**MACHINE LEARNING AND ITS TYPES**

**Machine learning** is a subfield of artificial intelligence that focuses on the development of algorithms and statistical models that enable computer systems to **learn patterns and make decisions or predictions** from data **without being explicitly programmed** for each specific task.

Thinking about our favourite streaming service or food delivery app, it suggests indie film we end up loving or knows we’re craving sushi on a Friday night—that’s machine learning at work. For example, Netflix combines different machine learning types in their recommendation system to [influence more than 80%](https://www.renascence.io/journal/how-netflix-uses-data-to-drive-hyper-personalized-customer-experience-cx#:~:text=According%20to%20Netflix%2C%20over%2080,extensively%20to%20optimize%20its%20platform.) of what their subscribers watch.

Building machine learning applications is **easier** said than done, though. For starters, simply choosing the right type of machine learning can be a roadblock. It’s not just about picking the most advanced or popular approach.



[**How does machine learning work?**](https://www.digitalocean.com/resources/articles/types-of-machine-learning#how-does-machine-learning-work)

**How does machine learning work?**

Machine learning is like **teaching your software** to get smarter over time. Instead of giving it a fixed set of rules, you let it learn from examples—just like people do.

Imagine you're building an app to catch fraudulent transactions. A traditional method would mean writing rules like:

* “Flag anything over $1,000”
* “Watch out for purchases from unfamiliar countries”

But fraudsters are clever—they constantly change tactics. So those rules quickly become outdated.

With machine learning, your app can look **at thousands of past transactions and learn patterns that even expert fraud analysts** might miss. It might notice that genuine users tend to shop in certain ways, or that fraud often happens when specific factors—like time, location, and purchase type—line up in a suspicious way. And the best part? As new fraud patterns emerge, the system can adapt automatically, without needing you to rewrite the rules.

This approach shines in complex situations. Take **image recognition: writing rules to identify cats in photos would be nearly impossible.** But with machine learning, you just show the system thousands of cat pictures, and it figures out what makes a cat a cat—fur texture, ear shape, whiskers, and so on.

**Types of Machine Learning**

Choosing the right type of machine learning shapes how your [**machine learning operations**](https://www.digitalocean.com/resources/articles/mlops) (MLOps) develop. **Each type has characteristics** that make it ideal for specific scenarios, from working **with labeled historical data to discovering hidden patterns** or learning through trial and error.

None of these approaches is necessarily better (or worse) than another. Each has a place in real-world applications. You just need to understand the ideal use case.

[**1. Supervised learning**](https://www.digitalocean.com/resources/articles/types-of-machine-learning#1-supervised-learning)

Supervised learning is a type of machine learning where an **algorithm learns from labeled training data** to predict outputs for new, unseen inputs. The model learns the relationship between **input features** and their **corresponding output** labels to help it make predictions on new data.

The goal is for the model to learn the mapping from inputs to outputs and generalize to unseen data. The **quality** and **quantity** of your labeled data is what makes or breaks your supervised learning approach. The more diverse and accurate examples you provide, the better your model becomes at handling real-world scenarios.

A diagram of a dog and cat

AI-generated content may be incorrect.

Let's understand it with the help of an example.

**Example:**Consider a scenario where you have to build an image classifier to differentiate between cats and dogs. If you feed the datasets of dogs and cats labelled images to the algorithm, the machine will learn to classify between a dog or a cat from these labeled images. When we input new dog or cat images that it has never seen before, it will use the learned algorithms and predict whether it is a dog or a cat. This is how **supervised learning** works, and this is particularly an image classification.

Supervised learning tasks typically fall into two categories:

1. **Classification:** Classification models predict discrete categories or labels. For example, a fraud detection system classifies transactions as either legitimate or fraudulent. Other common applications include spam detection, medical diagnosis, and image recognition.

Here are some classification algorithms:

* [**Logistic Regression**](https://www.geeksforgeeks.org/machine-learning/understanding-logistic-regression/)
* [**Support Vector Machine**](https://www.geeksforgeeks.org/machine-learning/support-vector-machine-algorithm/)
* [**Random Forest**](https://www.geeksforgeeks.org/machine-learning/random-forest-regression-in-python/)
* [**Decision Tree**](https://www.geeksforgeeks.org/machine-learning/decision-tree/)
* [**K-Nearest Neighbors (KNN)**](https://www.geeksforgeeks.org/machine-learning/k-nearest-neighbours/)
* [**Naive Bayes**](https://www.geeksforgeeks.org/machine-learning/naive-bayes-classifiers/)

1. **Regression:** Regression models predict continuous numerical values. A classic example is house price prediction, where the model considers factors like location, square footage, and number of bedrooms to estimate a property’s value. You’ll also find regression in stock market forecasting and demand prediction.

Here are some regression algorithms:

* [**Linear Regression**](https://www.geeksforgeeks.org/machine-learning/ml-linear-regression/)
* [**Polynomial Regression**](https://www.geeksforgeeks.org/videos/polynomial-regression-algorithm-machine-learning/)
* [**Ridge Regression**](https://www.geeksforgeeks.org/videos/lasso-ridge-regression-algorithm-machine-learning/)
* [**Lasso Regression**](https://www.geeksforgeeks.org/videos/lasso-ridge-regression-algorithm-machine-learning/)
* [**Decision tree**](https://www.geeksforgeeks.org/machine-learning/decision-tree-introduction-example/)
* [**Random Forest**](https://www.geeksforgeeks.org/machine-learning/random-forest-regression-in-python/)

A comparison of different colored dots

AI-generated content may be incorrect.

**Advantages of Supervised Machine Learning**

* **Supervised Learning** models can have high accuracy as they are trained on **labelled data**.
* The process of decision-making in supervised learning models is often interpretable.
* It can often be used in pre-trained models which saves time and resources when developing new models from scratch.

**Disadvantages of Supervised Machine Learning**

* It has limitations in knowing patterns and may struggle with unseen or unexpected patterns that are not present in the training data.
* It can be time-consuming and costly as it relies on**labeled**data only.
* It may lead to poor generalizations based on new data.

**Applications of Supervised Learning**

Supervised learning is used in a wide variety of applications, including:

* **Image classification**: Identify objects, faces, and other features in images.
* **Natural language processing:** Extract information from text, such as sentiment, entities, and relationships.
* **Speech recognition**: Convert spoken language into text.
* **Recommendation systems**: Make personalized recommendations to users.
* **Predictive analytics**: Predict outcomes, such as sales, customer churn, and stock prices.
* **Medical diagnosis**: Detect diseases and other medical conditions.
* **Fraud detection**: Identify fraudulent transactions.
* **Autonomous vehicles**: Recognize and respond to objects in the environment.
* **Email spam detection**: Classify emails as spam or not spam.
* **Quality control in manufacturing**: Inspect products for defects.
* **Credit scoring**: Assess the risk of a borrower defaulting on a loan.
* **Gaming**: Recognize characters, analyze player behavior, and create NPCs.
* **Customer support**: Automate customer support tasks.
* **Weather forecasting**: Make predictions for temperature, precipitation, and other meteorological parameters.
* **Sports analytics**: Analyze player performance, make game predictions, and optimize strategies.

**Limitations**

* Doesn’t perform well with unlabeled or ambiguous data
* Can’t discover hidden structures in data
* Limited adaptability to changing patterns unless retrained

### [**2. Unsupervised learning**](https://www.digitalocean.com/resources/articles/types-of-machine-learning#2-unsupervised-learning)

Unsupervised learning is a type of machine learning where algorithms discover hidden patterns or groupings in data **without labeled examples**. The model learns from the inherent structure of the data rather than from predefined outputs or correct answers.

Unlike supervised learning’s guided approach, unsupervised learning is more like turning your model loose to explore and discover patterns on its own. It’s great when you have data but don’t know exactly what you’re looking for, like understanding customer behavior or detecting anomalies in system performance.

A diagram of a computer program

AI-generated content may be incorrect.

Let’s consider an example

Consider a streaming service [analyzing viewing habits](https://redresscompliance.com/how-netflix-uses-ai-to-analyze-audience-viewing-habits-and-preferences/" \l ":~:text=Netflix%20uses%20machine%20learning%20algorithms,content%20tailored%20to%20individual%20preferences." \t "_blank). Instead of telling the algorithm which shows are similar, it discovers natural groupings based on viewing patterns. It might find that viewers who enjoy sci-fi documentaries also tend to watch post-apocalyptic survival series (a pattern that would almost *definitely* have been missed with predefined categories).

Unsupervised learning uses a couple of different approaches:

1. **Clustering:** Clustering algorithms group similar data points together. A retail business might use clustering to segment customers based on purchasing behavior, or a network security system might cluster traffic patterns to identify potential threats.

Here are some clustering algorithms:

* [**K-Means Clustering algorithm**](https://www.geeksforgeeks.org/machine-learning/k-means-clustering-introduction/)
* [**Mean-shift algorithm**](https://www.geeksforgeeks.org/machine-learning/ml-mean-shift-clustering/)
* [**DBSCAN Algorithm**](https://www.geeksforgeeks.org/machine-learning/dbscan-clustering-in-ml-density-based-clustering/)
* [**Principal Component Analysis**](https://www.geeksforgeeks.org/data-analysis/principal-component-analysis-pca/)
* [**Independent Component Analysis**](https://www.geeksforgeeks.org/machine-learning/ml-independent-component-analysis/)

1. **Dimensionality reduction:** This technique simplifies complex data while preserving important patterns. It’s great when dealing with high-dimensional data—like processing images or analyzing large datasets with many features.
2. **Association**

[Association rule learn](https://www.geeksforgeeks.org/machine-learning/association-rule/)ing is a technique for discovering relationships between items in a dataset. It identifies rules that indicate the presence of one item implies the presence of another item with a specific probability.

Here are some association rule learning algorithms:

* [**Apriori Algorithm**](https://www.geeksforgeeks.org/machine-learning/apriori-algorithm/)
* [**Eclat**](https://www.geeksforgeeks.org/machine-learning/ml-eclat-algorithm/)
* [**FP-growth Algorithm**](https://www.geeksforgeeks.org/machine-learning/frequent-pattern-growth-algorithm/)

A diagram of a learning process

AI-generated content may be incorrect.

**Advantages of Unsupervised Machine Learning**

* It helps to discover hidden patterns and various relationships between the data.
* Used for tasks such as**customer segmentation, anomaly detection,**and **data exploration**.
* It does not require labeled data and reduces the effort of data labeling.

**Disadvantages of Unsupervised Machine Learning**

* Without using labels, it may be difficult to predict the quality of the model's output.
* Cluster Interpretability may not be clear and may not have meaningful interpretations.
* It has techniques such as[autoencoders](https://www.geeksforgeeks.org/machine-learning/auto-encoders/) and [dimensionality reduction](https://www.geeksforgeeks.org/machine-learning/dimensionality-reduction/) that can be used to extract meaningful features from raw data.

**Applications of Unsupervised Learning**

Here are some common applications of unsupervised learning:

* **Clustering**: Group similar data points into clusters.
* **Anomaly detection**: Identify outliers or anomalies in data.
* **Dimensionality reduction**: Reduce the dimensionality of data while preserving its essential information.
* **Recommendation systems**: Suggest products, movies, or content to users based on their historical behavior or preferences.
* **Topic modeling**: Discover latent topics within a collection of documents.
* **Density estimation**: Estimate the probability density function of data.
* **Image and video compression**: Reduce the amount of storage required for multimedia content.
* **Data preprocessing**: Help with data preprocessing tasks such as data cleaning, imputation of missing values, and data scaling.
* **Market basket analysis**: Discover associations between products.
* **Genomic data analysis**: Identify patterns or group genes with similar expression profiles.
* **Image segmentation**: Segment images into meaningful regions.
* **Community detection in social networks**: Identify communities or groups of individuals with similar interests or connections.
* **Customer behavior analysis**: Uncover patterns and insights for better marketing and product recommendations.
* **Content recommendation**: Classify and tag content to make it easier to recommend similar items to users.
* **Exploratory data analysis (EDA)**: Explore data and gain insights before defining specific tasks.

**Limitations**

* May find patterns that aren’t meaningful
* Difficult to validate and explain results
* Not suitable for prediction tasks without further processing

### [3. Reinforcement learning](https://www.digitalocean.com/resources/articles/types-of-machine-learning#3-reinforcement-learning)

Reinforcement learning is a type of machine **learning where an agent learns to make sequences of decisions** by interacting with an environment. The agent receives **rewards or penalties** for its actions and learns to maximize long-term rewards through **experimentation and optimization.**

These algorithms are specific to a particular problem e.g. Google Self Driving car, AlphaGo where a bot competes with humans and even itself to get better and better performers in Go Game. Each time we feed in data, they learn and add the data to their knowledge which is training data. So, the more it learns the better it gets trained and hence experienced.

A diagram of a person's head

AI-generated content may be incorrect.

Let's understand it with the help of examples.

**Example:**Consider that you are training an [AI](https://www.geeksforgeeks.org/artificial-intelligence/what-is-artificial-intelligence-ai/) agent to play a game like chess. The agent explores different moves and receives positive or negative feedback based on the outcome. Reinforcement Learning also finds applications in which they learn to perform tasks by interacting with their surroundings.

Here are some of most common reinforcement learning algorithms:

* [**Q-learning:**](https://www.geeksforgeeks.org/machine-learning/q-learning-in-python/) Q-learning is a model-free RL algorithm that learns a Q-function, which maps states to actions. The Q-function estimates the expected reward of taking a particular action in a given state.
* [**SARSA (State-Action-Reward-State-Action):**](https://www.geeksforgeeks.org/machine-learning/sarsa-reinforcement-learning/) SARSA is another model-free RL algorithm that learns a Q-function. However, unlike Q-learning, SARSA updates the Q-function for the action that was actually taken, rather than the optimal action.
* [**Deep Q-learning**](https://www.geeksforgeeks.org/deep-learning/deep-q-learning/)**:** Deep Q-learning is a combination of Q-learning and deep learning. Deep Q-learning uses a neural network to represent the Q-function, which allows it to learn complex relationships between states and actions.

**Types of Reinforcement Machine Learning**

There are two main types of reinforcement learning:

**Positive reinforcement**

* Rewards the agent for taking a desired action.
* Encourages the agent to repeat the behavior.
* Examples: Giving a treat to a dog for sitting, providing a point in a game for a correct answer.

**Negative reinforcement**

* Removes an undesirable stimulus to encourage a desired behavior.
* Discourages the agent from repeating the behavior.
* Examples: Turning off a loud buzzer when a lever is pressed, avoiding a penalty by completing a task

A white background with black lines and dots

AI-generated content may be incorrect.

**Advantages of Reinforcement Machine Learning**

* It has autonomous decision-making that is well-suited for tasks and that can learn to make a sequence of decisions, like robotics and game-playing.
* This technique is preferred to achieve long-term results that are very difficult to achieve.
* It is used to solve a complex problems that cannot be solved by conventional techniques.

**Disadvantages of Reinforcement Machine Learning**

* Training Reinforcement Learning agents can be computationally expensive and time-consuming.
* Reinforcement learning is not preferable to solving simple problems.
* It needs a lot of data and a lot of computation, which makes it impractical and costly.

**Applications of Reinforcement Machine Learning**

Here are some applications of reinforcement learning:

* **Game Playing**: RL can teach agents to play games, even complex ones.
* **Robotics**: RL can teach robots to perform tasks autonomously.
* **Autonomous Vehicles**: RL can help self-driving cars navigate and make decisions.
* **Recommendation Systems**: RL can enhance recommendation algorithms by learning user preferences.
* **Healthcare**: RL can be used to optimize treatment plans and drug discovery.
* **Natural Language Processing (NLP)**: RL can be used in dialogue systems and chatbots.
* **Finance and Trading**: RL can be used for algorithmic trading.
* **Supply Chain and Inventory Management**: RL can be used to optimize supply chain operations.
* **Energy Management**: RL can be used to optimize energy consumption.
* **Game AI**: RL can be used to create more intelligent and adaptive NPCs in video games.
* **Adaptive Personal Assistants**: RL can be used to improve personal assistants.
* **Virtual Reality (VR) and Augmented Reality (AR):** RL can be used to create immersive and interactive experiences.
* **Industrial Control**: RL can be used to optimize industrial processes.
* **Education**: RL can be used to create adaptive learning systems.
* **Agriculture**: RL can be used to optimize agricultural operations.

**Limitations**

* Risky in real-world applications without simulation
* Not suitable for tasks without clear feedback signals
* May learn unintended behaviors if rewards are poorly designed

### [4.](https://www.digitalocean.com/resources/articles/types-of-machine-learning" \l "3-reinforcement-learning) **[Semi-Supervised Learning: Supervised + Unsupervised Learning](https://www.digitalocean.com/resources/articles/types-of-machine-learning" \l "3-reinforcement-learning)**

[Semi-Supervised learning](https://www.geeksforgeeks.org/machine-learning/ml-semi-supervised-learning/)is a machine learning algorithm that works between the supervised and unsupervised learning so it uses both **labelled and unlabelled** data. It's particularly useful when obtaining labeled data is costly, time-consuming, or resource-intensive. This approach is useful when the dataset is expensive and time-consuming. Semi-supervised learning is chosen when labeled data requires skills and relevant resources in order to train or learn from it.

We use these techniques when we are dealing with data that is a little bit labeled and the rest large portion of it is unlabeled. We can use the unsupervised techniques to predict labels and then feed these labels to supervised techniques. This technique is mostly applicable in the case of image data sets where usually all images are not labeled.

A diagram of a machine learning model

AI-generated content may be incorrect.

Let's understand it with the help of an example.

**Example**: Consider that we are building a language translation model, having labeled translations for every sentence pair can be resources intensive. It allows the models to learn from labeled and unlabeled sentence pairs, making them more accurate. This technique has led to significant improvements in the quality of machine translation services.

**Types of Semi-Supervised Learning Methods**

There are a number of different semi-supervised learning methods each with its own characteristics. Some of the most common ones include:

* **Graph-based semi-supervised learning:** This approach uses a graph to represent the relationships between the data points. The graph is then used to propagate labels from the labeled data points to the unlabeled data points.
* **Label propagation:** This approach iteratively propagates labels from the labeled data points to the unlabeled data points, based on the similarities between the data points.
* **Co-training:** This approach trains two different machine learning models on different subsets of the unlabeled data. The two models are then used to label each other's predictions.
* **Self-training:** This approach trains a machine learning model on the labeled data and then uses the model to predict labels for the unlabeled data. The model is then retrained on the labeled data and the predicted labels for the unlabeled data.
* [**Generative adversarial networks (GANs)**](https://www.geeksforgeeks.org/deep-learning/generative-adversarial-network-gan/)**:** GANs are a type of deep learning algorithm that can be used to generate synthetic data. GANs can be used to generate unlabeled data for semi-supervised learning by training two neural networks, a generator and a discriminator.

A diagram of a few types of training techniques

AI-generated content may be incorrect.

**Advantages of Semi- Supervised Machine Learning**

* It leads to better generalization as compared to **supervised learning,** as it takes both labeled and unlabeled data.
* Can be applied to a wide range of data.

**Disadvantages of Semi- Supervised Machine Learning**

* **Semi-supervised**methods can be more complex to implement compared to other approaches.
* It still requires some **labeled data** that might not always be available or easy to obtain.
* The unlabeled data can impact the model performance accordingly.

**Applications of Semi-Supervised Learning**

Here are some common applications of semi-supervised learning:

* **Image Classification and Object Recognition**: Improve the accuracy of models by combining a small set of labeled images with a larger set of unlabeled images.
* **Natural Language Processing (NLP)**: Enhance the performance of language models and classifiers by combining a small set of labeled text data with a vast amount of unlabeled text.
* **Speech Recognition:** Improve the accuracy of speech recognition by leveraging a limited amount of transcribed speech data and a more extensive set of unlabeled audio.
* **Recommendation Systems**: Improve the accuracy of personalized recommendations by supplementing a sparse set of user-item interactions (labeled data) with a wealth of unlabeled user behavior data.
* **Healthcare and Medical Imaging**: Enhance medical image analysis by utilizing a small set of labeled medical images alongside a larger set of unlabeled images

**Limitations**

* Not effective if labeled data is too sparse or noisy
* May propagate errors from mislabeled data
* Requires careful model design to balance learning