

Machine Learning Algorithms

Supervised Learning Algorithms

- Logistic Regression
- K-Nearest Neighbors (K-NN)
- Support Vector Machines (SVMs)
- Kernel SVMs
- Naive Bayes
- Decision Tree Classification
- Random Forest Classification

Unsupervised Learning Algorithms

- K-Means Clustering
- Hierarchical Clustering
- Probabilistic Clustering

Reinforcement Learning Algorithms

- Model-Free Reinforcement Learning
 - Policy Optimization
 - Q-Learning
- Model-Based Reinforcement Learning
 - Learn the Model
 - Given the Model

Logistics Regression

Logistic Regressing (also known as Logit Regression) is commonly used to estimate the probability that an input (sample) belongs to a particular class. For example, what is the probability that the patient is diabetic? If the probabily is greater than 50%, then the model predicts that the given input belongs to class 1 otherwise class 0 in binary classification problem.

K-Nearest Neighbors (K-NN)

K-NN algorithm can be used to solve both classification and regression problems. K-NN algorithm assumes the similarity between the new data point (sample) and available data point and put the new data into the category that is most similiar to the available categories. For example, we have an image of a creature that looks similar to cat and dog, but we want to know either it is a cat or dog. So for this purpose, we will use KNN algorithm, as it works on a similarity measure.

Support Vector Machines (SVMs)

A SVM is a powerful Machine Learning algorithm, it can be performed on clasification, regression, and even on outlier detection. SVMs are particularly well suited for classification of complex small or medium sized datasets. SVM can be used for face detection, image classification, text categorization, etch

Kernel SVM

Although linear SVM classifiers are efficient and work well mostly, but many datasets are not even close to being linearly separable. One approach to handling nonlinear datasets is to use different kernel functions that are used in SVM. These functions can be different types. For example, linear, nonlinear, polynomial, radial basis function (RBF), and sigmoid.

Naive Bayes

Naive Bayes classifiers are a collection of classification algorithms based on **Bayes' Theorem**. It assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature. For example, a fruit may be an orange if it is orange in color, round and have about 4 inches in diameter.

Decision Tree Classification

Like SVMs, Decision Trees are versatile Machine Learning algorithms that can perform both classification, and regression tasks, and even multioutput tasks. Decision Trees are capable to work with large datasets. For example, make house price predictions.

Random Forest Classification

Random Forest contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset. For example, suppose there is a dataset that contains multiple images of apples and orange and this dataset is given to the *Random Forest Classifier*. The dataset get processed through multiple decision trees and then based on the majority of results, the Random Forest classifier predicts the final decision.

K-Means Clustering

K-Means algorithm is a simple algorithm capable of clustering the unlabeled dataset very quickly and efficiently, often in just a few iterations. For example, Imagine we received data of cricket players which gives information on the runs scored by the player and the wickets taken by them. Based on this information, we need to group the data into two clusters, namely batsman and bowlers.

Hierarchical Clustering

Hierarchical clustering can be divided into two main types: **agglomerative and divisive**. Agglomerative clustering works in a bottom-up manner, whereas Divisive clustering works in a top-down manner.

Probabilistic Clustering

Probabilistic model assumes that the instances were generated from a mixture of several Gaussian distributions whose parameters are unknown. All the instances generated from a single Gaussian distribution form a cluster and each cluster can have different shape, size, density, and orientation.

Model-Free Reinforcement Learning

A model-free algorithm is an algorithm which does not use the *transition probability distribution* and the reward function which represents the problem to be solved. The transition probability distribution and the reward function are often collectively called the "model" of the environment, hence the name "model-free". Some Model-Free reinforcement learning algorithms are:

1. Policy Optimization
2. Q-Learning

1. Policy Optimization

Policy optimization methods are centered around the policy, the function that maps the agent's state to its next action. In these methods we optimize the expected reward with respect to the policy's parameters.

2. Q-Learning

Q-Learning is a model-free reinforcement learning algorithm o learn the value of an action in a particular state. It does not require a model of the environment (hence "model-free"), and it can handle problems with stochastic transitions and rewards without requiring adaptations.

Model-Based Reinforcement Learning

Model-based Reinforcement Learning refers to learning optimal behavior indirectly by learning a model of the environment by taking actions and observing the outcomes that include the next state and the immediate reward.