



# La prévision de défaillance des PME : une application neuronale sur données qualitatives

Muriel Perez

## ► To cite this version:

Muriel Perez. La prévision de défaillance des PME : une application neuronale sur données qualitatives. INTERNATIONAL MEETING OF FRENCH FINANCE AFFI, Jun 2003, LYON, France. <halshs-00522136>

**HAL Id: halshs-00522136**

**<https://halshs.archives-ouvertes.fr/halshs-00522136>**

Submitted on 29 Sep 2010

**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.

# **La prévision de défaillance des PME : une application neuronale sur données qualitatives**

**Muriel PEREZ**

MODEME UMR 5055  
Université Jean Moulin Lyon 3  
15, quai Claude Bernard 69007 LYON  
tél: 04 78 78 71 58 fax: 04 78 78 75 50  
Muriel.Perez@univ-lyon3.fr

---

**RESUME** : Les applications neuronales au domaine de la prévision de défaillance remontent au début des années 90. La plupart reposent sur l'utilisation d'un type particulier de réseaux de neurones artificiels, le Perceptron Multicouches. Cependant une autre catégorie de réseaux existe, capable de traiter des problèmes de classification. Nous avons donc construit une nouvelle expérimentation reposant sur les cartes auto-organisatrices de Kohonen. L'objectif premier de notre travail est d'étudier l'apport des données qualitatives dans le cadre de l'utilisation de cet outil au domaine de la prévision de défaillance des PME.

**MOTS CLES** : Prévision de défaillance, Réseaux de neurones artificiels, Cartes auto-organisatrices de Kohonen, classification.

---

**ABSTRACT** : The neuronal applications to the topic of bankruptcy forecasting appeared in the beginning of the 90's. The multilayer Perceptron is the most famous in this kind of application. Nevertheless, an other neural network can resolve classification problems : the self-organizing map of Kohonen. We have built a new application to bankruptcy forecasting with this kind of neural networks, including qualitative data. The first aim of this work, is to study the interest of qualitative data with Kohonen map in this field.

**KEY WORDS** : Bankruptcy forecasting, Artificial Neural Networks, Self-Organizing Map of Kohon, Classification

---

# **La prévision de défaillance des PME : une application neuronale sur données qualitatives**

Cette recherche s'inscrit dans la continuation d'un travail de thèse (PEREZ, 2002), portant sur les apports de la classification neuronale à la prévision de défaillance des PME où seules les données de type quantitatif avaient été mises en place. Dans le cadre d'une nouvelle expérimentation, nous utilisons des données qualitatives accessibles au tiers en grand nombre, disponibles sur les bases de données financières et comptables, afin de mesurer leurs apports. Nous mettons ainsi en place une batterie de tests basés sur les cartes auto-organisatrices de Kohonen afin de mener une classification sur des échantillons d'entreprises saines et défaillantes.

## **1. Mise en place de l'expérimentation**

La mise en place de notre expérimentation repose sur l'utilisation de deux outils principaux : la base de données Diane de la SCRL permettant de disposer, entre autres, des informations comptables des entreprises et le logiciel d'analyse de données Intelligent Miner d'IBM (1999) disposant d'une option de classification neuronale par cartes auto-organisatrices de Kohonen. Si les réseaux de neurones ont fait l'objet d'un grand nombre d'applications en finance (De BODT, COTTRELL et LEVASSEUR, 1995) à travers le Perceptron Multicouches, les cartes auto-organisatrices de Kohonen<sup>1</sup> restent encore peu utilisées. Leur capacité de classification les rend cependant particulièrement adaptées au problème de la prévision de défaillance.

### **1.1. Le choix des échantillons**

La construction des échantillons doit répondre à deux critères. Tout d'abord, afin de favoriser l'apprentissage de la technologie neuronale mise en place, l'échantillon ne doit pas être de taille trop restreinte. Malheureusement, il n'existe pas à ce jour, de théorie sur la taille d'un échantillon pour un apprentissage optimal. Une application neuronale relève donc d'une démarche itérative. L'échantillon doit également respecter une certaine homogénéité afin de limiter le nombre d'exemples nécessaires à la phase d'apprentissage. En effet, plus l'échantillon présente des entreprises aux profils hétérogènes, plus il devra être grand. Ces deux conditions ont ainsi guidé les différents choix réalisés dans le cadre de notre application, concernant notamment la notion de défaillance, le secteur d'activité, la taille des entreprises et le type de données d'entrée.

#### **a) La notion de défaillance**

La notion de défaillance retenue est celle de la disparition juridique de l'entreprise du parc national. Cette situation se manifeste à travers trois états dans la base de données Diane : la liquidation amiable, la liquidation judiciaire et le redressement judiciaire. Si la liquidation et le redressement représentent des étapes différentes dans le processus de disparitions des entreprises, dans le cadre des PME cette différence est moins marquée puisque la majorité des

---

<sup>1</sup> Sur les capacités des cartes auto-organisatrices de Kohonen et leur intérêt dans le cadre de la gestion le lecteur peut se référer à COTTRELL, FORT et PAGES (1998) et COTTRELL et GAUBERT (1999).

redressements judiciaires se solde en définitive par une liquidation. Les situations de cessation d'activité ou de cessation des paiements n'ont donc pas été retenues.

#### **b) Le secteur d'activité :**

En prenant en compte les deux critères de taille d'échantillon suffisante et d'homogénéité, deux secteurs d'activité ont été retenus. Le premier que nous nommerons secteur « Industrie » est construit artificiellement à partir des codes de la nomenclature NAF de 15 à 36. Le second, secteur « Travail des métaux » correspond à l'activité « travail des métaux » recensé sous le code NAF 28. La population du premier échantillon montre une population nettement supérieure en terme de taille d'échantillon mais également beaucoup plus hétérogène, en terme de taux de défaillance par sous-secteur d'activité.

#### **c) La taille des entreprises**

Afin de limiter l'hétérogénéité des échantillons nous ne retenons que des entreprises de type PME selon le critère de l'INSEE, à savoir les entreprises présentant un effectif compris entre 10 et 499 salariés. Néanmoins, cette notion contient encore intrinsèquement une forte hétérogénéité. En effet, à l'intérieur du concept même de PME cohabite déjà deux types d'entreprises aux caractéristiques distinctes à savoir les Petites Entreprises et les Moyennes Entreprises.

#### **d) Les données d'entrée**

Dans le cadre de nos travaux précédents (PEREZ, 2002) nous n'avions retenu que des données de type quantitatif. Nous reprenons ici les mêmes données en élargissant à quelques données de type qualitatif disponibles directement dans les bases de données financières et comptables. Ainsi les données quantitatives sont les ratios issus directement de la batterie du volet analyse financière de Diane. Elles sont au nombre de trente-cinq, classées en quatre groupes « Structure et Liquidité », « Gestion », « Productivité et Rentabilité » et « Marge et Valeur Ajoutée ».

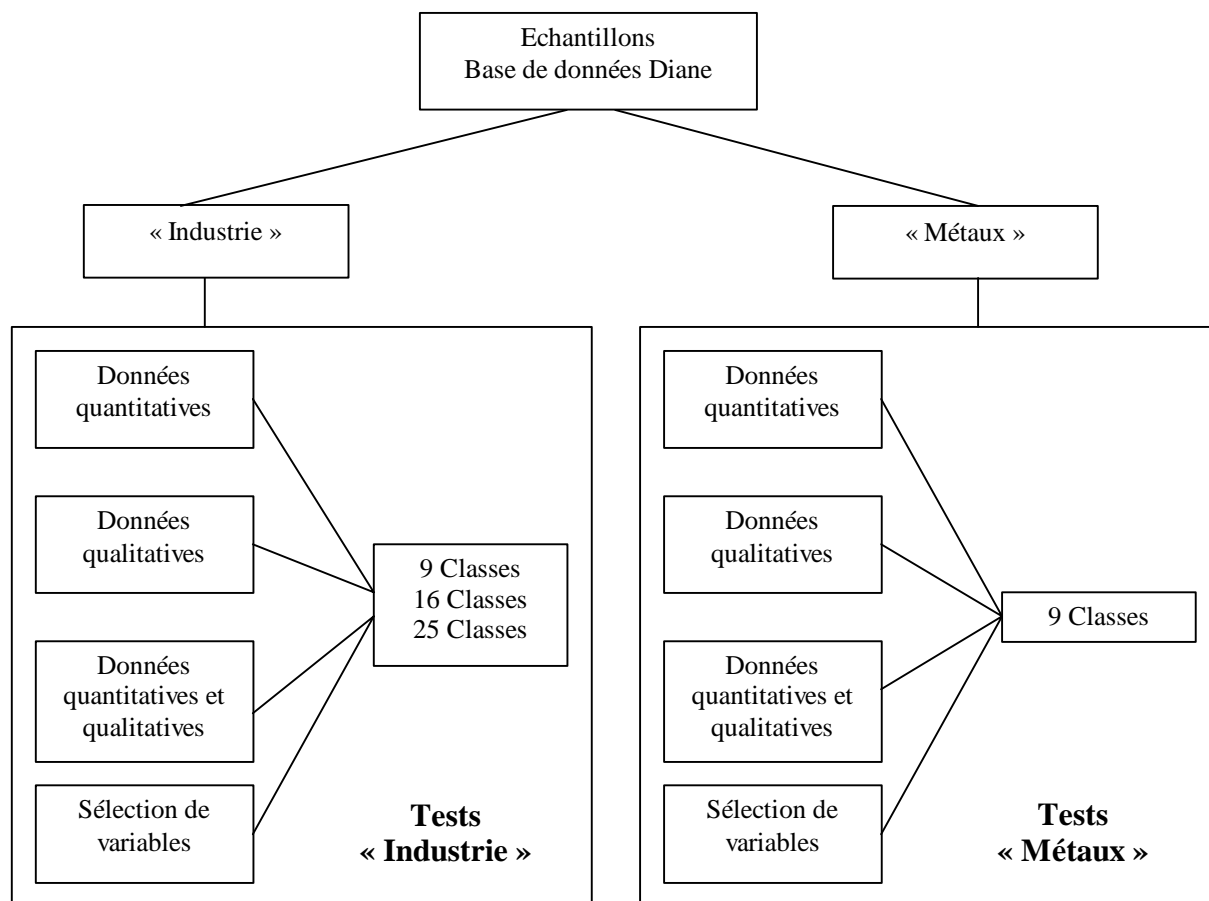
Les données qualitatives disponibles sur Diane et retenues sont le secteur d'activité (à travers le code NAF), la taille de l'entreprise (à travers l'effectif), la présence ou non d'actionnaires, l'existence ou non de filiales ainsi que la date de création. Il convient de préciser que le nombre de ces variables qualitatives est limité de par la contrainte de disponibilité sur support informatique. Il s'agit ici de mettre en place une application afin de tester l'intérêt de l'utilisation de ce type de variable.

Afin de limiter l'impact des entreprises nouvellement créées nous avons éliminé de notre échantillon les entreprises ayant moins de trois années d'existence. De plus, la date de défaillance porte sur 1996 et les données d'entrée sont celles de 1994, ainsi les données précédant directement la défaillance ne sont pas utilisées. Dans les faits ces données sont, en effet, rarement disponibles.

Au final le fichier d'entreprises du secteur « Travail des métaux » est constitué de 128 entreprises, soit 64 saines et 64 défaillantes. Le secteur « Industrie » quant à lui regroupe 750 entreprises, dont 375 saines et 375 défaillantes. Dans le cadre de notre expérimentation, l'échantillon est automatiquement limité par le nombre de défaillance d'un secteur.

## 1.2. Le protocole

Les données ainsi retenues sont extraites du cédérom Diane sous la forme de fichiers Excel. Les entreprises montrant un grand nombre d'informations manquantes sont ensuite éliminées de l'échantillon et les données qualitatives sont codées. Les fichiers sont ensuite exportés vers Statistica de Statsof (STATISTICA, 1998), (module analyse de données) afin d'être étudiés. Les échantillons sont enfin formatés à travers DB2 d'IBM (IBM, 1999), (gestionnaire de données) afin de pouvoir être traités par Intelligent Miner (IBM, 1999).



*Fig.1 – Les phases de l'expérimentation neuronale*

Etant donné l'étroitesse de l'échantillon du secteur « Travail des métaux », les tests ne sont menés que sur des cartes comportant neuf classes. L'échantillon du secteur « Industrie » étant plus conséquent, nous avons réalisé des tests avec neuf, seize et vingt-cinq classes. L'objectif premier de ce protocole expérimental est d'étudier les données les plus pertinentes dans le cadre d'une application de type neuronale à la prévision de défaillance de PME. Nous réalisons donc des apprentissages à partir uniquement de données quantitatives puis uniquement de données qualitatives, avant de mêler les deux types de données dans une nouvelle batterie. Enfin, nous réalisons un nouvel apprentissage combinant données qualitatives et quantitatives après sélection des variables. Le nombre d'indicateurs est alors considérablement réduit passant de près de quarante à seulement six après sélection. Les ratios retenus sont différents selon le secteur d'activité.

## 2. Les résultats

Nous présentons dans un premier temps les résultats sur données quantitatives avant d'aborder les résultats basés sur la combinaison des données quantitatives et qualitatives. Une distinction est également opérée selon les secteurs d'activité : « Travail des métaux » ou « Industrie ».

### 2.1. Résultats sur données quantitatives

#### a) Industrie

Suite à trois apprentissages en 9, 16 et 25 classes, la classification permettant d'obtenir les meilleurs taux de distinction entre entreprises saines et défaillantes est celle en 16 classes. Nous reprenons ci-après les résultats de la carte obtenue.

0	1	2	3
S 34%	S 23%	S 40%	S 51%
D 66%	D 77%	D 60%	D 49%
5.46	5.86	6.26	6.26
4	5	6	7
S 24%	S 30%	S 41%	S 66%
D 76%	D 70%	D 59%	D 34%
5.59	4.39	7.46	5.06
8	9	10	11
S 27%	S 47%	S 69%	S 84%
D 73%	D 53%	D 31%	D 16%
6.92	7.99	4.66	5.06
12	13	14	15
S 31%	S 61%	S 85%	S 88%
D 69%	D 39%	D 15%	D 12%
7.32	7.86	6.92	6.92

*Fig.2 – La carte Industrie 16 classes sur données quantitatives*

Les ratios privilégiés dans le cadre de la classification en 16 classes sont R17 en tête (couverture du chiffre d'affaires par le fonds de roulement), suivi des ratios de liquidités R07 et R08, puis viennent R10, R34, R06, R04 et R18. Une frontière nette apparaît entre les groupes d'entreprises saines au nombre de sept et les groupes d'entreprises défaillantes au nombre de neuf. Les classes montrant le plus de distinction entre les entreprises saines se situent loin de cette frontière, dans les angles de la carte.

La partie supérieure gauche, dont l'angle est formé par la classe 0, présente les classes ayant le taux de liquidité (R07, R08), la couverture du chiffre d'affaire (R17, R18) et l'autonomie financière (R04) les plus faibles. A l'opposé, on trouve dans la partie inférieure droite avec comme angle la classe 15, des entreprises qui montrent les valeurs les plus élevées pour l'ensemble de ces ratios. De manière générale, plus on monte dans la carte, plus les ratios se dégradent.

La classe 12 montre un profil ayant des ratios meilleurs que la classe 0, mais qui restent cependant inférieurs à la moyenne. La classe 3, quant à elle montre des ratios au-dessus de la moyenne, tout en étant largement inférieurs à ceux de la classe 15.

Afin de tester l'intérêt des indicateurs de type qualitatif, un apprentissage a été mené avec ces seules données, mais les résultats ne permettent pas de réaliser de distinction entre les entreprises saines et défaillantes. Nous abordons donc les tests réalisés avec la combinaison des deux batteries de variables mêlées, sans aucune sélection dans un premier temps, puis avec sélection dans un deuxième temps, dans le paragraphe 2.2.

### b) Travail des métaux

L'échantillon issu du secteur « Travail des métaux » étant nettement plus réduit, nous n'utilisons que des tests en neuf classes. Le nombre d'exemples par classe serait sinon bien trop réduit.

0	1	2
S 38%	S 33%	S 31%
D 62%	D 67%	D 69%
12.40	13.95	12.40
3	4	5
S 33%	S 38%	<b>S 23%</b>
D 67%	D 62%	<b>D 77%</b>
11.63	6.20	<b>10.08</b>
6	7	8
<b>S 89%</b>	<b>S 100%</b>	<b>S 73%</b>
<b>D 11%</b>	<b>D 0%</b>	<b>D 27%</b>
<b>14.73</b>	<b>10.08</b>	<b>8.53</b>

*Fig.3 – La carte Travail des métaux 9 classes sur données quantitatives*

Le nombre de classes représentant les entreprises défaillantes est supérieur au nombre de classes d'entreprises saines. Les taux de bons classements obtenus pour les trois classes « saines » sont élevés. La distinction paraît plus difficile au sein des entreprises défaillantes. Etant donnée la taille réduite de l'échantillon, ce test reste à prendre avec précaution. Nous avons donc pris l'option de réduire artificiellement le nombre de variables quantitatives en ne retenant que les trois plus pertinentes, à savoir R08, R17 et R24. Les résultats sont présentés ci-après.

### c) Travail des métaux avec sélection

Suite à la sélection, les variables retenues sont R08 (liquidité réduite), R17 (couverture du chiffre d'affaires par le fonds de roulement), R24 (rentabilité économique). Les résultats sont alors plus satisfaisants. Le nombre des données d'entrée ayant été réduit, les résultats sont du même coup moins sujet à caution que ceux présentés précédemment.

0	1	2
S 60%	S 22%	S 40%
D 40%	D 78%	D 60%
8	14	12
3	4	5
S 89%	S 41%	S 18%
D 11%	D 59%	D 82%
7	17	13
6	7	8
S 100%	S 67%	S 67%
D 0%	D 33%	D 33%
11	12	7

*Fig.4 – La carte Travail des métaux 9 classes sur données quantitatives sélectionnées*

A partir de trois indicateurs seulement les taux de classification sont alors meilleurs que ceux obtenus avec l'ensemble des indicateurs de type quantitatif. La faiblesse de la taille de l'échantillon nuit donc à la capacité d'apprentissage du réseau de neurones dans le cadre de notre application. La sélection par secteur d'activité, limitant la taille de l'échantillon, oblige à pratiquer une sélection sur les variables à retenir en entrée du réseau de neurones artificiels.

Les classes représentant en majorité des entreprises saines sont éloignées des classes à dominantes défaillantes. Les classes les moins distinctives se placent sur la diagonale de la carte. Nous reprenons ici les profils des classes les plus typées à savoir :

- Les classes 1 et 5 à dominante « saines » : Ces deux classes présentent toutes deux des valeurs légèrement inférieures à la moyenne pour les ratios de liquidités et de couverture et elles se distinguent par l'indicateur de rentabilité économique dans la moyenne pour la classe 5 et très faible pour la classe 1.
- Les classes 3 et 6 à dominante « défaillantes » : La classe 6 est l'exact opposée de la classe 2, avec des valeurs très élevées pour les trois indicateurs. Elle contient 100% d'entreprises saines. La classe 3, proche de la classe 6 sur la carte se distingue simplement par des valeurs légèrement inférieures, tout en restant très élevées pour le ratio de liquidités et de couverture et par un ratio de rentabilité économique proche de la moyenne.

De manière générale en partant de la classe 6, deux tendances sont à noter. Lors d'un déplacement vers la droite de la carte, les indicateurs de liquidité et de couverture se dégradent. Quand on se déplace vers le haut de la carte, l'indicateur de rentabilité économique se détériore.

Une fois les tests sur données quantitatives menés, nous avons introduit les données qualitatives dans notre expérimentation. De nouveaux tests ont alors été réalisés à partir de batteries d'indicateurs composés de données quantitatives et qualitatives.



## 2.2. Résultats sur combinaison données quantitatives et données qualitatives

Nous présentons dans un premier temps les résultats du secteur « Industrie » avant d'aborder les résultats du secteur « Travail des métaux ».

### a) Industrie

Sur les classifications en 9, 16 et 25 classes, la classification en 16 classes semble être encore la plus à même de distinguer clairement les entreprises saines des entreprises défailtantes. On obtient alors la carte suivante :

<b>0 (A)</b>	<b>1 (A)</b>	2 (A/F)	3 (F)
S 71%	S 86%	S 64%	S 100%
D 29%	D 14%	D 36%	
17.44	8.79	7.86	0.13
4 (A)	5 (A)	6	7 (F)
S 43%	S 37%		S 60%
D 57%	D 63%		D 40%
9.05	10.79		3.33
8 (A)	9	<b>10 (-)</b>	11(-)
S 32%		S 25%	S 62%
D 68%		D 75%	D 38%
10.92		7.06	6.66
12	<b>13 (-)</b>	14 (-)	15(-)
	S 20%	S 33%	S 40%
	D 80%	D 67%	D 60%
	5.99	4.39	7.59

*Fig.5 – La carte Industrie 16 classes sur données mixtes*

Deux groupes distincts apparaissent avec une frontière nette entre les deux. Les variables privilégiées sont la présence ou non d'actionnaires et de filiales, loin devant suivies par les ratios R06 (financement de l'actif circulant), R17 et R18 (ratios de couverture), puis les ratios de liquidités. Si l'on ne retient que des variables quantitatives, les taux de distinction sont similaires et on constate que l'apport des données qualitatives reste faible dans le cadre de ce test. La principale distinction sur la carte est réalisée à partir des indicateurs de filiales et d'actionnariat. Dans le tableau ci-dessus la présence du (A) et du (F) correspond à l'existence d'actionnaire ou de filiale. Les classes comprenant les plus forts taux de défaillance sont représentées par des classes ne disposant ni de filiale ni d'actionnaire (partie inférieure droite de la carte). Le positionnement sur la carte dépend principalement de ces deux variables.

### b) Industrie avec sélection

Ces variables sont ici choisies à partir d'un croisement entre l'étude des corrélations entre variables et les variables privilégiées par les cartes de Kohonen lors de l'apprentissage sans sélection. Au final six variables sont retenues R04 (autonomie financière), R17 (couverture du chiffre d'affaires par le fonds de roulement), R23 (productivité du capital investi), R34 (part des prêteurs) ainsi que l'effectif et la date de création. La classification en 9 classes montre peu de distinctions entre les saines et les défailtantes. La classification en 16 classes semble plus pertinente.

0	1	2	3
S 37%	S 49%	S 66%	<b>S 89%</b>
D 63%	D 51%	D 44%	<b>D 11%</b>
9	8	7	<b>7</b>
4	5	6	<b>7</b>
S 38%	S 49%	S 53%	<b>S 86%</b>
D 62%	D 51%	D 47%	<b>D 14%</b>
5	5	6	<b>6</b>
8	9	10	11
S 48%	S 41%	S 46%	S 63%
D 52%	D 59%	D 54%	D 37%
8	8	6	6
12	<b>13</b>	<b>14</b>	15
S 41%	<b>S 20%</b>	<b>S 22%</b>	S 47%
D 59%	<b>D 80%</b>	<b>D 78%</b>	D 53%
4	4	5	7

*Fig.6 – La carte Industrie 16 classes sur données mixtes sélectionnées*

Si quatre classes montrent des taux très marqués, il reste néanmoins beaucoup de classes avec des taux peu distinctifs. Dès que l'on augmente le nombre de classes jusqu'à 25, les capacités de distinctions entre classes redeviennent faibles.

Les variables privilégiées et largement dominantes dans le cadre du test en 16 classes sont la date de création et le ratio R34. L'effectif n'est pas mis en avant. Les entreprises les plus récentes se situent dans la partie supérieure gauche de la carte, les entreprises les plus anciennes dans la partie inférieure droite de la carte. Les dates évoluent donc le long de cette diagonale. La diagonale opposée montre une évolution du ratio R34 prenant des valeurs de plus en plus fortes. Ainsi la partie supérieure droite montre des entreprises à très faible taux d'endettement, tandis que la partie inférieure gauche montre des entreprises à très fort taux d'endettement.

Ainsi les groupes d'entreprises à dominante « saines » sont celles présentant une part des prêteurs, R34, extrêmement faible, combinée à une forte ancienneté. A l'inverse, les classes 13 et 14 représentant les entreprises à dominante « défaillantes » disposent d'une part des prêteurs très forte et de dates de création très anciennes pour la classe 14 et un peu plus récente pour la classe 13 (tout en restant ancienne). La distinction principale repose donc sur le ratio R34.

Nous reprenons ici les caractéristiques des classes les plus marquées :

- La classe 3 et 7, à dominante « saines » : Elles montrent toutes deux un ratio R23 (productivité du capital investi), un ratio R17 (couverture du chiffre d'affaires par le fonds de roulement) dans la moyenne et un ratio R34 largement inférieur à la moyenne. Par contre la classe 7 dispose d'un effectif faible et d'une création déjà très ancienne combinés à une autonomie financière (R04) très forte. La classe 3, quant à elle, dispose d'un effectif élevé et d'une création dans la moyenne avec une autonomie financière également supérieure à la moyenne mais de manière moins marquée que la classe 7.
- Les classes 13 et 14 à dominante « défaillantes » : La principale distinction avec les classes 3 et 7 vient du ratio part des prêteurs (R34) qui montrent ici des valeurs très élevées. Ces deux classes disposent de valeurs se situant pour la plupart près de la moyenne. La classe 13 prend des valeurs légèrement inférieures à la classe 14 pour les

ratios R23, R04 et R17. La distinction majeure entre les deux vient de la date de création très ancienne pour la classe 14 et dans la moyenne pour la classe 13.

Ces observations soulignent le fait que si la distinction entre entreprises saines et défaillantes se fait principalement à partir d'indicateurs de type endettement ou liquidité, il existe ensuite une multitude de situations se plaçant le long d'un continuum pour l'ensemble des ratios où des entreprises saines et des entreprises défaillantes peuvent présenter les mêmes caractéristiques.

### c) Travail des métaux

Etant donné l'étroitesse de l'échantillon issu du secteur « Travail des métaux », même en restant sur un test en neuf classes, nous avons donc pris l'option de limiter le nombre de variables en entrée en ne retenant que les trois indicateurs R08, R17 et R24, combinés au données qualitatives.

0	<b>1</b>	2
S 54 %	<b>S 74 %</b>	S 65 %
D 46 %	<b>D 26 %</b>	D 35 %
19	<b>18</b>	13
3	<b>4</b>	5
	<b>S 25 %</b>	S 33 %
	<b>D 75 %</b>	D 67 %
	<b>6</b>	5
6	7	8
S 48 %		S 32 %
D 52 %		D 68 %
18		22

*Fig.7 – La carte Travail des métaux 9 classes sur données mixtes sélectionnées*

Peu de distinctions apparaissent à travers les indicateurs utilisés, qu'ils soient qualitatifs ou quantitatifs. Dans l'ensemble, les classes à dominante « saines » disposent de ratios R08, R17 et R24 légèrement plus élevés que les classes à dominante « défaillantes ». Dans les trois classes « saines », celles disposant du plus faible taux de disparition montrent les dates de création les plus anciennes. Concernant les classes à dominante « défaillantes », celle présentant le plus faible taux d'entreprises saines montre un effectif inférieur à la moyenne.

Ces constats ne permettent pas d'améliorer la carte déjà obtenue à partir des données uniquement quantitatives pour le secteur « Travail des métaux ». A partir de cette nouvelle combinaison de variables quantitatives et qualitatives, un constat s'impose, l'apport des données qualitatives est moindre dans le cas de l'application à un secteur déjà délimité tel que le secteur « Travail des métaux », que dans le cadre d'un secteur moins clairement défini tel que le secteur « Industrie ».

## **Conclusion**

Nous ne retiendrons ici que les deux constats majeurs de cette expérimentation. Tout d'abord si le secteur est déjà clairement délimité, les données quantitatives semblent suffisamment à même de refléter les spécificités des entreprises saines et défaillantes. De plus, les cartes reflètent une multitude de situations possibles évoluant le long d'un continuum allant d'une forte probabilité de défaillance à une forte probabilité de survie des entreprises.

Dans l'ensemble l'apport des données qualitatives dans le cadre de notre expérimentation, reste faible en terme de taux de bon classement. Il permet néanmoins d'améliorer l'interprétation en sortant des données uniquement de type quantitatif. Nous rappelons qu'il s'agit ici simplement d'un essai sur quelques données disponibles directement sur Cédéroms. Il conviendrait de mettre en place une application à partir de données issues d'échantillons enrichis en informations qualitatives (changement de statut, existence de prêts, inscriptions de privilèges...). La mise en place d'une telle expérimentation sur un secteur d'activité nécessite la création d'un fichier comportant au moins 1000 entreprises et la collecte des données devrait s'étaler sur six mois environ.

Il n'en reste pas moins que les cartes de Kohonen représentent un outil prometteur dans le cadre d'applications en gestion. Basées sur une analyse de proximité entre classes, elles sont susceptibles d'apporter de nouveaux éclairages dans de nombreux domaines de la classification. De plus la visualisation de leurs résultats par l'intermédiaire de cartes représente un apport non négligeable dans l'étude du comportement des entreprises, comme dans d'autres applications. L'étude de C. PERRAUDIN et A. AKARCAY-GURBUZ (2002), qui réagit en terme de « logique de parcours », en intégrant la variable temps, pour les pays de l'Union Européenne peut se voir adapter au problème de la défaillance d'entreprise.

## **BIBLIOGRAPHIE**

COTTRELL M., FORT J.C. and PAGES G. (1998), « **Theoretical Aspects of the SOM Algorithm** », Neurocomputing, n° 21, p.119-138.

COTTRELL M. and GAUBERT P. (1999), « **Neural Network and Segmented Labour Market** » European Journal of Economics and Social Systems, vol. 13, n°1, p. 19-40.

De BODT E., COTTRELL M. et LEVASSEUR M. (1995), « **Les réseaux de neurones en finance : principes et revue de la littérature** », Finance, vol. 16, n° 1, p. 25-91.

DREYFUS G., MARTINEZ J.M., SAMUELIDES M., GORDON M.B., BADRAN F., THIRIA S. et HERAULT L. (2002), **Réseaux de neurones : méthodologie et applications**, Eyrolles.

IBM (1999), **IBM DB2 Intelligent Miner for Data, Using the Intelligent Miner for Data**, User Guide, Version 6, Release 1.

PERRAUDIN C. et AKARCA Y-GURBUZ A. (2002), **Comment situer l'économie de la Turquie parmi les économies de l'UE ? Une analyse exploratoire**, Conférence ACSEG 2002, Boulogne sur Mer, Novembre.

PEREZ (2002), **De l'analyse de performance à la prévision de défaillance : les apports de la classification neuronale**, Thèse Sciences de Gestion, Université Jean Moulin, Lyon 3, Décembre.

PEREZ (2002), **Performance des PME et classification neuronale : une application sur plusieurs secteurs d'activités**, Conférence ACSEG 2002, Boulogne sur Mer, Novembre.

STATISTICA (1998), **Guide utilisateur**, Statsoft.

## ANNEXE

### Les trente-cinq indicateurs de la base de données Diane de la Scrl

Code	Type	Ratios
Scrl_01	Structure et liquidité	Equilibre financier
Scrl_02		Indépendance financière (%)
Scrl_03		Endettement (%)
Scrl_04		Autonomie financière (%)
Scrl_05		Degré d'amortissements des immobilisations corporelles (%)
Scrl_06		Financement de l'actif circulant net
Scrl_07		Liquidité générale
Scrl_08		Liquidité réduite
Scrl_09	Gestion	Rotation des stocks (J)
Scrl_10		Crédit clients (J)
Scrl_11		Crédit fournisseurs (J)
Scrl_12		Chiffre d'affaires par effectif (kF/emploi)
Scrl_13		Taux d'intérêt financier (%)
Scrl_14		Endettement global (J)
Scrl_15		Capacité de remboursement
Scrl_16		Capacité d'autofinancement (%)
Scrl_17		Couverture du chiffre d'affaires par le fonds de roulement (J)
Scrl_18		Couverture du chiffre d'affaires par le besoin en fonds de roulement (J)
Scrl_19		Exportation (%)
Scrl_20	Productivité Et rentabilité	Efficacité économique (kF/emploi)
Scrl_21		Productivité du potentiel de production
Scrl_22		Productivité du capital financier
Scrl_23		Productivité du capital investi
Scrl_24		Rentabilité économique (%)
Scrl_25		Performance (%)
Scrl_26		Rendement brut des fonds propres nets (%)
Scrl_27		Rentabilité nette (%)
Scrl_28		Rendement des capitaux propres nets (%)
Scrl_29		Rendement des ressources durables nettes (%)
Scrl_30	Marge et valeur ajoutée	Taux de marge commerciale (%)
Scrl_31		Taux de valeur ajoutée (%)
Scrl_32		Part des salariés (%)
Scrl_33		Part de l'Etat (%)
Scrl_34		Part des prêteurs (%)
Scrl_35		Part de l'autofinancement (%)