

Supervised learning

This machine learning approach leverages labeled datasets, specifically crafted to instruct algorithms in the accurate classification or prediction of data. By utilizing labeled inputs and outputs, the model can continually refine its accuracy and acquire knowledge over time, encompassing both classification and regression tasks.

- **Classification** problems assign test data into specific categories. Types examples: Linear classifiers, support vector machines, decision trees and random forest.
- **Regression** tasks understand the relationship between dependent and independent variables. Regression models are helpful for predicting numerical values based on different data points. Some popular regression algorithms are linear regression, logistic regression and polynomial regression.

Unsupervised learning

It uses machine learning algorithms to analyze and cluster unlabeled data sets. These algorithms discover hidden patterns in data without the need for human intervention (hence, they are “unsupervised”).

Unsupervised learning models are used for three main tasks: clustering, association and dimensionality reduction.

Differences

The main distinction between the two approaches is the use of labeled datasets. To put it simply, supervised learning uses labeled input and output data, while an unsupervised learning algorithm does not.

In supervised learning, the algorithm “learns” from the training dataset by iteratively making predictions on the data and adjusting for the correct answer. While supervised learning models tend to be more accurate than unsupervised learning models, they require upfront human intervention to label the data appropriately.

Probabilistic Models

Probabilistic models are an essential component of machine learning, which aims to learn patterns from data and make predictions on new, unseen data. They are statistical models that capture the inherent uncertainty in data and incorporate it into their predictions.

A probabilistic model in machine learning is a mathematical representation of a real-world process that incorporates uncertain or random variables. The goal of probabilistic modeling is to estimate the probabilities of the possible outcomes of a system based on data or prior knowledge. These models identify uncertain relationships between variables in a data-driven manner while capturing the underlying trends or patterns in data³. Probabilistic models are an essential component of machine learning, which aims to learn patterns from data and make predictions on new, unseen data.

References:

- <https://www.geeksforgeeks.org/probabilistic-models-in-machine-learning/>
- <https://www.ibm.com/blog/supervised-vs-unsupervised-learning/>