All Machine Learning Algorithms

EXPLAINED IN ONE LINE

@RAMCHANDRAPADWAL

Supervised Learning Algorithms

LINEAR REGRESSION

Predicts a continuous output variable based on linear relationships between input features.

LOGISTIC REGRESSION

Classifies input data into discrete categories using a logistic function to model the probabilities.

DECISION TREES

Constructs a tree-like model by splitting data based on features to make decisions or predictions.

RANDOM FORESTS

Ensemble method that combines multiple decision trees to improve prediction accuracy and reduce overfitting.

SUPPORT VECTOR MACHINES

Separates data into different classes by finding an optimal hyperplane in a high-dimensional space.

NAIVE BAYES

Uses Bayes' theorem and assumes independence between features to classify data based on probability calculations.

K-NEAREST NEIGHBORS (K-NN)

Classifies data based on the majority vote of its k nearest neighbors in the feature space.

GRADIENT BOOSTING ALGORITHMS

Ensemble methods that sequentially build weak models, minimizing the errors of previous models to improve predictions.

Unsupervised Learning Algorithms

K-MEANS CLUSTERING

Divides data into k clusters based on similarity, aiming to minimize the intra-cluster variance.

HIERARCHICAL CLUSTERING

Builds a hierarchy of clusters by iteratively merging or splitting them based on similarity.

DBSCAN

Density-based clustering algorithm that groups together data points in high-density regions while marking outliers as noise.

GAUSSIAN MIXTURE MODELS (GMM)

Models data as a combination of Gaussian distributions to perform probabilistic clustering.

PRINCIPAL COMPONENT ANALYSIS (PCA)

Reduces the dimensionality of data by transforming it into a new set of uncorrelated variables called principal components.

T-DISTRIBUTED STOCHASTIC NEIGHBOR EMBEDDING (T-SNE)

Dimensionality reduction technique that visualizes high-dimensional data in a lower-dimensional space, emphasizing local structure.

Semi - Supervised Learning Algorithms

EXPECTATION-MAXIMIZATION (EM)

Iteratively estimates the parameters of a probabilistic model by alternately computing expected values and maximizing likelihood.

SELF-TRAINING

Uses a small amount of labeled data to train a model, which is then used to label a larger amount of unlabeled data for further training iterations.

CO-TRAINING

Simultaneously trains multiple models on different subsets of features or data instances, leveraging their agreement on the unlabeled data.

LABEL PROPAGATION

Propagates labels from labeled instances to unlabeled instances based on their similarity, utilizing the local structure of the data.

GENERATIVE MODELS WITH LABELED AND UNLABELED DATA

Combines generative models with both labeled and unlabeled data to estimate class distributions and make predictions.

Reinforcement Learning Algorithms

Q-LEARNING

Reinforcement learning algorithm that learns through trial and error, optimizing actions based on maximizing cumulative rewards.

DEEP Q-NETWORK (DQN)

Reinforcement learning algorithm that combines Q-learning with deep neural networks for improved performance in complex environments.

PROXIMAL POLICY OPTIMIZATION

Policy Optimization algorithm that iteratively updates policies to maximize rewards and improve sample efficiency.

MONTE CARLO TREE SEARCH (MCTS)

Search algorithm that simulates and evaluates possible moves in a game tree to determine optimal actions

ACTOR-CRITIC METHODS:

Reinforcement learning approach that combines a policy network (actor) and a value function (critic) to guide learning.

Deep Learning Algorithms

CONVOLUTIONAL NEURAL NETWORKS (CNN)

Deep learning models designed for image processing, using convolutional layers to extract meaningful features.

RECURRENT NEURAL NETWORKS (RNN)

Neural networks that can process sequential data by retaining and using information from previous inputs.

LONG SHORT-TERM MEMORY (LSTM)

A type of RNN that addresses the vanishing gradient problem and can retain information over longer sequences.

GENERATIVE ADVERSARIAL NETWORKS (GAN)

Neural network architecture consisting of a generator and a discriminator, trained in competition to produce realistic data.

TRANSFORMER NETWORKS

Architecture that employs self-attention mechanisms to process sequences, widely used in natural language processing tasks.

AUTOENCODERS

Neural networks designed to learn compressed representations of input data by training to reconstruct the original input from a reduced-dimensional representation.

Ensemble Learning Algorithms

BAGGING

Ensemble technique that combines multiple models trained on different subsets of the training data to make predictions.

BOOSTING

Ensemble method that combines weak learners sequentially, with each subsequent model focusing on instances that previous models struggled with.

STACKING

Ensemble approach that combines predictions from multiple models by training a meta-model on their outputs.

VOTING CLASSIFIERS

Ensemble method that combines predictions from multiple models by majority voting or weighted voting.

Dimensionality Reduction Algorithms

PRINCIPAL COMPONENT ANALYSIS (PCA)

Reduces the dimensionality of data by transforming it into a new set of uncorrelated variables called principal components.

LINEAR DISCRIMINANT ANALYSIS (LDA):

Maximizes class separability by finding linear combinations of features that best discriminate between classes.

INDEPENDENT COMPONENT ANALYSIS (ICA)

Separates a multivariate signal into additive subcomponents to discover underlying independent sources.

VARIATIONAL AUTOENCODERS (VAE)

Neural network-based generative models that learn low-dimensional representations and reconstruct original data with high fidelity.

Transfer Learning Algorithms

PRE-TRAINED DEEP NEURAL NETWORKS

Deep learning models that are trained on large-scale datasets for specific tasks, often used as a starting point for transfer learning.

FINE-TUNING

Technique where a pre-trained model is further trained on a specific task or dataset to improve its performance.

DOMAIN ADAPTATION

Technique that transfers knowledge from a source domain to a target domain with different distributions, improving generalization.

MULTI-TASK LEARNING

Simultaneously trains a model on multiple related tasks to improve overall performance by leveraging shared information.