

Machine Learning Algorithms

A Guide to **100+**
Machine Learning
Algorithms



Contents

- Clustering
- K Means
- K Medians
- BIRCH
- Fuzzy C-Means
- Fuzzy K-Modes
- Mini Batch K-Means
- DBSCAN
- Fuzzy Clustering
- Optics Algorithm
- Expectation Maximization
- Hierarchical Clustering
- Minimum Spanning Tree

Contents

- Dimensionality Reduction
- PCA Algorithm
- PCR Algorithm
- PILSR Algorithm
- Sammon Mapping
- MDS Algorithm
- Projection Pursuit
- Mean-Shift
- LDA Algorithm
- ICA Algorithm
- NMF Algorithm
- RDA Algorithm
- MDA Algorithm
- PLSDA Algorithm

Contents

- QDA Algorithm
- CCA Algorithm
- Diffusion Map
- Rule System
- Cubist
- OneR
- ZeroR
- Ripper
- ARL
- Apriori
- Eclat
- FP-Growth

Contents

- Deep Learning
- CNN
- ANN
- LSTMS
- Stacked Auto Encoders
- Deep Boltzman Machine
- Deep Belief Networks
- Decision Trees
- Conditional Decision Trees
- Decision Stump
- C4.5 & C5.0
- CHAID
- CARD
- M5

Contents

- 13
- Regression
- Least Squares
- Linear Regression
- Logistic Regression
- Stepwise Regression
- Mars Algorithm
- Locally-Estimated Scatterplot Smoothing
- Reinforcement Learning
- Q Learning
- Deep Q-Network
- Learning Automata
- DDPG Algorithm

Contents

- NAF Algorithm
- A3C Algorithm
- TRPO Algorithm
- PPO Algorithm
- Ensemble
- GBDT
- GBRT
- Gradient Boosting Machines
- Bagging
- AdaBoost
- Random Forest
- Boosting

Contents

- Bayesian
- AODE
- Naive Bayes
- Gaussian Naive Bayes
- Multinomial Naive Bayes
- Bayesian Belief Network
- Bayesian Network
- Regularization
- Ridge Regression
- Elastic Net
- LASSO
- LARS

Contents

- Other Algorithms
- ALOPEX
- GeneRec
- Leabra
- Rprop
- FastICA
- Evaluation
- LogitBoost
- Sparse PCA
- Structured KNN
- WMA Algorithm
- CN2 Algorithm
- Feature Selection
- Forward-Backward

Contents

- Algorithm Accuracy
- Performance Measures
- Optimization Algorithm
- Dynamic Time Warping
- Local Outlier Factor
- Logic Learning Machine
- Markov Chain Monte Carlo
- Wake Sleep Algorithm
- Linde Buzo Gray Algorithm

END

Clustering

K Means

A clustering algorithm that partitions data into k clusters based on the distances between data points and the means of each cluster.

K Median

It is a clustering algorithm that is similar to K-Means. However, it uses medians instead of means to define the cluster centroids.

BIRCH

A hierarchical clustering algorithm that uses CF-tree to reduce the number of data points and improve the clustering efficiency.



Clustering

Fuzzy C-Means

A clustering algorithm that assigns each data point a degree of membership to each cluster. The degree is represented by a fuzzy value.

Fuzzy K-Modes

A variant of Fuzzy C-Means that can handle categorical data by defining the cluster centroids as the modes of the data.

Mini Batch K-Means

A variant of K Means that uses random subsets (mini-batches) of data to update the cluster centroids, making it more scalable to large datasets.



Clustering

DBSCAN

A density-based clustering algorithm that identifies clusters as dense regions of data separated by areas of lower density.

Fuzzy Clustering

Fuzzy clustering is a type of clustering algorithm that allows one data point to belong to more than one cluster with varying degrees of membership.

Optics Algorithm

Optics algorithm is a density-based clustering algorithm that can discover clusters of arbitrary shape and size in a dataset.



Clustering

Expectation Maximization

A general algorithm used to estimate the parameters of probabilistic models, including Gaussian Mixture Models for clustering.

Hierarchical Clustering

A clustering algorithm that builds a hierarchy of clusters using a tree-like structure (dendrogram). It can be either agglomerative or divisive.

Minimum Spanning Tree

MST is a graph-based algorithm that constructs a tree-like structure connecting all data points with minimum total edge weight.



Dimensionality Reduction

PCA Algorithm

PCA is a dimensionality reduction technique used to reduce the dimensions of large datasets by identifying and extracting the most important features.

PCR Algorithm

Principal Component Regression (PCR) regression algorithm that uses the principal components obtained from PCA to fit a linear regression model.

PLSR Algorithm

A non-linear dimensionality reduction algorithm that seeks to preserve the pairwise distances between data points in the lower-dimensional space.



Dimensionality Reduction

Sammon Mapping

A non-linear dimensionality reduction algorithm that seeks to preserve the pairwise distances between data points in the lower-dimensional space.

MDS Algorithm

A dimensionality reduction algorithm that aims to preserve the pairwise distances between data points by projecting them onto a lower-dimensional space.

Projection Pursuit

A dimensionality reduction algorithm that finds low-dimensional projections of high-dimensional data that maximize some measure of interestingness.



Dimensionality Reduction

Mean-Shift

A clustering algorithm that finds clusters by iteratively shifting the centroid of a kernel density estimate of the data until convergence.

LDA Algorithm

LDA algorithm projects high-dimensional data onto a lower-dimensional space while maximizing the separation between classes.

ICA Algorithm

A blind source separation algorithm that recovers the underlying sources that generate observed data by assuming they are statistically independent.



Dimensionality Reduction

NMF Algorithm

A dimensionality reduction algorithm that factorizes a non-negative matrix into two low-rank matrices that are a set of basis vectors and coefficients.

RDA Algorithm

Supervised learning algorithm that combines elements of LDA and ridge regression to handle high-dimensional data with collinearity between features.

MDA Algorithm

MDA (Multiple Discriminant Analysis) is a supervised classification algorithm that extends LDA to handle more than two classes.



Dimensionality Reduction

PLSDA Algorithm

PLSDA (Partial Least Squares Discriminant Analysis) supervised classification algorithm that generalizes PLSR to handle classification problems.

QDA Algorithm

A supervised classification algorithm that models the class-conditional distributions of the data using quadratic functions.

CCA Algorithm

CCA algorithm finds linear combinations of the original features that maximize the correlation between two different sets of variables.



Dimensionality Reduction

Diffusion Map

Diffusion Map algorithm uses diffusion processes to construct a low-dimensional embedding of high-dimensional data based on its intrinsic geometry.



Rule System

Cubist

It is a decision tree-based algorithm that can also handle continuous data by creating a linear regression model at the leaves of the tree.

OneR

It is a simple algorithm that selects one feature and creates a rule based on the feature values that minimizes the error rate.

ZeroR

It is a very basic algorithm that always predicts the most frequent class in the training data. It does not use any predictor variables.



Rule System

Ripper

It is a rule-based algorithm that uses separate-and-conquer strategy. It creates rules from the data and can be used for classification or regression.



ARL

Apriori

Apriori is a classic data mining algorithm that scans the database and identifies frequent itemsets and association rules from a given dataset.

Eclat

It is a frequent itemset mining algorithm that uses a vertical data format and depth-first search strategy to find all frequent itemsets in the database.

FP-Growth

FP-Growth algorithm uses a divide-and-conquer strategy to compress the database into a FP-tree and mines frequent itemsets from the FP-tree.



Deep Learning

Convolutional Neural Networks

They are used for image and video processing and are particularly effective in tasks such as image classification, object detection, and face recognition.

Artificial Neural Networks

ANN is a machine learning model inspired by the structure and function of neural networks in the brain. Can be used for classification and regression.

Long Short-Term Memory

They are used for sequence modelling tasks such as language translation, speech recognition, and sentiment analysis.



Deep Learning

Stacked Auto Encoders

SAE's use an encoding process to compress the input data into a lower-dimensional space and a decoding process to reconstruct the data.

Deep Boltzmann Machines

They are a type of generative model that learns to represent and generate data by modeling the probability distribution of the data.

Deep Belief Networks

Generative model that uses unsupervised learning to extract meaningful features from input data and learn a hierarchical representation of the data.



Decision Trees

Conditional Decision Trees

Decision trees that incorporate if-then rules into the decision-making process. Domain experts provide these rules to improve the accuracy of the model.

Decision Stump

It is a simple decision tree model that consists of a single decision node and two leaf nodes. It is often used as a building block in more complex models.

C4.5 & C5.0

Decision tree algorithms that use entropy and information gain to select the best split at each node. These algorithms are developed by Ross Quinlan.

Decision Trees

CHAID

Chi-squared Automatic Interaction Detection is a decision tree algorithm that uses statistical tests to determine the significance of each split.

CART

It stands for Classification and Regression Trees and is a decision tree algorithm that can be used for both classification and regression tasks.

M5

It is a decision tree algorithm that uses a model-based approach to improve the accuracy of the model. It uses both regression and classification techniques.

Decision Trees

I3

It is a decision tree algorithm that uses a combination of information gain and minimum description length to select the best split at each node.

Regression

Least Squares

A method to find the best-fit line that minimizes the sum of the squared distances between the observed data points and predicted values by the model.

Linear Regression

It is a type of regression that models the relationship between a dependent variable and one or more independent variables using a linear equation.

Logistic Regression

It models the probability of the dependent variable (binary) taking on a particular value as a function of the independent variables.

Regression

Stepwise Regression

It is a statistical technique that iteratively adds or removes predictor variables in a linear regression model until an optimal model is obtained.

MARS Algorithm

A regression algorithm that models the relationship between the dependent and the independent variables using a series of piecewise linear functions.

LEOSS

Locally Estimated Scatterplot Smoothing is a non-parametric regression technique used for smoothing scatterplots. Can be used for interpolation and prediction.



Reinforcement Learning

Q Learning

It is a model-free reinforcement learning algorithm that learns the value of an action in a particular state based on the maximum expected future rewards.

Deep Q-Network

It is a neural network-based algorithm that uses Q-learning to learn to make optimal decisions by using a deep neural network.

Learning Automata

It is a type of reinforcement learning algorithm that is based on the theory of automata and used for solving multi-armed bandit problems.



Reinforcement Learning

DDPG Algorithm

It stands for Deep Deterministic Policy Gradient and is an actor-critic reinforcement learning algorithm that is used for continuous control tasks.

NAF Algorithm

It stands for Normalized Advantage Function and is a type of actor-critic reinforcement learning algorithm that is used for continuous control tasks.

A3C Algorithm

It stands for Asynchronous Advantage Actor-Critic and is a type of actor-critic reinforcement learning algorithm that is used for continuous control tasks.



Reinforcement Learning

TRPO Algorithm

It stands for Trust Region Policy Optimization and is a type of policy optimization algorithm that is used for reinforcement learning.

PPO Algorithm

It stands for Proximal Policy Optimization and is a type of policy optimization algorithm that is used for reinforcement learning.



Ensemble

Gradient Boosting Decision Tree

GBDT is an ensemble machine learning algorithm that combines multiple decision trees to form a more powerful and accurate prediction model.

Gradient Boosting Regression Tree

It is an ensemble machine learning algorithm that combines multiple regression trees to form a more powerful and accurate prediction model.

Gradient Boosting Machines

It is an ensemble machine learning algorithm that combines multiple weak prediction models to form a more powerful and accurate model.



Ensemble

Bagging (Bootstrap Aggregating)

An ensemble technique that creates samples of the dataset through bootstrapping, training, then combining to improve the predictive accuracy.

AdaBoost (Adaptive Boosting)

It is a boosting algorithm that assigns higher weights to misclassified samples to improve their importance in the next iteration.

Minimum Random Forest Tree

It is an ensemble method that uses a collection of decision trees trained on different subsets of the data and features.



Ensemble

Boosting

Boosting is an ensemble learning technique in which multiple weak models are combined to form a strong model through training on a subset of data.



Bayesian

AODE

A Bayesian algorithm for constructing classification models that uses a hybrid approach of decision trees and Naive Bayes.

Naive Bayes

A simple probabilistic algorithm based on Bayes' theorem that makes strong independence assumptions between features.

Gaussian Naive Bayes

A variant of Naive Bayes that assumes the probability distribution of each feature is Gaussian and computes the conditional probabilities accordingly.



Bayesian

Multinomial Naive Bayes

A variant of Naive Bayes that is suitable for discrete count data, where features represent the frequency with which certain events occur.

Bayesian Belief Network

A graphical model that represents variables and their dependencies through directed acyclic graphs, which allows for efficient probabilistic inference.

Bayesian Network

A probabilistic graphical model that represents variables, their dependencies through directed acyclic graphs. Allows for efficient probabilistic inference.



Regularization

Ridge Regression

Ridge Regression is a linear regression model with L2 regularization that adds a penalty term to the cost function to reduce overfitting.

Elastic Net

Elastic Net is a linear regression model with L1 and L2 regularization combined to overcome the limitations of LASSO and Ridge regression.

LASSO

A linear regression model with L1 regularization that adds a penalty term to the cost function to reduce overfitting and perform feature selection.



Regularization

LARS

A linear regression algorithm that iteratively adds variables to the model to find the best subset of variables that predict the target variable.



Other Algorithms

ALOPEX

ALOPEX is an unsupervised machine learning algorithm that combines principal component analysis and clustering.

GeneRec

A supervised learning algorithm that utilizes a genetic algorithm to optimize feature selection and classification models simultaneously.

Leabra

It is a biologically-inspired machine learning algorithm that simulates the functioning of the human brain and is used for classification tasks.



Other Algorithms

Rprop

Rprop is a supervised learning algorithm that adjusts the weight update step size based on the sign of the gradient.

FastiCA

FastiCA is an unsupervised machine learning algorithm that performs independent component analysis using a fast fixed-point algorithm.

Evaluation

Evaluation refers to the process of assessing the performance of a machine learning model using various metrics.



Other Algorithms

LogitBoost

LogitBoost is a boosting algorithm and fits base-learners in a stage-wise manner by minimizing the negative log-likelihood of logistic regression.

Sparse PCA

It is a variant of PCA that helps in extracting the most important features from a large set of variables leading to sparse representation.

Structured KNN

Structured KNN considers the relationships between feature dimensions and uses them to improve the classification performance.

Other Algorithms

WMA Algorithm

Weighted Moving Average is a method of time series forecasting that assigns weights to data points in a sliding window to calculate the average.

CN2 Algorithm

CN2 uses a divide-and-conquer approach to induce a set of classification rules from a given dataset, which are used to make predictions for new, unseen data.

Feature Selection

Feature selection is a process of selecting a subset of relevant features from the original set to improve the accuracy of a machine learning model.

Other Algorithms

Forward-Backward

Forward-Backward is an algorithm used to estimate the coefficients of a linear regression model with autoregressive errors.

Algorithm Accuracy

Algorithm Accuracy refers to the degree to which a machine learning algorithm's predictions match the actual outcomes.

Performance Measures

Performance Measures are metrics used to evaluate the performance of a machine learning model, such as precision, recall, and F1-score.



Other Algorithms

Optimization Algorithm

Optimization Algorithm is a method used to find the set of parameters that minimizes or maximizes an objective function.

Dynamic Time Warping

DTW is a distance-based similarity measure used for sequence alignment between two time-series datasets, which can be of unequal length.

Local Outlier Factor

Local Outlier Factor (LOF) is an unsupervised outlier detection algorithm used for identifying anomalies in a dataset.



Other Algorithms

Logic Learning Machine

Logic Learning Machine is a supervised learning algorithm that utilizes a logic-based approach to construct decision trees.

Markov Chain Monte Carlo

Markov Chain Monte Carlo (MCMC) is a computational algorithm used to obtain numerical approximations for complex statistical models.

Wake Sleep Algorithm

It is an unsupervised learning algorithm for training deep belief networks which alternates between a wake phase and a sleep phase.



Other Algorithms

Linde Buzo Gray Algorithm

This is a vector quantization algorithm which partitions data into groups and computes the centroids of those groups to represent the data.