INTRODUCTION TO MACHINE LEARNING



TABLE OF CONTENTS:

1.

Arthur Samuel coined the term Machine Learning in 1959. He defined it as “The field of study that gives computers the capability to learn without being explicitly programmed”. Machine Learning is a subset of Artificial Intelligence and it allows machines to learn from their experiences without any coding.

There are basically several types of machine learning. Based on the supervision needed, machine learning is divided into 4 types. They are,

1. **Supervised Learning:**

Supervised learning is basically learning from data to make prediction. In this type of supervision, machines are trained on labelled datasets and enabled to predict outputs based on the provided training.

It has two types,

**Classification:**

In classification, the output expected is categorical, for example, yes or no, true or false, male or female, etc**.**

**Regression:**

Regression trains on and predicts a continuous-valued response, for example predicting real estate prices. The model is typically represented by an equation that describes the relationship between the independent variables and the dependent variable, and it is often used to estimate the strength and direction of the relationship between variables.

Some of the supervised learning algorithms include random forest, decision tree, naïve bayes, SVM, Linear Regression

1. **Unsupervised Learning:**

Unsupervised learning is different from the Supervised learning technique; as its name suggests, there is no need for supervision. It means, in unsupervised machine learning, the machine is trained using the unlabelled dataset, and the machine predicts the output without any supervision. The main aim of the unsupervised learning algorithm is to group or categories the unsorted dataset according to the similarities, patterns, and differences.

Some of the popular clustering algorithms are given below:

K-Means Clustering algorithm

DBSCAN Algorithm

Principal Component Analysis

1. **Semi Supervised Learning:**

Semi-Supervised learning is a type of Machine Learning algorithm that lies between Supervised and Unsupervised machine learning. It represents the intermediate ground between Supervised (With Labelled training data) and Unsupervised learning (with no labelled training data) algorithms and uses the combination of labelled and unlabelled datasets during the training period.

1. **Reinforcement Learning:**

Reinforcement learning works on a feedback-based process, in which an AI agent (A software component) automatically explore its surrounding by hitting & trail, taking action, learning from experiences, and improving its performance. Agent gets rewarded for each good action and get punished for each bad action; hence the goal of reinforcement learning agent is to maximize the rewards.

There are many techniques that are described as types of learning. This includes multi-task, active, online, transfer, and ensemble learning.

Multi-Task Learning

Multi-task learning is a type of supervised learning that involves fitting a model on one dataset that addresses multiple related problems. It involves devising a model that can be trained on multiple related tasks in such a way that the performance of the model is improved by training across the tasks as compared to being trained on any single task.

Active learning is a type of supervised learning and seeks to achieve the same or better performance of so-called “passive” supervised learning, although by being more efficient about what data is collected or used by the model. The key idea behind active learning is that a machine learning algorithm can achieve greater accuracy with fewer training labels if it is allowed to choose the data from which it learns. An active learner may pose queries, usually in the form of unlabeled data instances to be labeled by an oracle (e.g., a human annotator).

Transfer learning is a type of learning where a model is first trained on one task, then some or all of the model is used as the starting point for a related task. It is a useful approach on problems where there is a task related to the main task of interest and the related task has a large amount of data.

1. **Online Learning**:

Online learning involves using the data available and updating the model directly before a prediction is required or after the last observation was made.

Traditionally machine learning is performed offline, which means we have a batch of data, and we optimize an equation […] However, if we have streaming data, we need to perform online learning, so we can update our estimates as each new data point arrives rather than waiting until “the end” (which may never occur).

1. **Ensemble learning:**

Ensemble learning is an approach where two or more modes are fit on the same data and the predictions from each model are combined.

There are three main classes of ensemble learning, they are,

Bagging, Boosting, Stacking.