



ARTIFICIAL INTELLIGENCE



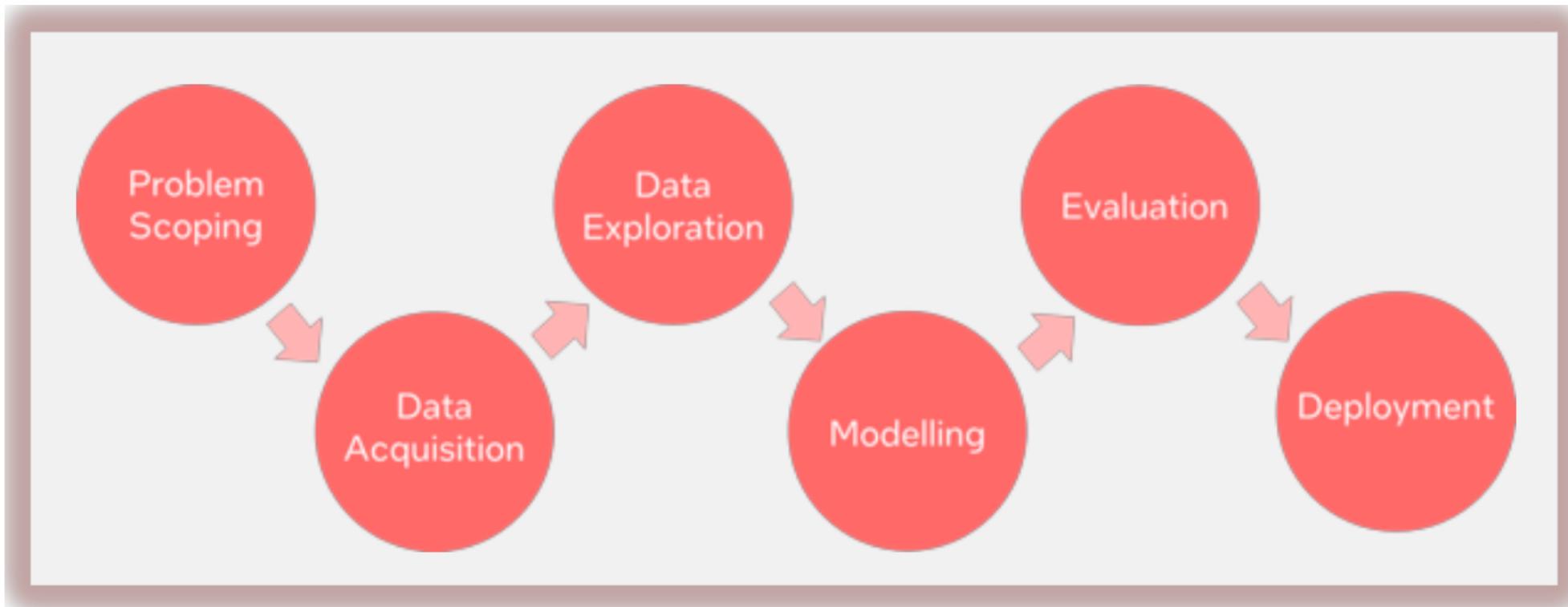


Revisiting AI Project Cycle & Ethical Frameworks for AI

Introduction

1.1 AI Project Cycle (The Overview of the six stages of the AI Project Cycle)

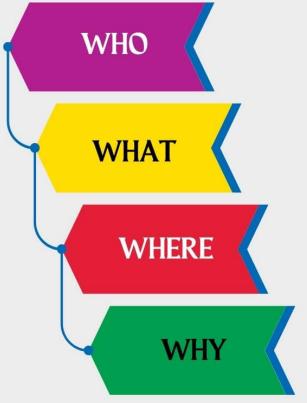
Our mind naturally creates plans to accomplish tasks, helping us gain clarity. Similarly, developing an AI project requires a structured approach. The AI Project Cycle offers a framework to guide this process, typically involving six key stages.



1. Problem Scoping

The first step is acquiring data, which forms the foundation of the project and helps determine the key factors involved in defining the problem.

What is 4W Canvas?



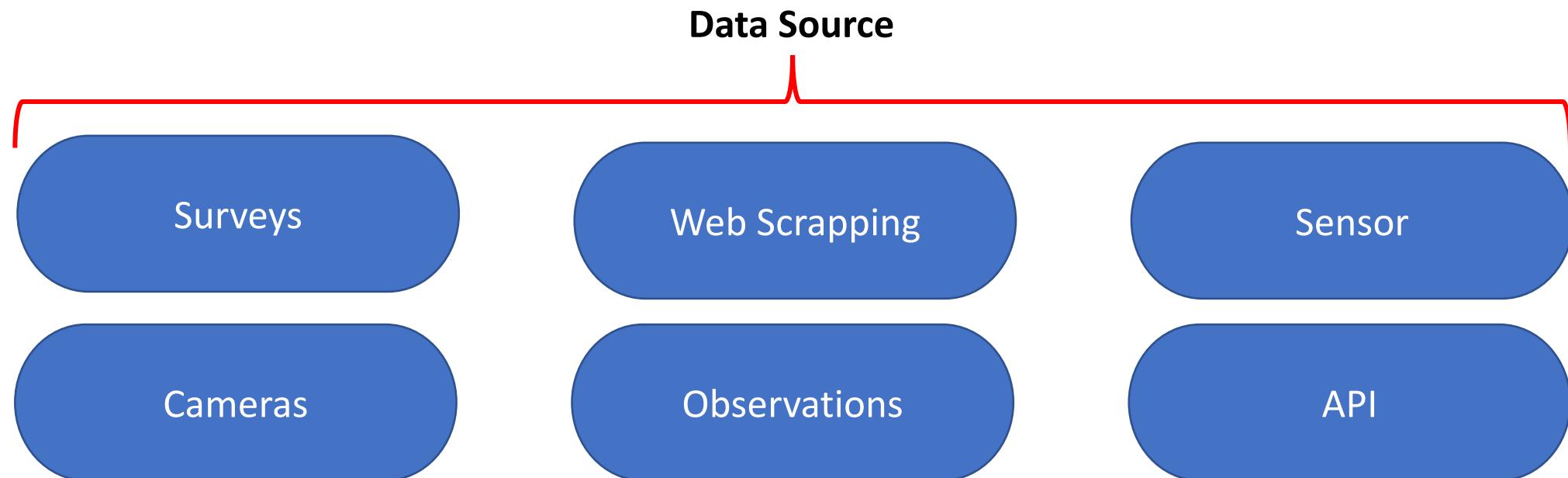
The 4Ws are:

- Who are you solving for?
- What is the problem?
- Where the problem exists?
- Why are you solving it?

OUR	[stakeholders] People who are using the new technology	WHO
HAS/ HAVE PROBLEM THAT	[issue, problem, need] Cyber security is the need when so much of the flow of data is not monitored or escapes the antiviruses/ firewall systems.	WHAT
WHEN/ WHILE	[context/situation] The problem is in the use of the latest technology where vast amounts of data is at risk.	WHERE
AN IDEAL SOLUTION WOULD	[benefit of solution to them] An effective AI system which is able to detect the flow of data and also report unusual activity	WHY

2. Data Acquisition

Data is collected from reliable sources and represented visually through graphs, charts, or maps to help identify patterns and make interpretation easier.

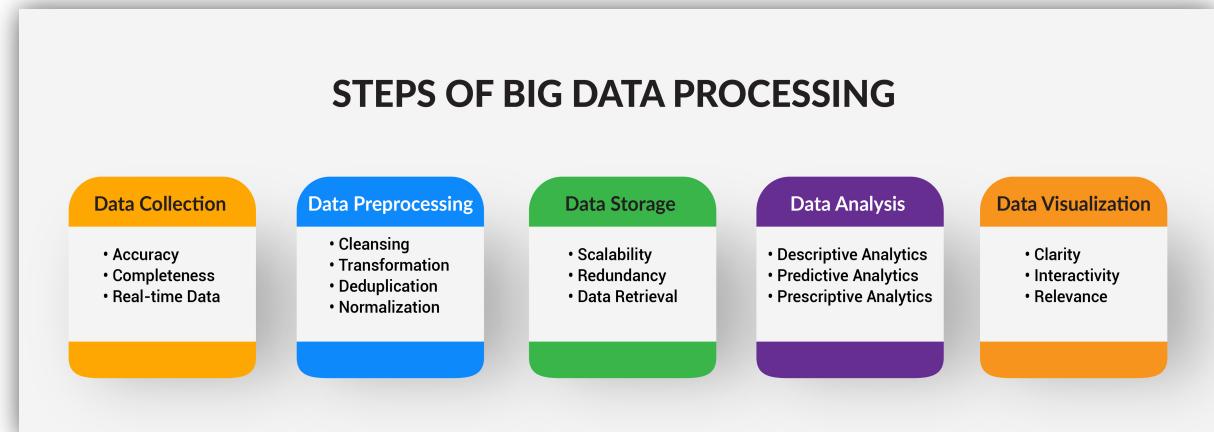
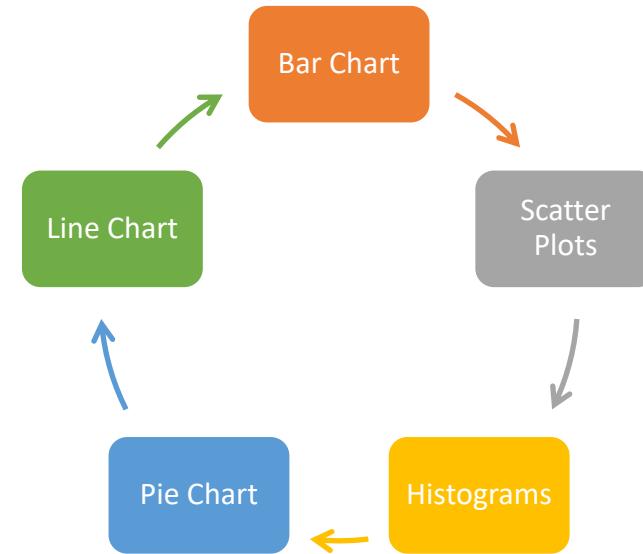


3. Data Exploration

After identifying patterns, you can choose the right model by exploring different options online and picking the one that gives the best results.

- ◆ In the data exploration stage of the AI project cycle, ***the main goal is to understand and visualize the collected data to spot patterns, trends, and relationships.***
- ◆ This involves using tools like ***charts, graphs, and basic analysis*** to learn more about the data's structure.
- ◆ It also helps identify any ***issues, such as missing values or outliers***, and guides the next steps, like selecting a suitable AI model or improving the problem definition.

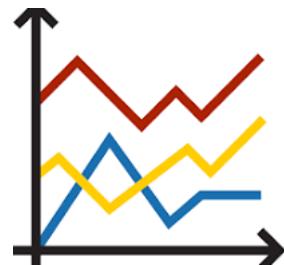
List of Graphs



3.1 Different types of Charts

Here's a quick overview of the common uses for each chart type:

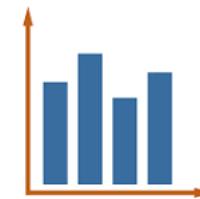
- ✧ **Line Chart**: Used to show trends and changes over time, *especially for continuous data*.
- ✧ **Bar Chart**: Used to compare data between different *groups or categories*.
- ✧ **Pie Chart**: Used to display *proportions or percentages* of a whole.
- ✧ **Histogram**: Used to show the *distribution of data within intervals or ranges*.
- ✧ **Scatter Plot**: Used to visualize the *relationship between two variables*.



Line Chart



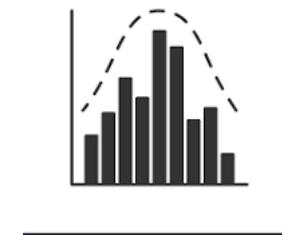
Pie Chart



Bar Chart



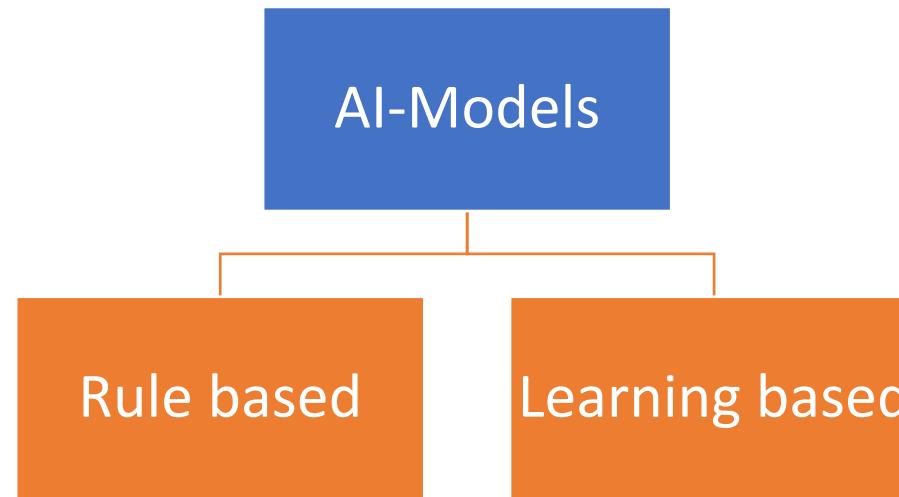
Scatter Plot



Histogram

4. AI-Model

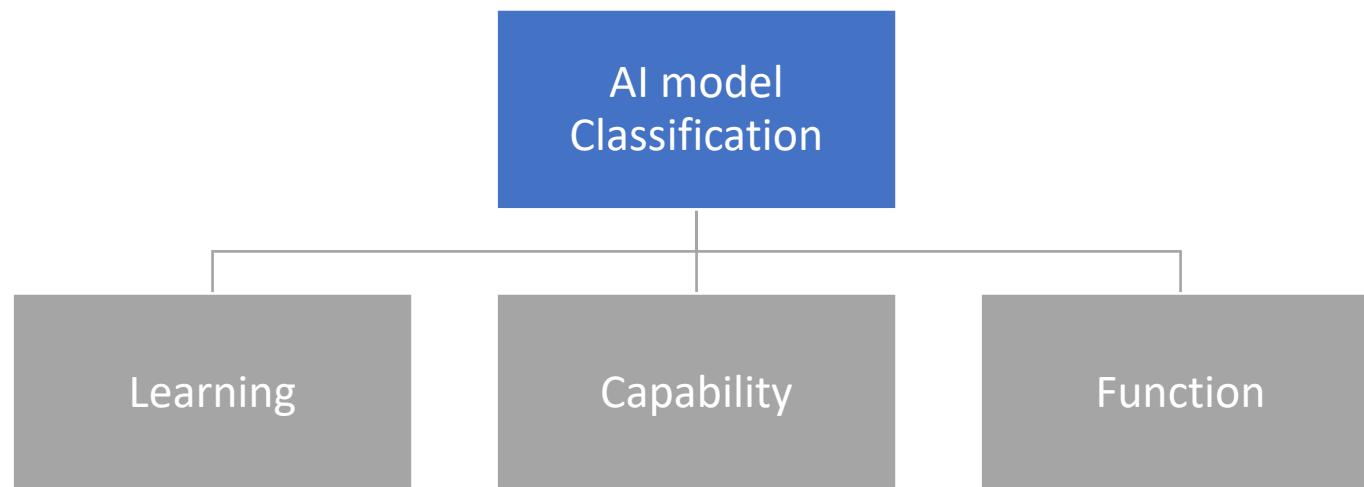
In this stage you can test the selected models and figure out which is the most efficient one



- ✧ Rule-based systems rely on ***predefined rules or knowledge bases*** to make decisions - Rule-based systems are ***pre-programmed with explicit rules***.
- ✧ Learning-based systems ***learn from data to adapt*** and make decisions, learning-based systems can ***improve their performance over time as they are exposed to more data***.

4.1 AI Model Classification

AI models can be classified in several ways. One common approach is based on the type of learning they utilize: ***Supervised, Unsupervised, and Reinforcement learning.***



4.2 Model Classification (by Learning)

- **Supervised Learning:**

These models learn from ***labeled data***, meaning the input data is paired with the correct output. Examples include classification (e.g., predicting email spam) and regression (e.g., predicting house prices).

- **Unsupervised Learning:**

These models learn from ***unlabeled data***, identifying patterns and structures within the data. Examples include clustering (e.g., grouping customers) and dimensionality reduction.

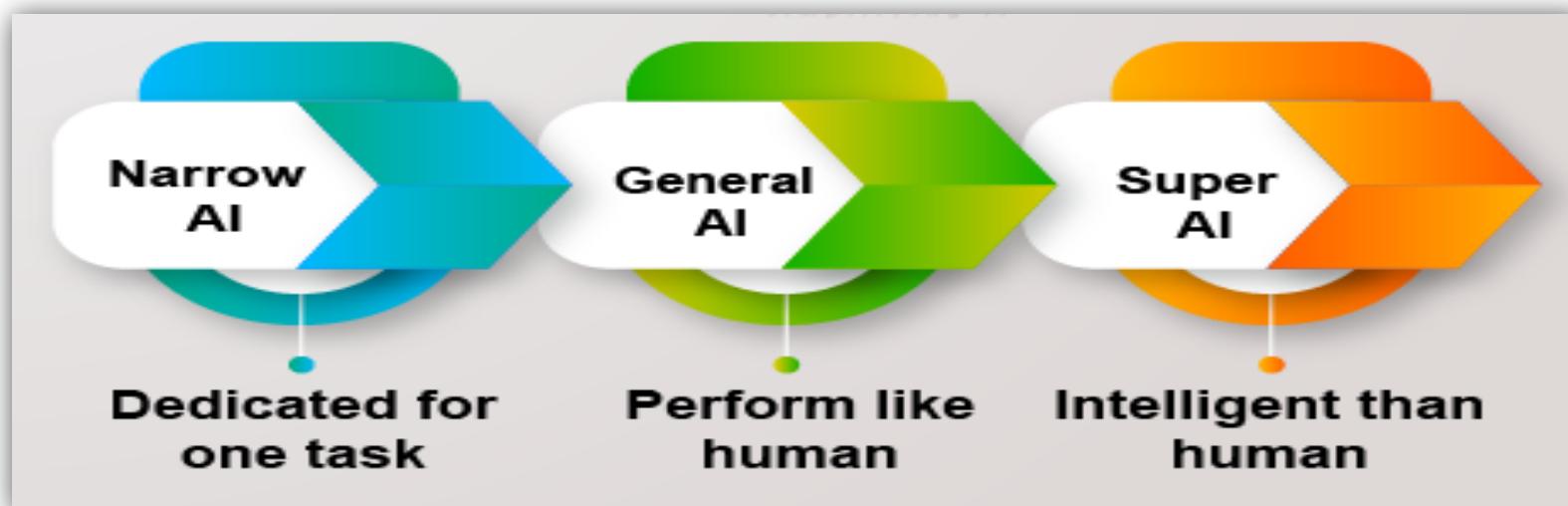
- **Reinforcement Learning:**

These models learn through trial and error, interacting with an environment and receiving rewards or penalties for their actions. Examples include training self-driving cars and game-playing agents.

- Labeled data has ***tags or labels*** that indicate what type of data it is or what category it belongs to like “cat” or “dog” for images.
- Unlabeled data, is ***raw data without these tags***, and the model has to learn patterns and features on its own.

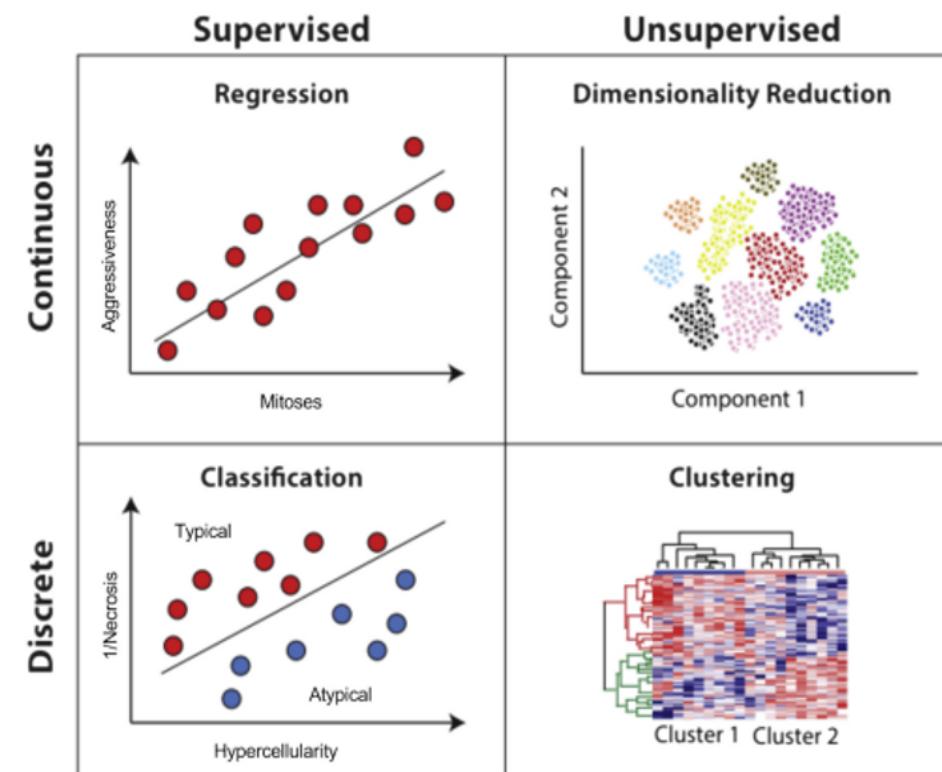
4.2 Model Classification (by Capability)

- **Narrow AI (Weak AI):**
Can perform specific tasks very well, but not general intelligence. Examples include image recognition and spam filtering.
- **General AI (Strong AI):**
Possesses human-level intelligence and can perform any intellectual task that a human can. Currently, this is a theoretical concept.
- **Super AI:**
Outperforms human intelligence in all areas. This is also theoretical concept.

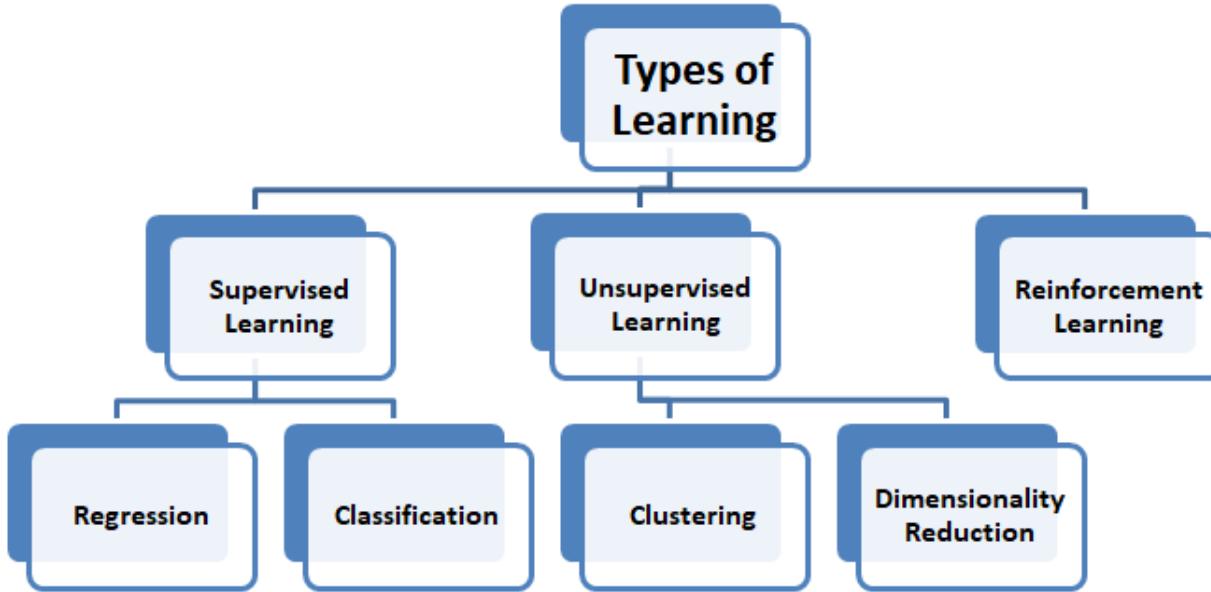


4.2 Model Classification (by function)

1. **Classification Models**: *Predict the category of an input.* Examples include determining if an *email is spam or identifying objects in an image.*
2. **Regression Models**: Predict a *continuous value.* Examples include predicting *house prices and stock prices.*
3. **Clustering Models**: *Group similar data points together.* Examples include customer segmentation and anomaly detection.
4. **Dimensionality Reduction Models**: Simplify data by reducing the number of input variables while keeping important information. Examples include improving model performance and visualizing high-dimensional data.



4.2 Types of Models



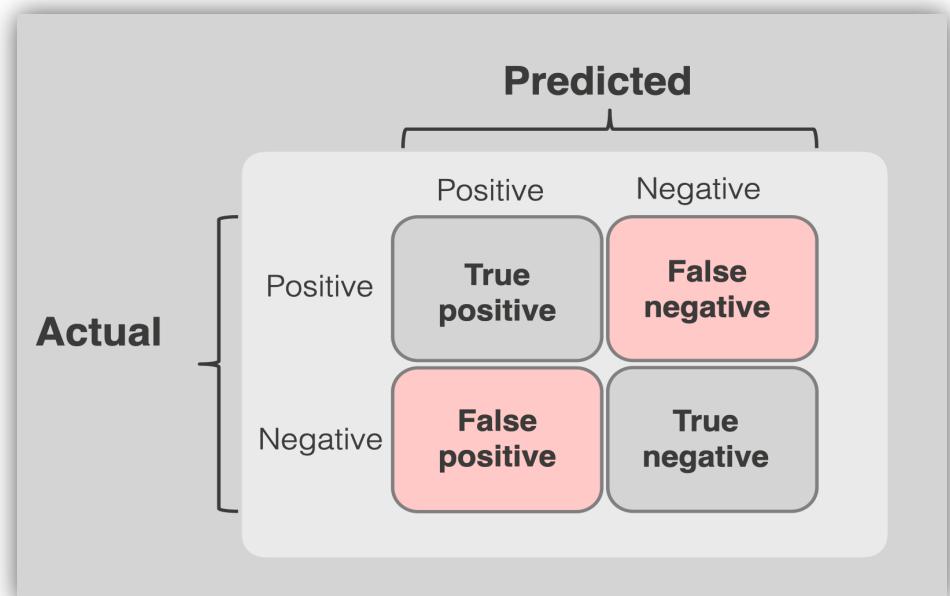
Supervised Learning	Unsupervised Learning	Reinforcement Learning
Has outcome information ("labels")	No outcome information available	Makes decisions based on trial and error
Finds patterns that relate to those outcomes	Analyzes or identifies groups without labels or human instruction	Decision-making algorithm is constantly refined based on "rewards"
Uses patterns to predict outcomes not yet known	Offers insights into characteristics that define groups	Excels in complex situations

5. Evaluation

After building the model, the next step is to test it using new data. This helps evaluate its performance and identify areas for improvement.

Key aspects of model evaluation:

- **Purpose:** To determine the model's effectiveness in making predictions and its ability to generalize to unseen data.
- **Metrics:** Various metrics are used to assess performance, such as **accuracy**, **precision**, **recall**, **F1-score**, and **AUC** (Area Under the Curve).
- **Methods:** Common methods include hold-out validation (splitting data into training (80%) and testing sets(20%)).
- **Benefits:** Model evaluation helps identify areas for improvement, guides model selection, and ensures the model is reliable and robust.



6. Deployment

- ✓ In an AI project cycle, ***deployment signifies moving a trained AI model from development to a real-world production environment*** where it can be used to solve practical problems.
- ✓ This ***involves integrating the model with existing systems, creating new applications***, or leveraging it for offline insights.

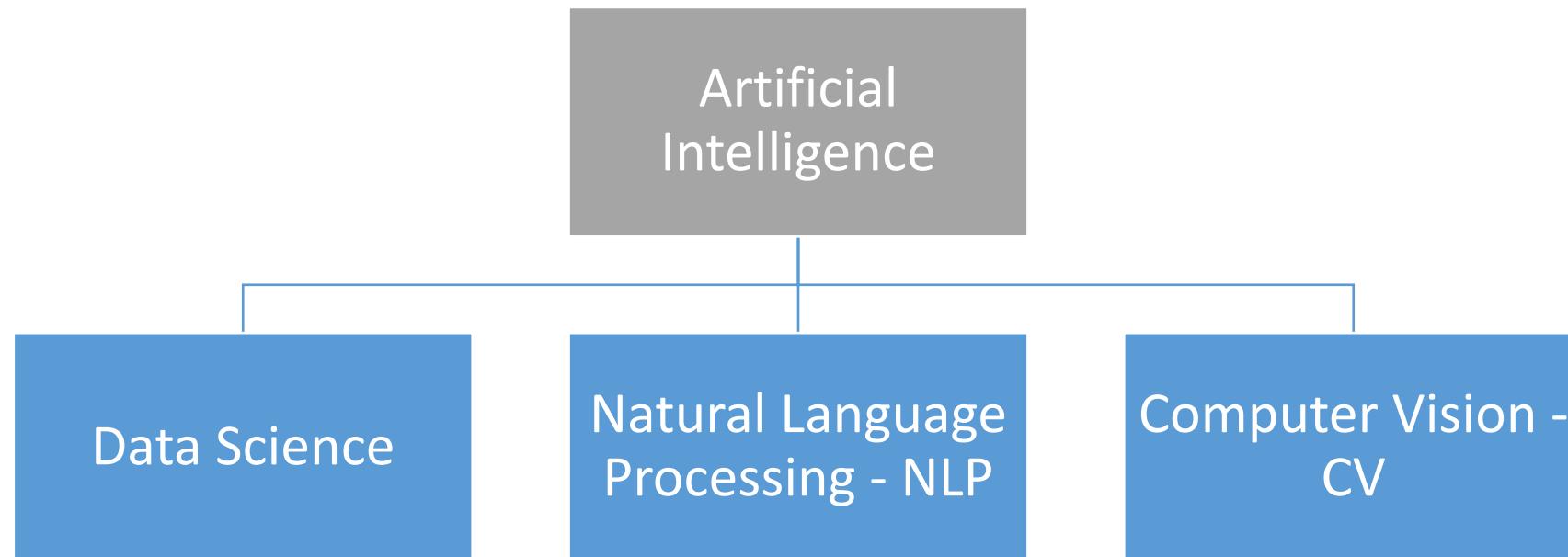


End of 1.1

1.1 AI Project Cycle (The Overview of the six stages of the AI Project Cycle)

AI Domains

1.2 Introduction to AI Domains (The three domains of AI and their applications)



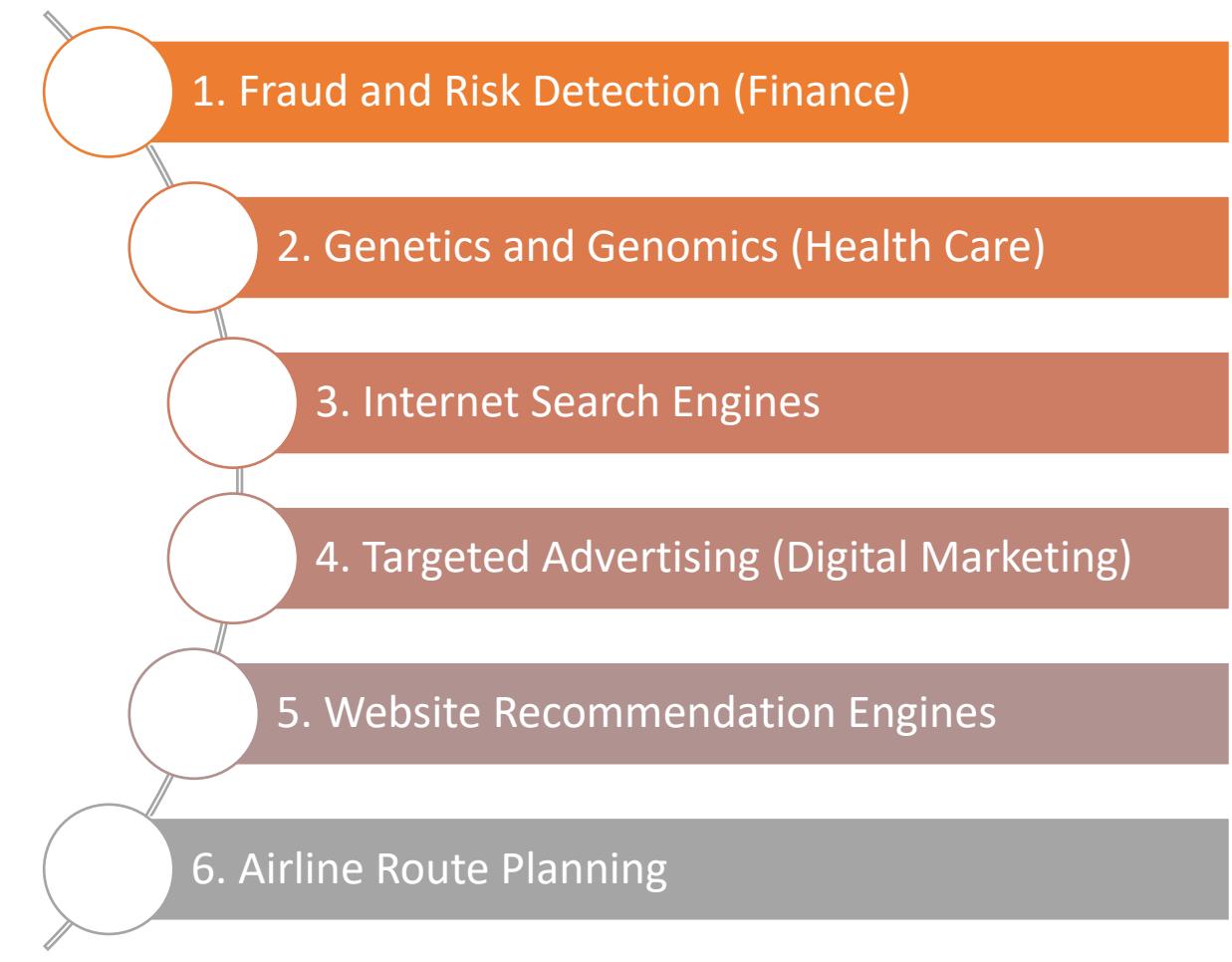
1. Data Science

Data Science is a part of AI that deals with collecting, organizing, and understanding data. It helps find useful information from data, which can be used to make better decisions.

Example:

Price Comparison Websites

These websites rely heavily on large amounts of data. They make it easy to compare prices for the same product from different sellers all in one place. Examples include PriceGrabber, PriceRunner, Jungle, Shopzilla, and DealTime. Today, such websites exist across various industries like technology, travel, automobiles, appliances, clothing, and more.

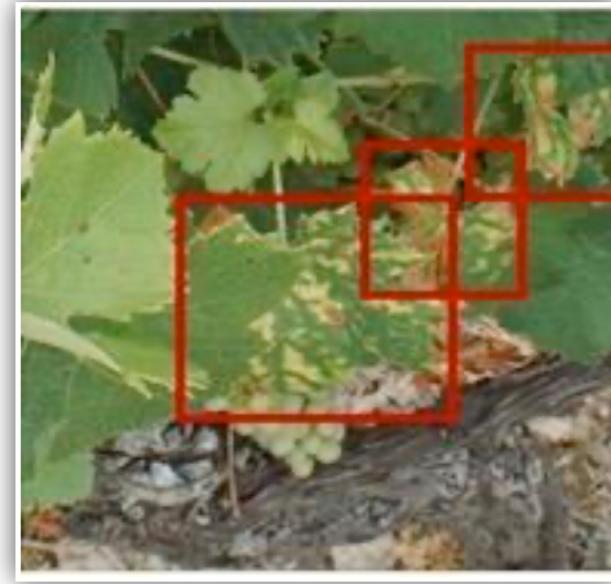


2. Computer Vision - CV

Computer Vision

Computer Vision (CV) is a field of AI that enables machines to ***understand and interpret visual information***. It involves capturing, processing, analyzing, ***and extracting useful data from images or videos to help machines make decisions***. The input can come from various sources like photos, videos, thermal or infrared sensors, and more.

The goal of computer vision is to teach machines how to gather information from visual data—such as ***pixels***—and convert it into a form they can understand. This allows systems to recognize patterns, objects, and actions, ultimately supporting decision-making processes.



Agricultural Monitoring

Computer vision is used in agriculture to monitor crops, detect pests, and estimate yields. Drones capture images of fields, which are analyzed to improve crop health and farming methods.

3. Natural Language Processing - NLP

Natural Language Processing (NLP)

NLP is a branch of AI that focuses on enabling computers to understand and process human language. It uses algorithms to extract meaning from spoken or written text, aiming to interpret and respond to language in a useful way.



Email filters

Email filters are one of the most basic and initial applications of NLP online. It started with spam filters, uncovering certain words or phrases that signal a spam message.



Machine Translation

NLP is used in translation tools like Google Translate to automatically convert text between languages by analyzing sentence structure and meaning.

Ethical Frameworks for AI

1.3 Ethical Frameworks of AI

Frameworks

Ethics are a set of values or morals which help us separate right from wrong.

- ✓ Ethical frameworks are frameworks which help us ensure that the choices we make do not cause unintended harm.

Why do we need Ethical Frameworks for AI?

- As we have seen how ***bias*** could result in unwanted outcomes in AI solutions.
 - ***Think of the hiring algorithm which was biased against women applicants!***
- AI is essentially being used as a decision-making / influencing tool.
 - As such we ***need to ensure that AI makes morally acceptable recommendations.***

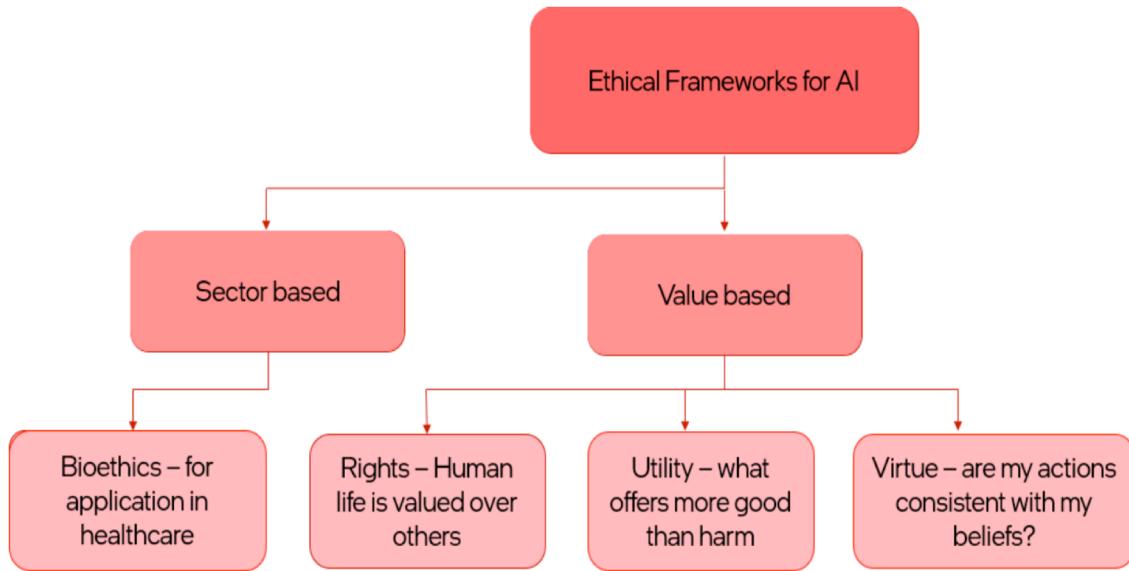
Ethical Considerations

(while developing a system)

Here are some key ethical considerations when developing AI models, briefly listed:

- 1. Bias and Fairness** – Ensure the model treats all users fairly and avoids discrimination.
- 2. Privacy** – Protect user data and avoid unauthorized data usage.
- 3. Transparency** – Make AI decisions understandable and explainable.
- 4. Accountability** – Clearly define responsibility for AI actions and outcomes.
- 5. Safety and Security** – Prevent harm and protect systems from misuse or attacks.
- 6. Consent** – Use data only with informed user consent.
- 7. Impact on Jobs** – Consider effects on employment and support fair transition.
- 8. Environmental Impact** – Minimize energy use and promote sustainable practices.

Types of Ethical Frameworks



Sector-based Frameworks

These are ethical guidelines designed for specific industries. For example, in healthcare, **bioethics focuses on patient privacy, data security, and the responsible use of AI**. Similar frameworks exist for finance, education, transport, agriculture, governance, and law enforcement.

Value-based Frameworks

These frameworks are based on core ethical principles and moral values. They help assess the rightness of actions and guide ethical decision-making and behavior.

- **Bioethics:** Bioethics is the study of ethical issues in healthcare and life sciences. It focuses on topics like patient rights, data privacy, medical research, and the responsible use of technology in medicine.
- ✓ **Rights-based:** Focuses on protecting human rights, dignity, and individual freedoms, ensuring AI does not harm or discriminate.
- ✓ **Utility-based:** Aims to maximize overall benefit and reduce harm, balancing AI's advantages with its risks to society.
- ✓ **Virtue-based:** Emphasizes the ethical character of individuals involved, encouraging honesty, compassion, and integrity in AI development and use.

Principles of bioethics:

- Respect for Autonomy.
- Do not harm.
- Ensure maximum benefit for all.
- Give justice.

Nonmaleficence is **the obligation of a physician not to harm the patient**. This simply stated principle supports several moral rules – do not kill, do not cause pain or suffering, do not incapacitate, do not cause offense, and do not deprive others of the goods of life.

- “**Non-maleficence**” refers to the ethical principle of avoiding causing harm or negative consequences. It emphasizes the obligation to minimize harm as much as possible and prioritize actions that prevent harm to individuals, communities, or the environment.
- “**Maleficence**” refers to the concept of intentionally causing harm or wrongdoing.
- “**Beneficence**” refers to the ethical principle of promoting and maximizing the well-being and welfare of individuals and society. It emphasizes taking actions that produce positive outcomes and contribute to the overall good, ensuring that the greatest benefit is achieved for all stakeholders involved.



Student

PORTFOLIO.



Case-Based Questions

1. Case Scenario:

In a corporate setting, a multinational company is facing scrutiny over its environmental practices, particularly regarding the disposal of industrial waste. The company has historically prioritized profit maximization and cost-cutting measures, leading to practices that result in environmental harm and negative impacts on local communities. As public awareness and concern about environmental sustainability grow, stakeholders, including investors, customers, and advocacy groups, are calling for the company to adopt more responsible and sustainable business practices.

Question

Drawing from the case presented, analyze the ethical considerations surrounding the company's environmental practices through the lens of value-based frameworks in ethics.

Submission date on or before 23rd April 2025 before 3pm

Student

PORTFOLIO.



Case-Based Questions

2. Case Scenario:

In a rural farming community, a group of small-scale farmers is faced with a dilemma regarding the use of pesticides on their crops. The farmers have traditionally relied on chemical pesticides to control pests and maximize crop yields. However, concerns have been raised about the potential environmental and health impacts of pesticide use, including soil contamination, water pollution, and adverse effects on human health. Additionally, neighboring communities and environmental advocacy groups have expressed opposition to the widespread use of pesticides, citing ecological damage and risks to biodiversity.

Question

Using the case provided, examine the ethical considerations surrounding pesticide use in the agricultural sector, applying ethical frameworks to analyze the competing interests and values at stake.

Submission date on or before 23rd April 2025 before 3pm

1.1. What is the purpose of defining the problem statement during the Problem Scoping stage in an AI project cycle?

- A) To collect data
- B) To understand the aim and objective of the project
- C) To train the model
- D) To process data

2. In what ways can AI models be categorized based on the type of data fed into them?

- A) Two domains
- B) Four domains
- C) Three domains
- D) Five domains

3. In Statistical Data, what is the primary function of the system in relation to data?

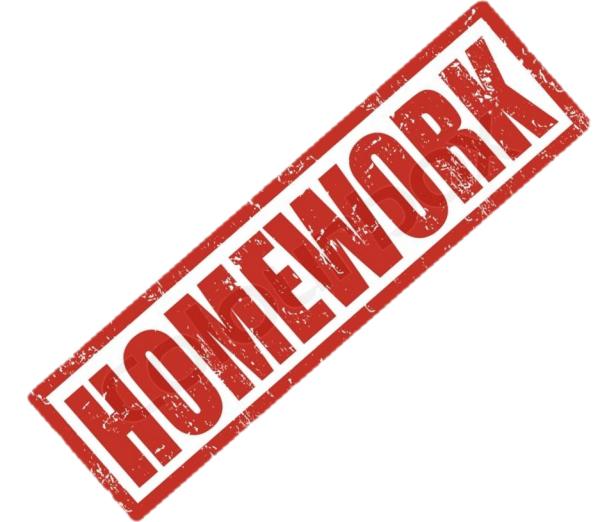
- A) Generating large datasets
- B) Analyzing data to extract insights
- C) Converting data into images
- D) Distributing data across networks

4. What is the main goal of Computer Vision projects?

- A) Translating audio data into visual descriptions
- B) Converting digital data into analogue signals
- C) Teaching machines to understand textual information
- D) Converting digital visual data into computer-readable language

5. What is the primary focus of NLP?

- A) Analyzing computer languages
- B) Interacting between computers and humans using artificial language
- C) Dealing with the interaction between computers and humans using natural language
- D) Enhancing human-to-human communication



6. What do frameworks provide in the context of problem-solving?

- A) Random solutions
- B) Step-by-step guidance
- C) Legal advice
- D) Ethical justifications

7. How are Ethical Frameworks for AI categorized?

- A) Into legal and illegal frameworks
- B) Into sector-based and value-based frameworks
- C) Into historical and contemporary frameworks
- D) Into theoretical and practical frameworks

8. What is the central focus of virtue-based value-based frameworks?

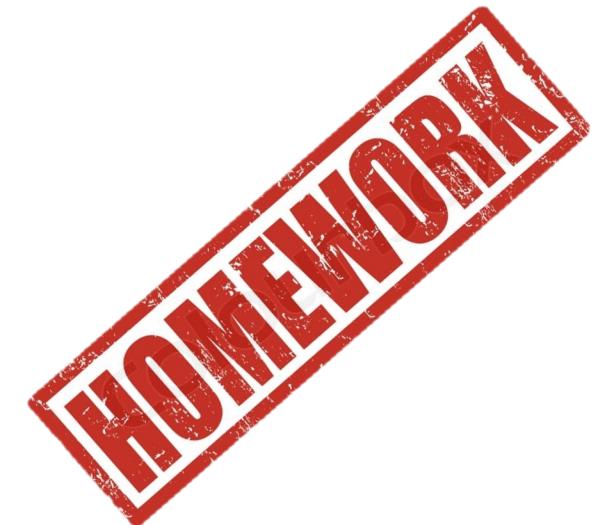
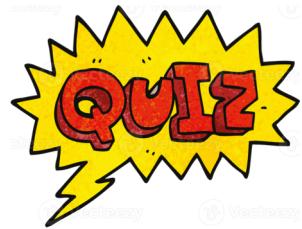
- A) Maximizing utility
- B) Protecting human rights
- C) Aligning actions with ethical principles and beliefs
- D) Ensuring compliance with legal regulations

9. Which of the following best describes rights-based value-based frameworks?

- A) Prioritizing human rights and dignity, valuing human life over other considerations
- B) Evaluating actions based on maximizing overall good and minimizing harm
- C) Centering on the character of the decision-maker and the alignment of actions with personal or societal virtues
- D) Focusing on achieving outcomes that offer the greatest benefit

10. What is the primary domain of application for Bioethics?

- A) Agriculture
- B) Healthcare and life sciences
- C) Information technology
- D) Environmental conservation





11. Assertion: Ethics provide guidance in distinguishing right from wrong.

Reasoning: Ethics consist of a set of values and morals that aid individuals in making moral judgments and decisions.

- A) Both Assertion and Reasoning are true, and Reasoning is the correct explanation of the Assertion.
- B) Assertion is true, but Reasoning is false.
- C) Both Assertion and Reasoning are true, but Reasoning is not the correct explanation of the Assertion.
- D) Assertion is false, but Reasoning is true.

12. Assertion: Value-based frameworks in ethics provide guidance by focusing on fundamental ethical principles and values.

Reasoning: These frameworks reflect different moral philosophies guiding ethical reasoning and are concerned with assessing the moral worth of actions.

- A) Both Assertion and Reasoning are true, and Reasoning is the correct explanation of the Assertion.
- B) Assertion is true, but Reasoning is false.
- C) Both Assertion and Reasoning are true, but Reasoning is not the correct explanation of the Assertion.
- D) Assertion is false, but Reasoning is true.





Reflection Time:

HOMEWORK

1. Outline the main steps in the AI Project Cycle briefly.
2. What roles does computer vision play in agricultural monitoring systems?
3. Mention the factors which knowingly or unknowingly influence our decision-making.
4. What is the necessity for Ethical Frameworks in AI development?
5. Mention the key characteristics of sector-based frameworks.
6. What do you mean by Bioethics?
7. What is Natural Language Processing? Explain any two real-life applications of NLP.
8. How do value-based frameworks contribute to ethical decision-making by emphasizing fundamental principles and values?



2025-2026