HAL
open science

APPROCHE CONNEXIONNISTE DE LA CLASSIFICATION DES ENTREPRISES: contribution au traitement d'informations incomplètes

Jean-François Casta, Bernard Prat

To cite this version:

Jean-François Casta, Bernard Prat. APPROCHE CONNEXIONNISTE DE LA CLASSIFICATION DES ENTREPRISES: contribution au traitement d'informations incomplètes. Recherches en comptabilité internationale, May 1994, France. pp.cd-rom. hal-00818712

HAL Id: hal-00818712

https://hal.science/hal-00818712v1

Submitted on 19 Sep 2014

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.

APPROCHE CONNEXIONNISTE DE LA CLASSIFICATION DES ENTREPRISES : contribution au traitement d'informations incomplètes

Jean-François CASTA

Bernard PRAT

L'évaluation du risque de défaillance des entreprises a fait l'objet depuis trois décennies, et plus particulièrement jusqu'au milieu des années 80, de très nombreux travaux de nature empirique utilisant, pour la plupart, l'analyse discriminante linéaire (ADL). Ces recherches ont suscité un grand nombre de critiques d'ordre méthodologique, favorisant l'émergence de diverses problématiques alternatives. La plus récente d'entre elles, l'approche connexionniste, se réfère à la dynamique des systèmes complexes et au paradigme neuro-cybernétique ; les travaux qui se situent dans cette perspective tendent à appliquer les propriétés d'apprentissage, de généralisation (ou d'auto-organisation) des réseaux de neurones au domaine de la prévision des défaillances d'entreprises.

Paradoxalement, eu égard au degré de raffinement des techniques statistiques utilisées, plus que dans toute autre démarche d'induction numérique, les travaux empiriques relatifs aux défaillances d'entreprises restent tributaires de la qualité des données et en particulier de leur incomplétude. Face à ce problème, deux solutions s'offrent : éliminer les entreprises dont les données sont incomplètes ou utiliser des méthodes économétriques afin de compléter les séries. L'élimination des entreprises de l'échantillon, procédé le plus fréquemment utilisé en matière de défaillances, introduit un biais méthodologique dès lors que les problèmes d'incomplétude affectent inégalement les entreprises en situation difficile et les autres.

Ces considérations nous ont amené à explorer les apports possibles de l'approche connexionniste, et en particulier des classifieurs neuronaux, afin de déterminer leur aptitude potentielle à traiter les informations incomplètes. Pour ce faire, nous avons élaboré une architecture spécifique de classifieurs neuronaux pour base d'exemples incomplète et testé ses performances par rapport à l'approche neuronale classique. La

base d'exemples ainsi que les jugements d'experts associés ont été mis à notre disposition par la société S.C.R.L.[1].

1 - L'évaluation du risque de défaillance : état de l'art

A partir d'une représentation de l'entreprise d'essence pragmatique, reposant sur les ratios financiers et comptables, les travaux relatifs à l'évaluation du risque de défaillance, guidés par la recherche d'une amélioration des performances des modèles de prévision, ont successivement exploré les possibilités offertes par différentes techniques statistiques d'induction numérique[2]. Dépassant les limites inhérentes à toute approche univariée (Beaver, 1966), les travaux pionniers ont très rapidement été caractérisés par l'introduction de méthodes d'analyse multivariée et plus particulièrement d'analyse discriminante linéaire [(Altman, 1968), en France : (Conan & Holder, 1979), (Ghesquière & alii, 1983) (Bardos, 1984)], puis accessoirement par le recours à la régression logistique [(Ohlson, 1980), en France : (Bardos, 1989)], ou par l'utilisation du partitionnement récursif [(Frydman & alii, 1985), en France : (Bardos, 1989)].

Entrés dans une phase d'exploitation opérationnelle depuis une dizaine d'années, ces travaux, se référant presque exclusivement à l'analyse discriminante linéaire (ADL) à deux classes, ont fait l'objet de nombreuses études critiques[3] visant : la démarche méthodologique adoptée, la nature ou la représentativité des informations traitées, la validité des indicateurs de mesure du pouvoir prédictif.

En premier lieu, souvent soulignée, l'absence ou la faiblesse de la modélisation théorique du processus de défaillance (Malécot, 1991) a eu pour corollaire le développement d'une problématique de recherche purement inductive. Affranchie de toute démarche de nature symbolique, qui consisterait à identifier au plan sémantique des concepts et à modéliser leurs relations, l'approche de la défaillance des entreprises a pu être considérée comme purement tautologique. En effet, le concept de défaillance lui-même n'est souvent défini qu'à partir d'une référence aux variables (les ratios) utilisées pour sa mesure et retenues pour spécifier la partition en deux classes.

En second lieu, de nombreuses critiques ont porté sur la méthode de constitution des échantillons (représentativité, tirage aléatoire ou non, taille de l'échantillon), sur la méthodologie de validation (échantillon d'étude, échantillon de test). En particulier, ces critiques ont souvent fait ressortir l'insuffisance notoire de la taille de l'échantillon d'entreprises en difficulté, phénomène dû à leur faible proportion dans la population, mais aussi à l'élimination d'entreprises en raison de l'impossibilité de réunir des données complètes sur l'ensemble de la période d'observation.

En troisième lieu, la nature même des variables utilisées (les ratios) par la quasi-totalité des recherches sur les défaillances d'entreprises pour spécifier la fonction

---

[1] Les auteurs tiennent particulièrement à remercier Monsieur Lartisien, Président de S.C.R.L. pour le soutien qu'il a accordé à ce travail de recherche.

[2] Sur l'évolution des modèles de prévision des faillites, ainsi que sur les problèmes méthodologiques, voir notamment : (Casta, 1985), (Fioleau, 1993). Pour une revue de la littérature relative aux modèles empiriques de prévisions de défaillance voir notamment : (Zavgren, 1983), (Barnes, 1987), (Dumontier, 1991).

[3] Pour une synthèse de ces critiques, voir notamment : (Casta & Zerbib, 1979), (Malécot, 1986), (Dumontier 1991), (Fioleau 1992).

discriminante linéaire pose le problème du choix de ceux-ci, de leur éventuelle corrélation et de son incidence sur la performance de la discrimination, enfin du traitement des données extrêmes ou aberrantes (Watson, 1990). Au-delà de ces considérations, la place exclusive laissée à des variables strictement quantitatives[4], comme les ratios financiers, accentue le caractère réducteur de l'approche empirique.

En dernier lieu, les critiques les plus nombreuses ont trait aux mauvaises conditions d'utilisation de l'instrument statistique[5] qui, en engendrant un risque d'interprétation erronée, biaiserait les résultats et les mesures de performances [(Joy & Tofelson, 1975), (Eisenbeis, 1977), (Ohlson, 1980)]. Par ailleurs, ces travaux se sont attachés à mieux définir les conditions de la mesure du pouvoir prédictif de la fonction discriminante. En particulier, elles ont mis en lumière l'influence de l'asymétrie existant entre les deux groupes tant au niveau du coût des erreurs de reclassement qu'au plan de la proportion d'entreprises (Eisenbeis, 1977).

Au regard de telles critiques, les travaux de la dernière décennie, tout en restant pour la plupart dans le cadre de l'ADL, se sont orientés dans des directions tendant à satisfaire aux conditions visant à atteindre une discrimination optimale : accroissement de la taille des échantillons, normalisation et élimination des données aberrantes, mise en œuvre de test (multicolinéarité des ratios, multinormalité des distributions ...). Malgré la lourdeur de ces études, la difficulté de satisfaire en pratique, même avec des échantillons très importants, aux exigences de multinormalité et d'équivariance - covariance reste toujours aussi importante (Malécot, 1986). Ces difficultés conduiraient logiquement à prescrire l'utilisation de techniques d'analyse discriminante non linéaire, par exemple avec fonction de séparation quadratique (Eisenbeis & Avery, 1972) bien que de telles techniques soient réputées peu robustes et d'une mise en œuvre très délicate.

Devant les problèmes posés par la mise en œuvre de l'ADL, quelques recherches ont développé des solutions alternatives[6] : soit en s'affranchissant de certaines contraintes statistiques par l'introduction, avec la régression logistique (Ohlson, 1980), d'une forme de non linéarité, soit plus radicalement, en recourant à une approche non paramétrique ne nécessitant plus aucune hypothèse sur les distributions [(Frydman & alii, 1985), (Barniv, 1989)]. Cependant les études comparatives menées à la Centrale de Bilans de la Banque de France ne permettent pas de mettre en évidence la supériorité de telles approches (Bardos, 1989).

A ce stade, il convient de se demander si les problèmes soulevés par la mise en œuvre pratique de l'ADL en matière de risque de défaillance ne relèvent pas plus de l'épistémologie que de la statistique. En d'autres termes, la représentation des formes (les entreprises) étant altérée par le caractère réducteur (et strictement quantitatif) des données retenues (les ratios), par le bruitage affectant les données ou par la présence de données incomplètes, n'est-il pas illusoire et coûteux de rechercher l'optimalité ? Ne serait-il pas opportun de réintroduire le principe de parcimonie dans une approche inductive : notamment en privilégiant la recherche de solutions simplement *satisfaisantes*, mais capables de généralisation, obtenues à partir d'échantillons *limités* mais représentatifs reposant sur une représentation des formes plus riche. C'est dans cette perspective que nous nous proposons d'exposer les potentialités offertes par

---

[4] Il existe des méthodes de discrimination sur variables qualitatives. Elles posent cependant de délicats problèmes de codage et de mise en œuvre.

[5] L'optimalité de la classification résultant de l'hyperplan obtenu par ADL est conditionnée par le respect des hypothèses définies par Fisher : multinormalité des distributions dans chacun des groupes, égalité de la matrice des variances-covariances des groupes, connaissance des moyennes, des variances-covariances et des probabilités a priori.

[6] Pour une synthèse, voir : (Fioleau, 1993).

l'approche connexionniste, et plus particulièrement par la technique des réseaux de neurones, à laquelle sont attachées les propriétés d'apprentissage à partir d'exemples, de généralisation, ainsi que de traitement de données bruitées, incomplètes, imprécises ou hétérogènes (Davalo & Naïm, 1992).

2 - L'apport des classifieurs neuronaux

2.1. L'approche connexionniste

L'approche connexionniste se réfère fondamentalement à la dynamique des systèmes complexes et renvoie au paradigme neuro-cybernétique (Lemoigne, 1991). Elle trouve dans les réseaux de neurones formels une forme féconde d'implémentation opérationnelle pour concevoir des classifieurs.

Un neurone formel i est essentiellement un automate de calcul. A partir des signaux d'entrée E_j qu'il reçoit des n neurones auxquels il est relié, à partir des poids synaptiques W_{ij} associés à chaque connexion et à l'aide de sa fonction d'activation F_i le neurone déterminera son nouvel état interne et produira un signal de sortie S_i tel que :

$$S_i = F_i\left(\sum_{j=1}^{n} W_{ij}\ E_j\right).$$

Un réseau de neurones est un ensemble de neurones interconnectés (totalement ou localement) et organisés selon une structure (étoile, couches, ...). Un tel réseau est défini par sa topologie, les caractéristiques des neurones et des méthodes d'apprentissage. Classiquement on distingue [(Fogelman & Gallinari, 1991), (Davalo & Naïm, 1992)] :

- les réseaux à apprentissage supervisé, dans lesquels l'apprentissage se fait par l'exemple, le réseau s'adaptant sans faire référence à un mode de raisonnement explicite, à l'aide d'algorithmes qui minimisent une fonction d'erreur entre les valeurs attendues et les valeurs estimées à la sortie du réseau,
- les réseaux non supervisés, dont l'apprentissage n'est guidé que par la seule recherche de formes stables ou de profils.

La détermination de l'architecture du réseau (nombre de neurones, organisation en couches ou non, nombre de couches, nombre de neurones par couche, connectivité locale ou totale ...), de la forme de la fonction d'activation et du type d'algorithme utilisé pour l'apprentissage est du ressort de l'analyste. De tels choix conditionnent totalement les performances du réseau pour une tâche déterminée.

2.2. Réseaux de neurones et défaillances d'entreprises

Les réseaux à apprentissage supervisé de type multicouches à connectivité locale et orientée (figure 1) ont été les premiers à être expérimentés dans le domaine de la classification et du diagnostic à la suite des travaux de Rummelhart & Mc Clelland (1986) sur le Multilayer Perceptron (MLP) et de la généralisation de l'algorithme de rétropropagation du gradient.

Par ailleurs, les architectures de réseaux à apprentissage non supervisé, utilisant les propriétés d'auto-organisation de l'information, sont l'objet d'un développement important et prometteur dans le domaine de la reconnaissance de formes complexes (patterns). En particulier, un réseau de type ART 2 (Carpenter & Grossberg, 1988) a récemment été expérimenté, et obtenu des résultats intéressants, en matière de

reconnaissance de profils d'entreprises (Couturier & Fioleau, 1993), en particulier dans le cas où l'expertise ne permettait pas de fixer a priori le nombre de classes.

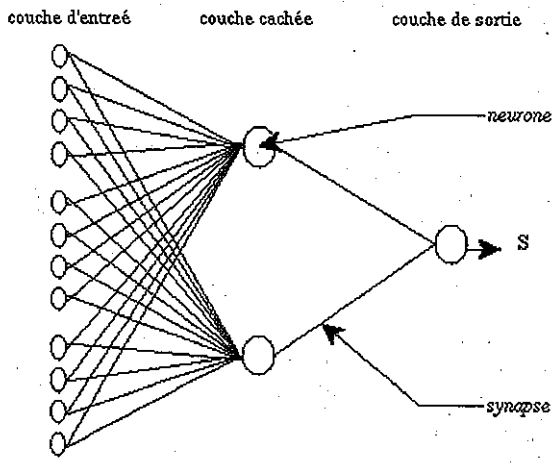

Figure 1 - Réseau Multicouches

Au contraire, faisant référence à une logique de classification dans laquelle le nombre de groupes est prédéterminé (par exemple : entreprises saines, entreprises en difficulté), les recherches traitant de l'évaluation du risque de défaillance ont généralement utilisé la technique bien stabilisée des réseaux de type MLP et, le plus souvent, retenu pour l'apprentissage l'algorithme de rétropropagation dans des versions plus ou moins sophistiquées. S'appuyant sur des résultats qui au plan théorique assurent des passerelles formelles avec les méthodes d'analyse de données, et en l'occurrence d'ADL (Gallinari et alii, 1991), ces recherches, menées depuis le début des années 90, ont eu pour objet d'explorer empiriquement les performances des réseaux neuronaux en fonction de différents paramètres (nombre de couches, nombre de neurones, type d'algorithme ...) en prenant pour référence les résultats obtenus par les méthodes devenues classiques dans le domaine de la prédiction du risque de défaillance.

De tels travaux cherchent à évaluer les capacités prédictives comparées des réseaux de neurones multicouches et de l'ADL [(Coats & Fant, 1992), (Varetto & Marco, 1993), (Baetge & Krause, 1992)] ou de la régression logistique [(Bell et alii, 1990), (De Almeida & Dumontier, 1993)], et à comparer les résultats obtenus par différents algorithmes d'apprentissage[7] [(Chung & Yan Tam, 1993), (Tam & Kiang, 1992), (Bodendorf & Maertens, 1992)]. Conformément aux résultats théoriques, ces études, bien que dissemblables et souvent préliminaires, montrent généralement une capacité prédictive très encourageante des réseaux de neurones du type MLP, même pour des configurations de réseaux assez simples.

[7] Rétropropagation du gradient (avec ses différentes variantes), Counterpropagation, Learning Vector Quantization (LVQ), Extended LVQ.

Cependant, bien que la démarche soit intéressante dans son principe, elle se limite le plus souvent, en l'état actuel, à l'élaboration de classifieurs neuronaux (linéaires ou non), suivie d'une comparaison de performances avec les méthodes classiques, ce qui constitue un bilan très modeste par rapport aux apports potentiels de l'approche connexionniste.

3 - Problématique de l'étude

Notre propos diffère de celui des recherches précédentes : il trouve son origine dans le décalage existant entre d'une part la technicité croissante des travaux empiriques relatifs à l'évaluation du risque de défaillance et d'autre part les problèmes de qualité ou d'incomplétude qui affectent les données disponibles (bilans, annexes, ...). Cette situation est paradoxalement mise en valeur par l'évolution récente de tels travaux, qui, pour la plupart, dans leur recours à l'approche neuronale n'ont recherché qu'un substitut aux méthodes de classification statistiques, sans exploiter les propriétés qui lui sont attribuées en matière de traitement des données bruitées ou imparfaites.

En effet, traditionnellement face à ce problème, on peut relever deux solutions utilisées dans les études empiriques : éliminer les entreprises dont les données sont incomplètes ou procéder à des traitements économétriques afin de reconstituer les séries[8]. La technique d'élimination, méthode très fréquente dans les recherches sur les défaillance, présente deux inconvénients majeurs : le premier, souvent souligné, a trait à la réduction de la taille des échantillons ; le second tient à la dissymétrie de la répartition des entreprises par rapport au critère d'incomplétude. De précédents travaux empiriques (Casta, 1985) nous conduisent à penser que la différence observable quant à la complétude des données entre les entreprises en difficulté et les autres entreprises constitue par elle-même un facteur de discrimination pertinent et donc à rechercher le moyen d'intégrer une telle information.

Dès lors, sans remettre en cause le mode de représentation de l'entreprise (les ratios financiers), prenant pour référence un type de classifieurs neuronaux courants (MLP avec rétropropagation du gradient), nous avons exploré empiriquement les possibilités de dégager une plus grande synergie entre classifieur et données par traitement de l'information relative aux données manquantes ou incomplètes.

Pour ce faire, nous avons suivi le protocole d'expérimentation suivant : sélection d'entreprises selon des critères de taille et d'activité, constitution de de bases pour l'apprentissage et la généralisation (avec répartition des entreprises par moitié de façon aléatoire), normalisation et codage des données, élaboration d'une première série de réseaux de neurones différant en fonction des données traitées (ratios de l'exercice ou des exercices antérieurs, données de nature comptable ou juridico-judiciaire, ...), mesure de performances en généralisation, modification de l'architecture du réseau afin d'utiliser toute l'information disponible (dont l'absence de données), mesure de performance en généralisation.

3.1. Les données

La base d'exemples concerne toutes les entreprises du secteur "Biens d'équipement" dont le chiffre d'affaires est supérieur à 10 millions de Francs. Elle est constituée de 1000 entreprises *jugées saines* et de 462 entreprises jugées en *situation difficile*,

---

[8] Sur ces techniques, voir : (Hachette & Mai, 1994).

classement [9] effectué à partir de l'expertise de la société SCRL pour l'année 1992. La répartition des entreprises en fonction du niveau de la cote d'expertise est présentée dans le Tableau 1.

	Nombre d'entreprises
cote = 1	38
cote = 2	324
Total défaillantes	**462**
cote = 3	1987
cote = 4	12
cote = 5	1
autres cotes	0
total saines	**1 000**

Tableau 1 - Répartition des entreprises de la base

Pour chaque entreprise de la base, nous avons extrait les 112 données suivantes[10] :

- 34 ratios classiques établis par la SCRL pour les exercices 1989, 1990 et 1991.
- le nombre et le montant des impayés, priviléges et contentieux pour les exercices 1990 et 1991,
- les cotes d'expertise SCRL pour les exercices 1989, 1990, 1991 et 1992,
- l'en-cours maximum de crédit pour 1992.

3.2. Classifieurs neuronaux sur données complètes

Dans une première étape, sur une base d'exemples restreintes aux entreprises disposant de données complètes, nous avons été amenés à construire à des fins exploratoires une série de six réseaux de type MLP, intégrant chacun des données de nature différente : R1 (utilisant les ratios financiers de 1991), R2 (ratios financiers de 1990), R3 (ratios financiers de 1989), R123 (ratios des trois années), RE (données extra-comptables), R123 (ratios de toutes les années) et R123E (utilisant tous les ratios et les données extra-comptables).

La technique utilisée se caractérise par l'architecture suivante : la taille de la couche d'entrée, variant selon le réseau, est fonction du vecteur de données ; la couche cachée est composée de 20 neurones[11], la couche de sortie de deux neurones (un pour chaque classe).

Si l'apprentissage a été mené avec le classique algorithme de rétropropagation, sa mise en oeuvre a fait l'objet d'un suivi (Figure 2) visant à éviter le risque de sur-spécialisation du réseau ("apprentissage par coeur") et à garantir le meilleur niveau de performance en phase de généralisation (Tableau 2).

---

[9] L'entreprise est considérée en situation difficile dès lors que sa cote est inférieure ou égale à 2.

[10] En théorie, car tous les champs ne sont pas nécessairement informés.

[11] Nous avons retenu cette solution après avoir expérimenté plusieurs configurations.

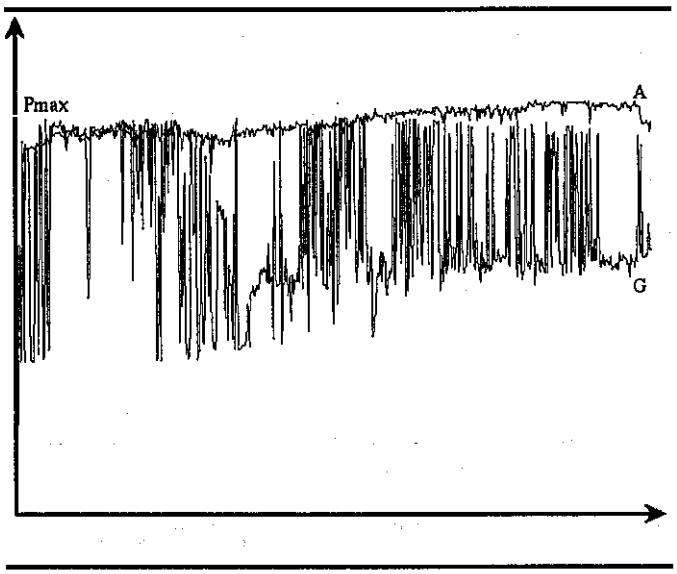

Figure 2 - Performances en phase d'apprentissage (A) et de généralisation (A)

	R123E	R123	R1	R2	R3	RE
défaillantes	84,4 %	85,7 %	84,4 %	81,8 %	60,2 %	38,1 %
saines	99,0 %	98,2 %	98,2 %	90,4 %	92,8 %	97,2 %
globales	94,4 %	94,2 %	93,8 %	87,7 %	82,5 %	78,5 %

Tableau 2 - Performances des différents réseaux

Par ailleurs, afin d'apprécier la différence qualitative des sorties du réseau par rapport à chacune des classes d'entreprises, nous avons été amenés à calculer un indice de certitude attaché à chaque réponse (Tableau 3).

		R123	R1	R2	R3	RE
mauvaises	90,6 %	87,4 %	59,3 %	82,0 %	90,7 %	40,7 %
bonnes	97,6 %	95,9 %	91,5 %	94,8 %	96,7 %	44,5 %
globales	97,3 %	95,4 %	89,5 %	93,2 %	95,6 %	43,7 %

Tableau 3 - Degré de certitude des réponses

3.3. Classifieurs neuronaux traitant l'incomplétude des données

Le modèle élaboré a pour objectif d'éviter l'élimination de la base d'entreprises dont les données sont incomplètes. Il doit permettre d'utiliser toute l'information disponible, même partielle, en la greffant sur un réseau de neurones existant, sans dégradation de ses performances antérieures.

Description du modèle

La structure de base du modèle est celle d'un réseau classique à 3 couches. La construction du réseau s'effectue en deux étapes distinctes. Dans un premier temps, un réseau à 3 couches, comportant toutes les entrées complètes (champs de la base toujours informés), est construit suivant l'algorithme de rétropropagation. La seconde étape consiste à greffer une nouvelle entrée sur ce réseau, entrée dédiée à l'adjonction d'éventuelles informations partielles. L'architecture du réseau est organisée selon le schéma ci-dessous (Figure 3). La partie du réseau qui correspond à la première étape est encadrée.

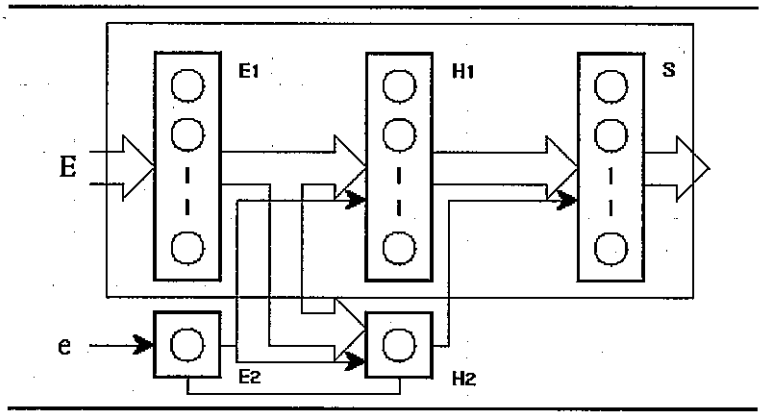

Figure 3 : architecture du réseau avec traitement de l'incomplétude des données

La couche d'entrée est composée de deux parties E1 et E2. La partie E1 correspond aux entrées du réseau pour lequel un apprentissage a déjà été réalisé à partir d'une base

d'exemples restreinte aux entreprises dont les données sont complètes. La partie E2 représente le neurone spécifique au traitement.
De la même façon, la couche cachée comprend deux blocs, H2, représente le ou les neurones ajoutés en couche cachée pour obtenir le meilleur gain. Les connexions entre E1 et H2, E2 et H1 sont créées de manière à obtenir un réseau entièrement connecté. La connexion particulière E2 ➔ H2 a pour objet d'inhiber H2 lorsque E2 l'est aussi. Son mode de fonctionnement apparaîtra lors de l'utilisation du réseau.
Enfin, la couche de sortie n'est pas modifiée. Seule une connexion de H2 ➔ S est ajoutée, de manière à recevoir l'apport de la partie du réseau dédiée au traitement de l'information incomplète.

La première partie de la phase d'apprentissage est menée avec une base d'exemples restreinte pour tenir compte des champs incomplets. Une fois le réseau optimal déterminé, il est procédé à l'adjonction de l'entrée la plus complète. Ensuite comme pour un réseau classique, il convient de déterminer le nombre de neurones dans la couche cachée[12]. Afin d'éviter la dégradation du niveau de performance obtenu sur le réseau initial, seules sont modifiées, durant l'apprentissage, les connexions liées à des neurones ajoutées (E2 ➔ H1, E1 ➔ H2 et E2 ➔ H2).

En généralisation, le réseau fonctionne en deux étapes. Il analyse dans un premier temps le vecteur d'entrée et détermine les entrées non informées. Il procède à leur inhibition, par l'intermédiaire de la liaison E2 ➔ H2, en neutralisant les neurones de la couche couchée associée aux entrées inactives, étape nécessaire dans la mesure où le réseau de base est optimisé dans ces conditions. Par ailleurs, l'inhibition de H2 permet d'intégrer l'information *absence de E2*. La dernière étape consiste, comme dans un réseau classique, à propager les entrées vers les sorties et à proposer la meilleure sortie possible quel que soit le vecteur d'entrée.

	R123E	R90	R89	R88	R87
nombre d'entreprises défaillantes	231	183	88	34	13
nombre d'entreprises saines	500	57	49	8	1
performances sur les entreprises défaillantes	86,1 %	86,1 %	87,4%	87,4%	87,4%
performances sur les entreprises saines	99,2 %	99,4 %	99,4 %	99,4 %	99,4 %
performances globales réelles	**98,3 %**	**98,5 %**	**98,6 %**	**98,6 %**	**98,6 %**
certitude	**96 %**	**96 %**	**96 %**	**95 %**	**96 %**

Tableau 4 - Performances des réseaux avec traitement de l'incomplétude des données

[12] Après divers essais, un seul neurone en couche cachée de la partie dédiée nous a permis d'obtenir les meilleurs résultats.

Le Tableau 4 présente les résultats obtenus en appliquant la procédure ci-dessus. Il illustre l'évolution des performances du réseau initial (R123E) au fur et à mesure que l'on adjoint les informations relatives aux entrées incomplètes des années 1990 (réseau R90), 1989 (réseau R89), 1988 (réseau R88). A contrario, l'évolution du nombre d'entreprises concernées par ces adjonctions successives est significatif des restrictions qui pesaient sur la base initiale.

Nous pouvons observer que le comportement de ce réseau intègre toute l'information disponible successivement présentée sans que l'on puisse noter l'apparition d'effet pervers. Cependant, si le gain de performance semble modeste, il bénéficie davantage aux entreprises en situation difficile qu'aux autres. Une telle remarque, qui semble généralisable en raison de la dissymétrie de l'incomplétude des données, est à replacer dans le contexte de la prise de décision financière (accords de crédit, ...). Dès lors, eu égard à la différence substantielle de coûts associés aux erreurs de classement des entreprises, très souvent soulignée dans la littérature relative au risque de défaillance (Eisenbeis, 1977), (Dumontier, 1991), (Fioleau, 1992), toute tentative visant à améliorer l'évaluation des entreprises en situation difficile peut sembler appréciable.

Conclusion

Au-delà de la présente étude, nous pensons que la récente introduction de l'approche connexionniste dans les recherches relatives à l'évaluation du risque de défaillance est de nature à en renouveler profondément les perspectives méthodologiques. En effet, comme nous avons essayé de le montrer, la technologie des réseaux neuronaux, actuellement très sous-utilisée, recèle d'importantes potentialités dans tous les domaines touchant à l'induction numérique. Elle est cependant victime, à juste titre, du syndrome de la "boite noire". Par rapport à ce problème, certains travaux récents touchant les défaillances se proposent "d'ouvrir la boite" de façon à faire émerger de l'induction numérique des règles à contenu plus symbolique (Nottola et alii., 1991).

BIBLIOGRAPHIE

ALTMAN (E. I.) : 1968, Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, **Journal of Finance**, Vol. XXIII, september, p. 589-609.

BAETGE (J.), KRAUSE (C.) : 1992, The classification of companies by means of neural networks, Working paper, Intitut für Revisionswesen, Münster, p. 1-38.

BEAVER (W.H.) : 1966, Financial ratios as predictors of failure, Empirical Research in Accounting : Selected Studies, **Journal of Accounting Research**, Supplement Vol 5, p. 71-111.

BARDOS (M.) : 1989, Ratios significatifs et détection du risque. Trois méthodes d'analyse discriminante, **Cahiers Economiques et Monétaires de la Banque de France**, n° 33.

BARDOS (M.) : 1984, Le risque de défaillance d'entreprise, **Cahiers Economiques et Monétaires de la Banque de France**, n° 19.

BARNES (P.) : 1987, The Analysis and use of financial ratios : a review article, **Journal of Business Finance and Accounting**, Vol. 14, n° 4, Winter, p. 449-461.

BARNIV (R). , RAVEH (A.) : 1989, Identifying financial Distress : a new non-parametric approach, **Journal of Business finance and Accounting**, Summer, p. 361-388.

BELL (T.), RIBAR (G.), VERCHIO (J.) : 1990, Neural nets vs. logistic regression : a comparison of each model's ability to predict commercial bank failure, Cash Flow Accounting Conference, Nice, décembre.

BODENDORF (F.), MERTENS (P.) : 1991, Früherkennung von Unternehmenskrisen, ein Vergleich von Neuronalen Netzen und Diskriminanzanalyse, Arbeits papier, n°4, Universität Erlangen-Nürnberg, p. 1-25.

CARPENTER (G.), GROSSBERG (S.) : 1987, ART 2 : Self-organization of stable category recognition codes for analog inputs patterns, **Applied optics**, Vol. 26, n°23, december, 4919-4930.

CASTA (J.F.) : 1985, Défaillances des entreprises, information financière et audit légal, Thèse, Université de Paris-Dauphine, CEREG.

CASTA (J.F.), ZERBIB (J.P.) : 1979, Prévoir la défaillance des entreprises ?, **Revue Française de Comptabilité**, n° 97, Octobre, p. 506-527.

CHUNG (H.), YAN TAM (K.) : 1993, A comparative analysis of inductive learning alogrithms, Intelligent Systems in Accounting, **Finance and Management**, janvier.

COATS (P.), FANT (F.) : 1992, Recognising financial distress patterns using a neural networks tool, **Financial Management**.

CONAN (J.), HOLDER (M.) : 1979, Variables explicatives de performance et contrôle de gestion dans les PMI, Thèse, Université de Paris-Dauphine, CEREG.

COUTURIER (A.), FIOLEAU (B.) : 1993, Performances des réseaux neuronaux comme outil de classification et de positionnement des entreprises au sein de leur secteur d'activité, Actes du XIVème Congrès de l'Association Française de Comptabilité, Vol. I, p. 355-364.

DAVALO (E.), NAIM (P.) : 1992, *Des réseaux de neurones*, Paris, Eyrolles.

De ALMEDIA (F.C.), DUMONTIER (P.) : 1993, Réseaux de neurones, information comptable et détection du risque de défaillance, Actes du XIVème Congrès de l'Association Française de Comptabilité, Vol. I, p. 269-286.

DEGOS (J.G.) : 1993, Les réseaux neuronaux supplanteront-ils les systèmes experts financiers ?, **Revue du Financier**, n° 90, p. 52-63.

DUMONTIER (P.) : 1991, Vices et vertus des modèles de prévision de défaillance, **Banque**, n° 514, mars, p. 275-282.

EISENBEIS (R.A.) : 1977, Pitfalls in the application of discriminant analysis in Business, Finance and Economics, **Journal of Finance**, Vol XXXII, June, 875-900.

EISENBEIS (R.A.), Avery (R.): 1972, *Discriminant Analysis and Classification Procedures*, Heath, Lexington Mass.

FIOLEAU (B.) : 1993, Histoire des modèles empiriques de prévision de défaillance d'entreprise : un essai de classification, Actes du XIVème Congrès de l'Association Française de Comptabilité, Vol II, p. 483-499.

FIOLEAU (B.) : 1992, Efficacité et risque d'exploitation. Application aux entreprises du secteur agricole. Thèse, Université de Nantes.

FOGELMAN (F.), GALLINARI (P.) : 1991, Neural Networks : from theory to industrial applications, Tutorial 7, Neuro-Nimes.

FRYDMAN (H.), ALTMAN (E.), DUEN-LI KAO (D.) : 1985, Introducing recursive partitioning for financial classification : the case of financial distress, **Journal of Finance**, Vol XL n°1, march, 269-291.

GALLINARI (P.), THIRIA (S.), BADRAN (F.), FOGELMAN-SOULIE (F.) : 1991, On the relations between discriminant analysis and Multilayer perceptrons, Neural Networks, Vol. 4, p. 349-360.

GHESQUIERE (S.), MICHA (B.), PACHOT (M.) : 1983, L'analyse des défaillances d'entreprises, Journée d'Etude des Centrales de Bilan, Juin.

HACHETTE (I), MAI (H.M.) : 1994, Le traitement des données manquantes pour la recherche en finance, in Recherches en Finance du CEREG, Hamon & Jacquillat (Ed), Paris, Economica.

JOY (O.M.), TOFFELSON (J.O.) : 1975, On the financial application of discriminant analysis, **Journal of financial and Quantitative Analysis**, Vol X, n° 15, december, 723-739.

LE MOIGNE (J.L.) : 1991, *Sur les fondements épistémologiques de la science de la cognition*, in Andreewsky (Ed.), Systémique et Cognition, Dunod, p. 11-50.

MALECOT (J.F.) : 1986, Sait-on vraiment prévoir les défaillances d'entreprises ? **Economie et Sociétés**, Série SG9, Tome XX, n°12, Décembre.

MALECOT (J.F.) : 1991, Analyses théoriques des défaillances d'entreprises : une revue de la littérature. Cahier de recherche n° 9108, CEREG-Université de Paris-Dauphine.

NOTTOLA (C.), CONDAMIN (L.), NAIM (P.) : 1991, Réseaux de neurones durs pour l'extraction de règles : une approche méthodologique appliquée à l'évaluation d'entreprises, Neuro-Nimes, p. 327-340.

OHLSON (J.) : 1980, Financial ratios and the probabilistic prediction of bankruptcy, **Journal of Accounting Research**, Vol 18, n° 1, Spring, p. 109-131.

RUMMELHART (D.E.), Mc CLELLAND (J.L.) : 1986 : *Parallel Distribued Processing*, Cambridge (M.A.), The MIT Press.

TAM (K.T.), KIANG (M.Y.) : 1992, Managerial applications of neural Networks : the case of Bank failure predictions, **Management Science**, Vol. 38, p. 926-947.

VARETTO (F.), MARCO (G.) : 1993, Bankruptcy diagnosis and neural networks. Experiments and comparisons using linear discriminant analysis. European Financial Statement Data Bases, Brixen, September.

WATSON (C.J.) : 1990, Multivariate distributional properties, outliers and transformation of financial ratios, **The Accounting Review**, Vol 65, n°3, July, 682-695.

ZAVGREN (C.) : 1983, The prevision of corporate failure : the state of the art, **Journal of Accounting Litterature**, Vol. 2.

ZMIJEWSKI (M.E.) : 1984, Methodological issues related to the estimation of financial distress prediction problems, **Journal of Accounting Research**, Vol XXII, Supplement, 59-82.