

---

---

# Machine Learning and AI

- Methods and Algorithms -

---

---

Personnal Notes  
François Bouvier d'Yvoire

CentraleSupélec & Imperial College  
Current Branch : LogicLearning  
Commit : 0997c3f2b0d5d05f4c8a2b7cebc16e2ad1998500

# Contents

<b>1</b>	<b>Logic-Based Learning</b>	<b>2</b>
1.1	Inductive Logic Programming (ILP) . . . . .	2
1.1.1	Predictive ILP . . . . .	3

# Todo list

 add ref to lecturer . . . . .	2
---	---

# Intro

This document will use the following classification for the machine learning algorithms. However their might be some changes. For exemple, some of them will be part of the commons algorithms and not from their real class.

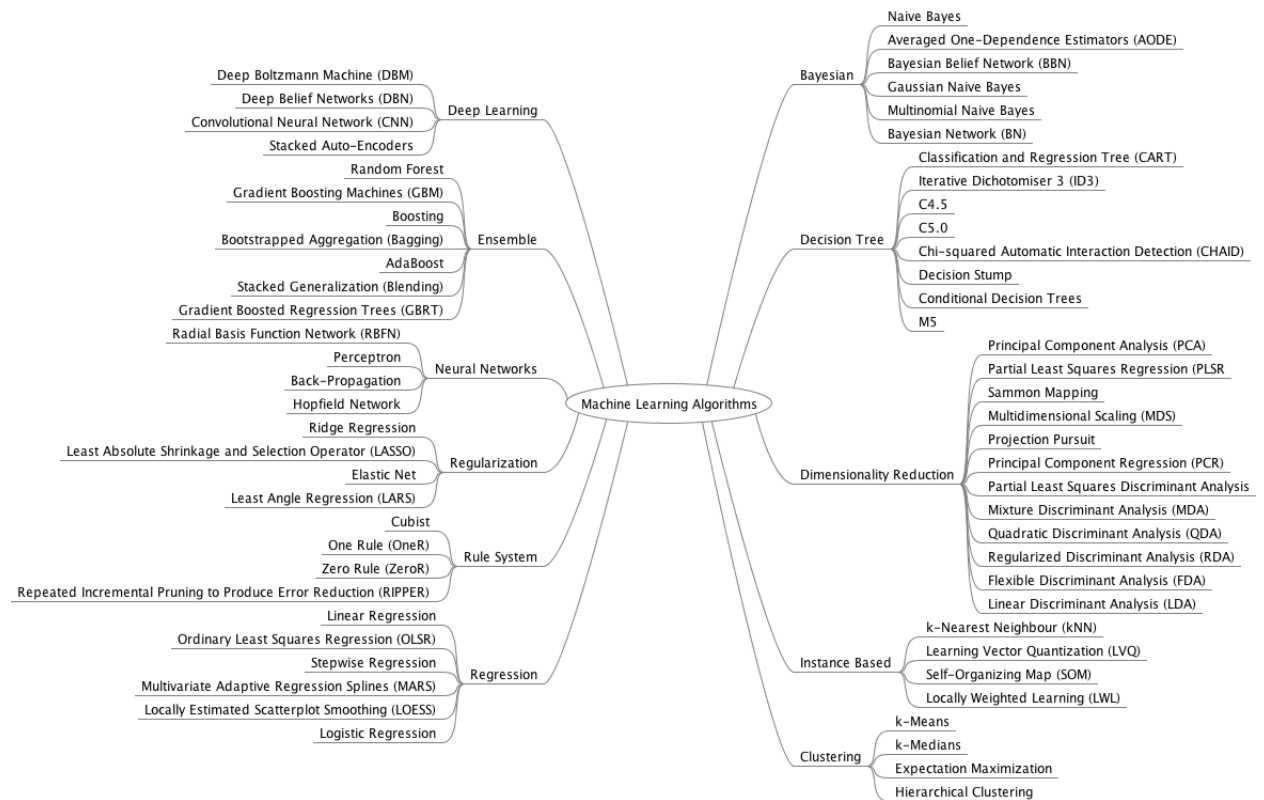


Figure 1 – Simple graph for algorithms classification in ML

# Chapter 1

## Logic-Based Learning

add ref to  
lecturer

### 1.1 Inductive Logic Programming (ILP)

**Introduction to Concept Learning** In Concept Learning we aim to compute a definition of a concept, expressed in a given language (called **hypothesis space**) that satisfies positive example and none of negatives ones. As an example, it can be a regexp which can match all words of a set, and none of an other set (see Regexp Golf)

**Machine Learning Task** The Inductive Logic Programming is a subset of Machine Learning where prior knowledge are expressed in declarative language. The task is then a search problem for an hypothesis that would minimise a loss function.

Given :

- A language of examples  $L_e$
- a language of hypotheses  $L_h$
- an unknown target function  $f : L_e \rightarrow Y$
- a set  $E$  of training examples  $E = \{(e_i, f(e_i))\}$
- a loss function

We want to find the hypothesis  $h \in L_h$  that minimises the loss function ( $h = \arg \min_{h_j \in L_h} \text{loss}(h_j, E)$ ). We want the hypothesis  $h$  to approximate as much as possible the function  $f$

Different loss can be choosed, such as  $l(h, E) = \frac{1}{|E|} \sum (f(e_i) - h(e_i))$  or the squared differenced

**Data Mining Task** In a Data Mining Task, the objective is also to discover hypothesis, but the loss is replaced by a quality criterion  $Q(h, E)$  such that the search is now to find  $h$  such as  $Q(h, E) = \frac{|c(h, E)|}{|E|} \geq \epsilon$ . This expresses a notion of coverage. This is essential for the concept learning, as we refer generally to the set of example covered by an hypothesis  $h$

### 1.1.1 Predictive ILP

Given

- A set of observation in  $L_e$  with positive examples  $E_+$  and negative examples  $E_-$