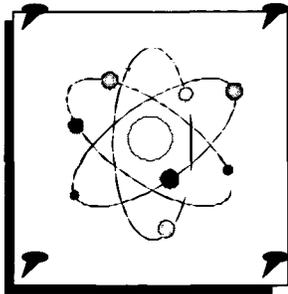


DES NEURONES DANS LES RÉSEAUX

Christian CALZADA,
Sophie MIDENET (INRETS-MAIA)*



Les domaines d'application potentiels ou déjà effectifs des techniques connexionnistes sont très vastes, tout comme l'ensemble des disciplines scientifiques concernées par leur étude à un titre ou à un autre. Nous nous proposons ici d'attirer l'attention sur les potentialités que ces méthodes laissent entrevoir dans le domaine des transports.

**Simuler
le fonctionnement
du cerveau humain**

Le terme **Réseau de Neurones Formels (RNF)** désigne un ensemble de modèles dont l'utilisation s'est largement répandue dans les quinze dernières années en touchant tous les domaines concernés par la modélisation et le traitement des données. Les premiers développements datent des débuts de l'ère informatique (années 1940) et sont menés principalement par des chercheurs issus de disciplines comme la neurobiologie ou la psychologie. Après une période de relative marginalité dans les années 1960, la discipline connexionniste a connu une véritable renaissance dans les années 1980 sous l'impulsion cette fois de physiciens, de mathématiciens et d'informaticiens.

**Applications
transports :
un domaine
en friche**

Parmi les **utilisations types des RNF**, citons les problèmes de reconnaissance de forme (classification, segmentation prédictive, détection), de prévision de trajectoire de systèmes dynamiques, de diagnostic, d'identification et de commandes non linéaires.

Les applications développées à ce jour dans le domaine des transports ont concerné surtout des **problèmes d'ingénierie transports**. Un exemple parmi d'autres : **la réflexion sur un système embarqué de résolution de conflits aériens entre deux aéronefs**¹.

Le rôle des RNF appris ici par **Algorithmes Génétiques** est de diriger au mieux un avion dans son espace aérien afin de résoudre les problèmes dits de conflits aériens (respect des normes de séparation entre avions) avec un maximum de sécurité et un minimum de retard à l'arrivée. Les RNF doivent être réutilisables dans des situations inconnues de configurations de vol. Les informations ne seraient plus uniquement à la charge des contrôleurs aériens au sol mais susceptibles d'être transmises directement par une liaison automatique de données (gestion automatique de contrôle aérien).

D'un **point de vue prédictif**, peu d'applications transports sont disponibles à ce jour.

Signalons quand même :

- * la prévision du trafic voyageur mensuel sur une ligne SNCF un mois à l'avance (la modélisation de séries temporelles peut dans certains cas se ramener à la modélisation d'un système dynamique),
- * la modélisation journalière du trafic sur un tronçon autoroutier en utilisant des informations sur le calendrier et les vacances scolaires,
- * la modélisation des relations entre flux aériens entrant et sortant d'un secteur,
- * la prévision de trafic autoroutier de très court terme (à la demi heure) basé sur le calibrage horaire d'un RNF.

**Connexionnisme
et Economie
des Transports**

A la question : comment appréhender ce que la théorie économique propose comme des objets non représentables sous forme de fonctions ? la logique connexionniste fournit un début de réponse.

* Cette réflexion fait suite à la "Journée de sensibilisation aux Méthodes Neuronales appliquées au domaine des transports", journée organisée par l'OEST le 29 septembre 1995. Les actes de cette journée sont disponibles sur simple demande écrite auprès de l'OEST.

¹ "Résolution de conflits aériens par réseaux de neurones appris par algorithmes génétiques" G. GAUDIERES, Juin 1995, Laboratoire d'optimisation stochastique de l'ENAC.

METHODES

Ces méthodes dont le domaine par excellence est celui de l'estimation de variétés non linéaires, possèdent des propriétés d'apprentissage, d'adaptation et de robustesse (résistance au bruit) qui sont chères aux macroéconomistes. Les problèmes rémanents en analyse des données comme le traitement des valeurs manquantes, la prise en compte de variables qui interagissent fortement entre elles, la prédiction simultanée de plusieurs choix individuels de comportements ont trouvé des solutions intéressantes grâce aux RNF.

L'aspect neuromimétique de ces modèles

Les modèles de RNF reposent sur un principe hérité directement des réseaux de neurones biologiques, puisqu'ils sont constitués d'un réseau de processeurs élémentaires à fonction de transfert non linéaire, appelés neurones formels, fortement connectés les uns aux autres et qui communiquent en transmettant des niveaux d'activation via des connexions dans le réseau. L'aspect neuromimétique de ces modèles ne se limite pas à l'architecture et au mode de propagation de l'information, car ces modèles disposent aussi de propriétés d'adaptation grâce auxquelles certains paramètres de l'architecture du réseau ne sont pas fixés a priori par le concepteur mais résultent d'un processus d'apprentissage à partir des données du problème particulier à traiter.

Différentes classes de RNF

De nombreux modèles de RNF sont étudiés et utilisés pour des applications, parmi lesquels on distingue deux catégories principales: les **réseaux à couches** et les **réseaux classifieurs**. Ces deux grandes familles de réseaux se différencient à la fois par leur fonction (association de formes d'entrée - sortie pour les réseaux à couches, classification de formes d'entrée pour le second) et par leur type d'apprentissage: **apprentissage supervisé pour les réseaux à couches**, c'est à dire utilisant l'information de sortie désirée au niveau de chaque forme d'entrée à traiter; **apprentissage non supervisé pour les réseaux classifieurs**.

Citons encore deux autres catégories de modèles classiques mais moins utilisées en pratique: les **réseaux à états d'équilibres** ou les **réseaux probabilistes**.

Principes fondamentaux

Les caractéristiques communes aux modèles connexionnistes peuvent se résumer de la façon suivante:

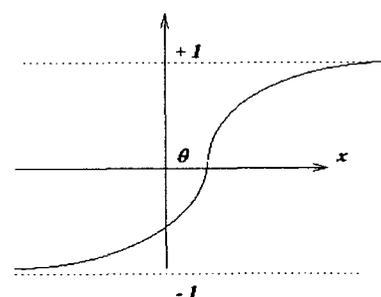
- non linéarité,
- **parallélisme** (répartition de la représentation de l'information et du traitement de l'information),
- **adaptation**.

L'exemple du Perceptron Multicouches

Le *Perceptron Multi-Couches* est le modèle le plus utilisé dans les applications pratiques. Il appartient à la catégorie des réseaux à couches. Son architecture comprend plusieurs couches de neurones formels successivement disposées entre couche d'entrée et couche de sortie.

Les niveaux d'activation des neurones formels de la couche d'entrée sont forcés aux valeurs des formes d'entrée à traiter, l'information se propage alors dans le réseau jusqu'à atteindre les neurones de la couche de sortie dont les niveaux d'activation constituent la forme de sortie du réseau. Entre les couches d'entrée et de sortie s'intercale au moins une couche de neurones dite "cachée". Les connexions n'existent qu'entre neurones appartenant à deux couches successives. La fonction de transfert des neurones cachés est de forme sigmoïdale. L'architecture multicouche de ces réseaux associée à la non linéarité de la fonction de transfert des neurones cachés donne au perceptron multicouche la propriété

f fonction sigmoïde

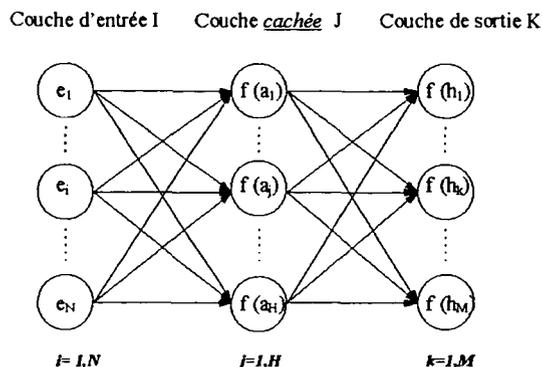


$$f(x) = \frac{\exp[k(x - \theta)] - 1}{\exp[k(x - \theta)] + 1}$$

avec k : raideur ou importance de la sigmoïde
 θ : seuil

METHODES

d'*approximateur universel* : un tel modèle permet de réaliser n'importe quelle fonction mesurable avec une précision fixée à l'avance. C'est une propriété théorique puissante, même s'il ne fournit aucune indication pratique sur l'architecture et la configuration de poids adéquate à une fonction donnée. L'intérêt pratique de ce modèle provient de l'existence d'un algorithme d'apprentissage appelé *rétro-propagation de l'erreur*, qui porte sur les poids des connexions et qui consiste à appliquer itérativement une *méthode d'optimisation par descente de gradient de l'erreur quadratique* observée en fonction des poids des connexions.



Formulation

w_{ij}^1 : poids d'entrée de la couche I à la couche cachée J
 w_{jk}^0 : poids de sortie de la couche cachée J à la couche K

(e_i) : entrées du neurone formel

Sortie h_j de la couche cachée J :

$$h_j = f \left(\sum_{i=1}^N w_{ij}^1 \cdot e_i \right) = f(a_j)$$

la fonction d'activation f est la fonction sigmoïde

Critère de qualité : Minimiser l'Erreur Quadratique Moyenne $e(W)$:

$$e(W) = (0,5) \parallel h^* - h(W) \parallel^2 = (0,5) \sum_{k=1}^M (h_k^* - h_k(W))^2 \quad (h^* : \text{sortie désirée})$$

Algorithme de rétro-propagation du gradient de l'erreur quadratique $e(W)$:

$$W^{n+1} = W^{n,d} + \lambda \cdot \Delta W \quad \text{où } \Delta W = - \text{grad}_W(e(W))$$

entre la couche d'entrée I et la couche cachée J

$$\Delta W_{ij} = \varepsilon \cdot d_j \cdot s_i$$

$$d_j = f'(a_j) \cdot \left(\sum_{k=1}^M w_{jk}^0 \cdot d_k \right)$$

Sortie h_k de la couche K :

$$h_k = f \left(\sum_{j=1}^H w_{jk}^0 \cdot h_j \right) = f(a_k)$$

entre la couche cachée J et la couche de sortie K

$$\Delta W_{jk} = \varepsilon \cdot d_k \cdot s_j$$

$$d_k = 2 \cdot f'(a_k) \cdot (a_k - h_k^*)$$

Apprendre à faire apprendre

La mise en oeuvre d'une application utilisant un modèle neuronal s'effectue suivant quatre phases successives:

1 - Dans un premier temps on fait le **choix du modèle à utiliser** (type d'architecture, de processus de propagation de l'activation et de méthode d'apprentissage). Ce choix s'accomplit en fonction de la nature du problème à traiter (association entrées - sorties ou classification, données de nature temporelle ou non ... etc) et des connaissances préalables dont le concepteur dispose sur le phénomène à modéliser. Ces connaissances sont aussi nécessaires à la détermination précise de l'architecture du réseau retenu (nombre de couches de neurones, nombre de neurones dans chaque couche, connectivité).

2 - Vient ensuite la **phase d'apprentissage** qui consiste à ajuster itérativement certains paramètres du réseau (typiquement le poids des connexions) en fonction d'exemples fournis.

3 - La **phase de test** permet alors de qualifier le résultat de cet apprentissage en confrontant le réseau ainsi adapté à de nouveaux exemples et en testant ses capacités de généralisation.

4 - Ce n'est que lorsque cette phase de test est jugée satisfaisante que le **réseau pourra véritablement être utilisé**.

On voit bien que la mise en oeuvre d'une application à base de réseaux de neurones nécessite de porter un soin très attentif à la constitution de la base d'exemples. La capacité de généralisation du réseau, qui détermine son utilité pratique in fine, ne sera satisfaisante que si la base d'apprentissage est soigneusement déterminée, de taille suffisante et bien représentative du domaine sur lequel on veut utiliser le réseau.

METHODES

L'utilisation de méthodes neuronales ne signifie donc pas que l'on s'affranchit des difficultés classiques de statistique inférentielle ("représentativité" de l'échantillon test), ni qu'on puisse s'abstenir d'un certain niveau d'expertise sur le domaine d'application (choix, codage des informations pertinentes). Remarquons également que si la mise au point et l'adaptation du réseau nécessitent en général un temps de réalisation non négligeable - ainsi qu'un certain savoir faire - la complexité du réseau n'est plus un obstacle en phase d'utilisation puisqu'une simple propagation d'information dans le réseau, opération quasi instantanée, fournit la réponse du système.

La démarche neuronale est sujette comme toute démarche algorithmique d'optimisation **aux problèmes classiques de convergence** : phase d'apprentissage souvent longue (choix du bon algorithme) - existence d'optima locaux, choix délicat des paramètres de gain.

Malgré tout **la facilité de mise en oeuvre** de ces méthodes dans tout système de production au prix d'une maintenance minimale en a fait un outil pérenne, d'autant plus que leur **caractère évolutif**, autrement dit la possibilité de relancer l'apprentissage sur de nouveaux cas sans avoir à tout recommencer a joué en leur faveur.

Les Réseaux de Neurons Formels ne sont pas intelligents !

Les techniques neuronales sont des méthodes statistiques non linéaires. Les contraintes d'architecture et de traitement de l'information qu'on définit dans les réseaux confèrent à ces modèles un caractère particulier et original. On dispose de quelques résultats théoriques qui permettent de mieux situer ces méthodes, mais il est clair qu'il **manque à ce jour une théorie mathématique unifiée**. Cependant l'introduction de neurones non plus seulement à valeurs réelles ou binaires mais à valeurs complexes ou vectorielles ainsi que la définition des réseaux logiques ont montré qu'une grande généralité tant sur le plan de la forme des données utilisées que sur celui de l'architecture s'avérait possible.

L'objection la plus couramment faite est celle du **manque de transparence** de l'outil qui en ferait une boîte noire de plus. Indiquons que des travaux portant sur **l'extraction de règles** ont montré que l'on peut traduire la connaissance issue de l'apprentissage sous une forme accessible bien que réductrice.

Insistons sur le fait que ces méthodes connexionnistes ne sont pas exclusives des méthodes classiques, on peut donc les concevoir comme un "sur-modèle" du modèle linéaire, dans le sens où on peut espérer faire mieux et en tout cas pas plus mal qu'avec un modèle classique linéaire, voire créer des **modèles composites** associant des modèles neuronaux séparés et des modèles conventionnels statistiques. ■

POUR EN SAVOIR PLUS

Ouvrages généraux :

- * "Des réseaux de neurones" - E. DAVALO, P. NAÏM, chez Eyrolles, 1990. [en français]
- * "Adaptive pattern recognition and neural networks" - PAO Y.H., Addison-Wesley, Reading MA, 1989.
- * "Self-organizing maps" - KOHONEN T., Springer, 1995.

R.N. et Statistique :

- * "Networks and chaos - statistical and probabilistic aspects" - édité par O.E. BARNDORFF-NIELSEN, J.L. JENSEN et W.S. KENDALL, Monographs on Statistics and Applied Probability 50 Chapman & Hall, 1993.

R.N. et Séries Temporelles :

- * "Neural networks in finance and investing" - TRIPPI R.R., TURBAN E., Chicago : Probus Publishing Co., 1992.

R.N. et l'Analyse Des Données :

- * "Neural nets and principal component analysis" - KORNIK K., Summer School : Mathematical Fondation of Artificial neural Networks, Sion (Suisse), 1992.
- * "Neural computation and self-organizing maps : an introduction" - RITTER H., MARTINEZ T., SCHULTEN K., Addison-Wesley, Reading, 1992.

R.N. et Analyse Spatiale :

- * "Neurocomputing and spatial information processing, from general considerations to a low dimensional real world application" - FISCHER M.M., GOPAL S. Eurostat 1993.
- * "Two applications of neural networks on spatial data" - VARFIS A. Eurostat 1993.

R.N. et Transports :

- * Actes de l'I.C.A.N.N. 95, "Conférence industrielle, session Transportation - A13", octobre 1995, EC2 Editions, Paris.