Opinion Article

An Overview of Machine Learning Algorithms: Techniques, Applications and Key Considerations

Ashrey Jue*

Department of Computer Science and Technology, University of Maryland, Maryland, United States of America

DESCRIPTION

Machine learning (ML) is a subset of Artificial Intelligence (AI) that focuses on developing systems capable of learning from data and improving their performance over time without explicit programming. Central to this process are algorithms that enable machines to recognize patterns, make predictions and automate decision-making tasks. This article provides an overview of the most commonly used ML algorithms, categorized by their learning approach and application.

Supervised learning

In supervised learning, algorithms are trained on labeled datasets, where the input-output relationship is predefined. The goal is to learn a mapping function that predicts outputs for new, unseen inputs.

Linear regression: Linear regression is used to predict continuous values by finding the relationship between dependent and independent variables. It works by fitting a line (or hyperplane) to the data that minimizes the difference between predicted and actual values.

Logistic regression: Despite its name, logistic regression is used for classification problems. It predicts probabilities that an input belongs to a particular class, often applied in binary classification tasks such as spam detection.

Support Vector Machines (SVM): SVM separates data points into different classes using a hyperplane. It is effective in high-dimensional spaces and is commonly used in text classification and image recognition.

Unsupervised learning

Unsupervised learning algorithms work with unlabeled data, aiming to explore hidden patterns or structures within the dataset.

Hierarchical clustering: This technique builds a hierarchy of clusters, often represented as a tree or dendrogram. It can reveal relationships between clusters at different levels of granularity.

Principal Component Analysis (PCA): PCA is a dimensionality reduction algorithm that transforms data into a smaller set of variables while retaining most of its variance. It is frequently used for preprocessing and visualization.

Autoencoders: Autoencoders are neural networks designed to learn efficient data representations by encoding and decoding the input data. They are often applied in anomaly detection and image denoising.

Reinforcement learning

Reinforcement learning focuses on training agents to make sequential decisions by interacting with an environment. The agent learns through trial and error, receiving feedback in the form of rewards or penalties.

Q-learning: Q-learning is a value-based algorithm where the agent learns a Q-value for each action-state pair, enabling it to select actions that maximize cumulative rewards.

Deep Q-Networks (DQN): DQN combines reinforcement learning with deep learning, allowing agents to handle complex environments with large state spaces, such as video games and robotic control.

Policy gradient methods: These methods optimize policies directly rather than estimating value functions, making them suitable for environments with continuous action spaces.

Choosing the right algorithm

Selecting the appropriate algorithm depends on the nature of the data, the problem's requirements and computational constraints. Factors to consider include,

Type of problem: Determine whether the problem is regression, classification, clustering or sequential decision-making.

Data size and quality: Some algorithms, like neural networks,

---

Correspondence to: Ashrey Jue, Department of Computer Science and Technology, University of Maryland, Maryland, United States of America, E-mail: jue_ash@hotmail.com

---

perform better with large datasets, while others, like decision trees, work well with smaller datasets.

Computational efficiency: Complex algorithms may require more computational power and time, which could be a limitation in resource-constrained environments.

Applications of ML algorithms

Healthcare: Predicting patient outcomes, diagnosing diseases and personalizing treatments using models like logistic regression and neural networks.

Finance: Fraud detection, credit scoring and algorithmic trading SVM and random forests.

E-commerce: Recommending products and analyzing customer behavior using clustering and neural networks.

Manufacturing: Predictive maintenance and quality control with reinforcement learning and decision trees.

Autonomous vehicles: Navigating environments using deep reinforcement learning and computer vision techniques.

CONCLUSION

ML algorithms form the foundation of intelligent systems that analyze data and provide actionable insights. Understanding their strengths, limitations and applications is key to supporting them effectively in solving real-world problems. As technology evolves, these algorithms will continue to shape industries and redefine what is possible in data-driven decision-making. ML algorithms are broadly divided into three categories based on how they learn from data: Supervised learning, unsupervised learning and reinforcement learning.