

Machine learning (ML) is a type of artificial intelligence (AI) that allows software applications to become more accurate at predicting outcomes without being explicitly programmed to do so. Machine learning algorithms use historical data as input to predict new output values.

Terms frequently used in this post:

* **Labeled data**: Data consisting of a set of training examples, where each example is a pair consisting of an input and the desired output value (also called the supervisory signal, labels, etc)
* **Classification:**The goal is to predict discrete values, e.g. {1,0}, {True, False}, {spam, not spam}.
* **Regression:**The goal is to predict continuous values, e.g. home prices.

Timeline

Description automatically generated with medium confidence

Machine learning methods:

Machine learning models fall into three primary categories.

* Supervised machine learning
* Unsupervised machine learning
* Reinforcement learning

Supervised Learning

* I like to think of supervised learning with the concept of function approximation, where basically we train an algorithm, and in the end of the process, we pick the function that best describes the input data, the one that for a given X makes the best estimation of y (X -> y). Most of the time we are not able to figure out the true function that always make the correct predictions and another reason is that the algorithm really upon an assumption made by humans about how the computer should learn, and these assumptions introduce a bias, Bias is topic I’ll explain in another post.
* Here the human experts act as the teacher where we feed the computer with training data containing the input/predictors and we show it the correct answers (output). From as the teacher, we feed the computer with training data containing the input/predictors and show it the correct answers (output). From the data, the computer should be able to learn the patterns.
* Supervised learning algorithms try to model relationships and dependencies between the target prediction output and the input features such that we can predict the output values for new data based on those relationships earned from the previous data sets.

List of Common Algorithms

1. Nearest Neighbour
2. Naive Bayes
3. Decision Trees
4. Linear Regression
5. Support Vector Machines (SVM)
6. Neural Networks
7. Nearest Neighbour:

K Nearest Neighbour is a simple algorithm that stores all the available cases and classifies the new data or case based on a similarity measure. It is mostly used to classifies a data point based on how its neighbours are classified. In this algorithm the only problem faced is finding a K value, there is no standard process to find out the best K value it’s a trial and error process.

1. Naive Bayes :

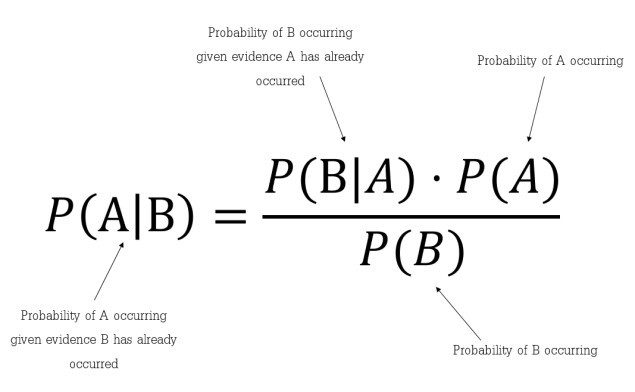
Naïve Bayes is a probabilistic machine learning algorithm based on the **Bayes Theorem**, used in a wide variety of classification tasks. In this article, we will understand the Naïve Bayes algorithm and all essential concepts so that there is no room for doubts in understanding.

# Bayes Theorem

Bayes’ Theorem is a simple mathematical formula used for calculating conditional probabilities.

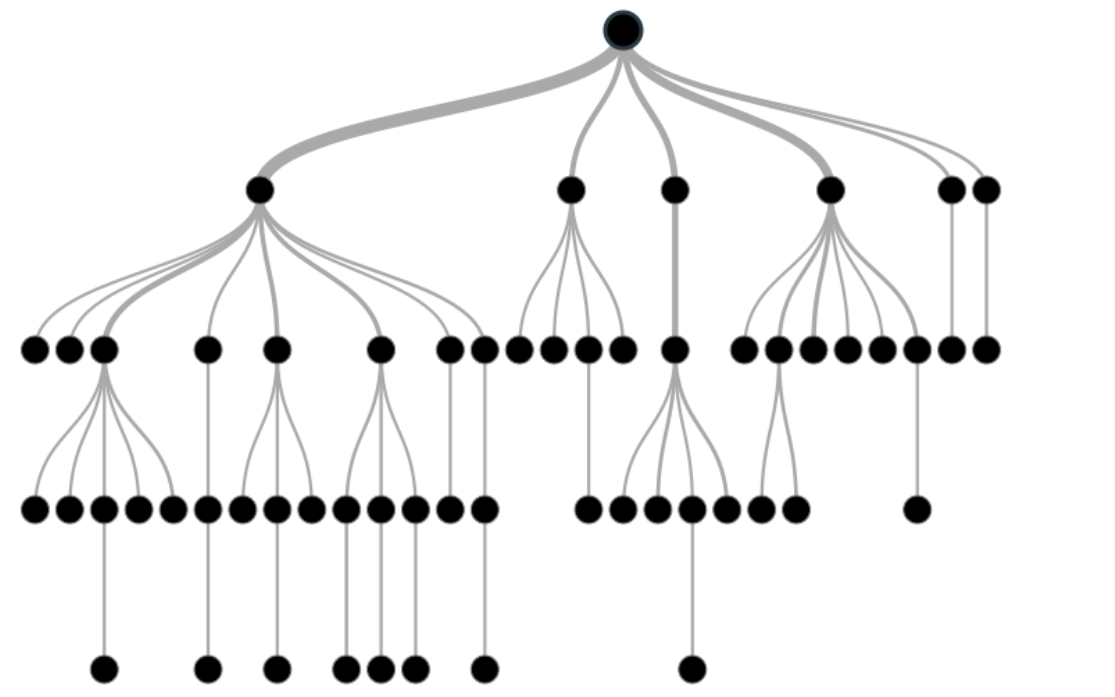
**Conditional probability** is a measure of the probability of an event occurring given that another event has (by assumption, presumption, assertion, or evidence) occurred.

The formula is: —



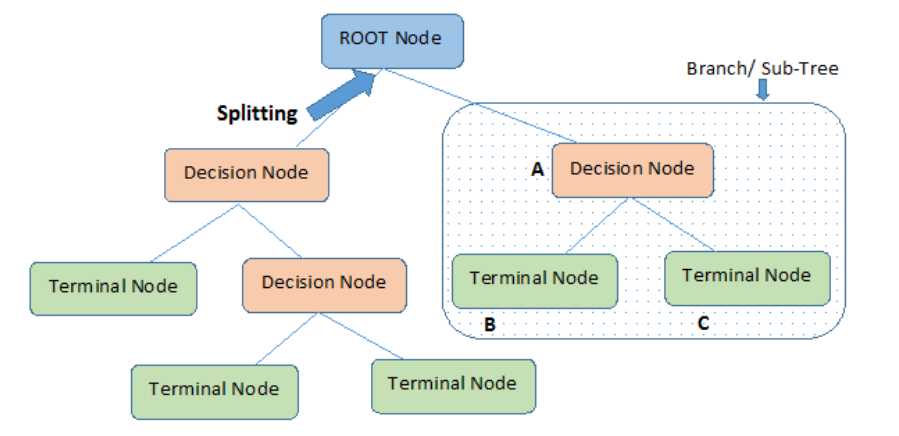
1. Decision Trees.

Decision trees can be used for classification as well as regression problems. The name itself suggests that it uses a flowchart like a tree structure to show the predictions that result from a series of feature-based splits. It starts with a root node and ends with a decision made by leaves.



Before learning more about decision trees let’s get familiar with some of the terminologies.

* ***Root Nodes***– It is the node present at the beginning of a decision tree from this node the population starts dividing according to various features.
* ***Decision Nodes*** – the nodes we get after splitting the root nodes are called Decision Node
* ***Leaf Nodes***– the nodes where further splitting is not possible are called leaf nodes or terminal nodes
* ***Sub-tree*** – just like a small portion of a graph is called sub-graph similarly a sub-section of this decision tree is called sub-tree.
* ***Pruning*** – is nothing but cutting down some nodes to stop overfitting.



1. Linear Regression:

Linear regression is a quiet and simple statistical regression method used for predictive analysis and shows the relationship between the continuous variables. Linear regression shows the linear relationship between the independent variable (X-axis) and the dependent variable (Y-axis), consequently called linear regression*. If there is a single input variable (x), such linear regression is called****simple linear regression****. And if there is more than one input variable, such linear regression is called****multiple linear regression*.** The linear regression model gives a sloped straight line describing the relationship within the variables.