

MACHINE LEARNING

By Amirthavarshini V

INTRODUCTION:

Machine learning (ML) is a branch of artificial intelligence (AI) that enables computers to “self-learn” from training data and improve over time. Machine learning algorithms are able to detect patterns in data and learn from them, in order to make their own predictions. ML algorithms, make calculations, process data, and identify patterns automatically in the training data and apply them the observations on the test data provided to produce desired outcomes. It gives computers the ability to learn.

TERMINOLOGY:

- 1) DATA: It can be any unprocessed fact, value, text, sound, or picture that is not being interpreted and analyzed. Data is the most important part of all Data Analytics, Machine Learning and Artificial Intelligence.
- 2) DATASETS: A dataset in machine learning is, quite simply, a collection of data pieces that can be treated by a computer as a single unit for analytic and prediction purposes.
- 3) TRAIN AND TEST DATA: Train/Test is a method to measure the accuracy of your model. It is called Train/Test because the data set is split into two sets: a training set and a testing set. 80% for training, and 20% for testing. The model is trained using train data and tested using test data.
- 4) ALGORITHMS: A machine learning algorithm is the method by which the AI system conducts its task, generally predicting output values from given input data. The two main processes of machine learning algorithms are classification and regression.
- 5) CLASSIFICATION: classification is the problem of identifying which of a set of categories an observation belongs to.
- 6) REGRESSION: regression is a machine learning technique where the model predicts the output as a continuous numerical value.

TRAIN – TEST PROCESS:



TYPES OF MACHINE LEARNING ALGORITHMS:

- 1) **SUPERVISED:** In supervised learning, the machine is taught by example. The user provides the machine learning algorithm with a known dataset that includes desired inputs and outputs, and the algorithm must find a method to determine how to arrive at those inputs and outputs. It involves classification, regression and forecasting.
- 2) **SEMI - SUPERVISED:** Semi-supervised learning is similar to supervised learning, but instead uses both labelled and unlabelled data. Labelled data is information that has meaningful tags so that the algorithm can understand the data, whilst unlabelled data lacks that information. By using this combination, machine learning algorithms can learn to label unlabelled data.
- 3) **UNSUPERVISED:** Here, the machine learning algorithm studies data to identify patterns. There is no answer key or human operator to provide instruction. Instead, the machine determines the correlations and relationships by analysing available data. The algorithm tries to organise data in some way to describe its structure. This might mean grouping the data into clusters or arranging it in a way that looks more organised.
- 4) **REINFORCEMENT LEARNING:** Reinforcement learning focuses on regimented learning processes, where a machine learning algorithm is provided with a set of actions, parameters and end values. By defining the rules, the machine learning algorithm then tries to explore different options and possibilities, monitoring and evaluating each result to determine which one is optimal. Reinforcement learning teaches the machine trial and error. It learns from past experiences and begins to adapt its approach in response to the situation to achieve the best possible result.

APPLICATIONS:

Some real life applications of Machine learning algorithms include:

- Virtual personal assistants: like Siri and Alexa
- Predictions: like traffic and price predictions by like Ola and uber.
- Surveillance: UAVs, drones
- Social media services like Youtube's recommendation algorithm
- Retail support: Customer services by virtual assistants
- Search engine optimization
- Product recommendations in online retail stores
- Self driving cars
- Image recognition