**Types of Machine Learning Algorithms You Should Know**

I particularly think that getting to know the types of Machine learning algorithms is like getting to see the Big Picture of AI and what is the goal of all the things that are being done in the field and put you in a better position to break down a real problem and design a machine learning system.

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 a 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.

Types of machine learning Algorithms

There some variations of how to define the types of Machine Learning Algorithms but commonly they can be divided into categories according to their purpose and the main categories are the following:

* **Supervised learning**
* **Unsupervised Learning**
* **Semi-supervised 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 other reason is that the algorithm rely upon an assumption made by humans about how the computer should learn and this assumptions introduce a bias, Bias is topic I’ll explain in another post.
* Here the human experts acts as the teacher where we feed the computer with training data containing the input/predictors and we show it the correct answers (output) and 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 which it learned from the previous data sets.

Draft

* Predictive Model
* we have labeled data
* The main types of supervised learning problems include regression and classification problems

List of Common Algorithms

* Nearest Neighbor
* Naive Bayes
* Decision Trees
* Linear Regression
* Support Vector Machines (SVM)
* Neural Networks

Unsupervised Learning

* The computer is trained with unlabeled data.
* Here there’s no teacher at all, actually the computer might be able to teach you new things after it learns patterns in data, these algorithms a particularly useful in cases where the human expert doesn’t know what to look for in the data.
* are the family of machine learning algorithms which are mainly used in *pattern detection* and *descriptive modeling*. However, *there are no output categories or labels* here based on which the algorithm can try to model relationships. These algorithms try to use techniques on the input data to *mine for rules*, *detect patterns*, and *summarize and group the data points* which help in deriving meaningful insights and describe the data better to the users.

Draft

* Descriptive Model
* The main types of unsupervised learning algorithms include *Clustering algorithms and Association rule learning algorithms.*

List of Common Algorithms

* k-means clustering, Association Rules

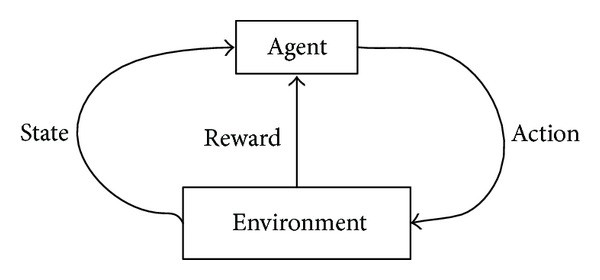
Semi-supervised Learning

In the previous two types, either there are no labels for all the observation in the dataset or labels are present for all the observations. Semi-supervised learning falls in between these two. In many practical situations, the cost to label is quite high, since it requires skilled human experts to do that. So, in the absence of labels in the majority of the observations but present in few, semi-supervised algorithms are the best candidates for the model building. These methods exploit the idea that even though the group memberships of the unlabeled data are unknown, this data carries important information about the group parameters.

Reinforcement Learning

method aims at using observations gathered from the interaction with the environment to take actions that would maximize the reward or minimize the risk. Reinforcement learning algorithm (called the agent) continuously learns from the environment in an iterative fashion. In the process, the agent learns from its experiences of the environment until it explores the full range of possible states.

**R**einforcement Learning is a type of *Machine Learning*, and thereby also a branch of *Artificial Intelligence*. It allows machines and software agents to automatically determine the ideal behavior within a specific context, in order to maximize its performance. Simple reward feedback is required for the agent to learn its behavior; this is known as the reinforcement signal.



**T**here are many different algorithms that tackle this issue. As a matter of fact, Reinforcement Learning is defined by a specific type of problem, and all its solutions are classed as Reinforcement Learning algorithms. In the problem, an agent is supposed decide the best action to select based on his current state. When this step is repeated, the problem is known as a *Markov Decision Process*.

In order to produce intelligent programs (also called agents), reinforcement learning goes through the following steps:

1. Input state is observed by the agent.
2. Decision making function is used to make the agent perform an action.
3. After the action is performed, the agent receives reward or reinforcement from the environment.
4. The state-action pair information about the reward is stored.

List of Common Algorithms

* Q-Learning
* Temporal Difference (TD)
* Deep Adversarial Networks

**Algorithm Foundations**

1: How to Talk About Data in Machine Learning

2: Principle That Underpins All Algorithms

3: Parametric and Nonparametric Algorithms

4: Bias, Variance and the Trade-off

**Linear Algorithms**

Linear Regression  
Logistic Regression  
Linear Discriminant Analysis

**Nonlinear Algorithms**

Classification and Regression Trees

Naive Bayes

k-Nearest Neighbors

Learning Vector Quantization

Support Vector Machines

**Ensemble Algorithms**

Bagging and Random Forest

Boosting and AdaBoost