Qualité des règles d'association : étude de données d'entreprise

Benoît Vaillant, Stéphanie Ménou, Sorin Moga, Philippe Lenca, Stéphane Lallich

Qualité des règles d'association : étude de données d'entreprise

Benoît Vaillant*, Stéphanie Menou**, Sorin Moga***,
Philippe Lenca***, Stéphane Lallich****

*Département STID / VALORIA
Université de Bretagne Sud
8, rue Montaigne, B.P. 561
56017 Vannes Cedex

**DIXID
4, rue Ampère
22300 Lannion

***UMR 2872 TAMCIC
GET / ENST Bretagne
Technopôle de Brest Iroise
CS 83818, 29238 Brest Cedex

****Laboratoire ERIC
Université Lumière - Lyon 2
5 avenue Pierre Mendès-France
69676 Bron Cedex

Résumé. L'extraction de connaissances à partir de données a pour objet la découverte de connaissances à partir de grandes quantités de données, par des méthodes d'apprentissage automatiques ou semi-automatiques, et l'utilisation industrielle ou opérationnelle de ces connaissances.

Nous explorons ici une application concrète de la fouille de données, sur un système de serveur vocal. Nous nous intéressons en particulier à l'étape cruciale de validation de motifs extraits via l'utilisation de mesures de qualité.

1 Introduction

Parmi les nombreuses définitions de l'extraction de connaissances à partir de données (ECD par la suite), on retrouve fréquemment celle proposée par Fayyad et al. (1996) : le *processus complexe permettant l'identification, au sein des données, de motifs valides, nouveaux, potentiellement intéressants et les plus compréhensibles possible*. Cette définition insiste sur la notion de qualité et d'intérêt, vis-à-vis d'une utilisation finale. La notion d'intérêt varie ainsi en fonction du contexte applicatif et de l'utilisateur expert des données.

La validation de ces connaissances est une étape cruciale du processus d'ECD (Hilderman et Hamilton, 2001). Elle devient particulièrement incontournable en extraction de règles d'association face au volume de règles automatiquement produit. Partant du constat que la sélection des *bonnes* connaissances passe aussi par l'utilisation des *bonnes* mesures (Tan et al., 2002; Lenca et al., 2003; Carvalho et al., 2005), nous présentons dans Vaillant (2006) une étude systématique des mesures d'intérêt des règles d'association selon différents axes d'analyse. Parmi ces axes, nous proposons une aide multicritères à la décision afin de sélectionner les bonnes mesures (Lenca et al., 2007). Nous présentons ici une application de cette approche pour des données issues de l'interaction d'utilisateurs avec un service d'assistance téléphonique.

Nous exposons dans un premier temps, en section deux, le processus d'ECD que nous avons suivi. Dans la troisième section on présente les données étudiées ainsi que les premiers résultats bruts. La quatrième section présente brièvement les critères de décision retenus ainsi que des éléments liés à leur importance au regard de préférences exprimées par l'utilisateur. Dans la section cinq nous analysons les résultats préliminaires de l'approche aide multicritère à la décision en vue d'orienter le choix de mesure(s) de qualité(s). Enfin, nous concluons et proposons des perspectives à nos travaux.

2 Présentation du processus de fouille de données

Un processus d'ECD se décompose en plusieurs étapes. Classiquement on retrouve les étapes suivantes : ciblage des données à fouiller, nettoyage, transformation et recodage, fouille, évaluation des résultats, intégration des résultats.

Chaque étape, spécifique au problème traité, est validée par un utilisateur (tantôt un expert métier – $E_{\mathcal{R}}$, tantôt un expert de l'ECD –$E_{\mathcal{A}}$). Nous distinguons clairement plusieurs intervenants dans le processus. L'étude de cas que nous présentons permet de mettre aussi en évidence les différents points de vue, les différentes expertises et par voie de conséquence les difficultés inhérentes à ce type de collaboration.

Dans cette étude, nous avons suivi toutes les étapes, exceptée bien évidemment celle d'intégration des résultats puisque nous n'en sommes qu'aux résultats préliminaires. Nous développons particulièrement les étapes de fouille et d'évaluation des résultats.

L'étape de fouille est au cœur du processus. C'est lors de son déroulement que l'on génère les motifs à interpréter. A l'issue de cette étape, les résultats produits par les algorithmes de fouille de données ne sont pas toujours exploitables. Il est alors nécessaire de les soumettre à une évaluation. La quantité d'information pouvant être générée étant prohibitive pour une évaluation manuelle, il est courant d'automatiser un filtrage ou un ordonnancement de ces motifs, afin d'assister l'utilisateur expert des données dans la tâche de prise de décision ou d'interprétation des résultats. Ceci peut se faire au moyen de mesures de qualité, fonctions mathématiques modélisant une catégorie de connaissances désirée, et mettant celle-ci en avant parmi l'ensemble des motifs extraits.

Nous utiliserons en pratique l'implémentation proposée par Borgelt et Kruse (2002) de l'algorithme APRIORI d'Agrawal et Srikant (1994) afin d'extraire des règles d'association. Nous présentons succinctement le principe de cette fouille en section 2.1.

2.1 L'extraction de règles d'association

En extraction de règles d'association, on suppose que les données à explorer sont binaires, *i.e.* qu'on peut décrire chaque objet (ou transaction) au moyen d'un ensemble fini d'attributs booléens $\mathcal{I} = \{i_1, \ldots, i_m\}$, également appelés *items*. Pour un ensemble X d'items (on parle généralement d'itemsets) de $\mathcal{I}$, on dira qu'une transaction t contient X si et seulement si $X \subseteq t$ et on appelle support d'un itemset X le rapport entre le nombre de transactions contenant X et le nombre total de transactions.

Une règle d'association est un couple (A,B), où A et B sont des itemsets non vides disjoints, *i.e.* $\mathtt{A} \neq \emptyset$, $\mathtt{B} \neq \emptyset$, et $\mathtt{A} \cap \mathtt{B} = \emptyset$. On note classiquement un tel couple sous la forme $\mathtt{A} \rightarrow \mathtt{B}$. L'itemset A est appelé *prémisse* et B *conclusion*. On définit le support d'une règle d'association

comme étant le support de l'itemset A∪B (*i.e.* la proportion de transactions contenant à la fois A et B). On définit de plus la confiance d'une règle $\mathtt{A} \rightarrow \mathtt{B}$, notée $\text{CONF}(\mathtt{A} \rightarrow \mathtt{B})$, comme étant le rapport entre le support de l'itemset A∪B et celui de A : $\text{CONF}(\mathtt{A} \rightarrow \mathtt{B}) = \text{SUP}(\mathtt{A} \cup \mathtt{B})/\text{SUP}(\mathtt{A})$.

Etant donnés des seuils minimaux de support et de confiance σ_s et σ_c fixés au préalable, l'objectif des algorithmes de la famille APRIORI est d'extraire **toutes** les règles $\mathtt{A} \rightarrow \mathtt{B}$ vérifiant $\text{SUP}(\mathtt{A} \rightarrow \mathtt{B}) \geq \sigma_s$ et $\text{CONF}(\mathtt{A} \rightarrow \mathtt{B}) \geq \sigma_c$. Nous nous limiterons par la suite aux règles ne faisant intervenir qu'un seul item en conclusion, sans que cela ne nuise à notre application pratique (voir section 3). Il découle de cette exploration exhaustive une explosion combinatoire, le nombre de règles extrait devenant trop important pour permettre une évaluation individuelle de la pertinence de chaque règle par l'expert des données.

Plusieurs voies permettent de répondre à ce problème. Nous nous sommes intéressés à l'une d'elle : l'évaluation de la qualité des règles par des mesures objectives.

2.2 Mesures de qualité de règles d'association

Etant donné une règle $\mathtt{A} \rightarrow \mathtt{B}$, il est courant de l'analyser en se rapportant à une matrice de contingence croisant deux variables binaires A et B. On obtient une description de la règle $\mathtt{A} \rightarrow \mathtt{B}$ sous l'une des formes listées dans le tableau 1, où n_x correspond au nombre de transactions contenant X (ou fréquence absolue) et p_x à la fréquence observée de X (ou fréquence relative).

Par convention, on note multiplicativement l'union de deux itemsets afin d'alléger l'écriture (ainsi, le nombre de transactions contenant $X \cup Y$ est n_{xy}). Ces notations sont équivalentes, on passe de l'une à l'autre simplement via la relation $p_x = n_x/n$.

$A \backslash B$	0	1	total
0	$n_{\bar{a}\bar{b}}$	$n_{\bar{a}b}$	$n_{\bar{a}}$
1	$n_{a\bar{b}}$	n_{ab}	n_a
total	$n_{\bar{b}}$	n_b	n

$A \backslash B$	0	1	total
0	$p_{\bar{a}\bar{b}}$	$p_{\bar{a}b}$	$p_{\bar{a}}$
1	$p_{a\bar{b}}$	p_{ab}	p_a
total	$p_{\bar{b}}$	p_b	1

TAB. 1 – *Notations usuelles associées à une règle* $\mathtt{A} \rightarrow \mathtt{B}$

Afin de quantifier l'intérêt d'une règle, il est fréquent d'avoir recours à des mesures objectives. De telles mesures sont des fonctions définies sur la table de contingence présentée dans le tableau 1. Elles sont dites objectives (Freitas, 1999; Hilderman et Hamilton, 2000) par opposition aux mesures subjectives (Silberschatz et Tuzhilin, 1995; Liu et al., 2000; Brisson, 2004) qui, en plus d'être basées sur des comptages fréquentiels sur les données, prennent aussi en compte des connaissances spécifiques au domaine fouillé.

On peut recenser un très grand nombre de mesures de qualité dans la littérature (voir par exemple Yao et Zhong (1999); Guillet (2004) pour une liste de plus de quarante mesures de qualité). Nous nous sommes restreints dans ce travail à vingt mesures objectives, et strictement décroissantes en fonction de $n_{a\bar{b}}$. Selon nous, ces conditions sont des critères d'éligibilité pour des mesures de qualité de règles d'association (Lenca et al., 2003).

On montre dans Vaillant et al. (2004) que ces mesures mettent en avant différentes règles. Afin de guider le choix d'une mesure adaptée au contexte et aux souhaits de l'expert, nous avons étudié les vingt mesures selon neuf propriétés pertinentes dans le cadre de la fouille de

règles et faisant sens pour un expert métier (voir tableau 2, et Vaillant (2006) pour la sémantique détaillée et l'évaluation de ces propriétés sur les vingt mesures). Certaines propriétés sont considérées comme étant normatives (*i.e.* les préférences sur les modalités sont indépendantes du contexte applicatif), d'autres comme subjectives et dépendantes de préférences utilisateurs. Afin de mettre en œuvre un processus d'aide à la décision, il sera alors nécessaire que l'expert métier précise un ordre de préférence sur les modalités des propriétés subjectives.

Critère	Sémantique	Nombre de modalités	Responsabilité
g_1	non symétrie selon A et B	2	$E_{\mathcal{A}}$
g_2	décroissance avec n_b	2	$E_{\mathcal{A}}$
g_3	situation à la règle logique	2	$E_{\mathcal{A}}$
g_4	situation à l'indépendance	2	$E_{\mathcal{A}}$
g_5	situation à l'indétermination	2	$E_{\mathcal{A}}$
g_6	tolérance aux premiers contre-exemples	3	$E_{\mathcal{R}}$
g_7	prise en compte de n	2	$E_{\mathcal{R}}$
g_8	facilité à fixer un seuil	2	$E_{\mathcal{A}}$ et $E_{\mathcal{R}}$
g_9	intelligibilité	3	$E_{\mathcal{R}}$

TAB. 2 – *Propriétés de mesures de qualité*

3 Contexte applicatif

Nous appliquons un processus d'ECD à des données fournies par la société DIXID, concernant un service d'assistance. Plus précisément, un utilisateur ayant un problème avec une application professionnelle sur son ordinateur peut appeler le serveur vocal concerné. Le serveur vocal lui demande alors quel est son problème afin de l'orienter vers un service spécifique d'assistance, dont les opérateurs sont formés pour répondre au mieux aux problèmes posés (au total, 9 services d'assistance différents existent, avec plusieurs opérateurs pour chaque service). Dans l'idéal, l'utilisateur expose son problème en prononçant le nom de l'application concernée (mais bien sûr cette situation est "idéale") et il est orienté vers le "bon" opérateur. Un appel dure en moyenne près de 40 secondes (avec une amplitude de 1 à 141 secondes). Nos objectifs sont multiples. Nous souhaitons analyser le système mis en œuvre et en juger les performances. De plus, la société DIXID travaillant dans le domaine de l'ergonomie, nous souhaitons voir s'il est possible d'améliorer celle du système vocal. C'est en vue d'atteindre ce dernier objectif que nous nous sommes intéressés en première approche aux règles concluant sur le résultat des appels (voir tableaux 3 et 7).

Les données à notre disposition portent sur 2006 appels dans leur globalité, et les caractérisent selon 8 attributs discrets listés dans le tableau 3. Il n'y a pas de données manquantes. D'autres données, quantitatives, décrivant de façon précise les différentes interactions entre l'utilisateur et le service vocal n'ont pas été prises en compte pour le moment. Pour autant, malgré les difficultés liées à leur exploitation, l'expert des données pense que ces dernières sont riches d'information.

On peut d'ores et déjà remarquer que certaines variables sont obligatoirement liées (par exemple, un appel avec résultat "raccrochage après flash" aura évidemment une modalité "oui"

Variable	modalités (fréquence)	sémantique
NB_FEEDBACK_OK	0 (680), 1 (1184), 2 (123), 3 (17), 4 (2)	Nombre de prompts du serveur montrant une compréhension correcte de la demande de l'utilisateur
NB_FEEDBACK_NOK	0 (1675), 1 (251), 2 (57), 3 (15), 4 (4), 5 (3), 6 (1)	Nombre de prompts du serveur montrant une compréhension incorrecte de la demande de l'utilisateur
EXPOSITION_PB	aucune_precision (14), avec_un_nom (1760), ne_dit_rien (204), probleme_hors_sujet (3), sans_nom (25)	Manière d'exposer le problème.
REPETITION	oui (283), non (1723)	L'utilisateur répète mot à mot, au moins une fois, une demande.
REFORMULATION	oui (181), non (1825)	L'utilisateur explique un même problème en utilisant des termes différents, au moins une fois.
NON_REPONSE	0 (1792), 1 (158), 2 (54), 4 (2)	Nombre de non-réponse de l'utilisateur à une question explicite posée par le serveur.
DIFFUSION_FLASH	oui (41), non (1965)	Lorsque l'utilisateur expose son problème, un message "flash" concernant une certaine application est automatiquement diffusé.
RESULTAT_APPEL	bon_operateur (1369), erreur_d_operateur (244), operateur_par_defaut (160), raccrochage_apres_description (102), raccrochage_immediat (28), raccrochage_sans_description (42), raccrochage_sans_rien_dire (39), raccrochage_apres_flash (22)	Résultat de l'appel.

TAB. 3 – *Attributs décrivant les données* DIXID

pour l'attribut DIFFUSION_FLASH...). De plus, certaines modalités de plusieurs attributs sont très peu fréquentes. Afin d'ignorer le moins d'information possible, nous avons ainsi fixé le seuil de support le plus bas possible, soit 1% avec l'implémentation d'APRIORI que nous utilisons. Nous fixons le seuil de confiance à 50%, afin de générer des règles ayant plus d'exemples que de contre-exemples. Enfin, nous avons précisé que les règles pouvaient faire intervenir entre 2 et 8 attributs, l'implémentation utilisée se limitant par défaut à une taille maximale de 5. Cette extension volontaire des règles à la plus grande taille possible est un choix en première approche, visant à vérifier qu'il n'existe pas de règle de grande taille porteuse d'information. La redondance entre règles apportée par ce choix, ainsi que la difficulté d'interprétation des règles longues incitera à revenir au paramètre par défaut d'APRIORI dans les études futures.

Avec ces paramètres, APRIORI génère 8921 règles, sur la base des 2006 objets. Si l'on se restreint aux règles faisant intervenir l'attribut RESULTAT_APPEL, qui est celui qui nous intéresse si l'on souhaite analyser l'ergonomie du serveur vocal, il reste 6061 règles à évaluer.

En étant encore plus strict et en ne considérant que les règles faisant intervenir cet attribut en conclusion, on retient 626 règles, ce qui est toujours trop volumineux pour une étude manuelle exhaustive. Les 626 règles retenues par la suite peuvent ainsi être vues comme un résultat d'un apprentissage supervisé sur les différentes modalités de l'attribut RESULTAT_APPEL.

Comme mentionné précédemment, certaines des règles générées sont sans grand intérêt, comme `DIFFUSION_FLASH='oui' -> RESULTAT_APPEL='raccrochage_apres_flash'` (support : 1.1%, confiance : 53.7%).

4 Choix de mesure de qualité

Afin de guider l'analyse de l'ensemble de règles généré, nous procédons à un filtrage par des mesures de qualité. En se basant sur la prise en compte des préférences de l'utilisateur expert des données, nous sélectionnons trois mesures parmi les vingt étudiées, et retenons les 20 règles les mieux classées par chacune d'entre elles.

La modélisation des préférences du décideur comporte deux aspects. Dans un premier temps, il est nécessaire de préciser l'ordre dans lequel les modalités des propriétés subjectives sont à considérer, afin de transformer ces propriétés en critères de choix. Puis, chaque critère est affecté d'un poids relatif à son importance, telle que exprimée par les différents intervenants. Un troisième aspect pourrait être considéré dans la méthode d'aide multicritère à la décision que nous utilisons, mais il ne fait pas sens dans notre cas : le choix d'une fonction de préférence (à valeurs dans $[0, 1]$), permettant de rendre commensurable les différents critères. N'ayant que des critères à deux ou trois modalités, nous avons retenu le modèle usuel de préférence.

Après explicitation des propriétés, l'expert des données a ordonné les modalités des propriétés contextuelles comme suit, exprimant ainsi ses préférences sur chaque critère pris individuellement :

g_6 N'étant pas dans une situation où les contre-exemples sont critiques, on préférera une décroissance linéaire de la mesure en fonction des contre-exemples autour de 0^+, puis une décroissance faible, et enfin une décroissance forte.

g_7 La prise en compte d'une modélisation statistique étant souhaitée, les mesures construites par cette approche seront préférées aux mesures descriptives.

g_8 On peut souhaiter opérer une validation statistique des règles retenues par la mesure. Une mesure se prêtant à un tel calcul sera donc préférée à une autre ne s'y prêtant pas.

g_9 L'interprétation de la valeur prise par la mesure ayant de l'importance afin de communiquer les résultats, plus la mesure est simple à comprendre plus elle est préférée.

Pour déterminer les poids de chaque critère, nous avons exploré deux scénarios. Le premier (SC1) consiste à ne reposer que sur l'avis de l'expert des données pour le fixer. Cette approche n'a pas abouti à des résultats satisfaisants à notre avis, mais permet de confirmer une assertion courante. Elle illustre aussi les difficultés liées à l'expression de besoins de l'utilisateur, particulièrement lorsque ce dernier n'est pas expert de ECD. Le second (SC2) consiste à prendre en compte l'avis de l'expert des données sur les propriétés subjectives, et à fixer les poids des propriétés normatives selon l'avis de l'expert de la fouille de données. Les jeux de poids ainsi déterminés sont listés dans le tableau 4. L'importance de chaque critère est évaluée sur une échelle entière entre 0 (sans importance) et 10 (forte importance).

	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9
SC1	8	6	2	0	7	4	2	7	3
SC2	9	7	6	6	5	4	2	7	3

TAB. 4 – *Jeux de poids utilisés pour sélectionner les mesures de qualité*

5 Résultats

A partir de ces préférences exprimées, nous avons utilisé la méthode d'aide multicritère PROMETHEE (Brans et al., 1984). Cette méthode de surclassement permet de constituer un flot agrégé sur les mesures afin de prendre en compte les préférences utilisateur. A partir de ce flot de surclassement, on peut alors disposer d'un ordre sur les mesures, de celles répondant au mieux aux souhaits exprimés à celles y répondant le moins. Le tableau 5 présente les trois meilleures mesures pour chaque jeu de poids, ainsi que les deux suivantes. La méthode PROMETHEE permet également de procéder à une étude de stabilité du jeu de poids. Le tableau 6 donne l'intervalle dans lequel on peut faire varier le poids associé à chaque critère pris individuellement, sans que cela ait un impact sur les trois premières mesures du rangement.

	rang 1	rang 2	rang 3	rang 4	rang 5	...
SC1 :	CONF (0.24)	LOE (0.17)	CONFCEN (0.16)	MOCO (0.14)	TEC (0.10)	
SC2 :	LOE (0.24)	FB (0.17)	INTIMP (0.16)	CONFCEN (0.15)	CONV (0.13)	

TAB. 5 – *Rangements et flots de préférence agrégés des premières mesures*

	g_1	g_2	g_3	g_4	g_5
SC1 :	$[0, +\infty]$	$[5.2, 8.8]$	$[1.8, 4.2]$	$[0, 2.8]$	$[4.2, 7.8]$
SC2 :	$[0, +\infty]$	$[4.45, +\infty]$	$[5.45, +\infty]$	$[3.45, +\infty]$	$[0, 7]$

	g_6	g_7	g_8	g_9
SC1 :	$[1.5385, +\infty]$	$[0, 5.45]$	$[3.55, +\infty]$	$[1.6667, 3.3333]$
SC2 :	$[1.3846, 4.5556]$	$[1.45, 2.25]$	$[5, +\infty]$	$[2.8148, 3.4074]$

TAB. 6 – *Stabilité des jeux de poids*

Le fait que la confiance soit classée première pour SC1 peut paraître surprenant, ou décevant. En effet, cette mesure ne prenant en compte que deux grandeurs et étant un filtre utilisé par APRIORI, on peut s'attendre à ne pas obtenir de résultats satisfaisants. Toutefois, ce résultat confirme une assertion courante : la confiance est très souvent utilisée en pratique pour sa simplicité et son intelligibilité. LOE est très bien classée dans les deux scénarios, elle réalise un bon compromis, comme déjà remarqué dans Lenca et al. (2007).

Lors de la sélection des 20 meilleures règles, CONF $= \frac{n_{ab}}{n_a}$ et LOE $= 1 - \frac{n n_{a\bar{b}}}{n_a n_{\bar{b}}}$ en sélectionnent effectivement 20, les autres mesures en retenant plus, à cause d'ex-aequo : 32 règles

sont retenues par CONFCEN $= \frac{nn_{ab}-n_an_b}{nn_a}$ et FB $= \frac{n_{ab}n_{\bar{b}}}{n_bn_{a\bar{b}}}$ (ce sont de plus les mêmes, les différences d'évaluation des règles par ces mesures n'apparaissent qu'en suite), et 98 règles sont retenues par INTIMP $= P\left[N(0,1) \geq \frac{n_an_b-nn_{ab}}{\sqrt{n}p_ap_{\bar{b}}}\right]$. Dans le cas de cette dernière, on observe la non-discriminance de la mesure. Il est à noter que de très nombreuses règles retenues sont des spécialisations d'autres règles. Ceci est dû au fait que nous ayons "forcé" l'algorithme à explorer totalement l'espace de recherche. En se limitant à 5 items par règle, un nombre important de telles règles ne seraient pas apparues.

Les seules modalités de conclusion retenues par nos mesures sont `bon_operateur` et `operateur_par_defaut`, ce qui ne nous permet donc de n'évaluer que les points positifs du système. En étudiant les règles non-supervisées, CONF, LOE et FB évaluent 2523 règles au maximum. Ce sont les règles n'ayant pas de contre-exemples. En revanche INTIMP ne sélectionne "que" 434 règles, qui ne sont pas nécessairement logiques ce qui cette fois illustre l'avantage d'avoir recours à un modèle statistique.

6 Conclusions et perspectives

Nous avons illustré dans cet article la mise en pratique d'une méthode d'aide multicritère à la décision afin de guider le choix d'une mesure de qualité en vue de quantifier la qualité de règles d'association, selon des préférences exprimées par des utilisateurs experts. Les résultats produits attestent de la difficulté de cette tâche, tant dans le choix des paramètres de l'algorithme de fouille de données que dans la retranscription du système de préférence expert.

Les règles sélectionnées par les mesures ont été soumises à l'expert des données afin d'affiner les paramètres choisis dans cette première approche. Il est apparu que certaines des mesures retenues n'étaient pas assez discriminantes, alors que d'autres retenaient le même ensemble de règles. Ces aspects expérimentaux pourraient être inclus au processus de sélection de mesure. En outre, des informations plus précises relatives aux différents appels sont à notre disposition. L'exploitation de ces données pourrait permettre de mieux analyser les appels. On se heurte toutefois à des problèmes de temporalité ou de causalité liés à la nature même de cette information.

Références

Agrawal, R. et R. Srikant (1994). Fast algorithms for mining association rules. In J. B. Bocca, M. Jarke, et C. Zaniolo (Eds.), *Proceedings of the 20th Very Large Data Bases Conference*, pp. 487–499. Morgan Kaufmann.

Borgelt, C. et R. Kruse (2002). Induction of association rules : APRIORI implementation. In *Proceedings of the 15th Conference on Computational Statistics*, Heidelberg, Germany. Physika Verlag.

Brans, J.-P., B. Mareschal, et P. Vincke (1984). PROMETHEE *: A New Family of Outranking Methods in MCDM*. IFORS 84.

Brisson, L. (2004). Mesures d'intérêt subjectif et représentation des connaissances. Technical Report ISRN I3S/RR-2005-35-FR, Université de Nice.

Carvalho, D., A. A. Freitas, et N. Ebecken (2005). Evaluating the correlation between objective rule interestingness measures and real human interest. In *Knowledge Discovery in Databases : Proceedings of PKDD 2005, LNAI 3731*, pp. 453–461. Springer Verlag.

Fayyad, U., G. Piatetsky-Shapiro, P. Smyth, et R. Uthurusamy (Eds.) (1996). *Advances in Knowledge Discovery and Data Mining*. AAAI/MIT Press.

Freitas, A. (1999). On rule interestingness measures. *Knowledge-Based Systems journal*, 309–315.

Guillet, F. (2004). Mesure de la qualité des connaissances en ECD. Tutoriel de la 4e Conf. Extraction et Gestion des Connaissances. 60 pages.

Hilderman, R. J. et H. J. Hamilton (2000). Applying objective interestingness measures in data mining systems. In *Proceedings of the 4th European Conference on Principles of Data Mining and Knowledge Discovery (PKDD'00)*, pp. 432–439. Springer-Verlag.

Hilderman, R. J. et H. J. Hamilton (2001). *Knowledge Discovery and Measures of Interest*, Volume 638 of *The International Series in Engineering and Computer Science*. Kluwer.

Lenca, P., P. Meyer, P. Picouet, B. Vaillant, et S. Lallich (2003). Critères d'évaluation des mesures de qualité en ECD. *Revue des Nouvelles Technologies de l'Information (Entreposage et Fouille de données)* (1), 123–134.

Lenca, P., P. Meyer, B. Vaillant, et S. Lallich (2007). On selecting interestingness measures for association rules : user oriented description and multiple criteria decision aid. *European Journal of Operational Research*, To appear.

Liu, B., W. Hsu, S. Chen, et Y. Ma (2000). Analyzing the subjective interestingness of association rules. *IEEE Intelligent Systems 15*(5), 47–55.

Silberschatz, A. et A. Tuzhilin (1995). On subjective measures of interestingness in knowledge discovery. In *Knowledge Discovery and Data Mining*, pp. 275–281.

Tan, P.-N., V. Kumar, et J. Srivastava (2002). Selecting the right interestingness measure for association patterns. In *Proceedings of the Eighth ACM SIGKDD International Conference on KDD*, pp. 32–41.

Vaillant, B. (2006). *Mesurer la qualité des règles d'association : études formelles et expérimentales*. Ph. D. thesis, École nationale supérieure de télécommunications de Bretagne/Université de Bretagne Sud.

Vaillant, B., P. Lenca, et S. Lallich (2004). A clustering of interestingness measures. In *Discovery Science*, pp. 290–297.

Yao, Y. Y. et N. Zhong (1999). An analysis of quantitative measures associated with rules. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pp. 479–488.

Summary

Knowledge discovery in databases aims at extracting information from large datasets through the use of automated tools, in order to re-use this information in industrial or decision oriented applications. In this paper, we explore a datamining process applied to the use of a vocal answering machine. We particularly focus on the crucial knowledge validation step, relying in our approach on the use of interestingness measures.

TAB. 7 – *Ensemble des* 20 *meilleures règles pour* CONF *(20 règles)*

id	prémisse	conclusion	CONF	n_a	n_b	n_{ab}	$n_{\bar{a}b}$
451	NB_FEEDBACK_OK=2 and NB_FEEDBACK_NOK=0 and REPETITION='non' and EXPOSITION_PB='avec_un_nom' and DIFFUSION_FLASH='non'	RESULTAT_APPEL='bon_operateur'	0.98611	72	1369	71	1
264	NB_FEEDBACK_OK=2 and NB_FEEDBACK_NOK=0 and REPETITION='non' and DIFFUSION_FLASH='non'	RESULTAT_APPEL='bon_operateur'	0.98611	72	1369	71	1
261	NB_FEEDBACK_OK=2 and NB_FEEDBACK_NOK=0 and REPETITION='non' and EXPOSITION_PB='avec_un_nom'	RESULTAT_APPEL='bon_operateur'	0.98611	72	1369	71	1
106	NB_FEEDBACK_OK=2 and NB_FEEDBACK_NOK=0 and REPETITION='non'	RESULTAT_APPEL='bon_operateur'	0.98611	72	1369	71	1
577	NB_FEEDBACK_OK=2 and NB_FEEDBACK_NOK=0 and REPETITION='non' and EXPOSITION_PB='avec_un_nom' and NON_REPONSE=0 and DIFFUSION_FLASH='non'	RESULTAT_APPEL='bon_operateur'	0.98551	69	1369	68	1
453	NB_FEEDBACK_OK=2 and NB_FEEDBACK_NOK=0 and REPETITION='non' and NON_REPONSE=0 and DIFFUSION_FLASH='non'	RESULTAT_APPEL='bon_operateur'	0.98551	69	1369	68	1
449	NB_FEEDBACK_OK=2 and NB_FEEDBACK_NOK=0 and REPETITION='non' and EXPOSITION_PB='avec_un_nom' and NON_REPONSE=0	RESULTAT_APPEL='bon_operateur'	0.98551	69	1369	68	1
262	NB_FEEDBACK_OK=2 and NB_FEEDBACK_NOK=0 and REPETITION='non' and NON_REPONSE=0	RESULTAT_APPEL='bon_operateur'	0.98551	69	1369	68	1
578	NB_FEEDBACK_OK=2 and NB_FEEDBACK_NOK=0 and REPETITION='non' and EXPOSITION_PB='avec_un_nom' and REFORMULATION='non' and DIFFUSION_FLASH='non'	RESULTAT_APPEL='bon_operateur'	0.98333	60	1369	59	1
454	NB_FEEDBACK_OK=2 and NB_FEEDBACK_NOK=0 and REPETITION='non' and REFORMULATION='non' and DIFFUSION_FLASH='non'	RESULTAT_APPEL='bon_operateur'	0.98333	60	1369	59	1
450	NB_FEEDBACK_OK=2 and NB_FEEDBACK_NOK=0 and REPETITION='non' and EXPOSITION_PB='avec_un_nom' and REFORMULATION='non'	RESULTAT_APPEL='bon_operateur'	0.98333	60	1369	59	1
263	NB_FEEDBACK_OK=2 and NB_FEEDBACK_NOK=0 and REPETITION='non' and REFORMULATION='non'	RESULTAT_APPEL='bon_operateur'	0.98333	60	1369	59	1
620	NB_FEEDBACK_OK=2 and NB_FEEDBACK_NOK=0 and REPETITION='non' and EXPOSITION_PB='avec_un_nom' and NON_REPONSE=0 and REFORMULATION='non' and DIFFUSION_FLASH='non'	RESULTAT_APPEL='bon_operateur'	0.98276	58	1369	57	1
579	NB_FEEDBACK_OK=2 and NB_FEEDBACK_NOK=0 and REPETITION='non' and NON_REPONSE=0 and REFORMULATION='non' and DIFFUSION_FLASH='non'	RESULTAT_APPEL='bon_operateur'	0.98276	58	1369	57	1
576	NB_FEEDBACK_OK=2 and NB_FEEDBACK_NOK=0 and REPETITION='non' and EXPOSITION_PB='avec_un_nom' and NON_REPONSE=0 and REFORMULATION='non'	RESULTAT_APPEL='bon_operateur'	0.98276	58	1369	57	1
452	NB_FEEDBACK_OK=2 and NB_FEEDBACK_NOK=0 and REPETITION='non' and NON_REPONSE=0 and REFORMULATION='non'	RESULTAT_APPEL='bon_operateur'	0.98276	58	1369	57	1
267	NB_FEEDBACK_OK=2 and NB_FEEDBACK_NOK=0 and EXPOSITION_PB='avec_un_nom' and DIFFUSION_FLASH='non'	RESULTAT_APPEL='bon_operateur'	0.97030	101	1369	98	3
110	NB_FEEDBACK_OK=2 and NB_FEEDBACK_NOK=0 and DIFFUSION_FLASH='non'	RESULTAT_APPEL='bon_operateur'	0.97030	101	1369	98	3
107	NB_FEEDBACK_OK=2 and NB_FEEDBACK_NOK=0 and EXPOSITION_PB='avec_un_nom'	RESULTAT_APPEL='bon_operateur'	0.97030	101	1369	98	3
28	NB_FEEDBACK_OK=2 and NB_FEEDBACK_NOK=0	RESULTAT_APPEL='bon_operateur'	0.97030	101	1369	98	3