

Introduction

Lille 3

Rapide
historique

Motivation

Régression

Classification

Clustering

Introduction

L'apprentissage automatique (2016-2017)

UFR MIME

Université Lille 3

7 décembre 2016

Sommaire

Introduction

Lille 3

Rapide
historique

Motivation

Régression

Classification

Clustering

1 Rapide historique

2 Motivation

- Régression
- Classification
- Clustering

Intelligence Artificielle

Introduction

Lille 3

Rapide
historique

Motivation

Régression
Classification
Clustering

Intelligence :

- Capacité de percevoir l'information
- Transformer en connaissance
- Appliqué dans un environnement ou un contexte spécifique

capacité à prendre de “bonnes” décisions étant donné un certain critère.

Intelligence Artificielle : (terme attribué à Minsky 1956)

Programmation des ordinateurs pour qu'ils réalisent des tâches qui nécessitent de l'intelligence lorsqu'elles sont réalisés par des êtres humains

Apprentissage automatique : une certaine forme d'IA qui privilégie la **généralisation à partir d'exemples**

Intelligence Artificielle

Introduction

Lille 3

Rapide historique

Motivation

Régression

Classification

Clustering

Considérer les tâches suivantes :

- Voiture autonome (Tesla, Google, BMW, Citroen...)
- Détection d'objets dans une image(scence)
- Traduction automatique de texte (google translate)
- Jeux d'échecs ou Go (deep blue, alpha-go)



Intelligence Artificielle

Introduction

Lille 3

Rapide
historique

Motivation

Régression
Classification
Clustering

Pour un ordinateur :

- Le problème est trop complexe compte tenu de plusieurs observations, estimations et décisions à faire
- Pourtant, nous savons conduire, identifier des visages, transcrire la voix, jouer aux échecs.

Points communs :

- On a eu des exemples et demos pour achever des tâches
- Instruction et exemples donner par dirigeant/parent
- Apprentissage par notre expérience

L'apprentissage consiste à construire un modèle à partir des exemples.

Intelligence artificielle aujourd'hui

Introduction

Lille 3

Rapide
historique

Motivation

Régression

Classification

Clustering

IA Classique : Symboliste

- Symbolique
- Système à la base des règles logiques construite à la main
- Succès moderne : Deep-Blue a battu Kasparov aux échecs en 1996
- Etudier **l'incertain** : Modèles graphiques probabilistes, d'inférence grammaticale probabiliste

Connexioniste

- Sub-symbolique : phénomènes peuvent être décrits à l'aide de réseaux d'unités simples interconnectées
- Origine des sciences cognitives
- Succès moderne : Apprentissage profond (deep learning)
- Apprentissage **statistique** (Statistical Machine Learning)

Panorama

Référence pour lire (pas de math) : Master Algorithm par Pedro Domingos

Introduction

Lille 3

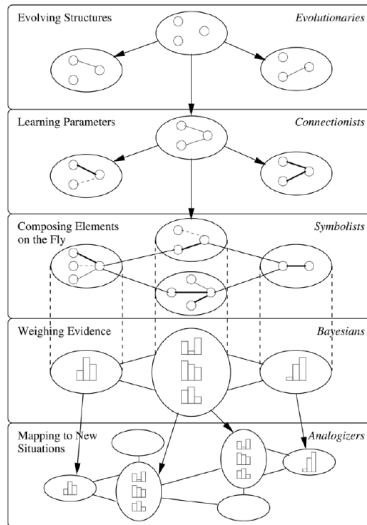
Rapide historique

Motivation

Régression

Classification

Clustering



Intelligence artificielle aujourd'hui

Introduction

Lille 3

Rapide historique

Motivation

Régression

Classification

Clustering

- Les algorithmes d'apprentissage automatiques d'aujourd'hui s'inspirent toujours des neurosciences/sciences cognitives, mais aussi de la “beauté” des mathématiques et de solutions d'ingénierie.
- Ils sont utilisés dans de multiples domaines : interface homme/machine, biologie, finance, neurosciences, etc.
- Domaines voisins :
 - fouilles de données (data mining)
 - reconnaissance de motifs (pattern recognition, computer vision)

Taxonomie

Introduction

Lille 3

Rapide
historique

Motivation

Régression
Classification
Clustering

- Apprentissage supervisée : Interaction avec environnement et l'utiliser pour expériences comme expertise
(Classification, Régression)
- Apprentissage non-supervise : résumer ou compresser l'information
(Clustering/estimation densité)

Deux histoires :

- "Pigeon Superstition" : lien
- "Poison shyness" dans les rats : (Garcia & Koelling 1996)

morale de l'histoire : Incorporation des *connaissance à priori*
bias d'apprentissage (pour des algorithmes)

Motivation

Exemple concret : Régression

Introduction

Lille 3

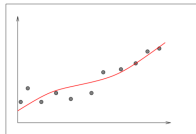
Rapide
historique

Motivation

Régression

Classification

Clustering



Détermination de l'âge d'un gastéropode :

- Par comptage au microscope du nombre d'anneaux sur le coquillage ← **tache laborieuse**.
- A travers d'autres mesures : sexe, diam'etre, longueur, poids total, poids du coquillage, etc. **facile a obtenir**.
- Exemples d'entrainement : $\{(mesures\ rapides, \text{âge})\}$.
- Tâche de **régression** : On veut prédire l'âge de l'animal uniquement à partir des mesures rapides.

Motivation

Exemple concret : Classification

Introduction

Lille 3

Rapide historique

Motivation

Régression

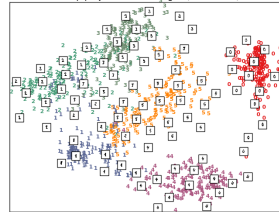
Classification

Clustering

A selection from the 64-dimensional digits dataset

0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
5	5	0	4	1	3	5	1	0	0	2	2	2	0	1	2	3	3	3	3
4	4	1	5	0	7	4	2	0	0	1	2	1	4	3	1	3	4	4	4
3	1	4	0	5	7	1	5	4	6	2	2	5	3	4	0	0	4	1	1
1	3	4	5	0	4	2	3	4	0	1	2	3	4	5	0	5	5	5	5
0	4	1	3	5	1	0	0	2	2	1	0	1	1	3	3	3	4	4	4
4	5	0	7	2	1	0	0	4	7	1	2	1	4	3	4	4	3	4	4
0	5	7	4	5	4	4	1	1	5	5	4	4	0	0	1	2	3	4	4
5	0	4	2	3	4	5	0	0	2	3	4	5	0	5	5	5	0	4	1
3	5	1	0	0	2	2	2	0	4	2	3	3	3	3	4	4	1	5	0
5	2	2	0	0	1	3	2	4	6	3	1	3	1	4	3	1	9	0	5
3	1	5	4	4	2	2	2	5	5	4	6	0	3	0	1	2	3	4	5
0	1	1	3	4	5	0	0	1	3	4	5	0	5	5	3	0	4	1	3
5	1	0	0	1	2	2	0	1	2	3	3	3	4	4	1	5	0	5	5
1	2	0	0	1	3	1	4	4	3	1	4	3	1	4	0	5	5	0	1
1	5	4	4	1	2	4	5	4	0	0	1	2	3	4	8	0	0	1	1
1	3	4	5	0	1	2	3	4	5	0	5	5	0	4	1	3	5	4	1
0	0	1	2	1	2	0	1	2	3	3	2	3	4	4	5	0	5	1	2
0	0	1	3	1	1	4	3	1	3	1	4	3	1	4	0	5	3	1	5
4	4	2	2	1	5	4	6	0	0	1	2	3	4	5	0	1	2	3	3

Isomap projection of the digits (time 1.10s)



Classification de chiffres écrits :

- Reconnaissance des codes postaux pour le tri automatique du courrier.
- La classification automatique : exemples d'entraînement = (image, chiffre).
- On veut prédire le chiffre correct la prochaine fois qu'on voit une image.

Motivation

Exemple concret : Estimation de densité/ Classement (Clustering)

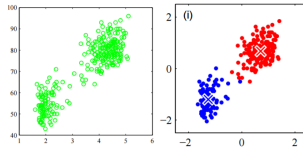
Introduction

Lille 3

Rapide
historique

Motivation

Régression
Classification
Clustering



Compression / Visualisation / Exploration de données :

- Temps entre deux éruptions par rapport à la durée de l'éruption précédente.
- Exemples d'entraînement = (mesures)
- Tache non-supervisée : On veut "organiser" l'information contenue implicitement dans ces données.