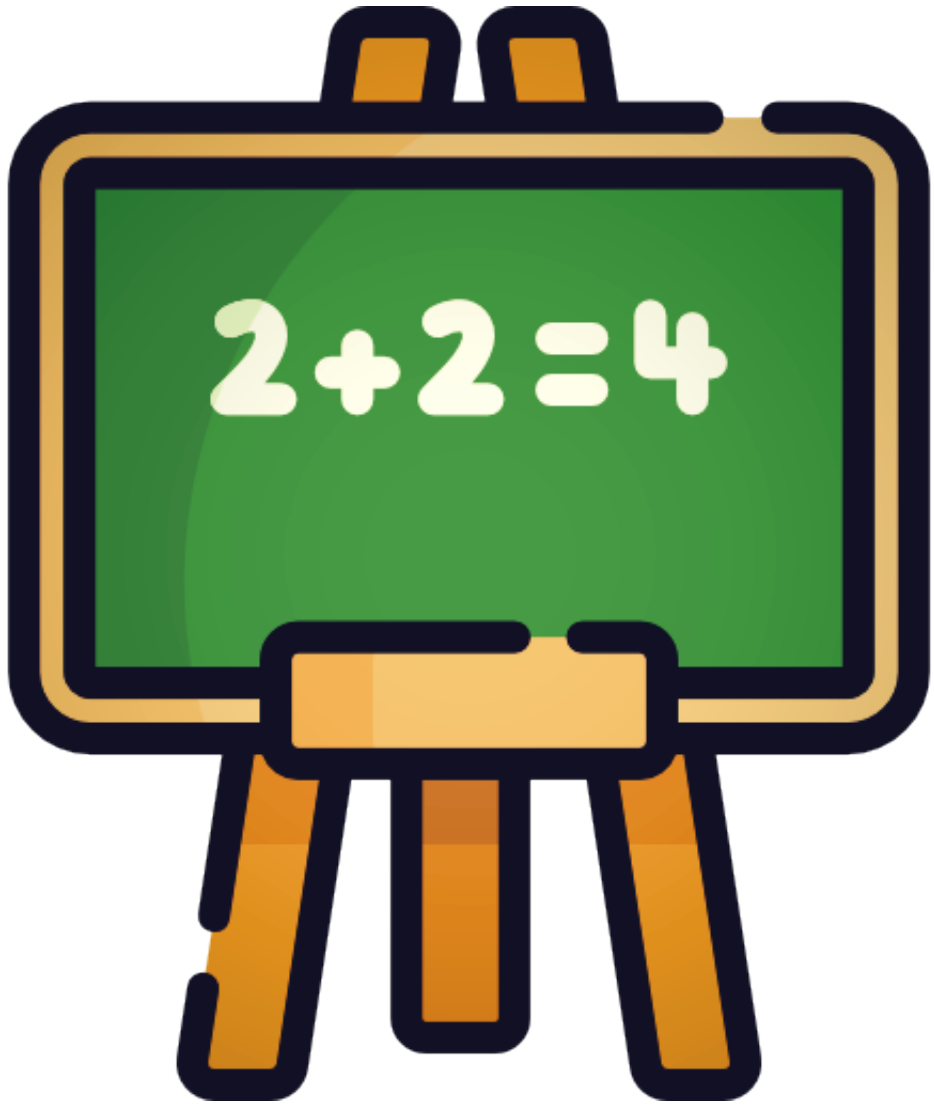
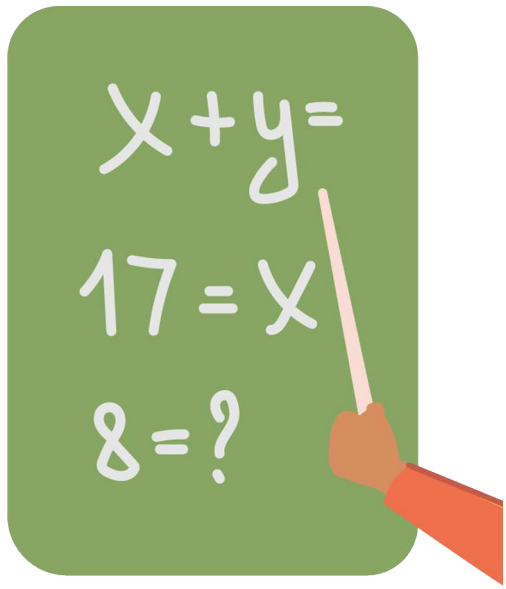
What is Supervised Learning?



DATARANCH.org
VISUALIZE | ANALYZE | CAPITALIZE

Supervised Learning



There are three main types of machine learning:

1. Supervised learning
2. Unsupervised learning
3. Reinforcement learning

In supervised learning, the model is trained on “labeled” data, where the input and output data are known. The goal is to use the labeled dataset to learn so that it can predict the output for new inputs.

Categories of Supervised Learning

Classification

A classification problem is where the output variable is a category, such as “green” or “red” , “disease” or “no disease”.

Regression

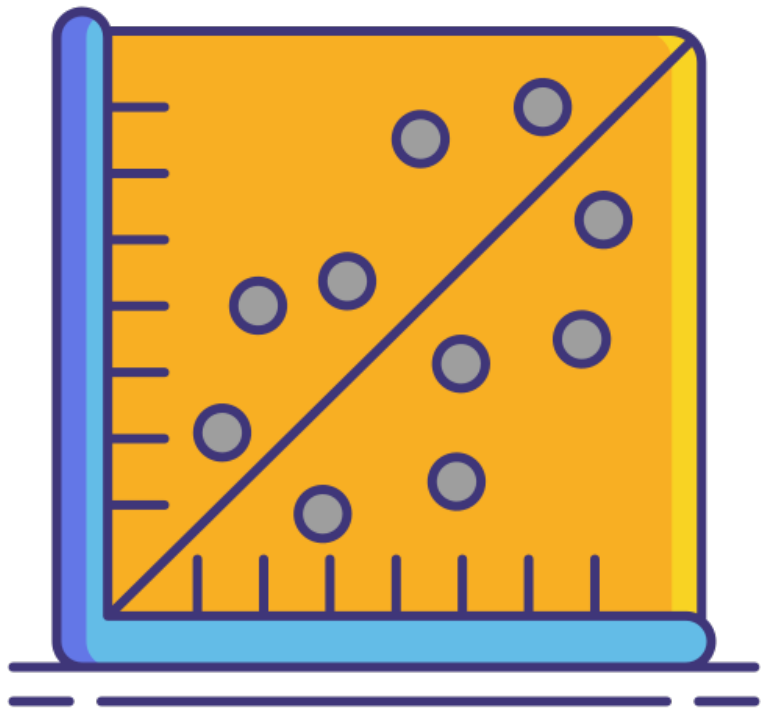
A regression problem is where the output variable is a real value, such as “amount” or “weight”.

Types of Supervised Learning Algorithms

DATARANCH.org

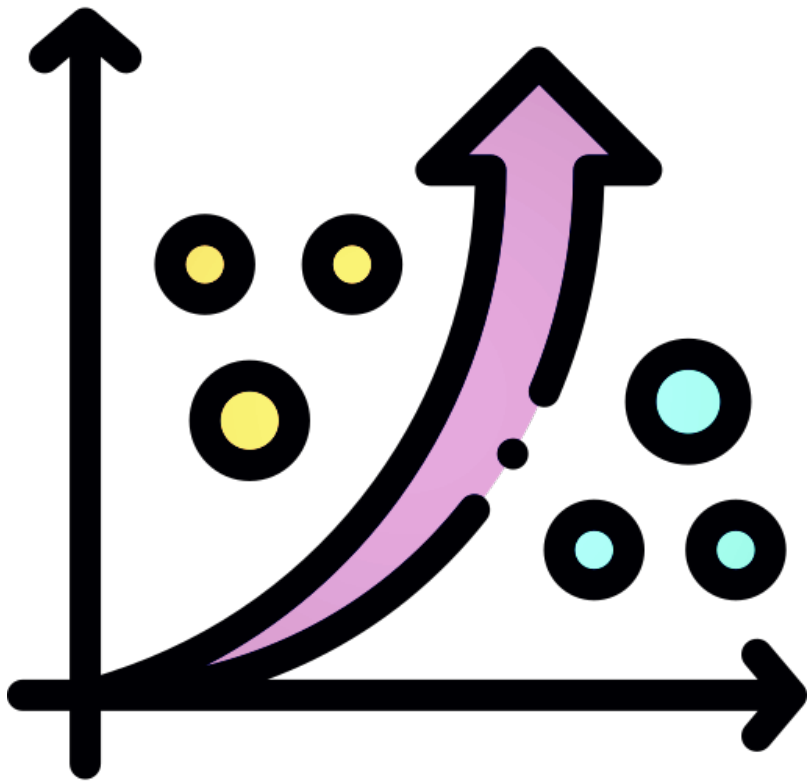
VISUALIZE | ANALYZE | CAPITALIZE

Linear Regression



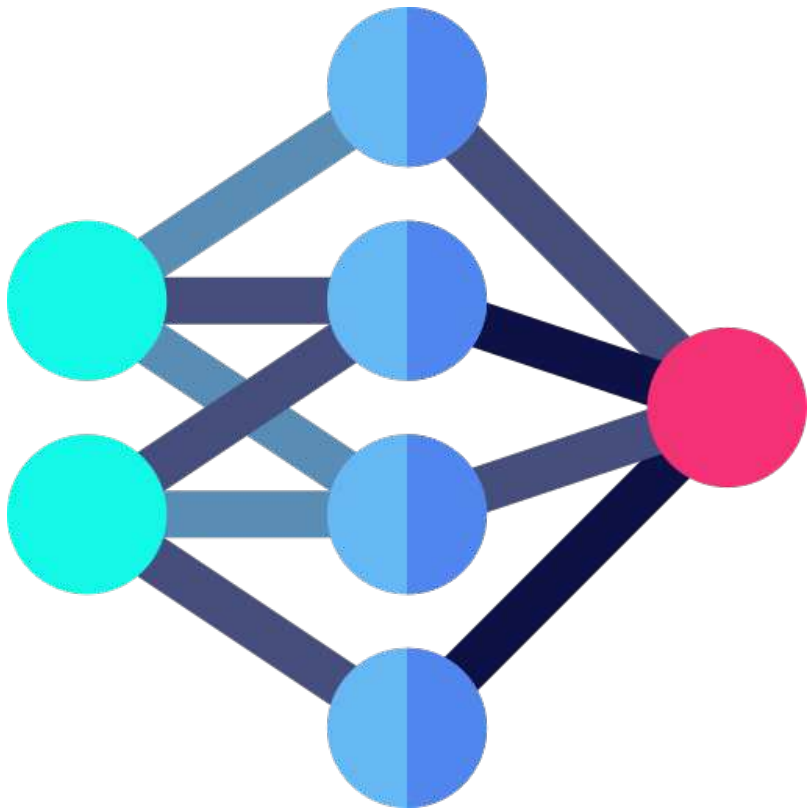
It is used to identify the relationship between a dependent variable and one or more independent variables and is typically used to make predictions about future outcomes. E.g. predicting the price of a house based on its size and location.

Logistic Regression



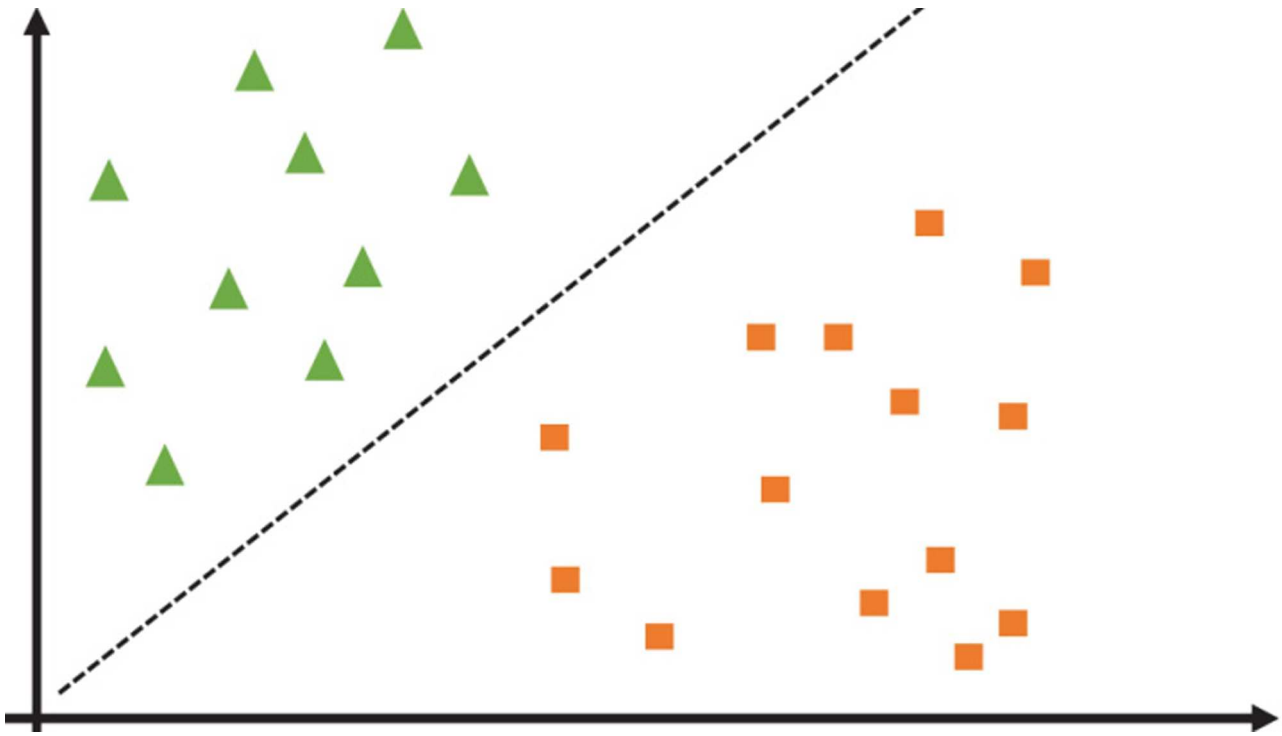
It is used for binary classification problems, where the output variable is categorical with two possible values. For example, predicting whether a customer will buy a product or not based on their demographics and purchase history.

Neural Networks



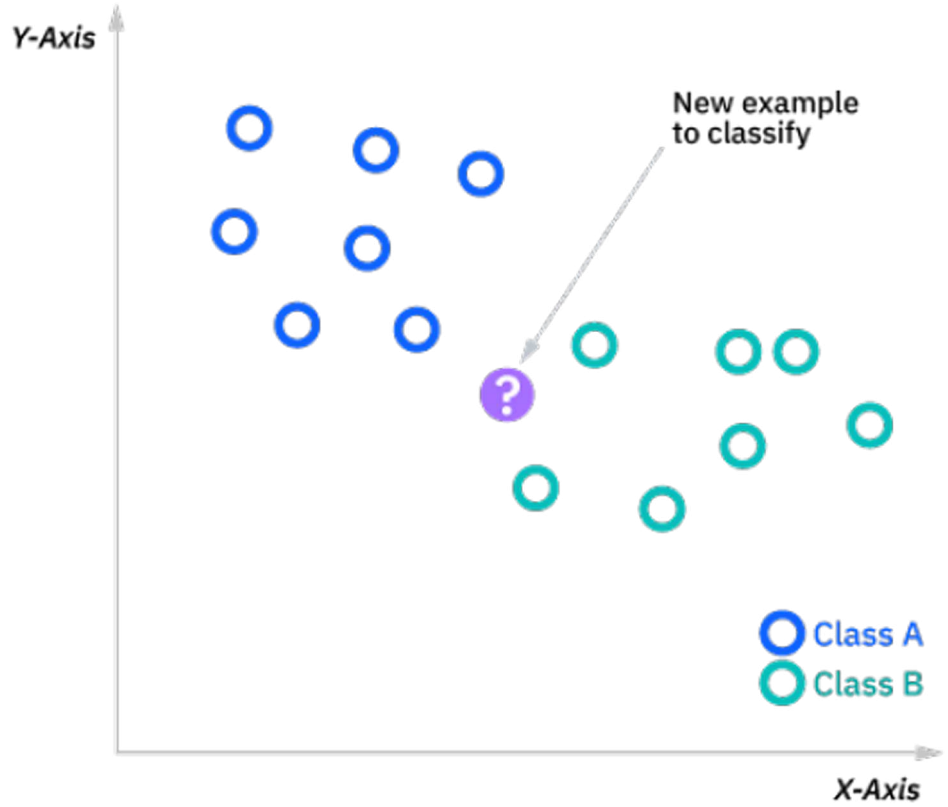
This algorithm is designed to cluster raw input, recognize patterns, or interpret sensory data. Despite their multiple advantages, neural networks require significant computational resources. It can get complicated to fit a neural network when there are thousands of observations.

Support Vector Machines



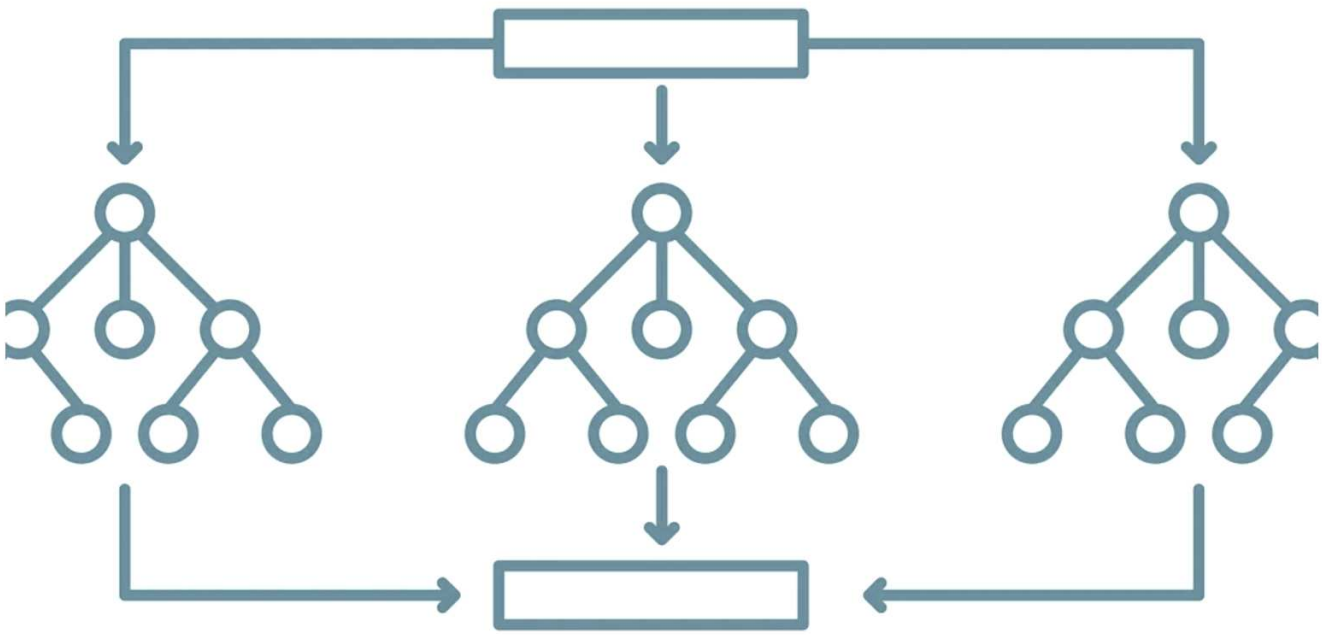
A support vector machine is typically used for classification problems by constructing a hyperplane where the distance between two classes of data points is at its maximum. This hyperplane is known as the decision boundary, separating the classes of data points (e.g., oranges vs. apples) on either side of the plane.

K-Nearest Neighbor



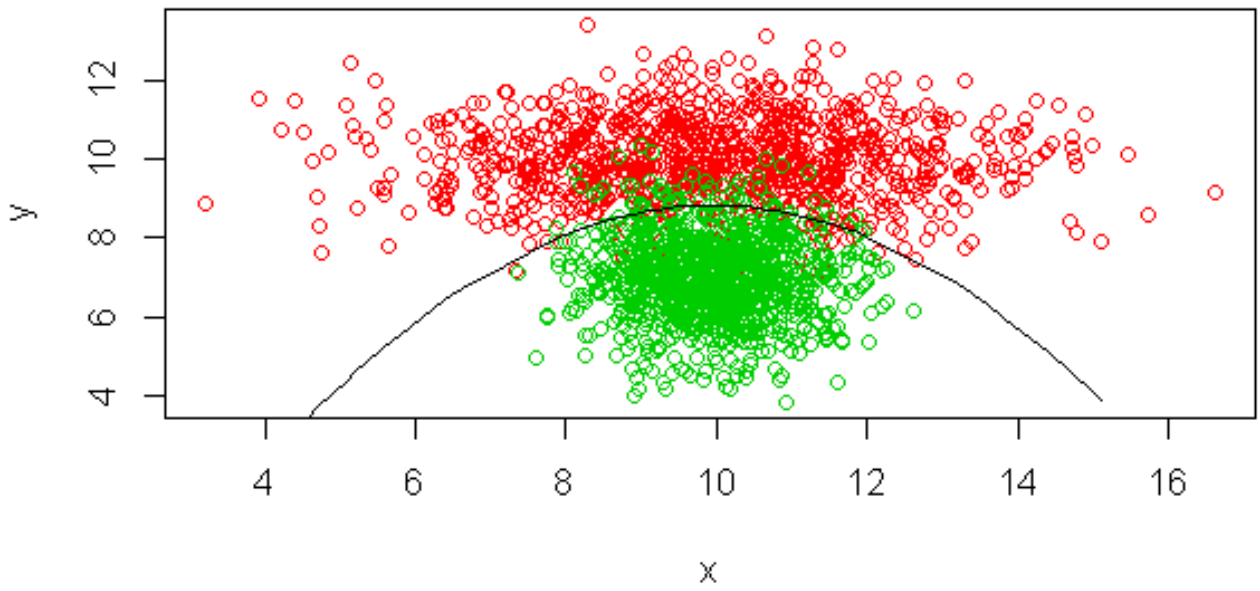
K-nearest neighbor, also known as the KNN algorithm, is a simple classification algorithm that assigns a new data point to the class that is most common among its k-nearest neighbors in the training set. KNN is typically used for recommendation engines and image recognition.

Random Forest



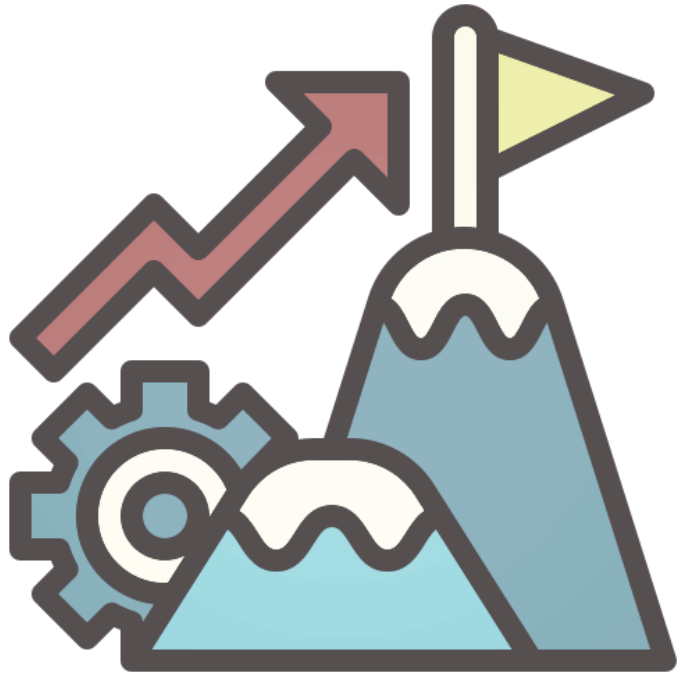
Random forest is another flexible supervised machine learning algorithm used for both classification and regression purposes. It is an ensemble learning algorithm that combines multiple decision trees to improve accuracy and reduce overfitting.

Naive Bayes



Naive Bayes assumes that the features (input variables) are conditionally independent of each other given the class label. This is a "naive" assumption because in reality, features may be correlated with each other. The three main types of Naive Bayes algorithms: Gaussian Naive Bayes, Multinomial Naive Bayes and Bernoulli Naive Bayes

Challenges of Supervised Learning



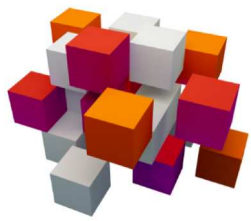
1. Supervised learning models require certain expertise levels to structure accurately.
2. Training supervised learning models can be very time intensive.
3. Datasets can have a higher likelihood of human error, resulting in algorithms learning incorrectly.
4. Unlike unsupervised learning models, supervised learning cannot cluster or classify data on its own.

Follow **#DataRanch** on LinkedIn for more...

Data Analysis Steps

Data Cleaning Steps

Data Wrangling Steps



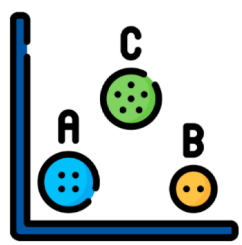
Essential Chart Types

Common Data Mistakes

What is Unsupervised Learning?

Principal Component Analysis

Clustering

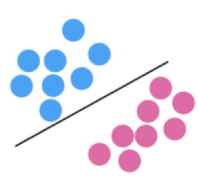


t-Distributed Stochastic Neighbour Embedding (t-SNE)

What is Supervised Learning?

Logistic Regression

Support Vector Machine

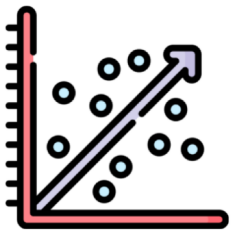


DATARANCH.org

VISUALIZE | ANALYZE | CAPITALIZE

Follow **#DataRanch** on LinkedIn for more...

Regression Analysis

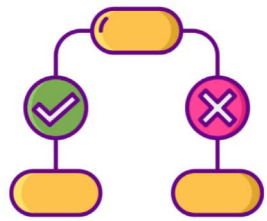


DATARANCH.org

Random Forest

DATARANCH.org

Decision Trees



DATARANCH.org

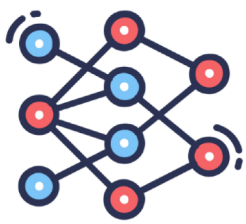
Deep Learning & Neural Networks

DATARANCH.org

Convolutional Neural Network (CNN)

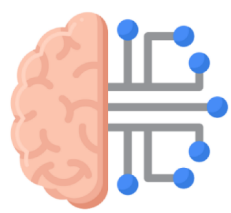
DATARANCH.org

Recurrent Neural Network (RNN)



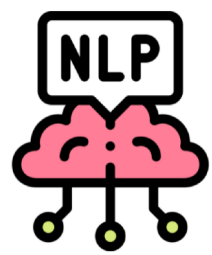
DATARANCH.org

Generative AI



DATARANCH.org

Natural Language Processing Models



DATARANCH.org

Deepfake

The purpose of this article is to spread awareness about Deepfake. The author is strongly against the spread of false information.

DATARANCH.org

DATARANCH.org

VISUALIZE | ANALYZE | CAPITALIZE

DATARANCH.org

VISUALIZE | ANALYZE | CAPITALIZE

info@dataranch.org

linkedin.com/company/dataranch