

MACHINE LEARNING

Machine learning is a field of study where science and technology are used together to develop various algorithms and techniques to enable computers to learn in a similar way to humans. First designed in 1952 by its father, Arthur Samuel, machine learning was developed in IBM's Poughkeepsie laboratory to play checkers. Arthur Samuel aimed for this program to improve itself every time it played checkers, identify and correct its errors, and discover stronger ways to win the game with the data it obtained. This auto-learning program was one of the first examples of machine learning.

Variable Types

- **Numerical Variable:** These are variables expressed with numerical values. They are divided into two as discrete or continuous.
- **Categorical Variables:** These are variables that express a class, that is, a category. They are divided into two as nominal and ordinal. Nominal categorical variables are variables such as Gender, Education Status, where there is no difference between classes. Ordinal categorical variables are variables such as education status, where there are differences between classes.

Learning Types

Machine learning is a subset of AI, which enables the machine to automatically learn from data, improve performance from past experiences, and make predictions. Machine learning contains a set of algorithms that work on a huge amount of data. Data is fed to these algorithms to train them, and on the basis of training, they build the model & perform a specific task

Based on the methods and way of learning, machine learning is divided into mainly four types, which are:

- Supervised Machine Learning
- Unsupervised Machine Learning
- Semi-Supervised Machine Learning
- Reinforcement Learning

Supervised Machine Learning: If there are Target variables and Dependent variables in the data set, this is called supervised learning. The main goal of the supervised learning technique is to map the input variable(x) with the output variable(y). Some real-world applications of supervised learning are Risk Assessment, Fraud Detection, Spam filtering.

Supervised machine learning can be classified into two types of problems, which are given below:

- Classification
- Regression

a) Classification

Classification algorithms are used to solve the classification problems in which the output variable is categorical, such as "Yes" or No, Male or Female, Red or Blue, etc. The classification algorithms predict the categories present in the dataset. Some real-world examples of classification algorithms are Spam Detection, Email filtering, etc.

Some popular classification algorithms are given below:

- Random Forest Algorithm
- Decision Tree Algorithm
- Logistic Regression Algorithm
- Support Vector Machine Algorithm

b) Regression

Regression algorithms are used to solve regression problems in which there is a linear relationship between input and output variables. These are used to predict continuous output variables, such as market trends, weather prediction, etc.

Some popular Regression algorithms are given below:

- Simple Linear Regression Algorithm
- Multivariate Regression Algorithm
- Decision Tree Algorithm
- Lasso Regression

Applications of Supervised Learning

Some common applications of Supervised Learning are given below:

Image Segmentation:

Supervised Learning algorithms are used in image segmentation. In this process, image classification is performed on different image data with pre-defined labels.

Medical Diagnosis:

Supervised algorithms are also used in the medical field for diagnosis purposes. It is done by using medical images and past labelled data with labels for disease conditions. With such a process, the machine can identify a disease for the new patients.

Fraud Detection - Supervised Learning classification algorithms are used for identifying fraud transactions, fraud customers, etc. It is done by using historic data to identify the patterns that can lead to possible fraud.

Spam detection - In spam detection & filtering, classification algorithms are used. These algorithms classify an email as spam or not spam. The spam emails are sent to the spam folder.

Speech Recognition - Supervised learning algorithms are also used in speech recognition. The algorithm is trained with voice data, and various identifications can be done using the same, such as voice-activated passwords, voice commands, etc.

Unsupervised Machine Learning: If there is no target variable/dependent variable in the data set, this situation is called an unsupervised learning problem.

Unsupervised Learning can be further classified into two types, which are given below:

- Clustering
- Association

1) Clustering

The clustering technique is used when we want to find the inherent groups from the data. It is a way to group the objects into a cluster such that the objects with the most similarities remain in one group and have fewer or no similarities with the objects of other groups. An example of the clustering algorithm is grouping the customers by their purchasing behaviour.

Some of the popular clustering algorithms are given below:

K-Means Clustering algorithm

Mean-shift algorithm

DBSCAN Algorithm

Principal Component Analysis

Independent Component Analysis

2) Association

Association rule learning is an unsupervised learning technique, which finds interesting relations among variables within a large dataset. The main aim of this learning algorithm is to find the dependency of one data item on another data item and map those variables accordingly so that it can generate maximum profit. This algorithm is mainly applied in Market Basket analysis, Web usage mining, continuous production, etc.

Some popular algorithms of Association rule learning are Apriori Algorithm, Eclat, FP-growth algorithm.

Applications of Unsupervised Learning

Network Analysis: Unsupervised learning is used for identifying plagiarism and copyright in document network analysis of text data for scholarly articles.

Recommendation Systems: Recommendation systems widely use unsupervised learning techniques for building recommendation applications for different web applications and e-commerce websites.

Anomaly Detection: Anomaly detection is a popular application of unsupervised learning, which can identify unusual data points within the dataset. It is used to discover fraudulent transactions.

Singular Value Decomposition: Singular Value Decomposition or SVD is used to extract particular information from the database. For example, extracting information of each user located at a particular location.

Semi-Supervised Machine Learning: Semi-Supervised learning is a type of Machine Learning algorithm that lies between Supervised and Unsupervised machine learning. It represents the intermediate ground between Supervised and Unsupervised learning algorithms and uses the combination of labelled and unlabeled datasets during the training period.

Reinforcement Learning: Models that learn by trial and error, by learning from wrong moves, are called reinforcement learning.

Reinforcement learning is categorized mainly into two types of methods/algorithms:

- **Positive Reinforcement Learning:** Positive reinforcement learning specifies increasing the tendency that the required behaviour would occur again by adding something. It enhances the strength of the behaviour of the agent and positively impacts it.
- **Negative Reinforcement Learning:** Negative reinforcement learning works exactly opposite to the positive RL. It increases the tendency that the specific behaviour would occur again by avoiding the negative condition.

Real-world Use cases of Reinforcement Learning

Video Games:

RL algorithms are much popular in gaming applications. It is used to gain super-human performance. Some popular games that use RL algorithms are AlphaGO and AlphaGO Zero.

Resource Management:

The "Resource Management with Deep Reinforcement Learning" paper showed that how to use RL in computer to automatically learn and schedule resources to wait for different jobs in order to minimize average job slowdown.

Robotics:

RL is widely being used in Robotics applications. Robots are used in the industrial and manufacturing area, and these robots are made more powerful with reinforcement learning. There are different industries that have their vision of building intelligent robots using AI and Machine learning technology.

Text Mining

Text-mining, one of the great applications of NLP, is now being implemented with the help of Reinforcement Learning by Salesforce company.