Introduction

Machine Learning is a branch of artificial intelligence that allows software applications to predict outcomes without being explicitly programmed. A machine learning algorithm should learn the patterns of the input data to train a model which can later be used for prediction on test data. The important components of machine learning are Data and algorithm. The better the data is pre-processed the better will be the performance. Broadly, there are 3 types of Machine Learning algorithms: Supervised, Unsupervised and Reinforcement Learning.

Supervised and unsupervised algorithms briefly include:

Supervised Learning Algorithm: A supervised learning algorithm acquires information from the training dataset and applies it on test datasets for predictions. It includes input variables and output variables. The algorithm learns the mapping function from the input to the output.

This mapping function is applied on new input data that can predict the output variables for data. The input and output variables are also called independent and dependent variables respectively.

Supervised learning problems can be further grouped into regression and classification problems.

Classification: A classification problem occurs when an output variable is a category. Some of the popular algorithms include logistic regression, decision tree, random forest and Support Vector Machines (SVM).

Regression: A regression problem occurs when the output variable is a real value. Some of the popular regression algorithms include linear and non-linear regressions.

Unsupervised Learning Algorithm: Unsupervised learning is where there is only input data and no corresponding output variables.

The goal for unsupervised learning is to extract the underlying structure in the data to learn more about the data. Unsupervised learning problems can be further grouped into clustering and association problems.

Clustering: A clustering problem discovers the inherent groupings in data. k-means is a popular clustering algorithm.

Association: An association rule learning problem discovers rules that describe large portions of the data. Apriori algorithm for association rule is very popular.

Machine Learning has transformed the way data extraction and interpretation are done by automating algorithms, thereby replacing traditional statistical techniques. The prerequisites to jump into this field are basic linear algebra, calculus and statistics. Machine Learning have huge applications in the domain of the financial sector, healthcare sector, retail sector and transportation sector.