



Classification Réseaux de neurones

Dr A. DJEFFAL

2^{ème} année Master Systèmes d'Information, Optimisation et Décision

2018-2019

www.abdelhamid-djeffal.net



Principe

RNA : Introduction

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Réseaux de neurones artificiels

- Inspirés de la méthode de travail du cerveau humain



Principe

RNA : Introduction

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Réseaux de neurones artificiels

- Inspirés de la méthode de travail du cerveau humain
- Cerveau humain



Principe

RNA : Introduction

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Réseaux de neurones artificiels

- Inspirés de la méthode de travail du cerveau humain
- Cerveau humain
 - Méthode totalement différente de celle d'un ordinateur.



Principe

RNA : Introduction

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Réseaux de neurones artificiels

- Inspirés de la méthode de travail du cerveau humain
- Cerveau humain
 - Méthode totalement différente de celle d'un ordinateur.
 - système de traitement très compliqué



Principe

RNA : Introduction

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Réseaux de neurones artificiels

- Inspirés de la méthode de travail du cerveau humain
- Cerveau humain
 - Méthode totalement différente de celle d'un ordinateur.
 - système de traitement très compliqué
 - parallèle et non linéaire



Principe

RNA : Introduction

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Réseaux de neurones artificiels

- Inspirés de la méthode de travail du cerveau humain
- Cerveau humain
 - Méthode totalement différente de celle d'un ordinateur.
 - système de traitement très compliqué
 - parallèle et non linéaire
 - organise ses composants pour traiter, d'une façon très performante et très rapide, des problèmes très compliqués tel que la reconnaissance des formes.

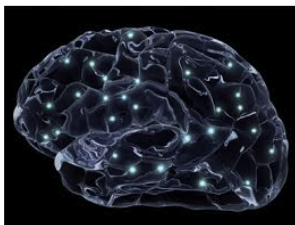
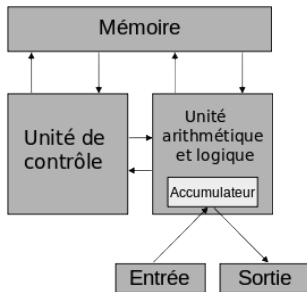


Principe

RNA : Introduction

Classification
Réseaux de
neurones

Dr A.
DJEFFAL



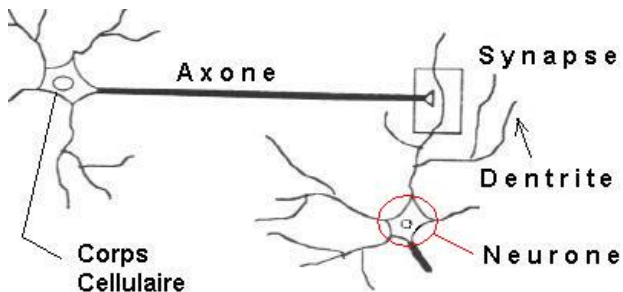


Principe

RNA : Introduction

Classification
Réseaux de
neurones

Dr A.
DJEFFAL





Principe

Physiologie du cerveau humain

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

- Les neurones reçoivent les signaux (impulsions électriques) par des extensions très ramifiées de leur corps cellulaire (les dendrites)
- et envoient l'information par de longs prolongements (les axones).
- Les impulsions électriques sont régénérées pendant le parcours le long de l'axone.
- La durée de chaque impulsion est de l'ordre d'1 ms et son amplitude d'environ 100 mV.
- Les contacts entre deux neurones, de l'axone à une dendrite, se font par l'intermédiaire des synapses



Principe

Physiologie du cerveau humain

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

- le cerveau contient environ 100 milliards de neurones.
- on ne dénombre que quelques dizaines de catégories distinctes de neurones.
- aucune catégorie de neurones n'est propre à l'homme.
- la vitesse de propagation des influx nerveux est de l'ordre de 100m/s, c'est à dire bien inférieure à la vitesse de transmission de l'information dans un circuit électronique.
- on compte de quelques centaines à plusieurs dizaines de milliers de contacts synaptiques par neurone. Le nombre total de connexions est estimé à environ 10^{15} .



Principe

Physiologie du cerveau humain

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

- la connectique du cerveau ne peut pas être codée dans un "document biologique" tel l'ADN pour de simples raisons combinatoires. La structure du cerveau provient donc en partie des contacts avec l'environnement. L'apprentissage est donc indispensable à son développement.
- le nombre de neurones décroît après la naissance. Cependant, cette affirmation semble remise en question.
- les synapses entre des neurones qui ne sont pas simultanément actifs sont affaiblis puis éliminés.
- il semble que l'apprentissage se fasse par un double mécanisme : des connections sont établies de manière redondantes et aléatoires puis seules les connexions entre des neurones simultanément actifs sont conservés (phase de sélection) tandis que les autres sont éliminés. On parle de stabilisation sélective.



Composition

Un neurone seul

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

- Un neurone est l'unité élémentaire de traitement d'un réseau de neurones.
- Il est connecté à des sources d'information en entrée (d'autres neurones par exemple) et renvoie une information en sortie.
- Un neurone artificiel reçoit en entrée des entrées numériques x_i ($1 \leq i \leq N$ où N est le nombre de données qu'il va recevoir) valorisée chacune par un coefficient w_i .



Composition

Un neurone seul

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

- Le neurone artificiel (qui est une modélisation des neurones du cerveau) effectue alors une somme pondérée de ses entrées et lui ajoute un coefficient w_0 dit de biais supposé lié à une donnée $x_0 = -1$.

$$S = w_1 \times x_1 + \dots + w_N \times x_N - w_0 = \sum_{i=1}^N w_i \times x_i - w_0$$

- Cette donnée est passée à une fonction f dite d'activation qui représente un filtre permettant d'adapter la valeur de la somme précédente aux caractéristiques de la sortie désirée.
- C'est une fonction qui, généralement, doit renvoyer un réel proche de 1 quand les "bonnes" informations d'entrée sont données et un réel proche de 0 quand elles sont "mauvaises".



Composition

Un neurone seul

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

- La valeur de la fonction d'activation est la sortie y du neurone.

$$y = f\left(\sum_{i=1}^N w_i \times x_i - w_0\right)$$



Composition

Un neurone seul

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

- La valeur de la fonction d'activation est la sortie y du neurone.

$$y = f\left(\sum_{i=1}^N w_i \times x_i - w_0\right)$$

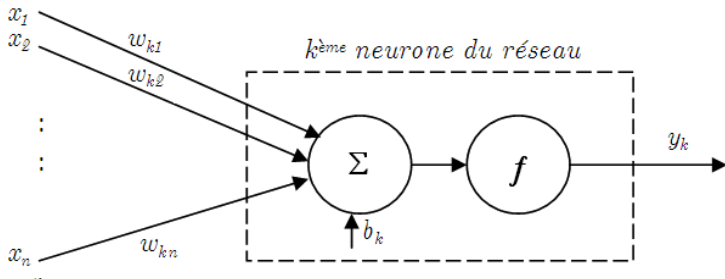


Composition

Un neurone seul

Classification
Réseaux de
neurones

Dr A.
DJEFFAL





Composition

Un neurone seul

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

- Un ensemble de connexions avec les différentes entrées x_i , pondérée chacune par un poids w_i ,
- Un additionneur permettant de calculer une combinaison linéaire des entrées x_i pondérées par les coefficients w_i ,
- Un biais w_0 qui permet de contrôler l'entrée de la fonction d'activation,
- Une fonction d'activation f permettant de délimiter la sortie y du neurone.



Composition

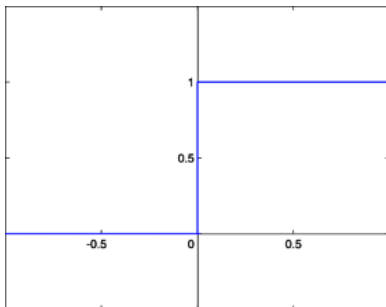
Un neurone seul

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

- **La fonction de Heaviside**

$$\forall x \in \mathbb{R} f(x) = 1 \text{ si } x \geq 0, 0 \text{ sinon}$$





Composition

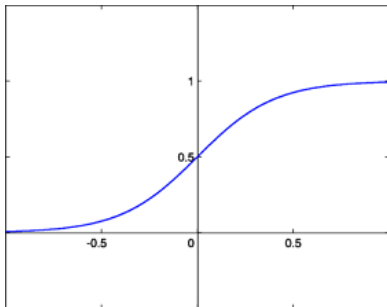
Un neurone seul

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

- **La fonction sigmoïde**

$$\forall x \in \mathbb{R} f(x) = \frac{1}{1+e^{-x}}$$





Architecture

Un réseau de neurones artificiels

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Architecture définie par :

- Nombre d'entrées,



Architecture

Un réseau de neurones artificiels

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Architecture définie par :

- Nombre d'entrées,
- Nombre de sorties,



Architecture

Un réseau de neurones artificiels

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Architecture définie par :

- Nombre d'entrées,
- Nombre de sorties,
- Nombre de nœuds



Architecture

Un réseau de neurones artificiels

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Architecture définie par :

- Nombre d'entrées,
- Nombre de sorties,
- Nombre de nœuds
- Façon d'interconnexion et d'organisation des nœuds.



Architecture

Un réseau de neurones artificiels

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Architecture définie par :

- Nombre d'entrées,
- Nombre de sorties,
- Nombre de nœuds
- Façon d'interconnexion et d'organisation des nœuds.
- Cette architecture est utilisée dans presque 90 % des applications commerciales et industrielles.



Architecture

Réseau de neurone monocouche (Perceptron)

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

- Un réseau de neurones est une structure de réseau constituée d'un nombre de nœuds interconnectés par des liaisons directionnelles.
- Chaque nœud représente une unité de traitement et les liaisons représentent les relations causales entre les nœuds. La figure suivante représente une schématisation d'un neurone.
- Le réseau le plus simple est celui monocouche appelé *le perceptron*



Architecture

Réseau de neurone monocouche (Perceptron)

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

- Il possède N informations en entrée ;
- Il est composé de p neurones, que l'on représente généralement alignés verticalement. Chacun peut en théorie avoir une fonction d'activation différente. En pratique, ce n'est généralement pas le cas ;
- Chacun des p neurones est connecté aux N informations d'entrée.

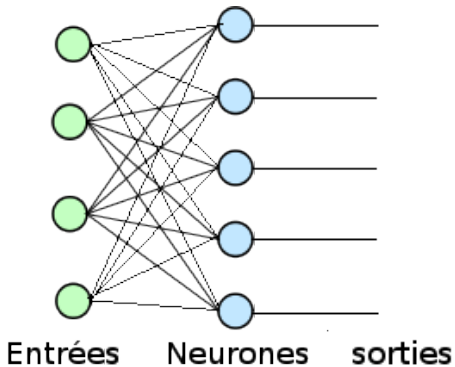


Architecture

Réseau de neurone monocouche (Perceptron)

Classification
Réseaux de
neurones

Dr A.
DJEFFAL





Architecture

Réseau de neurone monocouche (Perceptron)

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

- Le réseau de neurones possède ainsi N informations en entrée et P sorties,
- chaque neurone renvoyant sa sortie.
- Une utilisation courante est que chaque neurone de la couche représente une classe.
- Pour un exemple X donné, on obtient la classe de cet exemple en prenant la plus grande des p sorties.



Architecture

Réseau de neurone multicouche

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

- Dans ce réseau, les neurones de la première couche reçoivent toutes les informations entrées,
- ceux de la deuxième reçoivent toutes les sorties des neurones de la première couche,
- et ainsi de suite jusqu'au neurone de sortie qui reçoit celles de la dernière couche.

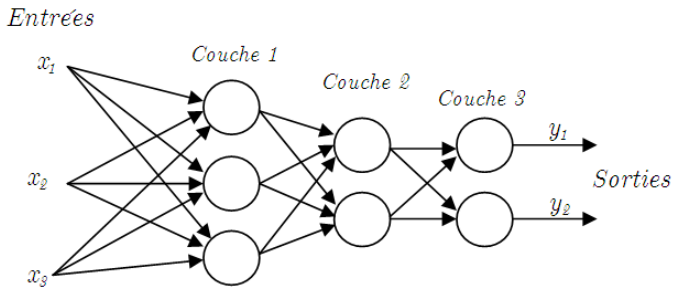


Architecture

Réseau de neurone multicouche

Classification
Réseaux de
neurones

Dr A.
DJEFFAL





Architecture

Apprentissage des RNA

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

- Dans ce réseau, les neurones de la première couche reçoivent toutes les informations entrées,
- ceux de la deuxième reçoivent toutes les sorties des neurones de la première couche,
- et ainsi de suite jusqu'au neurone de sortie qui reçoit celles de la dernière couche.



Principe

- Tâche principale : classification



Principe

- Tâche principale : classification
- Entraînement : Processus itératif



Principe

- Tâche principale : classification
- Entraînement : Processus itératif
- Obtenir les w_i permettant d'avoir $f(x_i) = y_i, \forall i = 1..N$



Principe

- Tâche principale : classification
- Entraînement : Processus itératif
- Obtenir les w_i permettant d'avoir $f(x_i) = y_i, \forall i = 1..N$
 - w_i initialisées aléatoirement



Principe

- Tâche principale : classification
- Entraînement : Processus itératif
- Obtenir les w_i permettant d'avoir $f(x_i) = y_i, \forall i = 1..N$
 - w_i initialisées aléatoirement
 - Corrigés selon les erreurs entre les y_i obtenus et attendus



Réseau de neurones multicouches

- Correction dans le sens inverse du sens de propagation des données



Réseau de neurones multicouches

- Correction dans le sens inverse du sens de propagation des données
- "*backpropagation*"



Réseau de neurones multicouches

- Correction dans le sens inverse du sens de propagation des données
- "*backpropagation*"
- À chaque présentation d'un exemple :



Réseau de neurones multicouches

- Correction dans le sens inverse du sens de propagation des données
- "*backpropagation*"
- À chaque présentation d'un exemple :
 - 1 Propagation : valeurs du vecteur exemple reçues dans la couche d'entrée puis propagées d'une couche à l'autre jusqu'à la sortie où un vecteur de sortie (les y_i) est obtenu.



Réseau de neurones multicouches

- Correction dans le sens inverse du sens de propagation des données
- "*backpropagation*"
- À chaque présentation d'un exemple :
 - 1 Propagation : valeurs du vecteur exemple reçues dans la couche d'entrée puis propagées d'une couche à l'autre jusqu'à la sortie où un vecteur de sortie (les y_i) est obtenu.
 - 2 Backpropagation, les w_i sont ajustés de la dernière couche jusqu'à la première de manière à rapprocher les y_i obtenus de ceux attendus.



Réseau de neurones multicouches

- Correction dans le sens inverse du sens de propagation des données
- "*backpropagation*"
- À chaque présentation d'un exemple :
 - 1 Propagation : valeurs du vecteur exemple reçues dans la couche d'entrée puis propagées d'une couche à l'autre jusqu'à la sortie où un vecteur de sortie (les y_i) est obtenu.
 - 2 Backpropagation, les w_i sont ajustés de la dernière couche jusqu'à la première de manière à rapprocher les y_i obtenus de ceux attendus.
- Répéter avec chaque exemple d'entraînement



Réseau de neurones multicouches

- Correction dans le sens inverse du sens de propagation des données
- "*backpropagation*"
- À chaque présentation d'un exemple :
 - 1 Propagation : valeurs du vecteur exemple reçues dans la couche d'entrée puis propagées d'une couche à l'autre jusqu'à la sortie où un vecteur de sortie (les y_i) est obtenu.
 - 2 Backpropagation, les w_i sont ajustés de la dernière couche jusqu'à la première de manière à rapprocher les y_i obtenus de ceux attendus.
- Répéter avec chaque exemple d'entraînement
- Obtenir un RNA entraîné



Réseau de neurones multicouches

- RNA entraîné (Réseaux + w_i + b_i)



Réseau de neurones multicouches

- RNA entraîné (Réseaux + $w_i + b_i$)
- Injection des valeurs du vecteur de l'exemple à classifier dans l'entrée



Réseau de neurones multicouches

- RNA entraîné (Réseaux + $w_i + b_i$)
- Injection des valeurs du vecteur de l'exemple à classifier dans l'entrée
- Propagation



Réseau de neurones multicouches

- RNA entraîné (Réseaux + $w_i + b_i$)
- Injection des valeurs du vecteur de l'exemple à classifier dans l'entrée
- Propagation
- Recevoir la classe à la sortie



Apprentissage

Méthode de descente de gradient

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Algorithme 4 ApprentissageRNA par descente de gradient

- 1: Créer n variables dw_i , pour $1 \leq i \leq n$, égales à 0
 - 2: Prendre un exemple e_k , pour $1 \leq k \leq N$
 - 3: Calculer la sortie obtenue avec les poids actuels, notée s_k
 - 4: Rajouter à dw_i , pour tout $1 \leq i \leq n$, le nombre $\alpha(y_k - s_k)x_i$
 - 5: Répéter (3) et (4) sur chacun des exemples
 - 6: pour $1 \leq i \leq n$, remplacer w_i par $w_i + dw_i$
-



Apprentissage

Méthode de descente de gradient

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Algorithme 4 ApprentissageRNA par descente de gradient

- 1: Créer n variables dw_i , pour $1 \leq i \leq n$, égales à 0
 - 2: Prendre un exemple e_k , pour $1 \leq k \leq N$
 - 3: Calculer la sortie obtenue avec les poids actuels, notée s_k
 - 4: Rajouter à dw_i , pour tout $1 \leq i \leq n$, le nombre $\alpha(y_k - s_k)x_i$
 - 5: Répéter (3) et (4) sur chacun des exemples
 - 6: pour $1 \leq i \leq n$, remplacer w_i par $w_i + dw_i$
-

- α est un nombre réel appelé taux (vitesse) d'apprentissage
- bien choisi, suffisamment petit
- Si trop grand : risque d'oscillation autour du minimum
- Si trop petit : nombre élevé d'itérations
- En pratique : on diminue graduellement α au fur et à mesure des itérations



Apprentissage

Méthode de descente de gradient

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

- Afin d'obtenir de bons résultats, il faudra passer plusieurs fois les exemples à chaque neurone, de sorte que les poids convergent vers des poids "idéaux".
- Le problème avec cette méthode est que l'on corrige sur la globalité des exemples,
- ce qui fait que le réseau ne s'adaptera aux exemples qu'après un certain moment.
- Il y a une autre méthode qui permet de corriger sur chacun des exemples, et qui se nomme méthode d'apprentissage de Widrow-Hoff.



Apprentissage Widrow-Hoff

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Algorithme 5 Apprentissage RNA par Widrow-Hoff

Entrée :

- n poids reliant les n informations à notre neurone ayant des valeurs quelconques
- N exemples (X_k, y_k) où X_k est un vecteur à n composantes x_i , chacune représentant une information de cet exemple
- Le taux d'apprentissage α

Sortie : les n poids modifiés

pour Tout exemple = (X_k, y_k) **faire**

 Calculer la sortie s_k du neurone

pour $1 \leq i \leq n$ **faire**

$$w_i = w_i + \alpha(y_k - s_k)x_i$$

fin pour

fin pour



Apprentissage

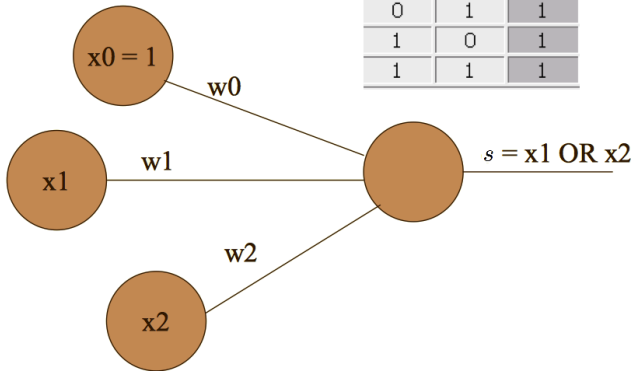
Widrow-Hoff (Exemple)

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Apprentissage du OU logique

a	b	OR
0	0	0
0	1	1
1	0	1
1	1	1





Apprentissage

Widrow-Hoff (Exemple)

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

- $\alpha = 1$
- x_0 vaut toujours 1
- Initialisation : $w_0 = 0$; $w_1 = 1$; $w_2 = -1$



Apprentissage Widrow-Hoff (Exemple)

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Étape	w_0	w_1	w_2	Entrée	$\sum_0^2 w_i x_i$	y	s	w_0	w_1	w_2
init								0	1	-1
1	0	1	-1	100	0	0	0	$0+0x1$	$1+0x0$	$-1+0x0$
2	0	1	-1	101	-1	0	1	$0+1x1$	$1+1x0$	$-1+1x1$
3	1	1	0	110	2	1	1	1	1	0
4	1	1	0	111	2	1	1	1	1	0
5	1	1	0	100	1	1	0	$1+(-1)x1$	$1+(-1)x0$	$0+(-1)x0$
6	0	1	0	101	0	0	1	$0+1x1$	$1+1x0$	$0+1x1$
7	1	1	1	110	2	1	1	1	1	1
8	1	1	1	111	3	1	1	1	1	1
9	1	1	1	100	1	1	0	$1+(-1)x1$	$1+(-1)x0$	$1+(-1)x0$
10	0	1	1	101	1	1	1	0	1	1



Apprentissage

Widrow-Hoff (Exemple)

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

- Donc : $w_0 = 0$; $w_1 = 1$; $w_2 = 1$
- Ce perceptron calcule le OU logique pour tout couple $(x_1 ; x_2)$



Apprentissage

Widrow-Hoff (Exemple)

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

- Si l'échantillon n'est pas linéairement séparable, l'algorithme ne converge pas.
- L'algorithme peut converger vers plusieurs solutions, selon :
 - les valeurs initiales des coefficients,
 - la valeur de α ,
 - l'ordre de présentation des exemples
- La solution n'est pas robuste : un nouvel exemple peut remettre en cause le perceptron appris.

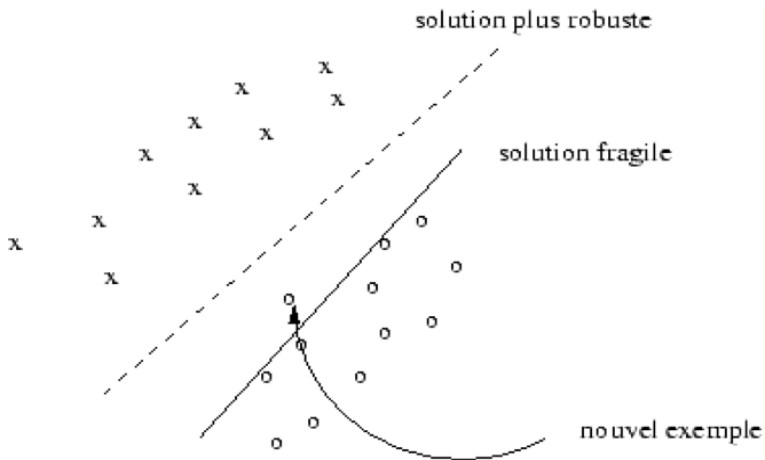


Apprentissage

Widrow-Hoff (Exemple)

Classification
Réseaux de
neurones

Dr A.
DJEFFAL



- Solution : RNA multicouches

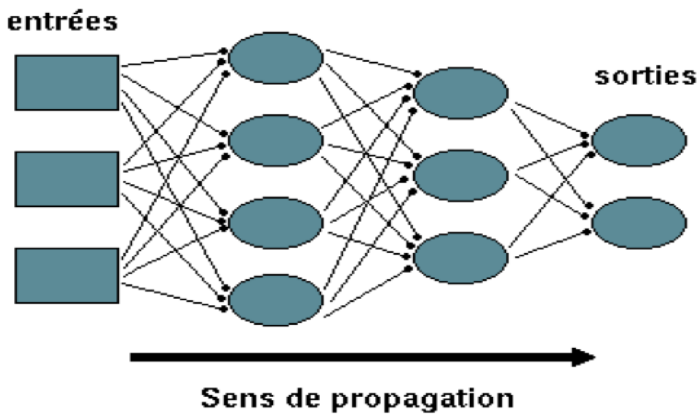


Apprentissage

RNA multicouches

Classification
Réseaux de
neurones

Dr A.
DJEFFAL





Apprentissage

RNA multicouches (entraînement)

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Algorithme 6 Apprentissage RNA multicouche

Initialiser aléatoirement les coefficients w_{ij} dans $[-0.5; 0.5]$

Répéter

Prendre un exemple (x, y)

Calculer la sortie s

pour toute cellule de sortie i **faire**

$$\delta_i = s_i \times (1 - s_i) \times (y_i - s_i)$$

fin pour

pour chaque couche de $q - 1$ à 1 **faire**

pour chaque cellule i de la couche courante **faire**

$$\delta_i = s_i \times (1 - s_i) \times \left[\sum_{(k \in \text{Succ}(i))} (\delta_k \times w_{ki}) \right]$$

fin pour

fin pour

pour Pour tout poids w_{ij} **faire**

$$w_{ij} = w_{ij} + \alpha \times \delta_i \times x_{ij}$$

fin pour

Fin Répéter.



Réseau de neurones multicouches

- Régression : Naturelle (y_i représentent la réponse de la fonction à estimer)



Réseau de neurones multicouches

- Régression : Naturelle (y_i représentent la réponse de la fonction à estimer)
- Classification : Discrétiser les y_i (fonction d'activation ou après)
 - Binaire : une seule sortie (0 ou 1)
 - Multiclasse : une sortie pour chaque classe



Types

Types RNAs

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Types

- Selon la manière d'interconnexion



Types

Types RNAs

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Types

- Selon la manière d'interconnexion
- RNA récurrents : propager les résultats au sens inverse de la propagation dans le calcul des w_i



Types

Types RNAs

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Types

- Selon la manière d'interconnexion
- RNA récurrents : propager les résultats au sens inverse de la propagation dans le calcul des w_i
- cartes auto-organisatrices de Kohonen : utilise la compétition pour prendre les meilleurs nœuds dans les calculs (généralement pour le clustering)



Avantages

Avantages des RNAs

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Avantages

- Surmonter le problème d'analyse d'un système donné pour le modéliser



Avantages

Avantages des RNAs

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Avantages

- Surmonter le problème d'analyse d'un système donné pour le modéliser
- On peut simuler son comportement uniquement à partir d'un certain nombre d'exemples observés



Inconvénients

Inconvénients des RNAs

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Inconvénients

- Boîte noire
 - très difficile (impossible) d'analyser et comprendre le fonctionnement en face d'un problème donné,
 - Difficulté de choisir la structure (type, nombre de nœuds, organisation, connexions,...etc) la mieux adaptée au problème.



Inconvénients

Inconvénients des RNAs

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Inconvénients

- Boîte noire
 - très difficile (impossible) d'analyser et comprendre le fonctionnement en face d'un problème donné,
 - Difficulté de choisir la structure (type, nombre de nœuds, organisation, connexions,...etc) la mieux adaptée au problème.
- Influence de l'ordre de présentation des exemples d'entraînement :
 - Solution : répéter l'entraînement avec des ordres différents mais temps d'apprentissage



Inconvénients

Inconvénients des RNAs

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Inconvénients

- Problème des attributs symboliques
 - Solution : encodage adapté



Inconvénients

Inconvénients des RNAs

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Inconvénients

- Problème des attributs symboliques
 - Solution : encodage adapté
- Sensibilité aux minimas locaux et divergence : La boite noire empêche la correction



Inconvénients

Inconvénients des RNAs

Classification
Réseaux de
neurones

Dr A.
DJEFFAL

Inconvénients

- Problème des attributs symboliques
 - Solution : encodage adapté
- Sensibilité aux minimas locaux et divergence : La boite noire empêche la correction
- Plusieurs applications favorisent les SVMs