Réseau de neurone

Intelligence Artificielle et Systèmes Formels
Master 1 I2L

Sébastien Verel
verel@lisic.univ-littoral.fr
http://www-lisic.univ-littoral.fr/~verel

Université du Littoral Côte d'Opale
Laboratoire LISIC
Equipe OSMOSE

Objectifs de la séance 09

- Savoir définir les notions d'apprentissage automatique
- Savoir définir apprentissage supervisé et non-supervisé
- Connaitre la notion de sur-apprentissage
- Connaitre les méthodes d'estimation de l'erreur (validation croisée, etc.)
- Savoir concevoir un réseau de neurones de type perceptron
- Savoir concevoir un réseau de neurones de type perceptron multi-couches
- Connaitre le principe d'apprentissage par rétropropagation de l'erreur

Plan

Intelligence Artificielle

5 domaines de l'IA

- Déduction logique
- Résolution de problèmes
- **Apprentissage automatique** (artificiel)
- Représentation des connaissances
- Systèmes multiagents

Apprentissage automatique (Machine Learning)

Définition informelle

Etude et conception de systèmes (méthodes exécutées par une machine) qui sont capables d'apprendre à partir de données.

Exemple

un système qui distinguent les courriels spam et non-spam.

Apprentissage automatique (Machine Learning)

E : l'ensemble de toutes les tâches possibles.
S : un système (une machine)

Définition un peu plus formelle [T.M. Mitchell, 1997]

$T \subset E$: ensemble de taches appelé *training set*
$P : \mathcal{S} \times E \rightarrow \mathbb{R}$: mesure de performance d'un syst. sur des tâches.

Un système S **apprend** lors d'une expérience Exp si
la performance de S sur les taches T, mesurée par P, s'améliore.
$$P(S_{\text{avant Exp}}, T) \leq P(S_{\text{après Exp}}, T)$$

Exemple

Taches T : Classifier des emails reçus durant une journée
Performance P : Taux de rejet correct des spams par S
Expérience Exp : 1 semaine exposition aux courriels d'un utilisateur

Capacité de généralisation

Définition (informelle)

Capacité d'un système à fonctionner correctement sur de nouvelles taches inconnues après avoir appris sur un ensemble d'apprentissage.

T : ensemble d'apprentissage (training set)
V : ensemble de test/validation (test set) avec $V \cap T = \emptyset$

Deux systèmes S_1 et S_2.

Supposons le résultat suivant :
$P(S_2, T) \leq P(S_1, T)$
$P(S_1, V) \leq P(S_2, V)$

Interprétations :
S_1 a mieux appris que S_2 sur l'ensemble d'apprentissage T
S_1 généralise moins bien que S_2 sur un ensemble indépendant V

Machine learning *vs.* Data mining

Finalités différentes *a priori*

- Machine learning :
  but de **prédiction** à partir de propriétés connues et apprises sur un ensemble d'apprentissage
- Data mining :
  but de **découverte** de propriétés pas encore connues dans les données.

Types d'apprentissage

- Apprentissage supervisé :
  Apprentissage sur un ensemble d'exemples étiquetés :
  (entrée, sortie désirée)
- Apprentissage non supervisé :
  Apprentissage sur un ensemble d'exemples non étiquetés (cf. clustering)
- Apprentissage semi-supervisé :
  Apprentissage sur un ensemble d'exemples étiquetés / non étiquetés
- Apprentissage par renforcement :
  Apprentissage où les actions sur l'environnement se mesurent par une récompense
- ...

Liste d'algorithmes d'apprentissage automatique

Liste non exhaustive

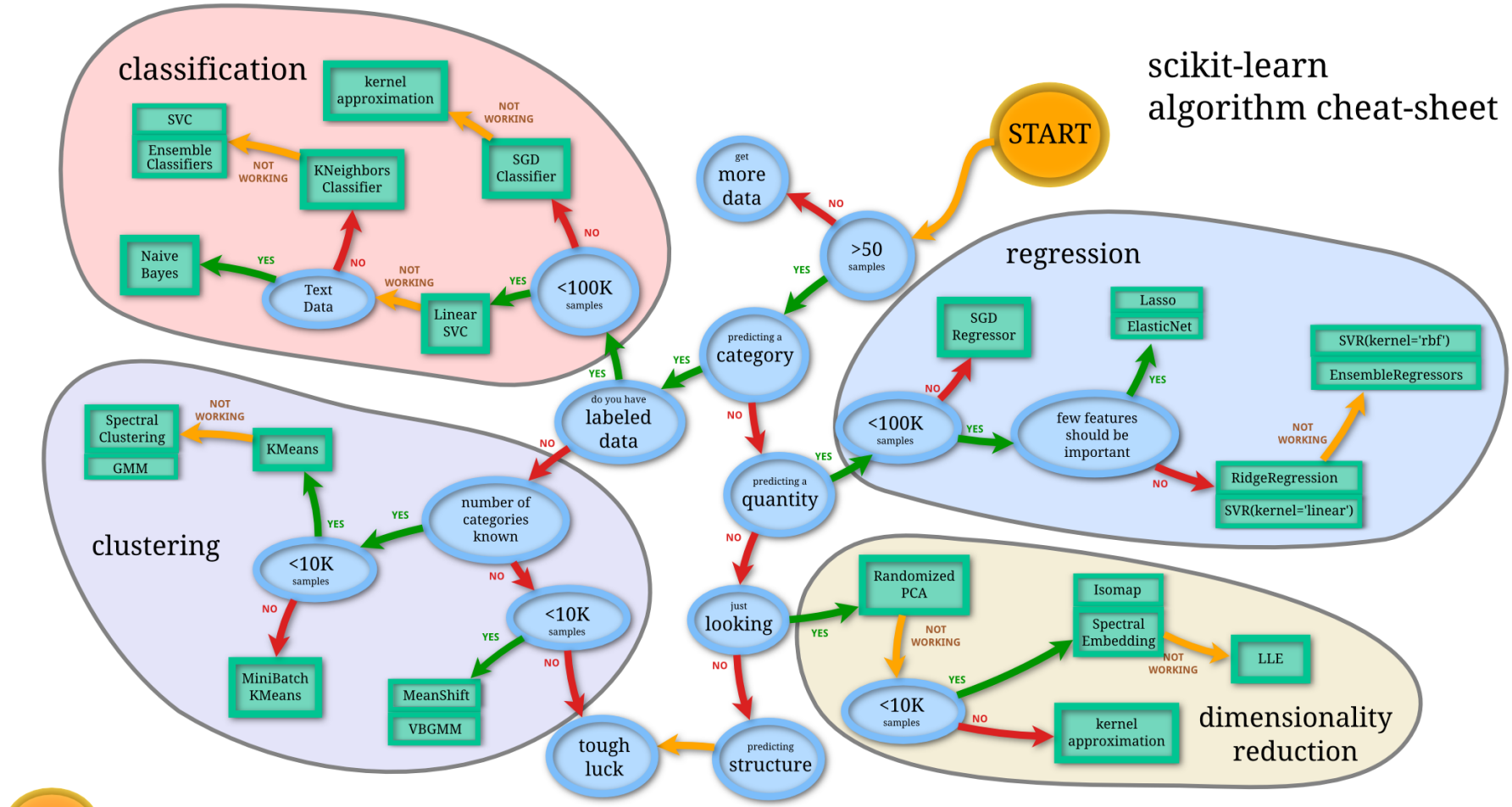
- Arbre de décision
- Régles d'association
- Réseau de neurones artificiels
- Support vector machine
- Clustering (classification)
- Inférence baysienne
- Réseaux baysiens
- Temporal difference (TD)
- etc.

Tâches d'apprentissage

- Classification :
  Classer les données
- Régression :
  Approximer les données par une fonction
- Partitionnement (Clustering) :
  Regrouper les données
- Réduction de dimension :
  Projeter les données dans un espace de dimension réduit

Choisir la bonne méthode...

source : scikit learn `http://scikit-learn.org/dev/index.html`



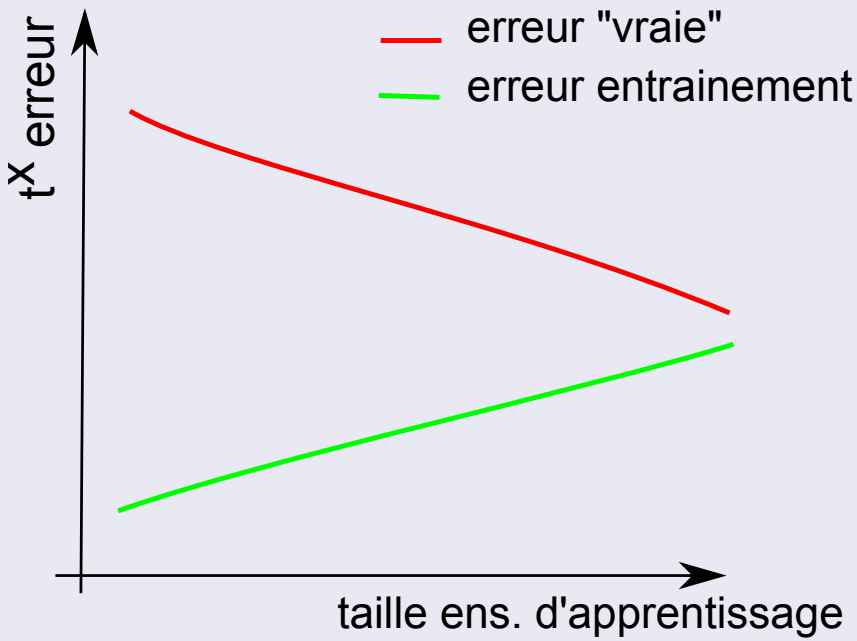
Logiciels et bibliothèques

- Weka : `http://weka.wikispaces.com/`
- R : `http://www.r-project.org`
- scikit learn (python) : `http://scikit-learn.org/dev/index.html`
- ...

Les erreurs

Relation entre erreurs

- Erreur d'apprentissage : taux d'erreur sur l'ensemble des exemples d'apprentissage
- Erreur "vraie" : erreur sur l'ensemble de tous les exemples possibles



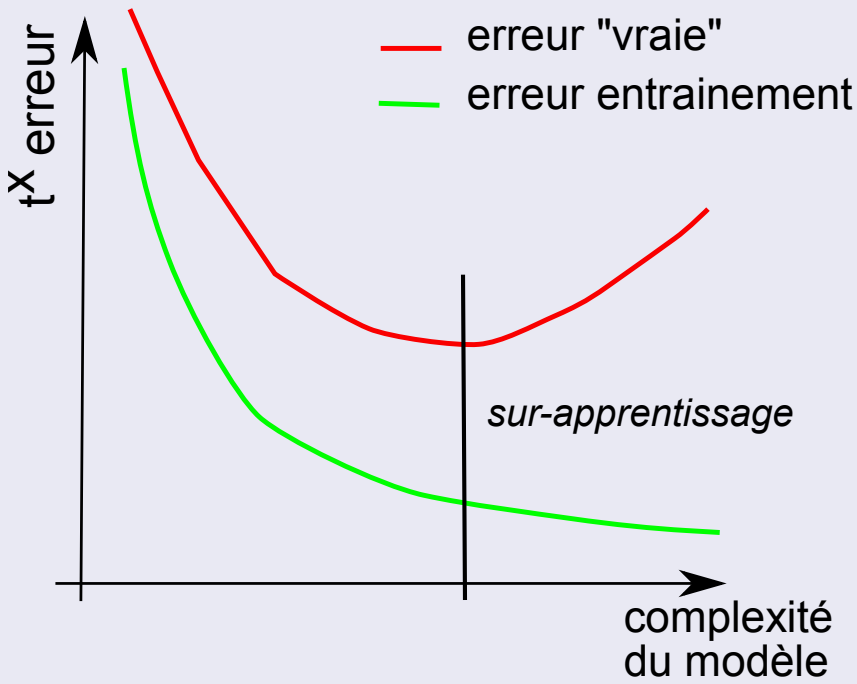
Sur-apprentissage

Exces d'apprentissage

Sur-spécialisation du modèle sur l'ensemble d'entrainement

$\Rightarrow$ Perte de capacité de généralisation

$\approx$ Apprentissage "par coeur"



Mesure de complexité d'un arbre de décision : nombre de feuilles

Evaluation d'un modèle d'apprentissage

Technique

Partitionner l'ensemble des exemples en :

- un ensemble d'apprentissage ($\approx 70\%$)
- un ensemble *indépendant* de test ($\approx 30\%$)

Le taux d'erreur est estimé (sans biais) sur l'ensemble de test.

Inconvénient

- Requiert un nombre important d'exemples
- Dilemme :
  - Plus on met d'exemples dans le test, plus l'estimation est précise
  - Plus on met d'exemples dans l'apprentissage, meilleur est le modèle (a priori)

Méthode de ré-échantillonnage

Permet d'estimer l'erreur de généralisation.

K-folds cross-validation

Partitionner aléatoirement l'échantillon en K blocs
Pour chaque bloc k,
    Construire le modéle sur les $k-1$ autres blocs
    Calculer l'erreur en test e_k sur le block k
Calculer l'erreur moyenne des erreurs e_k

Autres techniques :

- Leave-one-out ($K = n$)
- Bootstrap, bagging, etc.

Bibliographie

- Fabien Teytaud, Université du Littoral Côte d'Opale, http://www-lisic.univ-littoral.fr/~teytaud/enseignements.html
- Denis Robilliard, Université du Littoral Côte d'Opale, http://www-lisic.univ-littoral.fr/~robillia/index.html
- Marie Cottrell, Université Paris I - Sorbonne, http://samm.univ-paris1.fr/-Marie-Cottrell-
- Manuel Clergue, Université des Antilles et de la Guyanne http://lamia.univ-ag.fr/index.php?option=com_content&view=article&id=1296

Représentation

Les techniques d'apprentissage se distinguent par les représentations :

- Arbres de décision
- Règles d'association
- **Réseaux de neurones**
- ...

Une représentation est une structure de donnée (lecture/écriture). L'état de la structure permet la mémorisation.

Lors de la phase d'apprentissage,
l'état propre à la structure est modifiée : "le modèle apprend" pour augmenter la performance sur l'ensemble d'apprentissage et tout en gardant des capacités de généralisation

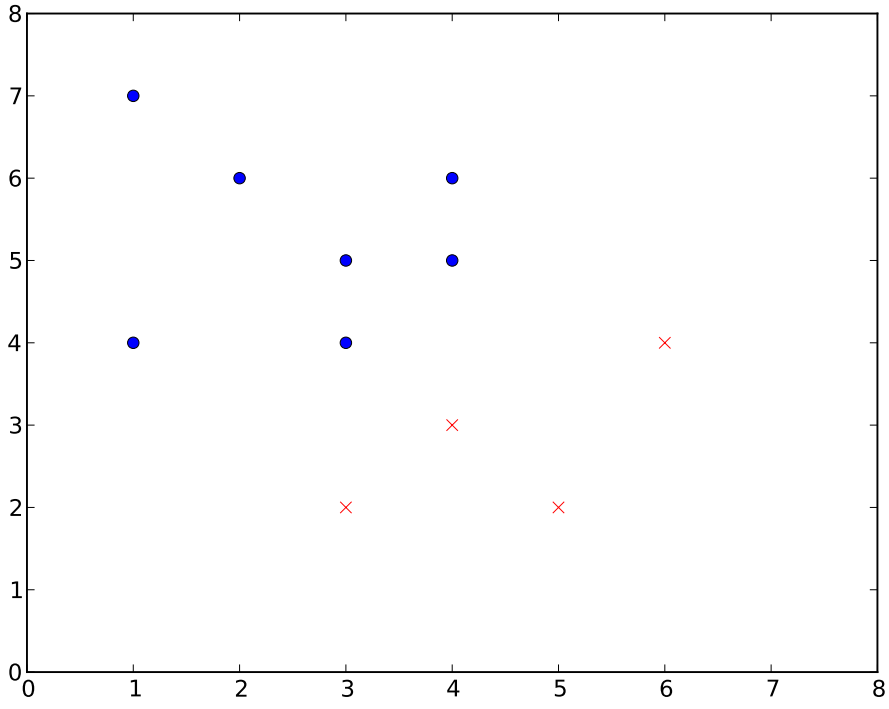
Réseau de neurones artificiels

Classe de fonctions non-linéaires inspirée du fonctionnement des neurones biologiques

Taches d'apprentissage :

- Classification
- Régression

Un exemple de classification



Question

Définir une fonction qui permet de classer les exemples bleus et rouges.

Suite du cours

cf. Marie Cottrell, Université Paris I - Sorbonne,
`http://samm.univ-paris1.fr/-Marie-Cottrell-`

Fabien Teytaud, Université du Littoral Côte d'Opale,
`http://www-lisic.univ-littoral.fr/~teytaud/enseignements.html`

Et tableau...