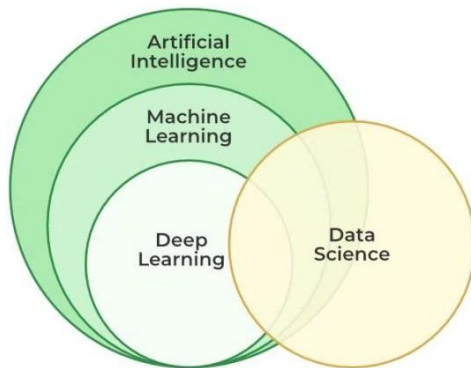


UNIT - I

Machine Learning

Machine learning is a branch of artificial intelligence that enables algorithms to uncover hidden patterns within datasets, allowing them to make predictions on new, similar data without explicit programming for each task. Traditional machine learning combines data with statistical tools to predict outputs, yielding actionable insights. This technology finds applications in diverse fields such as image and speech recognition, natural language processing, recommendation systems, fraud detection, portfolio optimization, and automating tasks.

For instance, recommender systems use historical data to personalize suggestions. Netflix, for example, employs collaborative and content-based filtering to recommend movies and TV shows based on user viewing history, ratings, and genre preferences. Reinforcement learning further enhances these systems by enabling agents to make decisions based on environmental feedback, continually refining recommendations.



Difference between Machine Learning and Traditional Programming

The Difference between Machine Learning and Traditional Programming is as follows:

Machine Learning	Traditional Programming	Artificial Intelligence
Machine Learning is a subset of artificial intelligence(AI) that focus on learning from data to develop an algorithm that can be used to make a prediction.	In traditional programming, rule-based code is written by the developers depending on the problem statements.	Artificial Intelligence involves making the machine as much capable, So that it can perform the tasks that typically require human intelligence.
Machine Learning uses a	Traditional programming	AI can involve many different

Machine Learning	Traditional Programming	Artificial Intelligence
data-driven approach, It is typically trained on historical data and then used to make predictions on new data.	is typically rule-based and deterministic. It hasn't self-learning features like Machine Learning and AI.	techniques, including Machine Learning and Deep Learning, as well as traditional rule-based programming.
ML can find patterns and insights in large datasets that might be difficult for humans to discover.	Traditional programming is totally dependent on the intelligence of developers. So, it has very limited capability.	Sometimes AI uses a combination of both Data and Pre-defined rules, which gives it a great edge in solving complex tasks with good accuracy which seem impossible to humans.
Machine Learning is the subset of AI. And Now it is used in various AI-based tasks like Chatbot Question answering, self-driven car., etc.	Traditional programming is often used to build applications and software systems that have specific functionality.	AI is a broad field that includes many different applications, including natural language processing, computer vision, and robotics.

Machine Learning works in the following manner.

A machine learning algorithm works by learning patterns and relationships from data to make predictions or decisions without being explicitly programmed for each task. Here's a simplified overview of how a typical machine learning algorithm works:

1. Data Collection:

First, relevant data is collected or curated. This data could include examples, features, or attributes that are important for the task at hand, such as images, text, numerical data, etc.

2. Data Preprocessing:

Before feeding the data into the algorithm, it often needs to be preprocessed. This step may involve cleaning the data (handling missing values, outliers), transforming the data (normalization, scaling), and splitting it into training and test sets.

3. Choosing a Model:

Depending on the task (e.g., classification, regression, clustering), a suitable machine learning model is chosen. Examples include decision trees, neural networks, support vector machines, and more advanced models like deep learning architectures.

4. Training the Model:

The selected model is trained using the training data. During training, the algorithm learns patterns and relationships in the data. This involves adjusting model parameters iteratively to minimize the difference between predicted outputs and actual outputs (labels or targets) in the training data.

5. Evaluating the Model:

Once trained, the model is evaluated using the test data to assess its performance. Metrics such as accuracy, precision, recall, or mean squared error are used to evaluate how well the model generalizes to new, unseen data.

6. Fine-tuning:

Models may be fine-tuned by adjusting hyperparameters (parameters that are not directly learned during training, like learning rate or number of hidden layers in a neural network) to improve performance.

7. Prediction or Inference:

Finally, the trained model is used to make predictions or decisions on new data. This process involves applying the learned patterns to new inputs to generate outputs, such as class labels in classification tasks or numerical values in regression tasks.

Machine Learning lifecycle:

The lifecycle of a machine learning project involves a series of steps that include:

1. Study the Problems:

The first step is to study the problem. This step involves understanding the business problem and defining the objectives of the model.

2. Data Collection:

When the problem is well-defined, we can collect the relevant data required for the model. The data could come from various sources such as databases, APIs, or web scraping.

3. Data Preparation:

When our problem-related data is collected, then it is a good idea to check the data properly and make it in the desired format so that it can be used by the model to find the hidden patterns. This can be done in the following steps:

- Data cleaning
- Data Transformation
- Explanatory Data Analysis and Feature Engineering
- Split the dataset for training and testing.

4. Model Selection:

The next step is to select the appropriate machine learning algorithm that is suitable for our problem. This step requires knowledge of the strengths and weaknesses of different algorithms. Sometimes we use multiple models and compare their results and select the best model as per our requirements.

5. Model building and Training:

- After selecting the algorithm, we have to build the model.
- In the case of traditional machine learning building mode is easy it is just a few hyperparameter tunings.
- In the case of deep learning, we have to define layer-wise architecture along with input and output size, number of nodes in each layer, loss function, gradient descent optimizer, etc.
- After that model is trained using the preprocessed dataset.

6. Model Evaluation:

Once the model is trained, it can be evaluated on the test dataset to determine its accuracy and performance using different techniques. like classification report, F1 score, precision, recall, ROC Curve, Mean Square error, absolute error, etc.

7. Model Tuning:

Based on the evaluation results, the model may need to be tuned or optimized to improve its performance. This involves tweaking the hyperparameters of the model.

8. Deployment:

Once the model is trained and tuned, it can be deployed in a production environment to make predictions on new data. This step requires integrating the model into an existing software system or creating a new system for the model.

9. Monitoring and Maintenance:

Finally, it is essential to monitor the model's performance in the production environment and perform maintenance tasks as required. This involves monitoring for data drift, retraining the model as needed, and updating the model as new data becomes available.

Supervised Machine Learning:

- Supervised learning is a type of machine learning in which the algorithm is trained on the labeled dataset. It learns to map input features to targets based on labeled training data. In supervised learning, the algorithm is provided with input features and corresponding output labels, and it learns to generalize from this data to make predictions on new, unseen data.
- There are two main types of supervised learning:
- Regression: Regression is a type of supervised learning where the algorithm learns to predict continuous values based on input features. The output labels in regression are continuous values, such as stock prices, and housing prices. The different regression algorithms in machine learning are: Linear Regression, Polynomial Regression, Ridge Regression, Decision Tree Regression, Random Forest Regression, Support Vector Regression, etc
- Classification: Classification is a type of supervised learning where the algorithm learns to assign input data to a specific category or class based on input features. The output labels in classification are discrete values. Classification algorithms can be binary, where the output is one of two possible classes, or multiclass, where the output can be one of several classes. The different Classification algorithms in machine learning are: Logistic Regression, Naive Bayes, Decision Tree

Unsupervised Machine Learning:

- Unsupervised learning is a type of machine learning where the algorithm learns to recognize patterns in data without being explicitly trained using labeled examples. The goal of unsupervised learning is to discover the underlying structure or distribution in the data.
- There are two main types of unsupervised learning:
- Clustering: Clustering algorithms group similar data points together based on their characteristics. The goal is to identify groups, or clusters, of data points that are similar to each other, while being distinct from other groups. Some popular clustering algorithms include K-means, Hierarchical clustering, and DBSCAN.
- Dimensionality reduction: Dimensionality reduction algorithms reduce the number of input variables in a dataset while preserving as much of the original information as possible. This is

useful for reducing the complexity of a dataset and making it easier to visualize and analyze. Some popular dimensionality reduction algorithms include Principal Component Analysis (PCA), t-SNE, and Autoencoders

Reinforcement Machine Learning

- Reinforcement learning is a type of machine learning where an agent learns to interact with an environment by performing actions and receiving rewards or penalties based on its actions. The goal of reinforcement learning is to learn a policy, which is a mapping from states to actions, that maximizes the expected cumulative reward over time.
- There are two main types of reinforcement learning:
- Model-based reinforcement learning: In model-based reinforcement learning, the agent learns a model of the environment, including the transition probabilities between states and the rewards associated with each state-action pair. The agent then uses this model to plan its actions in order to maximize its expected reward. Some popular model-based reinforcement learning algorithms include Value Iteration and Policy Iteration.
- Model-free reinforcement learning: In model-free reinforcement learning, the agent learns a policy directly from experience without explicitly building a model of the environment. The agent interacts with the environment and updates its policy based on the rewards it receives. Some popular model-free reinforcement learning algorithms include Q-Learning, SARSA, and Deep Reinforcement Learning.

Applications of Machine Learning

- Now in this Machine learning tutorial, let's learn the applications of Machine Learning:
- Automation: Machine learning, which works entirely autonomously in any field without the need for any human intervention. For example, robots perform the essential process steps in manufacturing plants.
- Finance Industry: Machine learning is growing in popularity in the finance industry. Banks are mainly using ML to find patterns inside the data but also to prevent fraud.
- Government organization: The government makes use of ML to manage public safety and utilities. Take the example of China with its massive face recognition. The government uses Artificial intelligence to prevent jaywalking.
- Healthcare industry: Healthcare was one of the first industries to use machine learning with image detection.
- Marketing: Broad use of AI is done in marketing thanks to abundant access to data. Before the age of mass data, researchers develop advanced mathematical tools like Bayesian analysis to estimate the value of a customer. With the boom of data, the marketing department relies on AI to optimize customer relationships and marketing campaigns.
- Retail industry: Machine learning is used in the retail industry to analyze customer behavior, predict demand, and manage inventory. It also helps retailers to personalize the shopping experience for each customer by recommending products based on their past purchases and preferences.
- Transportation: Machine learning is used in the transportation industry to optimize routes, reduce fuel consumption, and improve the overall efficiency of transportation systems. It also plays a role in autonomous vehicles, where ML algorithms are used to make decisions about navigation and safety.