

Comment décider de l'attribution de crédits en utilisant les nouvelles techniques ? (Méthode scoring et réseaux de neurones)

Mme SADI Khadidja

INTRODUCTION

Décider ou intervenir dans un processus de décision c'est trouver une solution à un problème. L'entreprise est confrontée quotidiennement à des situations et des enjeux complexes qui oblige le décideur à rechercher une aide à la décision allant au-delà du bon sens, de l'expérience ou de la mise en œuvre de techniques de calcul élémentaire. La majorité des problèmes industriels, financiers ou économiques vécus par l'entreprise sont du ressort de la décision multicritères. A titre d'exemple mentionnons : les problèmes de gestion financière, l'évaluation des projets d'investissement, la constitution d'un portefeuille de titres, et l'octroi de crédits bancaires.

De nos jours, plusieurs méthodes d'aide multicritères à la décision sont proposés telles que : ELECTRE et PROMETHEE et leurs variantes. L'objectif de ces méthodes est d'amener le décideur à réduire la part de l'arbitraire et de trouver le meilleur compromis.

Le but de cette communication est de présenter, aux économistes, un nouvel outil d'aide multicritère à la décision, qui consiste en l'application des réseaux artificiels.

Afin de montrer la puissance et l'apport de la technique neuronale, on procèdera en fait à une étude comparative entre les réseaux de neurones et la méthode du scoring. Cette dernière est une technique classique d'attribution de crédits et la utilisée au niveau des banques.

Par contre, la technique neuronale relève du domaine de l'I.A(intelligence artificielle) qui commence à se frayer un chemin dans la prise de décision en milieu financier.

Dans un but illustratif, nous présenterons notre travail sur la base d'une étude comparative qui mettra en évidence la robustesse du modèle neuronal.

De longue date, l'octroi de crédit était basé sur l'analyse financière par ratios, et presque dans tous les cas elle reposait sur un indicateur unique (décision uni critère) et ne permettait pas de faire de réelles prévisions. Mais très vite, cette voie fût délaissée pour laisser place aux indicateurs synthétiques financiers (construits sur la base de plusieurs ratios considérés simultanément). Ainsi l'analyse multicritères voie le jour, se basant sur une science : '*les mathématiques*'.

Avec le développement de l'informatique, plusieurs méthodes se sont développées, mettant ainsi à la portée du décideur tous les moyens de détection de défaillance des entreprises sollicitant un crédit.

D'une manière générale, l'octroi de crédit reste et sera un problème de classement. Autrement dit, il convient d'analyser la santé financière, et calculer leur risque de défaillance. Ainsi, les banques se trouvent confrontées à un problème majeur qui est celui de la maîtrise du risque, permettant d'assurer leur rentabilité, et par conséquent faire face à la concurrence.

La présente étude se compose de trois parties :

- la première partie est une présentation de la méthode score avec ses avantages et ses limites.
- la deuxième partie est une présentation de la méthode des réseaux de neurones artificiels (R.N).
- la troisième partie consistera en l'application des deux techniques sur un exemple d'illustration. Ainsi, la robustesse du modèle neuronal sera mise en évidence par la comparaison des résultats obtenus par les deux techniques.

I. PRESENTATION DE LA METHODE SCORING

La construction de la fonction score repose sur une méthode statistique: l'analyse factorielle discriminante. Elle permet d'établir un indicateur de classement simple. Comme il n'y a que deux classes, la fonction score se présente ainsi:
$$Z = \sum_{i=1}^n \alpha_i * .R_i + c$$

Où R_i sont des ratios financiers, α_i les coefficients numériques d'ajustement, et c une constante.

" C'est en fait le signe de la fonction Z qui permet de d'affecter l'entreprise à l'une de des deux classes:

- dans la zone de (Z) négatif se trouve le maximum d'entreprises défailtantes sachant qu'il en a toujours quelques unes qui se trouvent mal classées dans la zone de score positif.
- dans la zone de (Z) positif se trouve à l'inverse le maximum d'entreprises saines bien classées à partir de leur score.

La pertinence de la fonction (Z) s'apprécie par rapport au nombre relatif d'entreprises bien classées à partir de leur score. On utilise pour cela une table d'affectation qui fait apparaître le nombre de cas où il y a identité entre le groupe d'origine et le groupe d'affectation déterminé la fonction score (Z) par rapport par rapport au nombre total d'entreprises.

Pratiquement la fonction (Z) est construite en sélectionnant progressivement les ratios financiers les plus discriminants. En règle générale, l'on procède de la façon suivante:

- Observation de la normalité de distribution des ratios financiers de départ (il importe que ceux-ci soient en nombre suffisamment important pour éviter une très grande subjectivité et pour obtenir le plus vaste éventail d'interprétation) et élimination de ceux ne correspondant pas au test correspondant.
- Examen du pouvoir de discrimination individuel des ratios retenus et classement de ceux-ci par catégorie.
- Evaluation des corrélations existant entre ses ratios et constitution de typologies: élimination des ratios dont le pouvoir discriminant est le plus faible.
- Observation de la précision de discrimination des différentes combinaisons et sélection par itération de la combinaison discriminante qui donne le meilleur classement (fonction Z).

Par rapport à l'examen isolé ou successif d'une batterie de ratios financiers, un indicateur tel que la fonction (Z) privilégie une approche multidimensionnelle car:

- elle utilise une sélection de variables qui ont conjointement un pouvoir discriminant plus élevé;
- chaque ratio de la fonction est affecté d'un coefficient pondérateur rendant compte de l'importance relative des variables retenues"¹

¹ Bernard GUILHOT. LA REVUE BANQUE N°466-NOV 1986. P976.

Avantages et limites de la fonction score :

- ***Avantages :***

La construction de la fonction score se fait L'analyse factorielle discriminante (A.F.D) procède à une pondération des ratios qui entrent dans les calculs.

Elle accorde à chaque ratio une pondération qui convient à son pouvoir explicatif dans la séparation des groupes (défaillant et sain). Elle fait aussi disparaître l'effet taille entre les ratios dont les ordres de grandeurs sont fondamentalement différents.

- ***Inconvénients ou limites :***

N'intégrant que les ratios financiers (variables de même nature), la fonction score ne peut prendre en considération certaines informations qualitatives. De plus, les ratios doivent suivre une loi multi normale au sein de chaque groupe (sain, défaillant), et ne permettent pas toujours une séparation linéaire des deux groupes. Enfin la qualité prédictive du score n'est jamais parfaite à 100%.

II. PRESENTATION DES RESEAUX DE NEURONES ARTIFICIELS

Conçus à l'origine comme des modèles mathématiques très simplifiés de notre cerveau, les réseaux de neurones formels ou artificiels (R.N) sont des systèmes d'apprentissage par exemple, imitant en cela la façon dont nous apprenons par essais, erreurs et corrections successives à effectuer des tâches complexes. Concrètement, ces techniques neuronales sont des outils mathématiques et algorithmiques qui s'avèrent être puissants et commodes pour résoudre des problèmes non linéaires et complexes. Ils nourrissent des contributions provenant des mathématiques (optimisation, statistiques, systèmes non linéaires, théorie de la décision), de la physique statistique, de l'informatique (intelligence artificielle, reconnaissance des formes). En modélisation, il existe une interaction importante entre créateurs de modèles et chercheurs en neurosciences et science cognitives.

Les modèles, ainsi développés sont permis de proposer de nouvelles méthodes de traitement des données, de nouveaux systèmes de traitement de l'information. Au fur et à mesure du développement des moyens de calcul et de l'augmentation de la puissance des ordinateurs, les réseaux de neurones ont marqué les années 90 en proposant de nouveaux paradigmes pour le calcul, de nouvelles méthodes de modélisation, et en réintroduisant les méthodes numériques au sein de l'I.A. le domaine s'est développé extrêmement vite en proposant des méthodes pratiques pour résoudre des problèmes de grande taille et des applications ont été réalisées dans la plupart des domaines traditionnels de l'ingénierie (diagnostic, contrôle, reconnaissance de caractères et d'images, prévision de série temporelles, discrimination, etc.)¹.

Les réseaux de neurones artificiels sont donc des modèles mathématiques, se présentant comme un ensemble d'unités de calcul fortement connectés appelé neurone formel ou artificiel. A ce titre ils peuvent être décrit par leur architecture et leurs composants.

Les composants du réseau de neurones :

1/ Structure du neurone formel ou artificiel

La figure 1 montre la structure d'un neurone artificiel. Chaque neurone artificiel est un processeur élémentaire. il reçoit un nombre variable d'entrées en provenance de neurones amonts. A chacune de ces entrées est associé un poids w_i représentatif de la force de la connexion. Chaque processeur élémentaire est doté d'une sortie unique, qui se ramifie ensuite pour alimenter un nombre variable de neurones avals. A chaque connexion est associée un poids.

¹ Pour plus de détails, voir les ouvrages spécialisés.

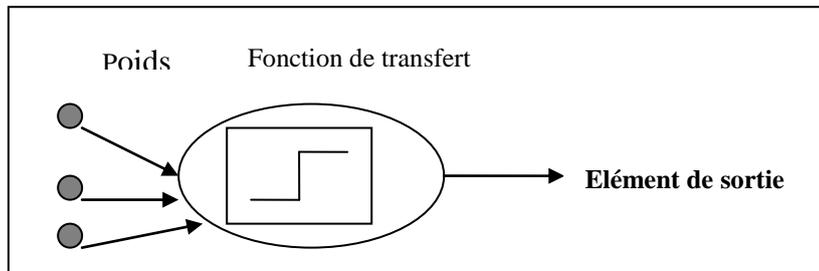


Figure 1. Neurone artificiel

2/ Le fonctionnement du neurone formel

On distingue deux phases. La première est habituellement le calcul de la somme pondérée des entrées (a) selon l'expression suivante: $a = \sum (w_i \cdot e_i)$

A partir de cette valeur, une fonction de transfert (fonction d'activation) calcule l'état la valeur du neurone. C'est cette valeur qui sera transmise aux neurones avals. Il existe de nombreuses formes possibles pour la fonction de transfert. La plupart de ces fonctions sont continues, offrant une infinités de valeurs possibles comprises dans l'intervalle $[0,+1]$ (ou $[0,+1]$).

3/ Architecture du modèle

C'est la structure des connexions entre les neurones qui composent le réseau, elle décrit la topologie du modèle. Elle peut être quelconque, mais le plus souvent il est possible de distinguer une certaine régularité.

On distingue le réseau multicouche (c'est le modèle utilisé dans notre cas); les neurones sont arrangés par couches. Il n'y a pas de connexions entre neurones d'une même couche (figure 2). Il existe d'autres types de réseaux tels que

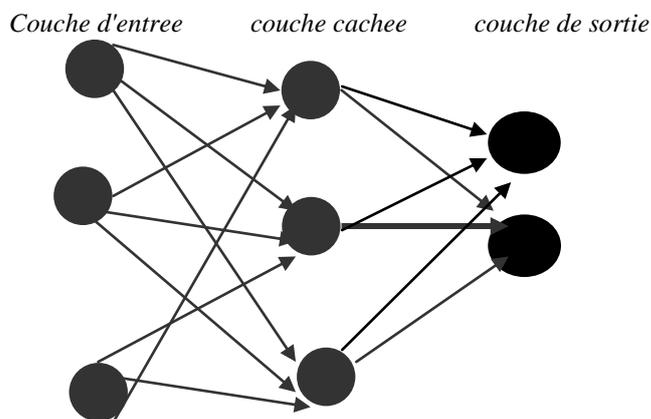


Figure 2. Réseau multicouche.

4/ Apprentissage du réseau

L'apprentissage est la propriété la plus intéressante des réseaux neuronaux. Elle ne concerne cependant pas tous les modèles, mais les plus utilisés. ***L'apprentissage est une phase du développement d'un réseau de neurones durant laquelle le comportement du réseau est modifié jusqu'à obtention du comportement désiré. L'apprentissage neuronal fait appel à des exemples¹³.***

Dans le cas des réseaux de neurones artificiels, on ajoute souvent à la description du modèle l'algorithme d'apprentissage.

L'apprentissage est la modification des poids du réseau dans l'optique d'accorder la réponse du réseau aux exemples et à l'expérience.

Au niveau des algorithmes d'apprentissage, il a été défini deux grandes classes selon que l'apprentissage est dit supervisé ou non supervisé. Cette distinction repose sur la forme des exemples d'apprentissage. Dans le cas de l'apprentissage supervisé, les exemples sont des couples (entrée, sortie associée) alors que l'on ne dispose que des valeurs (entrée) pour l'apprentissage non supervisé.

III. APPLICATION DES DEUX TECHNIQUES ET COMPARAISON DES RESULTATS

1. Information de base

L'échantillon de base est constitué de données réelles figurant sur des dossiers de clients d'une banque algérienne. Il comporte quatre vingt dix (90) entreprises de type P.M.E provenant de secteurs d'activité variés. Ces entreprises ont bénéficié d'un crédit d'exploitation, et décrites à l'aide de quatorze (14) ratios économiques et financiers déduit d'un traitement comptable, et de trois (03) variables spécifiques. Au total nous disposant de dix sept (17) variables explicatives. Les entreprises sont ainsi réparties en deux (02) groupes que l'on cherche à discriminer : le groupe **DI** des entreprises saines ou solvables et le groupe **D0** des entreprises défailtantes.

Logiciels utilisés :

- Le logiciel statistique SPSS version 8.0 pour la méthode du scoring.
- Le logiciel mathématique MATLAB version 4.1 pour la méthode neuronale.

2. Application de la méthode du score :

L'analyse discriminante a conduit à une réduction du nombre de variables sans pour autant perdre de l'information. Ainsi les 10 premiers facteurs véhiculent près de 90% de l'information initiale et présentent l'avantage d'être corrélés deux à deux.

Les facteurs issus de l'analyse factorielle ont été utilisés conjointement dans la construction linéaire de Fisher, appelée couramment **fonction score** :

- Si la fonction prend une valeur positive pour une entreprise donnée, celle-ci est considérée saine.
- Si la fonction prend une valeur négative pour une entreprise donnée, celle-ci est considérée défailtante.

La fonction score permet ainsi de classer les entreprises de la plus saine à la plus défailtante. Une fois le score calculé pour chaque entreprise, la probabilité à posteriori est déterminée permettant au décideur de juger de la qualité de prévision.

Résultats de l'analyse discriminante

a) Classement de l'échantillon de construction

Entreprises	Bien classées	Mal classées
Groupe D0	75 %	25 %
Groupe D1	91.7 %	8.3 %

b) Classement de l'échantillon de validation

<i>Entreprises</i>	Bien classées	Mal classées
Groupe D0	65 %	35 %
Groupe D1	80 %	20 %

3. Application de réseaux de neurones :

Nous avons utilisé un réseau multicouche (figure 3) composé des éléments suivants:

- Une couche d'entrée constituée de 17 neurones correspondants aux variables explicatives.
- Une couche de sortie constituée d'un seul neurone (neurone de décision) dont la valeur représente le groupe d'appartenance : si la valeur est zéro l'entreprise appartient au groupe D0, et si la valeur est 1 alors l'entreprise appartient au groupe D1.
- Une couche cachée constituée de 15 neurones. Ce nombre a été atteint progressivement après plusieurs tests sur l'échantillon d'apprentissage jusqu'à atteindre l'optimum d'apprentissage et de généralisation.

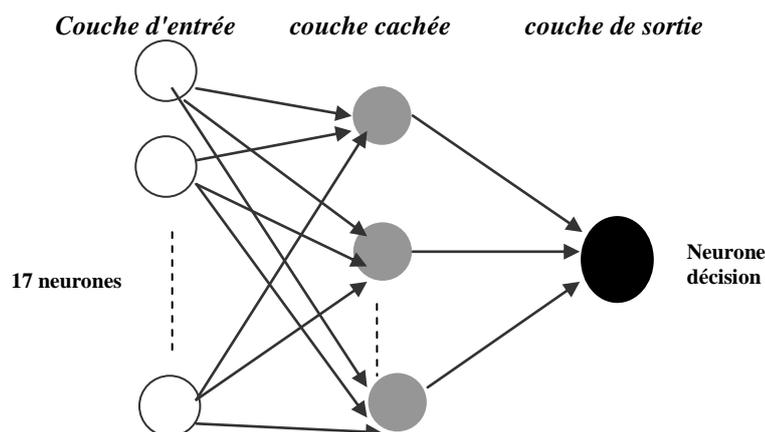


Figure 3. Modèle du réseau utilisé.

Deux fonctions d'activation ont été utilisées :

- La fonction logistique pour la couche de sortie.
- La fonction tangente hyperbolique pour la couche cachée.

L'apprentissage d'un tel réseau est supervisé et utilise l'algorithme de rétro-propagation du gradient de l'erreur.

Résultats de l'analyse discriminante

a) Classement de l'échantillon de construction

L'échantillon d'apprentissage ou de construction du modèle neuronal est présenté plusieurs fois au réseau de manière itérative jusqu'à convergence de l'erreur vers un minimum fixé au préalable.

Les résultats obtenus sont comme suit :

<i>Entreprises</i>	Bien classées	Mal classées
Groupe D0	97 %	3 %
Groupe D1	100 %	0 %

*Le nombre d'itérations est de 668.
Le taux d'apprentissage est de 84.3 %.
La somme quadratique des erreurs est de 0.0083.*

b) Classement de l'échantillon de validation ou test :

Dans cette phase on présente au réseau des exemples non connus et on teste sa capacité de génération et donc sa robustesse de prévision.

<i>Entreprises</i>	<i>Bien classées</i>	<i>Mal classées</i>
Groupe D0	85 %	15 %
Groupe D1	93.7 %	6.6 %

4. Comparaison des résultats :

Contrairement aux réseaux de neurones qui permettent un taux de bon classement de près de 90% sur l'échantillon de construction, l'analyse discriminante linéaire n'en est qu'à 75%. Ceci est dû au fait que tous les ratios utilisés dans l'analyse discriminante ne permettent pas une discrimination linéaire.

Aussi le modèle de R.N est capable d'identifier les exemples qui ne lui en été jamais présentés (phase test). Ainsi la performance de cet outil est toujours supérieure ou égale à celle obtenue par des techniques classiques.

L'apport supplémentaire de l'analyse discriminante par rapport aux R.N est de fournir des probabilités d'appartenance à chacun des 2 groupes connaissant du score.

CONCLUSION

Les R.N présentent un avantage sur la fonction score, car opérant une séparation non linéaire des deux groupes, prennent ainsi la non linéarité des relations entre les variables. En effet, comparé au score, c'est un outil qui donne des résultats sensiblement meilleurs, et donc permet de mieux prévoir le comportement du demandeur de crédit.

Aussi performant qu'il parait, il présente cependant un inconvénient car il fonctionne comme une **boîte noire** et par conséquent, il est impossible de contrôler la sélection effectuée.

Enfin ceci ne constitue qu'un essai de construction d'un outil de détection précoce de défaillance des entreprises qui associe la statistique et l'outil informatique.

Références

1. C.Touzet, réseaux de neurones artificiels introduction au connexionnisme, juillet 1992.
2. J.F.Joudouin, les R.N principes et définitions, Ed Hermes 1994.
3. E.Duvalo et P.Naim, des réseaux de neurones , Ed Eyrolles 1989.
4. J.S.Taur et S.Y.Kung, prediction-based networkwith E.C.G application, IEEE1993.
5. B.Guilhot, méthode des scores intérêts et limites (1 partie), revue banque N°466 nov.1986.
6. B.Guilhot, méthode des scores intérêts et limites (2 partie), revue banque N°468 jan.1987.
7. P.Saint-Server, les méthodes scores sont elles efficaces ? revue banque, N°475 sept. 1987.
8. P.Damel, score d'entreprise: mieux apprécier le risque de défaillance, revue banque, N°575 nov. 1996.
9. P.Canarelli et H.Ivanoff, intelligence artificielle et scoring bancaire, revue banque, N°577 jan. 1997.

