Code 417

Under NSQF

artificial intelligence

A Textbook of Class X

Part A : Employability Skills
Part B : Vocational Skills

SUMITA ARORA

DHANPAT RAI & Co.

VOCATIONAL SKILLS

PART B

UNIT 1

Introduction to AI

IN THIS UNIT

SESSION 1

Foundational Concepts of AI

- What is Intelligence ?
- Role of Intelligence in Our Lives
- Types of Intelligence
- What is Decision-Making ?
- Role of Intelligence in Decision-Making
- What AI Is ?
- What AI Is Not ?
- How Machines become Intelligent ?

1.2 WHAT IS INTELLIGENCE ?

Intelligence refers to the ability to acquire and apply knowledge and skills in various domains. In other words, intelligence is the :

- faculty of understanding
- ability to interact with the real world
- capacity of learning, reasoning and understanding *e.g.*, recognising speech, recognising objects and images
- application of acquired knowledge *e.g.*, ability to take action: to have an effect
- aptitude in grasping truths, feels, meaning etc. *i.e.*, continuous learning and adapting graph
- mental alertness
- using mental capacity and knowledge for decisions, such as :
  - Modelling the external world, given input
  - Solving new problems, planning and making decisions
  - Ability to deal with unexpected problems, uncertainties

Whosoever has the above mentioned qualities, would be termed *intelligent* – be it a human, an animal or a machine. These days you see even machines are getting intelligent – they can learn from facts, increase their knowledge and perform intricate decision-making.

Intelligence

Intelligence refers to the ability to understand, distinguish, question things/objects/feelings/situations/people along with acquiring and applying knowledge and skills in various domains.

1.4 TYPES OF INTELLIGENCE

Intelligence can be of various forms, such as intelligence with numbers or music or understanding own and others' feelings and so on. In 1983, an American developmental psychologist *Howard Gardener* described *nine* types of intelligence :

1. **Naturalist Intelligence** *(Environmental Skills)*. It refers to the human ability to identify and categorise among living things (plants, animals etc.) as well as understanding other features of the natural world (clouds, rock configurations etc.).

2. **Musical Intelligence** *(Music Skills)*. It refers to the ability to differentiate between *pitch*, *rhythm*, *timbre*, and *tone*. This type of intelligence enables to recognize, create, reproduce, and reflect on music. The *music composers, conductors, musicians, vocalists*, and *sensitive listeners* exhibit this intelligence.

Ravi Shankar (celebrated Sitar player), *A.R. Rehman* (great music composer), *Ustad Zakir Hussain* (celebrated Tabla Player) are some great Indians to possess musical intelligence.

3. **Logical-Mathematical Intelligence** *(Numerical and Logical Skills)*. It refers to the ability to *calculate, quantify, use propositions and hypotheses*, and carry out *complete mathematical operations*. It also includes *reasoning skills, inductive and deductive thinking patterns, logical intelligence* etc.

Dr. Neena Gupta was awarded the prestigious Shanti Swarup Bhatnagar Prize in 2019, at the age of 35, for having solved a 70-year-old mathematics puzzle called *the Zariski Cancellation Problem*. She possesses great logical-mathematical intelligence.

4. **Existential Intelligence** *(Religious and Spiritual Skills)*. It refers to that form of intelligence which people use to ponder over deep questions such as reason of human existence, meaning of life, why we die, and how we got here, and so forth.

India is a land of Sufis and Saints who possess existential intelligence in abundance.

5. **Interpersonal Intelligence** *(People Skills)*. It refers to the ability to understand and interact effectively with others. It includes *effective communication (verbal and non-verbal), sensitivity to the moods and temperaments of others,* and the *ability to understand multiple perspectives.*

6. **Bodily-Kinesthetic Intelligence** *(Mind Body Skills)*. It refers to the intelligence that helps manipulate objects and use a variety of physical skills, such as skills involving right timing with mind–body union. *Athletes, dancers, surgeons*, and *crafts people* exhibit well-developed bodily-kinesthetic intelligence.

Indian golden boy *Neeraj Chopra* has excellent bodily-kinesthetic intelligence.

7. **Linguistic Intelligence** *(Language Processing Skills)*. This is the ability to think in words and to use language to express and appreciate complex meanings. *Poets, novelists, journalists*, and *effective public speakers* are said to have this form of intelligence. This intelligence is required for activities like writing, reading, telling stories or doing crossword puzzles.

8. **Intra-personal Intelligence** *(Self-awareness Skills)*. It refers to the ability to understand oneself and one's thoughts and feelings, and to use such knowledge in planning and directing one's life. This form of intelligence is evident in *psychologists, spiritual leaders*, and *philosophers*.

9. **Spatial Intelligence** *(Visual World Perceiving Skills)*. This form of intelligence makes good use of *three dimensions, mental imagery, 3-D reasoning, image manipulation, graphic* and *artistic skills*, and an active *imagination*. *Sailors, pilots, sculptors, painters*, and *architects* all exhibit spatial intelligence.

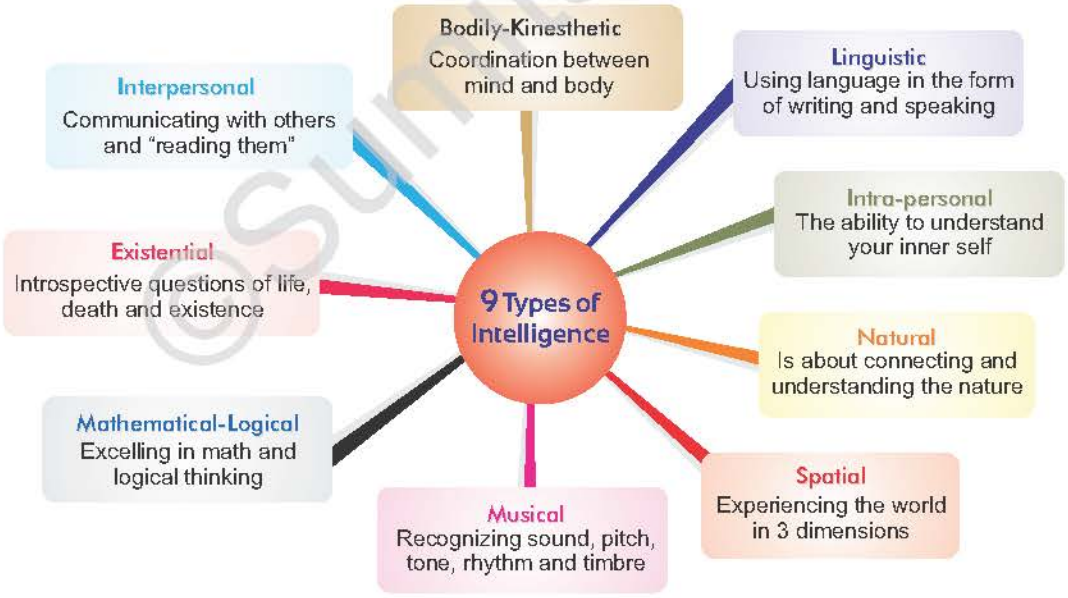


Figure 1.1 Different types of Intelligence

Note

A person can possess multiple types of intelligence simultaneously.

Figure 1.2 depicts this, *i.e.*, the role of intelligence in decision-making process.

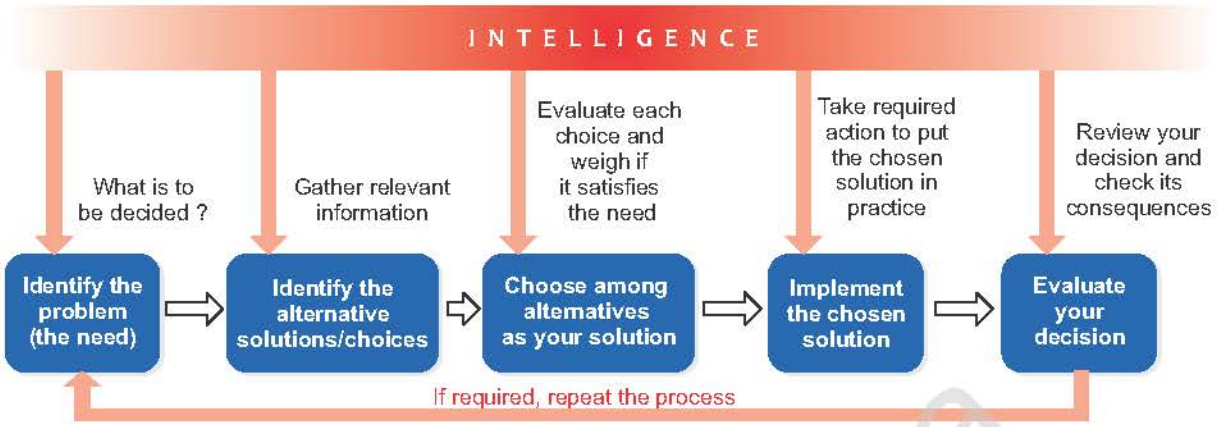


Figure 1.2 Role of Intelligence in Decision-Making

Consider the scenarios in the following activity, where intelligence played an important role in decision-making.

1.7 WHAT AI IS ?

You know that the ***Intelligence*** is basically the faculty of reasoning and integration of knowledge. Artificial Intelligence is something which is not natural yet mimics intelligence.

Let us make it clearer. "*Artificial intelligence*" refers to a human-made interface (a machine or framework or application) having the power to reason and integrate knowledge and behave in the same way a naturally intelligent entity (human/animal etc.) would behave.

Artificial Intelligence

Artificial Intelligence refers to the ability of a human-made interface (machine or application) to mimic human-like intelligence, *i.e.*, reasoning and integration of knowledge.

Some examples of *what AI is and can do*, are :

- Like humans, AI based systems can discover patterns from the available information and attempt to make estimates for future, *e.g.*, farmers have discovered the pattern of higher average temperatures leading to higher rainfall and thus they make estimates accordingly.
- Like humans, AI can make decisions by picking up from the available options, *e.g.*, upon falling a robot may decide to get up or turn direction or even crawl.
- Like humans, AI based self-driven cars can make decisions like if there is a human or animal or an object in front of it; to stop at red-signals; to activate windshield wiper in case of the rain.
- Like humans, AI based technologies can recognise and read from images, *e.g.*, AI based cameras can identify the traffic violators and initiate the process of challan. AI based cameras can identify human faces from live pictures or photos.
- Like humans, AI based systems can converse in natural language. Siri, Alexa, Google Assistant, Cortana etc. are all examples.

and many more.

PART B

What can't AI systems do yet ?

- Understand natural language robustly (*e.g., read and understand articles in a newspaper*)
- Surf the web
- Interpret an arbitrary visual scene
- Learn a natural language
- Play well
- Construct plans in dynamic real-time domains
- Refocus attention in complex environments
- Perform life-long learning

1.9 HOW MACHINES BECOME INTELLIGENT ?

In order to understand how machines become (artificially) intelligent, it is important for you to know about various AI domains and branches. AI mainly has the following domains and branches :

- **Machine Learning.** The branch of AI that teaches a machine how to make inferences and decisions based on past experience. The learning of a machine can take place in supervised/unsupervised/semi-supervised or reinforced manner.
- **Deep Learning.** The branch of AI that teaches a machine to process inputs through layers in order to classify, infer and predict the outcome.

- **Neural Networks.** The branch of AI that works similar to human neural networks where multiple layers capture the relationships among data and process them as per the need.
- **Natural Language Processing.** The branch of AI that teaches a machine to read, understand and interpret a natural language, and provide a response in a natural language.
- **Computer Vision.** The branch of AI that helps machines recognise an image by breaking down an image and studying different parts of the objects.

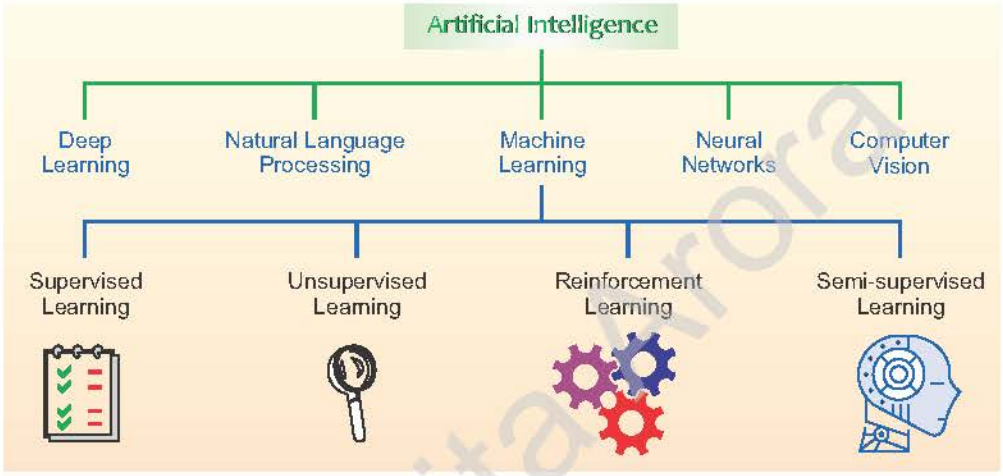


Figure 1.3 AI Domains and Branches

LET US REVISE

- *Intelligence refers to the ability to understand, distinguish, question things/objects/feelings /situations/people along with acquiring and applying knowledge and skills in various domains.*
- *Animals and humans possess multiple types of intelligence.*
- *Psychologist **Howard Gardener** described 9 types of intelligence.*
- *The **nine** types of intelligence are : Naturalistic intelligence, Musical intelligence, Logical-mathematical intelligence, Existential intelligence, Interpersonal intelligence, Linguistic intelligence, Bodily-kinesthetic intelligence, Intra-personal intelligence and Spatial intelligence.*
- *Decision-Making is the process of identifying and picking a final choice/action/item/belief for a need, from an available set of choices, after carefully assessing the available options.*
- *In order to make good decisions, we need proper information about the situation, needs, available factors and challenges.*
- *Intelligence is an outcome of the data/information available, past experience, knowledge, intuition and many other factors.*
- *Artificial Intelligence refers to the ability of a human-made interface (machine or application) to mimic human-like intelligence, i.e., reasoning and integration of knowledge.*
- *AI means mimicking human intelligence.*
- *AI does not mean just the automation.*
- *AI is not a single entity or some digital magic.*
- *AI is the result of data, math, algorithm and continuous improvement.*
- *Machines keep updating and learning from more data and past experiences through training, validation and testing and thereby become intelligent.*

SESSION 2

AI Domains and Technologies

- AI Technologies
- Domains of AI

2.3 DOMAINS OF AI

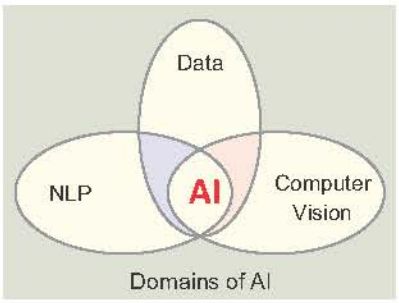
Artificial Intelligence is not a single technology. It is a mix of technologies. Broadly there are *three* domains of AI :

(*i*) Data (Data Science) (*ii*) Computer Vision

(*iii*) NLP (Natural Language Processing)

(*i*) Data / Data Science

This domain of AI may deal with data (structured and unstructured data) and specific methods to use it in AI applications. Data science is a field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from many structural and unstructured data.



Domains of AI

So, if your AI powered refrigerator knows in advance that what types of food you would prefer in summers, it is because it has all your past data and it applied data science methods to determine your choice this summer.

Data science is related to *data mining, machine learning* and *big data*.

- ***Data Mining*** is a process of finding potentially useful patterns from huge data sets.
- ***Machine Learning (ML)*** is a process of automatic learning of computers and machines by discovering insights from data with experience. ML is also used to make predictions about future outcomes.

Data Science

Data Science is a field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from many structural and unstructured data to apply in AI applications.

- ***Big Data*** refers to huge amounts of data, which is regularly growing at an exponential rate, *e.g.*, data of social media (posts, pictures, responses, users, etc.). Big data cannot be handled without AI.

(ii) Computer Vision (CV)

Computer vision is a field of artificial intelligence that trains computers to interpret and understand the visual world. Using digital images from cameras and videos and deep learning models, computers and machines can accurately identify and classify objects — and then react to what they "see".

Computer Vision

Computer Vision refers to the training of computers to have a vision somewhat like humans so that machines can accurately identify and classify objects and produce specific 'reactions'.

So, if your AI powered lock with camera can identify people and then opens door only for authentic members, it is all because of computer vision.

Note

DeepFace (Facebook's deep learning AI specialising in computer vision) has 97.25% accuracy, regardless of lighting conditions or angles.

PART B

(iii) NLP (Natural Language Processing)

Natural Language Processing is a subfield of linguistics and AI concerned with the interactions between computers and human languages, in particular how to program computers to process and analyse large amounts of natural language data.

NLP

Natural Language Processing (NLP) is a subfield of AI, mainly concerned with the interactions between computers and human languages in the form of speech (natural language) input and speech (natural language) output.

It has two sub-categories :

- **Natural-Language Understanding**, which helps machine recognise natural speech (*e.g.*, we instruct *Google Assistant* or *Siri* in our speech) and understand its meaning.
- **Natural-Language Generation**, which helps in producing a response in natural language *e.g.*, speech response by a virtual assistant like *Cortana*, *Google Assistant* or *Siri*. It is closely related to Automated Speech Recognition (ASR).

So, if you buy a robot to assist you in work and it is capable of understanding your language and speech, and can carry out work and also talk to you – it is all thanks to NLP.

PART B

LET US REVISE

- *Artificial Intelligence (AI) refers to ability of modern machines and computers to mimic human intelligence. AI machines work through data, algorithms and models.*
- *Most used of AI technologies are Machine Learning and Deep Learning.*
- *Machine Learning (ML) is a branch of AI that enables machines to automatically learn and improve at tasks with experience and by the use of data.*
- *Deep Learning (DL) is a subset of machine learning where learning takes place through examples by filtering the input data using layers and rules-based algorithms to predict and classify information.*
- *AI has three domains :* ***Data and Data Science, Computer Vision*** *and* ***Natural Language Processing.***
- ***Data science*** *is a field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from many structural and unstructured data to apply in AI applications.*
- ***Computer vision*** *refers to the training of computers to have a vision somewhat like humans so that machines can accurately identify and classify objects – and then react to what they "see".*
- ***Natural Language Processing (NLP)*** *is a subfield of AI, mainly concerned with the interactions between computers and human languages in the form of speech (natural language) input and speech (natural language) output.*
- *FaceBook's* ***DeepFace*** *technology is based on Computer Vision AI.*
- *Modern day chatbots and virtual assistants are based on NLP.*

SESSION 3

Real Life AI Applications

AI Applications

3.2 REAL LIFE AI APPLICATIONS

Artificial Intelligence has impacted many fields of our lives. We are already benefitting from AI applications in real life in various ways. Listed below are some common real life applications of AI.

1. AI in Healthcare

Artificial Intelligence has transformed patient care and diagnoses in many ways through robotics, data and many others. AI's ability with data has revolutionised healthcare by suggesting timely treatment, predicting ailments, future medical needs and so on.

Note

Artificial Intelligence can analyse data on patient visits to the clinic, medications prescribed, lab tests, and procedures performed, data outside the health system – such as social media, purchases made using credit cards, census records, Internet search activity logs that contain valuable health information, and comes up with precise diagnosis, prediction, and suggests about timely requirements.

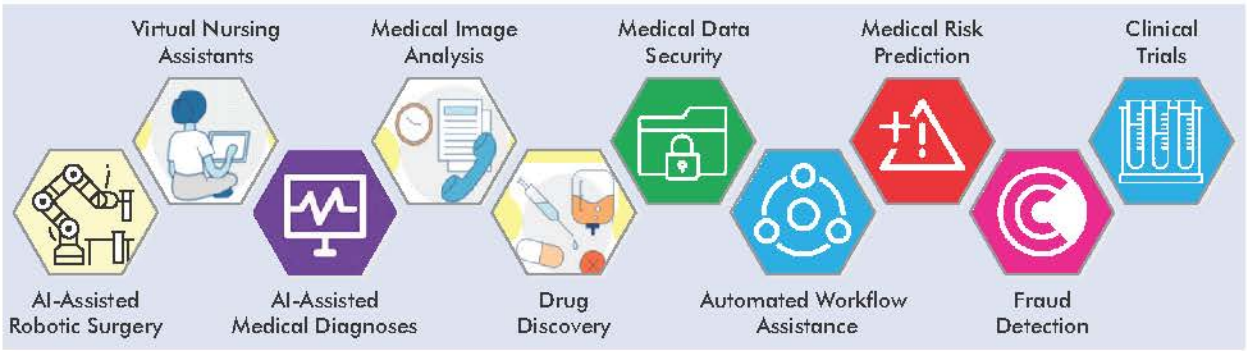


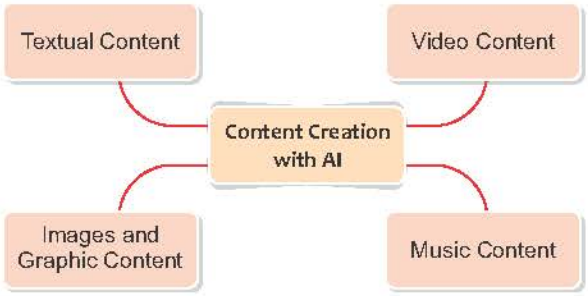
Figure 3.1 AI in Healthcare

2. AI in Content Creation

Artificial Intelligence based applications of today can generate text content in the form of articles, blogs etc., video content, images and graphics, music content and so forth (Fig. 3.2).

Note

AI has great capability of processing and analysing **humongous** data and predict trends.



"AI can write just like me... I've seen how OpenAI's GPT2 system can produce a column in my style."

– Guardian journalist Hannah Jane Parkinson

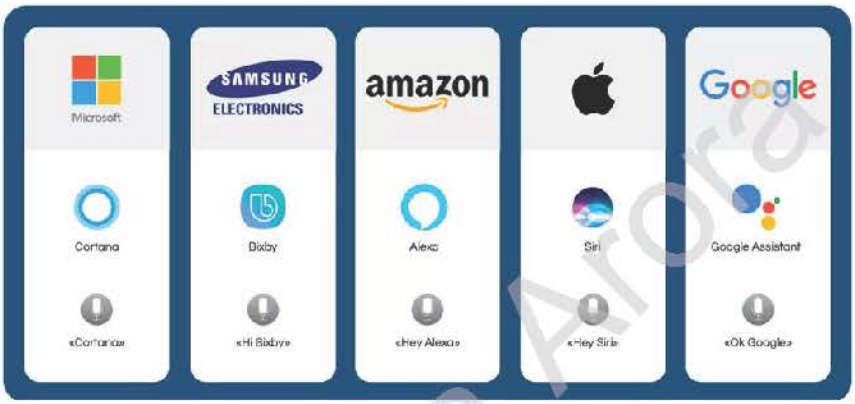
Figure 3.2 Content Creation with AI

PART B

3. AI in Chatbots and Virtual Assistants

Virtual Assistant

A **virtual assistant** is an application program that understands natural language voice commands and completes tasks for a user, while providing a variety of remote services to a business.



Chatbot

A **chatbot** is an artificial intelligence product or software which can simulate a real-life conversation on the chat which happens between the user and the company.

4. AI in Autonomous Vehicles

Autonomous vehicles or the driverless vehicles, which was beyond imagination until a few years ago, is happening in reality now - all thanks to AI which using ***computer vision*** and ***data*** accomplishes this.

5. AI in GPS (Global Positioning Systems)

These days smartphones double as car navigation devices (**Google Maps**), smartwatches can be a hiking trail guide and so on.

AI is helping navigation very efficient considering other data such as weather, environment, electromagnetic interference, trees, buildings, line-of-sight and so on.

6. AI in Sales and Marketing

These days online shopping is so common that we keep buying real and digital things online, *e.g.*, groceries from ***Blinkit*** or ***Big Basket***, clothes and shoes from ***Myntra***, household items and books from ***Amazon***, digital content from ***Netflix*** or ***HotStar***, music from ***Spotify*** and so on. Based on our history of past purchase and searches, these sites/apps keep recommending us items which are due for purchase or are of our liking. AI plays a huge role in such recommendations as our data is processed through AI to reach at these recommendations.

7. AI in Logistics and Supply Chain

When we buy physical items online (such as from ***Amazon*** or ***FlipKart*** or ***Myntra***), these are delivered to us through a supply chain. AI plays an important role in logistics and supply chain management.

AI helps in warehouse management through computer vision, can provide optimised routes for transportation of goods, can predict demand as per seasons and thereby requirement of warehouses for such things and so on.

8. AI in Gaming

Artificial Intelligence (AI) is taking the gaming experience to the next level by making the game having better and better visual experience with high-end graphics and giving an immersive and interactive game experience.

9. AI in Agriculture

These days harvesting robots are available that can harvest crops at a considerably higher volume and a faster pace as compared to humans. Other than this, AI is now playing another important role in agriculture through its deep learning and computer vision algorithms to process data

10. AI in Robotics

Artificial intelligence (AI) and robotics are a powerful combination for automating tasks inside and outside of the factory setting. In recent years, AI has become an increasingly common presence in robotic solutions, introducing flexibility and learning capabilities in previously rigid applications.

Let Us Revise

- *AI has found applications in nearly all fields and arenas, such as healthcare, content creation, chatbots and virtual assistants, autonomous vehicles, navigation through GPS, Sales and marketing, logistics and supply chains, gaming, robotics and so forth.*
- *AI has great capability of processing and analysing humongous data and predict trends.*
- *Artificial Intelligence based applications of today can generate text content in the form of articles, blogs etc., video content, images & graphics, music content and so forth.*
- *A virtual assistant is an application program that understands natural language voice commands and completes tasks for a user, while providing a variety of remote services to a business.*
- *A chatbot is an artificial intelligence product or software which can simulate a real-life conversation on the chat which happens between the user and the company.*
- *Driverless cars are a reality because of AI's using computer vision and data processing capabilities.*
- *GPS, or the Global Positioning System, is a global navigation satellite system (made up of at least 24 satellites) that provides location, velocity and time synchronization.*
- *Data science domain of AI plays an important role in sales and marketing.*
- *AI provides immersive and interactive gaming experience to the players after analysing their playing style, past choices and preferences.*
- *AI's all domains such as NLP (Natural Language Processing), CV (Computer Vision) and Data play important roles in creating human like robots.*

SESSION 4

AI Ethics

- Ethical Issues Around AI
- AI Bias and AI Access

4.2 ETHICAL ISSUES AROUND AI

The dictionary defines ethics as, *"the moral principles that govern a person or a group's behaviour or actions"* **Or** *"the moral correctness of a conduct or action"*. In short, **Ethics** are the moral responsibility of anyone or anything that can impact others.

Figure 4.1 lists various ethical issues surrounding AI.

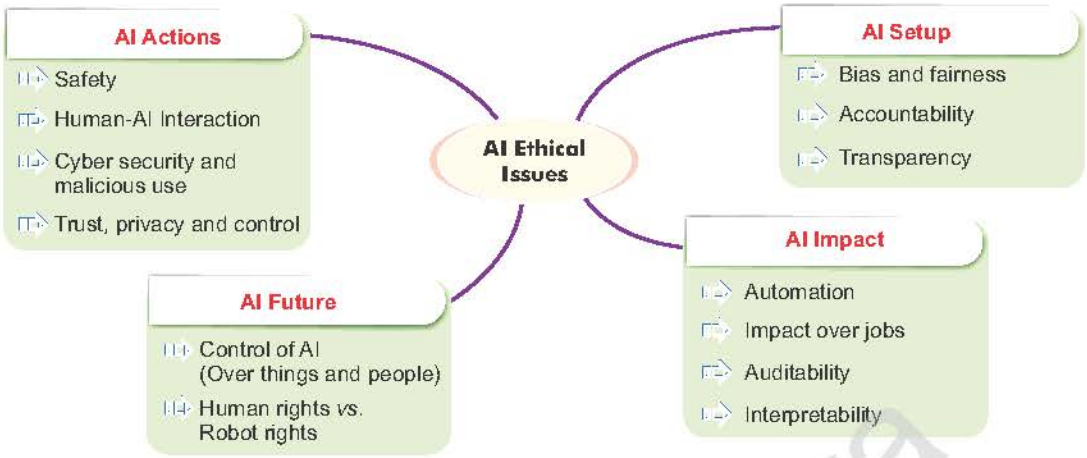


Figure 4.1 Ethical Issues of AI

4.2.1 Examples of AI Ethical Issues

Let us discuss some examples of AI ethical issues.

1. Bias and Fairness

Ethically an AI system should be free from all types of biases and be fair, *e.g.*, an AI system designed for picking candidates for a job must not be biased against any gender, race, colour or sexuality and so forth. It should be free from all such things and be totally fair.

2. Accountability

AI learns and evolves over time and data. What if an evolved algorithm makes some big mistake? Who would be accountable for it? For instance, when an autonomous Tesla car hit a random pedestrian during a test, Tesla was blamed and not the human test driver sitting inside, and certainly not the algorithm itself. But what if the program was created by dozens of different people and was also modified with each incident and more data available? Can the developers or the testers be blamed then?

3. Transparency

Transparency means nothing is hidden and everything that AI performs is explainable. Transparency ensures that there is full information and knowledge about these :

- data used, its range, interval and sources etc.
- models used are appropriate for the context make sense,
- models are thoroughly tested
- why particular decisions are made

4. Safety

AI technology, tools and practices should be so implemented such that they cause no direct or indirect harm to *data, people* and the *outcomes*. AI practices must be safe to ensure the well being of individual persons and the public welfare. AI practices must uphold public trust through the responsible use of technologies.

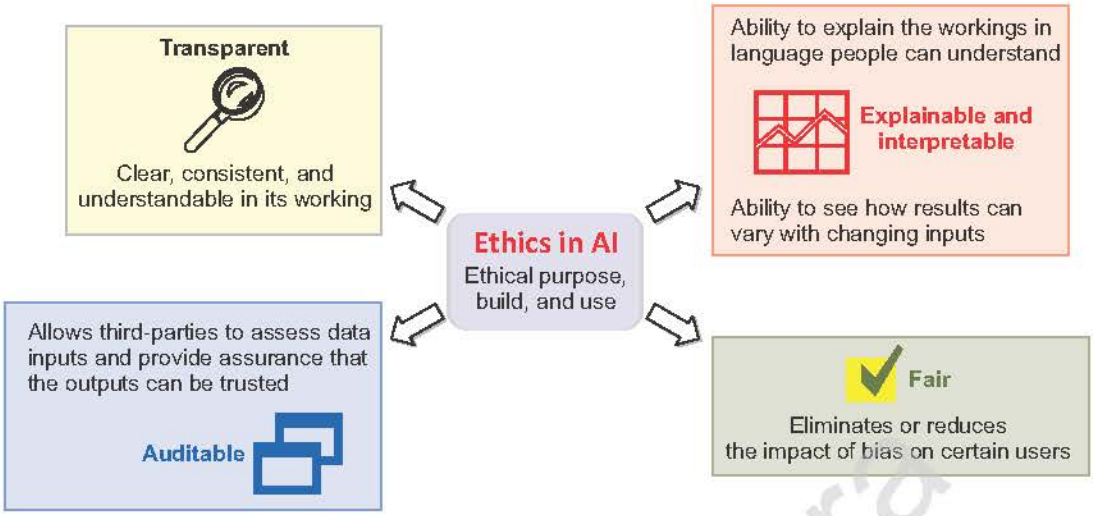


Figure 4.2

5. Human AI Interaction

AI must not deceive humans or other living beings, and it must not threaten or violate human dignity in any way.

PART B

6. Trust, Privacy and Control

Improved AI "faking" technologies make what once was reliable evidence — into unreliable evidence – this has already happened to *digital photos*, *sound recordings*, and *video*.

Thus, it is the ethical responsi- bility of the creator and user of AI to ensure that these are not misused.

7. Cyber Security and Malicious use

It is the ethical responsibility of an organisation to have human control over AI usage in terms of its span and control so that it is not available to hackers for malicious use.

8. Automation and Impact over Jobs

AI does not mean that jobs are reduced, it just means that the nature of jobs and work is predominantly changing.

It is the ethical responsibility of an organisation to upgrade the skillset of its workers so that they upgrade their skillset and be ready for futuristic AI oriented jobs. It is ethical responsibility of governments too (equally and even more) to bring appropriate changes.

9. Human Rights in the Age of AI

AI has generated new form of threats

- huge risk to data privacy and protection — *violates human right to privacy.*
- biased decision and *violates human right to fair chance and justice.*

- As AI can process humongous sets of data, it can analyse huge symptoms dataset of a person and can predict may possible future ailments and disease. Using this analysis, the health insurance companies may deny insurance to some people and thus would *violate human right to affordable healthcare.*

AI Ethics

AI ethics is a set of values, principles, and techniques that employ widely accepted standards of right and wrong to guide moral conduct in the development and use of AI technologies.

PART B

AI Bias

AI bias is an anomaly (irregularity of abnormality) in the result produced through AI based programs and algorithms because of prejudiced (discriminatory) assumptions made during the algorithm development process or prejudices in the training data.[1]

Training Data is used to store people characteristics in the form of feature values. Bias in a data collection happens if some data representing a *feature*, *group*, *ethnicity* etc. is under-represented or over-represented. For instance, consider this — a minority group's preferred colour for car is *'Red'* because of their cultural influence. Now if another dataset stores that 'Red' is also a preferred choice of colour for aggressive drivers, then without much representation, this dataset may link the minority group with aggressive driving — an AI bias, here.

Training Data

Training data is a huge collection of labelled information that's used to build an AI model (*e.g.*, machine learning model).

Reasons for AI Bias in Data

Other than the over- and under-representation, there are many more reasons that cause or contribute to AI bias.

These are :

(*i*) Human bias in decisions
(*ii*) Flawed and unbalanced data collection
(*iii*) Under- or over-representation of specific features
(*iv*) Wrong assumptions
(*v*) No proper bias testing
(*vi*) No bias mitigation (*i.e.*, reducing the severity of bias)

Bias in Data Collection

Bias in data collection refers to flawed or unbalanced data with over- or under-representation of data related to specific features or groups or ethnicity etc. in the final data-collection.

PART B

Ensuring Data Fairness

It is very important for the collectors to do the following for the fairness of data and thereby, fair decision-making

- Identifying the correlation of features with data (should be diverse)

- Identifying the correlation among sets of data, while studying the impact of data and minimising the impact so as to have fair decisions
- Observing biases in human decisions and the data collected
- Ensuring "balanced" data
- Supervised decision-making
- Regular Bias testing by learning about biases and inducing fairness by thorough and repeated testing of data and modification in training data

PART B

Reducing and Mitigating AI Bias

Let us now learn how the AI people can reduce the AI biases in data collections and decisions :

(*i*) **Thorough Research.** The data collector must research their users or subjects in advance about which the data are being collected. They should be aware of general results and odd results of the data.

(*ii*) ***Diversity of Team.*** The team working for data collection or algorithm development must be diverse so that one person or team does not have major influence on data and algorithm of decision-making.

(*iii*) ***Data Diversity.*** Combine inputs from multiple sources to ensure data diversity.

(*iv*) ***Standardised Data Labelling.*** The team must have standardised way of labelling so that accurate, consistent and standardised data labels are used in data collection.

(*v*) ***Identify Bias-proneness.*** The team should identify the possible occurrences of biases among data sets and use multi-pass annotations, *i.e.*, multiple set of annotators label the data so as to minimise the possible bias.

(*vi*) ***Data Review.*** Enlist the help of someone with domain expertise to review collected and/or annotated data. Someone from outside of the team may see biases that the team has overlooked.

(vii) ***Regular Data Analysis.*** The team should keep track of errors and problem areas so as to respond to and resolve them quickly.

(viii) ***Regular Bias Testing.*** The team must test the collected data, training data and the overall performance of the algorithm against biases and use approaches and tools to mitigate the biases.

PART B

LET US REVISE

- *Various ethical issues of AI are : Bias and Fairness, Accountability, Transparency, Safety, Human – AI interaction, Trust, Privacy and Control, Malicious use of AI, Impact over jobs, Human and Robot rights and so forth.*
- *Deepfake is a technology that can generate fake digital photos, sound recordings, and video, which look just as original as possible.*
- *AI ethics is a set of values, principles, and techniques that employ widely accepted standards of right and wrong to guide moral conduct in the development and use of AI technologies.*
- *AI bias is an anomaly (irregularity of abnormality) in the result produced through AI based programs and algorithms because of prejudiced (discriminatory) assumptions made during the algorithm development process or prejudices in the training data.*
- *Training data is a huge collection of labelled information that's used to build an AI model (e.g., machine learning model).*

- *Some possible reasons of AI bias are : Human bias in decisions; flawed and unbalanced data collection; Under- or over-representation of specific features; Wrong assumptions; No proper bias testing; No bias mitigation (i.e., reducing the severity of bias).*
- *Bias in data collection refers to flawed or unbalanced data with over- or under-representation of data related to specific features or groups or ethnicity etc., in the final data-collection.*
- *A given AI model is fair if the outputs are independent of sensitive parameters (e.g., gender, race, sexuality, religious faith, disability, etc.) for a specific task that is already affected by social discrimination.*

UNIT 2

AI Project Cycle

IN THIS UNIT

SESSION 1

Introduction to AI Project Cycle

- Stages of AI Project Cycle :
  Problem Scoping/Data Acquisition/Data Exploration Modelling/Evaluation

1.2 STAGES OF AI PROJECT CYCLE

An artificial intelligence based project undergoes broadly five (below given) stages. In other words, the stages of an *AI Project Cycle* are :

(*i*) Problem Scoping (*ii*) Data Acquisition (*iii*) Data Exploration
(*iv*) Modelling (*v*) Evaluation

(*i*) Problem Scoping

In this stage, broadly the aim and scope of the project undertaken are decided. In other words, during this stage, following things are decided :

- strategic business objectives
- expected outcomes of the project
- stakeholders' expectations
- key resources and steps
- success metrics

Note

In the ***problem scoping stage***, broadly the *aim* and *scope* of the project undertaken are decided.

(ii) Data Acquisition

This stage is a crucial stage of AI project cycle. In this stage, the data are acquired keeping in mind the scope and parameters of the previous stage.

Standardised forms of data are extracted from dissimilar sources of data (bulk of sources) and produced in the most appropriate input form, most suitable form to produce the desired form of outcome, *e.g.*, for an AI based insurance project, the data will be heavily numbers and figures, while for a social-media based AI project it will be social media posts, or for providing security to a bank, the data may be mostly in visual form. However, to interpret the patterns and trends in the data, data in visual forms such as *charts*, *graphs*, *maps*, *histograms* and so forth is preferred.

Thus, the data acquired during this stage must be :

- as per the scope and parameters of the project
- accurate
- reliable

In the ***data acquisition stage***, the data are acquired keeping in mind the scope and parameters of the previous stage.

Note

To interpret the patterns and trends in the data, which is in large quantities, data in visual forms such as ***charts***, ***graphs***, ***maps*** (*e.g.*, ***heat maps***, ***density maps*** etc.), ***Tree diagrams***, ***Venn diagrams***, ***histograms*** and so forth is preferred.

(iii) Data Exploration

The purpose of the third stage is to explore data and its patterns so as to :

- choose the types of models for the project that can solve the problem.
- clean and normalize the data to standardise and correlate the data.

The *'cleaning and normalizing dirty data'* requires data scientists to make decisions on data they may not understand, like what to do with *missing data*, *incomplete data*, and *deviating data* (called *outliers*).

Thus, during this phase after deciding about the type of models that may serve :

- data from multiple sources is aggregated into a format suitable for the AI project's model, *e.g.*, for Computer Vision images and videos are the best suited format.
- the form (*e.g.*, database or visual forms of data or metadata etc.) and structure of data is such chosen so that it is likely to produce the desired outcome.

In the ***data exploration stage***, the data are explored to choose the possible models and clean & normalise the data.

(iv) Modelling

The next phase of an AI project cycle is to model the data that will be used for the prediction. So, the following activities take place in this phase :

- The selected models (from the previous stage) are tested and analysed and the most suitable AI-model is chosen that matches the requirements of the project.
- Once the most efficient model is chosen, AI algorithm(s) is developed around it.

- During this stage only, the ***training data*** and ***testing data*** are also decided.
  - A **training data (model)** is a dataset that is used to train an AI algorithm. The ***testing data*** is used with the **AI algorithm** to correlate the input data with the *processed output*. The result from this correlation is used to modify the model so that the processed output becomes more and more matching to the *sample output*.
  - The **testing dataset** is a set of observations used to validate the developed models after training is complete using some performance metric.

In the ***modelling stage***, the selected models are tested and analysed and the most suitable AI-model is chosen and AI-algorithms are developed around it.

(v) Evaluation

In this final phase, now the developed model is actually evaluated for accuracy and performance using new data. Then the results are evaluated to determine if the model can be deployed or requires some improvement prior to the final deployment.

Before a model can be deployed,

- it must meet some minimum accuracy standard, required for the deployment.
- there must be a clear agreement on what level of risk is acceptable for inaccuracy.

In the ***evaluation stage***, the developed model is actually evaluated for accuracy and performance using new data so as to determine if the developed model is deployable or not.

PART B

LET US REVISE

- *An AI project undergoes some phases from its initiation to closure, which together are known as AI Project Cycle.*
- *The five stages of AI project cycle are : Problem Scoping, Data Acquisition, Data Exploration, Modelling, and Evaluation.*
- *In the problem scoping stage, broadly the aim and scope of the project undertaken are decided.*
- *In the data acquisition stage, the data are acquired keeping in mind the scope and parameters of the previous stage.*
- *In the data exploration stage, the data are explored to choose the possible models and clean and normalise the data.*
- *In the modelling stage, the selected models are tested and analysed and the most suitable AI-model is chosen and AI-algorithms are developed around it.*
- *In the evaluation stage, the developed model is actually evaluated for accuracy and performance using new data so as to determine if the developed model is deployable or not.*

SESSION 2

Understanding Problem Scoping

- Choosing a Theme and Topic of AI Project
- Identify the Problems Around the Selected Topic
- Define Problem Statement and Set Actions

2.4 DEFINE PROBLEM STATEMENT AND SET ACTIONS

Setting goals for a project is not enough. In order to solve a problem, the *problem statement* must be clearly defined. It is very important to know how this problem is *affecting whom; what is the cause* and *why should it be solved* ? For this purpose, a useful tool — **4Ws Canvas** is used.

2.4.1 4Ws Canvas

The 4Ws Canvas tool (also known as **problem statement template**) explores four W-based questions – WHO, WHAT, WHERE and WHY – in the context of problem to be solved to clearly understand various aspects of the problem.

4Ws Canvas

The **4Ws Canvas** is a tool to explore a problem space in terms of WHO, WHAT, WHERE & WHY and formulate a problem statement with clarity.

The method consists of collectively answering questions gathered in 4 categories :

- **Who.** (**Find the stakeholders** - *Stakeholders are the people/persons that are directly affected by the problem*) Who is facing the problem and would benefit from its future solution?
- **What.** Are you sure it's a problem ? What is the nature of the problem ? Can you explain it simply? How do you know it's a problem ? What is the evidence to support the problem? Is it openly stated or unconscious?
- **Where.** In which context or situation do people experience the problem ? Have you observed the problem in context ? Can you describe that context ?
- **Why.** Why do you believe it is a problem worth solving ? Is it an acute problem for the people experiencing it ? How acute ?

4Ws Canvas for AI Project for *Face Mask Detection*

WHO has the problem? *Who else face the problem?*	**WHAT** is the problem? *What are the elements of the problem?* (The ***Pain-Points***)	**WHERE/WHEN** does the problem arise? *What is the context/location of it?*
Agencies ensuring public health and well-being	**1.** People do not wear face masks. **2.** People wear masks wrongly.	**1.** When people step out of their homes without wearing face masks or wearing face masks wrongly.
WHY the problem should be solved? *Is it actually worth solving? Why hasn't it been solved earlier ? (Is it hard or low-priority ?)*		
1. It is important to ensure that the people wear face masks properly at public spaces so that infection spreading is reduced and a clear and safe environment is made available to people.		

Problem Statement of AI Project of *Face Mask Detection*

Government and non-government agencies responsible for public health and well-being (WHO) ensure that a safe and healthy environment is made available to people. For this, certain rules and laws are in place – one such rule by law is to wear face masks on public places, especially during pandemic times.

But, some **people violate this law and do not wear face masks (WHAT) at public places (WHERE) or wear them wrongly (WHAT).**

It is important to ensure that face masks are worn properly **to ensure a safe, infection-free public environment** for people (WHY).

2.4.2 Set Actions Around the Goal

Now set the actions around the goal in order to achieve it

- to list the ***pain-points***, *i.e.*, specific, persistent problematic issues,
- define desired and expected outcomes of the project,
- to anticipate the key resources and steps and outline multiple possible alternatives to solve the problem,
- to compare and pick the best solution.

LET US REVISE

- *For your AI project, first decide about a theme and its subtopics.*
- *Then choose a subtopic and list a set of problems in the chosen subtopic.*
- *From the list of chosen problems. Choose one or two problems as the Goal of AI project.*
- ***4Ws Canvas** is a framework to define the problem statement articulately (i.e., effective and clear definition) by finding factors like 'Who', 'What', 'Where'/'When' and 'Why'.*
- *The WHO canvas lists about the stakeholders for which the solution is to be developed.*
- *The WHAT canvas lists what exactly the nature of the problem is.*
- *The WHERE/WHEN canvas lists the location and context of the problem.*
- *The WHY canvas lists if the reasons are strong enough to find the solution of the problem.*
- *A problem statement is a short, clear description in words of the problem listing its stakeholders, their issue(s), context, and reasons to solve the problem.*

SESSION 3

Simplifying Data Acquisition

- Significance of Data
- Type of Data used in AI Projects
- Data Acquisition

3.4 DATA ACQUISITION

Data acquisition begins when you acquire required data in quality form. Before actual data acquisition happens, some preprocessing is required where following questions are answered.

- What are the data features needed ?
- Where can you get the data ?
- How frequent do you have to collect the data ?
- What happens if you don't have enough data ?
- What kind of analysis needs to be done ?
- How will it be validated ?
- How does the analysis inform the action ?

3.4.1 Identifying Data Requirements

After answering all the questions mentioned above, you can finalise the data requirement by doing these :

(*i*) Group together the **relevant data features** in logically related structures.

(*ii*) Be clear about the relationship of data in and outside the logical data structure.

(*iii*) Use consistent and standardised terminology and format.

3.4.2 Finding Reliable Data Sources

There can be many different sources wherefrom data may be collected. Most commonly used data sources are being discussed below.

1. Interview

It is one of the most effective sources of data gathering. In this method, an analyst talks to the users and clients who know about the system, its functions and flaws.

Interview

An **interview** refers to a one-on-one conversation between an analyst and the users and clients to find out about the systems, its functions, shortcomings and flaws.

2. Survey

In Surveys, first the goal of the survey is ascertained and thereafter the questionnaires are drafted accordingly. Then using this, the responses of all users and stakeholders are documented.

Survey

A **survey** refers to a study of the opinions, responses, etc. of a group of stakeholders.

3. Observation

Under the observation method, the responsible person observes the team in a real working environment and gets ideas about the required data and its form, and subsequently documents the observation.

Observation

The **observation** method refers to human or mechanical watching, noticing or perceiving of what people actually do or what events take place in a specific working environment.

4. Application Programming Interface (API)

API is a specialized technique in which specific type of data is collected through the use of a programming interface, such as using social media programs' interface, data like people's most preferred game, most liked post, most used time etc. may be gathered.

API

An **API** refers to Application Programming Interface that works behind a popular software program or game to collect specific type of data pertaining to users' way of using that program.

5. Web Scraping

Web scraping, web harvesting, or web data extraction is data scraping used for extracting data from websites. A web scraper is a specialized tool designed to carry the web scraping.

Web Scraping

Web Scraping refers to a data collection technique using a tool called web scraper that extracts data from websites.

6. Sensors

Sensors or electronic sensors can measure various different parameters such as *weather, humidity, body temperature, blood pressure, heart beat, weight* and many more. For instance, you can see that modern medical diagnosis and wearables like *Fitbit*, *'Apple watch'* make good use of sensors.

Internet of Things (IoT) cannot function without sensors.

Sensors

Sensors are mini devices that can collect data about an environment or a body or a specific task.

7. Cameras

Cameras, because of their video recording and image capturing features have proven to be good data collection tools in various situations such as *traffic rules violations, automatic detection of flaws in design and outlook of products, places, buildings* etc.

Cameras

The method of **data collection using Cameras** is a way to collect data graphically or in video form about the look, design or action as per the requirements.

8. The Internet

Searching the Internet for data as per one's requirements is a commonly used technique. However, you should not take data directly from the Internet for the following *two* reasons :

(*i*) The data might not be authentic as its accuracy cannot be proved. Studies have shown that more than half of the data of the Internet comes from unreliable sources or is inaccurate.

(*ii*) Even if the data is reliable, it cannot be directly taken if it is copyright protected because of IPR (Intellectual Property Rights).

3.4.3 Acquiring Data

After identifying the data requirements, required data features and appropriate and reliable data sources, finally data is collected in required form. That is, in data acquisition, data is **understood, gathered, filtered, cleaned and finally stored** in a data storage system.

Data Acquisition

Data Acquisition refers to understanding, gathering, filtering, cleaning data as per the requirement of the AI system so as to train it using the collected data.

Let us Revise

- *Quality data has these characteristics : accuracy, relevance, reliability, timeliness, validity and completeness.*
- *The data used for AI systems can be structured as well as unstructured data.*
- *The **structured data** conforms to some predefined data model and has well-defined relationships among its elements.*
- *The **unstructured data** does not conform to any pre-existing data model and has undefined relationships among its elements.*
- ***Data Acquisition** refers to processes, methods or systems that are used to collect information related to a certain theme or objective, to document or analyse some phenomenon.*
- *Data must be gathered from reliable sources.*
- *Most commonly used data sources are Interview, Survey, Observation, API, Web Scraping, Sensors, Cameras, the Internet, Problem reports etc.*
- *An interview refers to a one-on-one conversation between an analyst and the users and clients to find out about the systems, its functions, shortcomings and flaws.*
- *A survey refers to a study of the opinions, responses etc. of a group of stakeholders.*
- *The observation method refers to human or mechanical watching, noticing or perceiving of what people actually do or what events take place in a specific working environment.*
- *An API refers to Application Programming Interface that works behind a popular software program or game to collect specific type of data pertaining to users' way of using that program.*
- *Web Scraping refers to a data collection technique using a tool called web scraper that extracts data from websites.*
- *Sensors are mini devices that can collect data about an environment or a body or a specific task.*
- *The method of data collection using Cameras is a way to collect data graphically or in video form about the look, design or action as per the requirements.*

SESSION 4

Data Exploration with Data Visualisation

- What is Data Exploration ?
- Data Visualisation
- Ways to Visualise Data

4.2 WHAT IS DATA EXPLORATION ?

Data exploration is the phase after data acquisition wherein the collected data is cleaned by removing redundant data and handling missing values[1] and then analysed using data visualisation and statistical techniques to understand the nature of data before it can be converted into AI models.

Data Exploration

Data Exploration is the phase of exploring data with an intention of understanding the nature of data.

1. There are special techniques used which help clean data, remove redundant data, handle missing values by filling some legal agreed upon values. This is done so that data becomes analysable. Covering these techniques here is beyond the scope of the book, however, data visualisation techniques are being discussed here.

The data exploration techniques are applied mainly :

- to visually **explore and identify relationships between different data variables,**
- to understand **the structure of the dataset,**
- to identify the presence of data points that differ significantly from other observations (called **outliers**).
- to obtain the distribution of data values in order **to reveal trends and patterns and points of interest.**

4.3 DATA VISUALISATION

Data visualisation refers to the process of representing data visually or graphically, by using visual elements like charts, graphs, diagrams and maps etc.

The importance of data visualisation is summarised as follows :

(*i*) Data visualisation is a powerful way to represent a bulk of data in a collective visual form.

(*ii*) It is a way to explore data with presentable results.

(*iii*) Data visualisation makes it easy to interpret and comprehend data.

(*iv*) It becomes easier to see the trends, relationships and trends of data through data visualisation.

(*v*) Data visualisation is useful for combining categories of data and thereby reducing data for processing.

(*vi*) Data visualisation helps in defining strategy for using data for AI model to be developed at later stage.

1. SCATTER CHART (Used with numeric type of data)

An XY (scatter) chart either shows the relationships among the numeric values in several data series or plots two groups of numbers as one series of *XY* coordinates.

How to draw?

The scatter chart is drawn by plotting the independent variable on the horizontal axis *X*, the dependent variable on the vertical axis *Y* and then by marking data points as per their *XY* values.

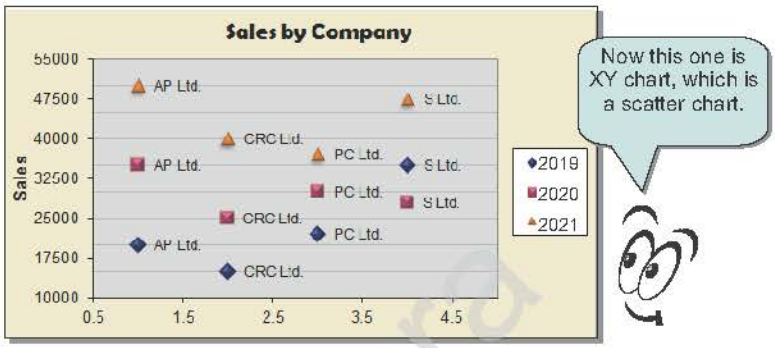


Figure 4.1 A Scatter Chart

2. BUBBLE CHART (Used with numeric type of data)

A bubble chart is primarily used to depict and show relationships between numeric variables with marker size as additional dimension. Bigger marker means bigger value.

How to draw?

The bubble chart is drawn by plotting the independent variable on the horizontal axis (*X*), the dependent variable on the vertical axis (*Y*) and then by marking bubbles at their *XY* values. The *Y* values will determine the bubble size.

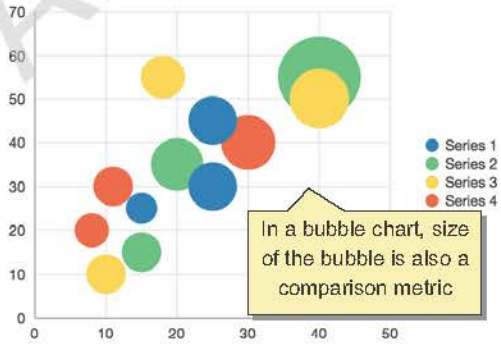


Figure 4.2 A Bubble Chart

3. LINE GRAPH (Used with numeric type of data)

A line chart shows trends in data at equal intervals. Line charts are useful for depicting the change in a value over a period of time.

How to draw?

The line chart is drawn by plotting the independent variable on the horizontal axis (*X*), the dependent variable on the vertical axis (*Y*) and then by marking data points as per their *XY* values. Then a line is drawn by joining the marked data points.

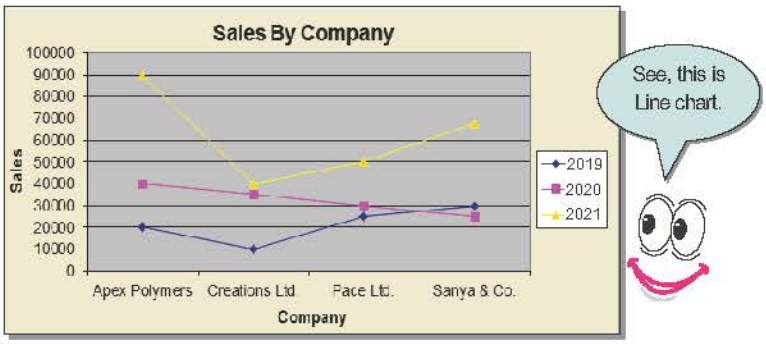


Figure 4.3 A Line Chart

4. PIE GRAPH (Used with numeric type of data)

A pie chart shows the proportional size of items that make up a single data series to the sum of the items.

How to draw?

The pie chart represents single data series, whole of which represents full circle (360°). Each data value is calculated as a percentage of whole and drawn as a pie of the circle.

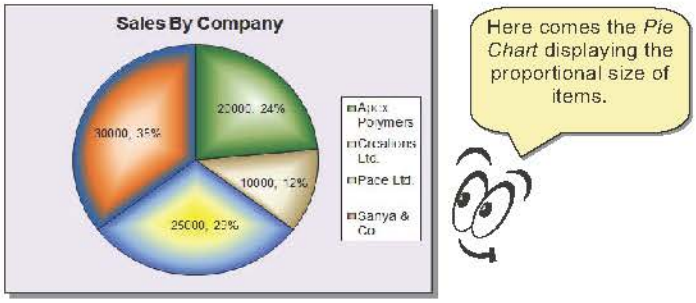


Figure 4.4 A Pie Chart

5. BAR GRAPH (Used with numeric type of data)

A bar chart illustrates comparisons among individual items, mainly of number types.

How to draw?

The bar chart is drawn by plotting the independent variable on the horizontal axis (X), the dependent variable(s) on the vertical axis (Y) and then by marking bars for their Y values.

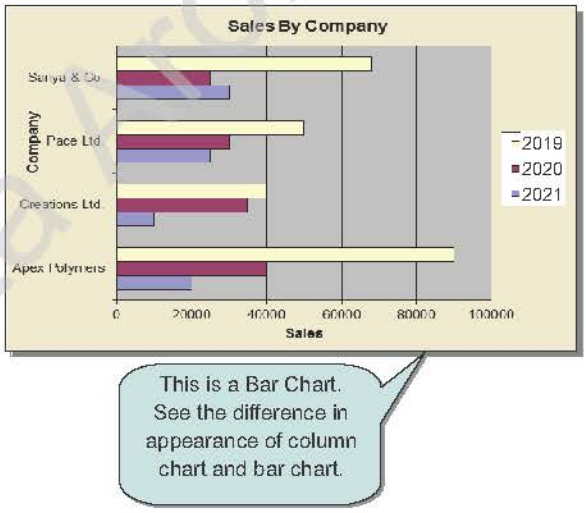


Figure 4.5 A Bar Chart

6. HISTOGRAM (Used with numeric type of data)

A histogram is used to summarize discrete or continuous data by showing the number of data points that fall within a specified range of values (called "**bins**"). Unlike a bar chart, there are no gaps in between in a histogram.

How to draw?

Like bar chart, rectangles of varying height are used to represent the frequency of different values of the continuous variable (Y values). There are no spaces between the rectangles.

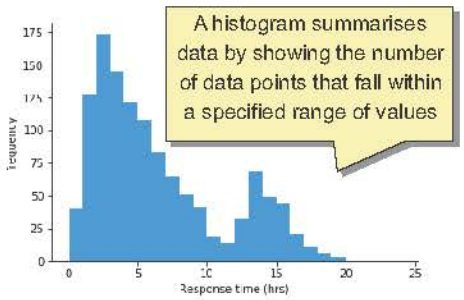


Figure 4.6 A Histogram Chart

PART B

7. CHOROPLETH (Used with processed numeric data linked with textual units)

Choropleth maps are used with statistical data (numeric, processed data) attached to enumeration units (textual data *e.g.*, countries, provinces, states etc.) to depict data for geographic regions. For example,

- world map of income tax rates, country wise.
- world map of Covid 19 spread, country wise.
- map showing the percentage increase in real estate value, state wise.

Covid 19 spread all over the world

Figure 4.7 Choropleth Chart

How to draw ?

In the region map, firstly the statistical values are written in the sub-regions. Then the sub-regions are filled with the corresponding colour for that value.

8. HEAT MAP (Used with numeric data depicted through colour codes)

A **heat map** is a graphic representation of data in which values are represented by colours. Some examples of heat maps are :

- A geographical heatmap representing areas of high and low density of a certain parameter (population density, network density, etc.) by displaying data points on a map through different colours.
- A stock index heatmap depicting prevailing trends in the market through colours, *e.g.*, cold-to-hot colour scheme to indicate which stock options are bullish and which are bearish.

India Climate heat map

Figure 4.8 Heat Map

How to draw?

The region is divided into smaller squares. Then each square is filled with the colour code as per the data it is storing.

9. TIMELINE (Used to represent all types of data against time)

A Timeline Chart shows a series of events in chronological order. It can be used to depict the order of historical events, critical milestones of a project schedule, and so on.

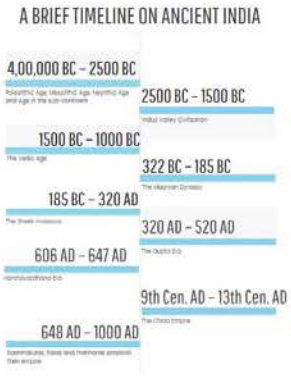


Figure 4.9 Timeline Chart

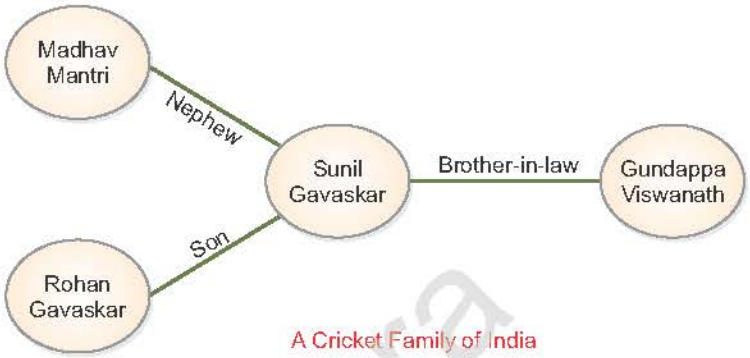
How to draw?

Draw a horizontal/vertical line with ends marking the start and end dates. Mark the points on the line for each of the events. Mark at each point the date and event.

10. NODE LINK DIAGRAM (Used with all types of data)

A node-link diagram shows how things are interconnected through the use of nodes/vertices and link lines to represent their connections and the type of relationships between a group of entities.

These are used in many applications, for example, for analysis of social networks or mapping product sales across geographic areas and many similar ones.



A Cricket Family of India

Figure 4.10 Node Link Diagram

How to draw?

Points are represented as nodes (vertices) that are linked by lines (edges).

11. WORD CLOUD (used with textual data)

The word cloud data visualisation technique represents the frequency of a word within a body of text with its relative size in the cloud. This technique is used on unstructured data as a way to display high- or low-frequency words.

How to draw?

The frequency of each word determines its weight, which determines its priority. The words with the highest priority get drawn first, and will be drawn with larger font-size.

Figure 4.11 Word Cloud

Popular Data Visualisation Tools

There are many software available that let you visualise data in various forms. Most commonly used such software are **Ms Excel**, **Google Charts** etc. However, these days there are many other software available that are so strong with data visualisation techniques that they have gained equal popularity. Some such popular, open source data visualisation tools are shown in the adjacent figure 4.12.

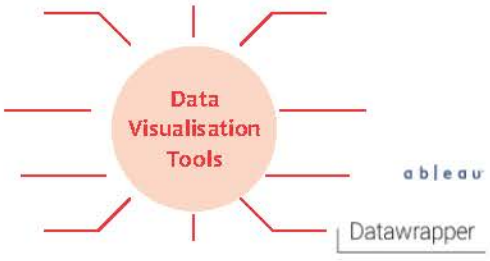


Figure 4.12 Some Popular Open-source Data Visualisation Tools

Let Us Revise

- *Data Exploration is the phase of exploring data with an intention of understanding the nature of data.*
- *Data visualisation refers to the process of representing data visually or graphically, by using visual elements like charts, graphs, diagrams and maps etc.*
- *Data visualisation is important and useful for understanding the information stored in data.*
- *An XY (scatter) chart either shows the relationships among the numeric values in several data series or plots two groups of numbers as one series of XY coordinates.*
- *A bubble chart is primarily used to depict and show relationships between numeric variables with marker size as additional dimension. Bigger marker means bigger value.*
- *A line chart shows trends in data at equal intervals. Line charts are useful for depicting the change in a value over a period of time.*
- *A pie chart shows the proportional size of items that make up a single data series to the sum of the items.*
- *A bar chart illustrates comparisons among individual items, mainly of number types.*
- *A histogram is used to summarize discrete or continuous data by showing the number of data points that fall within a specified range of values (called "bins").*
- *Choropleth maps are used with statistical data (numeric, processed data) attached to enumeration units (textual data e.g., counties, provinces, states etc.) to depict data for geographic regions.*
- *A heat map is a graphic representation of data in which values are represented by colours.*
- *A Timeline Chart shows a series of events in chronological order.*
- *A node-link diagram shows how things are interconnected through the use of nodes/vertices and link lines to represent their connections and the type of relationships between a group of entities.*
- *The word cloud data visualisation technique represents the frequency of a word within a body of text with its relative size in the cloud.*

SESSION 5

Modelling

- What is Modelling ?
- Categories of AI Models
- Supervised/Unsupervised Learning
- Semi-Supervised Learning/Reinforcement Learning

5.2 WHAT IS MODELLING ?

Modelling is the phase during which the AI model for the desired outcome is trained using the collected data repeatedly until it starts producing the desired results. Mostly, for this, some publically available pre-trained AI models are picked, tuned as per requirements and then trained using own data.

5.3.1 Rule Based AI Model

Also known as ***model-driven AI model***, it is a system which derives decisions about the right answer through explicit representation and rules. For example, in such a system, a cat would be explicitly represented as a four-legged animal, with two eyes, a nose and a mouth that is furry (except when not) and that is relatively small (except when not), etc.

Note

The rule based AI is used when we have known or labelled dataset.

Model Driven AI

Rule Based AI (Model driven AI) refers to the branch of AI where models are developed using the algorithms having pre-defined labels, rules, patterns and relationships.

In this, an AI system is developed using the predefined *labels, rules, patterns or relationships* as given by the developer in the algorithm.

Thus, the machine would follow the algorithm's rules or instructions and perform the given tasks or take the decisions accordingly.

5.3.2 Learning Based AI Model

Also known as ***data-driven AI model***, it is a system in which a lot of data is shown and questions/answers are asked in order to **train** the system about the right answer(s). For example, to train a system about recognising cats, it would be shown lost and lots of images of cats, and other animals and letting it know when it "guessed" it correctly or not. After many (millions) of training cycles, it will "learn" to get it increasingly right.

Note

The *learning based approach* is used when data is unknown or random or unlabelled.

Data Driven AI

The **Learning based AI (Data Driven AI)** refers to that branch of AI where models are trained to learn by inputting them tons of data. Here, there are no patterns, rules and relationships predefined by the developer, rather machine learns with each new input and comes up with own algorithm.

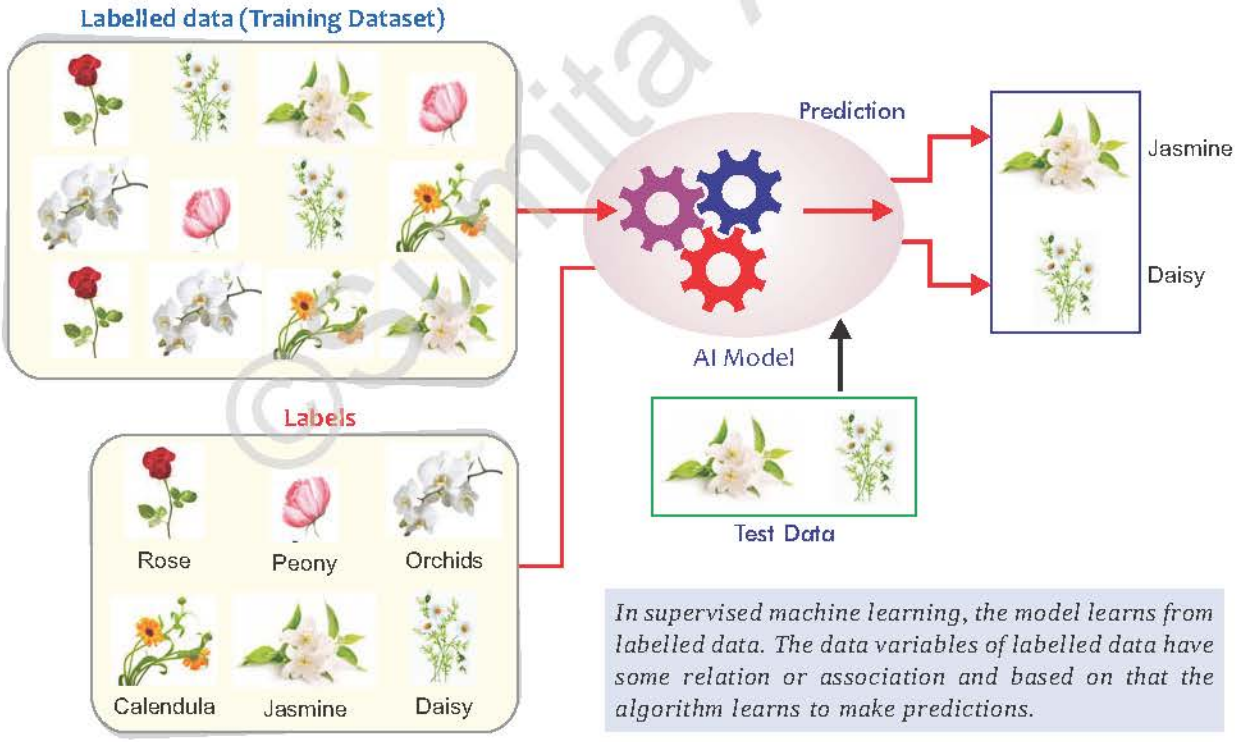
Using this style when an AI system is developed then the relationship or patterns in data are not defined by the developer. Rather, random data is repeatedly input to the machine. The machine analyses each input and tries to figure out patterns and trends out of the input collectively. This style is appropriate when the data to be processed is **unlabelled** and **too random** to fit into a frame of common rules and patterns.

5.4 SUPERVISED LEARNING

Supervised learning is a machine learning approach in which a machine, with the help of an algorithm (called *the model*), learns from a labelled dataset and desired outputs. Using this dataset, it learns to identify the type or class of the data given to it. Later some data is shown to it to test if it can clearly identify the data or not. For example, a labelled dataset of ***flower images*** would contain photos of roses tagged as *roses*, photos of daisies tagged as *daisies* and so on for other flowers. When shown a new image, the model compares it to the training examples to predict the correct label (*see* figure below). The model gets feedback about its result as per the desired outputs and this way, it learns to classify correctly and that is why it is supervised learning.

Supervised Learning

Supervised Learning is a learning approach of machine where the machines with the help of an algorithm (the model) learns on a labelled dataset and is later tested with some unlabelled data whose answers are pre-known to evaluate its accuracy on training data.

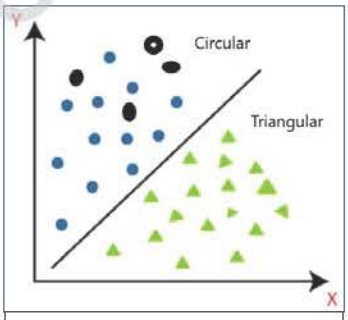


In supervised machine learning, the model learns from labelled data. The data variables of labelled data have some relation or association and based on that the algorithm learns to make predictions.

Figure 5.2 Supervised Learning

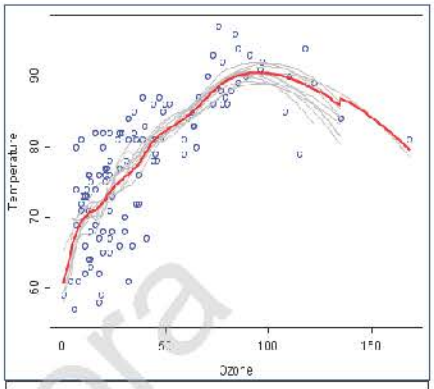
5.4.1 Types of Supervised Learning

(*i*) ***Classification.*** *Classification* means, as the name suggests, identifying the class of input value. For example, in the supervised learning model explained above, the training dataset was based on labelled images of flowers. Now upon receiving a new unlabelled image of a flower, if it can identify the class type (which flower type) of the image, it has accurately performed classification.



Classification AI model

(*ii*) ***Regression.*** Regression refers to a mathematical approach used to find the relationship between two or more variables. Regression works with continuous data. For example, 'how would the prices of a specific fruit be affected if its production is increased and there is overall dip in production cost' can easily be determined through a regression model.



Regression AI model

PART B

5.5 UNSUPERVISED LEARNING

In unsupervised learning, as the name suggests, there is no supervision, no feedback, no pre-known/desired outputs and not even any labelled data. Unsupervised Machine Learning discovers patterns within an existing set of unlabelled data, *i.e.*, the data without having any pre-existing labels or categories. Unsupervised learning based AI model (the algorithm) finds the patterns, trends and features, and clubs the data having same patterns in one group (or cluster). This way,

Unsupervised Learning

Unsupervised Learning is a learning approach of machine where the machines with the help of an algorithm (the model) learns on an unlabelled dataset where it categorises data on the basis of common characteristics, features and patterns.

unsupervised learning creates clusters with a similar set of values sufficiently different than other clusters. For every new input (also unlabelled), it tries to put it in a cluster as per its pattern or characteristics and then enabling new data to be categorised into an existing cluster. (*see* figure below)

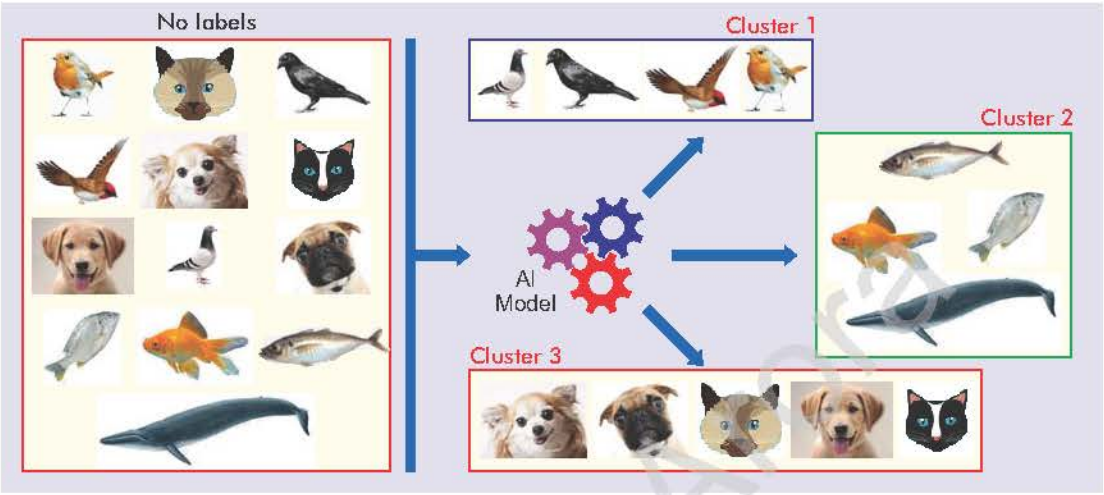


Figure 5.3 Unsupervised Learning

5.5.1 Types of Unsupervised Learning

5.5.1A Clustering

Clustering is an unsupervised learning approach of AI models, which groups unlabelled data based on their similarities or differences. Thus, we can say that it is because of

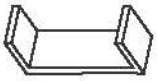
clustering algorithms, an AI model without even being an expert *ornithologist* (bird specialist), can take a collection of bird photos and separate them roughly by species, relying on cues like *feather colour*, *size* or *beak shape*.

Clustering

Clustering is an unsupervised-learning approach of AI models, which groups unlabelled data based on their similarities or differences.

5.5.1B Dimensionality Reduction

Dimensionality reduction broadly means representing an object in smaller dimensions.

For example, a 3-dimensional object  requires more variables to represent all its sides and dimensions. For visualisation purposes, we may need to view it in 2-dimensional views, *i.e.*, as  (*front view*) or as  (*side view*).

In other words, we may need to represent it in lesser number of variables for visualisation purposes. Thus, dimensional reduction is required in cases where the goal is to summarise the data in a reduced number of dimensions, *i.e.*, by using a reduced number of variables.

PART B

5.5.1C Association

Association is another unsupervised learning technique that finds important relations between variables or features in a data set. For example, if you pick some home decor items such as lamps or shelves in an online shopping cart, it will start suggesting the related items such as furniture, rugs and even interior designing firms. This is an example of association, where certain features of a data sample correlate with other features. By looking at a couple key attributes of a data point, an unsupervised learning model can predict the other attributes with which they're commonly associated.

Association

Association is an unsupervised-learning technique that finds important relations between variables or features in a data set.

5.7 REINFORCEMENT LEARNING

In reinforcement learning approach, the **AI model** (the *algorithm*, also called the **agent**) iteratively attempts to accomplish a particular goal, or improve performance on a specific task, in the best possible way known to it. If the action/step of the agent is helpful toward achieving the goal, it is given a reward. The overall aim of the agent is to predict the best next step to take to earn the biggest final reward.

To make its choices in reinforcement learning, the agent relies on :

- the learnings from past **feedback**
- exploration of new tactics that may present a larger payoff
- a **long term strategy** keeping in mind the final goal and best possible reward

The main elements of an RL system are (Fig 5.6) :

(*i*) The **agent** or the learner.

(*ii*) The **environment** the agent interacts with

(*iii*) The **policy** that the agent follows to take **actions.**

(*iv*) The **reward** signal that the agent observes upon taking actions.

'Learning to make correct decisions' is the core of reinforcement learning.

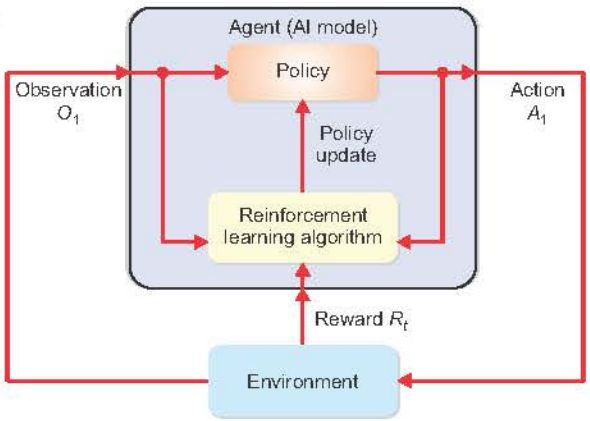


Figure 5.6 Reinforcement Learning

Let Us Revise

- *Modelling (developing and training an AI model) refers to mathematically analysing the data and its inside relationships and with the parameters passed, and finding ways through algorithm & repeated training to reach to desired and expected intelligent outcomes.*
- *The ability to mathematically describe the relationships between data and parameters forms the core of every AI model.*
- ***Rule Based AI (Model driven AI)** refers to the branch of AI where models are developed using the algorithms having pre-defined labels, rules, patterns and relationships.*
- *The **Learning based AI (Data Driven AI)** refers to that branch of AI where models are trained to learn by inputting them tons of data. Here there are no patterns, rules and relationships predefined by the developer, rather machine learns with each new input and comes up with own algorithm.*
- *Rule based models are often preferred for limited scale projects that require limited efforts, cost, and updates.*
- *The rule based systems are static and not largely scalable.*
- *Learning based models constantly "adapt" and "evolve" their performance in accordance with the continuous streams of training data.*
- *Learning based systems are dynamic and scalable.*
- *Learning based systems are categorised into three types : supervised learning, unsupervised learning and reinforcement learning.*
- *Data without any explanatory tags or names, is unlabelled data.*
- *Data with explanatory tags and names, is labelled data.*
- *Supervised Learning is a learning approach of machine where the machines with the help of an algorithm (the model) learns on a labelled dataset with corresponding outputs and is later tested with some unlabelled data whose answers are pre-known to evaluate its accuracy on training data.*
- *Discrete data is countable whereas Continuous data is the opposite of discrete data as it is measurable.*
- *Supervised learning based problems are mainly of two types : classification and regression.*
- *A Classification AI model refers to a type of Supervised Learning technique, which can classify the category of new unlabelled inputs on the basis of training data.*
- *The classification models use non-continuous, i.e., discrete data.*

- *Regression AI model refers to a type of Supervised Learning technique, which is based on a mathematical approach used to find the relationship between two or more variables and predict and outcome.*
- *The regression models use continuous data.*
- *Unsupervised Learning is a learning approach of machines where the machine with the help of an algorithm (the model) learns on an unlabelled dataset without any corresponding outputs and where it categorises data on the basis of common characteristics, features and patterns.*
- *Most commonly used unsupervised learning applications are based on clustering, association and dimensionality reduction.*
- *Clustering is an unsupervised learning approach of AI models, which groups unlabelled data based on their similarities or differences.*
- *Association is another unsupervised learning technique that finds important relations between variables or features in a data set.*
- *Dimensionality reduction is an unsupervised learning approach that uses techniques for reducing the number of input variables in training data while retaining its sense and meaning.*
- *Dimensionality reduction reduces the complexity of a problem by reducing the number of variables involved.*
- *Semi-supervised learning refers to a learning approach where the training dataset has both labelled and unlabelled data and the AI model uses a combination of supervised and unsupervised learning techniques. Using labelled data supervised learning techniques are used and new/unexplored features are extracted from the unlabelled data using unsupervised learning techniques.*
- *Reinforcement learning refers to an AI learning approach that trains algorithms using a system of reward and penalty. The learning system (called agent) learns in an interactive environment where the agent iteratively selects and performs actions and receives rewards by performing correctly and penalties for performing incorrectly.*

SESSION 6

Neural Networks

- Biological Neural Networks
- Artificial Neural Networks (ANN)
- Features of Neural Networks
- Advantages and Disadvantages of Neural Networks
- Applications of Neural Networks

6.2 BIOLOGICAL NEURAL NETWORKS

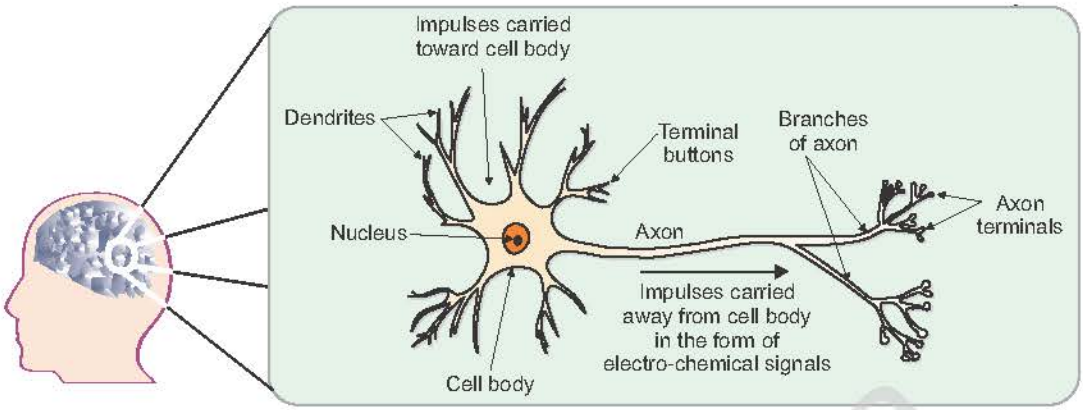
Our brain has a large number ($\approx 10^{11}$) of highly connected elements ($\approx 10^4$ connections per element) called **neurons**. A **neuron** has *three* major components (Fig. 6.1) :

- **Soma.** It is a cell body that contains nucleus, and sums and thresholds all incoming signals.
- **Axon**. The axon is a long fibre that carries signal from the cell body out to other neurons.
- numerous **Dendrites**. Dendrites are tree-like receptive networks of nerve fibres that carry electrical signals into the cell body.

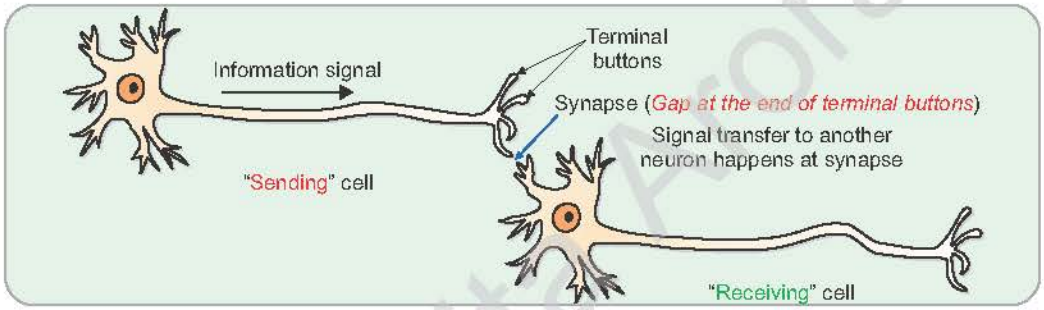
Along with the three components listed above, an important role is played by *synapse*.

- **Synapse.** The point of contact between an axon of one cell and a dendrite of another cell is called a synapse.

Figures 6.1(*a*), (*b*) show the structure and working of human brain neurons.



(*a*) Structure of human brain neurons



(*b*) Information transfer happens at synapse

Figure 6.1

PART B

6.3 ARTIFICIAL NEURAL NETWORKS (ANN)

An **Artificial Neural Network (ANN)** is a software- or circuit-based simulation of a biological neural network. An ANN can be thought of as an interconnected assembly of simple processing elements, called **units** or **nodes**, whose functionality is loosely based on a biological neuron. The components of an artificial neural network (ANN) are :

Neuron	equivalent of neuron (also called **node** or **neurode**)
Weighted inputs	equivalent to dendrites
Activation function	equivalent to soma (It defines how the weighted sum of the inputs is transformed into an output from a node or nodes in a layer of the network.)
Synapse	connection from a neuron to another that carry the information
Axon	output

6.3.1 Structure of ANN

A Neural Network is divided into multiple layers. Each layer of an ANN consists of several ***artificial neurons*** called **nodes**. Each node has to perform a specific task and pass the information to the next layer.

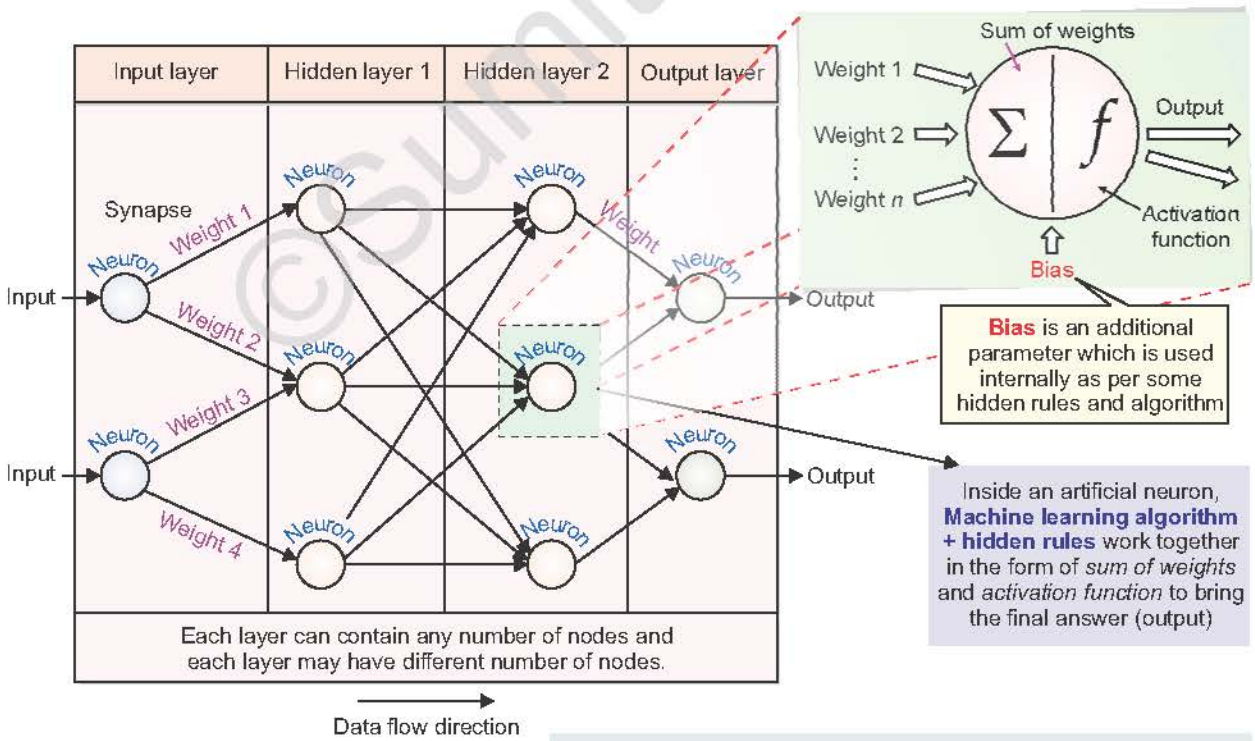


Figure 6.2 Artificial Neural Network

In an ANN, a **node** (or **neurode**) is the artificial equivalent to a neuron. It consists of a set of weighted inputs (**dendrites**), an activation function (**soma**) and one output (**axon**).

There are *three* types of layers in an ANN :

(i) Input Layer

The first layer of a Neural Network is called the *input layer*, whose role is to acquire data and feed it to the Neural Network. The input layer carries out no processing, it just takes the input data and passes it on to the next connected layer.

(ii) Hidden Layer

Input layer is connected to a *hidden layer*, which is further connected to other hidden layers or to the final *output layer*. The role of hidden layers is to process the inputs and carry out a task. The processing at the hidden layers is carried out as :

Sum of weighted inputs
+ activation function (*i.e.*, machine learning algorithm)
+ hidden rules (such as getting additional parameters such as a bias)

There can be multiple hidden layers in an ANN, depending upon the complexity of the task(s) being performed. Hidden layers are not visible to the user. The processed output of a hidden layer is then fed to the subsequent hidden layer of the network.

Each layer of an ANN can contain any number of nodes and each layer may have different number of nodes.

(iii) Output Layer

After processed data travels through multiple hidden layers, it (the final processed data) is finally fed to the final layer known as the **output layer**. The output layer simply provides the final output to the user. At the output layer also, no processing takes place; it only provides user-interface for the output.

This point onwards in this session, the term **"neural networks"** will refer to *Artificial Neural Networks*.

A Neural Network is divided into multiple layers (*hidden layers* in between *input and output layers*). Each layer of an ANN consists of several **artificial neurons** called **nodes**. Each node has to perform a specific task and pass the processed information to the next layer.

In a neural network, information transfers from the input layer nodes to the connected layer(s), which, after required *processing and decision-making* (using **activation function**) reaches to the final layer as output. Neural networks are initially trained with some input data and then they keep learning with every new input and feedback.

PART B

6.4 FEATURES OF NEURAL NETWORKS

From the above discussion, we can list the features of neural networks as follows :

(*i*) Neural networks have been developed to mimic the structure and working of human brain.

(*ii*) Neural networks evolve and automatically learn with each input and each new attempt.

(*iii*) Neural networks can work with big data sets.

(*iv*) The neural networks employ machine learning techniques to function and evolve.

(*v*) The Neural Networks (NNs) are said to exhibit the following *two* abilities :

(*a*) Ability to learn

- NN's can figure out how to perform their function on their own.
- NNs can determine their function based only upon sample inputs

(*b*) Ability to generalize

- NNs can produce reasonable outputs for those inputs for which it has not been taught how to deal with, based on its past learning.

(*vi*) ANNs take ample time to train and need lots of data to train.

6.6 APPLICATIONS OF NEURAL NETWORKS

Neural Networks have found their applications in multiple fields and areas. Some common applications of neural networks are being listed below.

- **Character Recognition.** Character Recognition, is a process of recognizing text inside images and converting it into an electronic form. These images could be of *handwritten text, printed text like documents, receipts, name cards*, etc., or even a *natural scene photograph*. Neural networks can be used to recognise characters in various forms, such as in images, printed text or even handwritten characters.
- **Speech Recognition.** NNs have found their successes in speech or communication recognition. You can see these applications around you in terms of *Siri, Alexa, Google assistant* and so on.
- **Computer Vision.** With the help of NNs, computers can accurately understand and process visual data efficiently like videos and images.
- **Image Compression.** Neural networks can receive and process vast amounts of information at once, making them useful in image compression. Image compression means storing and processing images with lesser amount of data.

- **Stock Market Prediction.** The day-to-day business of the stock market is extremely complicated. Many factors weigh in whether a given stock will go up or down on any given day. Since neural networks can examine a lot of information quickly and sort it all out, they can be used to predict stock prices.

- **Pattern Recognition.** NNs are also very useful in recognition of a pattern or pattern recognition (PR). Patterns are basically repeated trends in various forms of data. For example, a pattern could be a *fingerprint image, a handwritten cursive word, a human face, or a speech signal* and so forth.

- **Detection.** Detection in medical diagnosis, security, image objects, financial irregularity, a fault in a system, are being enhanced through ANNs application. Thus, ANN plays an essential role in the detection and diagnosis, such as diagnosis of breast cancer, crime detection using DNA and so on.

- **Travelling Salesman's Problem.** The travelling Salesman problem is defined as - *Given a set of cities and distance between every pair of cities, the problem is to find the shortest possible route that visits every city exactly once and returns to the starting point.* Interestingly enough, neural networks can solve the travelling salesman problem, but only to a certain degree of approximation.

- **Miscellaneous Applications.** These are some very interesting applications of neural networks in finance, loan applications, medicine etc. where using a neural network they will decide whether or not to grant a loan whether this security lapse is repeatable; whether this ailment will relapse and so forth.

Let Us Revise

- *Artificial neural network (ANN) refers to a collection of connected computational units or nodes (called neurons) along with biologically inspired computer programs designed to simulate the way in which the natural neural networks of human brain process information.*
- *The message transfer from a neuron is called* ***action potential.***
- *In an ANN, a node (or neuron) is the artificial equivalent to a neuron. It consists of a set of weighted inputs (dendrites), an activation function (soma) and one output (axon). Information signal travels through multiple layers of connecting neurons before it is transformed in the form on an output.*
- *A Neural Network is divided into multiple layers (hidden layers in between input and output layers).*
- *Each layer of an ANN consists of several artificial* ***neurons*** *called* ***nodes.***
- *Each* ***neuron*** *node has to perform a specific task and pass the information to the next layer.*
- *An activation function in a neural network defines how the weighted sum of the input is transformed into an output from a node or nodes in a layer of the network.*
- *Providing feedback about the difference from the correct output is known as back propagation.*
- *After initial training, Neural networks keep updating their internal mechanism of reaching to an output with every new feedback received.*
- *The NNs have ability to learn and ability to generalize.*
- *The NNs are massively parallel and fault-tolerant.*
- *The supervised learning is the most commonly used with neural networks, although other ways of training neural networks are also used.*

SESSION 7

Evaluation

- What is Model Evaluation ?
- Model Evaluation Metrics

7.2 WHAT IS MODEL EVALUATION ?

Evaluation, in general, refers to evaluating a system or device in a systematic way to check its merit, correctness and performance as per a set of standards. In AI project cycle, we can say that, **Evaluation** refers to systematically checking and analysing the merit, correctness and reliability of an AI model based on the outputs produced by it.

Evaluation

Evaluation refers to systematically checking and analysing the merit, correctness and reliability of an AI model based on the outputs produced by it.

7.2.2 Causes Behind Performance of AI Model

1. Overfitting

Overfitting means the AI model performs so well against only known data, *i.e.*, the training data or the data very much similar to it (*i.e.*, fits the data). However, the AI model fails to fit the unknown data, *i.e.*, cannot predict the reliable results for unknown data. An overfitted model will appear to have a higher accuracy when you apply it to the training data. The model developers may take it as the success of the AI model, thinking it as highly accurate model, whereas in reality it will underperform in production when given new data.

Overfitting

Overfitting refers to a situation when an AI model performs so well as the test data it got, fitted exactly against its training data and thus AI model always produced correct result.

2. Underfitting

Underfitting is the opposite of overfitting. It happens when AI model is not complex enough to accurately capture the structure and relationships of training data so as to use the dataset's features for producing a specific result. An underfitted model results in problematic or erroneous outcomes on new data, or data not same as its training data, and often performs poorly even on training data.

Underfitting

Underfitting refers to a situation when an AI model is not complex enough to capture the structure and relationships of its training data and predict effective outcomes.

PART B

7.3 MODEL EVALUATION METRICS

7.3.1 Confusion Matrix

*A **Confusion Matrix** is a technique using a chart or table for summarizing the performance of a classification based AI model by listing the predicted values of an AI model and the actual/correct outcome values.*

A confusion table includes both **predictive** and **actual values** in context of AI model, which are :

- the Actual Value represents the actual result (observed or measured).

Actual Values	
True	False

- the Predicted Value is the value of the outcome/result of the AI model, produced on the basis of its algorithm and learning.

Predicted Values	
Positive	Negative

Before we proceed to how to create and use confusion matrices, it is important to discuss some terms associated with it. These are :

(*i*) True Positive (TP). True positive refers to an instance for which both *predicted* value of the AI model and *actual* value are positive. For example, while testing a patient for Covid, if the test also produced the result (**predicted value**) as *positive* and the actual result (**actual value**) is also *positive*, it is ***True positive***.

(*ii*) True Negative (TN). True negative refers to an instance for which both *predicted* value of the AI model and *actual* value are negative. For example, while testing a patient for Covid, if the test also produced the result (**predicted value**) as *negative* and the actual result (**actual value**) is also *negative*, it is ***True negative***.

(*iii*) False Positive (FP) **(also called Type I Error).** False positive refers to an instance for which *predicted* value of an AI model is positive but *actual* value is negative. For example, while testing a patient for Covid, if the test produced the result (**predicted value**) as *positive* and the actual result (**actual value**) is *negative*, it is *False positive*.

(*iv*) False Negative (FN) **(also called Type II Error).** False negative refers to an instance for which *predicted* value of an AI model is negative but *actual* value is positive. For example, while testing a patient for Covid, if the test produced the result (**predicted value**) as *negative* and the actual result (**actual value**) is *positive*, it is ***False negative***.

The first two terms signify when both the actual and predicted values match. In the 3rd and 4th terms listed above, the latter part (*Positive/Negative*) represents the predicted value and the *False* means the actual value is opposite of the predicted value.

For example, if looking at picture, the AI model has to identify if it is the picture of vegetable Lady finger, then *True Positive/Negative* and *False Positive/Negative* will be identified as :

True Positive  **Result.** Lady finger Vegetable	False Negative  **Result.** NOT Lady finger Vegetable
False Positive  **Result.** Lady finger Vegetable	True Negative  **Result.** NOT Lady finger Vegetable

PART B

Example Case 1

Let us assume that you developed an *AI model that tests pooled specimen (blood/urine/mucus/cell tissues etc.) to diagnose some ailment (say Covid). After its training with sample collection of specimens whose accurate results were known to you, post testing, you are now ready to evaluate your AI model. For your AI model, you conduct about 630 tests and the confusion matrix with these 630 tests results looked like :*

Table 7.2 *Confusion Matrix for an AI model being evaluated after 630 tests*

N = 630 Actual Values	Predicted Values	
	Positive	*Negative*
Positive	110 (*TP*)	50 (*FN*)
Negative	60 (*FP*)	410 (*TN*)

It means, out of $N = 630$ tests :

True Positives (TP) = 110
False Positives (FP) = 60
False Negatives (FN) = 50
True Negatives (TP) = 410
Total = 630

Let us now understand each of the evaluation metrics mentioned in the previous section and learn how to calculate it.

1. Accuracy

A prediction by an AI model is considered correct only when the predicted result (the outcome) matches the actual value (the reality). Accuracy of an AI model is determined as **the percentage of correct predictions** (all True cases, *i.e.*, *TP* + *TN*) **out of all the observations or tests-conducted** (*N*, *i.e.*, *TP* + *FP* + *TN* + *FN*)

The Formula to determine Accuracy is :

$$\text{Accuracy} = \frac{\text{Number of correct predictions } (TP+TN)}{\text{Total number of predictions made } (TP+TN+FP+FN)} \times 100\%$$

Thus, the **accuracy** for the AI model, as per the Confusion Matrix given in Table 7.2 would be :

$$\textbf{Accuracy} = \frac{110(TP)+410(TN)}{110(TP)+50(FN)+60(FP)+410(TN)} \times 100\%$$

$$\textbf{Accuracy} = \frac{520}{630} \times 100\% = 82.5396825\%$$

Thus, the **Accuracy** of our sample AI model is 82.5396825%

2. Precision

Precision refers to **how accurate it truly predicted positive out of all its positive predictions**. In other words, it is the percentage of ***True Positive*** cases out of all the cases where the prediction is true (***True positives and False Positives***). A model with high precision is trustworthy.

Thus, the formula for **Precision** rate is : $\frac{TP}{(TP+FP)}$

In percentage, **Precision** rate is : $\frac{TP}{(TP+FP)} \times 100\%$

Thus, precision for our sample AI model as per table 7.2's confusion matrix is :

$$\textbf{Precision} = \frac{110}{110+50} \times 100\% = \frac{110}{170} \times 100\% = 64.7058824\%$$

A model with high precision is considered trustworthy.

Thus, the precision rate for our sample AI model is 0.64706 ; in percentage 64.7058824%

3. Recall

Recall indicates **out of all actually positive values; how many are predicted positive.** ***Recall*** measures fraction of positive cases that are correctly identified. It is a ratio of correct positive predictions to the overall number of positive instances in the dataset.

The formula to compute Recall is

$$\text{Recall} = \frac{\text{Predictions actually positive}}{\text{Actual positive values in the dataset}} = \frac{TP}{(TP+FN)}$$

$$\text{In percentage} = \frac{TP}{(TP+FN)} \times 100\%$$

For our example case given above (Table 7.2),

$$\textbf{Recall} = \frac{110}{110+50} \times 100\% = \frac{110}{160} \times 100\% = 68.75\%$$

Thus, the Recall for our sample AI model is 0.6875 and in percentage 68.75%.

As we can see, the ***Precision*** ($\approx$64.71%) and ***Recall*** ($\approx$68.75%) are both lower than ***Accuracy*** ($\approx$82.54%), for our example case.

Significance of *Precision or Recall* Metrics

Carefully examining the formulas of Precision and recall, you will observe that **Precision** counts the ***False Positives*** while **Recall** takes ***False Negatives*** into consideration. How is this information useful ? **Precision** is used as a metric when our objective is ***to minimize false positives*** and **Recall** is used when the objective is ***to minimize false negatives***.

PART B

4. F1 Score (F-Measure)

When avoiding both *False Positives* and *False Negatives,* is equally important for our problem, we need a trade-off between ***Precision*** and ***Recall***, which **F1 Score** metric provides. **F1 Score refers to a metric that balances *Precision* and *Recall*** and hence balances the impact of *False Positives* and *False Negatives*. It is computed as per the following formula :

$$F1 = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} = \frac{TP}{TP + \frac{1}{2}(FP + FN)}$$

So $F1$ score for our example case 1 (Confusion matrix in Table 7.2) will be :

$$F1 = \frac{110}{\left(110 + \frac{1}{2} \times (60 + 50)\right)} = 0.66666666666$$

So, for our example case, the $F1$ score is 0.6666666.

In an ideal situation both *Precision* and *Recall* will be 100% (*i.e.*, value of 1). In that case, the $F1$ score would also be an ideal 1 (100%), known as the perfect value for $F1$ Score. The metrics ***Precision***, ***Recall*** and ***F1* score** range from 0 to 1.

F1 Score refers to a metric that balances *Precision* and *Recall* and hence balances the impact of *False Positives* and *False negatives*.

Let Us Revise

- *Evaluation refers to systematically checking and analysing the merit, correctness and reliability of an AI model based on the outputs produced by it.*
- *Evaluation metrics refers to the measures used to test the quality of the AI model.*
- *Overfitting refers to a situation when an AI model performs so well as the test data it got, fitted exactly against its training data and thus AI model always produced correct result.*
- *Underfitting refers to a situation when an AI model is not complex enough to capture the structure and relationships of its training data and predict effective outcomes.*
- *Overfitting and underfitting are the two biggest causes for poor performance of AI models.*
- *Ideally, an AI model should be balanced between underfitting and overfitting to be a good fit.*
- *True positive refers to an instance for which both predicted values by the AI model and actual values are positive.*
- *True negative refers to an instance for which both predicted by the AI model and actual values are negative.*
- *False positive refers to an instance for which predicted value by the AI model is positive but actual value is negative.*
- *False negative refers to an instance for which predicted value by the AI model is negative but actual value is positive.*
- *Accuracy of an AI model is determined as the percentage of correct predictions out of all the observations or tests-conducted.*
- *Precision refers to how accurate it truly predicted positive out of all its positive predictions.*
- *Recall indicates out of all actually positive values; how many are predicted positive.*
- *Precision is used as a metric when our objective is to minimize false positives and Recall is used when the objective is to minimize false negatives.*
- *F1 Score refers to a metric that balances Precision and Recall and hence balances the impact of False Positives and False negatives.*
- *For all the AI models developed, the AI model with the higher F1 score is chosen.*