Machine Learning: A Very General Guide

Nicholas P. Ross

March 6, 2017

Abstract

This is a simple document which will make some v. general notes on things connected to "Machine Learning". This document can be found at:
`https://github.com/d80b2t/Research_Notes/tree/master/MachineLearning`.

1 Introduction

From Wikipedia, retrieved, Mon Dec 5 16:28:49 PST 2016:

Machine learning is the subfield of computer science that "gives computers the ability to learn without being explicitly programmed" (Arthur Samuel, 1959). Evolved from the study of pattern recognition and computational learning theory in artificial intelligence, machine learning explores the study and construction of algorithms that can learn from and make predictions on data such algorithms overcome following strictly static program instructions by making data driven predictions or decisions, through building a model from sample inputs. Machine learning is employed in a range of computing tasks where designing and programming explicit algorithms is unfeasible; example applications include spam filtering, detection of network intruders or malicious insiders working towards a data breach, optical character recognition (OCR), search engines and computer vision.

Broadly, there are 3 types of Machine Learning Algorithms:
1. Supervised Learning
2. Unsupervised Learning
3. Reinforcement Learning

[1] *Supervised learning* is the machine learning task of inferring a function from labeled training data.[1] The training data consist of a set of training examples. In supervised learning, each example is a pair consisting of an input object (typically a vector) and a desired output value (also called the supervisory signal). A supervised learning algorithm analyzes the training data and produces an inferred function, which can be used for mapping new examples. An optimal scenario will allow for the algorithm to correctly determine the class labels for unseen instances. This requires the learning algorithm to generalize from the training data to unseen situations in a "reasonable" way (see inductive bias).

The parallel task in human and animal psychology is often referred to as concept learning.

[2] *Unsupervised machine* learning is the machine learning task of inferring a function to describe hidden structure from unlabeled data. Since the examples given to the learner are unlabeled, there is no error or reward signal to evaluate a potential solution this distinguishes unsupervised learning from supervised learning and reinforcement learning.

Unsupervised learning is closely related to the problem of density estimation in statistics.[1] However, unsupervised learning also encompasses many other techniques that seek to summarize and explain key features of the data. Approaches to unsupervised learning include: clustering (k-means; mixture models;

[1] `https://en.wikipedia.org/wiki/Supervised_learning`
[2] `https://en.wikipedia.org/wiki/Unsupervised_learning`

hierarchical clustering); anomaly detection; Neural Networks.

[3] *Reinforcement learning* is an area of machine learning inspired by behaviorist psychology, concerned with how software agents ought to take actions in an environment so as to maximize some notion of cumulative reward. The problem, due to its generality, is studied in many other disciplines, such as game theory, control theory, operations research, information theory, simulation-based optimization, multi-agent systems, swarm intelligence, statistics, and genetic algorithms. In the operations research and control literature, the field where reinforcement learning methods are studied is called approximate dynamic programming. The problem has been studied in the theory of optimal control, though most studies are concerned with the existence of optimal solutions and their characterization, and not with the learning or approximation aspects. In economics and game theory, reinforcement learning may be used to explain how equilibrium may arise under bounded rationality.

In machine learning, the environment is typically formulated as a Markov decision process (MDP) as many reinforcement learning algorithms for this context utilize dynamic programming techniques. The main difference between the classical techniques and reinforcement learning algorithms is that the latter do not need knowledge about the MDP and they target large MDPs where exact methods become infeasible.

"The 10 Algorithms Machine Learning Engineers Need to Know"

`http://www.kdnuggets.com/2016/08/10-algorithms-machine-learning-engineers.html`

Supervised Learning:

- Decision Trees
- Nave Bayes Classification
-
-
-
-

Unsupervised Learning:

- Clustering Algorithms:
- Principal Component Analysis:
- Singular Value Decomposition:
- Independent Component Analysis:

And techniques/methods of ML:

- Linear Regression
- K-means
- Decision Trees
- Random Forest
- PCA
- SVM: `Concise technical overview`.
- Artificial Neural Networks (ANN)

With links from Machine Learning Algorithms: A Concise Technical Overview, Matthew Mayo.

[3] `https://en.wikipedia.org/wiki/Reinforcement_learning`

2 Deep Learning

3 Useful URLs and References

http://www.zdnet.com/topic/how-to-implement-ai-and-machine-learning/