Algorithms

*Websites sourced*

[Tutorials List - Javatpoint](https://www.javatpoint.com/)

[Support Vector Machine - an overview | ScienceDirect Topics](https://www.sciencedirect.com/topics/neuroscience/support-vector-machine)

[Towards Data Science](https://towardsdatascience.com/)

Linear Regression

* **Linear Regression is a** machine learning algorithm based on**supervised learning.**
* **It is a Machine Learning algorithm that allows us to map numeric inputs to numeric outputs, fitting a line into the data points. Linear regression is a linear model, e.g., a model that assumes a linear relationship between the input variables (x) and the single output variable (y).**
* ![Chart, scatter chart

  Description automatically generated]()
* It is mostly used for finding out the relationship between variables and forecasting.
* Linear regression makes predictions for continuous/real or numeric variables such as **sales, salary, age, product price,** etc.

Logistic Regression

* Logistic regression is a Supervised Learning technique
* It is a machine learning algorithm used to predict the probability that an observation belongs to one of two possible classes.
* **Logistic regression is used for solving the classification problems**
* **It can be used to b**uild an email classifier to tell us whether an incoming email should be marked as “spam” or “not spam” or check radiological images to predict whether a tumor is benign or malignant.
* In Logistic regression, instead of fitting a regression line, we fit an "S" shaped logistic function, which predicts two maximum values (0 or 1). The curve from the logistic function indicates the likelihood of something



Decision Tree

* Decision Tree is a **Supervised learning technique**that can be used for both classification and Regression problems, but mostly it is preferred for solving Classification problems.
* In a Decision tree, there are two nodes, which are the **Decision Node** and**Leaf Node.** Decision nodes are used to make any decision and have multiple branches, whereas Leaf nodes are the output of those decisions and do not contain any further branches.
* The decisions or the test are performed on the basis of features of the given dataset.
* 

•Is it Supervised/Unsupervised/Reinforcement learning?•What does the algorithm do?•In which situations will it be most useful?•(Optional) Can you find any examples of where this algorithm has been used?

SVM (Support Vector Machine)

* SVM is a supervised training algorithm that can be useful for the purpose of classification and regression
* A support vector machine (SVM) is a supervised [machine learning](https://monkeylearn.com/machine-learning/) model that uses [classification algorithms](https://monkeylearn.com/blog/machine-learning-algorithms/) for two-group classification problems
* To separate the two classes of data points, there are many possible hyperplanes that could be chosen. Our objective is to find a plane that has the maximum margin, Hyperplanes are decision boundaries that help classify the data points. Data points falling on either side of the hyperplane can be attributed to different classes.

Chart

Description automatically generated

* a machine learning method that has become exceedingly popular for neuroimaging analysis in recent years.

Naive Bayes

* Naïve Bayes algorithm is a supervised learning algorithm, which is based on **Bayes theorem** and used for solving classification problems.
* It is mainly used in *text classification* that includes a high-dimensional training dataset.
* Naïve Bayes Classifier is one of the simple and most effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions.
* **It is a probabilistic classifier, which means it predicts on the basis of the probability of an object**.
* Some popular examples of Naïve Bayes Algorithm are **spam filtration, Sentimental analysis, and classifying articles**.

KNN (K-Nearest Neighbours)

* K-Nearest Neighbour is based on Supervised Learning technique and assumes the similarity between the new case/data and available cases and put the new case into the category that is most similar to the available categories. This means when new data appears then it can be easily classified into a well suite category by using K- NN algorithm.
* It can be used for Regression but mostly it is used for the Classification problems.



K-Means

* K-Means Clustering is an unsupervised learning algorithm that it groups unlabelled dataset into different clusters. Here K defines the number of pre-defined clusters that need to be created in the process, as if K=2, there will be two clusters, and for K=3, there will be three clusters etc.
* Hence each cluster has datapoints with some commonalities, and it is away from other clusters.



Random Forest

* Random Forest is a supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of **ensemble learning,** which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model.
* Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output.
* **The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting.**

