

L'utilisation des réseaux de neurones artificiels en finance

Philippe Paquet

▶ **To cite this version:**

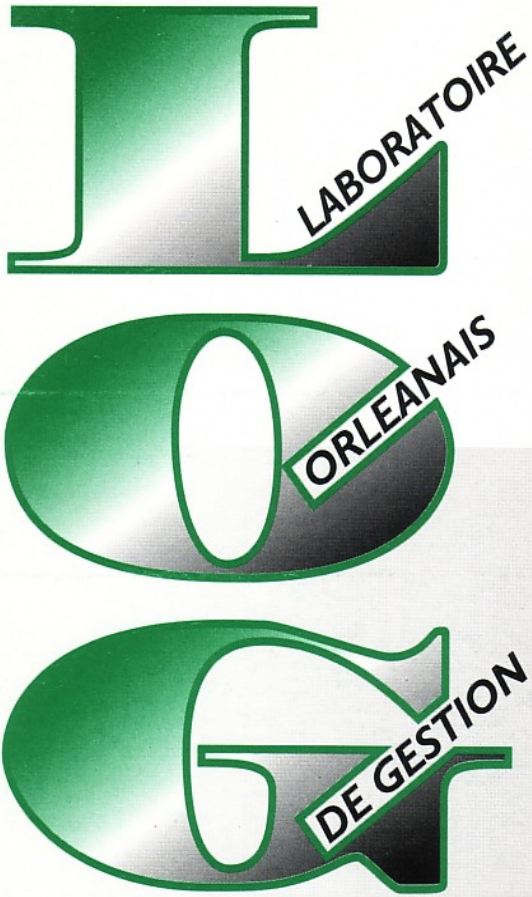
Philippe Paquet. L'utilisation des réseaux de neurones artificiels en finance. 1997. halshs-02096266

HAL Id: halshs-02096266
https://shs.hal.science/halshs-02096266

Submitted on 11 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



INSTITUT
D'ADMINISTRATION
DES ENTREPRISES

Document de recherche

N° 1997 - 1

L'utilisation des réseaux de neurones artificiels en finance.

Philippe PAQUET

LABORATOIRE ORLEANAIS DE GESTION

I.A.E
Faculté de Droit d'Economie et de Gestion
rue de Blois - B.P. 6739
45067 Orléans Cedex 2

Tél. : **02 38 41 70 28**
Fax : **02 38 41 73 60**
E.Mail : **iae@univ-orleans.fr**

L’utilisation des réseaux de neurones artificiels en finance

Philippe PAQUET
Professeur de Gestion

Résumé

Depuis le début de la décennie 1990, les réseaux de neurones artificiels habituellement utilisés en physique appliquée font leur entrée dans les sciences de gestion en tant que méthode quantitative de prévision, à côté des méthodes statistiques classiques. Ils sont en particulier utilisés en finance, mais d'autres champs de la gestion sont aussi concernés.

L'objet du présent article est d'abord de présenter succinctement l'architecture et le mode de fonctionnement de la classe de réseaux les plus couramment utilisés en finances : les réseaux à couches. Il est ensuite de montrer l'intérêt de cet outil pour les applications de finance, face aux méthodes statistiques classiques, à travers un balayage des champs d'application déjà explorés.

Enfin, la dernière partie de l'article s'attache à recenser les imperfections dont souffre encore cet outil, aujourd'hui en plein développement.

INTRODUCTION

Depuis le début de la décennie 1990, les réseaux de neurones artificiels habituellement utilisés en physique appliquée font leur entrée dans les sciences de gestion en tant que méthode quantitative de prévision, à côté des méthodes statistiques classiques. Ils sont en particulier utilisés en finance, mais d'autres champs de la gestion sont aussi concernés.

Il existe sans doute deux raisons principales qui ont conduit les chercheurs en Sciences de Gestion à s'intéresser à cet outil. La première est que contrairement aux méthodes statistiques classiques les réseaux de neurones artificiels ne nécessitent aucune hypothèse sur les variables. La seconde est qu'ils sont tout à fait adaptés pour traiter des problèmes complexes non structurés, c'est-à-dire des problèmes sur lesquels il est impossible *a priori* de spécifier la forme des relations entre les variables utilisées.

C'est grâce à des algorithmes, que ces systèmes apprennent par eux-mêmes les relations entre les variables à partir d'un ensemble de données, à l'image de ce que ferait le cerveau humain. Ainsi, le réseau se paramètre de lui-même à partir des exemples qui lui sont fournis.

Cette approche dénommée approche connexionniste se démarque des systèmes experts en ce sens que dans ces derniers, l'utilisateur est tenu d'élaborer une base de connaissances qui renferme les règles de déduction qui doivent être appliquées. La construction de cette base est souvent fastidieuse et demande un long travail de formalisation des connaissances de l'expert. L'attrait de l'approche connexionniste face aux systèmes experts réside donc bien dans le fait que les réseaux de neurones sont capables d'apprendre par eux-mêmes les relations entre les variables.

Dans le domaine financier, ces systèmes peuvent être utilisés pour traiter différentes questions. Dès 1990, Hawley et al. envisageaient leur application à des problèmes aussi divers que la détection des entreprises en difficultés, la gestion de portefeuille, l'évaluation des introductions en bourse, l'identification des opportunités d'arbitrage, l'analyse technique ou encore, l'analyse fondamentale. Depuis, un certain nombre de travaux ont montré qu'ils pouvaient être utiles dans bien d'autres domaines (Refenes, 1995). On pourrait citer, l'évaluation d'actifs, le rating des obligations ou encore l'analyse des cours boursiers. Mais il semble que ce soit dans le domaine de la détection préventive des entreprises en difficultés qu'ils fournissent aujourd'hui les meilleurs résultats.

L'objet de cet article est de s'interroger sur l'utilité de ces systèmes dans les applications de finance.

I] PRESENTATION DES RESEAUX DE NEURONES ARTIFICIELS HABITUELLEMENT UTILISES.

Les réseaux de neurones artificiels s'inspirent du fonctionnement du cerveau et du système nerveux ou, plus exactement, de la représentation que l'on s'en fait.

Il existe deux grandes catégories de réseaux :

- les réseaux à apprentissage supervisé dans lesquels le système apprend à reconnaître des formes à partir d'un échantillon d'apprentissage qui associe les modalités portées par des variables censées caractériser une forme, et la forme elle-même. Sur cet échantillon les résultats correspondant aux divers ensembles d'informations données au système sont connus. C'est à partir de là que le système se paramètre ;

- les réseaux à apprentissage non supervisé, qui sont utilisés lorsque l'utilisateur du réseau n'est pas en mesure de présenter au système un échantillon mettant en regard une somme d'informations et la forme qu'elle est censée représenter. Le réseau s'auto-organise de façon à découvrir des formes récurrentes dans les informations qu'il reçoit, mais il le fait sans aide extérieure, contrairement aux réseaux à apprentissage supervisé. Le plus connu de cette catégorie de réseaux est celui de Kohonen (1984)

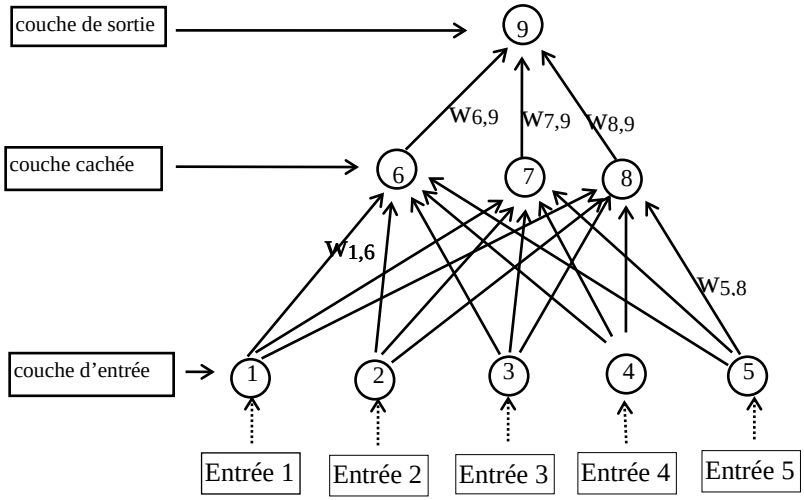
Dans le domaine de la finance, c'est le premier type de réseau que l'on utilise, le plus généralement, appelé réseau à couches.

1) Architecture des réseaux à couches

Un réseau de neurones artificiel est organisé en couches, chacune de ces couches comportant plusieurs neurones. Chacun de ces neurones qui se présente comme une unités de calcul autonome, est relié à la totalité ou à certains neurones de la ou des couches précédentes.

La figure 1 ci-dessous donne un exemple d'un réseau complètement connecté.

FIGURE 1 : Architecture d'un réseau à couches



Dans la figure 1, les cercles représentent les neurones disposés en couches. Le réseau représenté ici comporte trois couches, la couche d'entrée qui reçoit les informations sur cinq neurones, une couche de sortie comportant un seul neurone et donnant le résultat du calcul interne. Entre ces deux

couches figure une couche non visible de l'extérieur, dite couche cachée servant à effectuer des calculs intermédiaires.

Les neurones sont reliés entre eux par l'intermédiaire de poids synaptiques, notés $w_{i,j}$. L'algorithme d'apprentissage aura pour tâche d'évaluer ces poids en fonctions des données présentées lors de la phase d'apprentissage.

Il est à noter que certains réseaux plus complexes peuvent rajouter des connexions directes entre la couche d'entrée et la couche de sortie.

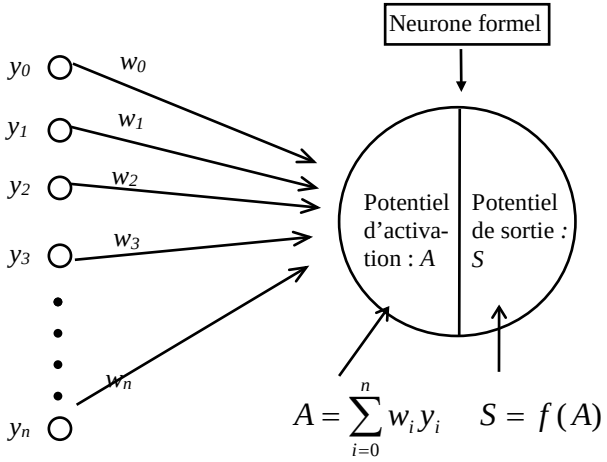
Dans le domaine qui nous intéresse, les neurones disposés sur la couche d'entrée reçoivent les informations censées expliquer le phénomène à analyser. Pour ce qui concerne la détection préventive d'entreprises en difficulté, par exemple, il s'agira le plus souvent de ratios financiers. Le neurone de sortie prendra une valeur binaire, 0 ou 1, selon que l'entreprise est considérée comme saine ou susceptible de défaillance.

2) La fonction d'activation des neurones

Chaque neurone collecte les informations fournies par les neurones de la couche précédente avec lesquels il se trouve en relation et calcule alors son potentiel d'activation. Celui-ci est ensuite transformé par une fonction pour déterminer l'impulsion envoyée aux neurones de la couche suivante (potentiel de sortie), comme cela se trouve schématisé dans la figure 2.

L'activation d'un neurone est donnée par la somme des potentiels de sortie de ses prédécesseurs, pondérée par les poids synaptiques. Ce potentiel d'activation est ensuite transformé par une fonction afin de déterminer le potentiel de sortie.

FIGURE 2 : Schéma d'un neurone artificiel

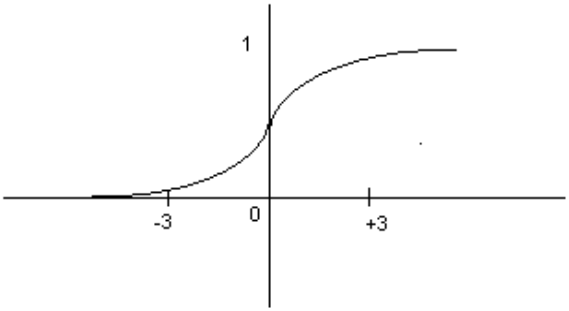


y_i, désigne les sorties des prédécesseurs du neurone,
w_i, désigne les poids synaptiques.

En principe, toute fonction choisie croissante et impaire peut être utilisée, mais le plus souvent on fait appel à des fonctions ramenant le résultat à l'intérieur de bornes prédéfinies. Plusieurs fonctions répondant à cet impératif peuvent être envisagées, mais la plus communément utilisée est la fonction sigmoïde représentée sur la figure 3, et dont l'expression analytique est la suivante :

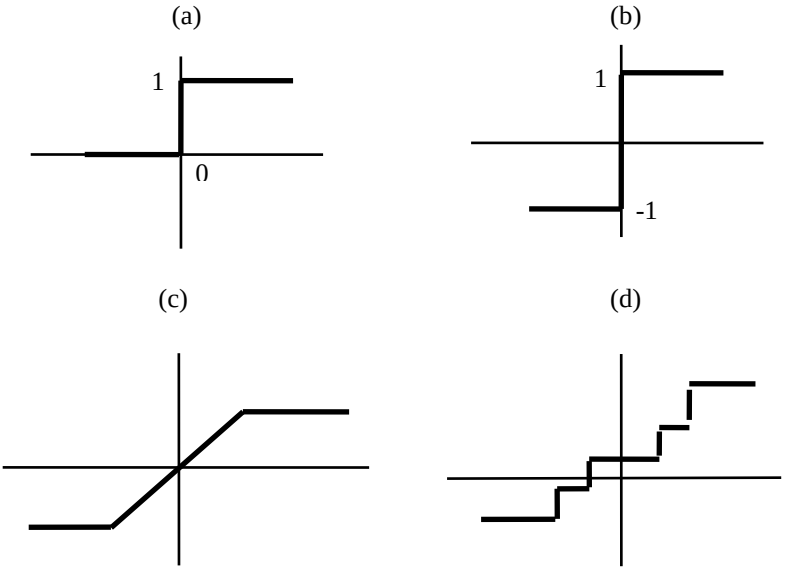
$$S = f(A) = \frac{1}{1 + e^{-A}}$$

FIGURE 3 : Fonction sigmoïde



D'autre formes de fonctions sont parfois utilisées, telles que des fonctions binaires à seuil (figure a et b) ou, linéaires à seuil (figure 4, c) ou encore des fonctions en escalier (figure 4, d).

FIGURE 4 : Exemples de fonctions d'activation



3) L'algorithme d'apprentissage

Le réseau se paramètre à l'aide d'un échantillon d'apprentissage, qui associe des formes présentées à un résultat désiré. C'est l'algorithme d'apprentissage qui ajuste les poids synaptiques en cherchant à minimiser une fonction de coût.

C'est sans aucun doute la découverte de l'algorithme de rétropropagation du gradient de l'erreur par Rumelhart, Hinton et Williams en 1986, qui a marqué le pas décisif dans l'application des réseaux de neurones artificiels. C'est encore aujourd'hui le plus utilisé dans les applications de finance.

La fonction de coût, E minimisée par cet algorithme n'est autre que la somme des carrés des erreurs produites par le réseau eu égard au résultat désiré.

Ainsi, si l'échantillon d'apprentissage comporte s exemples décrits chacun par un vecteur d'entrée de dimension m, $X_i = (x_{i,1}, x_{i,2}, ..., x_{i,m})$ et un vecteur de sortie de dimension n,
$D_i = (d_{i,1}, d_{i,2}, ..., d_{i,n})$, la fonction à minimiser s'écrit :

$$E = \sum_{i=1}^{s} \sum_{j=1}^{n} \frac{(y_{i,j} - d_{i,j})^2}{2}$$

$y_{i,j}$ représente les sorties produites par le réseau alors que $d_{i,j}$ représente les sorties désirées.

Le calcul de la quantité E, permet de déterminer dans un deuxième temps, la variation des poids synaptiques, $\Delta w_{i,j}$ de la façon suivante :

$$\Delta w_{i,j} = -\frac{\partial E}{\partial w_{i,j}} \varepsilon$$

$0<\varepsilon<1$, est un paramètre permettant de contrôler la rapidité de convergence de l'algorithme.

Après avoir initialisé les poids synaptiques de façon aléatoire, un premier calcul est effectué dans l'ordre topologique du réseau. La sortie obtenue est alors comparée à la sortie désirée et la fonction de coût, E est évaluée.

L'erreur totale commise par le système est ensuite rétropropagée de la couche de sortie vers la couche d'entrée et les poids synaptiques sont modifiés selon la formule ci-dessus, ce qui permet d'initier un nouveau calcul. Le processus se poursuit ainsi jusqu'à ce que l'utilisateur intervienne pour y mettre fin ou jusqu'à ce qu'une valeur de la fonction de coût, fixée *a priori* soit atteinte.

4) La mise en oeuvre d'un réseau de neurones artificiels

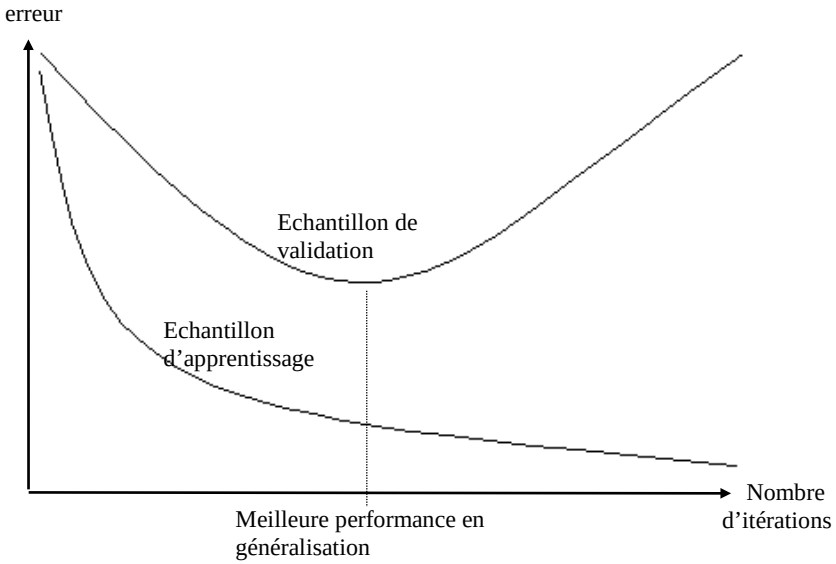
Après avoir déterminé l'architecture du réseau, sa mise en oeuvre nécessite trois échantillons de données. Le premier, destiné à l'apprentissage servira au paramètrage, le second servira à la validation et le troisième sera un échantillon test destiné à évaluer les capacités de généralisation du réseau.

Durant la phase d'apprentissage, l'erreur commise diminue, jusqu'à tendre asymptotiquement vers zéro si l'architecture du réseau a été correctement choisie. Mais il convient de remarquer que plus l'erreur est faible, c'est-à-dire plus le réseau apprend à reconnaître les formes qui lui sont présentées, moins il risque d'être capable de généralisation. En effet au-delà d'un certain point, si le réseau reconnaît de mieux en mieux les formes appartenant à l'échantillon d'apprentissage, il donne de piètres résultats sur un échantillon inconnu. On dit alors qu'il se spécialise. Or le but d'un tel système est bien de reconnaître des formes inconnues jusqu'alors. Il convient donc d'arrêter la phase d'apprentissage lorsque le réseau semble donner les meilleurs résultats en généralisation. C'est l'échantillon de validation qui permet de suivre la capacité du système à généraliser.

Typiquement, les courbes décrivant l'erreur commise sur l'échantillon d'apprentissage et sur l'échantillon de validation durant la phase d'apprentissage se présentent comme l'indique la figure 5.

Enfin, le réseau étant supposé correctement paramétré, il est appliqué à l'échantillon test de façon à mesurer sa capacité à reconnaître des formes qui n'ont jamais été vues auparavant.

FIGURE 5 : Evolution des courbes d'erreur durant la phase d'apprentissage



II] LES APPLICATIONS DES RESEAUX DE NEURONES ARTIFICIELS EN FINANCE

Avant de mentionner les domaines déjà explorés à l'aide des réseaux de neurones, on s'intéressera d'abord aux atouts qu'ils présentent.

1) l'intérêt de ces modèles

Par rapport aux modèles classiques d'analyse statistique, les réseaux de neurones artificiels possèdent plusieurs avantages :

a - ils permettent, comme cela a déjà été signalé, de traiter des problèmes non structurés, c'est-à-dire des problèmes sur lesquels on ne dispose d'aucune information *a priori*. En particulier, on notera qu'il n'est pas nécessaire de connaître la distribution de probabilité des variables, ce qui n'est pas le cas dans la plupart des modèles d'analyse statistique sauf s'il s'agit d'analyses non paramètriques ;

b - en découvrant eux-mêmes les relations entre les variables, ils sont tout à fait adaptés pour traiter des problèmes non linéaires éminemment complexes. Cet aspect est tout à fait intéressant car il n'oblige pas à s'interroger sur la forme de la fonction à estimer ;

c - ils peuvent travailler sur des données incomplètes ou des données bruitées. L'incomplétude des données peut être prise en compte par l'ajout de neurones supplémentaires. Casta et Prat (1994) montrent que des architectures spécifiquement réfléchies pour traiter les problèmes d'incomplétude des données peuvent améliorer sensiblement les résultats obtenus sur un modèle de détection préventive d'entreprises en difficulté ;

d - l'arrêt du processus itératif lorsque le réseau produit les meilleurs résultats sur l'échantillon de validation lui confère une certaine robustesse. On peut en effet considérer que de cette manière, seule l'information pertinente est intégrée dans le système. Les bruits qui pourraient concourir à la spécialisation du réseau ne sont pas intégrés ;

e - ils permettent la prise en compte de variables qualitatives à travers des neurones recevant des valeurs binaires.

Ces avantages ont conduit les chercheurs à tester cet outil sur divers problèmes financiers.

2) Quelques applications

Depuis le début de la décennie 1990, les applications financières mettant en oeuvre des réseaux de neurones artificiels se sont multipliées. D'une façon générale, elles tombent dans l'une des trois catégories suivantes : la prévision des séries temporelles, les substituts aux systèmes experts, et les applications de classification.

2.1 - La prévision des séries temporelles

La prévision des séries temporelles semble constituer aujourd'hui un champ d'investigation privilégié, à tel point que Azoff (1994) lui consacre un ouvrage entier. Typiquement, dans les prévisions univariées, il s'agit d'utiliser le passé d'une variable afin d'en extraire des relations permettant de prédire sa valeur future. Mais, à supposer que ces relations existent

réellement, le problème le plus délicat à résoudre est d'en déterminer leur forme qui, le plus souvent n'est pas linéaire. Cela a d'ailleurs conduit à développer ces dernières années, des modèles de prévision statistiques non linéaires. Mais, les chercheurs qui se tournent vers les réseaux de neurones artificiels comptent sur leur capacité à découvrir des formes récurrentes dans les séries, dont la complexité pourrait échapper aux modèles statistiques.

L'arrivée de ce nouvel instrument d'analyse permet en particulier de relancer la question de l'efficience des marchés. On s'interroge, en effet, sur la question de savoir si la puissance des réseaux de neurones ne serait pas de nature à mettre en lumière certaines formes d'inefficience qui auraient, jusqu'à maintenant, échappées aux outils d'analyse classiques.

En utilisant simplement le signe des taux de rendement passés de six titres américains, Tsibouris et Zeidenberg (1995) reconnaissent un certain pouvoir prédictif au réseau testé. Cependant, si les résultats sont satisfaisants sur l'échantillon d'apprentissage, ils sont moins convaincants sur l'échantillon test, comme l'avait déjà remarqué White (1988) à partir du titre IBM. Plus pertinents sont les résultats obtenus par Avouyi-Dovi et Caulet (1995) qui obtiennent des performances remarquables en termes de prévision d'indices boursiers (DAX, CAC40) et de taux de change (Mark/Dollar), face à des modèles statistiques traditionnels de type ARMA.

Plus récemment, les réseaux de neurones ont été utilisés pour prévoir la volatilité des indices boursiers des grandes places financières (Donaldson et Kamstra, 1996).

Mais, c'est naturellement en multipliant les variables d'entrée que les réseaux fournissent les meilleurs résultats. Ainsi, Kimoto et Yoda (1993) rapportent qu'à partir d'un réseau comprenant 5 variables en entrée, ils arrivent à prédire de façon relativement satisfaisante l'évolution du TOPIX (Tokyo Stock Index) sur la période 1989-1992. Une simulation de stratégie d'investissement dérivée des résultats du modèle aurait rapporté une rentabilité moyenne annuelle de 1,60% alors que dans le même temps l'indice chutait de 20,63% !

Des tests furent aussi entrepris avec quelques succès à partir des contrats à terme sur l'or et sur l'indice S&P (Grudnitski et Osburn, 1993 ; Grudnitski et Quang Do, 1995) en utilisant un réseau comportant des variables contextuelles.

Enfin, pour rester dans le domaine des marchés financiers, signalons encore que des travaux ont porté sur les options. En particulier, Hutchinson et al. (1994) montrent qu'un réseau entrainé sur un échantillon de données journalières observées sur deux années, retrouve la formule de Black and Scholes (1973).

La question de la prévision des taux de change, a aussi été exploré à plusieurs reprises (Rawani et al., 1993 ; Refenes et al., 1993 ; Azoff, 1994 ; Mehta, 1995 ; Avouyi-Dovi et Caulet, 1995).

En particulier, un travail original de Refenes et Zaidi (1995) montre comment les prévisions de taux de change peuvent être améliorées en combinant les recommandations issues de méthodes classiquement utilisées par les professionnels, dans un réseau. Ainsi, sur le taux de change USD/DEM, l'application des recommandations du réseau donne une rentabilité de 18% sur une base annuelle, lorsque les stratégies de moyenne mobile et valeur moyenne donnent des rentabilités de 12,3% et de 13,1%.

2.2 - Les réseaux de neurones, substituts des systèmes experts

La facilité d'utilisation des réseaux de neurones artificiels et leur capacité à reconnaître des formes après apprentissage, en font un outil qu'il est tentant de tester sur des problèmes habituellement traités par des systèmes experts. Si leur capacité prédictive se révélait satisfaisante, cela représenterait un bénéfice considérable car il ne serait plus nécessaire de dépenser un temps précieux à l'élaboration de la base de règles qui constitue le coeur d'un système expert.

Les travaux qui peuvent être classés dans cette catégorie sont encore assez peu nombreux. Ils portent sur les prévisions de fusions-acquisition d'entreprises et surtout sur la notation des obligations.

Dans le premier de ces domaines, Sen et al. (1995) testent les capacités prédictives d'un réseau de neurones contre celles d'une régression logistique en utilisant les variables habituellement rencontrées dans la littérature portant sur le sujet. Il apparaît que si le réseau, mathématiquement, ajuste mieux les données que la régression logistique, il se révèle tout aussi incapable de prévoir correctement les fusions-acquisitions.

S'agissant de la notation des obligations, les travaux effectués ont cherché à reproduire les notations affectées par les agences de rating américaine, Moody's et Standard and Poor's (Dutta et Shashi, 1988 ; Singleton et Surkan, 1991 ; Moody et Utans,1995). Les systèmes construits apprennent, à partir d'un échantillon, à reproduire les notations des experts. Si les résultats obtenus sont plus encourageant que dans le domaine des fusions-acquisitions, ils montrent en même temps que les modèles ont encore besoin d'être affinés.

Il faut en particulier remarquer que dans ces deux domaines de recherche, il est vraisemblable que les seules variables internes qui ont été retenues possèdent un pouvoir explicatif très insuffisant. La complexité du processus amenant une entreprise à devenir la cible d'une opération de fusions-acquition ou la complexité du raisonnement mené par les experts des agences de rating nous laisse penser que ces résultats seraient sans doute améliorés par l'introduction de variables contextuelles, y compris qualitatives à l'instar de celles qui sont habituellement intégrées dans les systèmes experts. Or, le réseau se prête beaucoup mieux que les techniques d'analyses statistiques à la prise en compte de telles informations, ce qui laisse entrevoir des amélioration futures, alors que ce champ d'investigation n'en est encore qu'à ses balbutiements.

Il semblerait d'ailleurs qu'un modèle intégrant à la fois, réseau de neurones et représentation symbolique de la connaissance de l'expert, soit déjà opérationnel chez Nikko Securities depuis octobre 1992. Ce système produirait des réponses exactes dans 96% des cas sur des données hors de l'échantillon d'apprentissage (Asakawa, 1993).

2.3 - La classification

C'est sans doute dans ce domaine que les réseaux de neurones artificiels donnent aujourd'hui les meilleurs résultats lorsqu'ils sont comparés aux techniques classiques habituellement utilisées : l'analyse discriminante et la régression logistique. Plusieurs études ont été réalisées, portant sur la détection préventive des entreprises en difficulté.

Le premier travail visant à comparer la performance d'un réseau de neurones artificiels avec l'analyse discriminante remonte à Odom et Sharda (1990). A partir d'un échantillon de 65 firmes considérées comme défaillantes entre 1975 et 1982 et de 64 firmes considérées comme saines,

appariées par secteur d'activité et par année, ils calculent 5 ratios financiers, censés caractériser les entreprises.

Dans cette étude, le réseau utilisé donne de meilleurs résultats que l'analyse discriminante sur l'échantillon test. En effet il classe correctement 81,81% des entreprises contre seulement 74,28% pour l'analyse discriminante.

Avec un échantillon de 94 entreprises, Coats et Fant (1992) concluent dans le même sens, avec un réseau comprenant une couche cachée et 5 ratios financiers observés sur trois années, soit 15 neurones en entrée.

Il faut cependant remarquer que la taille de l'échantillon dans ces deux études apparaît ici comme étant relativement faible, pour tirer des conclusions solides.

A partir d'un échantillon beaucoup plus important, plus de 800 entreprises italiennes, Altman et al. (1994) testent des réseaux de complexités différentes, avec plusieurs batteries de ratios. Ils obtiennent des taux de reconnaissance de 97,7% pour les entreprises saines et de 97,0% pour celles qui sont considérées comme se trouvant en difficulté. Ces résultats apparaissent nettement meilleurs que ceux obtenus par l'analyse discriminante qui donne des taux de reconnaissance, respectivement, de 90,3% et de 86,4%. Poddig (1995) rapporte des résultats un peu moins probants à partir de tests effectués sur des entreprises françaises, mais note qu'en moyenne, les réseaux de neurones artificiels dominent l'analyse discriminante. Les résultats obtenus par Bardos et Zhu (1995), travaillant sur des données de PME françaises, apparaissent aussi encourageants de ce point de vue.

Avec 300 entreprises pour l'échantillon d'apprentissage, 100 pour l'échantillon test et 16 ratios financiers, Udo (1993) conclut à la supériorité des réseaux de neurones artificiels face à une technique de régression multiple. De façon similaire, de Almeida et Dumontier (1993), montrent à partir d'un plan d'expérience construit sur des entreprises françaises du secteur des transports, qu'un modèle de régression de type LOGIT est généralement dominé par les réseaux de neurones artificiels.

Dans le domaine des institutions financières, Tam et Kiang (1992) s'intéressent aux banques du Texas ayant connues des difficultés sur la période 1985-1987. Chacune d'entre elles est caractérisée par 19 ratios. Les

résultats obtenus montrent, à une exception près, qu'un réseau comprenant une couche cachée de 10 neurones réalise de meilleures prédictions que l'analyse discriminante ou la régression logistique. Salchenberger et al. (1992) réalisent une étude portant sur les Savings and Loans Associations aux Etats Unis. La période observée s'étend de janvier 1986 à décembre 1987. Le réseau utilisé comporte 5 neurones d'entrée et une couche cachée de 3 neurones. Les résultats obtenus à partir de prévisions faites à plus ou moins longue échéance confirment ceux de l'étude précédente : le réseau donne toujours des prévisions au moins aussi bonnes que le meilleur modèle de régression logistique.

Pourtant, si au vu de ces applications diverses, l'outil semble prometteur, il n'est pas aujourd'hui exempt de critiques.

III) LES CRITIQUES HABITUELLEMENT ADRESSEES AUX RESEAUX DE NEURONES

Plusieurs limites dans l'utilisation des réseaux de neurones peuvent être retenues.

1) L'architecture optimale

On mentionnera qu'il n'existe pas encore de théorie permettant de déterminer la structure optimale d'un réseau. En particulier la détermination du nombre de couches cachées et du nombre de neurones sur chacune d'entre elle relève encore largement de l'intuition de l'utilisateur et de sa capacité à expérimenter plusieurs architectures afin de retenir celle qui donne les résultats les meilleurs. Plus le réseau est complexe, c'est-à-dire plus il comporte de neurones cachés et de liens synaptiques, plus il est capable de reconnaître les formes qui lui sont présentées à travers l'échantillon d'apprentissage. Mais malheureusement, l'accroissement de la complexité n'améliore pas nécessairement le taux de reconnaissance sur l'échantillon test.

Il faut cependant remarquer que cette critique doit être nuancée, car plusieurs méthodes ont été proposées pour aider l'utilisateur dans cette tâche. Celles-ci s'appuient sur des techniques analytiques faisant intervenir la dimension du vecteur d'entrée, ou sur des algorithmes qui permettent soit la construction automatique pas à pas des couches cachées, soit l'élagage

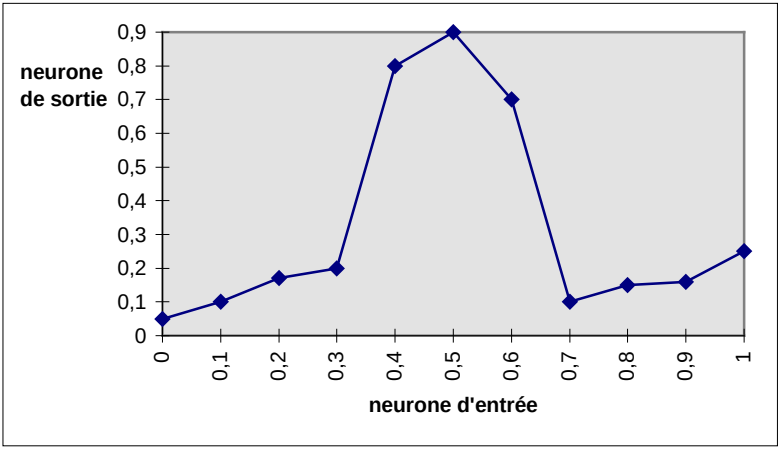
des connexions les moins pertinentes dans un réseau surdimensionné au départ (Refenes, 1995).

2) Le pouvoir explicatif des réseaux

Un réseau de neurones reste encore en partie aujourd'hui une « boîte noire » de laquelle il reste difficile d'extraire les relations pertinentes entre les variables. De ce point de vue, le système n'a donc qu'un pouvoir explicatif médiocre contrairement aux systèmes experts qui sont capables de retracer le cheminement suivi pour atteindre le résultat.

Pourtant, le phénomène de « boîte noire » n'est pas total puisque des analyses de sensitivité sont possibles, en faisant varier l'une des entrées pas à pas, de sa valeur minimale vers sa valeur maximale, les autres entrées du réseau restant figées à leur valeur moyenne. Cette méthode a notamment été utilisée par Refenes et al. (1994) et Sen et al. (1995). Elle permet de visualiser successivement l'impact de chacune des variables d'entrée sur le ou les neurones de sortie. La figure 6 donne un exemple de résultat obtenu par une analyse de sensibilité. Le neurone d'entrée testé ainsi que le neurone de sortie prennent des valeurs d'un minimum de zéro, jusqu'à un maximum de 1. La courbe montre ici que la variable d'entrée produit une activation maximale du neurone de sortie lorsque sa valeur se situe autour du point milieu (0,5). Les valeurs extrêmes ont au contraire peu d'influence sur le résultat.

FIGURE 6 : Exemple de résultat produit par une analyse de sensibilité



3) Le retraitement des données

Il faut signaler que les données d'entrée d'un réseau doivent fréquemment être retraitées afin de les normaliser, et de les stationnariser sous peine d'obtenir des résultats décevants. La normalisation vise à ramener la plage d'évolution des valeurs prises par les variables à l'intérieur d'un intervalle standardisé, fixé *a priori*. Elle est souhaitable car elle évite au système de se paramétrer sur une plage de valeurs particulières, ignorant ainsi les valeurs extrêmes.

Ce phénomène peut se comprendre en examinant les résultats que produirait une fonction d'activation de type sigmoïde telle que celle décrite par la figure 3, p. 8. Toute valeur en deçà de -3 ou au-delà de + 3 produira toujours un résultat très proche de 0 ou de 1. La réponse de la fonction devient très peu sensible à l'intensité de l'activation : une activation de + 4 donnera sensiblement la même réponse qu'une activation de + 15, c'est-à-dire un résultat très proche de l'unité. C'est dans la zone [-3, +3] que la sensibilité est la plus forte. On dit que la région opératoire normale de la fonction se situe alors entre ces bornes.

Deux retraitements sont fréquemment utilisés. Le premier consiste à centrer et réduire la variable. Chaque observation est transformée de la façon suivante :

$$x_1 = \frac{x - \mu_x}{\sigma_x}$$

x, représente l'observation d'origine, μ_x la moyenne des observations de la variable et σ_x son écart type.

Le deuxième retraitement utilisé consiste en une remise à l'échelle de la variable de la façon suivante :

$$x_1 = \frac{x - \min_x}{\max_x - \min_x}$$

$\min_x$ et $\max_x$, représentent respectivement, le minimum et le maximum de la variable.

La stationarisation des séries est souvent aussi une nécessité, comme elle l'est dans les modèles d'analyse statistiques classiques car les trends intégrés dans les variables seront appris par le réseau, bien qu'ils n'aient aucun pouvoir prédictif.

4) L'intervention humaine

Comme cela a été signalé pour ce qui est du choix de l'architecture, ces systèmes font toujours trop largement appel à l'intuition de l'utilisateur. En particulier, l'apprentissage est guidé par des paramètres qu'il convient de régler manuellement. Un taux d'apprentissage (paramètre réglant la vitesse de convergence de l'algorithme) trop important peut aboutir à une oscillation du système alors qu'un taux trop faible se traduira par une convergence très lente, ce qui peut se révéler prohibitif en temps de calcul. Par ailleurs, il revient encore à l'utilisateur de choisir le point d'arrêt de l'apprentissage de façon à ce que le réseau conserve ses capacités de généralisation.

CONCLUSION

Au terme de cet article il apparaît que les réseaux de neurones artificiels donnent aujourd'hui de bon résultats dans les applications de finance. Il semblerait que ce soit dans le domaine de la classification, en particulier, dans le champ de la détection préventive des entreprises en difficulté qu'ils soient les plus performants. Les travaux menés sur le sujet, tant en France qu'aux Etats Unis semblent s'accorder sur ce point.

Cependant, ils présentent encore certaines limites : la construction du réseau, le retraitement des variables d'entrée et le réglage des paramètres d'apprentissage relèvent encore trop de l'intervention humaine. Par ailleurs ils n'ont encore qu'un pouvoir explicatif réduit. Mais tous ces problèmes font aujourd'hui l'objet de recherches dont on peut espérer qu'elles apportent des réponses satisfaisantes dans les années qui viennent. Mais, comme le signale Hoptroff (1993), si l'outil est intéressant, comme toutes les méthodes de prévisions, les réseaux de neurones ne pourront jamais prévoir l'imprévisible.

BIBLIOGRAPHIE

ALTMAN, E. I., MARCO, G. and VARETTO, F. (1994), Corporate Distress Diagnosis : Comparisons Using Linear Discriminant Analysis and Neural Networks (The Italian Experience*), Journal of Banking and Finance*, 18, pp. 505-529.

ASAKAWA, K. (1993), Neurocomputers and their Applications, *Fujitsu Scientific and Technical Journal*, 29, N° 3, pp. 195-200.

AVOUYI-DOVI, S. et CAULET, R. (1995), *Les réseaux de neurones artificiels : une application à la prévision des prix des actifs financiers*, document de travail, N° 1995-19/T, Caisse des dépôts et consignations, 23 p.

AZOFF, E., M. (1994), *Neural Network Time Series Forecasting of Financial Markets*, Wiley Finance Edition, New York, 196 p.

BARDOS, M. et ZHU, W. (1995) Comparaison de l'analyse discriminante linéaire et des réseaux de neurones : application à la détection de défaillance d'entreprises, *in Actes de la deuxième rencontre internationale sur l'approche neuronale en Sciences de Gestion*, Poitiers, pp. 55-73.

BLACK, F. and SCHOLES, M. (1973), The Pricing of Options and Corporate Liabilities, *Journal of Political Economy*, 81, pp. 637-659.

CASTA, J.-F., et PRAT, B. (1994), *Approche connexionniste de la classification des entreprises : contribution au traitement d'informations incomplètes*, Document dactylographié, CEREG, Université de Paris-Dauphine

COATS, P., K. and FANT, L. (1992) A Neural Network Approach to Forecast Financial Distress, *The Journal of Business Forecating*, 10, Winter 1991-92, pp. 9-12.

De ALMEIDA, F., C. and DUMONTIER, P. (1993), Réseaux de Neurones, information comptable et détection du risque de défaillance, in *Actes du XIV Congrès de l'Association Française de Comptabilité*, Vol. I, pp. 269-286.

DONALDSON, R., G.and KAMSTRA, M. (1996), Forecast Combining with Neural Networks, *Journal of Forecasting*, 15, pp. 49-61.

DUTTA, S. and SHASHI, S. (1988), Bond Rating : A Non-Conservative Application of Neural Networks, in *Proc. ICNN-88, San Diego*, July 1988, Vol II, pp. 443-450.

GRUDNITSKI, A. and OSBURN, L. (1993), Forecasting S&P and Gold Futures Prices : An Application of Neural Networks, *The Journal of Futures Markets*, 13, N° 6, pp. 631-643.

GRUDNITSKI, G. and QUANG DO, A. (1995), Important Factors in Neural Networks's Forecasts of Gold Futures Prices, in Refenes A.-P. (ed), *Neural Network in the Capital Markets*, Wiley Finance Edition, New York, pp. 163-173.

HAUWLEY, D., D., JOHNSON, J. D., and RAINA, D. (1990), Artificial Neural Systems : A New Tool for Financial Decision-Making, *Financial Analyst Journal*, Nov.-Déc. pp. 63-72.

HOPTROFF, R., G. (1993), The Principles and Practice of Time series Forecasting and Business Modelling Using Neural Nets, *Neural Computing and Applications*, 1, 1993, pp. 59-66.

HUTCHINSON, J., M., LO, A., W. and POGGIO, T., (1994), A Nonparametric Approach to Pricing and Hedging Derivative Securities Via Learning Networks, *The Journal of Finance*, Vol. XLIX, pp. 851-889.

KIMOTO, T. and YODA, M. (1993), Buying and Selling Timing Prediction System for Stocks Based on Modular Neural Networks, *Fujitsu Scientific and Technical Journal*, 29, N° 3, pp. 257-264.

KOHONEN, T. (1984), *Self-Organisation and Associative Memory*, Springer-Verlag, Berlin.

METHA, M. (1995), Foreign Exchange Market, , in Refenes A.-P. (ed), *Neural Network in the Capital Markets*, Wiley Finance Edition, New York, pp. 177-198.

MOODY, J. and UTANS, J. (1995), Architecture Selection Strategies for Neural Networks : Application to Corporate Bond Rating Prediction, in Refenes A.-P. (ed), *Neural Network in the Capital Markets*, Wiley Finance Edition, New York, pp. 277-307.

ODOM, M. D. and SHARDA, R. (1990), A Neural Network Model for Bankruptcy Prediction, in *International Joint Conference on Neural Networks, San Diego*, CA, 2, pp. 163-167.

PODIG, T. (1995) Bankruptcy Prediction : A Comparison with Discriminant Analysis, in Refenes A.-P. (ed), *Neural Network in the Capital Markets*, Wiley Finance Edition, New York, pp. 311-323.

RAWANI, A., M., MOHAPATRA, D., K., SRINIVASAN, S., MOHAPATRA, P., K., MEHTA, M., S. and RAO, G., P. (1993), Forecasting and Trading Strategy for the Foreign Exchange Market, *Information and Decision Technologies*, 19, pp. 55-62.

REFENES, A.-P. (1995), *Neural Networks in the Capital Markets*, Apostolos-Paul Refenes Editor, Wiley Finance Edition, New York, 376 p.

REFENES, A.-P., AZEMA-BARAC, M., CHEN, L. and KAROUSSOS, S., A. (1993), Currency Exchange Rate Prediction and Neural Network Design Strategies, *Neural Computing and Applications*, 1, pp. 46-58.

REFENES, A.-P. and ZAIDI, A. (1995), Managing Exchange-Rate Prediction Strategies with Neural Networks, in Refenes A.-P. (ed), *Neural Network in the Capital Markets*, Wiley Finance Edition, New York, pp. 213-219.

REFENES, A.-P., ZAPRANIS, A. and FRANCIS, G. (1994), Stock Performance Modeling Using Neural Networks : A Comparative Study with Regression Models, *Neural Networks*, 7, N°2, pp. 375-388.

RUMELHART, D. E., HINTON, G. AND WILLIAMS, R. (1986), Learning Representation by Back-Propagation Errors, *Nature*, 323, pp. 533-536.

SALCHENBERGER, L., M., CINAR, E., M. and LASH, N., A. (1992), Neural Network : A New Tool for Predicting Thrift Failures, *Decision Sciences*, 23, pp. 899-916.

SEN, T., K., OLIVER, R. and SEN, N. (1995), Predicting Corporate Mergers, in Refenes A.-P. (ed), *Neural Network in the Capital Markets*, Wiley Finance Edition, New York, pp. 325-340.

SINGLETON, J. C. and SURKAN, A., J. (1991), Modelling the Judgement of Bond Rating Agencies : Artificial Intelligence Applied to Finance, *Journal of the Midwest Finance Association*, 20, pp. 72-80.

TAM, K., Y. and KIANG, M., Y. (1992), Managerial Applications of Neural Networks : The Case of Bank Failure Predictions, *Management Science*, 38, N° 7, pp. 926-947.

TSIBOURIS, G. and ZEIDENBERG, M. (1995), Testing Efficient Market Hypothesis with Gradient Descent Algorithms, in Refenes, A.-P. (ed), *Neural Network in the Capital Markets*, Wiley Finance Edition, New York, pp. 127-136.

UDO, G. (1993), Neural Network Performance on the Bankruptcy Classification Problem, *Computer and Industrial Engineering*, 25, N° 4, pp. 377-380.

WITHE, H. (1988), Economic Prediction Using Neural Networks : The Case of IBM Daily Stock Returns, in *Proc. of the IEEE International Conference on Neural Networks*, San Diego, Vol. 2, pp. 451-459.