**Linear regression**

This is a supervised learning algorithm used to predict avalues within a continuous range for example the GDP of a particular country. Linear regression relies on two independent variables to produce new insights into data and develop better predictions.

**Logistic regression**

This is another supervised learning algorithm used for binary classification. This type of algorithm can be used in tumour classification when treating cancer patients.

This type of regression stemmed from statistics and is used to predict the probability that an input can be categorized into a single primary class. In practice, however, this can be used to group outputs into one of two categories: either the input fits the description, or it does not. Hence why it is used for binary classification.

**Decision Tree**

This is a supervised learning algorithm used for classification and predictive modelling.

A decision starts with a root node, which consists of specific question of data and then sends it down a branch depending on the answer. These branches each lead to an internal node, which contains another question of the data before directing it toward another branch depending on the answer. This continues until the data reaches an end node, also called a leaf node, that doesn’t branch any further.

Decision trees are common in machine learning because they can handle complex data sets with relative simplicity.

**Support Vector Machines**

This is a supervised learning system, and this usually used for classification problems. By supplying the algorithm with labelled data, the system will output values which will act as a boundary or a range for our target data to fall into. If the value is within the range, the value will be accepted however if it does meet the criteria, the value is rejected.

**Naive Bayes**

Thisis a set of supervised learning algorithms used to create predictive models for either binary or multi-classification. Based on **Bayes’ theorem,**Naive Bayes operates on conditional probabilities, which are independent of one another but indicate the likelihood of a classification based on their combined factors.

For example, a program created to identify plants might use a naive Bayes algorithm to categorize images based on particular factors, such as perceived size, colour, and shape. While each of these factors is independent of one another, the algorithm would note the likelihood of an object being a particular plant using the combined factors.

**K-nearest neighbor (KNN) algorithm**

This is a supervised learning algorithm used for classification and predictive modelling.

True to its name, KNN algorithms classify an output by its proximity to other outputs on a graph. For example, if an output is closest to a cluster of blue points on a graph rather than a cluster of red points, then it would be classified as a member of the blue group. This approach means that KNN algorithms can be used to either classify known outcomes or predict the value of unknown ones.

### K means algorithm

Thisis an unsupervised algorithm used for classification and predictive modelling.

Much like KNN, K means uses the proximity of an output to a cluster of data points to identify it. Each of the clusters is defined by a centroid, a real or imaginary centre point for the cluster. K means is useful on large data sets, especially for clustering, though it can falter when handling outliers.