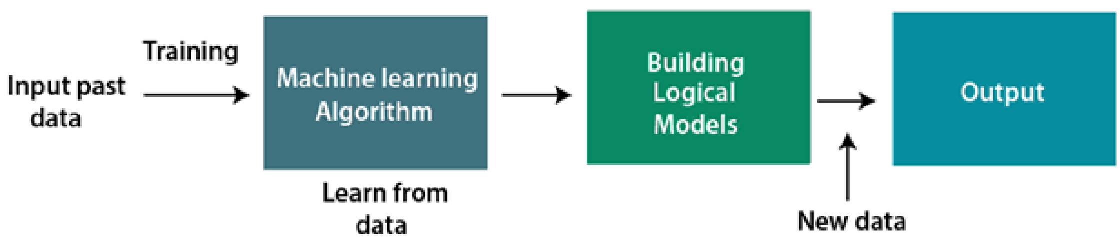
UNIT-I: Introduction

Table of Content

1. Introduction

1.1 What Is Machine Learning?

Machine learning is programming computers to optimize a performance criterion using example data or past experience. We have a model defined up to some parameters, and learning is the execution of a computer program to optimize the parameters of the model using the training data or past experience. The model may be *predictive* to make predictions in the future, or *descriptive* to gain knowledge from data, or both.
Arthur Samuel, an early American leader in the field of computer gaming and artificial intelligence, coined the term “Machine Learning” in 1959 while at IBM. 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 for machine learning. Different authors define the term differently.



Definition of learning

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

Examples
i) Handwriting recognition learning problem
- Task T: Recognising and classifying handwritten words within images
- Performance P: Percent of words correctly classified
- Training experience E: A dataset of handwritten words with given classifications ii) A robot driving learning problem
- Task T: Driving on highways using vision sensors
- Performance measure P: Average distance traveled before an error
- training experience: A sequence of images and steering commands recorded while observing a human driver iii) A chess learning problem
- Task T: Playing chess
- Performance measure P: Percent of games won against opponents
- Training experience E: Playing practice games against itself

Definition

A computer program which learns from experience is called a machine learning program or simply a learning program. Such a program is sometimes also referred to as a learner.

1.2 Components of Learning

Basic components of learning process

The learning process, whether by a human or a machine, can be divided into four components, namely, data storage, abstraction, generalization and evaluation. Figure 1.1 illustrates the variouscomponents and the steps involved in the learning process.

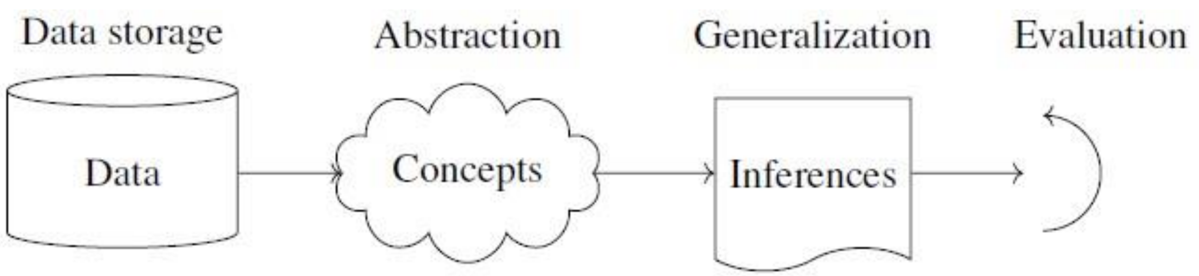


Figure 1.1: Components of learning process

1. Data storage

Facilities for storing and retrieving huge amounts of data are an important component of the learning process. Humans and computers alike utilize data storage as a foundation for advanced reasoning.

- In a human being, the data is stored in the brain and data is retrieved using electrochemical signals.
- Computers use hard disk drives, flash memory, random access memory and similar devices to store data and use cables and other technology to retrieve data.

2. Abstraction

The second component of the learning process is known as abstraction.

Abstraction is the process of extracting knowledge about stored data. This involves creating general concepts about the data as a whole. The creation of knowledge involves application of known models and creation of new models.

The process of fitting a model to a dataset is known as training. When the model has been trained, the data is transformed into an abstract form that summarizes the original information.

3. Generalization

The third component of the learning process is known as generalisation.

The term generalization describes the process of turning the knowledge about stored data into a form that can be utilized for future action. These actions are to be carried out on tasks that are similar, but not identical, to those what have been seen before. In generalization, the goal is to discover those properties of the data that will be most relevant to future tasks.

4. Evaluation

Evaluation is the last component of the learning process.

It is the process of giving feedback to the user to measure the utility of the learned knowledge. This feedback is then utilised to effect improvements in the whole learning process

Applications of machine learning

Application of machine learning methods to large databases is called data mining. In data mining, a large volume of data is processed to construct a simple model with valuable use, for example, having high predictive accuracy.

The following is a list of some of the typical applications of machine learning.

1. In retail business, machine learning is used to study consumer behaviour.
2. In finance, banks analyze their past data to build models to use in credit applications, fraud detection, and the stock market.
3. In manufacturing, learning models are used for optimization, control, and troubleshooting.
4. In medicine, learning programs are used for medical diagnosis.
5. In telecommunications, call patterns are analyzed for network optimization and maximizing the quality of service.
6. In science, large amounts of data in physics, astronomy, and biology can only be analyzed fast enough by computers. The World Wide Web is huge; it is constantly growing and searching for relevant information cannot be done manually.
7. In artificial intelligence, it is used to teach a system to learn and adapt to changes so that the system designer need not foresee and provide solutions for all possible situations.
8. It is used to find solutions to many problems in vision, speech recognition, and robotics.
9. Machine learning methods are applied in the design of computer-controlled vehicles to steer correctly when driving on a variety of roads.
10. Machine learning methods have been used to develop programmes for playing games such as chess, backgammon and Go.

1.3 Learning Models

Machine learning is concerned with using the right features to build the right models that achieve the right tasks. The basic idea of Learning models has divided into three categories.
For a given problem, the collection of all possible outcomes represents the **sample space or instance space**.

- Using a Logical expression. (**Logical models**)
- Using the Geometry of the instance space. (**Geometric models)**
- Using Probability to classify the instance space. (**Probabilistic models**) Grouping and Grading

1.4 Types of Learning

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

- Supervised learning
- Unsupervised learning
- Reinforcement learning

1.4.1 Supervised learning

A training set of examples with the correct responses (targets) is provided and, based on this training set, the algorithm generalises to respond correctly to all possible inputs. This is also called learning from exemplars. Supervised learning is the machine learning task of learning a function that maps an input to an output based on example input-output pairs.

In supervised learning, each example in the training set is a pair consisting of an input object (typically a vector) and an output value. A supervised learning algorithm analyzes the training data and produces a function, which can be used for mapping new examples. In the optimal case, the function will correctly determine the class labels for unseen instances. Both classification and regression problems are supervised learning problems. A wide range of supervised learning algorithms are available, each with its strengths and weaknesses. There is no single learning algorithm that works best on all supervised learning problems.

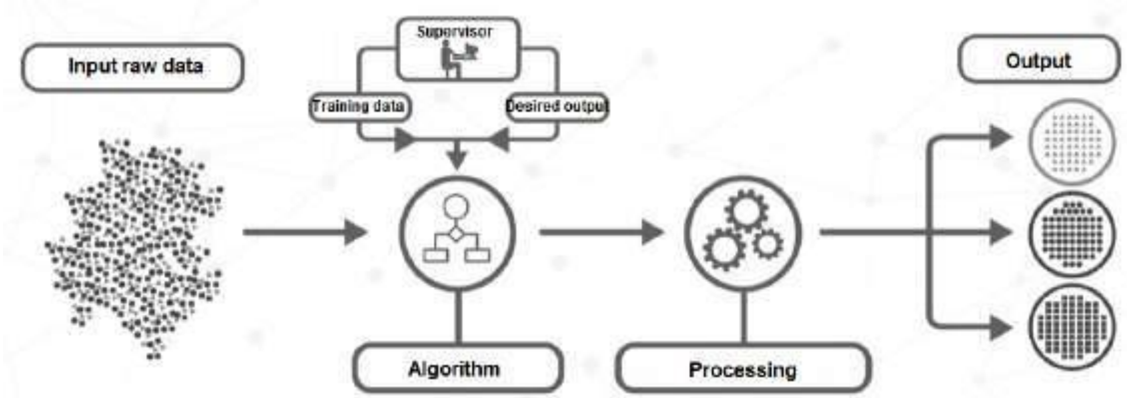


Figure 1.4: Supervised learning

Remarks

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. We know the correct answers (that is, the correct outputs), the algorithm iteratively makes predictions on the training data and is corrected by the teacher. Learning stops when the algorithm achieves an acceptable level of performance.

Example

Consider the following data regarding patients entering a clinic. The data consists of the gender and age of the patients and each patient is labeled as “healthy” or “sick”.

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

1.4.2 Unsupervised learning

Correct responses are not provided, but instead the algorithm tries to identify similarities between the inputs so that inputs that have something in common are categorised together. The statistical approach to unsupervised learning is known as density estimation.

Unsupervised learning is a type of machine learning algorithm used to draw inferences from datasets consisting of input data without labeled responses. In unsupervised learning algorithms, a classification or categorization is not included in the observations. There are no output values and so there is no estimation of functions. Since the examples given to the learner are unlabeled, the accuracy of the structure that is output by the algorithm cannot be evaluated. The most common unsupervised learning method is cluster analysis, which is used for exploratory data analysis to find hidden patterns or grouping in data.

Example

Consider the following data regarding patients entering a clinic. The data consists of the gender and age of the patients.

gender	age
M	48
M	67
F	53
M	49
F	34
M	21

Based on this data, can we infer anything regarding the patients entering the clinic?

1.4.3 Reinforcement learning

This is somewhere between supervised and unsupervised learning. The algorithm gets told when the answer is wrong, but does not get told how to correct it. It has to explore and try out different possibilities until it works out how to get the answer right. Reinforcement learning is sometime called learning with a critic because of this monitor that scores the answer, but does not suggest improvements.

Reinforcement learning is the problem of getting an agent to act in the world so as to maximize its rewards. A learner (the program) is not told what actions to take as in most forms of machine learning, but instead must discover which actions yield the most reward by trying them. In the most interesting and challenging cases, actions may affect not only the immediate reward but also the next situations and, through that, all subsequent rewards.

Example

Consider teaching a dog a new trick: we cannot tell it what to do, but we can reward/punish it if it does the right/wrong thing. It has to find out what it did that made it get the reward/punishment. We can use a similar method to train computers to do many tasks, such as playing backgammon or chess, scheduling jobs, and controlling robot limbs. Reinforcement learning is different from supervised learning. Supervised learning is learning from examples provided by a knowledgeable expert.

	Supervised Learning	Unsupervised Learning
Input Data	Uses Known and Labelled Data as input	Uses Unknown Data as input
Computational Complexity	Less Computational Complexity	More Computational Complex
Real-Time	Uses off-line analysis	Uses Real-Time Analysis of Data
Number of Classes	The number of Classes is known	The number of Classes is not known
Accuracy of Results	Accurate and Reliable Results	Moderate Accurate and Reliable Results
Output data	The desired output is given.	The desired, output is not given.
Model	In supervised learning it is not possible to learn larger and more complex models	In unsupervised learning it is possible to learn larger and more complex models than with unsupervised learning

	Supervised Learning	Unsupervised Learning
	than in, supervised learning	
Training data	In supervised learning training data is used to infer model	In unsupervised learning training data is not used.
Another name	Supervised learning is also called classification.	Unsupervised learning is also called clustering.
Test of model	We can test our model.	We can not test our model.
Example	Optical Character Recognition	Find a face in an image.

Well-Posed Learning Problems

1. **Definition**:
   - A **well-posed learning problem** arises when a computer program learns from experience **E** within the context of a specific task **T** and a performance measure **P**.
   - The program's performance on task **T**, as measured by **P**, should improve with accumulated experience **E**.
2. **Key Traits of Well-Posed Learning Problems**:
   - To classify a problem as well-posed, it must exhibit the following three features:
     - **Task**: Clearly defined task or objective.
     - **Performance Measure**: A quantifiable measure to evaluate the program's performance.
     - **Experience**: Accumulation of relevant data or observations.
3. **Examples of Well-Posed Learning Problems**:
   - Let's explore some examples that efficiently illustrate well-defined learning problems:

1. **Email Spam Classification**:
   - **Task**: Classifying emails as spam or not.
   - **Performance Measure**: Fraction of emails accurately classified.
   - **Experience**: Observing labeled emails as spam or not spam.
2. **Checkers Game Learning**:
   - **Task**: Playing checkers.
   - **Performance Measure**: Percentage of games won against opponents.
   - **Experience**: Playing implementation games against itself.
3. **Handwriting Recognition**:
   - **Task**: Recognizing handwritten words.
   - **Performance Measure**: Accuracy in classifying words.
   - **Experience**: A directory of handwritten words with known classifications.
4. **Robot Driving**:
   - **Task**: Driving on public highways using sensors.
   - **Performance Measure**: Average distance traveled before errors.
   - **Experience**: Capturing images and steering instructions while observing a human driver.
5. **Fruit Prediction**:
   - **Task**: Predicting different fruits.
   - **Performance Measure**: Ability to predict a wide variety of fruits.
   - **Experience**: Training with a large dataset of fruit images.
6. **Face Recognition**:
   - **Task**: Identifying different types of faces.
   - **Performance Measure**: Ability to predict various face types.
   - **Experience**: Training with diverse face image datasets.
7. **Automatic Translation of Documents**:
   - **Task**: Translating one language to another.
   - **Performance Measure**: Efficient conversion between languages.
   - **Experience**: Training with a large dataset of different languages.

In summary, a well-defined learning problem involves understanding the problem, gathering relevant data, selecting appropriate learning algorithms, training models, and evaluating results. It's the foundation for successful machine learning endeavors!

Designing a Learning System

1. **Choosing the Training Experience**:
   - The first and essential task is selecting the **training data** or **training experience** that will be fed to the machine learning algorithm.
   - The quality and relevance of this data significantly impact the success or failure of the resulting model.
   - Consider the following attributes when choosing training data:
     - **Representativeness**: Ensure that the data represents the problem domain adequately.
     - **Completeness**: The dataset should cover a wide range of scenarios and examples.

- **Quality**: High-quality data leads to better model performance.
- **Balance**: Maintain a balance between different classes or categories.
- **Relevance**: Focus on data relevant to the specific task.
- **Consistency**: Ensure consistency across data points.

2. **Defining the Task (T)**:
   - Clearly define the task that the learning system needs to perform.
   - For example, in **spam email detection**:
     - **Task (T)**: Classify emails as "Spam" or "Not Spam."
     - **Performance measure (P)**: The percentage of correctly classified emails.
     - **Experience (E)**: The set of labelled emails (e.g., "Spam" or "Not Spam").
3. **Choosing the Target Function** :
   - The target function represents what the system aims to learn.
   - In our spam email detection example:
     - **Target function ©**: The decision boundary that separates spam from non-spam emails.
4. **Selecting a Representation**:
   - Decide how to represent the target function.
   - Common representations include:
     - **Rules**: Explicit rules or conditions.
     - **Mathematical Models**: Equations or algorithms.
     - **Neural Networks**: Complex architectures for deep learning.
5. **Choosing a Learning Algorithm**:
   - Select an appropriate learning algorithm to infer the target function from the training data.
   - Algorithms can include decision trees, neural networks, support vector machines, etc.
6. **Model Evaluation and Iteration**:
   - Evaluate the model's performance using validation data.
   - Iterate by adjusting hyperparameters, modifying the representation, or selecting a different algorithm.

Issues in Machine Learning

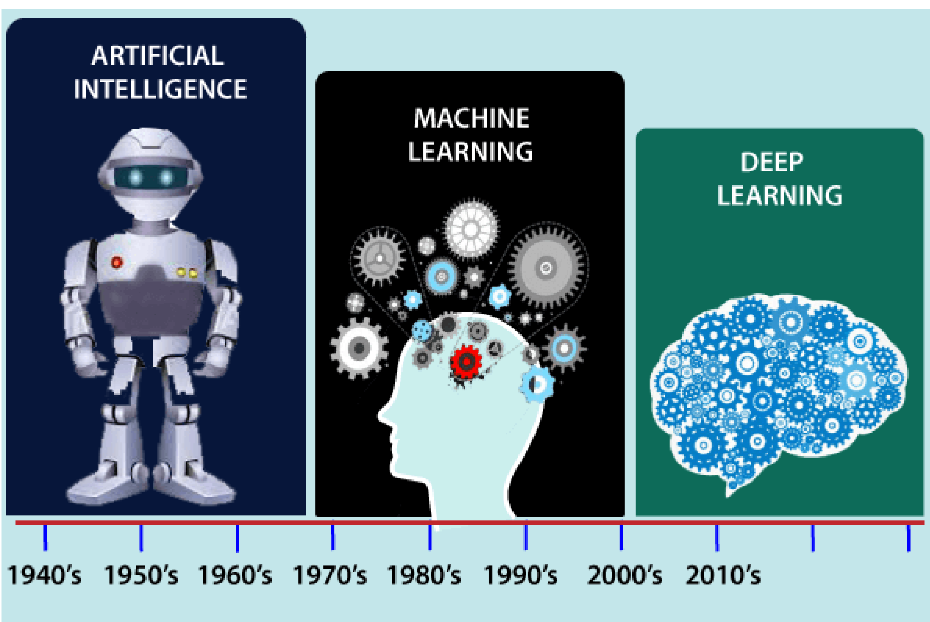
Our checkers example raises a number of generic questions about machine learning. The field of machine learning, and much of this book, is concerned with answering questions such as the following:

- What algorithms exist for learning general target functions from specific training examples? In what settings will particular algorithms converge to the desired function, given sufficient training data? Which algorithms perform best for which types of problems and representations?
- How much training data is sufficient? What general bounds can be found to relate the confidence in learned hypotheses to the amount of training experience and the character of the learner's hypothesis space?

- When and how can prior knowledge held by the learner guide the process of generalizing from examples? Can prior knowledge be helpful even when it is only approximately correct?
- What is the best strategy for choosing a useful next training experience, and how does the choice of this strategy alter the complexity of the learning problem?
- What is the best way to reduce the learning task to one or more function approximation problems? Put another way, what specific functions should the system attempt to learn? Can this process itself be automated?
- How can the learner automatically alter **its** representation to improve its ability to represent and learn the target function?

History of Machine Learning

Before some years (about 40-50 years), machine learning was science fiction, but today it is the part of our daily life. Machine learning is making our day to day life easy from self-driving cars to Amazon virtual assistant "Alexa". However, the idea behind machine learning is so old and has a long history. Below some milestones are given which have occurred in the history of machine learning:



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

1834: In 1834, Charles Babbage, the father of the computer, conceived a device that could be programmed with punch cards. However, the machine was never built, but all modern computers rely on its logical structure.

1936: In 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: In 1940, the first manually operated computer, "ENIAC" was invented, which was the first electronic general-purpose computer. After that stored program computer such as EDSAC in 1949 and EDVAC in 1951 were invented.

1943: In 1943, a human neural network was modeled with an electrical circuit. In 1950, the scientists started applying their idea to work and analyzed how human neurons might work.

Computer machinery and intelligence:

1950: In 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, who was the pioneer of machine learning, created a program that helped an IBM computer to play a checkers game. It performed better more it played.

1959: In 1959, the term "Machine Learning" was first coined by Arthur Samuel.

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.

In this duration, failure of machine translation occurred, and people had reduced their interest from AI, which led to reduced funding by the government to the researches.

Machine Learning from theory to reality

1959: In 1959, the first neural network was applied to a real-world problem to remove echoes over phone lines using an adaptive filter.

1985: In 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.

Machine Learning at 21st century

2006:

Geoffrey Hinton and his group presented the idea of profound getting the hang of utilizing profound conviction organizations.

The Elastic Compute Cloud (EC2) was launched by Amazon to provide scalable computing resources that made it easier to create and implement machine learning models.

2007:

Participants were tasked with increasing the accuracy of Netflix's recommendation algorithm when the Netflix Prize competition began.

Support learning made critical progress when a group of specialists utilized it to prepare a PC to play backgammon at a top-notch level.

2008:

Google delivered the Google Forecast Programming interface, a cloud-based help that permitted designers to integrate AI into their applications.

Confined Boltzmann Machines (RBMs), a kind of generative brain organization, acquired consideration for their capacity to demonstrate complex information conveyances.

2009:

Profound learning gained ground as analysts showed its viability in different errands, including discourse acknowledgment and picture grouping.

The expression "Large Information" acquired ubiquity, featuring the difficulties and open doors related with taking care of huge datasets.

2010:

The ImageNet Huge Scope Visual Acknowledgment Challenge (ILSVRC) was presented, driving progressions in PC vision, and prompting the advancement of profound convolutional brain organizations (CNNs).

2011:

On Jeopardy! IBM's Watson defeated human champions., demonstrating the potential of question-answering systems and natural language processing.

2012:

AlexNet, a profound CNN created by Alex Krizhevsky, won the ILSVRC, fundamentally further developing picture order precision and laying out profound advancing as a predominant methodology in PC vision.

Google's Cerebrum project, drove by Andrew Ng and Jeff Dignitary, utilized profound figuring out how to prepare a brain organization to perceive felines from unlabeled YouTube recordings.

2013:

Ian Goodfellow introduced generative adversarial networks (GANs), which made it possible to create realistic synthetic data.

Google later acquired the startup DeepMind Technologies, which focused on deep learning and artificial intelligence.

2014:

Facebook presented the DeepFace framework, which accomplished close human precision in facial acknowledgment.

AlphaGo, a program created by DeepMind at Google, defeated a world champion Go player and demonstrated the potential of reinforcement learning in challenging games.

2015:

Microsoft delivered the Mental Toolbox (previously known as CNTK), an open-source profound learning library.

The performance of sequence-to-sequence models in tasks like machine translation was enhanced by the introduction of the idea of attention mechanisms.

2016:

The goal of explainable AI, which focuses on making machine learning models easier to understand, received some attention.

Google's DeepMind created AlphaGo Zero, which accomplished godlike Go abilities to play without human information, utilizing just support learning.

2017:

Move learning acquired noticeable quality, permitting pretrained models to be utilized for different errands with restricted information.

Better synthesis and generation of complex data were made possible by the introduction of generative models like variational autoencoders (VAEs) and Wasserstein GANs.

These are only a portion of the eminent headways and achievements in AI during the predefined period. The field kept on advancing quickly past 2017, with new leap forwards, strategies, and applications arising.

Machine Learning at present:

The field of machine learning has made significant strides in recent years, and its applications are numerous, including self-driving cars, Amazon Alexa, Catboats, and the recommender system. It incorporates clustering, classification, decision tree, SVM algorithms, and reinforcement learning, as well as unsupervised and supervised learning.

Present day AI models can be utilized for making different expectations, including climate expectation, sickness forecast, financial exchange examination, and so on.

Machine Learning Approach:

1. Artificial Neural Network (ANN):
   - Artificial Neural Networks (ANNs) are computational models inspired by the structure and functioning of biological neural networks, such as the human brain.
   - ANNs consist of interconnected nodes arranged in layers, including an input layer, one or more hidden layers, and an output layer.
   - Each connection between nodes is associated with a weight that adjusts during the learning process.
   - ANNs are trained using algorithms like backpropagation, which adjusts the weights to minimize the difference between predicted and actual outputs.
   - They are versatile and widely used for tasks such as classification, regression, pattern recognition, and time series prediction.

2. Clustering:
   - Clustering algorithms group similar data points together based on their characteristics or features, without requiring predefined categories.
   - The goal of clustering is to identify natural groupings or clusters within the data.
   - Popular clustering algorithms include K-means, which partitions the data into K clusters based on centroids, and hierarchical clustering, which creates a tree of clusters by recursively merging or splitting them.
   - Clustering is applied in various domains, including customer segmentation, anomaly detection, image segmentation, and document categorization.

3. Reinforcement Learning:

- Reinforcement Learning (RL) is a type of machine learning where an agent learns to make decisions by interacting with an environment.
- The agent takes actions to maximize cumulative rewards over time while exploring different strategies.
- RL involves learning from feedback received in the form of rewards or penalties.
- Popular RL algorithms include Q-learning, Deep Q-Networks (DQN), and policy gradient methods.
- RL finds applications in autonomous systems, game playing, robotics, recommendation systems, and optimizing resource allocation.

4. Decision Tree Learning:

- Decision tree learning constructs a tree-like model of decisions based on features of the data.
- Each internal node of the tree represents a decision based on a feature, and each leaf node represents the outcome or prediction.
- Decision trees are built recursively by selecting the best feature to split the data at each node.
- Popular algorithms include ID3 (Iterative Dichotomiser 3), C4.5, and CART (Classification and Regression Trees).
- Decision trees are interpretable and suitable for both classification and regression tasks, with applications in medicine, finance, and customer relationship management.

5. Bayesian Networks:

- Bayesian networks, also known as belief networks or probabilistic graphical models, represent probabilistic relationships among variables using a directed acyclic graph.
- Nodes in the graph represent random variables, and edges represent dependencies between them.
- Bayesian networks encode conditional probability distributions, allowing inference and reasoning under uncertainty.
- They are employed in decision support systems, medical diagnosis, risk assessment, and anomaly detection.

6. Support Vector Machine (SVM):

- Support Vector Machine is a supervised learning algorithm used for classification and regression tasks.
- SVM aims to find the hyperplane that best separates data points into different classes while maximizing the margin between the classes.
- It works well in high-dimensional spaces and is effective even when the number of dimensions exceeds the number of samples.
- SVM can handle linear and nonlinear classification tasks using different kernel functions such as linear, polynomial, and radial basis function (RBF) kernels.
- Applications of SVM include text categorization, image classification, and bioinformatics.

7. Genetic Algorithm:

- Genetic algorithms (GAs) are optimization algorithms inspired by the principles of natural selection and genetics.
- GAs iteratively search for optimal solutions by evolving a population of candidate solutions through processes like selection, crossover, and mutation.
- Solutions are represented as chromosomes or strings of parameters that undergo genetic operations to produce offspring.
- Genetic algorithms are applicable to a wide range of optimization problems, including function optimization, scheduling, and parameter tuning in machine learning algorithms.

Machine Learning (ML) V/S **Deep Learning (DL)**

Aspect	**Machine Learning (ML)**	**Deep Learning (DL)**
Definition	Subfield of AI that develops algorithms and models based on data.	Subset of ML that focuses on neural networks with multiple layers.
Learning Approach	Learns from data without explicit programming.	Learns hierarchical features from raw data.
Data Representation	Structured data (features with labels).	Raw data (e.g., images, audio).
Complexity	Moderate complexity.	High complexity due to deep architectures.
Applications	Predictive modeling, recommendation systems, fraud detection.	Image recognition, natural language processing, speech synthesis.
Key Component	Algorithms (e.g., linear regression, decision trees).	Neural networks (e.g., CNNs, RNNs).

Aspect	Machine Learning (ML)	Deep Learning (DL)
Training Data Size	Works well with small to medium-sized datasets.	Requires large datasets for effective training.
Interpretability	Models are interpretable (e.g., feature importance).	Less interpretable due to complex architectures.
Hardware Requirements	Can run on standard CPUs.	Requires GPUs or specialized hardware for training.

data science V/S **machine learning** in a tabular format:

Aspect	Data Science	Machine Learning
Definition	A multidisciplinary field that extracts value from large data sets, creating meaning and insights.	A subset of artificial intelligence (AI) that focuses on learning from data to make predictions.
Focus	Extracting insights from data using tools like statistics, data analytics, and data modeling.	Creating algorithms that learn from data and improve performance or inform predictions.
Components	Includes mining, statistics, data analytics, modeling, and programming.	Requires data cleaning, preparation, and analysis before learning from the data.
Problem Solving	Defines business problems; machine learning techniques help solve them.	Learns from data to create insights and improve performance.

Aspect	Data Science	Machine Learning
Human Intervention	Requires human understanding of the problem and data analysis.	Minimal human intervention; algorithms learn from data through experience.
Data Processing	Processes raw data to create meaning and insights.	Learns from data to make predictions.
Challenges	Identifying pertinent business issues, communicating results, data security, KPI metrics.	Cleaning and preparing data, collaboration between data scientists and engineers, KPI determination.
Evolution	Emerged with the increase in data from various sources.	Part of the broader field of data science, evolving as more data is processed.

key differences between **machine learning (ML)** and **traditional programming**:

Aspect	Machine Learning	Traditional Programming
Definition	A branch of artificial intelligence that develops algorithms by learning from data to make predictions on new similar type data, without being explicitly programmed for each task.	The process of providing explicit instructions to a computer using predefined rules and inputs to perform a specific task.

Aspect	Machine Learning	Traditional Programming
Approach	**Data-Driven**: Learns patterns from data and makes predictions based on those patterns.	**Rule-Based**: Manually defines rules for the computer to follow.
Flexibility	Handles complex tasks (e.g., natural language processing, image recognition) where explicit rules may not suffice.	Effective for tasks with well-defined rules and inputs (e.g., sorting algorithms, mathematical computations).
Human Intervention	Minimal human intervention; models learn from data through iterative training.	Requires explicit human understanding and coding of rules.
Examples	Recommender systems, natural language processing, image recognition, fraud detection.	Mathematical computations, sorting algorithms, simple rule-based tasks.

Resource

TEXT BOOKS:

1. Ethem Alpaydin, "Introduction to Machine Learning", MIT Press, Prentice Hall of India, 3rd Edition2014.
2. Mehryar Mohri, Afshin Rostamizadeh, Ameet Talwalkar " Foundations of Machine Learning", MIT Press,2012.
3. Tom Mitchell, "Machine Learning", McGraw Hill, 3rdEdition, 1997.
4. MACHINE LEARNING - An Algorithmic Perspective, Second Edition, Stephen Marsland, 2015.

REFERENCE BOOKS:

1. CharuC.Aggarwal,"DataClassificationAlgorithmsandApplications",CRCPress,2014.

2. Charu C. Aggarwal, "DATA CLUSTERING Algorithms and Applications", CRC Press, 2014.
3. Kevin P. Murphy "Machine Learning: A Probabilistic Perspective", The MIT Press, 2012
4. Jiawei Han and Micheline Kambers and JianPei, "Data Mining Concepts andTechniques",3rd edition, Morgan Kaufman Publications, 2012.