# Machine Learning and AI

- Methods and Algorithms -

Personnal Notes François Bouvier d'Yvoire

CentraleSupélec & Imperial College

# **Contents**

1	Con	nmon N	fachine Learning algorithms	3
	1.1	Linear	Regression	3
		1.1.1	Maximum Likelihood Estimation (MLE)	3
	1.2	Gradi	ent Descent	3
		1.2.1	Simple Gradient Descent	3
		1.2.2	Gradient Descent with Momentum	3
		1.2.3	Stochastic Gradient Descent	3
	1.3	Mode	Selection and Validation	3
		1.3.1	Cross-Validation	3
		1.3.2	Marginal Likelihood	3
	1.4	Bayesi	an Linear Regression	3
		1.4.1	Mean and Variance	3
		1.4.2	Sample function	3
2	Reir	nforcen	nent Learning	5
	2.1	Marko	ov Reward and Decision Process	5
		2.1.1	State Value Function Closed-form	5
		2.1.2	Iterative Policy Evaluation Algorithm	5
	2.2	Dynar	nic Programming in RL	5
		2.2.1	Policy Iteration Algorithm	5
		2.2.2		5
		2.2.3	Assynchronous Backup in RL	5
		2.2.4	Properties and drawbacks of Dynamic Programming	5
	2.3	Mode	-Free Learning	5
		2.3.1	Monte-Carlo Algorithms	5
		2.3.2	Monte-Carlo Control Algorithms	6
		2.3.3	Temporal Difference Learning	6
		2.3.4	Temporal Difference Learning Control Algorithm	6
	2.4	Reinfo	orcement Learning with Function Approximation	6
		2.4.1	Exemple of features	6
		2.4.2	Monte-Carlo with Value Function Approximation	7

iv

	2.4.3	Temporal Difference Learning with Value Function Approxi-
		mation
	2.4.4	Q-Learning with FA
	2.4.5	subsection name
2.5	Deep :	Learning Reinforcement Learning
	2.5.1	Experience Replay
	2.5.2	Target Network
	2.5.3	Clipping of Rewards
	2.5.4	Skipping of Frames

# **Todo list**

Add bibtex reference	3
Find better paragraph layout	3
Add "you cannot backup death" explanations	5
Don't forget Starting to explore	6
Add Comparison between MC and TD learning	6

## 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.

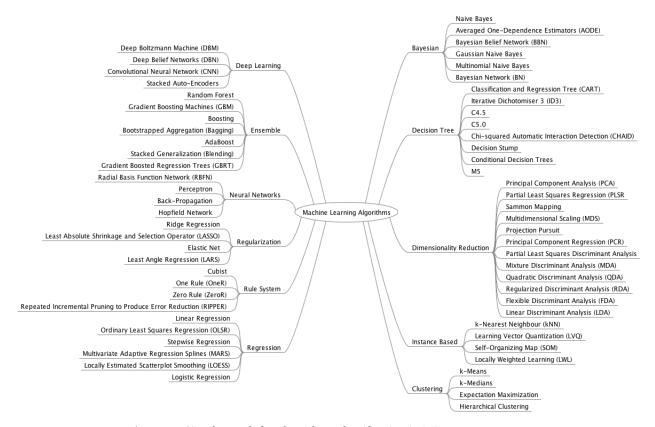


Figure 1 – Simple graph for algorithms classification in ML

## Chapter 1

# Common Machine Learning algorithms

This chapter is dedicated to the most common ML algorithms, a major part of the notes come from the mml-books.com

Find better paragraph layout

Add bibtex reference

### 1.1 Linear Regression

1.1.1 Maximum Likelihood Estimation (MLE)

**Closed-Form Solution** 

Maximum A Posteriori Estimation (MAP)

- 1.2 Gradient Descent
- 1.2.1 Simple Gradient Descent
- 1.2.2 Gradient Descent with Momentum
- 1.2.3 Stochastic Gradient Descent
- 1.3 Model Selection and Validation
- 1.3.1 Cross-Validation
- 1.3.2 Marginal Likelihood
- 1.4 Bayesian Linear Regression
- 1.4.1 Mean and Variance
- 1.4.2 Sample function

## Chapter 2

# **Reinforcement Learning**

$\sim$	-	3.7 1	<b>D</b> 1	1	<b>D</b> • •	<b>T</b>
٠,		Markov	RATIONA	วทศ	Decision	Proces
<b>∠</b> .		IVIAINUV	newaiu	anu	Decision	110003

- 2.1.1 State Value Function Closed-form
- 2.1.2 Iterative Policy Evaluation Algorithm
- 2.2 Dynamic Programming in RL
- 2.2.1 Policy Iteration Algorithm
- 2.2.2 Value Iteration Algorithm
- 2.2.3 Assynchronous Backup in RL

**Prioritised Sweeping** 

**Real-time Dynamic Programming** 

- 2.2.4 Properties and drawbacks of Dynamic Programming
- 2.3 Model-Free Learning
- 2.3.1 Monte-Carlo Algorithms

(First Visit) Monte-Carlo Policy Evaluation

**Every Visit Monte-Carlo Policy Evaluation** 

Add "you cannot backup death" explanations

**Batch vs Online Monte-Carlo** 

Incremental Monte-Carlo Update

Runing Mean for Non-Stationnary World

### 2.3.2 Monte-Carlo Control Algorithms

**Monte-Carlo Policy Improvement** 

**Greedy Policy Improvement over State Value Function** 

**Greed Policy Improvement over State-Action Value Function** 

Don't forget Starting to explore

**Exploring Starts Problem** 

On-Policy  $\epsilon$ -greedy first-visit Monte-Carlo control Algorithm

Monte-Carlo Batch Learning to Control

Monte-Carlo Iterative Learning to Control

#### 2.3.3 Temporal Difference Learning

Add Comparison between MC and TD learning

Temporal Difference Value Function Estimation Algorithm

#### 2.3.4 Temporal Difference Learning Control Algorithm

SARSA - On Policy learning Temporal Difference Control

SARSA-Lambda

Hindsight Experience Replay

Q-Learning: Off-Policy Temporal Difference Learning

## 2.4 Reinforcement Learning with Function Approximation

### 2.4.1 Exemple of features

**Coarse Coding** 

**Tile Coding** 

**Radial-Basis Function** 

### Deep Learning

- 2.4.2 Monte-Carlo with Value Function Approximation
- 2.4.3 Temporal Difference Learning with Value Function Approximation
- 2.4.4 Q-Learning with FA
- 2.4.5 subsection name
- 2.5 Deep Learning Reinforcement Learning
- 2.5.1 Experience Replay
- 2.5.2 Target Network
- 2.5.3 Clipping of Rewards
- 2.5.4 Skipping of Frames