**SOFTWARE USED**

**3.1. ANACONDA**

The developer of Anaconda is Anaconda Inc. (Continuum Analytics).it is officially released on 17 July 2012, six years ago. The stable version was released on 28 September 2018, 24 days ago. It was mainly a python development console. It operates on all kinds of platforms. It is programming language software that is mainly used in Machine learning and data science. The official website of Anaconda is [www.anaconda.com](http://www.anaconda.com). Conda install or pip install commands are given in the cmd to install the open source data packages from the Anaconda repository.

**3.1.1. ANACONDA NAVIGATOR**

Anaconda navigator is a virtual environment manager that comes with anaconda distributors and 1,000 other data packages. It eliminates the need of downloading specific library and studying about it.it is a desktop graphical user interference that permits the users to run and manage anaconda packages. It can search for packages on cloud in local repository. It is available for Windows, Linux, and Mac os.

**3.1.2. ANACONDA CLOUD**

It is a package management service where you can surf public and private notes, environments, and packages. Cloud hosts useful python packages, notes and environment for a large variation of applications. We can built new packages through CLI (Anaconda Client Command Line Interference), then it will upload it to the cloud.

**3.1.3. OTHER ANACONDA PRODUCTS**

The following applications are defaultly available in the anaconda navigator

* JupyterLab
* Jupyter Notebook
* QtConsole
* Spyder
* Glueviz
* Orange
* Rstudio

**3.2. TENSORFLOW**

It is an open source software library which is used for a wide range of tasks like dataflow modelling. It is also used in Deep learning like CNN (Convolution Neural Networks). In Google, it is also used as a source of both research and production. It is developed by Google Brain Team. It was initially released on 9 November 2015, two years ago. The stable version of the software is released on 27 September 2018, 29 days ago. It was written in Python, C++, and CUDA. The platforms on which it works are Linux, Windows, Mac os and Android. It is a Machine Learning Library. The official website of Tensorflow is [www.tensorflow.org](http://www.tensorflow.org).

**3.2.1. TENSORFLOW LITE**

The Google had announced a software specifically for Android operating systems called Tensorflow Lite starting from Android Oreo. It was officially announced on May 2017.

**3.3. CLASSIFICATION ALGORITHMS**

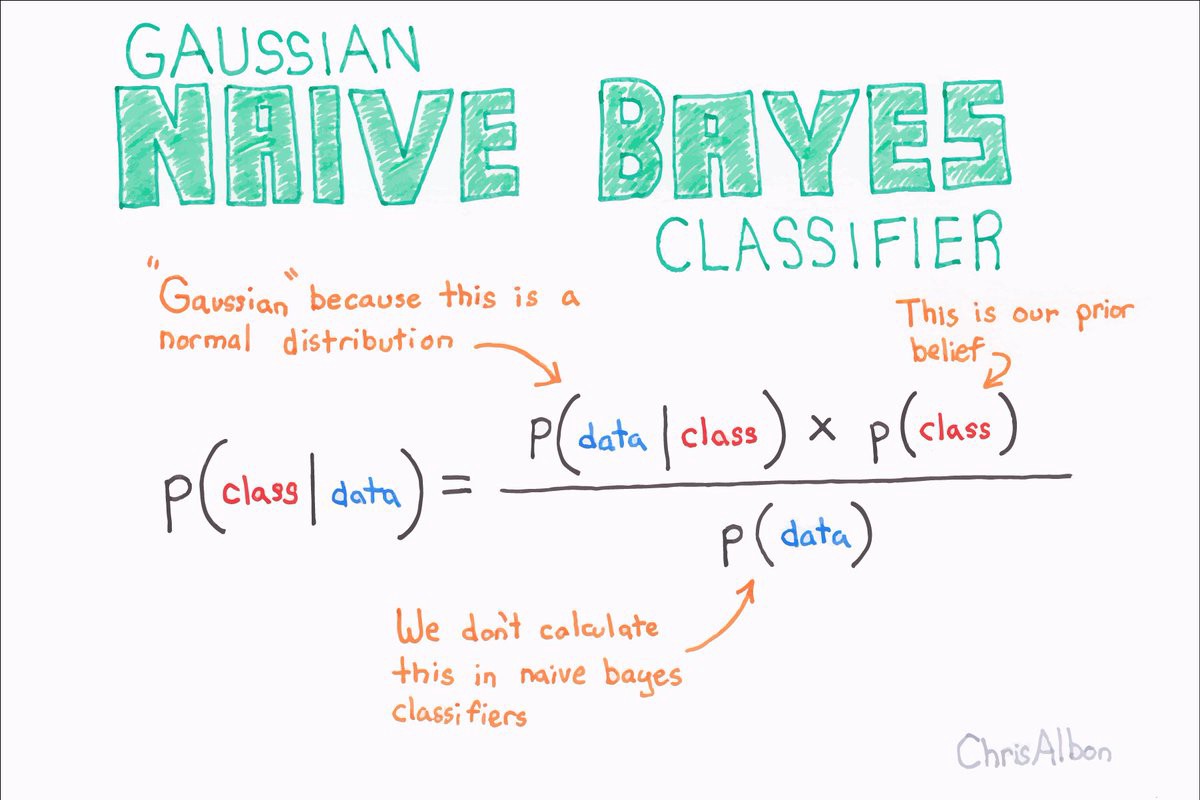
The types of classification of algorithms in Machine Learning are stated below:

Linear classifier: Logistic Regression, Naive Bayes classifier

* Support Vector Machines (SVM)
* Decision Trees
* Boosted Trees
* Random Forest
* Neural Networks
* Nearest Neighbour

**3.3.1. NAIVE BAYES CLASSIFIER**

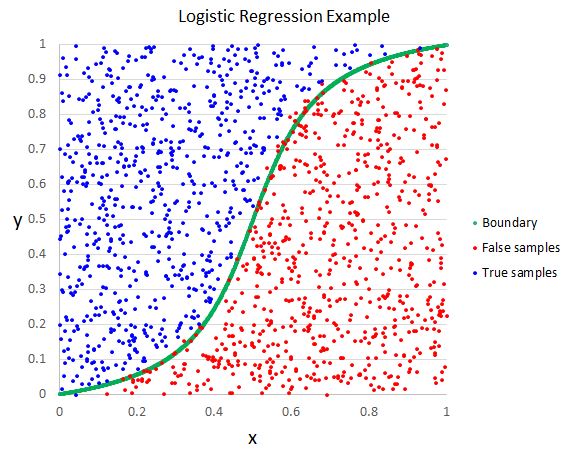
It is a classification method used depending on Bayes Theorem with an consideration of independence among predictors. In easy terms, a Naïve Bayes Classifier considers that the presence of a designated feature in a class is not related to the presence of any other feature. Even though these features depend on each other’s features, all of them have properties that independently contribute to the probability. It is easy to construct and is very useful in huge data types.



**FIG 3.1**

**3.3.2. LOGISTIC REGRESSION**

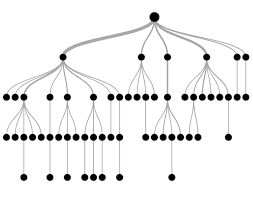
It is a statistical method for data set analyses in which it has one or more independent variables that determine the output. The output is determined with a dichotomous value (i.e., only two possible outcomes). The main aim of logistic regression is to determine the best fitting model that describes the connection between his dichotomous characteristics of interest and a pair of independent variable.



**FIG 3.2**

**3.3.3. DECISION TREE**

It builds regression or classification models that are represented in a tree form. It splits the dataset in tiny subsets while at the same time a decision tree is also increasing in size or length or sub trees. The final outcome is a tree with decision nodes and leaf nodes. A decision tree has two branches or more than two branches and a leaf node represent a decision. The top most nodes is the root node. It can handle categorical as well as numerical values.



**FIG 3.3**

**3.3.4. RANDOM FOREST**

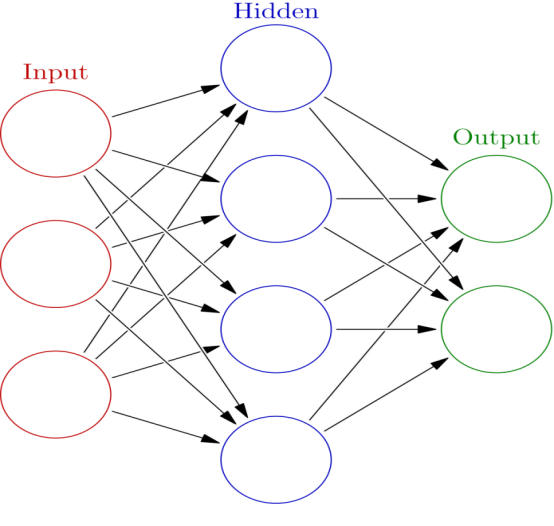
It is also called as random decision forest. They are a learning method for classification, regression and other tasks. It operates by establishing a multiple number of decision trees at training period and output the mode of the classes or means prediction of single trees.



**FIG 3.4**

**3.3.5. NEURAL NETWORKS**

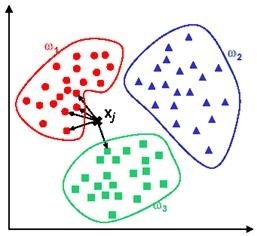
It consists of neurons, arranged in layers, which transform an input into some output. Each unit takes an input, applies a nonlinear function and then passes the output to next level. Weightage is applied to the passing signals from one to another and these are tuned in the training sector to get a neural network to the particular problem given.



**FIG 3.5**

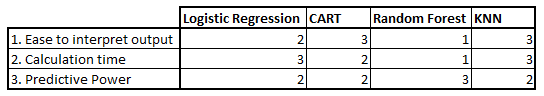
**3.3.6. NEAREST NEIGHBOUR**

The k nearest neighbour is an algorithm used for classification and in supervised learning. To label a new pointer, it looks at the next point that is closest to it (i.e. the nearest neighbour). The k in the name indicates that k is the number of neighbours it checks.



**FIG 3.6**

**3.4. COMPARISION OF EFFICIENCY OF ALGORITHMS**



**FIG 3.7**