The Three Types of Machine Learning

Broadly speaking, Machine Learning algorithms are of three types- Supervised Learning, Unsupervised Learning, and Reinforcement Learning.

A) *Supervised Learning*

We provide the machine with a ton of information about a case and also provide it with the case outcome. The outcome is called the labelled data while the rest of the information is used as input features. The machine is supervised by the labelled data to learn what the relationships and dependencies are between the outcome and the input.

Models are fit on training data comprised of inputs and outputs and used to make predictions on test sets where only the inputs are provided and the outputs from the model are compared to the withheld target variables (withheld output/labels) and used to estimate the skill of the model.

There are two main types of supervised learning problems: they are classification that involves predicting a class label and regression that involves predicting a numerical value.

- **Classification**: Supervised learning problem that involves predicting a class label.
- **Regression**: Supervised learning problem that involves predicting a numerical label.

Both classification and regression problems may have one or more input variables and input variables may be any data type, such as numerical or categorical.

Common supervised learning algorithms include: *Linear regression; Naïve Bayes, Nearest Neighbours, Decision Trees, Support Vector Machines and Neural Networks.*

B) *Unsupervised Learning*

As the name suggests, in case of unsupervised learning, there is no help from the user for the computer to learn. In the lack of labelled training sets, the machine identifies patterns in the data that is not so obvious to the human eye. So, unsupervised learning is extremely useful to recognize patterns in data and help us take decisions. Unsupervised learning is often used for anomaly detection, like to uncover fraudulent transactions or payments.

What makes unsupervised learning such an interesting area is that an overwhelming majority of data in this world is unlabelled. Having intelligent algorithms that can take our terabytes and terabytes of unlabelled data and make sense of it is a huge source of potential profit for many industries. Because unsupervised learning is based upon the data and its properties, we can say that unsupervised learning is data-driven. The outcomes from an unsupervised learning task are controlled by the data and the way it's formatted.

There are many types of unsupervised learning, although there are two main problems that are often encountered by a practitioner: they are clustering that involves finding groups in the data and density estimation that involves summarizing the distribution of data.

- **Clustering:** Unsupervised learning problem that involves finding groups in data.
- **Density Estimation**: Unsupervised learning problem that involves summarizing the distribution of data.

An example of a clustering algorithm is k-Means where *k* refers to the number of clusters to discover in the data. An example of a density estimation algorithm is Kernel Density Estimation that involves using small groups of closely related data samples to estimate the distribution for new points in the problem space.

The most common use of unsupervised learning is in clustering problems, with the most talked about algorithms being *k-means and hierarchical clustering*, though other algorithms like *Hidden Markov models, Self-Organizing Maps or Gaussian Mixture models* are also often used.

C) *Reinforcement Learning*

Reinforcement learning is probably the closest to how we humans learn. In this case, the algorithm or the agent learns continually from its environment by interacting with it. It gets a positive or a negative reward based on its action.

For any reinforcement learning problem, we need an agent and an environment as well as a way to connect the two through a feedback loop. To connect the agent to the environment, we give it a set of actions that it can take that affect the environment. To connect the environment to the agent, we have it continually issue two signals to the agent: an updated state and a reward (our reinforcement signal for behavior).

The use of an environment means that there is no fixed training dataset, rather a goal or set of goals that an agent is required to achieve, actions they may perform, and feedback about performance toward the goal.

▼ Related Course/Tutorial Name

- Machine Learning: Algorithms in the Real World Specialization, C01 - Introduction to Applied Machine Learning

▼ Related Blogs & Articles

- https://pioneerlabs.io/insights/the-three-types-of-machine-learning-algorithms/
- https://machinelearningmastery.com/types-of-learning-in-machine-learning/
- https://towardsdatascience.com/what-are-the-types-of-machine-learning-e2b9e5d1756f