UNE METHODE DE CLASSIFICATION DE DONNEES BINAIRES BASEE SUR LA DECOMPOSITION RECTANGULAIRE

Hédia Mhiri Sellami

Institut Supérieur de Gestion; 41 Av de la liberté. Bouchoucha.Tunis
Email : hedia.mhiriàisg.rnu.tn

Ali Jaoua

King Fahd University of Petroleum and minerals, Information and Computer Science Department. P.O. Box: 126,Dhahran 3126
Emai l: ajaoua@ccse.kfupm.edu.sa

Introduction

Les tableaux binaires sont utilisés dans différents domaines et notamment en intelligence artificielle, on a souvent recours par exemple en apprentissage à ce type de tableaux pour représenter la connaissance et l'exploitation de cette connaissance fait souvent appel à des techniques de partitionnement, de classification supervisée, l'extraction des rectangles maximaux d'une relation binaire finie fait aussi partie de ces techniques. Ce principe d'extraction a d'ailleurs fait l'objet d'études par des mathématiciens dans le contexte de la théorie des treillis et dont la pertinence a été démontrée dans plusieurs autres domaines d'application [RIG48]. Le but initial de notre travail était de comparer les résultats de cette décomposition rectangulaire à ceux générés par la classification. Le problème est que l'on parle de classe en classification alors que la décomposition rectangulaire génère des recouvrements. Nous avons alors proposé une transformation à cette décomposition afin qu'elle génère des classes. Le résultat est une nouvelle stratégie de classification de données binaires. Dans ce travail nous commençons par présenter certains concepts théoriques que nous utilisons. Nous exposons ensuite les principes de la décomposition rectangulaire ainsi que notre stratégie puis nous terminons par une comparaison des résultats obtenus par notre stratégie à ceux fournis par certaines méthodes de classification.

1. Définitions et propriétés

1.1. Définition d'une partition

Une partition de Ω est un ensemble $P = (P_1, \cdots, P_K)$ de parties non vides de Ω d'intersections vides deux à deux et dont la réunion forme Ω ayant les propriétés suivantes:

1) $\forall \ell = 1, \cdots, K$ on a $P_\ell \neq \varnothing$

2) $\bigcup_{\ell=1}^{K} P_l = \Omega$

3) $\forall \ell, m = 1, \cdots, K$ et $\ell \neq m$ alors $P_\ell \cap P_m = \varnothing$

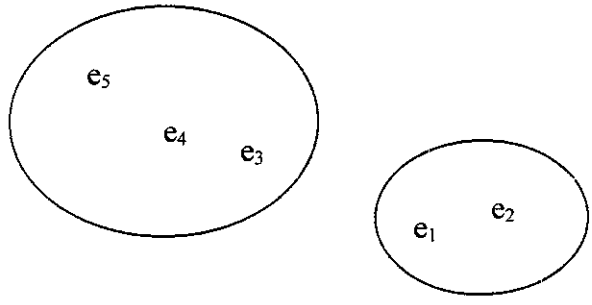


figure 1 : représentation d'une partition

1-2 Définition d'un recouvrement

Un recouvrement de Ω est un ensemble $P = (P_1, \cdots, P_K)$ de parties non vides de Ω vérifiant :

1) $\forall \ell = 1, \cdots, K$ on a $P_\ell \neq \varnothing$

2) $\bigcup_{\ell=1}^{K} P_l = \Omega$

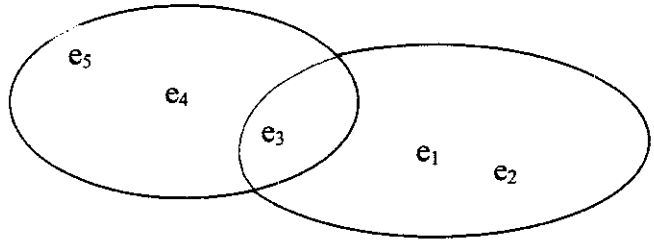


figure 2 : représentation d'un recouvrement

Une partition est donc un cas particulier de recouvrement [CEL89].

1.3. Définition d'une relation binaire

Une relation binaire d'un ensemble E dans un ensemble F est un sous-ensemble du produit cartésien ExF [BEL94].

Un élément d'une relation R est noté (x,y). On dit que x est un argument de R et y est une image de x par R. On note donc xRy le fait qu'un élément x de E soit lié à un élément y de F.

Parmi les relations binaires nous distinguons la relation d'identité I. Ainsi si S est un ensemble quelconque alors $I(S) = \{(x,x) \in R / x \in S\}$. A une relation R de E dans F nous pouvons aussi associer les sous-ensembles suivants de E ou de F:

- l'ensemble image de x défini par: $x.R = \{ y/ xRy \}$;

- les antécédents de y définis par: $R.y = \{ x/ xRy\}$;

- le domaine de R défini par: $dom(R) = \{ x/ \exists y : xRy \}$

- le co-domaine de R défini par : $cod(R) = \{y/ \exists x : xRy \}$;

- l'inverse d'une relation R est la relation $R^{-1} = \{ (x,y) / yRx\}$.

1.4. Définition d'un rectangle

Soit R une relation binaire définie de E dans F. Un rectangle de R est un couple de deux ensembles (A,B) tel que $A \subseteq E$, $B \subseteq F$ et $AxB \subseteq R$. A est le domaine du rectangle et B est le co-domaine.

Cette notion de rectangle existe aussi dans d'autres domaines tels que la théorie des graphes [GAR79] où une relation binaire R de E dans F définit les arcs d'un graphe bipartie sur

$(E \cup F,R)$ et un rectangle est un sous-graphe bipartie complet. La notion de rectangle occupe aussi une place dans la théorie de l'apprentissage à partir d'exemple où elle se retrouve sous le nom de couple complet [GOD89].

1.5. Définition d'un rectangle maximal

Soit R une relation binaire définie de E dans F. Un rectangle (A,B) est dit maximal si et seulement si $A*B \subseteq A'*B' \subseteq R \Rightarrow A=A'$ et $B=B'$.

1.6. Définition d'une relation élémentaire

Soient R une relation binaire finie et $(a,b) \in R$. L'union des rectangles qui contiennent l'élément (a,b) est égale à la relation $\Phi R(a,b) = I(b.R^{-1}) \circ R \circ I(a.R)$. La relation $\Phi R(a,b)$ est la relation élémentaire contenant l'élément (a,b) et le symbole "o" représente l'opération de composition des relations. Nous montrons ci-dessous une relation R ainsi que le $\Phi R(a,1)$ contenant l'élément (a,1):

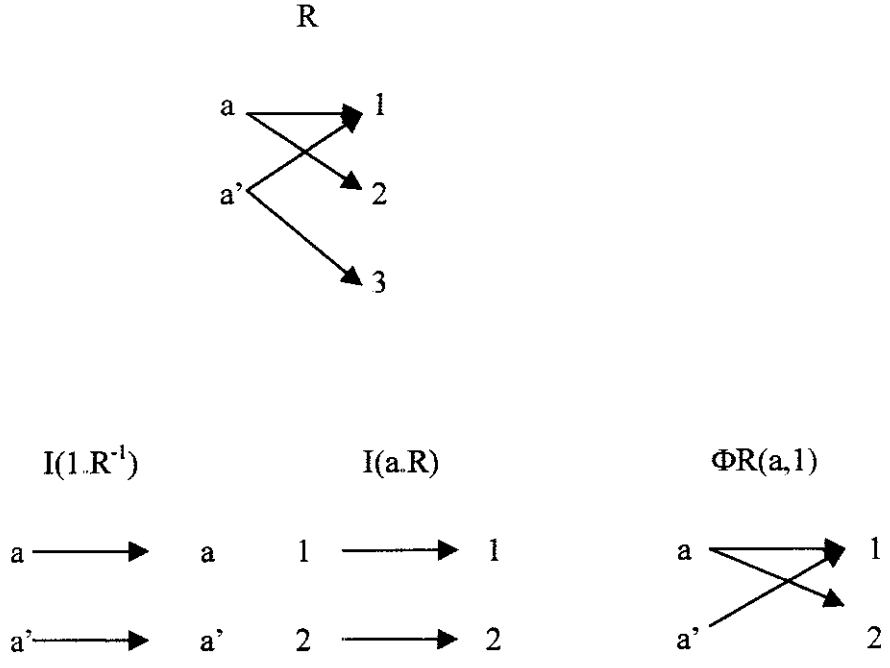
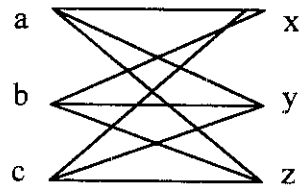


figure 3 : représentation de la notion de relation élémentaire

On remarque qu'un rectangle est une relation élémentaire particulière (l'inverse n'est pas toujours vrai).

En travaillant sur les concepts de rectangle les auteurs [BEL94] proposent de remplacer chaque rectangle RE=(A,B) ayant i*j couples par i+j couples (i = card(A) et j = card(B)). En effet soit la relation RE suivante formée de (3*3) couples ainsi que sa représentation condensée



a	1
b	1

x	1
y	1

équivalente:

figure 4 : représentation équivalente d'un rectangle

En représentant tous les éléments du rectangles par les couples (a,u) et (u,y), où u est une constante qui identifie le rectangle (ici u=1), a représente un élément quelconque de l'ensemble A et y un élément quelconque de l'ensemble B on a réalisé un gain de l'économie de codage de l'information noté gij et défini par gij= (i*j) - (i+ j). Nous pouvons montrer cette notion de gain sur la relation R suivante ainsi que les trois rectangles maximaux contenant l'élément (y,3):

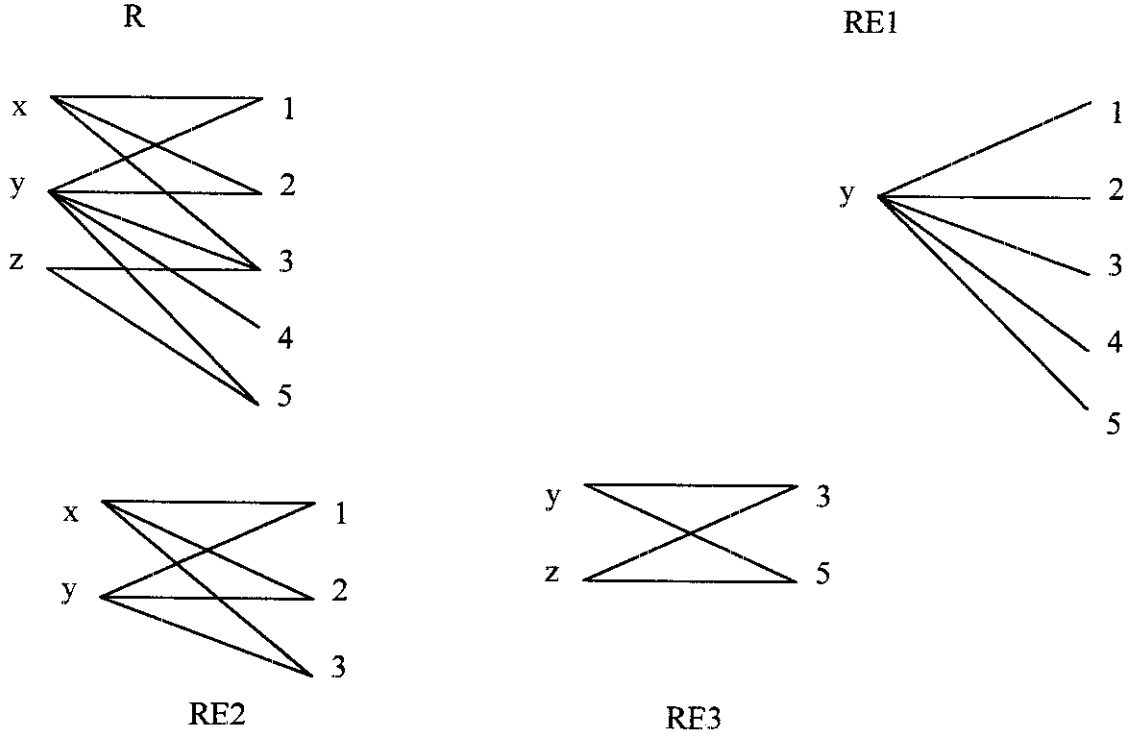


figure 5 : représentation de rectangles maximaux

La valeur du gain du rectangle RE1 est de -1, pour RE2 on trouve 1 et pour RE3 on a 0.

1.7. Définition d'un rectangle optimal

Un rectangle contenant un élément (a,b) d'une relation R est dit optimal s'il réalise le maximum de gain parmi tous les rectangles maximaux qui contiennent (a,b).

Dans le dernier exemple le rectangle optimal contenant l'élément (y,3) est RE2 puisqu'il réalise le maximum de gain à savoir 1.

1.8. Définition d'une couverture

Nous appelons couverture d'une relation R un ensemble de rectangles C = { RE1, RE2, ... REn} de R tel que tout élément (a,b) de R appartienne à au moins l'un des rectangles de C.

Le problème de la recherche du rectangle optimal contenant un élément (a,b) d'une relation binaire peut aussi consister à chercher le sous-graphe bipartie complet de cardinalité maximale contenu dans un graphe bipartie [GAR79]. Les auteurs [BEL94] proposent un algorithme de décomposition dont le principe est de chercher des rectangles recouvrant une relation et ce en procédant comme suit:

- séparer la relation R en *p* relations élémentaires PR1, ... PRp.
- effectuer sur chacune des relations élémentaires PRi la recherche de tous les rectangles optimaux contenant un élément de PRi
- pour chaque relation élémentaire PRi sélectionner le minimum de rectangles optimaux qui couvrent PRi
- éliminer au maximum les éléments qui se répètent à travers les rectangles optimaux.

Pour cela les auteurs utilisent donc une heuristique basée sur la méthode "branch and bound" [BRA87] et évaluent le gain mesuré par la fonction de gain suivante (qui se substitue à la fonction gij précédente) :g(PR') = (r/d*c)(r - (d+c)); avec r = cardinal(PR'); d = cardinal(dom(PR')) et c = cardinal(cod(PR')).

2. Présentation de la stratégie de classification basée sur la décomposition rectangulaire

2.1. La décomposition rectangulaire

L'extraction des rectangles d'une relation binaire finie est un problème qui a fait l'objet de plusieurs études par des mathématiciens dans le contexte de la théorie des treillis [RIG48] et [JAO92]. La stratégie de décomposition rectangulaire qui en découle est utilisée dans plusieurs domaines tel que l'apprentissage, la recherche documentaire [GUE90], [WIL90]. Le principe est de trouver une couverture d'une relation binaire R par un nombre minimal de rectangles. Actuellement plusieurs travaux permettent de générer à partir d'un tableau binaire l'ensemble des rectangles associés et bien entendu les couvertures qui en découlent [KHC97]. Pour illustrer cela nous prenons une relation R représentée par un tableau binaire composé de dix micro-ordinateurs en ligne (a, b .... j) et de dix variables (v1, ... v10). Une valeur de "1" indique que le micro possède la variable en question [CEL89]:

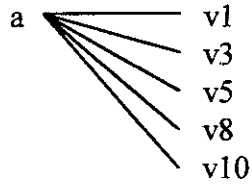
a	1	0	1	0	1	0	0	1	0	1
b	0	1	0	1	0	1	1	0	1	0
c	1	0	0	0	0	0	0	1	1	0
d	1	0	1	0	0	0	0	1	0	0
e	0	1	0	1	0	1	1	0	1	0
f	0	1	0	0	0	1	1	0	1	0
g	0	1	0	0	0	0	0	1	0	1
h	1	0	1	0	1	1	0	1	1	1
i	1	0	0	1	0	0	0	0	0	1
j	0	1	0	1	0	0	1	0	0	0

Un exemple de rectangle associé a ce tableau peut être R1qui a pour domaine { b,e,f} et pour co-domaine {v2,v6,v7,v9}.

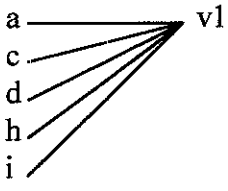
En s'inspirant des travaux relatifs à la décomposition rectangulaire nous avons introduit des changements à cette méthode pour aboutir à une stratégie de classification des tableaux binaires. Etant donné que la décomposition rectangulaire génère des recouvrements les suggestions que nous avons proposées ont permis à notre algorithme de générer des classes.

2.1. Principes d'une stratégie de classification basée sur la décomposition rectangulaire

Etant donnée une relation R, notre travail consiste à déterminer pour chaque couple (a,b) de R, la relation élémentaire ΦR(a,b) correspondante. Nous allons utiliser le tableau binaire présenté dans la section précédente pour illustrer les différentes étapes de notre stratégie. Nous commençons donc par déterminer a.R, qui correspond à toutes les images de la première observation et donc à cod(ΦR(a,v1)) et ayant pour cardinal c=5:



En seconde étape nous cherchons les antécédents de v1 ou encore dom(ΦR(a,v1)), son cardinal d a pour valeur 5:



Il nous reste à déterminer les autres éléments (arcs) qui lient ces deux ensembles relativement à R pour avoir la valeur de r. Avec ces deux ensembles nous déterminons la valeur de r qui correspond à card(ΦR(a,v1)) et qui a pour valeur 17. La valeur du gain g(ΦR(a,v1)) est ensuite calculée en utilisant la dernière formule du paragraphe précédent qui est : gain = (r/d*c)(r - (d+c)) , elle génère comme valeur 4 pour ce cas. Ce processus est répété pour tous les couples (a,vi) de R puis nous cherchons la plus grande valeur de r [KHC97]. Le domaine associé à cette valeur correspond à une classe de E, et l'ensemble des observations qui composent cette classe est alors supprimé de l'échantillon. Le processus est réitéré jusqu'à épuisement des observations. L'ensemble des différentes classes que nous obtenons correspond à une partition.

2.2. Description de l'algorithme

T un tableau binaire qui contient la relation R et ayant n lignes et p colonnes.
Algorithme Classer (T);

Début
i :=0;
Tant que (i < n) faire
Début
j:=0;
Tant que (j < p) faire
Début
Déterminer Φ(T[i, j]);
Calculer (g (Φ(T[i, j]));
Ecrire [dom(Φ(T[i, j])), g (Φ(T[i, j]))];
j=j+1;
Fin
i :=i+1;
Fin
Classe := Chercher_max g (Φ(T[i, j]));
T:= T - Classe;
Fin

Chercher_max est une fonction qui cherche la valeur du maximum de gain et rend comme résultat l'ensemble des observations dom(Φ(T[i, j])) qui est à l'origine de ce gain. La fonction Ecrire permet d'écrire dans un fichier les observations qui ont généré la valeur maximale de g et correspondent donc à une classe.

Ce programme se poursuivra jusqu'à ce que le tableau T soit vide, c'est à dire que toutes les observations soient classées. L'ensemble des différentes classes constitue notre partition.

Nous avons transcrit cet algorithme en utilisant le langage C++, et le résultat est un programme de 256 lignes qui tourne actuellement sur un Pentium.

3. Comparaison des résultats

3.1. Critère de comparaison

Pour valider les résultats de notre stratégie nous devons comparer ses résultats avec ceux générés par des méthodes de classification. Le critère de comparaison que nous utilisons est

celui utilisé par les auteurs [CEL89] et qui correspond au rapport entre l'inertie interclasse et l'inertie totale. Nous présentons ce rapport par: R=B/T où **B** correspond à l'inertie interclasse, **T** correspond à l'inertie totale du nuage et R représente la part d'inertie conservée ou expliquée en assimilant les individus aux k centres de gravité (k correspond au nombre des classes). Plus la valeur de R est forte plus la partition constitue une bonne représentation des individus. La formule générale de l'inertie totale T relativement à un nuage de points est la suivante:

$$I = \sum_{i=1}^{n} p_i d^2(x_i, g)$$

où g est le centre de gravité de tout le nuage, d est la distance euclidienne simple et p_i est le poids associé à l'observation x_i. Lorsqu'on associe à une population une partition en k classes $P = (P_1, \cdots, P_K)$, on peut définir :

$$\mu_l = \sum_{x_i \in P_l} p_i$$

avec $\sum_{l=1}^{k} \mu_l = 1$. Le centre de gravité g_l relatif à chaque classe a comme formule générale:

$$g_l = \frac{1}{\mu_l} \sum_{xi \in P_l} p_i x_i .$$

L'inertie interclasse B peut ainsi être définie comme:

$$B = \sum_{l=1}^{k} \mu_l d^2(g_l, g)$$

Il s'agit de l'inertie du nuage des centres de gravité g_l munis des poids μ_l. Remarquons que l'on a généralement T=W+B où l'inertie intraclasse a la formule suivante:

$$W = \sum_{l=1}^{k} \sum_{x_i \in P_l} p_i d^2(x_i, g_l) .$$

Cela s'énonce aussi comme : l'inertie totale est la somme de l'inertie interclasse (B) et de l'inertie intraclasse (W). Pour dire qu'une partition est "meilleure" qu'une autre (au sens du critère sélectionné), il suffit de prendre la partition qui génère la plus grande valeur de R. Maximiser R revient à alors à minimiser l'inertie intraclasse W et maximiser l'inertie interclasse B. Les classes que nous obtenons sont donc comparées à d'autres classes générées par d'autres méthodes de classification.

3.2. Premier test

Le premier échantillon sur lequel nous avons évalué notre stratégie correspond au tableau des dix observations et dix variables présenté dans la section précédente [CEL89]. En appliquant une classification croisée sur ce tableau l'auteur a obtenu l'ensemble des trois classes suivantes: { {a,d,h}; {b,e,f, j}; {c,g, i} }. Nous avons calculé la valeur de R associée à cette partition, elle est de 0.61731. La partition extrait donc 61% de l'inertie totale. En appliquant notre méthode sur le même tableau de données nous obtenons une partition en quatre classes: { {a,d,h}; {b,e,f, j}; {c, i} ; {g} } et la valeur du R associée est de 0.703147. Notre partition en quatre classes extrait donc 70% de l'inertie totale.

3.3. Deuxième test

Le deuxième échantillon utilisé correspond à un tableau binaire de 40 fonctions élémentaires (lignes) caractérisant 12 systèmes de fichiers (colonnes) [DID82]. Les classes obtenues par une classification croisées sont au nombre de quatre et la valeur de R associée est de 0.472020. La partition correspondante extrait donc 47% de l'inertie totale. En appliquant notre programme sur le même échantillon nous obtenons une partition en sept classes et la valeur de R est de 0.475639. Notre partition extrait donc un taux d'inertie supérieur à la précédente (48%).

3.4. Troisième test

L'échantillon correspond à un ensemble de 59 observations représentant des plaques-boucles de ceintures damasquinées du nord-est de la France s'échelonnant entre la fin du VI ème et le VIII ème siècle. Les archéologues ont observé sur ces plaques la présence et l'absence de 26 critères relatifs à la technique de fabrication, à la forme, au décor..... [PER80]. Le but de ces archéologues est d'établir une typologie de ces plaques-boucles et de mettre en évidence l'évolution dans les techniques de fabrication. Les chercheurs de cette équipe ont eu recours à plusieurs techniques de classification et ce relativement au même échantillon initial.

3.4.1. Classification hiérarchique

Pour obtenir une typologie des données les auteurs ont eu recours à deux classifications hiérarchiques, l'une sur les critères et l'autre sur les plaques à la suite desquelles ils ont généré six classes d'observations correspondant au premier niveau de la hiérarchie. La valeur de R correspondant aux classes générées par cette technique est 0.764042; soit donc un taux de 76%.

3.4.2. Traitement graphique

Les auteurs ont aussi réalisé un traitement graphique sur la matrice de données et ce par permutation manuelle des lignes et des colonnes [LEP90]. Le résultat de ce traitement fût une mise sous forme diagonale de la matrice de données qui a permis de distinguer dix classes. La valeur de R correspondant aux classes générées par cette technique est 0.723848, soit donc un pourcentage d'inertie de 72%.

3.4.3. Traitement par l'analyse factorielle

Le tableau de données initiales a aussi subit un autre traitement. Il s'agit d'étudier ces mêmes données à l'aide d'une analyse factorielle des correspondances [DID82]. Les projections simultanées des plaques et des critères dans le premier plan factoriel s'organisent sous forme d'une parabole. Les auteurs ont donc réorganisé les lignes et les colonnes du tableau primitif en suivant les projections des plaques et des critères le long du premier axe factoriel. Le résultat était donc un tableau à partir duquel les auteurs ont distingué neuf classes. La valeur de R correspondant aux classes générées par cette technique est 0.762482, soit donc un pourcentage d'inertie de 76%.

3.4.4. Classification automatique

Le dernier traitement réalisé par les auteurs du même papier fût l'utilisation d'un algorithme de classification automatique qui a permis de mettre en évidence une typologie et une évolution très proche de celles déterminées à l'aide des techniques graphiques matricielles. Les auteurs ont distingué onze classes. La valeur de R correspondant aux classes générées par cette technique est de 0.811843, soit donc un pourcentage d'inertie de 81%.

3.4.5. Notre stratégie

Nous avons appliqué notre programme sur le fichier de données initiales qui contient les 59 plaques décrites par les 26 variables. A la suite de la génération de la partition qui contient sept classes nous avons appliqué notre programme de calcul des paramètres d'aide à l'interprétation et ce essentiellement pour avoir la valeur de R. La valeur de R correspondant aux classes que notre programme a générées est égale à 0.760584; notre partition explique donc 76% du nuage initial.

La valeur de R correspondant aux classes générées à la suite de l'application de deux classifications hiérarchiques est 0.764042; soit donc un taux de 76% égal à celui généré par notre programme. Remarquons que la classification hiérarchique a été appliquée à deux reprises; une sur les observations et une autre sur les variables. Notre programme, quant à lui génère le même taux à la suite d'une seule application.

La valeur de R correspondant aux classes générées à la suite de l'application d'une permutation manuelle des lignes et des colonnes est 0.723848, soit donc un pourcentage d'inertie de 72%, valeur inférieure à celle générée par notre stratégie.

La valeur de R correspondant aux classes générées à la suite des résultats de l'analyse factorielle est 0.762482, soit donc un pourcentage d'inertie de 76%. Remarquons que dans ce cas il y a eu récupération des résultats de l'analyse factorielle puis regroupement des observations à la suite de leur projection sur les axes factoriels. Par contre notre stratégie génère un taux égal à celui de l'analyse factorielle et ce suite à une seule étape.

La valeur de R correspondant aux classes générées à la suite de l'application de deux classifications automatiques sur les lignes et les colonnes est 0.811843, soit donc un pourcentage d'inertie de 81%. Ce taux est certes supérieur à ce que génère notre programme cependant pour l'avoir il faut appliquer deux classifications. Notre stratégie génère un taux très proche à partir d'une seule application.

3.5. Quatrième test

Le quatrième échantillon est théorique. L'auteur [CAR84] traite un tableau théorique de 25 observations décrites par 45 variables. L'auteur génère une partition en utilisant des permutations manuelles. Le principe est de permuter les lignes et les colonnes pour avoir le

maximum de groupements de "zéro" ainsi que des groupements de "un". Cette technique génère une partition identique à celle générée par notre stratégie, nous n'avons donc pas calculé le R correspondant.

3.6. Cinquième test

Les données relatives à ce test correspondent à un tableau de 26 observations et 26 variables, et pour le manipuler l'auteur [CAR84] a utilisé indice de Jaccard [JAC08] pour grouper les entités. Dans ce cas aussi la partition présentée dans le papier est identique à celle générée par notre programme.

3.7. Sixième test

Le sixième test est aussi relatif à des données théoriques. Il s'agit de 25 observations et de 25 variables. L'auteur de ces données [CAR84] a aussi utilisé l'indice de Jaccard [JAC08] pour calculer les distances entre les lignes. Ces valeur ont été utilisées pour le groupement de ces lignes. Encore une fois la partition présentée par l'auteur est identique à celle générée par notre programme.

Le tableau suivant résume les différentes valeurs de R, les cas où cette valeur est la même pour notre stratégie et une méthode de classification sont absents:

	échantillon 10 observ	échantillon 40 observ	échantillon 59 observ
Notre stratégie	70%	48%	76%
Classification croisée	61%	47%	
2 classif. hiérarchiques			76%
Traitement graphique			72%
Analyse factorielle			76%
Classif. automatique			81%

4. Conclusion

Les techniques de groupement supervisées nécessitent souvent l'intervention de l'utilisateur qui doit avoir une idée sur le domaine étudié, et même si ce problème est partiellement résolu par l'utilisation des méthodes de partitionnement on reste dépendant du paramètre "nombre de classes" qui est aussi souvent choisi par l'utilisateur [CEL89]. Aussi avons-nous essayé de présenter une stratégie de classification de tableaux binaires qui ne nécessite pas la précision du nombre de classes. Notre stratégie est inspirée d'une technique utilisée dans le domaine de l'intelligence artificielle à savoir la décomposition rectangulaire. Ainsi en partant de concepts associés à la décomposition rectangulaire sur les tableaux binaires qui génère des recouvrements nous avons pu aboutir à une stratégie de classification qui permet de générer des partitions. Comparée à des méthodes de classification sur certains échantillons, notre stratégie offre des résultats similaires sans pour autant exiger ce paramètre "nombre de classe". En effet notre stratégie s'arrête dès l'épuisement de toutes les observations. Le nombre de classes généré par notre programme pour les différents échantillons reste très proche de celui généré par les méthodes de classification. Notre stratégie offre donc des performances similaires aux stratégies de classification qui existent dans la littérature tout en restant simple puisqu'elle a une complexité en $O(n^2)$.

Ces résultats sont fort encourageants, cependant nous cherchons à les valider sur des échantillons plus volumineux, ce qui sera notre prochaine étape. En une seconde étape nous cherchons aussi à nous intéresser d'avantage aux variables. Nous projetons de comparer l'influence des variables dans les partitions générées par notre stratégie par rapport aux autres partitions tout en utilisant les critères d'inertie [CEL89]. Finalement nous cherchons aussi à généraliser ce travail sur des variables non binaires.

BIBLIOGRAPHIE

[BEL94] Belkhiter, N. ; Bourhfir, C.; Gammoudi, M.M.; Jaoua, A.; Thanh, Le.; Reguig M. (1994) Décomposition rectangulaire optimale d'une relation binaire: Application aux bases de données documentaires. Revue INFOR vol 32 Feb 1994.

[BRA87] Brassard, G.; Bratley, P. Algorithmique, conception et analyse. Masson, Les Presses de l'Université de Montréal, 1987.

[CAN93] Can, F. Incremental clustering for dynamic information processing. CAM Trans. Inform. Syst. 11 , 143-164., 1993.

[CAR84] Caraux, G. Réorganisation et représentation visuelle d'une matrice de données numérique: un algorithme itératif. Revue de Statistique Appliquée, 1984, Vol XXXII, n°4.

[CEL89] Celeux, G.; Diday, E.; Govaert, G.; Lechevallier, Y.; Ralambondrainy, H. Classification automatique des données. Dunod informatique, 1989.

[DID82] Diday, E.; Lemaire, J. ; Pouget, J. ; Testu, F. Eléments d'analyse de données. Dunod, 1982.

[EVE80] Everitt, B. Cluster Analysis. Second Edition Halsted Press, 1980.

[GAR79] Garey, M.R.; Johnson, D.S. Computers and Interactability: A guide to the Theory of NP-Completeness. W.H. Freeman, 1979.

[GOD89] Godin. R; Gecsei. J; Pichet. C. Design of Browsing Interface for Information Retrieval. In Proceeding of the 12th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, N.J. Belkin and C.J. Van Rijsbergen (Ed), Cambridge, M.A: ACM SIGIR Forum, pp 32-39, 1989.

[GUE90] Guenoche, A. Construction de treillis de Galois d'une relation binaire". Mat. Inform. Sci Hum. 28 41-53, 1990.

[HUB85] Hubert, L. ; P. Arabie. Comparing Partitions. Journal of Classification. Springer Verlag., 1985.

[JAC08] Jaccard. P. Nouvelles recherches sur la distribution florale. Bull.Soc. Vaud.Sci. Nat., t. 44, pp. 223-270, 1908.

[JAO92] Jaoua, A.; Beaulieu, N. ; Desharnais, J. ; Reguig, M. Optimal Rectangular decomposition of a finite binary Relation. Sixth SIAm Conference on Discreta Mathematics, Vancouver, 1992.

[KHC97] Khcherif, R.; Jaoua, A. Rectangular Decomposition Heuristics for Documentary Databases. Review Informatics and Computer science, Intelligent Systems, Applications, 1997.

[LEP90] Le Prince, V. La méthode des pôles d'attraction et des pôles d'agrégation. Thèse de troisième cycle. Université de Paris VI, 1990.

[MAD96] Maddouri, M; Jaoua, A. Incremental rule production: toward a uniform approach for knowledgr organisation". Proc Ninth international conference on Industrial and engineering applications of artificial intelligence and expert systems, 1996.

[MHI99a] Mhiri Sellami, H.; Jaoua, A. Classifying binary data. The IX international Symposium on Applied Stochastic Models and Data Analysis. Lisbon, Portugal, 1999.

[MHI99b] Mhiri Sellami, H.; Jaoua, A. Non supervised rectangular classification of binary data ". The twelfth international conference on industrial and engineering applications of artificial intelligence and expert systems (IEA/AIE-99)., Le Caire, Egypte, 1999.

[MHI00] Mhiri Sellami, H.; Partitioning binary data; Compstat. Hollande, 2000.

[PER80] Perin, P; Leredde, H. Les plaques-boucles Mérovingiennes, Les dossiers de l'archéologie,. N° 42, Mars-Avril, 1980.

[RIG48] Riguet, J. Relations binaires, fermetures et correspondances de Galois, Bulletin de la Société mathématiques de France 76, 114-155, 1948.

[WIL90] Wille, R. Concept lattice & Conceptual knowledge systems. Computer Mathematic. Applied, 23 (493-515), 1990.