ML Algorithms

KNN

K-Nearest Neighbour is a Supervised Learning-based Machine Learning algorithm. The algorithm maintains all the existing data and uses similarity to classify data points. i.e when new data is obtained it is instantly classified into the best-suited category using the algorithm. Both classification and regression problems can be solved by this method, but it is more commonly utilized for classification tasks. It is widely used due its easy interpretation and low calculation time.

Logisticxregression

Chart, line chart

Description automatically generated

Logistic Regression is also termed as the logit model. It is a type of statistical analysis which is frequently used for predictive analytics, and machine learning.

This type of analysis helps to predict the likelihood of an event or a choice being made.

It works best when the task is based on two values or binary classification.

Based on the dataset, a threshold can be established to forecast which class a data will belong. The derived estimated probability is categorized into classes based on this threshold.

The aim of this algorithm is to discover a link between the different characteristics and the find the likelihood of a specific outcome.

AdaBoost Classifier

AdaBoost is an ensemble learning algorithm which was initially developed to improve the performance and efficiency of the existing binary classifiers. It learns from weak classifiers and iterate them to make them strong. It is usually called the "best out-of-the-box classifier" for this reason.

A single classifier might not accurately anticipate the class of an object. A classifier that outperforms random guessing but is poor at designating classes to objects is called a weak classifier. However, when multiple weak classifiers are grouped together where each one gradually learns from the others' incorrectly classified objects, one robust model can be developed. The classifier could be any of the basic classifiers, from Decision Trees (often the default) through Logistic Regression.

Naive Bayes classifier

It is generally based on the Bayes theorem and falls under the category of supervised learning algorithm which is used to solve classification related problems. It is a very fast model which is capable of making quick predictions and is simple to use while making the effective classification.It is based on probability, which means it makes predictions based on an object's likelihood.

Here “Naïve ” terms tell that it’s outcome is independent of the occurrence of variables and “Bayes ” shows us that it is uses Bayes theorem

Bayes' Theorem concept.

**P(A|B)=P(B|A)xP(A)/P(B)**…………………………………..

Where,

P(A|B) - Posteriorxprobability

P(B|A) - LikelihoodXprobability

P(A) - PriorXProbability

P(B) - MarginalxProbability

Gradient Boosting Classifier

It is a collection of algorithms where several weak learning models are combined together to generate a strong and powerful predictive model. Generally, a decision tree is used to combine the weak model. . it has also the ability to classify datasets efficiently

Gradient Boosting classifier’s main aim is to reduce the loss. On training when a weak learner is added to a Gradient Boosting Machine, the weights of the prior model are fixed, and the new layers are injected with no changes.

Stochastic Gradient Descent

Stochastic gradient descent is an optimization algorithm for determining model parameters that correspond to the best fit between expected and actual outputs. It's an inexact but effective method.

Gradient descent technique modified using stochastic gradient descent algorithms. In stochastic gradient descent, the gradient is calculated using only a random small sample of the observations rather than all of them. This method can minimise computing time in some circumstances.

Online stochastic gradient descent is a type of stochastic gradient descent in which the gradient of the cost function is estimated for each observation and the decision variables are updated accordingly. If the objective function is convex, this can help you identify the global minimum.

Batch stochastic gradient descent is a hybrid of traditional gradient descent and online methods. Iteratively, with subsets of all observations termed minibatches, the gradients are calculated and the decision variables are updated. For neural network training, this variation is particularly popular..