

Machine Learning

FOUNDATIONS, KEY CONCEPTS AND LEARNING PARADIGMS



DEFINITION

Machine Learning (ML) is a subfield of Artificial Intelligence (AI) in which computer systems learn from data and use the discovered patterns to make predictions/decisions on new, unseen data. During training, a model analyzes the given examples and adjusts its internal parameters to improve the own performance on a specific task, rather than being explicitly programmed with rules.

TERMINOLOGY

- ➔ **Artificial Intelligence (AI):** General term for systems that perform tasks associated with human intelligence, such as reasoning, problem-solving and language processing.
- ➔ **Machine Learning (ML):** A subfield of AI in which systems learn patterns from data and improve their performance through experience.
- ➔ **Deep Learning (DL):** A subfield of ML based on multi-layer neural networks, particularly effective for processing images, speech and text.

CORE IDEA

- ➔ **Data:** Raw examples such as images, text or numbers.
- ➔ **Features:** Relevant attributes extracted or selected from the data that the model uses as input signals.
- ➔ **Model:** A mathematical function with learnable parameters (weights) that maps inputs to outputs.
- ➔ **Loss:** An error signal that measures how incorrect the model's predictions are. It guides the learning process.
- ➔ **Optimization:** An algorithm (e.g. gradient descent) that iteratively adjusts weights to minimize the loss.

KEY METRICS FOR EVALUATION

- ➔ **Accuracy:** Percentage of correct predictions overall.
- ➔ **Precision:** Proportion of correct positive predictions out of all predicted positives → measure of prediction quality.
- ➔ **Recall:** Proportion of correctly identified positives out of all actual positives → measure of prediction completeness.
- ➔ **F1-Score:** Balanced measure combining precision and recall.
- ➔ **Loss Function:** Error measure used during training, such as MSE, MAE or cross-entropy loss.
- ➔ **Confusion Matrix:** Table of all prediction outcomes (true positives, true negatives, false positives and false negatives).

SUPERVISED LEARNING

The model learns a mapping from input data to known output labels. Each training example comes with the correct target labels provided in the dataset and the model then adjusts itself to minimize the error between its predictions and the true labels. Once trained, it generalizes this mapping to unseen data.

This approach is widely used for classification and prediction tasks.

UNSUPERVISED LEARNING

The model discovers patterns and underlying structure in unlabeled data. No target labels are provided, so the system must organize the data on its own based on similarities, differences and statistical relationships within the data.

This approach is useful for clustering, dimensionality reduction, anomaly detection and as a preprocessing step for downstream supervised tasks.

REINFORCEMENT LEARNING

An agent interacts with an environment by taking actions and receiving rewards or penalties based on the outcomes. The goal is to learn a strategy that maximizes long-term cumulative reward by balancing exploration (trying new actions) and exploitation (using known successful actions).

This approach is inspired by behavioral psychology and used in robotics and autonomous decision-making.

BIAS / VARIANCE

Underfitting (high bias): The model is too simple to capture the structure in the data.

Overfitting (high variance): The model is too complex and memorizes training data instead of generalizing.

Goal: Find the sweet spot!

TRAIN / VAL / TEST

The data is typically divided into three sets: The training set (~70%) to fit the model, the validation set (~15%) to tune hyperparameters and select the best model and the test set (~15%) as unbiased final performance estimate.

Note: Never train & validate on test data!

HYPERPARAMETERS

These settings (e.g. learning rate, number of layers) are chosen before training and determine how the model learns. Unlike model weights, they are not learned from the data but selected through systematic search methods such as grid/random search or Bayesian optimization.

NEURAL NETWORKS

Inspired by the brain, neural networks consist of layers of connected nodes (neurons). Each connection carries a learnable weight and each neuron applies a non-linear activation function.

Deep networks can learn complex patterns directly from raw data.

REFERENCES

- [1] C.M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2009.
- [2] M. P. Deisenroth, A. A. Faisal, and C. S. Ong. *Mathematics for Machine Learning*. Cambridge University Press, 1st edition, 2020.