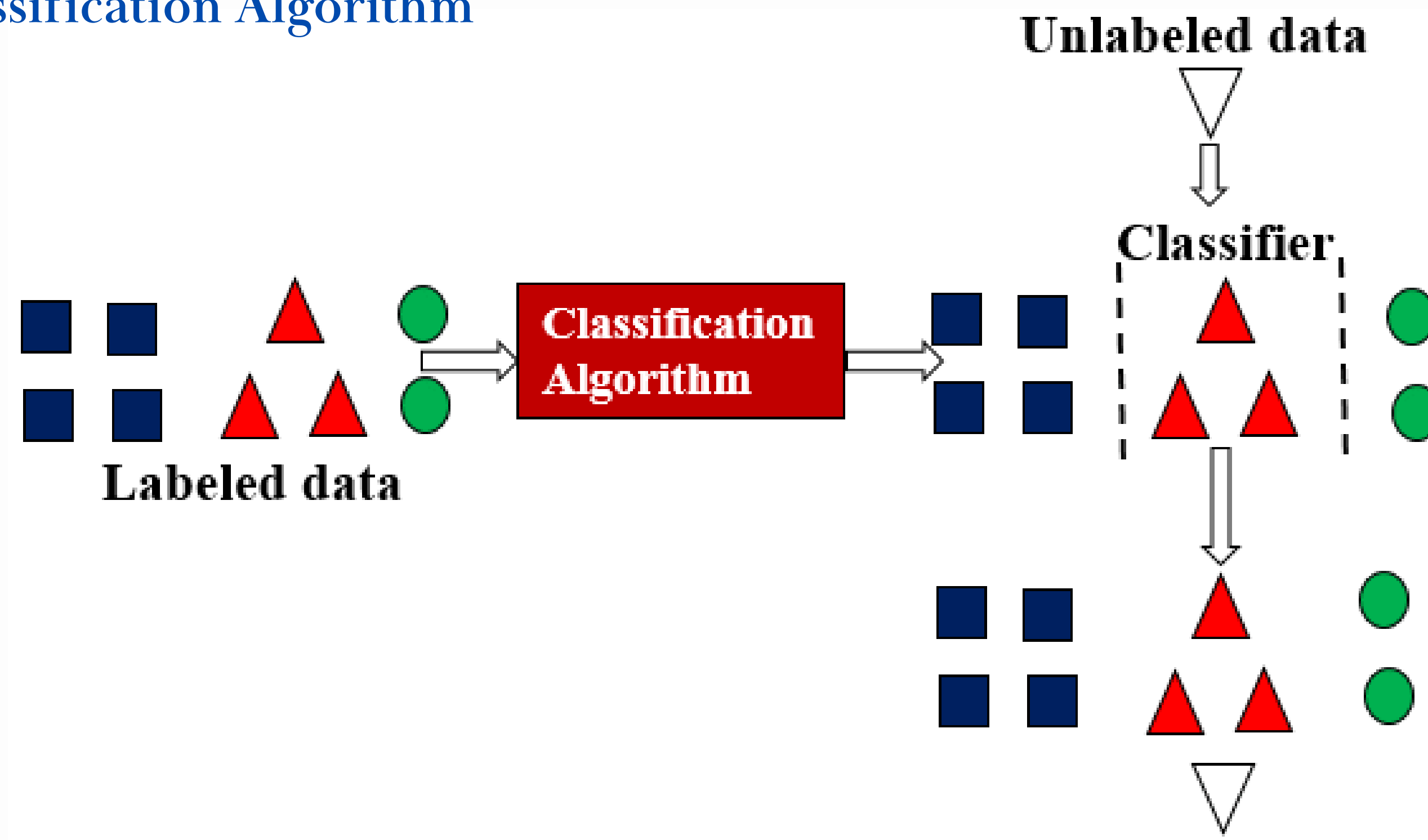




Module 3: Machine Learning in AI

2 Machine Learning Model - Supervised Machine Learning Algorithms

1. Classification Algorithm





Module 3: Machine Learning in AI

2 Machine Learning Model - Supervised Machine Learning Algorithms

1. Classification Algorithm

Applications of Classification Algorithms:

1. Medical Imaging:

- Example: Predicting whether an X-ray image shows a benign or cancerous tumor.

2. Speech Recognition:

- Example: Classifying spoken words into specific commands like "turn on" or "turn off".

3. Handwriting Recognition:

- Example: Identifying handwritten characters and classifying them as A, B, or C.

4. Credit Scoring:

- Example: Classifying loan applicants into groups like high risk or low risk based on their financial data.

5. Email Classification:

- Example: Predicting whether an incoming email is spam or not spam based on previous examples.





Module 3: Machine Learning in AI

2 Machine Learning Model - Supervised Machine Learning Algorithms

2. Regression Algorithm

A Regression Algorithm is used in machine learning to predict a real value (a continuous number) based on the data it has learned from. Unlike classification algorithms, which predict discrete categories, regression algorithms predict continuous values.

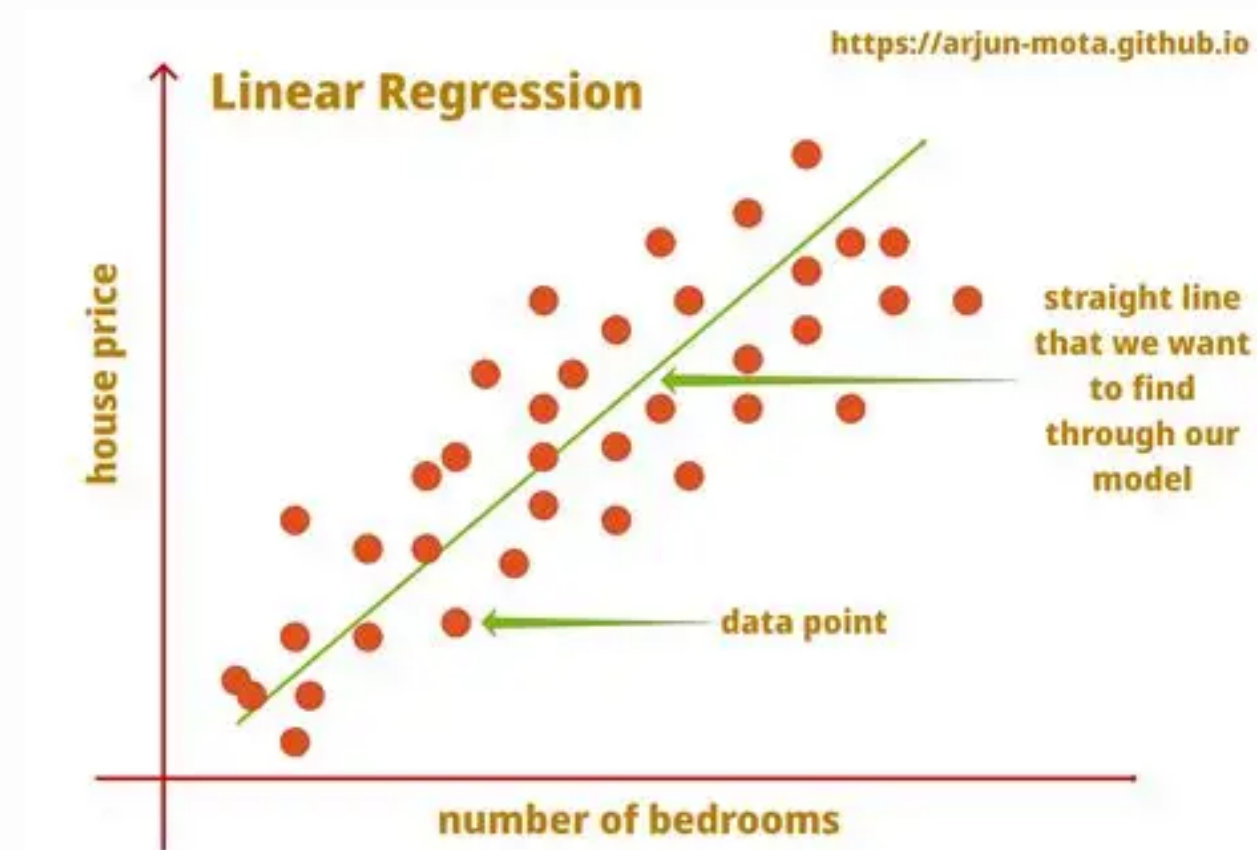
How Regression Works:

1. Training Phase:

- a. During training, the algorithm is provided with labeled data that includes input variables (features) and their corresponding real values (output).
 - **Example:** You have data on house features (size, location, number of rooms) and their prices. The model learns the relationship between these features and the price.

2. Prediction:

- a. After learning, the model can predict the real value (continuous output) for new, unseen data.
 - **Example:** Given the size, location, and number of rooms of a new house, the model can predict its price.





Module 3: Machine Learning in AI

2 Machine Learning Model - Supervised Machine Learning Algorithms

Categories of Supervised Learning:

Supervised learning can be classified into two main types:

1. Classification:

- When the output is a category or label.
- Example: Identifying whether an email is spam or not spam.

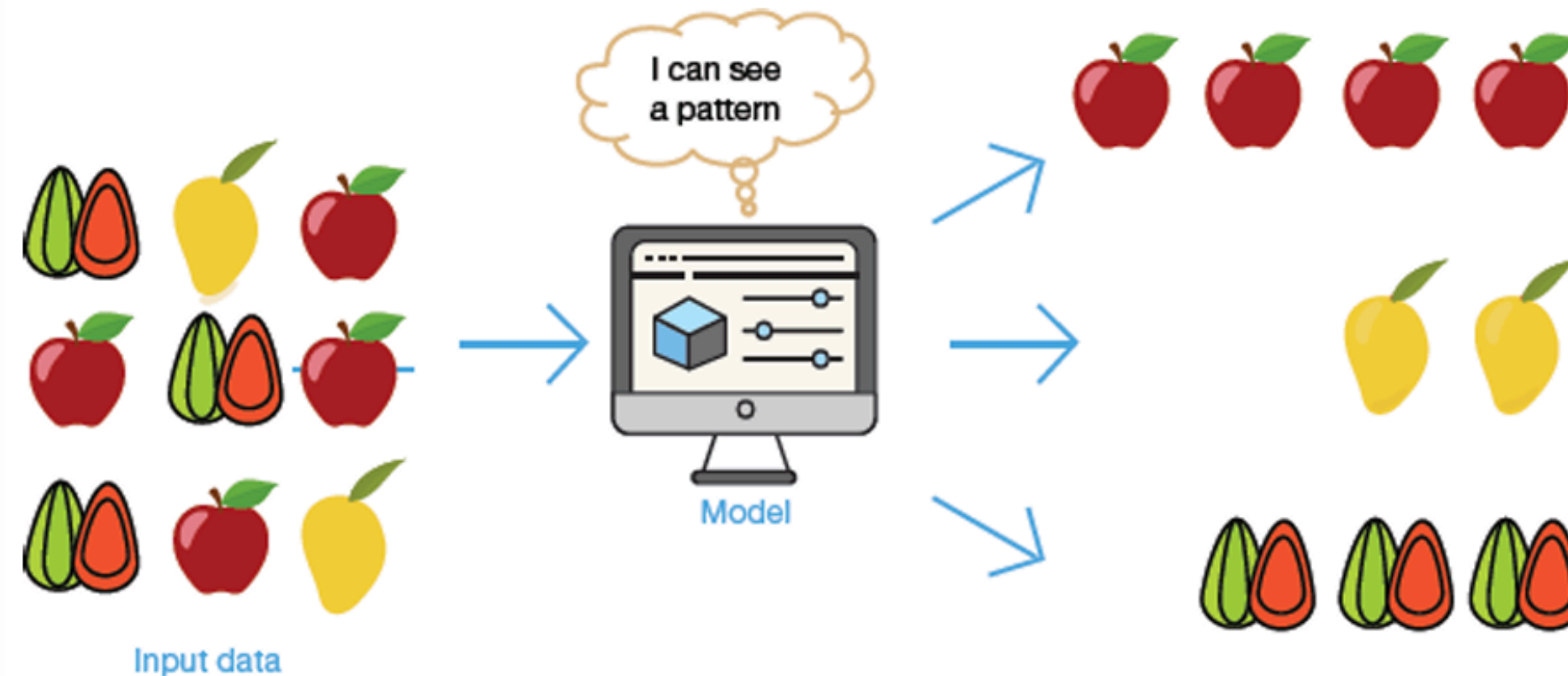
2. Regression:

- When the output is a continuous value.
- Example: Predicting the price of a house based on its size, location, and features.



Module 3: Machine Learning in AI

2 Machine Learning Model - Un-Supervised Machine Learning Algorithms



Unsupervised Learning is a type of machine learning where the model is trained using data that is neither labeled nor classified. This means the machine doesn't have any predefined answers to learn from and must discover patterns or relationships in the data on its own.



Module 3: Machine Learning in AI

2 Machine Learning Model - Un-Supervised Machine Learning Algorithms

How Unsupervised Learning Works:

1. No Labeled Data:

- In unsupervised learning, there is no guidance or labels provided to the machine. It only has the data and must find patterns or groupings within it.
- Example:
- If we give the machine images of fruits (like mangoes and oranges) without labels, it doesn't know what a mango or orange is. It must figure out how to group the fruits based on similarities like color, shape, or texture.

2. Finding Patterns and Grouping:

- The machine looks for patterns, similarities, and differences in the data and creates groups or clusters.
- Example: The machine might notice that all mangoes are yellow and oranges are orange, so it will group the images into two clusters based on color and shape.

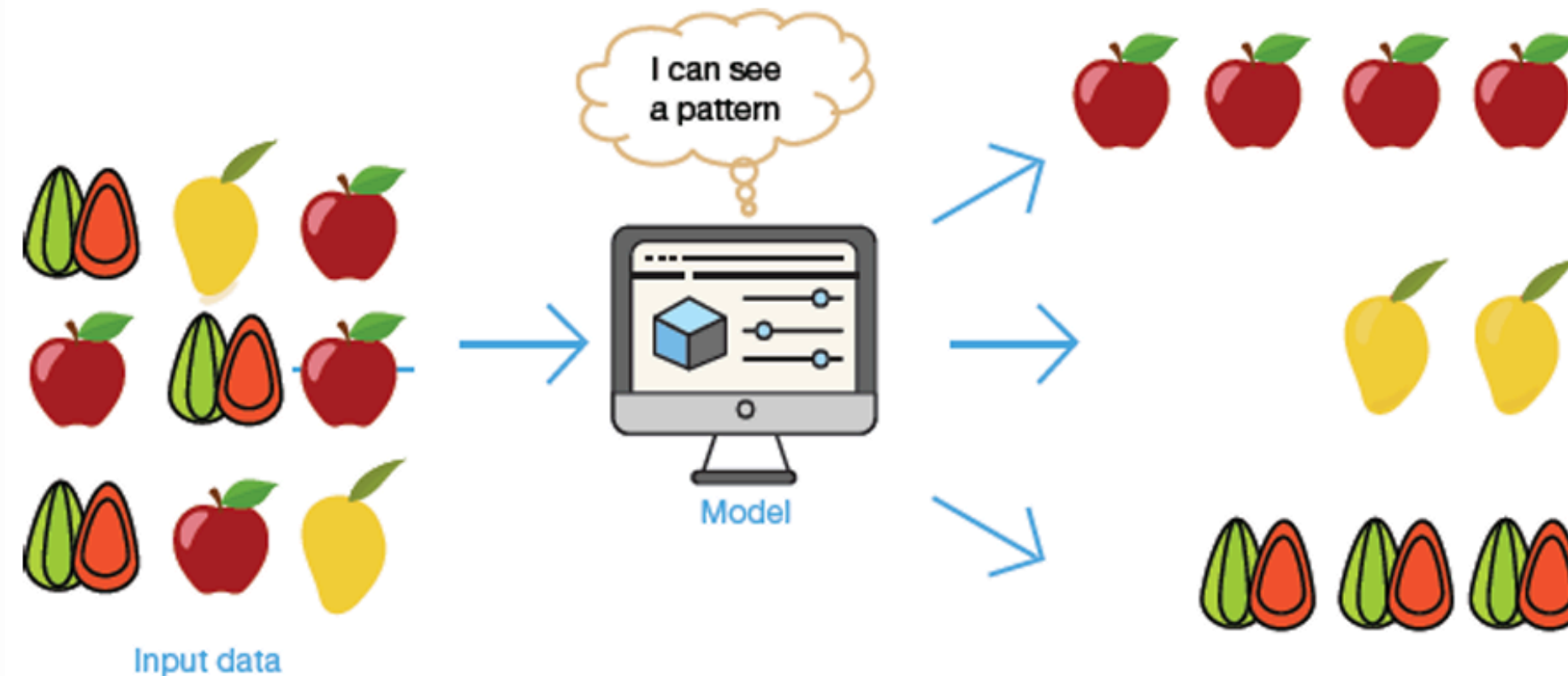
3. Clustering:

- Unsupervised learning works by creating clusters where similar data points are grouped together.
- Example: If the machine is given a dataset of fruits, it could automatically create two clusters—one for mangoes and one for oranges—without knowing what these fruits are called.



Module 3: Machine Learning in AI

2 Machine Learning Model - Un-Supervised Machine Learning Algorithms



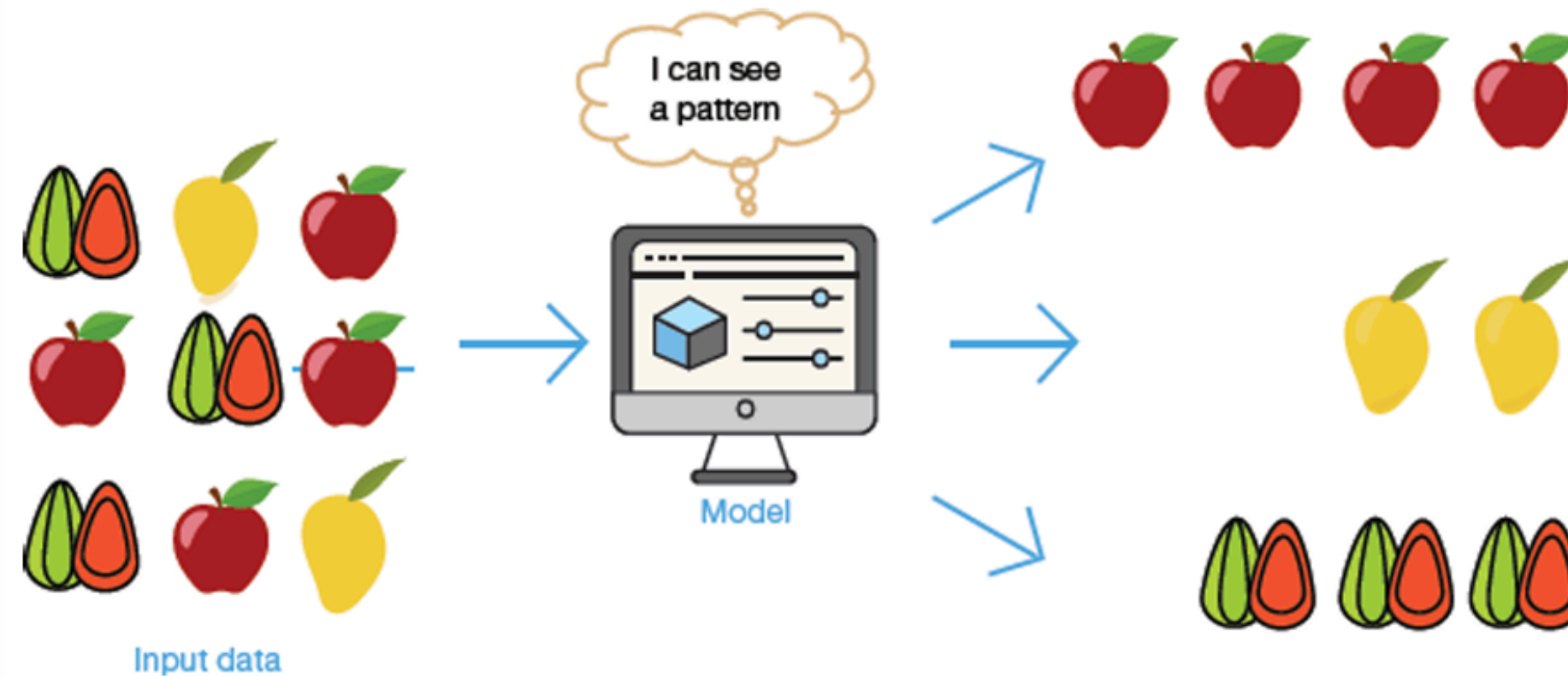
Unsupervised Learning in Action:

1. Clustering Fruits:

- Imagine the machine receives pictures of mangoes and oranges, but it has no idea what fruits these are.
- The model will analyze the features of the images (like color, shape, texture) and group similar images together.
- Result: It will create two clusters—one for mangoes and one for oranges—but it won't label them as "mango" or "orange." The machine simply groups similar items together.

Module 3: Machine Learning in AI

2 Machine Learning Model - Un-Supervised Machine Learning Algorithms



Unsupervised Learning in Action:

2. Finding Patterns in Data:

Example: In a dataset of customer behaviors, unsupervised learning could identify patterns and group customers with similar shopping habits, even if no labels (like “high spender” or “low spender”) are given.

○



Module 3: Machine Learning in AI

2 Machine Learning Model - Un-Supervised Machine Learning Algorithms

What Unsupervised Learning Cannot Do:

- **No Labels:**
 - Unlike supervised learning, unsupervised learning does not provide labels or categories. It can only group data based on patterns or similarities.
 - Example: The machine can group mangoes and oranges, but cannot label them as such. It just knows these items belong together in one group.

Applications of Unsupervised Learning:

- **Customer Segmentation:**
 - Unsupervised learning can help businesses group customers with similar purchasing behaviors for targeted marketing campaigns.
- **Image Segmentation:**
 - It can be used in image processing to automatically group pixels that form certain objects or regions in an image.
- **Anomaly Detection:**
 - It helps detect outliers or unusual data points in a dataset, which can be useful for fraud detection or network security.

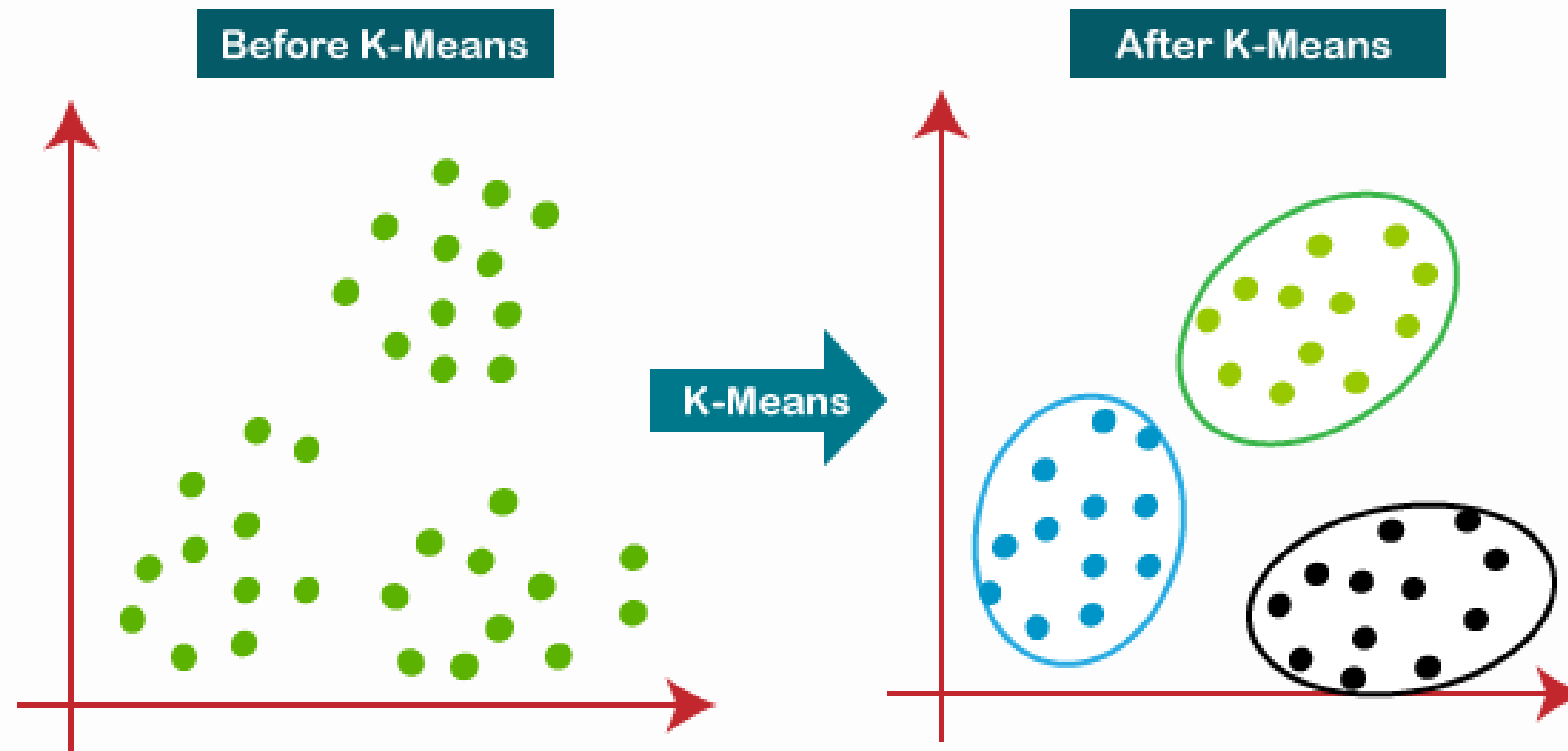




Module 3: Machine Learning in AI

2 Machine Learning Model - Un-Supervised Machine Learning Algorithms

1. Clustering



Clustering is a technique in unsupervised machine learning where the goal is to group similar data points together into clusters. This technique helps to discover hidden patterns or inherent groupings in the data, which might not be obvious at first glance.





Module 3: Machine Learning in AI

2 Machine Learning Model - Un-Supervised Machine Learning Algorithms

1. Clustering

How Clustering Works:

1. Grouping Data Based on Similarities:

- The main idea of clustering is to group data points that are similar to each other.
- Each group is called a cluster, and the goal is for data within the same cluster to be similar while data in different clusters is distinct or different.
- Example:
 - In a dataset of customer purchases, clustering might group customers who buy similar products into the same cluster. One cluster might include people who buy tech gadgets, while another might include people who buy home appliances.

2. Finding Patterns in Data:

- Clustering identifies patterns or relationships in the data by grouping similar points together.
- Example:
 - A cell phone company can use clustering to find where most of its customers live and use that information to decide the best locations to build new cell phone towers.





Module 3: Machine Learning in AI

2 Machine Learning Model - Un-Supervised Machine Learning Algorithms

1. Clustering

Real-World Applications of Clustering:

1. Customer Segmentation:

- Companies can use clustering to group customers based on purchasing behavior, allowing them to target marketing campaigns to specific groups.
- Example: A retail store can identify one cluster of customers who frequently buy sportswear, while another cluster may buy formal clothing.

2. Location-based Decisions (e.g., Cell Tower Placement):

- Companies can use clustering to find the best locations for services like cell towers, based on where the most people are located.
- Example: A mobile company could use clustering to identify cities with the highest customer density, helping them plan where to build more towers.

3. Gene Sequence Analysis:

- In biology, clustering is used to group similar gene sequences, helping researchers to understand genetic relationships and identify specific patterns.
- Example: Researchers can analyze gene data to identify groups of genes that behave similarly across different species.





Module 3: Machine Learning in AI

2 Machine Learning Model - Un-Supervised Machine Learning Algorithms

2. Association Analysis

Association Analysis is a technique used in data mining to find relationships or patterns between different items in large datasets. The goal is to discover rules that describe these patterns, showing how items or events are associated with each other.

How Association Analysis Works:

1. Finding Relationships:

- Association analysis tries to find patterns that describe how different items or actions are linked together.
- Example: If customers frequently buy bread, they might also buy butter. The analysis finds that these two products are often purchased together.

2. Creating Association Rules:

- The main goal is to create association rules in the form of:
 - "If X happens, then Y happens".
- These rules help businesses understand which items tend to be bought together.
- Example: "If a customer buys a laptop (X), then they are likely to buy a mouse (Y)."
 - This rule can be used to recommend related products to customers.





Module 3: Machine Learning in AI

2 Machine Learning Model - Un-Supervised Machine Learning Algorithms

2. Association Analysis

Applications of Association Analysis:

1. Retail and E-commerce:

- Businesses use association analysis to discover which products are often bought together.
- Example: In a supermarket, if many customers buy diapers and baby wipes together, the store might place these items next to each other to increase sales.

2. Market Basket Analysis:

- Commonly used in retail, this technique finds associations between products purchased together.
- Example: "If a customer buys shampoo, they are likely to buy conditioner."

3. Recommender Systems:

- Used in online platforms (like Amazon or Netflix) to suggest products or movies based on what others with similar preferences have liked.
- Example: "Customers who bought this book also bought these other books."

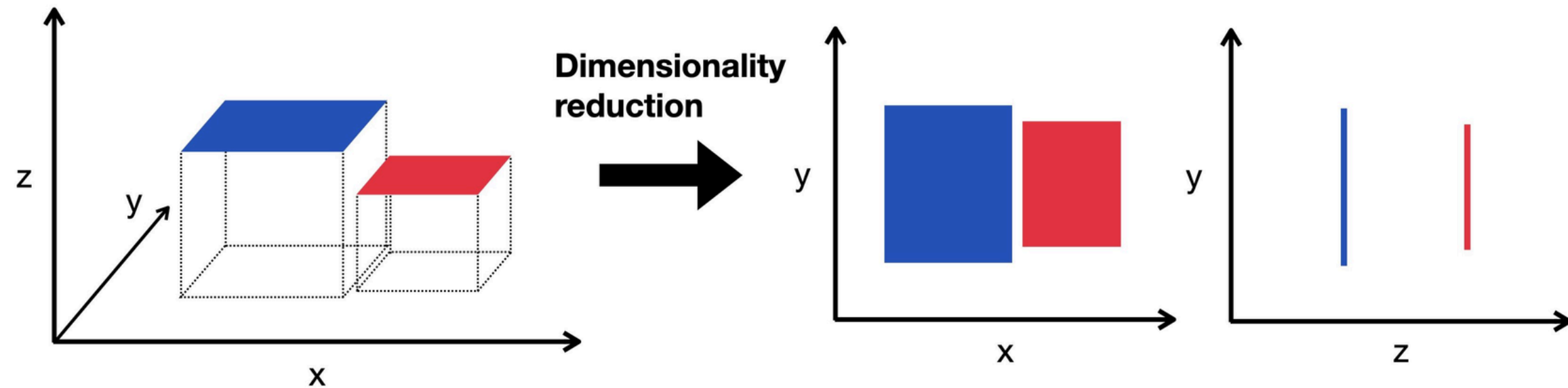




Module 3: Machine Learning in AI

2 Machine Learning Model - Un-Supervised Machine Learning Algorithms

3. Dimensionality reduction



Dimensionality Reduction is a technique used to simplify a dataset by reducing the number of features (variables) without losing important information. This is especially useful when you have datasets with a large number of features, often in the millions, making it difficult to process and analyze the data effectively.





Module 3: Machine Learning in AI

2 Machine Learning Model - Un-Supervised Machine Learning Algorithms

3. Dimensionality reduction

Why is Dimensionality Reduction Important?

1. Simplifies Data:

- When datasets have too many features (variables), it can be overwhelming and computationally expensive to work with them.
- Dimensionality reduction reduces the number of features, making the dataset more manageable.

2. Reduces Complexity:

- Fewer features make models easier to train, faster to process, and often lead to better performance.
- Reducing features also helps to avoid problems like overfitting, where a model is too complex and doesn't generalize well to new data.

3. Improves Visualization:

- With too many features, it's hard to visualize the data. Dimensionality reduction makes it easier to visualize the dataset in 2D or 3D.





Module 3: Machine Learning in AI

2 Machine Learning Model - Un-Supervised Machine Learning Algorithms

4. Outlier Detection

Outlier detection, also known as anomaly detection, is a technique used to find rare or unusual events in a dataset that do not follow the normal pattern. These events or observations are called outliers and can indicate important issues like fraud, errors, or new trends.

How Outlier Detection Works:

1. Identifying Anomalies:

- Outlier detection identifies data points that differ significantly from the rest of the data.
- Example: In credit card transactions, a sudden large withdrawal made in an unusual location could be flagged as an anomaly (possible fraud).

2. Clustering (KNN):

- K-Nearest Neighbors (KNN) is a common technique used to detect anomalies by checking the distance between data points. If a data point is far from its neighbors, it's considered an outlier.
- Example: In a KNN-based anomaly detection, if a customer transaction is much higher than the usual range, it may be flagged as suspicious.





Module 3: Machine Learning in AI

2 Machine Learning Model - Un-Supervised Machine Learning Algorithms

4. Outlier Detection

Applications of Outlier Detection:

1. Fraud Detection:

- In banking or credit card systems, detecting unusual transactions (like purchases in unexpected locations) is done using anomaly detection.

2. Healthcare:

- Detecting rare diseases or unusual symptoms in patients' data can be done using anomaly detection.

3. Manufacturing:

- In a production line, detecting defects or malfunctioning machinery through unusual patterns in data can prevent larger issues.





Module 3: Machine Learning in AI

2 Machine Learning Model - Semi-Supervised Learning

Semi-Supervised Learning is a type of machine learning that combines the strengths of supervised and unsupervised learning. It uses both labeled data (data with known answers) and unlabeled data (data without known answers) to train the machine. This approach helps to improve the accuracy of the model while requiring less labeled data than supervised learning alone.

How Semi-Supervised Learning Works:

1. Using Labeled and Unlabeled Data:

- The algorithm is provided with a small amount of labeled data (where the answers are known) and a large amount of unlabeled data (where the answers are unknown).
- The algorithm learns from both the labeled and the unlabeled data to improve its predictions.





Module 3: Machine Learning in AI

2 Machine Learning Model - Semi-Supervised Learning

Two Approaches to Semi-Supervised Learning:

1. Supervised Model + Unsupervised Model Approach:

- **Step 1:** Start by building a supervised model using a small amount of labeled data.
- **Step 2:** Apply the trained model to a large amount of unlabeled data to generate predictions.
- **Step 3:** Use the predicted labels from the unlabeled data to add more labeled data.
- **Step 4:** Iterate this process multiple times to improve the model's accuracy by generating more labeled data from the initial small set.
 - **Example:** If you have a small dataset of labeled images (e.g., some images labeled as "cat" or "dog"), the model is first trained using these images. Then, it can predict labels for a large set of unlabeled images, and you can use the predicted labels to expand your training data.





Module 3: Machine Learning in AI

2 Machine Learning Model - Semi-Supervised Learning

Two Approaches to Semi-Supervised Learning:

2. Unsupervised Clustering + Annotation Approach:

- **Step 1:** Use unsupervised learning techniques to group similar unlabeled data into clusters.
- **Step 2:** Annotate or label these clusters based on patterns observed.
- **Step 3:** Use the labeled clusters to train the model.
 - **Example:** In a set of unlabeled customer data, you can cluster customers with similar purchasing habits. Then, based on the cluster characteristics, you can assign labels to these groups and use that information to train the model.





Module 3: Machine Learning in AI

2 Machine Learning Model - Reinforcement Learning (RL)

Reinforcement Learning (RL) is a type of machine learning where an agent learns how to make decisions by interacting with an environment. Unlike supervised learning, where the model is trained on labeled data, RL involves an agent that learns by trial and error, with the goal of maximizing rewards and minimizing penalties.





Module 3: Machine Learning in AI

2 Machine Learning Model - Reinforcement Learning (RL)

How Reinforcement Learning Works:

1. Agent-Environment Interaction:

- An agent interacts with its environment by taking actions (decisions) and receiving feedback in the form of rewards or penalties.
- Example: In a game scenario, the agent might be a robot and the environment is the game world.

2. Rewards and Penalties:

- Rewards are given for performing correct actions, while penalties are given for incorrect actions.
- The agent's goal is to maximize its total reward over time by learning from experience.
- Example: If a robot picks the right path in the game (avoiding fire), it earns a reward (like a diamond). If it chooses the wrong path (touches fire), it loses some reward.

3. Trial and Error:

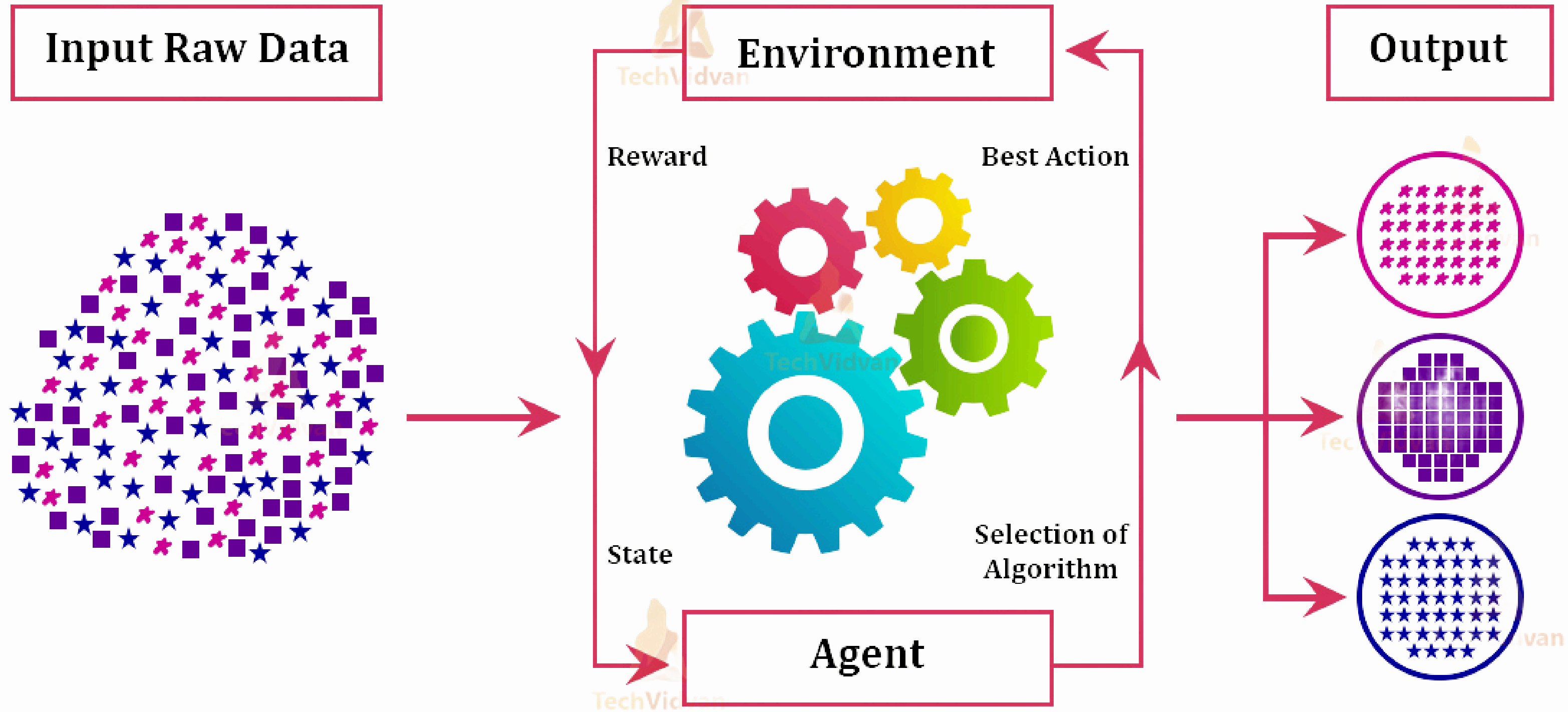
- The agent starts with no knowledge and learns by trying different actions, observing the results, and adjusting its strategy.
- Over time, the agent learns which actions lead to the highest rewards and which actions lead to penalties.





Module 3: Machine Learning in AI

2 Machine Learning Model - Reinforcement Learning (RL)





Module 3: Machine Learning in AI

2 Machine Learning Model - Reinforcement Learning (RL)

Key Features of Reinforcement Learning:

1. Input (State):

- The initial state from which the agent begins its learning.
- Example: In the game, the starting point of the robot is the initial state.

2. Output (Action):

- The action the agent can take in response to the current state.
- Example: The robot can choose to move forward, turn left, or turn right based on the environment.

3. Training Process (Learning from Experience):

- The agent learns by interacting with the environment and getting feedback. It keeps adjusting its actions based on whether it receives a reward or penalty.
- Example: If the robot's action leads to a reward, it will try to repeat that action next time; if it leads to a penalty, it will try to avoid it.

4. Maximizing Reward:

- The agent's goal is to maximize its total reward over time, finding the best path to reach the final goal.
- Example: The robot tries to find the best path to reach the diamond with the least number of hurdles (fires).





Module 3: Machine Learning in AI

2 Machine Learning Model - Reinforcement Learning (RL)

Reinforcement Learning vs. Supervised Learning:

- **Supervised Learning:**
 - The model learns from labeled data, where the correct answers are already provided.
 - Example: In supervised learning, you train the model with a dataset where the correct answers (labels) are known.
- **Reinforcement Learning:**
 - The agent learns from its own experience by receiving feedback in the form of rewards and penalties.
 - Example: The robot learns the best path to the diamond by experimenting and learning from each step.



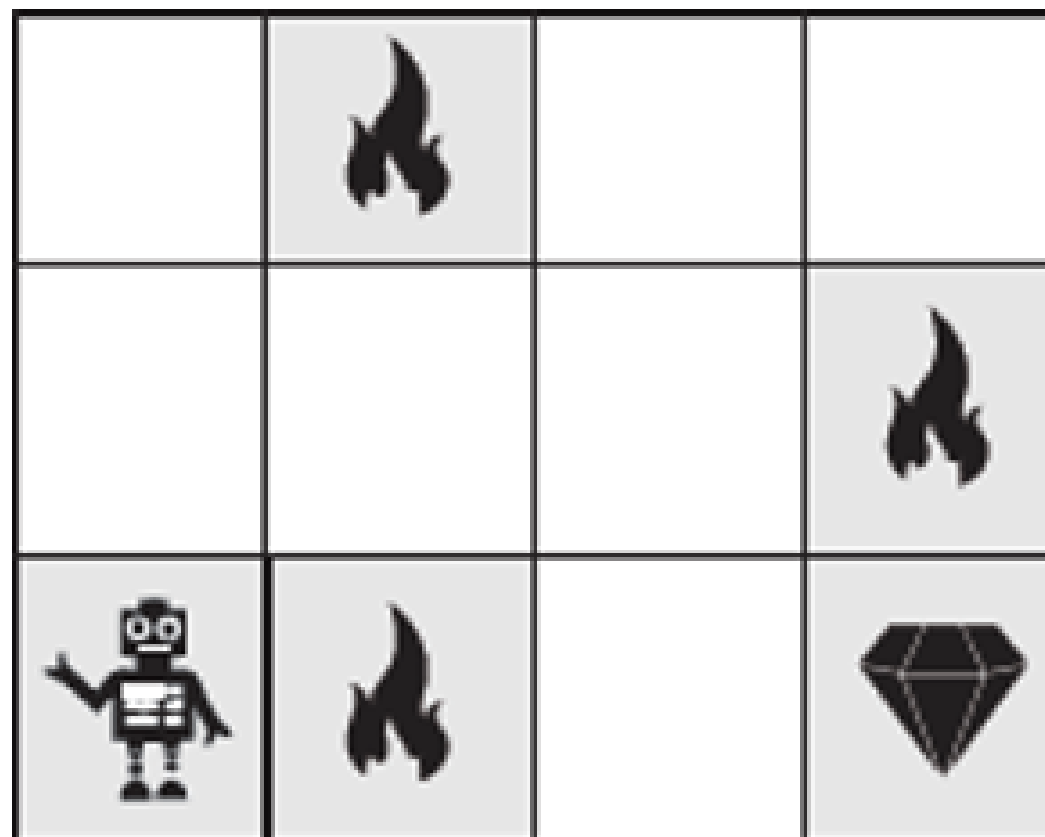


Module 3: Machine Learning in AI

2 Machine Learning Model - Reinforcement Learning (RL)

Example: Reinforcement Learning in a Game

- Scenario:



- A robot needs to reach a diamond (reward) while avoiding fires (penalties) in a game.
 - The robot starts at a random point and tries different paths.
 - Right path = Reward (diamond), Wrong path = Penalty (fire).
 - The robot learns over time by repeating the process and adjusting its behavior to maximize rewards and avoid penalties.
- Final Goal: The robot learns the best path to take in the game and successfully reaches the diamond, maximizing its total reward.





Module 3: Machine Learning in AI

2 Machine Learning Model - Reinforcement Learning (RL)

Types of Reinforcement in Reinforcement Learning

- In Reinforcement Learning (RL), reinforcement refers to the process of giving feedback to an agent based on its actions to encourage or discourage certain behaviors.
- There are two types of reinforcement:
 - Positive Type
 - Negative Type





Module 3: Machine Learning in AI

2 Machine Learning Model - Reinforcement Learning (RL)

1. Positive Reinforcement:

Positive Reinforcement is when a reward is given to the agent for performing a desired action. This increases the strength and frequency of the behavior, motivating the agent to repeat that action in the future.

- **How It Works:**

- When the agent does something right, it receives a positive reward (like points, praise, or a desirable outcome), encouraging the agent to continue performing that action.

- **Example:**

- Imagine a robot in a game. If the robot successfully avoids a fire and reaches the diamond (goal), it gets a reward (like 10 points or a success message). This positive feedback makes the robot more likely to avoid fires in future attempts.





Module 3: Machine Learning in AI

2 Machine Learning Model - Reinforcement Learning (RL)

2. Negative Reinforcement:

Negative Reinforcement is when a negative condition is removed or avoided after the agent performs the desired action. The removal of the negative condition strengthens the behavior and encourages the agent to repeat the action.

- **How It Works:**

- The agent performs an action to avoid a negative outcome or to stop an unpleasant event from happening.

- **Example:**

- In a driving simulator, if the agent (car) drives safely and avoids crashing into objects, it avoids a penalty (like losing points or a time delay). The removal of the penalty encourages the agent to drive safely in the future.





Module 3: Machine Learning in AI

2 Machine Learning Model - Reinforcement Learning (RL)

Type	Definition	Example	Effect on Behavior
Positive Reinforcement	Increases behavior by giving a reward for a correct action.	A robot gets a reward for avoiding a fire in a game.	Strengthens the behavior, encouraging it to be repeated.
Negative Reinforcement	Increases behavior by removing an unpleasant condition.	A car avoids a penalty by driving safely.	Strengthens the behavior by avoiding negative consequences.





Module 3: Machine Learning in AI

2 Machine Learning Model - Regression Analysis in Machine Learning

Regression Analysis is a statistical method used to study the relationship between a dependent (target) variable and one or more independent (predictor) variables. The primary goal of regression analysis is to predict the value of the dependent variable based on the given predictors.

How Regression Analysis Works:

1. Identifying Relationships:

- Regression analysis aims to determine how changes in the independent variables affect the dependent variable.
- Example: A company may want to find out how advertising expenditure affects sales.

2. Predicting Continuous Values:

- Unlike classification, which predicts discrete categories, regression predicts continuous values like price, salary, or temperature.
- Example: Predicting the price of a house based on its size, location, and age.





Module 3: Machine Learning in AI

2 Machine Learning Model - Regression Analysis in Machine Learning

Steps in Regression Analysis:

1. Data Collection:

- Collect data for the independent and dependent variables.

2. Model Creation:

- Create a mathematical equation that defines the dependent variable as a function of the independent variables.
- **Example:**
 - If Y is the price of a house, and X1 and X2 are size and location, then the model might be:
 - $Y = f(X1, X2)$.

3. Finding the Best Fit:

- A regression line or curve is plotted on a graph to find the best fit through the data points.
- The goal is to minimize the vertical distance between the data points and the regression line.

4. Evaluating the Model:

- After building the regression model, evaluate its accuracy by measuring how well it predicts new data.





Module 3: Machine Learning in AI

2 Machine Learning Model - Regression Analysis in Machine Learning

Types of Regression:

1. Linear Regression:

- The simplest form of regression where the relationship between the dependent and independent variables is assumed to be linear (a straight line).
- Equation for Linear Regression:
 - $Y=b_0+b_1X+e$
 - b_0 is the intercept, b_1 is the coefficient of the predictor variable (X), and e is the error term.
- Example:
 - A company can calculate sales based on advertising expenditure using a linear regression model.

2. Multiple Regression:

- When there are multiple independent variables, the equation extends to:
 - $Y=b_0+b_1X_1+b_2X_2+e$
 - This is used when there are multiple predictors.





Module 3: Machine Learning in AI

2 Machine Learning Model - Regression Analysis in Machine Learning

Examples of Regression Analysis:

- **Case Study 1 - Auto Fare Calculation:**

- **Problem:** The cost of an auto fare depends on a fixed charge plus a per-kilometer rate.

- **Linear Equation:**

- $y=11x+30$

- Where y is the total cost, x is the distance traveled (in km), and the fixed charge is 30.

- For 10 km, the fare would be:

- $y=11 \times 10 + 30 = 140$ Rs.

- **Case Study 2 - Monthly Rental Cost:**

- **Problem:** A company's rental cost is based on a fixed cost plus a per-employee charge.

- **Linear Equation:**

- $y=10000x+20000$

- Where x is the number of employees and y is the total rental cost.

- For 20 employees, the monthly rental is:

- $y=10000 \times 20 + 20000 = 220000$ Rs.





Module 3: Machine Learning in AI

2 Machine Learning Model - Regression Analysis in Machine Learning

Summary:

- Regression Analysis helps predict continuous values (e.g., price, salary) based on the relationship between dependent and independent variables.
- Linear Regression is the most common technique, but multiple regression, non-linear regression, and other techniques are used for more complex tasks.
- Applications include cost predictions, profit calculations, and sales forecasting.





Module 3: Machine Learning in AI

2 Machine Learning Model - Regression Analysis in Machine Learning

Model Evaluation Metrics for Regression

To evaluate the performance of a regression model, we focus on the prediction error and use various metrics.

1. Key Metrics:

1. Root Mean Squared Error (RMSE):

- Measures the difference between observed and predicted values.
- Lower RMSE indicates a better model.
- Formula: $RMSE = \sqrt{(\sum(\text{observed} - \text{predicted})^2 / n)}$

2. Adjusted R-Square (R^2):

- Represents the proportion of variation in the data explained by the model.
- Higher R^2 indicates a better model.





Module 3: Machine Learning in AI

2 Machine Learning Model - Regression Analysis in Machine Learning

Model Evaluation Metrics for Regression

2. Evaluation Methods:

1. Train-Test Split:

- Split data into 80% for training and 20% for testing to evaluate model performance.

2. K-Fold Cross-Validation:

- **Step 1:** Split the data into k subsets (e.g., 5 subsets for k=5).
- **Step 2:** Train the model on k-1 subsets, and test on the remaining subset.
- **Step 3:** Repeat for all subsets and calculate the average prediction error.
- Best model is the one with the lowest error.





Module 3: Machine Learning in AI

2 Machine Learning Model - Regression Analysis in Machine Learning

Types of Regression

Regression analysis involves different types of models, each suited for specific types of data and relationships between variables.

1. Linear Regression
2. Logistic regression
3. Ridge regression
4. Lasso (Least Absolute Shrinkage Selector Operator) regression
5. Polynomial regression
6. Stepwise regression
7. ElasticNet regression





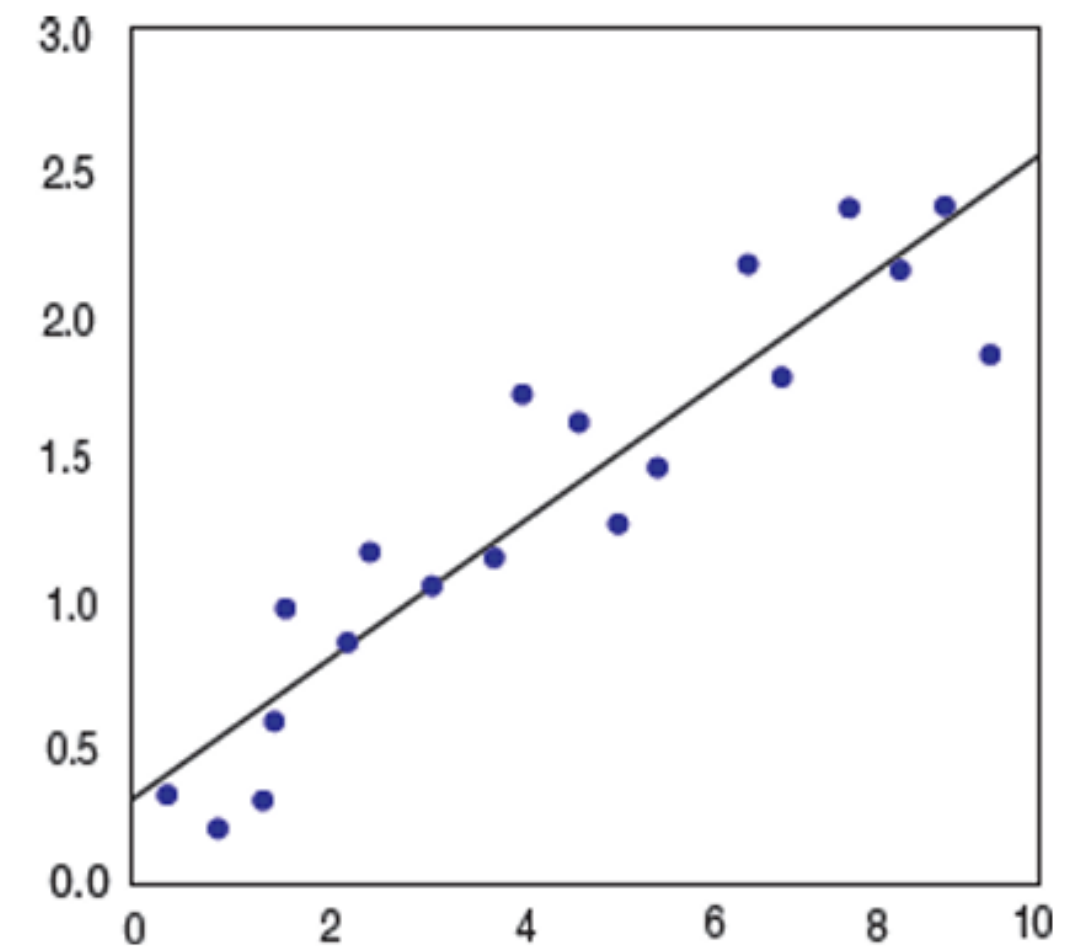
Module 3: Machine Learning in AI

2 Machine Learning Model - Regression Analysis in Machine Learning

Types of Regression

1. Linear Regression:

- **Use Case:** Predicting a continuous dependent variable with a linear relationship to one or more independent variables.
- **Example:** Predicting house prices based on size, location, and number of rooms.
- **Equation:** $Y=bX+CY$
 - Y is the dependent variable (e.g., house price).
 - X is the independent variable (e.g., size of the house).
 - b is the slope (relationship between X and Y).
 - C is the intercept.
- **Advantages:** Simple, fast, and easy to understand.
- **Limitations:** Can suffer from overfitting if data is too complex.



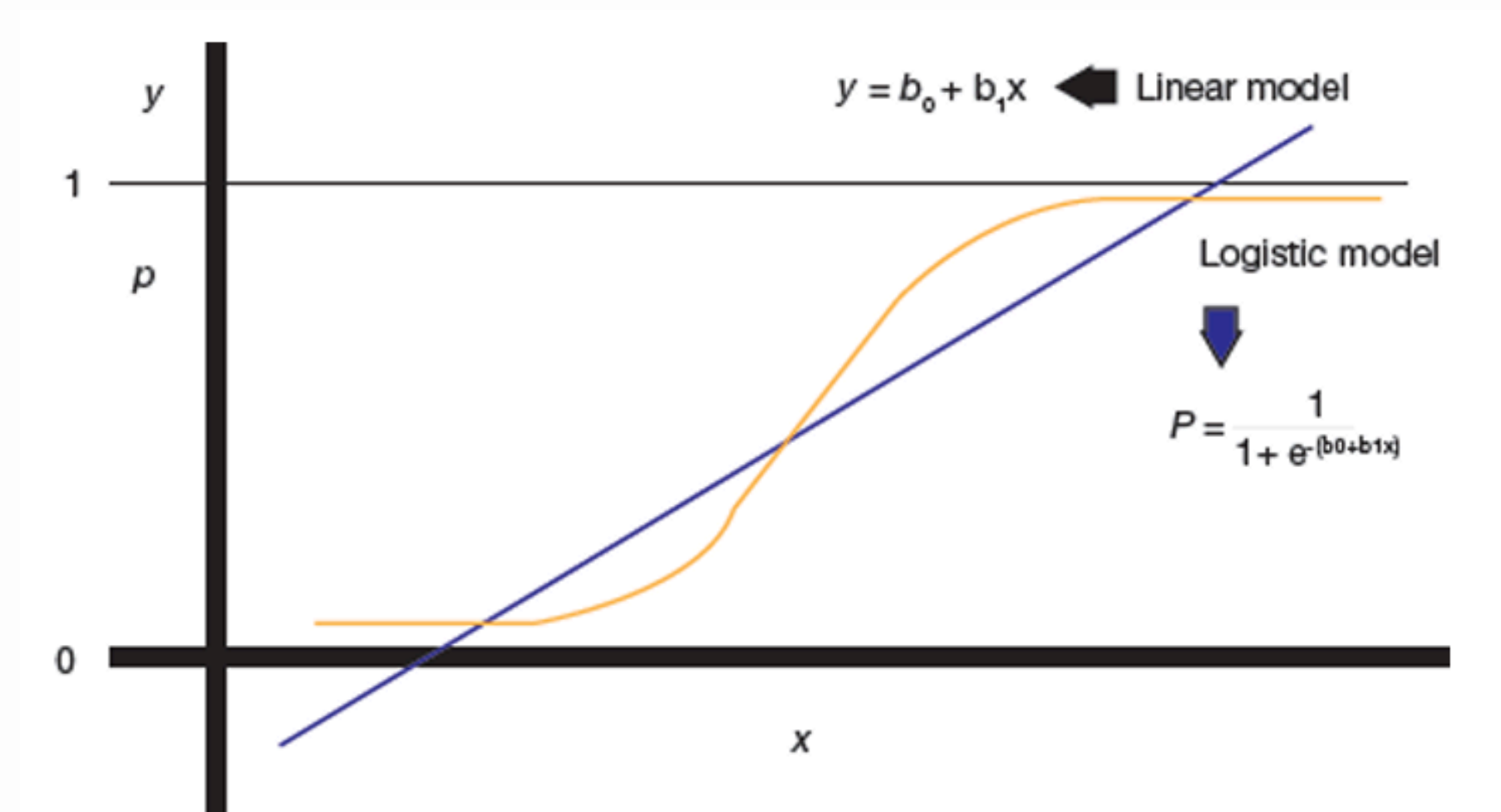


Module 3: Machine Learning in AI

2 Machine Learning Model - Regression Analysis in Machine Learning

2. Logistic Regression:

- **Use Case:** Predicting a binary outcome (yes/no, true/false) or the probability of an event occurring.
- **Example:** Predicting whether an email is spam or not spam based on certain features.
- **Equation:** The output is transformed using the logit function (log-odds).
- **Applications:** Categorical data, such as in classification tasks (e.g., disease diagnosis, fraud detection).





Module 3: Machine Learning in AI

2 Machine Learning Model - Regression Analysis in Machine Learning

3. Ridge Regression:

- **Use Case:** Used when there is multicollinearity (high correlation between independent variables) in the data. It helps prevent overfitting by reducing the size of coefficients.
- **How It Works:** It adds a penalty to the size of coefficients to reduce their impact and stabilize the model.
- **Example:** Predicting sales, where multiple variables are highly correlated (e.g., price, promotion, and distribution).





Module 3: Machine Learning in AI

2 Machine Learning Model - Regression Analysis in Machine Learning

4. Lasso Regression (Least Absolute Shrinkage and Selection Operator):

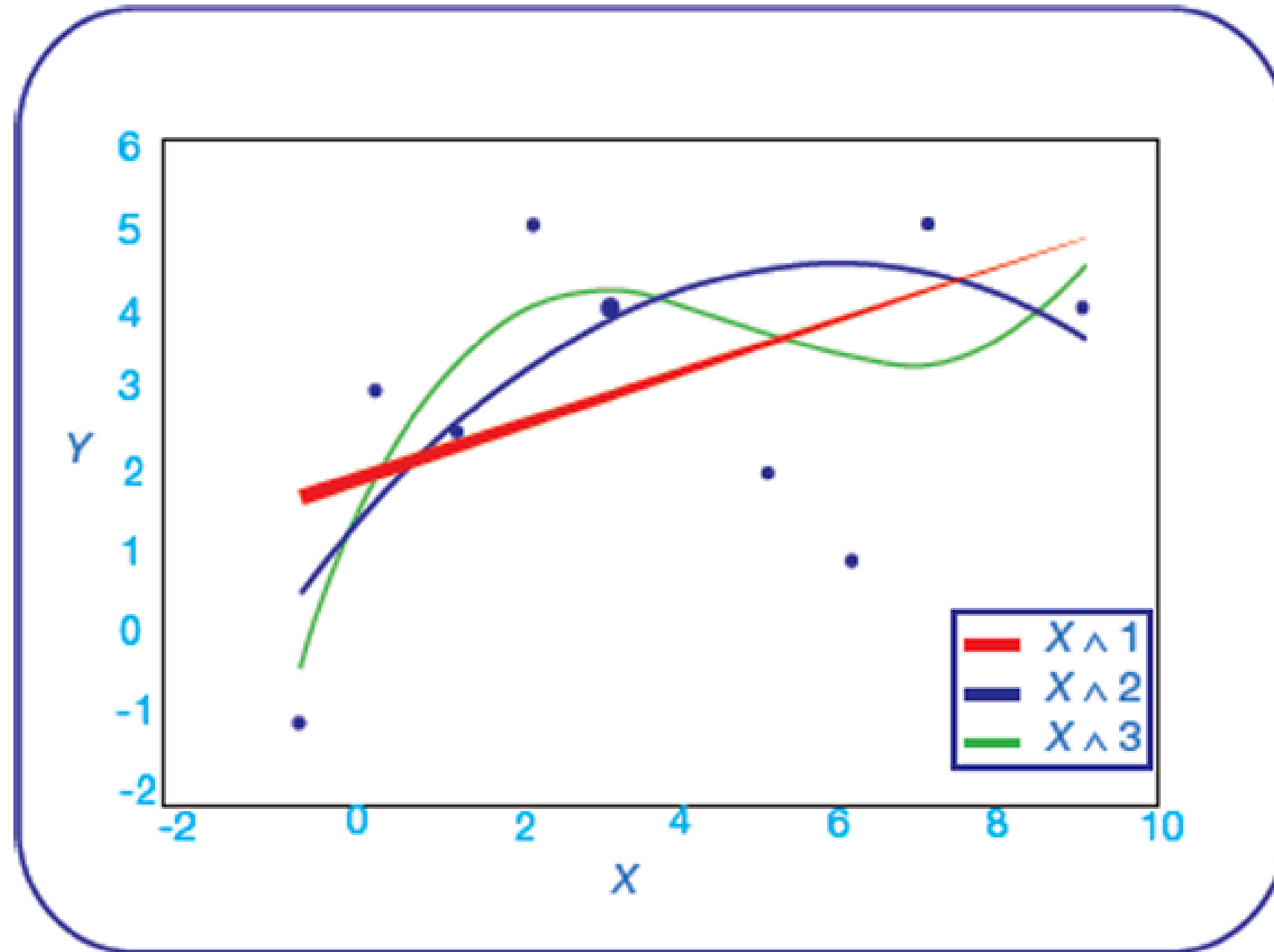
- **Use Case:** Used for variable selection and regularization in regression models.
- **How It Works:** Lasso reduces the coefficients of less important predictors to zero, effectively eliminating them from the model.
- **Example:** In a dataset with many features, Lasso can help select the most important ones, improving model efficiency.
- **Also Known As:** L1 regularization.





Module 3: Machine Learning in AI

2 Machine Learning Model - Regression Analysis in Machine Learning



5. Polynomial Regression:

- **Use Case:** When the relationship between the independent and dependent variables is non-linear.
- **Example:** Predicting temperature changes with respect to time, where the relationship is curved (e.g., seasonal variations).
- **Equation:**
 - $Y = b_0 + b_1X_1 + b_2X_2^2 + \dots + b_nX_n^n$
- **Best Fit:** This technique is used when data cannot be fit by a straight line, but a curve is needed.
- **Applications:** Modeling curvilinear data (e.g., trends in population growth).





Module 3: Machine Learning in AI

2 Machine Learning Model - Regression Analysis in Machine Learning

6. Stepwise Regression:

- **Use Case:** Builds the regression model by adding or removing variables step by step based on performance.
- **Approaches:**
 - **Forward Selection:** Starts with no variables and adds them one at a time.
 - **Backward Elimination:** Starts with all variables and removes them step by step.
 - **Bidirectional Elimination:** Combines both approaches.
- **Example:** Selecting the most significant predictors in a dataset for predicting sales growth.





Module 3: Machine Learning in AI

2 Machine Learning Model - Regression Analysis in Machine Learning

7. ElasticNet Regression:

- **Use Case:** A combination of Ridge and Lasso regression, especially useful when there are many predictors compared to observations.
- **How It Works:** ElasticNet combines the L1 penalty of Lasso and the L2 penalty of Ridge to create a more balanced model.
- **Example:** In Support Vector Machines (SVM) or document optimization, where there are many features but not enough data points.





Module 3: Machine Learning in AI

2 Machine Learning Model - Regression Analysis in Machine Learning

7. ElasticNet Regression:

- **Use Case:** A combination of Ridge and Lasso regression, especially useful when there are many predictors compared to observations.
- **How It Works:** ElasticNet combines the L1 penalty of Lasso and the L2 penalty of Ridge to create a more balanced model.
- **Example:** In Support Vector Machines (SVM) or document optimization, where there are many features but not enough data points.





Module 3: Machine Learning in AI

2 Machine Learning Model - Regression Analysis in Machine Learning

Type	Use Case	Key Feature	Example
Linear Regression	Predict continuous data with a linear relationship.	Simple, fast, used for straight-line relationships.	Predicting house prices based on size.
Logistic Regression	Predict binary outcomes or probabilities.	Used for classification tasks (yes/no, true/false).	Spam vs. non-spam email classification.
Ridge Regression	Handle multicollinearity in regression.	Adds a penalty to reduce overfitting.	Predicting sales with correlated features.
Lasso Regression	Variable selection and regularization.	Reduces coefficients to zero to eliminate irrelevant features.	Feature selection in a large dataset.
Polynomial Regression	Non-linear relationships between variables.	Fits a curve rather than a line.	Predicting temperature variations.
Stepwise Regression	Build models by adding/removing variables.	Automates variable selection based on model performance.	Selecting predictors for sales forecasting.
ElasticNet Regression	Combines Ridge and Lasso regression.	Balances L1 and L2 penalties for better performance.	Used in SVM and document optimization.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

- Classification is a type of machine learning task where the goal is to predict which category or class an observation belongs to.
- The model is trained using labeled data, where each input is already tagged with the correct class or category. Based on this, the model learns the patterns and uses them to predict the class of new, unseen data
- **How Classification Works:**
 - The classification model takes input data (features) and assigns it to one of the predetermined categories.
 - Example: In an email classification task, the model predicts whether an email is spam or not spam based on the features (like subject, sender, content, etc.).





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

Key Classification Algorithms:

- **Decision Trees:**
 - **How It Works:** Decision trees split the data into branches based on feature values, creating a tree-like structure. Each leaf node represents a predicted class.
 - **Example:** A decision tree might classify an email as spam or not spam by asking questions like "Is the sender known?" and "Does the subject contain certain keywords?"
- **Random Forest:**
 - **How It Works:** A random forest is an ensemble of multiple decision trees. Each tree gives a prediction, and the final class is determined by a majority vote.
 - **Example:** Random forests are often used for more accurate classification, such as predicting whether a customer will buy a product based on their demographic information.
- **K-Nearest Neighbors (KNN):**
 - **How It Works:** KNN classifies data points based on the majority class of their nearest neighbors in the feature space.
 - **Example:** If a new email has features similar to several spam emails, KNN will classify it as spam.
- **Support Vector Machines (SVM):**
 - **How It Works:** SVM finds the hyperplane that best separates the data into different classes. It maximizes the margin between the classes to ensure the best separation.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

Key Classification Algorithms:

- **Naive Bayes:**
 - **How It Works:** Naive Bayes is based on Bayes' Theorem and assumes that features are independent. It calculates the probability of each class and assigns the class with the highest probability.
 - **Example:** Naive Bayes is often used for text classification, such as classifying news articles into topics like sports, politics, or technology.
- **Logistic Regression:**
 - **How It Works:** Despite its name, logistic regression is a classification algorithm used to predict the probability of a binary outcome (yes/no, 0/1).
 - **Example:** Predicting whether a customer will buy a product (1) or not buy (0) based on features like age, income, and browsing history.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

Summary:

- Classification algorithms are used to predict the category or class of new data based on learned patterns.
- Examples of widely used classification techniques include:
 - **Decision Trees:** Splits data based on feature values.
 - **Random Forest:** Ensemble of decision trees.
 - **K-Nearest Neighbors (KNN):** Classifies based on nearest neighbors.
 - **Support Vector Machines (SVM):** Finds a hyperplane to separate classes.
 - **Naive Bayes:** Uses probabilities for classification.
 - **Logistic Regression:** Used for binary classification.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

1 K-Nearest Neighbors (KNN)

- The K-Nearest Neighbors (KNN) algorithm is a supervised learning method that classifies data points based on the similarity of nearby data points. It is used for classification and regression tasks.
- The KNN algorithm works by finding the k nearest neighbors to a new data point and making predictions based on the majority class or average value of those neighbors.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

1 K-Nearest Neighbors (KNN)

How KNN Algorithm Works:

1. Data Points and Labels:

- KNN requires labeled data (training data) to classify new, unseen data points.
- It looks at the 'k' nearest neighbors in the training data and assigns a class or predicts a value based on those neighbors.

2. Distance Measure:

- The distance between data points is calculated using a metric like Euclidean distance.
- **Euclidean Distance Formula:**

$$\text{Distance} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

- This helps in determining the closeness of the new data point to the training data points.=





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

1 K-Nearest Neighbors (KNN)

Steps in KNN Algorithm:

1. Choose 'k' Value:

- 'k' is the number of nearest neighbors to consider. For example, $k=3$ means looking at the three closest neighbors to classify a new data point.

2. Calculate Distance:

- Measure the distance between the new data point and all other data points in the training set.

3. Find Nearest Neighbors:

- Identify the 'k' closest data points to the new data point.

4. Classify or Predict:

- **For classification:** Assign the class that is the most common among the 'k' neighbors.
- **For regression:** Calculate the average of the 'k' neighbors' values.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

1 K-Nearest Neighbors (KNN)

Choosing the Appropriate 'k' Value:

- Small 'k' (e.g., $k=1$):
 - Higher variance: Model might be more sensitive to noise.
 - Example: A single neighbor determines the class, which could be misleading if there are outliers.
- Large 'k' (e.g., $k=10$):
 - Lower variance: More stable predictions, but could ignore smaller patterns.
 - Example: More neighbors help smooth out the prediction, but smaller or rare patterns might be missed.
 -





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

1 K-Nearest Neighbors (KNN)

Example of KNN Algorithm:

Example 1: Classifying Students

- **Data:** A table with students' academic score and extracurricular score.
- **New Student:** Academic Score = 7, EC Score = 6.
- **K=3:** Find the 3 closest students and classify based on majority.
- Calculate distance (using Euclidean formula) between the new student and all other students.
- **Result:** The new student belongs to the group with highest number of similar scores.

Student	Academic Score	EC Score
A	8	7
B	6	5
C	9	8
D	5	4





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

1 K-Nearest Neighbors (KNN)

Step 1: Calculate Euclidean Distance

The Euclidean distance between two points (x_1, y_1) and (x_2, y_2) $Distance = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$

- Where:
 - (x_1, y_1) is the new student's data (7, 6).
 - (x_2, y_2) is the data of each of the other students.
- Now, let's calculate the distance between the new student and each of the existing students (A, B, C, and D):

Student	Academic Score	EC Score
A	8	7
B	6	5
C	9	8
D	5	4





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

1 K-Nearest Neighbors (KNN)

Student	Academic Score	EC Score
A	8	7
B	6	5
C	9	8
D	5	4

Distance between New Student (7, 6) and Student A (8, 7):

$$\text{Distance}_A = \sqrt{(8 - 7)^2 + (7 - 6)^2} = \sqrt{1^2 + 1^2} = \sqrt{2} \approx 1.41$$

Distance between New Student (7, 6) and Student B (6, 5):

$$\text{Distance}_B = \sqrt{(6 - 7)^2 + (5 - 6)^2} = \sqrt{(-1)^2 + (-1)^2} = \sqrt{2} \approx 1.41$$

Distance between New Student (7, 6) and Student C (9, 8):

$$\text{Distance}_C = \sqrt{(9 - 7)^2 + (8 - 6)^2} = \sqrt{2^2 + 2^2} = \sqrt{8} \approx 2.83$$

Distance between New Student (7, 6) and Student D (5, 4):

$$\text{Distance}_D = \sqrt{(5 - 7)^2 + (4 - 6)^2} = \sqrt{(-2)^2 + (-2)^2} = \sqrt{8} \approx 2.83$$





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

1 K-Nearest Neighbors (KNN)

Step 2: Sort the Distances

Now that we have calculated the distances, we can sort them in increasing order:

1. Distance to Student A: 1.41
2. Distance to Student B: 1.41
3. Distance to Student C: 2.83
4. Distance to Student D: 2.83

Student	Academic Score	EC Score
A	8	7
B	6	5
C	9	8
D	5	4

Step 3: Find the 3 Nearest Neighbors (K=3)

The 3 closest neighbors (with the smallest distances) are:

- Student A (Distance: 1.41)
- Student B (Distance: 1.41)
- Student C (Distance: 2.83)





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

1 K-Nearest Neighbors (KNN)

Step 4: Classify the New Student

Now, we classify the new student based on the majority class of the 3 nearest neighbors.

- Student A belongs to Group 1 (let's say it's "Outstanding").
- Student B belongs to Group 2 (let's say it's "Sporty").
- Student C belongs to Group 1 (let's say it's "Outstanding").

Student	Academic Score	EC Score
A	8	7
B	6	5
C	9	8
D	5	4

Majority Class:

- Group 1 (Outstanding) appears twice (from Student A and Student C).
- Group 2 (Sporty) appears once (from Student B).

Final Classification:

Since Group 1 (Outstanding) is the majority class among the 3 nearest neighbors, the new student is classified as "Outstanding".





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

1 K-Nearest Neighbors (KNN)

Pros and Cons of KNN Algorithm:

Pros:

1. **Simple and Easy to Understand:** KNN is a simple algorithm that doesn't require much training time.
2. **No Assumptions:** It makes no assumptions about the underlying data distribution.
3. **Works for Both Classification and Regression:** It can be applied to both types of tasks.
4. **Good for Multi-class Problems:** It works well when there are more than two classes.

Cons:

1. **High Prediction Time:** For large datasets, predicting for new data points can be slow since the algorithm needs to compare it to all the training data.
2. **Sensitive to Data Scaling:** If the data features are on different scales, the algorithm may be biased. Data needs to be standardized.
3. **Memory Intensive:** KNN stores all the training data, requiring large memory usage.
4. **Outlier Sensitivity:** KNN can be affected by noisy data and outliers.
5. **Not Ideal for High-Dimensional Data:** KNN may not perform well when the number of features (dimensions) increases significantly.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

2. Decision Trees

A Decision Tree is a supervised learning algorithm used for classification and regression tasks. It works by creating a tree-like structure that splits data based on different features to make predictions.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

2. Decision Trees

How Decision Trees Work:

1. Root Node:

- This is the starting point of the tree where the first split happens. The tree will choose the best feature to split the data based on a specific criterion.

2. Decision Nodes:

- These nodes represent the decision points where data is split into branches based on conditions.

3. Leaf/Terminal Nodes:

- The leaf nodes contain the final prediction or class for the data (e.g., "Yes" or "No", "Cancerous" or "Benign").

4. Branches:

- These are the edges that connect nodes, showing the decision rules (e.g., "Is Age > 30?").





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

2. Decision Trees

Steps to Build a Decision Tree:

1. Select a Feature to Split:

- The first decision is to select the feature that best separates the data into different classes (e.g., Age, Salary, etc.).

2. Split the Data:

- Split the data based on the chosen feature, creating two or more sub-nodes. Each sub-node represents a further division of the data.

3. Repeat:

- Continue splitting the data at each node using the most relevant feature until the data cannot be divided further or a stopping criterion is met.

4. Assign Class to Each Leaf:

- Once the tree reaches a leaf node, assign the class based on the majority class of the data at that leaf.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

2. Decision Trees

Example of Decision Tree:

- Task: Predict whether a person will buy a product based on age and income.
- Step 1: Root Node: Split based on Age.
 - If Age > 40, predict Yes (Person C).
 - If Age ≤ 40, go to the next node.
- Step 2: Decision Node: Split based on Income for people aged ≤ 40.
 - If Income > 40K, predict Yes (Person A).
 - If Income ≤ 40K, predict No (Person B).
 - Leaf Nodes: The leaf nodes at the end of each branch contain the predictions ("Yes" or "No").

Person	Age	Income	Bought Product
A	30	50K	Yes
B	22	30K	No
C	45	70K	Yes
D	60	20K	No





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

2. Decision Trees

Types of Decision Trees:

1. Classification Trees:

- Used when the target variable is categorical (e.g., "Yes" or "No").
- Example: Predicting whether an email is spam or not spam.

2. Regression Trees:

- Used when the target variable is continuous (e.g., predicting house prices).
- Example: Predicting the price of a house based on features like size, location, and age.





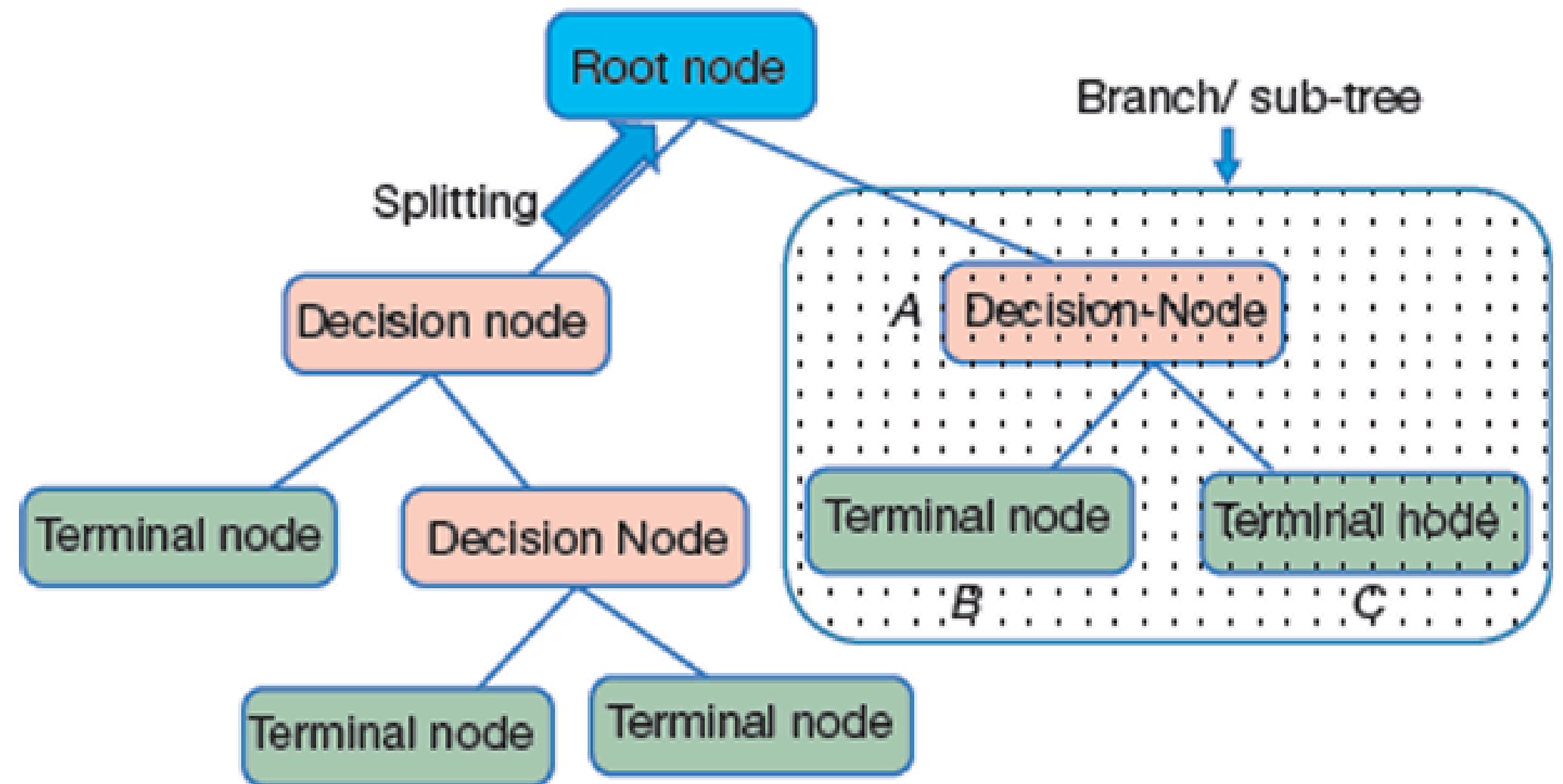
Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

2. Decision Trees

Key Terminology in Decision Trees:

1. **Root Node:** The first node in the tree.
2. **Terminal (Leaf) Node:** Nodes where the final decision or classification is made.
3. **Branches:** Edges connecting nodes, showing decisions or rules.
4. **Splitting:** Dividing the data into sub-groups based on features.
5. **Parent Node:** A node that splits into sub-nodes (has children).



Note:- A is parent node of B and C





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

2. Decision Trees

Advantages of Decision Trees:

1. **Easy to Understand and Implement:** Visual and interpretable model.
2. **No Need for Data Scaling:** It works with both numerical and categorical data.
3. **Handles Missing Data:** Some decision tree algorithms can deal with missing data.
4. **Non-Linear Relationships:** Can model complex, non-linear relationships.

Disadvantages of Decision Trees:

1. **Overfitting:** Decision trees can easily overfit the data, especially if the tree is too deep.
 - Solution: Pruning is used to trim the tree and prevent overfitting.
2. **Instability:** A small change in the data can lead to a completely different tree structure.
3. **Bias Toward Features with More Levels:** If a feature has many possible values, it can dominate the splitting process.
 - Solution: Random Forests (ensemble method) can be used to handle this issue.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

2. Decision Trees

Pruning the Tree:

- Pruning removes branches that provide little predictive power to avoid overfitting.
- **Goal:** Simplify the tree and make it more general.

Real-World Applications of Decision Trees:

1. Medical Diagnosis: Predicting whether a tumor is cancerous or benign.
2. Finance: Predicting whether a loan will be approved or denied.
3. Marketing: Predicting if a customer will buy a product based on demographics.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

3. Random Forests

Random Forests are an ensemble learning method used for both classification and regression tasks. They combine multiple decision trees to improve the overall performance by reducing the overfitting issue seen in individual decision trees.

Steps in Random Forest:

1. **Step 1:** Randomly sample K cases from the dataset for training each decision tree.
2. **Step 2:** Select m features from p available features at each node.
3. **Step 3:** Grow each tree as deep as possible, without pruning.
4. **Step 4:** For each new data point, each tree gives a prediction.
5. **Step 5:** Voting or Averaging to decide the final prediction.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

3. Random Forests

How Random Forest Works:

1. Bootstrap Sampling:

- From the dataset, random samples (with replacement) are taken to train multiple decision trees. This is similar to bagging (Bootstrap Aggregating), but random forests introduce another layer of randomness.

2. Random Feature Selection:

- For each decision tree, a subset of features (m predictors) is randomly selected at each node. The best split is found from these m features. This prevents trees from being overly correlated with each other.
- Difference from Bagging: In bagging, all features are considered at every split, but in random forests, only a random subset is used.

3. Tree Growth:

- Each decision tree is grown to its full depth without pruning, ensuring a more complex model with more diversity in trees.

4. Final Prediction:

- For classification, each tree in the forest votes on the class, and the class with the most votes is chosen as the final prediction.
- For regression, the predictions of all trees are averaged to get the final prediction.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

3. Random Forests

Example of Random Forest:

- **Task:** Predicting whether a customer will buy a product based on their age, income, and browsing history.
 - Step 1:** Randomly select subsets of the data and train decision trees on each.
 - Step 2:** For each decision tree, a random subset of features (e.g., age and income) is used at each node.
 - Step 3:** Grow all trees to full depth.
 - Step 4:** Each tree gives a prediction: Buy or Not Buy.
 - Step 5:** The majority vote is taken, and the final prediction is made.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

3. Random Forests

Advantages of Random Forests:

1. Handles Missing Data:

- Random forests are effective at estimating missing values in the dataset.

2. Resilient to Overfitting:

- Even with noisy or imbalanced data, random forests maintain high accuracy.

3. Works with Large Datasets:

- It can handle large datasets with many features (dimensions).

4. Flexible:

- Can be used for both classification and regression tasks.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

3. Random Forests

Limitations of Random Forests:

1. Overfitting on Noisy Data:

- Although less prone to overfitting than a single decision tree, random forests can still overfit if the data is very noisy, especially in regression tasks.

2. Slow Prediction Time:

- Since multiple decision trees are constructed and predictions are made by majority vote, it can be slow when generating predictions, especially on large datasets.

3. Interpretability:

- Random forests are harder to interpret compared to a single decision tree, as they involve many trees with many decisions.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

4. Clustering Techniques

- Clustering is a machine learning technique used to group similar data points together based on their features.
- The goal is to partition the data into clusters where:
 - Data points within the same cluster are similar to each other.
 - Data points in different clusters are dissimilar.
- This helps identify meaningful patterns in data, and is used in various applications like healthcare, business, and marketing.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

4. Clustering Techniques

Overview of Clustering Techniques:

- There are three main types of clustering algorithms, each suited to different types of data:
 - a. Partitional Clustering
 - b. Hierarchical Clustering
 - c. Density-Based Clustering





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

4. Clustering Techniques

1. Partitional Clustering:

- Partitional clustering divides the dataset into non-overlapping groups (clusters) where each data point belongs to exactly one cluster.
- **Example Algorithms:**
 - **K-Means:** Divides data into k clusters where k is predefined by the user.
 - **K-Medoids:** Similar to K-Means but chooses actual data points as the center of each cluster (medoids).
- **Advantages:**
 - Works well when clusters are spherical in shape.
 - Scalable and efficient with large datasets.
- **Limitations:**
 - Struggles with complex shapes (non-spherical clusters).
 - Cannot handle clusters of different densities.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

4. Clustering Techniques

2. Hierarchical Clustering:

- Hierarchical clustering builds a tree-like structure (dendrogram) that represents the hierarchy of data clusters.
 - **Agglomerative (Bottom-up):** Starts with individual data points and merges them into clusters.
 - **Divisive (Top-down):** Starts with all data points in one cluster and splits them into smaller clusters.
- **Advantages:**
 - Shows relationships between data points at different levels.
 - The resulting clusters are easy to interpret.
- **Limitations:**
 - Computationally expensive, especially with large datasets.
 - Sensitive to noise and outliers.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

4. Clustering Techniques

3. Density-Based Clustering:

- Clusters are formed based on the density of data points in a region. It doesn't require you to specify the number of clusters beforehand.
- Popular Algorithms:
 - **DBSCAN (Density-Based Spatial Clustering of Applications with Noise)**: Forms clusters where data points are close enough to each other, and handles outliers (points that don't belong to any cluster).
 - **OPTICS (Ordering Points to Identify the Clustering Structure)**: Similar to DBSCAN but can handle varying densities of clusters.
-





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

4. Clustering Techniques

3. Density-Based Clustering:

- Clusters are formed based on the density of data points in a region. It doesn't require you to specify the number of clusters beforehand.
- **Popular Algorithms:**
 - **DBSCAN (Density-Based Spatial Clustering of Applications with Noise):** Forms clusters where data points are close enough to each other, and handles outliers (points that don't belong to any cluster).
 - **OPTICS (Ordering Points to Identify the Clustering Structure):** Similar to DBSCAN but can handle varying densities of clusters.
- **Advantages:**
 - Works well with non-spherical shapes and outliers.
 - No need to specify the number of clusters.
- **Limitations:**
 - Not ideal for high-dimensional data.
 - Struggles with clusters of different densities.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

4. Clustering Techniques

Clustering Type	Description	Advantages	Limitations
Partitional Clustering	Divides data into non-overlapping groups (e.g., K-Means).	Works well with spherical clusters; scalable with large data.	Struggles with complex shapes or varying densities.
Hierarchical Clustering	Builds a tree-like structure showing relationships.	Results are easy to interpret; shows relationships at various levels.	Computationally expensive; affected by noise and outliers.
Density-Based Clustering	Forms clusters based on density (e.g., DBSCAN).	Can handle non-spherical clusters and outliers; no need to specify clusters.	Struggles with high-dimensional data and varying densities.

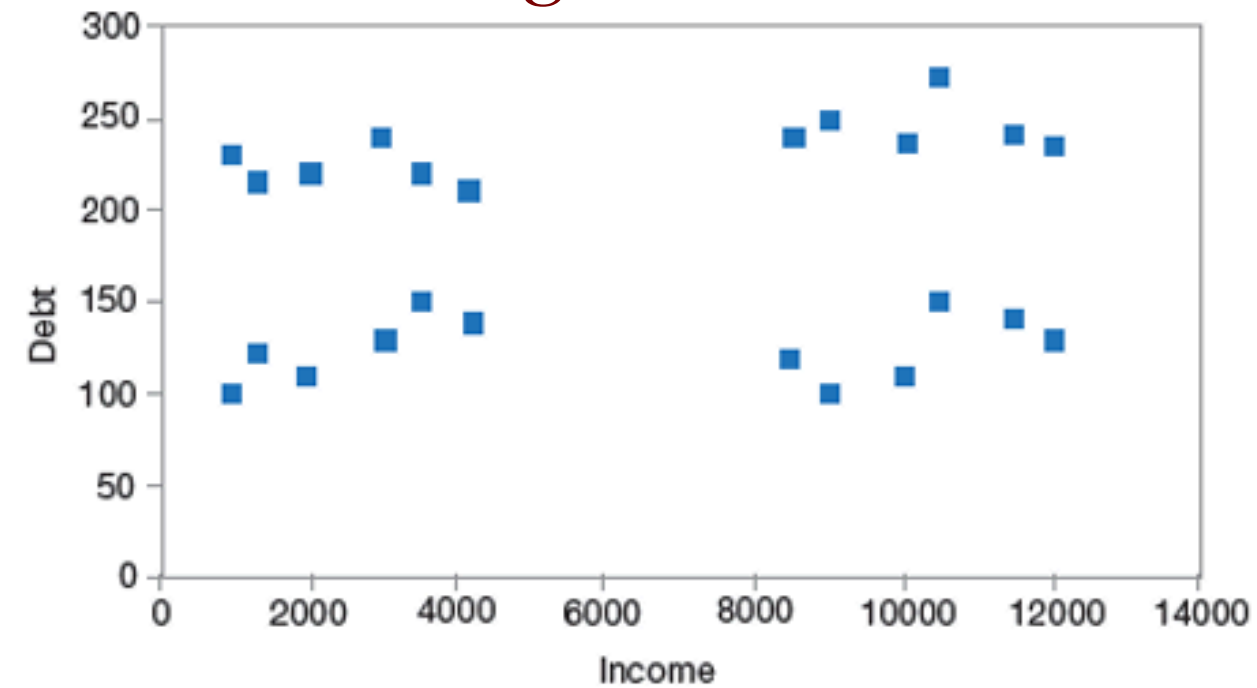




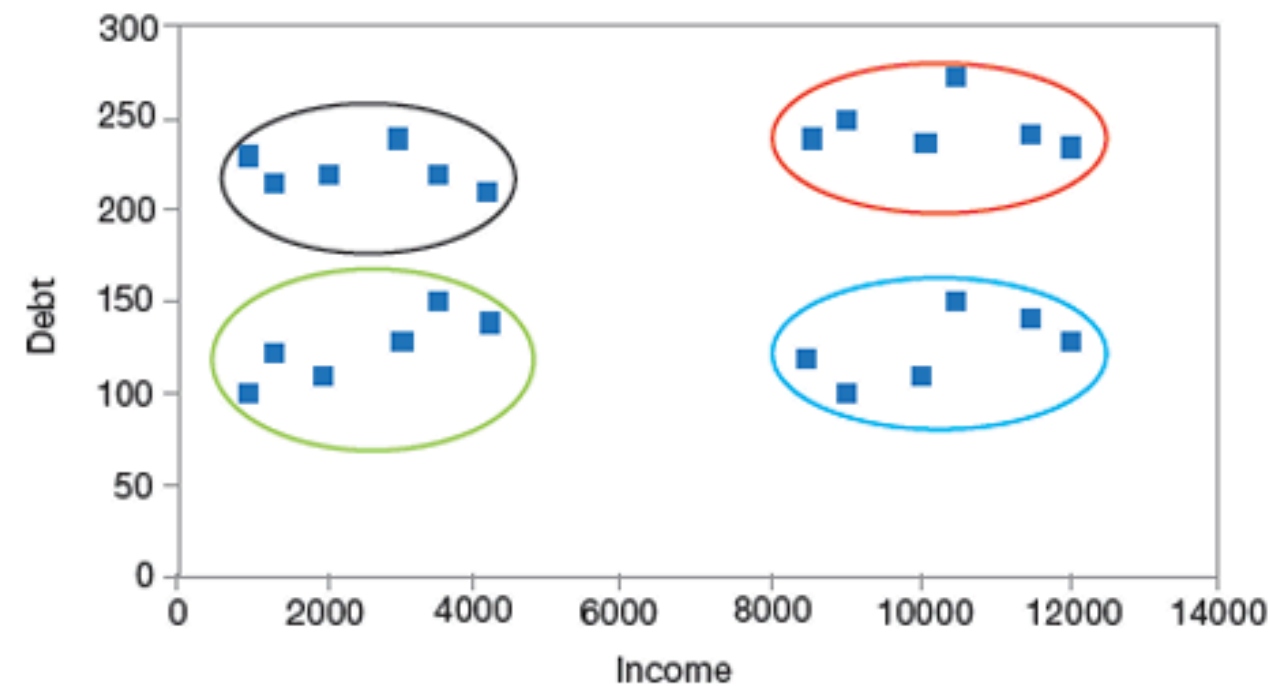
Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

5. K-Means Algorithm



K-Means is a popular unsupervised learning algorithm used for clustering. It groups data into k clusters where each cluster contains data points that are more similar to each other than to those in other clusters.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

5. K-Means Algorithm

How K-Means Works:

1. Choose the Number of Clusters (k):

a. First, decide how many clusters you want (k). For example, $k = 3$.

2. Initialize Centroids:

a. Randomly pick k points in the dataset as the initial centroids (the center of each cluster).

3. Assign Points to Clusters:

a. For each data point, calculate the Euclidean distance from each centroid and assign the point to the nearest centroid.

4. Update Centroids:

a. After all points are assigned to clusters, recalculate the centroid of each cluster by finding the mean of all the points in that cluster.

5. Repeat:

a. Repeat the process of assigning points to clusters and updating the centroids until the centroids stop changing (convergence is achieved).





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

5. K-Means Algorithm

Example of K-Means Algorithm:

Let's say we have a dataset of customers with two features: Income and Debt

We want to group customers into 2 clusters (k=2).

Customer	Income	Debt
A	50K	10K
B	40K	15K
C	70K	25K
D	90K	20K
E	60K	30K

Step-by-Step Process:

1. Update Centroids:

Recalculate the centroids:

- Cluster 1 (near A): Average of A, B, E → New centroid at 55K, 15K
- Cluster 2 (near D): Average of C, D → New centroid at 80K, 22K

2. Reassign Points to New Clusters:

- Recalculate distances using the new centroids and reassign customers to clusters.
- Repeat the process until the centroids no longer change.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

5. K-Means Algorithm

Properties of Clusters:

- **Intra-Cluster Similarity:** Data points within a cluster should be as similar as possible.
- **Inter-Cluster Dissimilarity:** Data points in different clusters should be as different as possible.

Real-World Applications of Clustering:

- **Customer Segmentation:** Group customers by buying behavior for targeted marketing.
- **Image Segmentation:** Group pixels into segments for image processing.
- **Recommendation Systems:** Suggest products based on similar user preferences.
- **Document Clustering:** Group similar documents for better organization.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

5. K-Means Algorithm

K-Means Algorithm Evaluation Metrics:

- 1. Inertia:** Measures the sum of squared distances between data points and their respective centroids. Lower inertia indicates better clustering.
- 2. Dunn Index:** Measures the inter-cluster distance and intra-cluster distance. Higher values indicate better clusters.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

5. K-Means Algorithm

Pros of K-Means Algorithm:

1. **Simple and Fast:** Easy to implement and runs efficiently.
2. **Scalable:** Works well with large datasets.
3. **Works Well for Spherical Clusters:** Performs best when clusters are circular/spherical.

Cons of K-Means Algorithm:

1. **Needs Predefined 'k':** The number of clusters (k) must be specified in advance.
2. **Sensitive to Outliers:** Outliers can distort the centroids.
3. **Random Initialization:** The final clusters can vary based on the initial centroids.
4. **Non-Optimal for Non-Spherical Clusters:** Struggles with clusters of different shapes or densities.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

5. K-Means Algorithm

Example Use Case:

Imagine you work in marketing for an e-commerce store, and you have a dataset of customer purchases. By using K-Means clustering, you could group customers based on their buying habits (e.g., high spenders vs. low spenders), and then tailor marketing campaigns for each group.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

6. Naïve Bayes Classification

Naïve Bayes is a probabilistic machine learning algorithm based on Bayes' Theorem. It is simple, fast, and widely used for classification tasks, such as spam detection, sentiment analysis, and text classification.

Key Concepts:

1. Bayes' Theorem:

- Bayes' Theorem helps calculate the probability of an event occurring given prior knowledge about related events.
- It is the foundation of Naïve Bayes classification.

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)}$$

- **P(A|B)**: The probability of A occurring given that B is true (posterior probability).
- **P(B|A)**: The probability of B occurring given that A is true (likelihood).
- **P(A)**: The probability of A occurring (prior probability).
- **P(B)**: The probability of B occurring (evidence).

○





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

6. Naïve Bayes Classification

Understanding Conditional Probability:

To understand Naïve Bayes, consider the following conditional probability example:

Example:

- If you want to know the probability of getting a Queen of Spades in a deck of cards:
 - There are 52 cards in total.
 - There are 13 Spades in total, and 1 of them is the Queen.
- So, the probability of drawing a Queen of Spades is:

$$P(\text{Queen of Spades}) = \frac{1}{13}$$

- This is an example of conditional probability, where we find the probability of an event (Queen of Spades) given a condition (it's a spade).





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

6. Naïve Bayes Classification

Naïve Bayes Assumption:

- Naïve means that we assume all features (variables) are independent of each other.
- In reality, features are often dependent, but the "naïve" assumption simplifies the calculation, which makes the model faster and easier to implement.
 - For example, in spam email detection, we treat the presence of each word as independent, though in reality, some words may often appear together





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

6. Naïve Bayes Classification

How Naïve Bayes Works:

1. Given a dataset with labeled categories (e.g., spam or not spam emails), we calculate the probability of each feature (word) belonging to each class (spam or not spam).
2. For a new email, we compute the probabilities of it being spam or not spam based on its features (words) and classify it to the class with the highest probability.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

6. Naïve Bayes Classification

Example of Naïve Bayes in Action:

Suppose we want to classify whether an email is spam or not spam based on the words "offer", "money", and "free".

- **Step 1:** Calculate the probability of each word appearing in spam and not spam emails. For example:
 - $P(\text{offer}|\text{spam})$ = Probability of "offer" appearing in spam emails.
 - $P(\text{offer}|\text{not spam})$ = Probability of "offer" appearing in non-spam emails.
- **Step 2:** Using Bayes' Theorem, calculate the probability that the email is spam or not spam, given the presence of these words.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

6. Naïve Bayes Classification

Laplace Correction:

- Sometimes, a word may never appear in a category, leading to a zero probability. This is called the Zero Conditional Probability Problem. To solve this, we use Laplace correction, where we add 1 to the count of each word to avoid the probability from becoming zero.
- **For example:**
 - If "offer" appears in 0 spam emails, $P(\text{offer}|\text{spam})$ would be zero. Using Laplace correction, we add 1 to the count, ensuring the probability is not zero.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

6. Naïve Bayes Classification

Pros and Cons of Naïve Bayes Algorithm:

Pros:

1. **Simple and Fast:** Easy to implement and works well with a large dataset.
 - a. Works well with categorical and continuous data.
 - b. Performs well with missing data as it can handle unknown values.
 - c. Can work with small datasets.

Cons:

1. **Assumption of Independence:** The assumption that all features are independent is rarely true in real-world datasets.
2. **Zero Probability Problem:** If a class is never observed with a certain feature, the probability can become zero, which is handled by Laplace correction.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

6. Naïve Bayes Classification

Applications of Naïve Bayes:

1. **Text Classification:** Classifying whether a document is spam or not.
2. **Sentiment Analysis:** Analyzing tweets or reviews to classify them as positive, negative, or neutral.
3. **Recommendation Systems:** Predicting whether a user will like a product based on their past behaviors.
4. **Medical Diagnosis:** Classifying whether a patient has a certain disease based on medical features.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

6. Naïve Bayes Classification

Real-World Example:

Let's say you have a dataset of emails and you want to classify them as spam or not spam using the words in the email as features.

1. **Step 1:** Count how many times each word appears in spam and non-spam emails.
2. **Step 2:** Use Bayes' Theorem to calculate the probability that a new email with certain words is spam or not spam.
3. **Step 3:** Classify the new email based on the higher probability.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

6. Naïve Bayes Classification

Example 1: Spam Email Classification

We have a dataset of emails and want to classify whether an email is Spam or Not Spam based on the words in the email.

Word	Spam Count	Non-Spam
offer	3	1

- **Dataset:**

- We know from the dataset that the word "offer" appears in 3 spam emails and 1 non-spam email.
- Let's say we are given a new email that contains the word "offer", and we want to classify it as Spam or Not Spam.

Step 1: Apply Bayes' Theorem

- **Bayes' Theorem is:**
$$P(\text{Spam}|\text{offer}) = \frac{P(\text{offer}|\text{Spam}) \cdot P(\text{Spam})}{P(\text{offer})}$$

- Where:
- $P(\text{Spam} | \text{offer})$ is the posterior probability of an email being spam given the word "offer."
- $P(\text{offer} | \text{Spam})$ is the likelihood of the word "offer" occurring in spam emails.
- $P(\text{Spam})$ is the prior probability of an email being spam.
- $P(\text{offer})$ is the evidence (probability of "offer" occurring in any email).





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

6. Naïve Bayes Classification

Step 2: Calculate Probabilities

- **P(offer|Spam)**: This is the probability of the word "offer" appearing in a spam email. From our data, there are 3 spam emails with "offer" and a total of 5 spam emails, so:

$$P(\text{offer}|\text{Spam}) = \frac{3}{5} = 0.6$$

- **P(Spam)**: This is the probability that an email is spam. We don't have the exact number of spam and non-spam emails, but let's assume that 60% of the emails are spam. So: $P(\text{Spam})=0.6$
- **P(offer)**: This is the overall probability that the word "offer" appears in any email. From our dataset, there are 4 instances of "offer" (3 in spam and 1 in non-spam) out of 10 total emails, so:

$$P(\text{offer}) = \frac{4}{10} = 0.4$$





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

6. Naïve Bayes Classification

Step 3: Calculate the Posterior Probability

Now, we can plug these values into Bayes' Theorem:

$$P(\text{Spam}|\text{offer}) = \frac{0.6 \times 0.6}{0.4} = 0.9$$

Thus, the probability of the email being Spam given that it contains the word "offer" is 0.9 or 90%.

Step 4: Interpretation

Since the probability of the email being Spam is 90%, we classify this email as Spam.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

6. Naïve Bayes Classification

Conclusion:

- Naïve Bayes is a simple, efficient, and powerful algorithm for classification tasks.
- Bayes' Theorem is used to calculate the probability of a class given a set of features.
- It assumes features are independent, which simplifies the calculation but might not always reflect real-world data accurately.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

7. Deep Learning and Neural Networks

Introduction to Neural Networks:

- **Neural Networks (NN)** are inspired by the human brain and consist of layers of interconnected neurons.
- **Key Concept:** Neurons process input data and learn to make predictions or classifications based on the patterns they detect.
- **Training:** The neural network adjusts its internal structure (called weights) to minimize errors during training.
- **Example:** Think of recognizing a dog in an image; the neural network learns to identify features like shapes, edges, and colors by processing the image through layers of neurons.

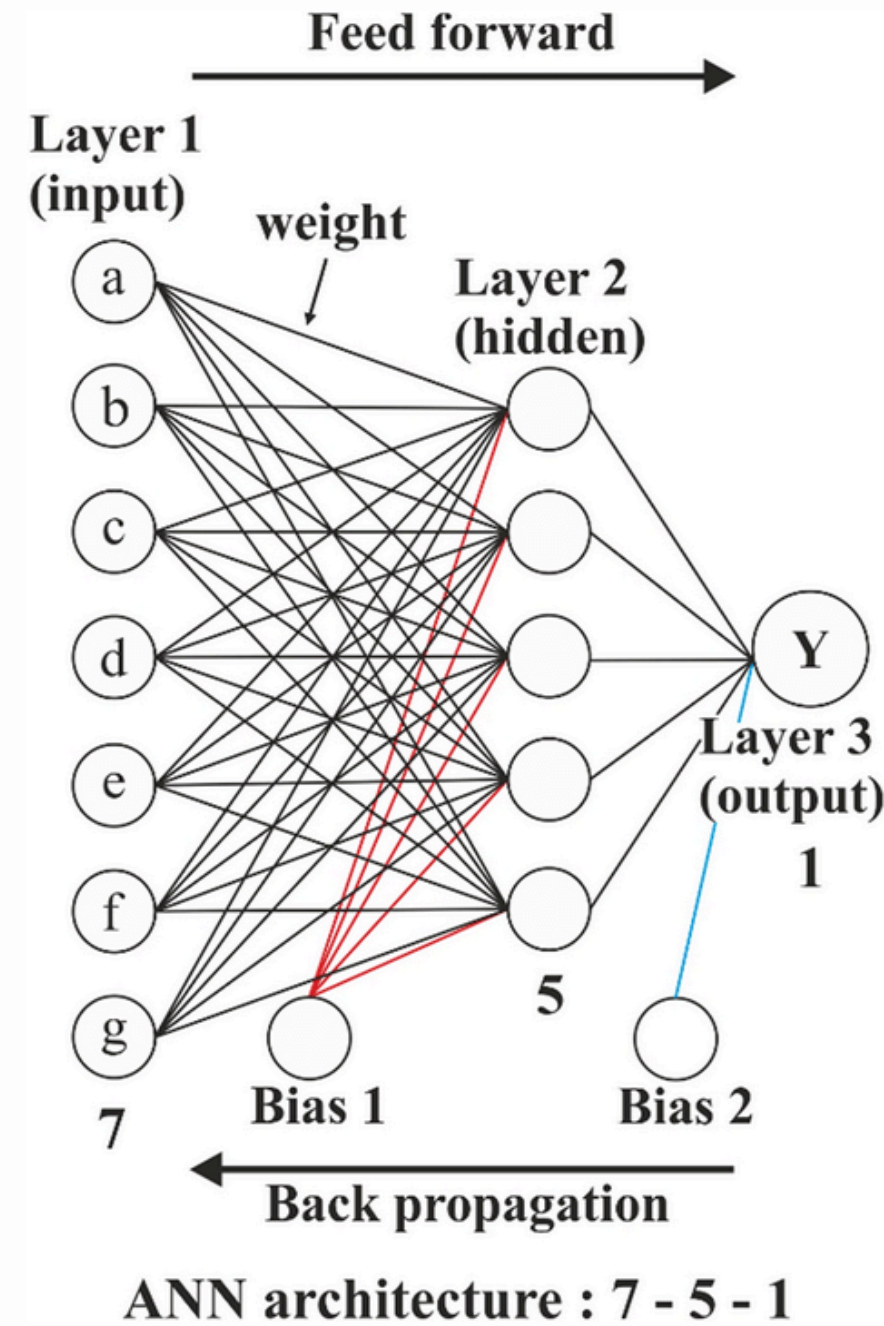




Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

7. Deep Learning and Neural Networks



Steps in Training a Neural Network:

- 1. Input Layer:** Input data is fed to the first layer of neurons.
- 2. Processing in Hidden Layers:** Each layer processes the data using its weights and applies the activation function.
- 3. Output Layer:** The final result is output after all layers have processed the data.
- 4. Backpropagation:** Errors from the output layer are passed backward through the network, adjusting the weights for better future predictions.

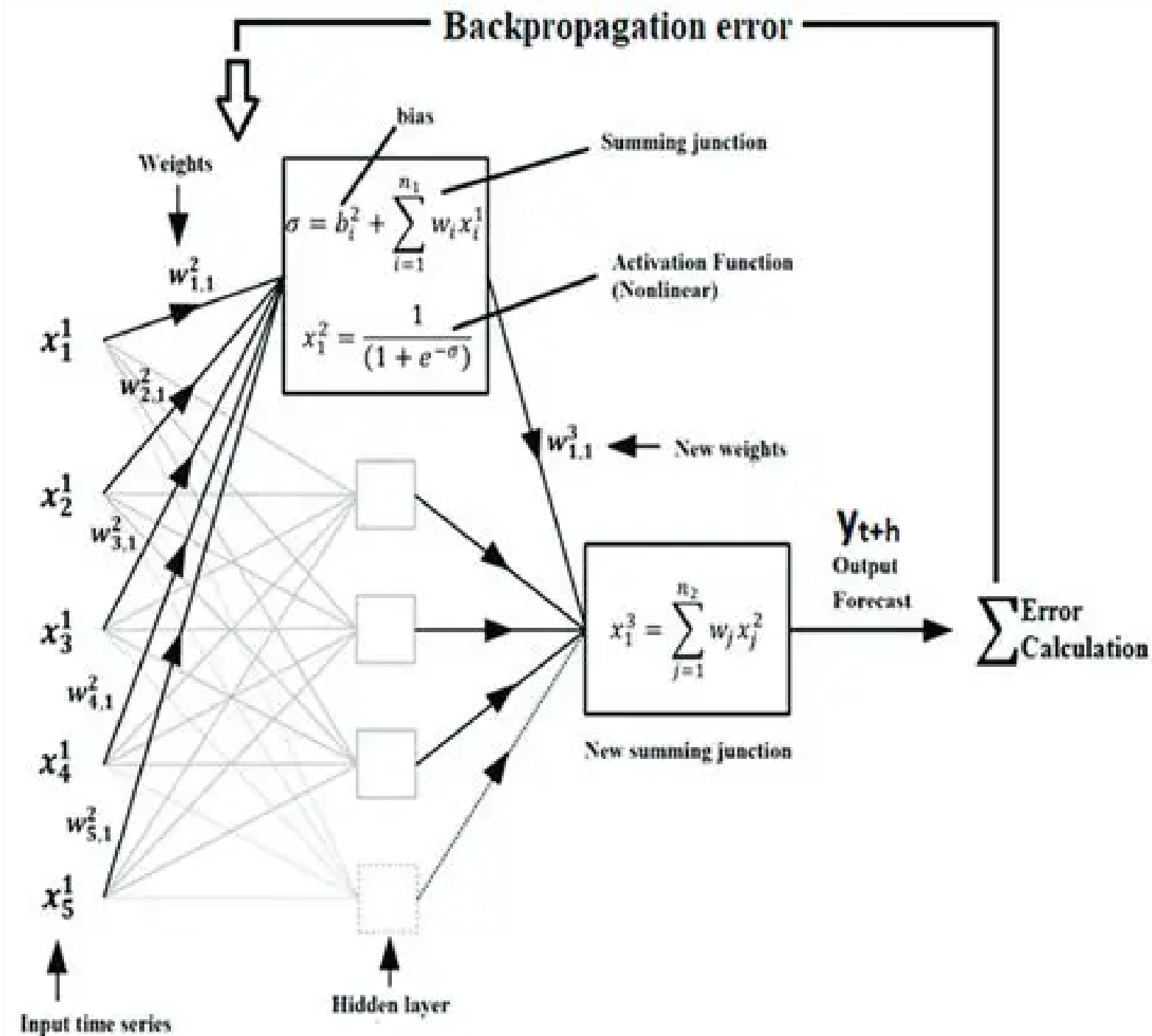




Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

7. Deep Learning and Neural Networks



2. Working of Neural Networks:

- **Neurons and Layers:** A neural network has layers:
 - **Input Layer:** Receives raw data.
 - **Hidden Layers:** Processes the data.
 - **Output Layer:** Gives the final prediction or classification.
- **Activation Function:**
 - **Purpose:** Transforms the input into an output within a certain range, typically between 0 and 1 (for binary classification).
 - **Example Activation Functions:** Sigmoid, Tanh, ReLU.
- **Forward Propagation:** Data flows through layers to predict the output.
- **Backpropagation:** Error is calculated at the output and then passed backward to adjust weights, minimizing prediction errors.



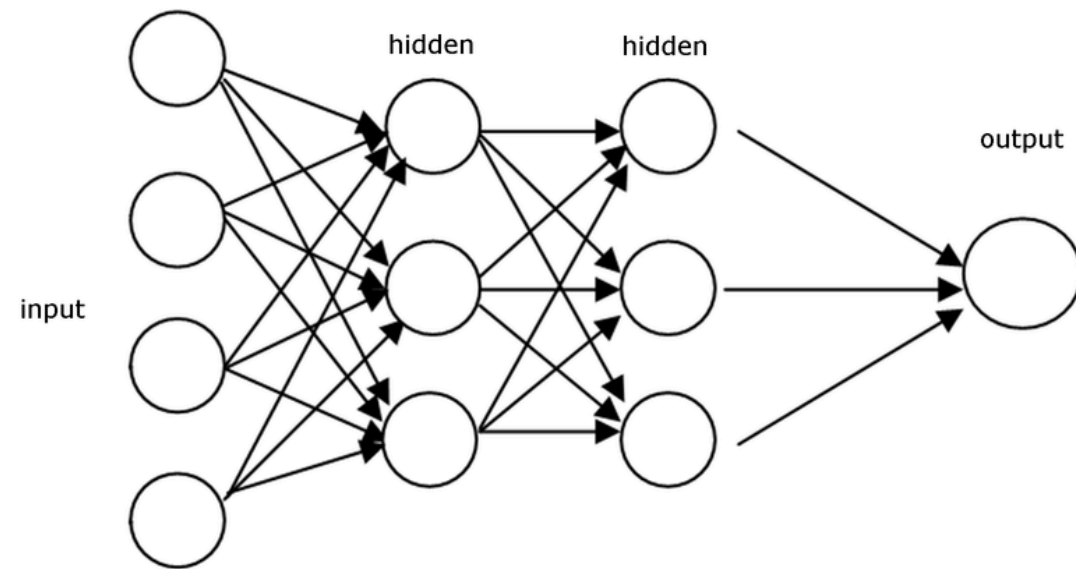


Module 3: Machine Learning in AI

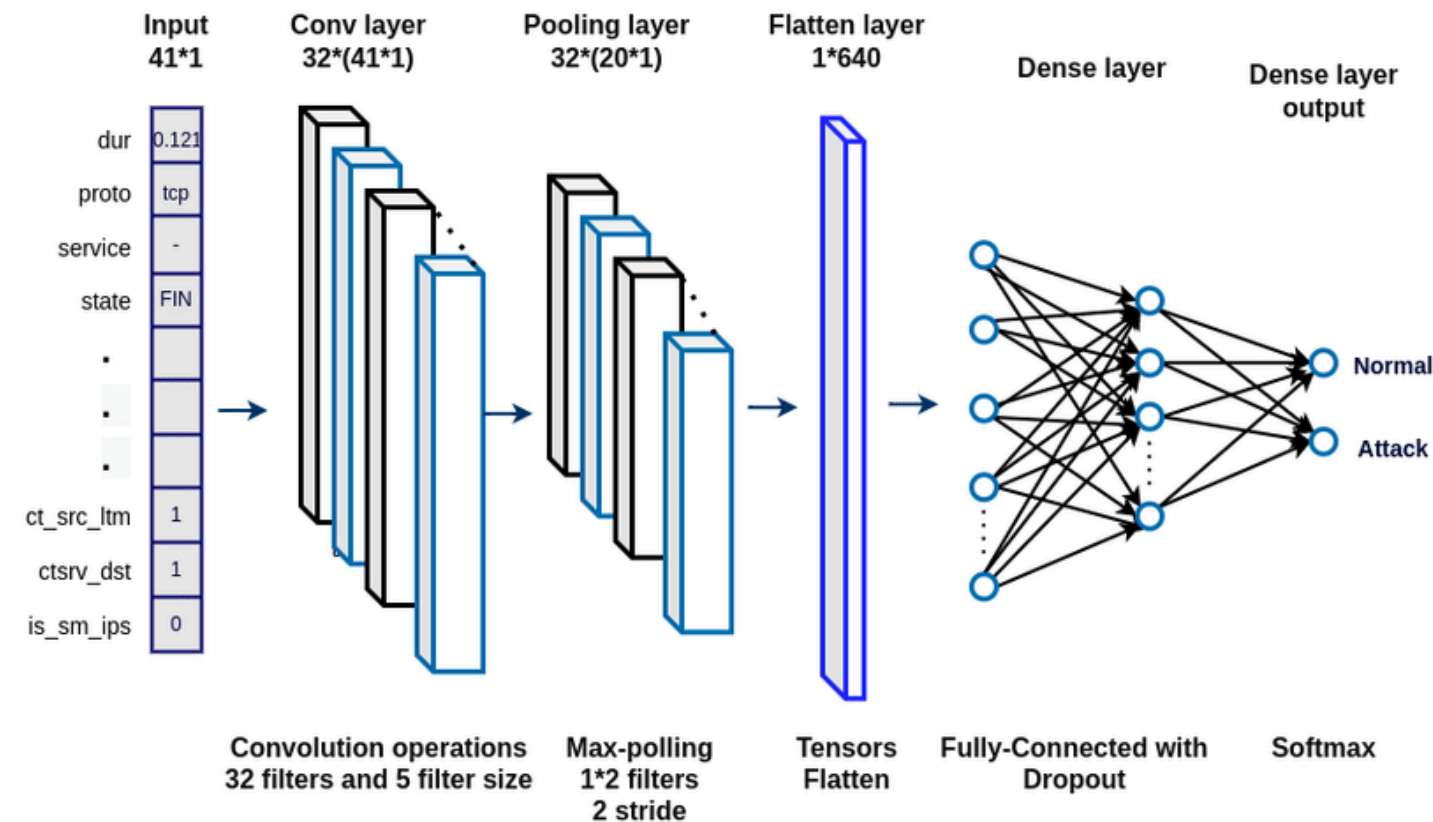
2 Machine Learning Model - Classification Techniques

7. Deep Learning and Neural Networks

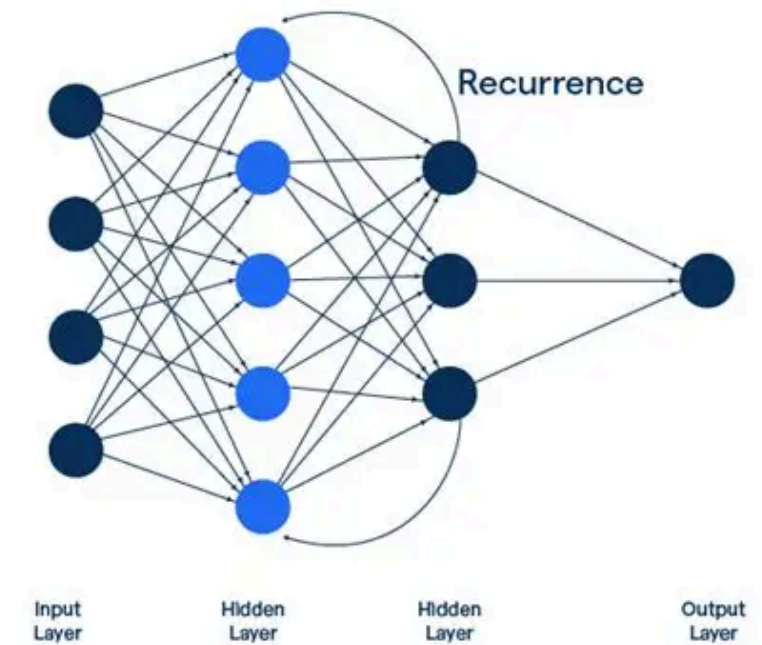
Types of Neural Networks:



Feedforward Neural Network:
The most basic type, where information moves in one direction from input to output.



Convolutional Neural Networks (CNNs):
Specialized for image recognition and processing.



Recurrent Neural Networks (RNNs):
Used for sequence-based data, like time series or speech.



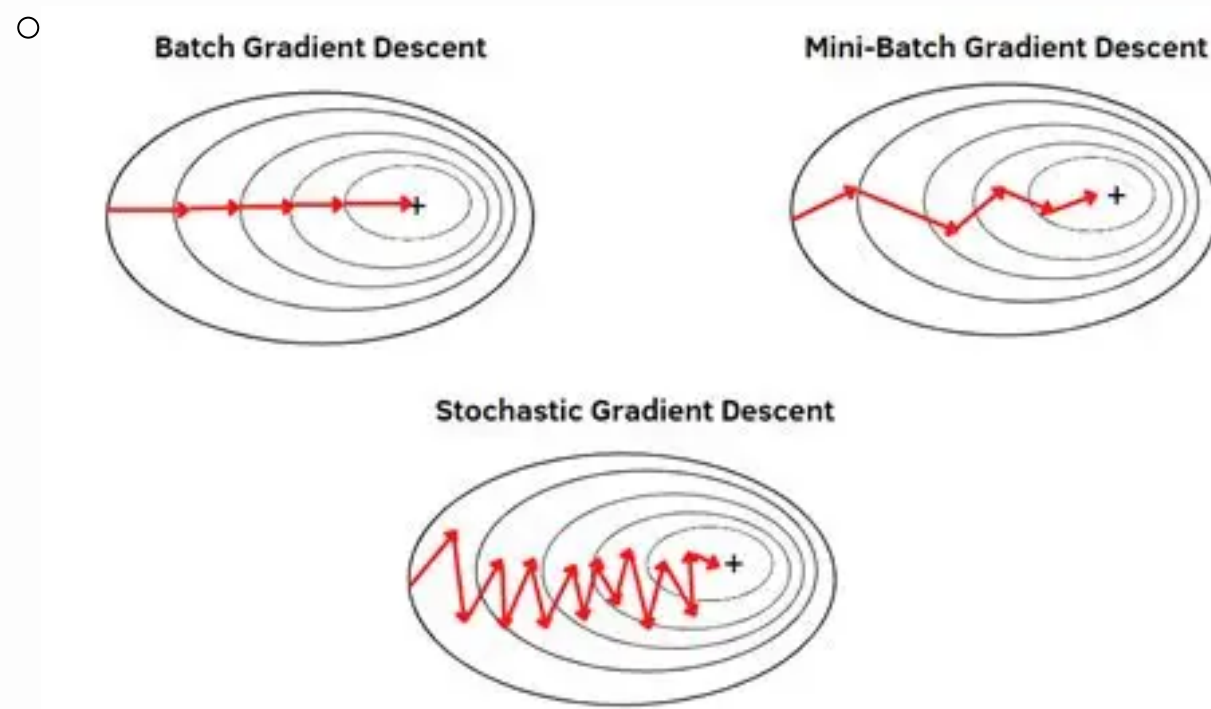
Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

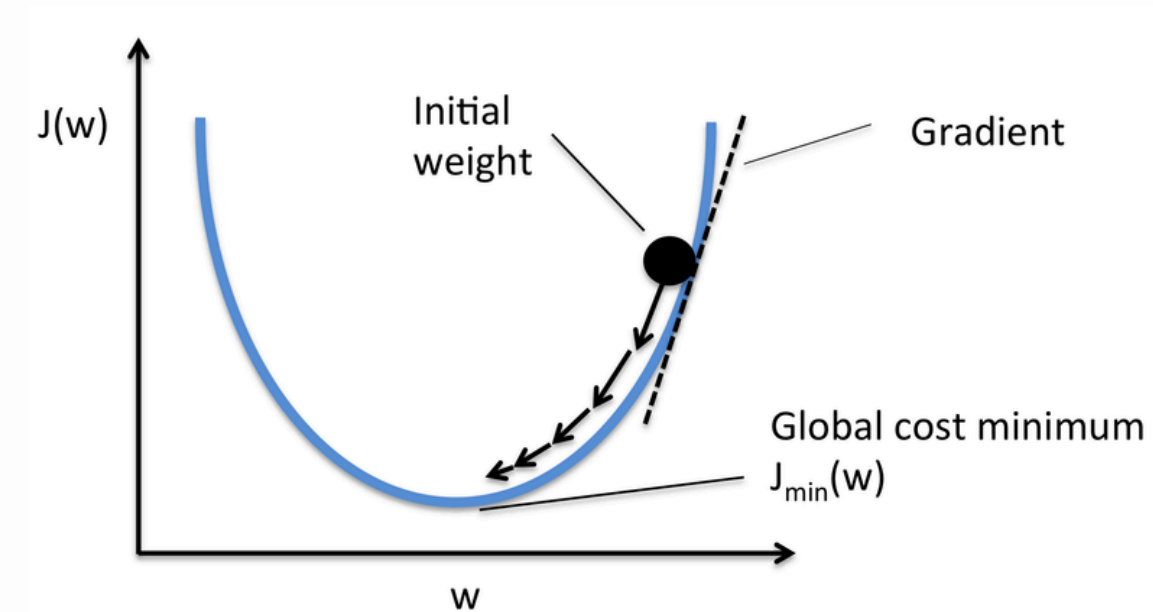
7. Deep Learning and Neural Networks

Training a Neural Network:

- **Gradient Descent:** The most common optimization technique, where the algorithm minimizes the error by adjusting weights. The two types are:



Batch Gradient Descent:
Uses all data to update weights.



Stochastic Gradient Descent (SGD):
Updates weights after every individual data point.



Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

7. Deep Learning and Neural Networks

Pros and Cons of Neural Networks:

Pros:

- **Flexible:** Can be used for both regression and classification problems.
- **Works well with complex and non-linear data:** Ideal for tasks like image recognition and natural language processing.
- **Scalable:** Works with a large number of inputs and hidden layers.

Cons:

- **Complex and computationally expensive:** Requires significant computing resources.
- **Training can take time:** Especially for large datasets.
- **Requires a lot of data:** Performance improves as more data is provided.
- **Interpretability issues:** It's often seen as a "black box" due to its complexity.



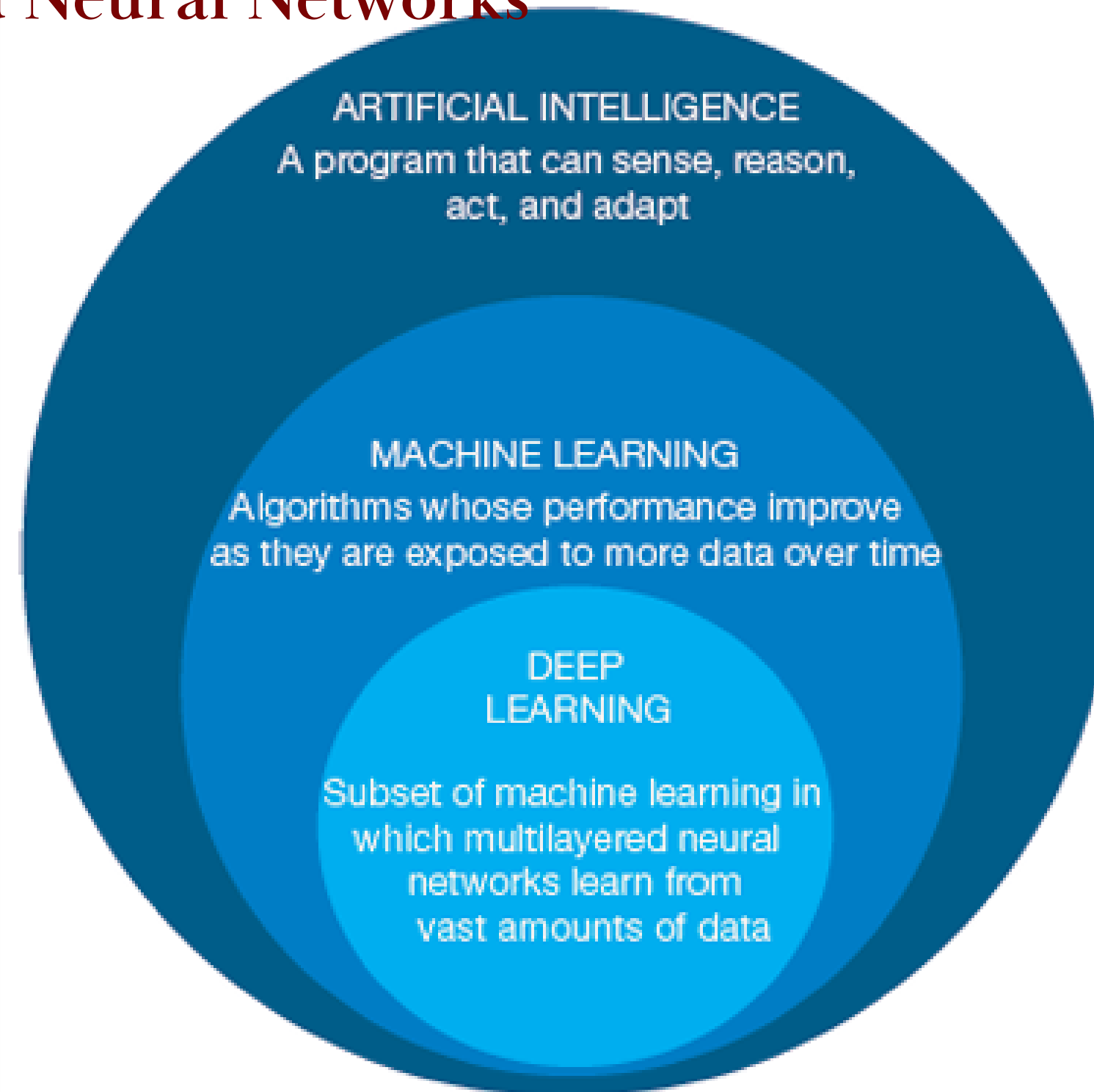


Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

7. Deep Learning and Neural Networks

Deep Learning



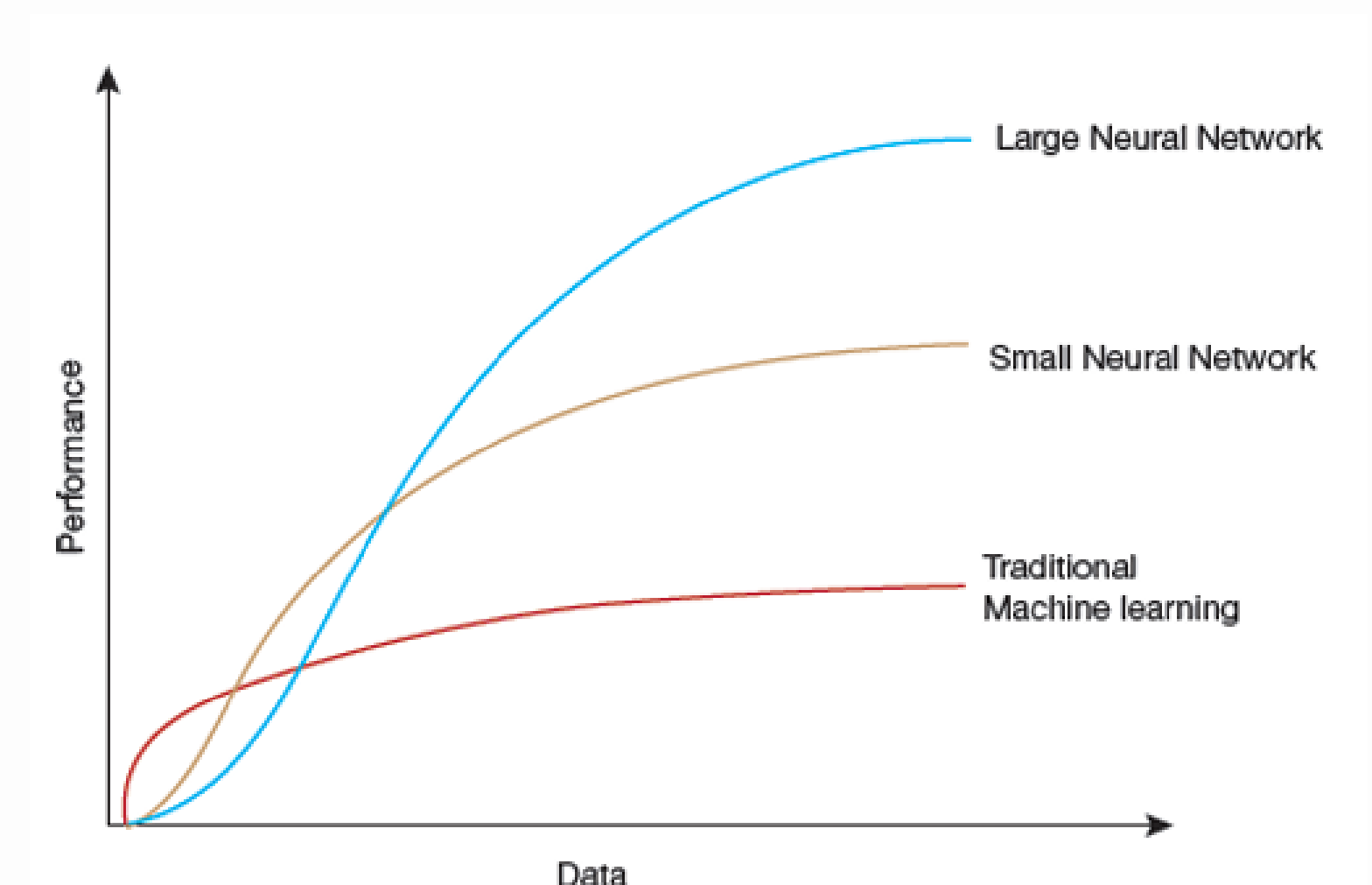
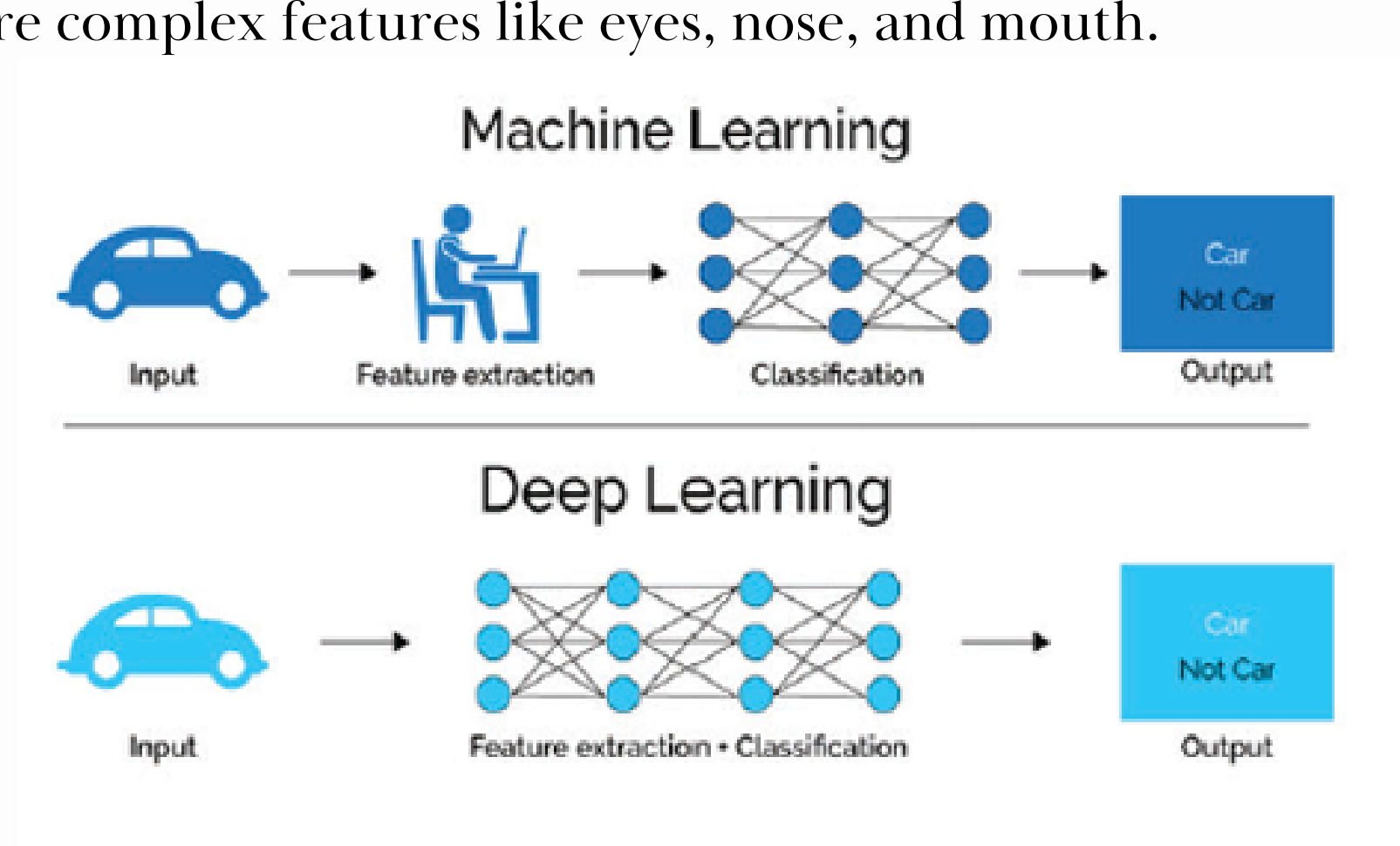
Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

7. Deep Learning and Neural Networks

Deep Learning:

- Deep Learning is a subset of machine learning, specifically using deep neural networks with multiple hidden layers.
- **Benefit:** Can automatically extract features from raw data, eliminating the need for manual feature engineering.
- **Example:** In a facial recognition system, deep learning models can detect edges in the first layers, then move on to more complex features like eyes, nose, and mouth.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

7. Deep Learning and Neural Networks

Example: Fraud Detection with Deep Learning:

- **Task:** Identify fraudulent transactions in a dataset.
- **Input:** Transaction data (e.g., amount, location, time).
- **Output:** Fraudulent or not fraudulent.
- **Process:**
 - a. **Input Layer:** Transaction details are fed into the network.
 - b. **Hidden Layers:** Detect patterns and relationships, such as unusual spending behavior.
 - c. **Output Layer:** Classifies the transaction as fraudulent or not.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

7. Deep Learning and Neural Networks

Example: Fraud Detection with Deep Learning:

- Imagine a bank using deep learning to detect fraudulent transactions. The process could go as follows:
- **Input Layer:**
 - Data like transaction amount, time, user's IP address, and location are input into the neural network.
 - For example, if someone suddenly tries to spend a large sum of money in a different country, that could be suspicious.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

7. Deep Learning and Neural Networks

Example: Fraud Detection with Deep Learning:

- **First Hidden Layer:**
 - The first layer processes the transaction amount and passes it as output to the next layer.
 - The neural network starts learning basic patterns, like: "Large transactions in unfamiliar locations may indicate fraud."
- **Second Hidden Layer:**
 - Now, additional data like IP address is processed.
 - The system checks if the IP address matches the expected location of the user.
 - If the transaction comes from an unexpected IP address, the system starts to raise a red flag.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

7. Deep Learning and Neural Networks

Example: Fraud Detection with Deep Learning:

- **Third Hidden Layer:**
 - The neural network adds geographic location into the equation.
 - It compares the user's usual location with where the transaction is being made.
 - If the location is too far away from usual activity, the system sees this as a potential fraud trigger.
- **Final Layer:**
 - The final layer generates a result.
 - Based on the accumulated data, the system may classify the transaction as fraudulent or not fraudulent.
 - If it's flagged as fraud, the system may alert the bank to freeze the user's account or block the transaction.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

7. Deep Learning and Neural Networks

Example: Fraud Detection with Deep Learning:

- **Why This Works:**
 - **Learning from Data:** The deep learning model learns from vast amounts of historical data to detect patterns in transactions that humans may miss.
 - **Improvement Over Time:** As more data is processed, the model becomes more accurate at identifying fraudulent activities.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

7. Deep Learning and Neural Networks

Applications of Neural Networks and Deep Learning:

1. **Pattern Recognition:** Used in facial recognition, object detection, and handwriting recognition.
2. **Anomaly Detection:** Identifying fraudulent activities or rare events in data.
3. **Time-Series Prediction:** Predicting stock prices, weather forecasting.
4. **Natural Language Processing:** Applications like sentiment analysis, machine translation, and speech recognition.
5. **Recommendation Systems:** Suggesting products or content based on user preferences.
6. **Medical Diagnosis:** Identifying diseases in medical images or predicting patient conditions.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

8. Support Vector Machine (SVM)

- A Support Vector Machine (SVM) is a popular supervised learning algorithm primarily used for classification tasks.
- Its main goal is to classify data points into one of two categories based on features in the dataset.
- An SVM is a classification algorithm that works by finding the best hyperplane that separates two classes of data points in an n-dimensional space (where n is the number of features).
- The hyperplane is the decision boundary that helps classify data points into their respective classes.

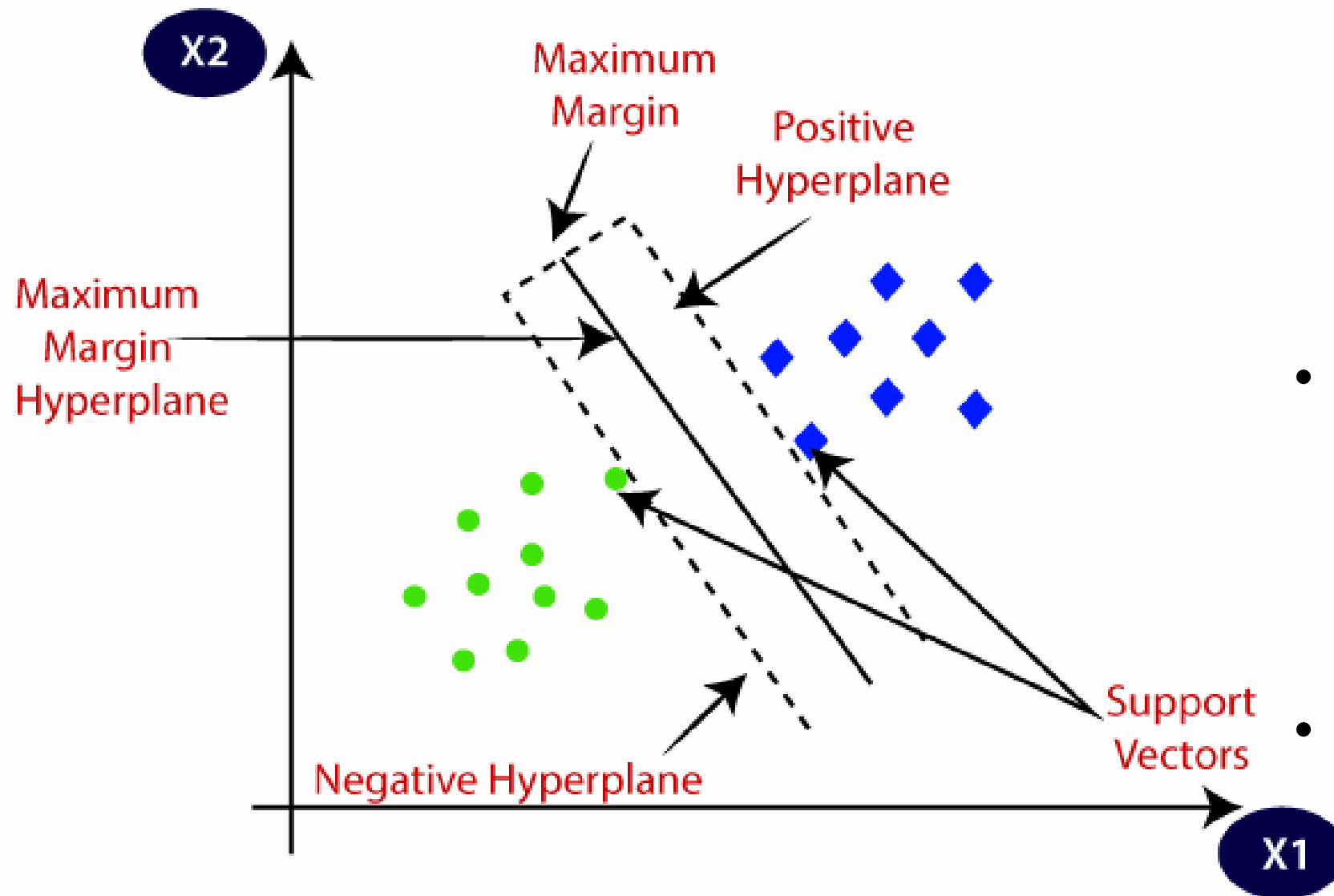




Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

8. Support Vector Machine (SVM) Key Concepts in SVM:



- **Hyperplane:**
 - A hyperplane is simply a line (in 2D), a plane (in 3D), or a general decision boundary in higher dimensions that separates the data into two classes.
 - **Example:** In a 2D space, the hyperplane would be a line that divides the data into two parts (one part for Class 1 and another for Class 2).
- **Support Vectors:**
 - Support vectors are the data points that are closest to the hyperplane. These points are critical in determining the optimal hyperplane.
 - The margin is the distance between the hyperplane and the closest support vectors. The larger the margin, the better the classifier.
- **Margin:**
 - The margin is the distance between the hyperplane and the closest points from each class. The SVM tries to maximize this margin, as a larger margin leads to a better separation between the classes and a more accurate model.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

8. Support Vector Machine (SVM)

How Does SVM Work?

1. Linear Classification:

- For linearly separable data, SVM finds a hyperplane that maximizes the margin between two classes.
- **Example:**
 - We have a set of data points (e.g., red and blue circles), and we need to classify them. SVM draws a line (hyperplane) between the red and blue circles to separate them.
 - SVM chooses the line that maximizes the distance (margin) between the red and blue points.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

8. Support Vector Machine (SVM)

How Does SVM Work?

2. Non-Linear Classification (Kernel Trick):

- When the data is not linearly separable, SVM uses a method called the kernel trick to map the data into higher dimensions, where it becomes easier to find a separating hyperplane.
- **Example:**
 - In the case of non-linearly separable data, imagine you have two classes (red and blue), but they can't be separated by a straight line. By transforming the data into a higher-dimensional space (e.g., using polynomial or Gaussian kernels), the data points become separable by a hyperplane.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

8. Support Vector Machine (SVM)

Types of Kernels in SVM:

1. Linear Kernel:

- Used when the data is linearly separable. The decision boundary is a straight line or hyperplane.

2. Polynomial Kernel:

- Used when the data has a non-linear relationship, but we can represent it as a polynomial.

3. Radial Basis Function (RBF) Kernel:

- One of the most commonly used kernels, especially for non-linear problems. It maps data points to a higher-dimensional space and creates a curved decision boundary.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

8. Support Vector Machine (SVM)

Example:

Let's consider an example where we need to classify whether a cricket player is a batsman or a bowler based on their runs and wickets:

- Data points (Runs, Wickets):
 - Batsman: (250, 10), (300, 12), (400, 8)
 - Bowler: (50, 30), (60, 40), (80, 20)

We want to classify new players based on these features.

Step 1: Plot the Data Points:

- Plot the runs on the X-axis and wickets on the Y-axis.
- You'll see that the batsmen and bowlers form two distinct groups in the 2D space.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

8. Support Vector Machine (SVM)

Step 2: Find the Optimal Hyperplane:

- The SVM will find the line (hyperplane) that best separates the batsmen from the bowlers by maximizing the margin (distance between the two classes).

Step 3: New Data Point Classification:

- When a new player comes with runs and wickets, say (350, 15), we can use the learned model (hyperplane) to classify the player as a batsman or bowler based on which side of the hyperplane the data point lies.
-





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

8. Support Vector Machine (SVM)

Problem:

We have the following data points that we need to classify as Batsmen or Bowlers based on their Runs and Wickets:

- Batsmen:
 - (250, 10)
 - (300, 12)
 - (400, 8)
- Bowlers:
 - (50, 30)
 - (60, 40)
 - (80, 20)

We want to classify a new player with the following data:

- New player: (350, 20) (Runs = 350, Wickets = 20)





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

8. Support Vector Machine (SVM)

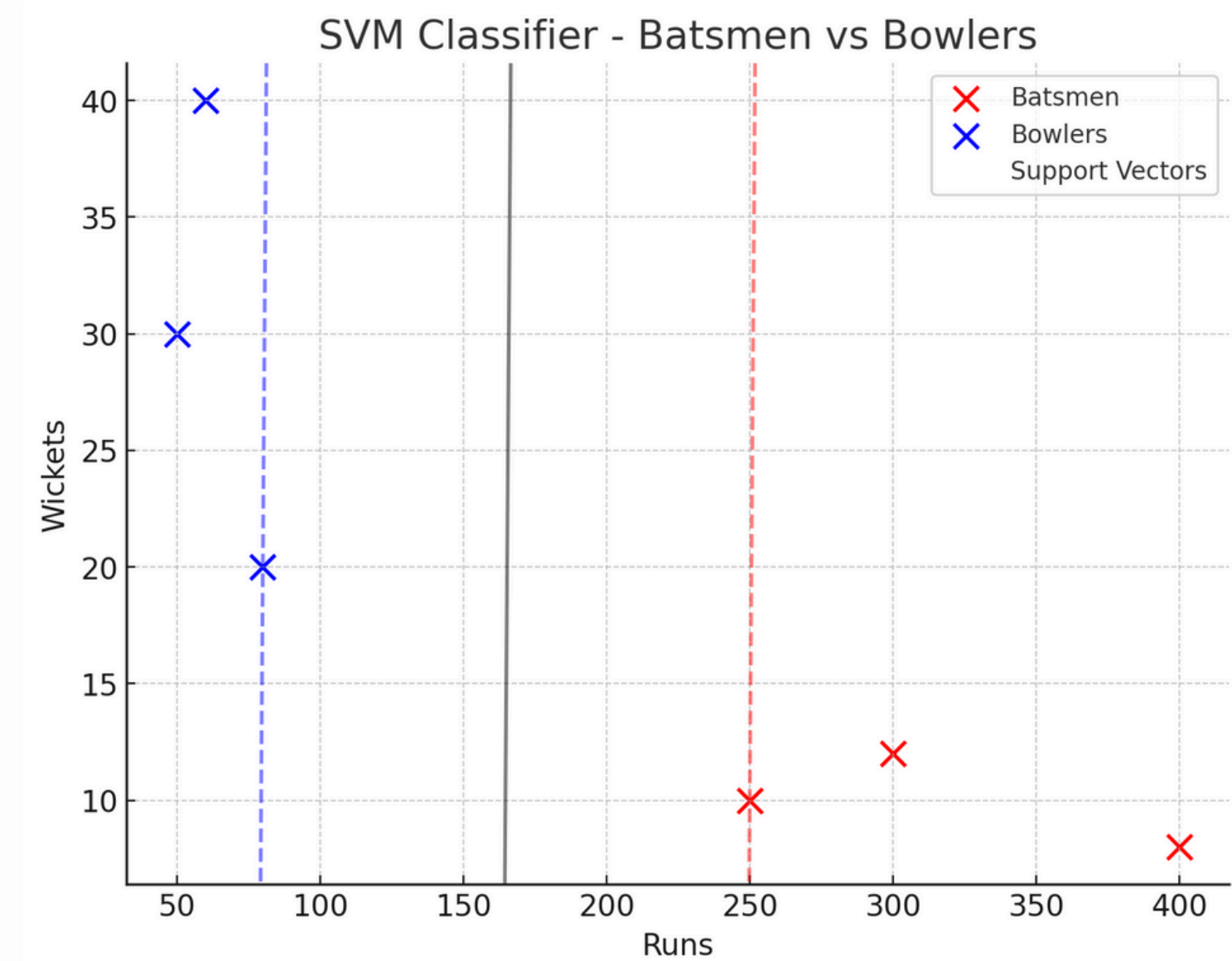
Problem:

We have the following data points that we need to classify as Batsmen or Bowlers based on their Runs and Wickets:

- Batsmen:
 - (250, 10)
 - (300, 12)
 - (400, 8)
- Bowlers:
 - (50, 30)
 - (60, 40)
 - (80, 20)

We want to classify a new player with the following data:

- New player: (350, 20) (Runs = 350, Wickets = 20)





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

8. Support Vector Machine (SVM)

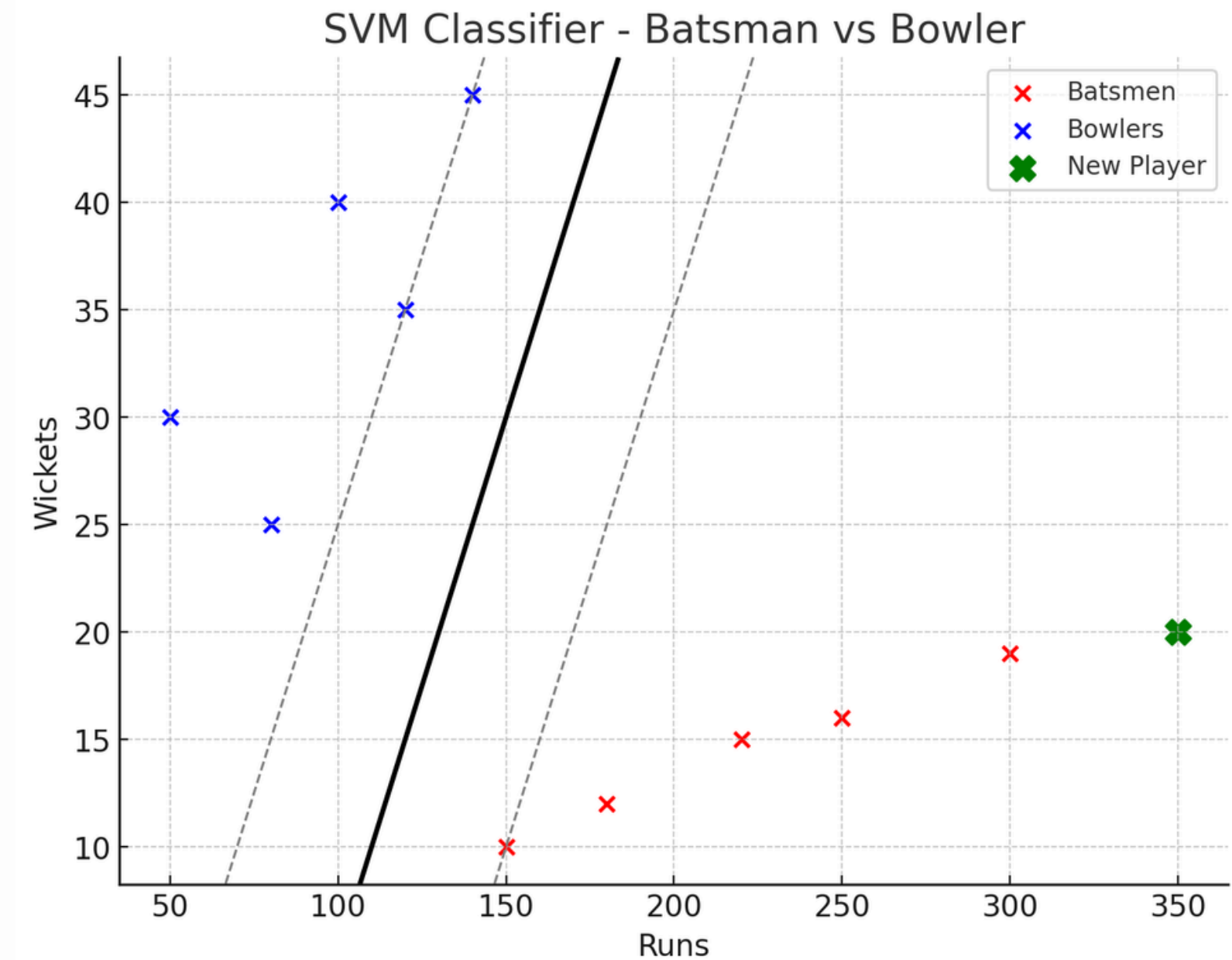
Problem:

Find the Optimal Hyperplane

- The SVM will compute the best possible hyperplane that separates the batsmen and bowlers based on the given data.
- The margin, which is the distance between the hyperplane and the closest points (support vectors), is maximized.

4. Classify the New Player

Once the hyperplane is defined, we classify the new player based on which side of the hyperplane they lie.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

8. Support Vector Machine (SVM)

Tuning Parameters in SVM:

1. Cost Parameter (C):

- This parameter controls the trade-off between maximizing the margin and minimizing classification error.
- A high C value results in a narrower margin but fewer misclassifications (risk of overfitting).
- A low C value results in a wider margin but more misclassifications (risk of underfitting).

2. Gamma (for RBF Kernel):

- Gamma defines the influence of a single training example. A low gamma means the influence is far, and a high gamma means the influence is local.
- High gamma can lead to overfitting, and low gamma can lead to underfitting.





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

8. Support Vector Machine (SVM)

Advantages of SVM:

1. Works well with high-dimensional data.
2. Effective in both linear and non-linear classifications.
3. Memory efficient, as it uses only support vectors for classification.

Disadvantages of SVM:

1. It can be slow with large datasets.
2. It does not perform well with overlapping classes.
3. It is sensitive to the choice of kernel and parameter tuning.

○





Module 3: Machine Learning in AI

2 Machine Learning Model - Classification Techniques

8. Support Vector Machine (SVM)

Applications of SVM:

1. **Image Recognition:** SVMs are used to recognize patterns and objects in images.
2. **Text Classification:** SVMs are used for spam email classification or document categorization.
3. **Face Detection:** SVMs are used to detect faces in images by classifying pixel data as face or non-face.
4. **Bioinformatics:** SVMs are used for classifying protein sequences and gene expression data.

○





Outline

Module 3: Machine Learning

Reema Thareja, Artificial Intelligence: Beyond Classical AI, Pearson Education, 2023.

1. Machine Learning in AI

- Overview of Machine Learning Techniques
- Introduction to Machine Learning Models

2. Regression Analysis in Machine Learning

- Basics of Regression
- Linear and Non-linear Regression Techniques

3. Classification Techniques

- Overview of Classification Algorithms
- Naïve Bayes Classification
- Support Vector Machine (SVM)

4. Clustering Techniques

- Introduction to Clustering
- Types of Clustering Algorithms

5. Neural Networks

- Basics of Neural Networks
- Types and Applications of Neural Networks

6. Support Vector Machine (SVM).





Thank You

Module 3 - Completed