Decision Tree

A decision tree is a structure that is used to predict the value of a target value based on certain input variables. Each node of the tree represents a decision that will affect the outcome of the target value. One advantage of using a decision tree to predict behavior is that decision trees are easy to understand and follow. It is easy to follow the conditional logic that decision trees use.

Random Forest

Decision trees that grow deep might lead to some incorrect results, random forests are a way to deal with this. Random forests work by building a number of decision trees and merging them together and taking the averages of the test variables. One advantage of random forests is that it reduces the variance of using decision trees. A disadvantage is that making multiple decision trees and merging them together might be run slow in some cases.

Logistic Regression

Logistic regression is a classification algorithm. It predicts the probability of an input belonging to a certain set by separating data into two regions. Logistic regression is used when the response variable will be binary, for example, pass/fail.

K-Nearest Neighbors

K-Nearest Neighbors is an algorithm that can be used for both classification and regression. It works by taking in training data and then seeing which data points are close to a data point and then classifying that data point as part of the same class as the majority of the k-nearest data points. An advantage of k-nearest neighbors is that it is usually pretty accurate and works well for non-linear data. A disadvantage is that it has to store all the training data which can lead to memory and runtime issues.

Naive Bayes

Naive Bayes is a collection of algorithms that are based on the Bayes Theorem. It works by classifying features of input data. It can then predict the class of input by seeing which features match between classes. Some advantages are that it is fast and easy to train. A disadvantage is that it assumes every feature is independent of the other which is not always true for every class.

Perceptron

A perceptron is used for classifying data. It is a linear classifier so it classifies data into two categories. It works by calculating a weighted sum of input values and returning 1 if the sum is greater than a certain value or it returns 0 otherwise.

Gradient Descent

Gradient Descent is an algorithm that minimizes a cost function. Gradient descent is important in machine learning because it optimizes how well a machine learning algorithm is working by minimizing that algorithm’s cost function. It works by calculating gradients and using those gradients to change weights for predictions.

Support Vector Machine

Support Vector Machine is a linear model that can be used for classification or regression. It works by creating a line or plane that separates data into classes. It finds a line between points from the two classes and then maximizes the distance between the line and the points. SVMs usually have fast performances.