Chapter -1 Overview of ML
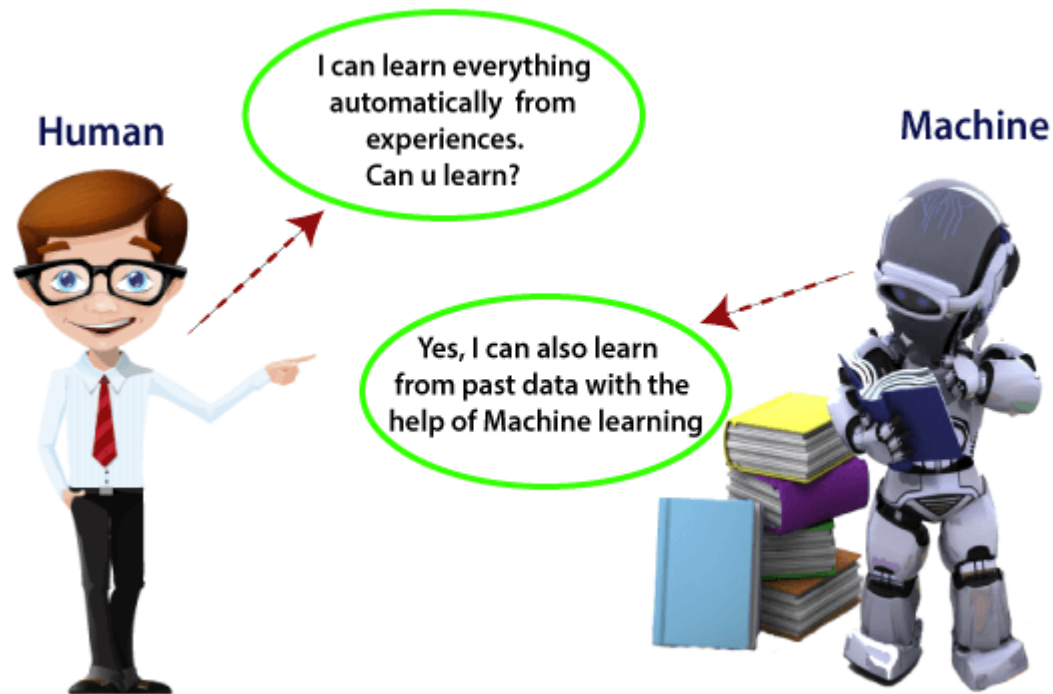
What is Machine Learning?

▪Machine learning is a **growing technology** which enables computers to learn automatically **from past data**

▪we have **computers or machines** which work on **our instructions.** But can a machine also learn from experiences or past data like a human does**?**

What is Machine Learning?

- The term machine learning was first introduced by **Arthur Samuel in 1959**

- **He defined machine learning as "**the field of study that **gives computers the ability** to learn **without being explicitly** programmed**."**

- However, there is **no universally accepted definition**. Different authors define the term differently.

- **Tom Mitchell (1998):**

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

What is Machine Learning?

Breaking Down

- **Experience (E):** This refers to the **data or examples** the system is exposed to

- **Tasks (T):** These are the **specific activities or problems** that the machine is **learning** to perform or solve.

- **Performance Measure (P):** This is the **metric used to evaluate** how well the machine is doing in performing tasks from T.   accuracy, error rate, or another relevant metrics.

- Example
Handwriting recognition learning problem,

 Task T: Recognizing and classifying handwritten words within images,

 Performance P: Percent of words correctly classified,

 Experience E: A dataset of handwritten words

What is Machine Learning?

Email Spam Detection Problem:

Task T: classify email as spam or not-spam

Performance P: number of correctly predicted spams

Experience E: A dataset of labelled emails
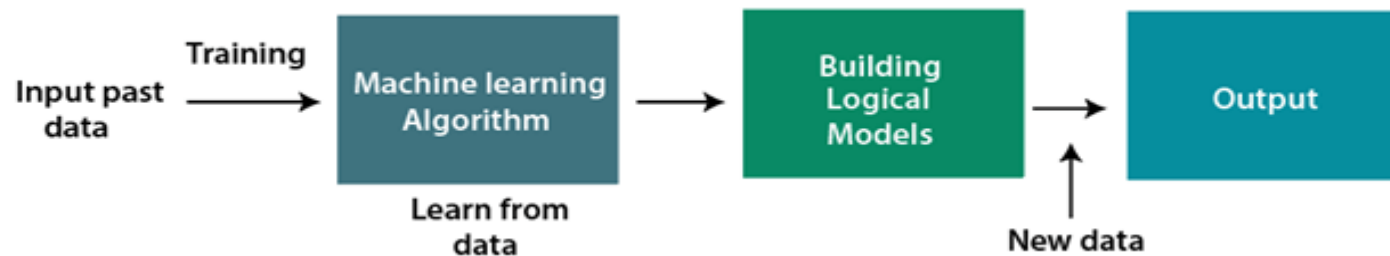
Other Def:

▪Machine learning is **programming computers** to **optimize a performance** criterion using example data or past experience.

▪It is a **branch of Artificial Intelligence and computer science** that helps **build a model based on training data** and make predictions and decisions without being constantly programmed.

How ML Help Us?

- It is very **hard to write programs** that solve problems like **recognizing a face.**

  - We **don't know what program** to write because we **don't know how our brain does it**.

  - Even if we **had a good idea about how to do it**, the program might be horrendously complicated.

A. Instead of writing a program by hand, **we collect lots of examples** that specify the correct output for a given input.

B. A machine learning algorithm then **takes these examples and produces** a program that does the job.

How ML works?

- A Machine Learning system **learns from historical data**, builds the prediction models(Using Algorithms), and **whenever it receives new data,** predicts the output for it.

- The **accuracy of predicted output depends** upon the amount of data, as the huge amount of data helps to build a better model

Features of ML

- **uses data to detect various patterns in a given dataset.**

- **It can learn from past data and improve automatically.**

- It is a data-driven technology

- **is much similar to data mining as it also deals with the huge amount of the data.**
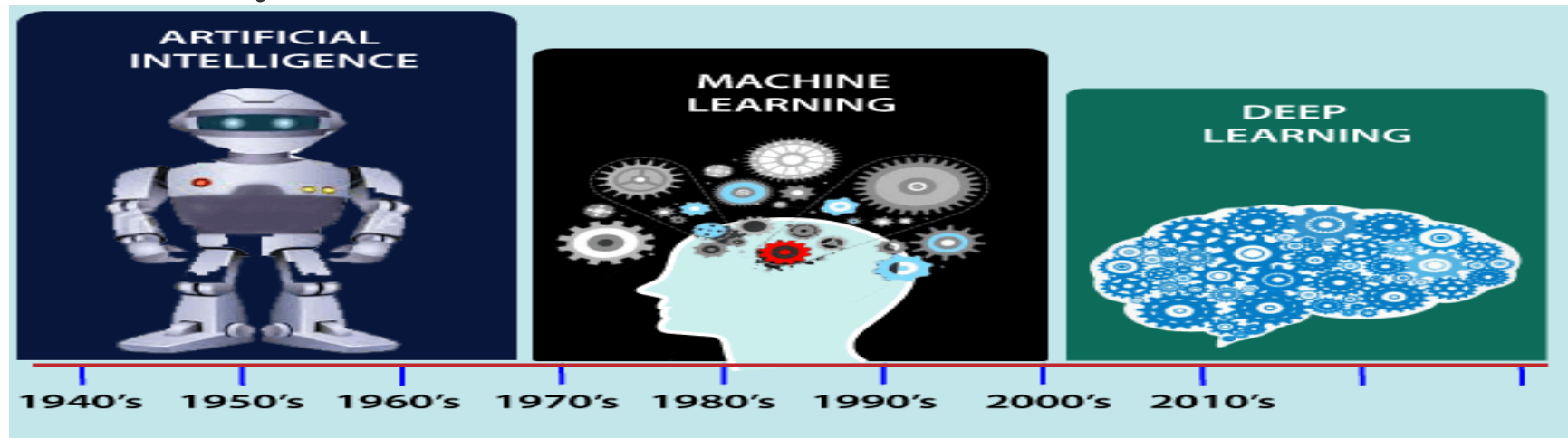
Need of ML

The need for machine learning is **increasing day by day.** The reason behind the need for machine learning is that it is **capable of doing tasks that are too complex** for a person to implement directly.

- Rapid increment in the **production of data**

- **Solving complex problems,** which are difficult for a human

- **Decision making** needs in various sector including finance

- **Finding hidden patterns** and extracting useful information from data.

History of ML

Before some years , **machine learning was science fiction,** but today it is the part of our daily life.

Has old history

History of ML

The early history of Machine Learning (Pre-1940):

- 1834: Charles Babbage, the father of the computer, conceived a device that could be programmed with punch cards.

- 1936: Alan Turing gave a theory that how a machine can determine and execute a set of instructions.

The era of stored program computers:

1940: the first manually operated computer, "ENIAC" was invented, which was the first electronic general-purpose computer.

1950, the scientists started applying their idea to work and analyzed how human neurons might work.

Computer Machinery and Intelligence:

1950: Alan Turing published a seminal paper, "Computer Machinery and Intelligence," on the topic of artificial intelligence. In his paper, he asked, "Can machines think?"

Machine intelligence in Games:

1952: Arthur Samuel,created a program that helped an IBM computer to play a checkers game

History of ML...

The first "AI" winter:

The duration of 1974 to 1980 was the tough time for AI and ML researchers, and this duration was called as AI winter.

Machine Learning from theory to reality

- 1959: the first neural network was applied to a real-world problem to remove echoes over phone lines

- 1985: Terry Sejnowski and Charles Rosenberg invented a neural network **NETtalk,** which was able to teach itself how to correctly pronounce 20,000 words in one week.

- 1997: The IBM's Deep blue intelligent computer won the chess game against the chess expert Garry Kasparov, and it became the first computer which had beaten a human chess expert.

History of ML ...

Machine Learning at 21$_{st}$ century

Deep learning

Deep Face

Chat boat technologies and others invented

Machine Learning at Present

Now machine learning has got a **great advancement** in its research, and **it is present everywhere around us**, such as self-driving cars, Catboats, recommender system,Various AI technologies and many more
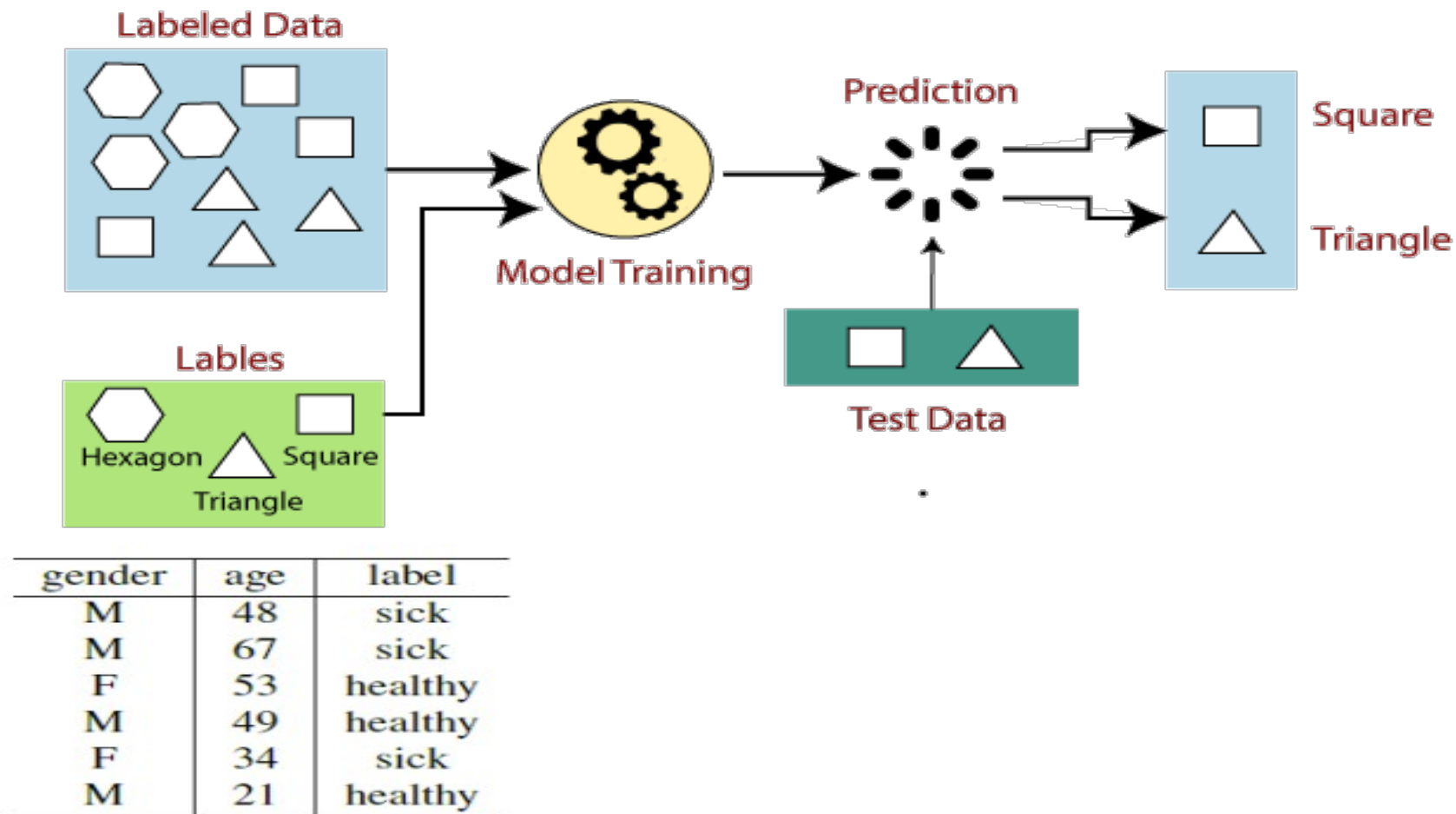
Classification of ML

In general, machine learning algorithms can be classified into three types.

- ☐ Supervised learning
- ☐ Unsupervised learning
- ☐ Reinforcement learning

Supervised Learning

▪ Supervised learning involves training an algorithm on a labeled dataset, where input data is paired with corresponding output labels.

▪ The goal is to learn a mapping from input to output based on provided labelled examples.

• A "supervised learning" is so called because the process of an algorithm learning from the training dataset can be thought of as a teacher supervising the learning process

Supervised Learning

Labeled Data

Lables

Hexagon Square
Triangle

Model Training

Prediction

Test Data

Square

Triangle

gender	age	label
M	48	sick
M	67	sick
F	53	healthy
M	49	healthy
F	34	sick
M	21	healthy

Supervised Learning

Types of Problem To Solve

✔ Classification:

Predicting whether an email is spam or not based on features like subject, sender, and content.
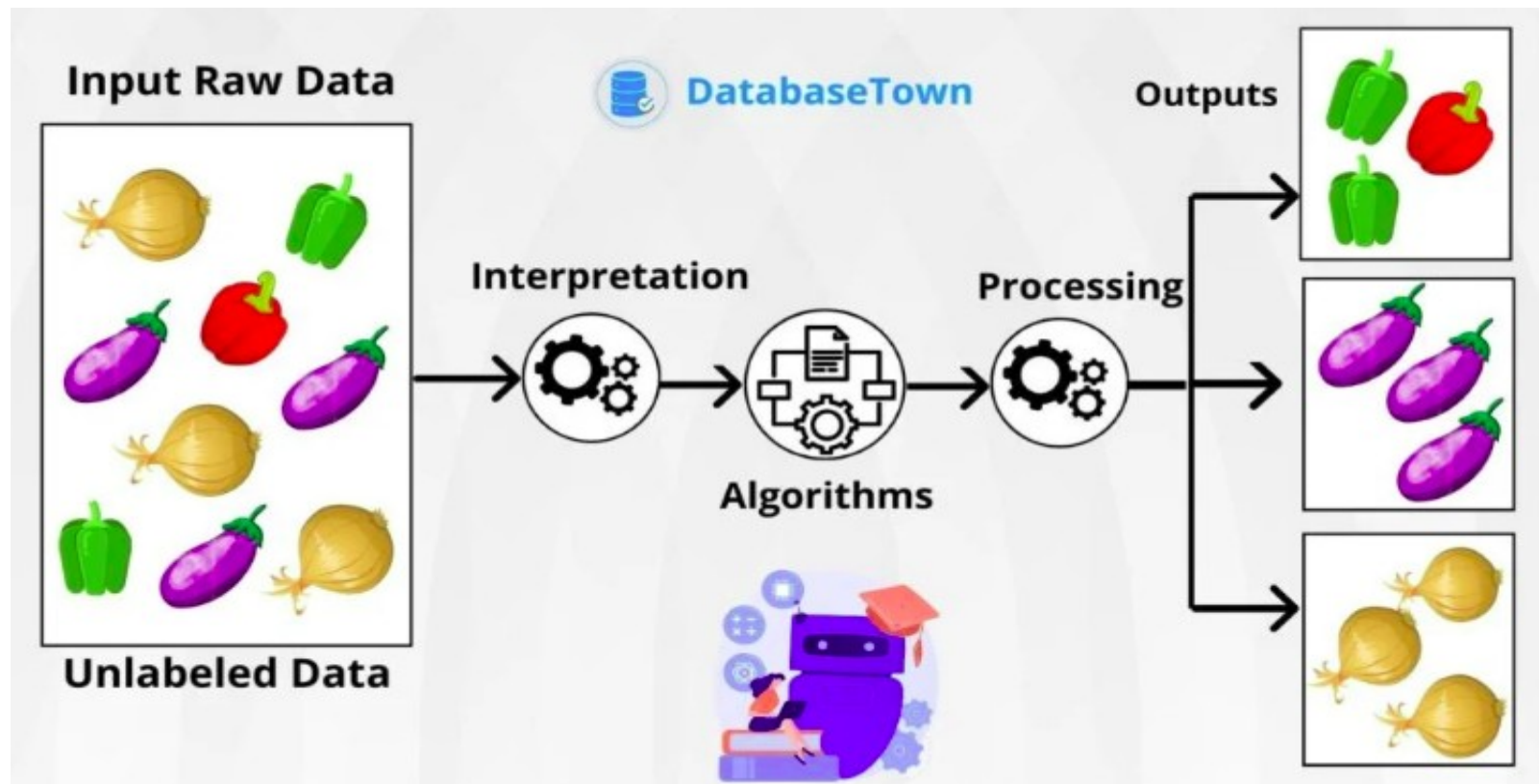
✔ Regression:

Predicting the price of a house based on features like square footage, location, and the number of bedrooms.

Health care, financial, sentiment....

Unsupervised Learning

- Unsupervised learning involves training a model on **not labeled, classified, or categorized, data** and the algorithm needs to act on that data without any supervision,

- The algorithm aims to **explore the inherent structure** in the data.

- models itself find the **hidden patterns and insights** from the given data.

- The goal is to **restructure the input data** into **new features or a group of objects** with similar patterns.

- It can be compared to learning which takes place in the human brain while learning new things

Unsupervised Learning

Unsupervised ML Problem Types

A. Clustering:

Grouping similar customers based on **purchasing behaviour** without predefined categories.

- Market Segmentation: Identifying **distinct customer segments** for targeted marketing campaigns.

- Anomaly Detection: Detecting **unusual patterns in network traffic** for cybersecurity.

B. Association :

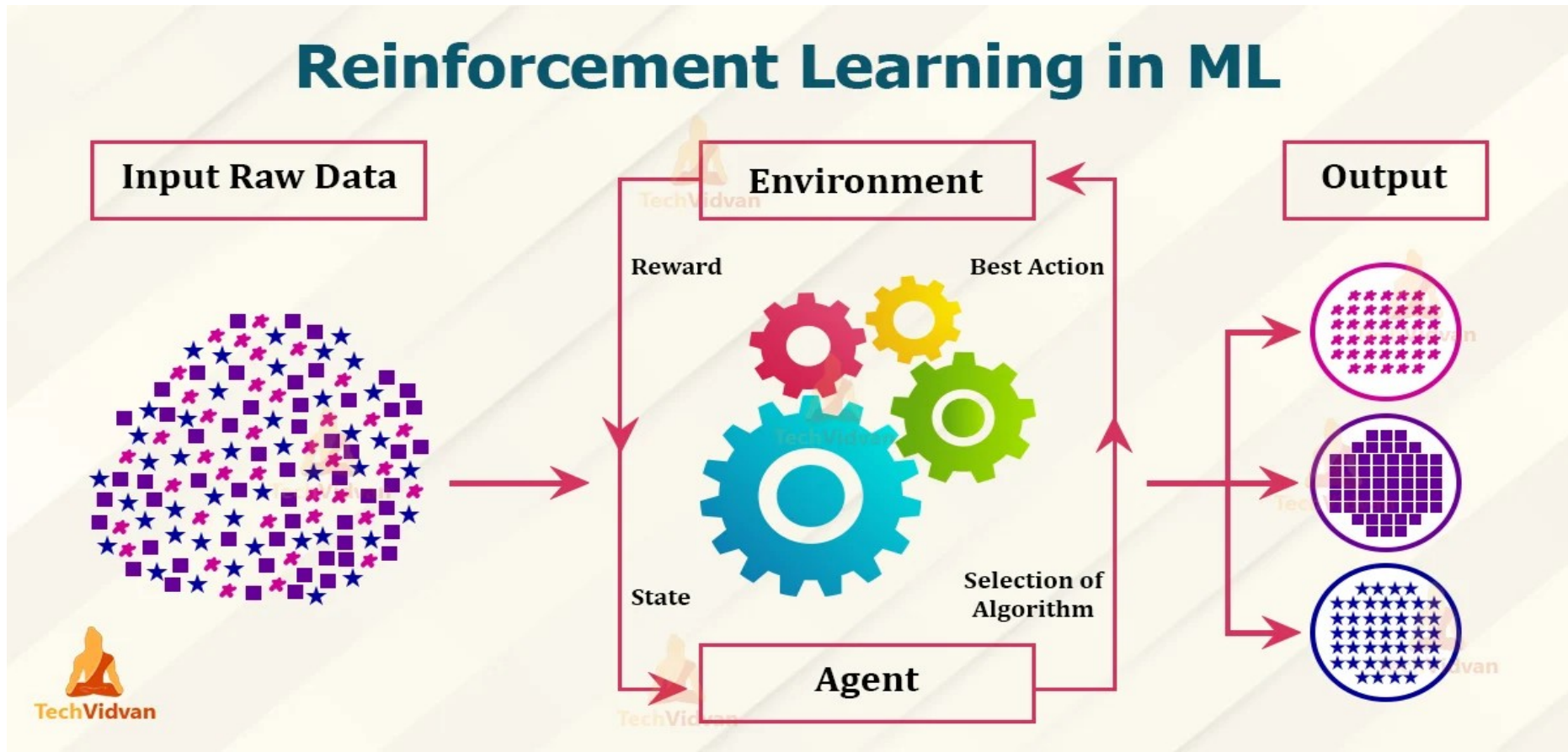   find associations and patterns in the data ex : dimensionality reduction

More examples

Re-Inforcement Learning

- Reinforcement learning involves an **agent learning** to **make decisions** by **interacting** with an environment.

- The agent **receives feedback** in the form of rewards or penalties based on the actions it takes.

- The agent **learns automatically** with **these feedbacks** and improves its **performance**

- Reinforcement learning is applicable when an agent needs to learn a sequence of actions to achieve a goal in a dynamic environment, receiving feedback to guide its learning process.

21

Re-Inforcement Learning

Re-Inforcement Learning ex:

Controlling A Walking Robot

Agent:

The **program controlling** a walking robot.

Environment:

The real world.

Action:

One out of four moves (1) forward; (2) backward; (3) left; and (4) right.

Reward:

Positive when it **approaches the target destination**; negative when it wastes time, goes in the wrong direction or falls down.

Example 2: The robotic dog, which automatically learns the movement of his arms,
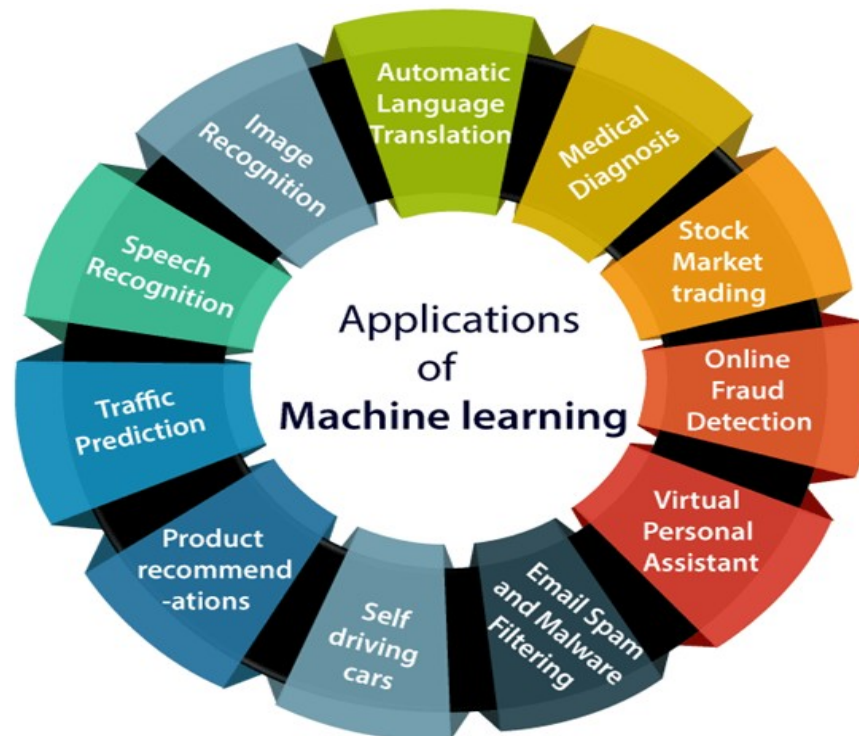
Challenges In Machine Learning

- Machine learning projects **encounter challenges** such as

- **data quality** issues,

- **insufficient** data

- the need for **robust model evaluation.**

Limitations of current machine learning approaches:

While powerful, current ML approaches have limitations, including

interpretability issues, susceptibility to bias, and challenges in handling

complex, unstructured data.

Understanding these limitations is crucial for **responsible and effective deployment.**

Application Areas of ML

Application areas of ML

A. Image Recognition

- It is used to identify objects, persons, places, digital images

- Facebook Deepface - Automatic friend tagging suggestion

B. Speech Recognition

- process of converting v**oice instructions into text**, and it is also known as **"Speech to text"**, or **"Computer speech recognition."**

- Google Search by voice,

    Example: **Google Assistant, Siri, Cortana(Microsoft), and Alexa(amazon)**

C. Traffic prediction : google map

D. Product recommendations:

- e-commerce and entertainment companies such as **Amazon, Netflix**

Application areas of ML

E. **Self-driving cars**

Tesla, the most popular car manufacturing company is working on **self-driving car**

F. **Email Spam and Malware Filtering**

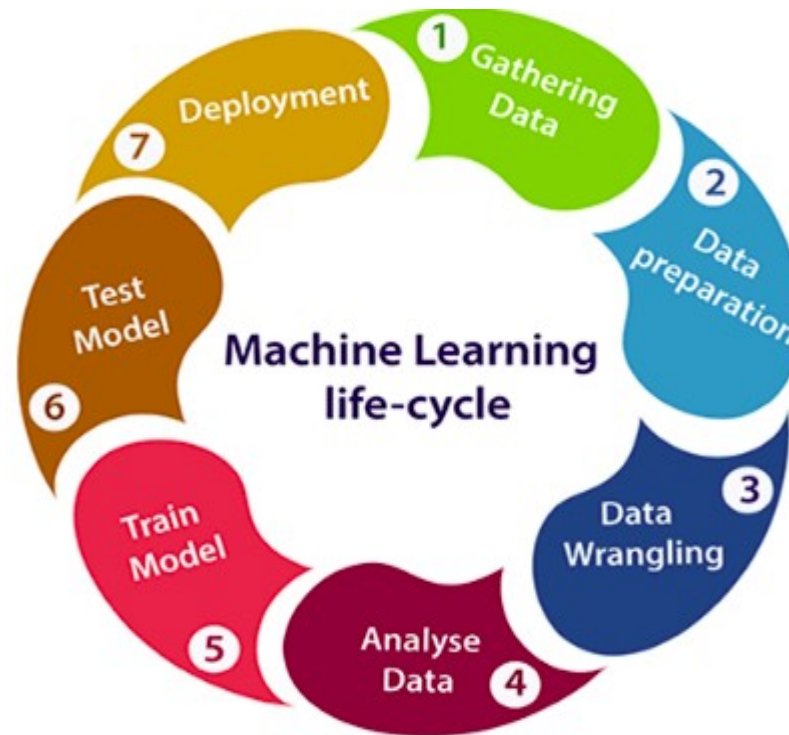- When we receive new email, it is filtered automatically as **important, normal, and spam**

G. **Virtual Personal Assistant**

- Help us using **voice instructions** such as Play music, call someone, Open an email, Scheduling an appointment,

- various virtual personal assistants such as **Google assistant, Alexa, Cortana, Siri.**

Machine learning Life cycle

Machine learning life cycle is a **cyclic process** to build an efficient machine learning project

☐ Gathering Data

☐ Data preparation

☐ Data Wrangling

☐ Analyse Data

☐ Train the model

☐ Test the model

☐ Deployment

Steps of ML

1. Gathering Data

→ Identify various data sources ,Collect data and Integrate the data obtained from different sources-> **datasets**

→ Open data sources: UCI ML repository,kaggle,OpenML,google and Microsoft Research Open Data

→ **2.Data Preparation:** a step where we **put our data into a suitable place and prepare it to use in our machine learning** training ex: **Data exploration,EDA**

→ **3.Data Wrangling :** is the process of **cleaning and converting raw** data into a useable format.

Missing Values ,Duplicate data ,Invalid data , Noise removal

Steps of ML

4. **Data Analysis :** Now the **cleaned and prepared data** is passed on to the **analysis step.** This step involves: Selection of **analytical techniques ,Building models ,Review the result**

5. **Train Model :** -in this step we **train our model to improve its performance** for better outcome of the problem.

6. **Test Model :** once trained we check for the **accuracy of our model by** providing **a test dataset to it.**

7. **Deployment**

we deploy the model in the **real-world system.**

Data Science

What is Data science

"A field of **deep study of data** that includes **extracting useful insights from the data**, and processing that information using different tools, statistical models, and Machine learning algorithms.

- Concept that is used to **handle big data that** includes **data cleaning, data preparation, data analysis, and data visualization**.
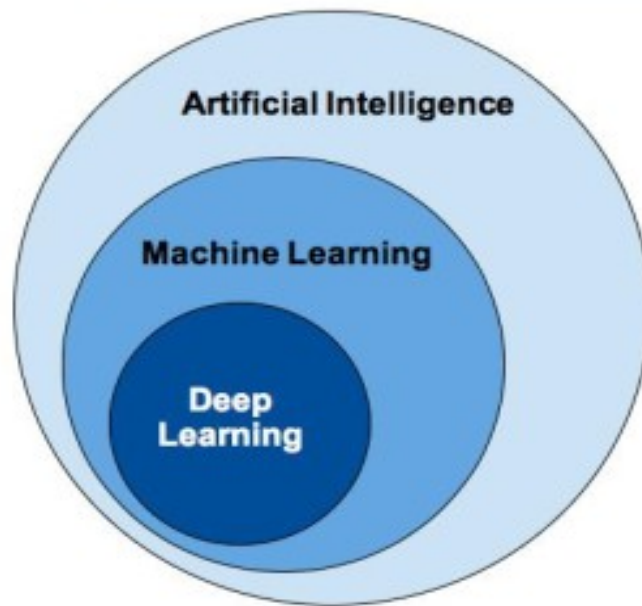
Skills needed to be Data Scientist

☐ An excellent programming **knowledge of Python, R, SAS, or Scala.**

☐ Experience in **SQL database Codin**g.

☐ Knowledge of **Machine Learning Algorithms**.

☐ Deep Knowledge of **Statistics concepts.**

☐ **Data Mining, cleaning, and visualizing** skills.

☐ Skills to use **Big data tools such as Hadoop.**

Data Science VS ML

Data Science	Machine Learning
It deals with understanding and finding hidden patterns or useful insights from the data, which helps to take smarter business decisions.	It is a subfield of data science that enables the machine to learn from the past data and experiences automatically.
It is used for discovering insights from the data.	It is used for making predictions and classifying the result for new data points.
It is a broad term that includes various steps to create a model for a given problem and deploy the model.	It is used in the data modeling step of the data science as a complete process.
A data scientist needs to have skills to use big data tools like Hadoop, Hive and Pig, statistics, programming in Python, R, or Scala.	Machine Learning Engineer needs to have skills such as computer science fundamentals, programming skills in Python or R, statistics and probability concepts, etc.
It can work with raw, structured, and unstructured data.	It mostly requires structured data to work on.
Data scientists spent lots of time in handling the data, cleansing the data, and understanding its patterns.	ML engineers spend a lot of time for managing the complexities that occur during the implementation of algorithms and mathematical concepts behind that.

AI vs ML vs DL

- **AI** : broader concept (science+Engineering) to create intelligent machines that can simulate human thinking capability.

- **ML** : subset of AI that allows machines to learn from data without being programmed explicitly.

- **DL:** subset of ML, that uses the neural networks to analyze different factors with a structure that is similar to the human neural system.

AI,ML and DL Summary

- **Artificial intelligence** is the overarching field focused on **creating intelligent systems.**

- **Machine learning** is a subset of artificial intelligence that emphasizes **learning from data to make predictions or decisions.**

- **Deep learning** is a **subfield of machine learning** that uses **neural networks** with **multiple layers to learn hierarchical representations** of data.

End of Chapter-one