Fouille de données

▷ Mesures d'intérêt pour les règles d'association

Philippe Lenca & Stéphane Lallich

philippe.lenca@telecom-bretagne.eu

Telecom Bretagne & Université Lyon

2016-2017

Outline

Outline

Basic Definitions [PCKW89], [FPSSU96], [KNZ01]...

KDD is a non-trivial (decision aid interactive and iterative) process where user(s) seek to identify valid, novel, potentially useful, and ultimately understandable **patterns in data**.

KDD must be considered as a process of contextualization : in practice exact definitions of all concepts are required.

$\hookrightarrow$ Our focus is interestingness (valid, novel, useful, understandable).

Classical troubles...

- [BMUT97] found $6,732$ rules with a maximum conviction value (obvious) in Census data : five years old don't work, men don't give birth, etc.
- [Tsu00] found $29,050$ rules, out of which only 220 (less than 1%) were considered interesting or surprising by the user
- [WL00] found rules with 40-60% confidence (i.e. low confidence) which were considered novel and more accurate than some doctor's domain knowledge

$\hookrightarrow$ Qualitative and quantitative troubles.

Retaining the right rules

How can we help the end-user to select the good patterns (from his point of view) ?

Some common approaches :

- measurement :
  - objective interestingness measures [Fre99], [HH00]
  - subjective interestingness measures [ST95]
- redundancy analysis [LGB98]
- visualization tools & human centered processes [BGB03], [DP03], [MBY10]
- domain-driven [Cao10]

↪ Our subject : objective interestingness measures.

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Interestingness is perhaps a broad concept [GH06]

- validity (on new data with some degree of certainty)
- novelty (at least to the system and preferably to the user)
- utility (that is, lead to some benefit to the user or task)
- understandability (if not immediately then after some post-processing)
- but also conciseness, coverage, peculiarity, diversity, surprisingness, and actionability.

$\hookrightarrow$ Can we define quantitative measures for evaluating extracted patterns ?

Evaluating extracted patterns

Is it easy to define such quantitative measures ?

- validity : measures of certainty, robusteness
  - estimated prediction accuracy, confidence on new data
- utility : gain
  - in money saved because of better predictions
  - speedup in response time
- novelty, surprising : more subjective
  - if the pattern contradicts a user expectation
  - with a stochastic model
- understandability : more subjective
  - estimated by simplicity (size of the pattern)

$\hookrightarrow$ Objective measures for rule-based patterns.

Discovering comprehensible/understandable knowledge

Production rules based models

if conditions then conclusion

Popular models

- association rules [AS94] (unsupervised paradigm)
- class association rules [LHM98] (supervised paradigm)
- decision trees [BFSO84] (supervised paradigm)

↪ Objective interestingness measures for such models.

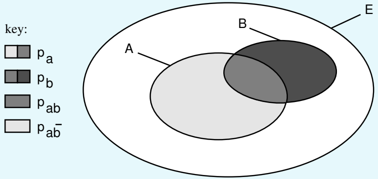
Outline

Objective interestingness measures

Definition

An objective interestingness measure (OIM) is a function from the space of rules $\{\mathtt{A} \rightarrow \mathtt{B}\}$ to the space of extended real numbers $(\mathbb{R} \cup \{-\infty, +\infty\})$.
OIM are based on the rules cardinalities (data-driven) :

A\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



$\hookrightarrow$ Usually functions $\mu(n, p_a, p_b, p_{ab})$ or $\mu(n, p_a, p_b, p_{a\bar{b}})$.

Objective interestingness measures

OIM provide numerical information on the quality of a rule

A rule $\mathtt{A} \rightarrow \mathtt{B}$ is said "of quality" when evaluated by an interestingness measure μ, if its evaluation by μ is greater than a user defined threshold σ_μ.

$$\mu(\mathtt{A} \rightarrow \mathtt{B}) \geq \sigma_\mu \Longleftrightarrow \mathtt{A} \rightarrow \mathtt{B} \text{ is of quality}$$

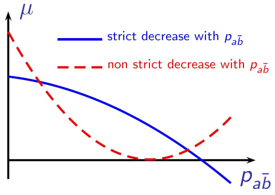
where σ_μ has to be fixed by the user.

$\hookrightarrow$ Measures are used to rank, to filter the rules i.e. to select most pertinent rules.

Objective interestingness measures

Eligibility property

Common assertion : the fewer counterexamples (A true and B false) to the rule there are, the higher the interestingness of the rule is.
Focus on decreasing measures wrt. $p_{a\bar{b}}$, all marginal frequencies being fixed.



$\hookrightarrow$ Measures like χ^2, Pearson's r^2, J-measure or Pearl's measure are excluded. . .

A huge number of measures...

	Absolute definitions	Relative definitions
BF	$\frac{n_{ab} n_{\bar{b}}}{n_b n_{a\bar{b}}}$	$\frac{p_{b/a}/p_{\bar{b}/a}}{p_b/p_{\bar{b}}} = \frac{p_{a/b}}{p_{a/\bar{b}}}$
CONFCEN	$\frac{n_{ab}}{n_a} - \frac{n_b}{n}$	$p_{b/a} - p_b$
CONF	$\frac{n_{ab}}{n_a}$	$p_{b/a}$
CONV	$\frac{n_a n_{\bar{b}}}{n n_{a\bar{b}}}$	$\frac{p_a p_{\bar{b}}}{p_{a\bar{b}}}$
TEC	$\frac{n_{ab} - n_{a\bar{b}}}{n_{ab}}$	$1 - \frac{p_{a\bar{b}}}{p_{ab}}$
GI	$\log(\frac{n n_{ab}}{n_a n_b})$	$\log \frac{p_{ab}}{p_a p_b}$
-INDIMP	$\frac{n_a n_b - n n_{ab}}{\sqrt{n n_a n_{\bar{b}}}}$	$\sqrt{n}\frac{p_{a\bar{b}} - p_a p_{\bar{b}}}{\sqrt{p_a p_{\bar{b}}}}$
INTIMP	$P\left[Poisson\left(\frac{n_a n_{\bar{b}}}{n}\right) \geq n_{a\bar{b}}\right]$	
IQC	$2\frac{n n_{ab} - n_a n_b}{n n_a + n n_b - 2 n_a n_b}$	$2\frac{p_{ab} - p_a p_b}{p_a + p_b - 2 p_a p_b}$
MOCO	$\frac{n_{ab} - n_{a\bar{b}}}{n n_b}$	$\frac{p_{ab} - p_{a\bar{b}}}{p_b}$
...	...	...

A huge number of measures. . .

	Absolute definitions	Relative definitions
...	...	...
LIFT	$\frac{nn_{ab}}{n_a n_b}$	$\frac{p_{b/a}}{p_b}$
LOE	$\frac{nn_{ab}-n_a n_b}{n_a n_{\bar{b}}}$	$\frac{p_{b/a}-p_b}{p_{\bar{b}}}$
IPD	$P\left[\mathcal{N}(0,1) > \textsc{IndImp}^{CR/B}\right]$	
PS	$n_{ab} - \frac{n_a n_b}{n}$	$n(p_{ab} - p_a p_b)$
SEB	$\frac{n_{ab}}{n_{a\bar{b}}}$	$\frac{p_{ab}}{p_{a\bar{b}}}$
SUP	$\frac{n_{ab}}{n}$	p_{ab}
TIIE	$[i(\mathtt{A} \to \mathtt{B}) \times \textsc{IntImp}(\mathtt{A} \to \mathtt{B})]^{1/2}$	
ZHANG	$\frac{nn_{ab}-n_a n_b}{\max\{n_{ab}n_{\bar{b}}, n_b n_{a\bar{b}}\}}$	$\frac{p_{ab}-p_a p_b}{\max\{p_{ab}p_{\bar{b}}; p_b(p_a-p_{ab})\}}$
...	...	...

Different measures. . . different rankings

measure	rank 1	rank 2	rank 3	rank 4	rank 5	rank 6	rank 7	rank 8	rank 9	rank 10	rank 11
LIFT	r_{18}	r_1 r_2 r_{17}	r_{21}	r_6 r_7	r_3 r_{16} r_{20}	r_{12} r_{13}	r_4 r_8 r_9 r_{10} r_{11}	r_{19}	r_5 r_{14} r_{15}		
CONF	r_1 r_3 r_{16} r_{17}	r_6 r_8 r_{11}	r_{18} r_{19} r_{20} r_{21}	r_5 r_7 r_{12} r_{13} r_{14} r_{15}	r_2 r_4 r_9 r_{10}						
SUP	r_5 r_6 r_7 r_8 r_{11} r_{12} r_{13} r_{14} r_{15}	r_1 r_2 r_3 r_4 r_9 r_{10} r_{16} r_{17} r_{18} r_{19} r_{20} r_{21}									
INTIMP	r_1 r_{17}	r_{18}	r_2	r_3 r_{16}	r_6 r_{21}	r_7	r_{20}	r_{12} r_{13}	r_4 r_8 r_9 r_{10} r_{11}	r_{19}	r_5 r_{14} r_{15}
CONV	r_1 r_3 r_{16} r_{17}	r_{18}	r_6 r_{21}	r_2	r_{20}	r_7	r_{12} r_{13}	r_4 r_8 r_9 r_{10} r_{11}	r_{19}	r_5 r_{14} r_{15}	
GI	r_{18}	r_1 r_2 r_{17}	r_{21}	r_6 r_7	r_3 r_{16} r_{20}	r_{12} r_{13}	r_4 r_8 r_9 r_{10} r_{11}	r_{19}	r_5 r_{14} r_{15}		
LAP	r_1 r_3 r_{16} r_{17}	r_6 r_8 r_{11}	r_{18} r_{19} r_{20} r_{21}	r_5 r_7 r_{12} r_{13} r_{14} r_{15}	r_2 r_4 r_9 r_{10}						
PS	r_{18}	r_1 r_2 r_{17}	r_6 r_7 r_{21}	r_3 r_{16} r_{20}	r_{12} r_{13}	r_4 r_8 r_9 r_{10} r_{11}	r_{19}	r_5 r_{14} r_{15}			
SEB	r_1 r_3 r_{16} r_{17}	r_6 r_8 r_{11}	r_{18} r_{19} r_{20} r_{21}	r_5 r_7 r_{12} r_{13} r_{14} r_{15}	r_2 r_4 r_9 r_{10}						
ZHANG	r_1 r_3 r_{16} r_{17}	r_{18}	r_2	r_{21}	r_6	r_7 r_{20}	r_{12} r_{13}	r_4 r_8 r_9 r_{10} r_{11}	r_{19}	r_5 r_{14} r_{15}	

$\hookrightarrow$ r_{18} seems to be "good", r_{19} seems to be "bad", but r_2 ?

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Outline

Does the measure help to distinguish $\mathtt{A} \rightarrow \mathtt{B}$ and $\mathtt{B} \rightarrow \mathtt{A}$ [Fre99] ?

Asymmetric processing of $\mathtt{A}$ and $\mathtt{B}$

Since the antecedent and the consequent of a rule may have very different significations, it is desirable to make a distinction between measures that evaluate rules $\mathtt{A} \rightarrow \mathtt{B}$ differently from rules $\mathtt{B} \rightarrow \mathtt{A}$ and those which do not.

Sex\coat	Non red	red	total
H	48	2	50
F	32	18	50
total	80	20	100

`if sex=F then red coat`
($\text{SUP} = 0.18$ $\text{CONF} = 0.36$ $\text{LIFT} = 1.8$)

`if red coat then sex=F`
($\text{SUP} = 0.18$ $\text{CONF} = 0.90$ $\text{LIFT} = 1.8$)

where $\text{SUP}(\mathtt{A} \rightarrow \mathtt{B}) = p_{ab}$, $\text{CONF}(\mathtt{A} \rightarrow \mathtt{B}) = p_{ab}/p_a$, $\text{LIFT}(\mathtt{A} \rightarrow \mathtt{B}) = \frac{p_{ab}}{p_a p_b}$

Does the measure penalizes large B ? [PS91]

Decrease with p_b

The interest of A $\rightarrow$ B should be decreasing with the size of B when p_{ab} and p_a are given.

beer $\rightarrow$ bread,
washing machine $\rightarrow$ bread.

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Does the measure help recognizing independence ? [PS91]

A\B	0	1	
0	2	18	20
1	8	72	80
	10	90	100

$$\textsc{Sup}(\mathtt{A} \rightarrow \mathtt{B}) = p_{ab} \quad = 0.72$$
$$\textsc{Conf}(\mathtt{A} \rightarrow \mathtt{B}) = p_{b/a} \quad = 0.9$$
$$\text{but } p_b = 0.9$$

Situation at the independence

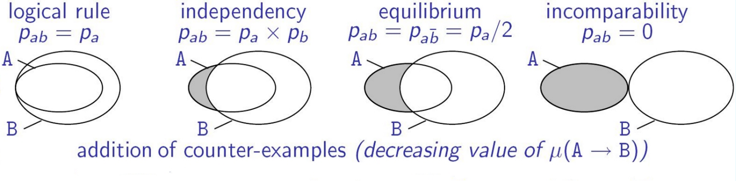
- [PS91] proposed that $\mu(\mathtt{A} \rightarrow \mathtt{B}) = 0$
  $\mathrm{PS} = n(p_{ab} - p_a p_b)$
- [LMVL08] proposed that $\mu(\mathtt{A} \rightarrow \mathtt{B}) = C$
  $\textsc{Lift} = \frac{p_{b/a}}{p_b} = \frac{n n_{ab}}{n_a n_b} = \frac{\textsc{Conf}}{p_b} = \frac{p_{ab}}{p_a p_b}$ (which values 1)

$\hookrightarrow$ It is a first steep but not sufficient.

Does the measure help recognizing reference situations ?

Reference situations

Remember : $\mu(\mathtt{A} \rightarrow \mathtt{B}) \geq \sigma_\mu$ $\mathtt{A} \rightarrow \mathtt{B}$ is of quality, how to choose σ_μ ?
From value for a logical rule to incompatibility [LMVL08] through independency [PS91] or indetermination [BGBG05]. . ., not too high (harsh selective), not too low (weak selective).



$\hookrightarrow$ Take into account some reference value wrt. the user's goal : targeting or prediction, or any other threshold [LVL07] ?

Value at reference situations

For 3 classical measures :

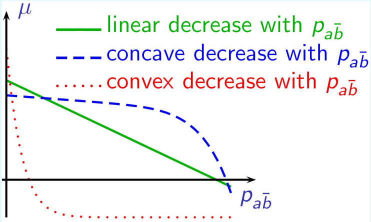
- $\textsc{Sup}(\mathtt{A} \rightarrow \mathtt{B}) = p_{ab}$
- $\textsc{Conf}(\mathtt{A} \rightarrow \mathtt{B}) = p_{ab}/p_a$
- $\textsc{Lift}(\mathtt{A} \rightarrow \mathtt{B}) = \frac{p_{ab}}{p_a p_b}$

	logical rule $p_{a\bar{b}} = 0$	independency $p_{ab} = p_a \times p_b$	equilibrium $p_{ab} = p_a/2$
Sup	p_{ab}	$p_a p_b$	$p_a/2$
Conf	1	p_b	$1/2$
Lift	$1/p_b$	1	$\frac{1}{2p_b}$

Tolerance to counter-examples [GKCG01]

Behavior with $p_{a\bar{b}}$ around 0^+

The user may tolerate a few counter-examples without significant loss of interest. However, the opposite choice could also be preferred to increase the sensitivity to a false positive.



↪ Robustness issue and origin of the counter-examples ? [LLV06].

Statistical or descriptive measures [Lal02]

Sensitivity to n (total number of records)

The user can prefer to have a measure which is invariant (descriptive measure) or not with the dilatation of data (statistical measures).

A statistical measures will increases with n (for constant rates of $\mathtt{A}$, $\mathtt{A} \rightarrow \mathtt{B}$, $\mathtt{B}$).

$\hookrightarrow$ Discrimination power issue [Lal02].

Statistical measures and increasing n [Lal02]

Discrimination power

If a measure depends of n, it loses its discriminating power.

SOLARFLARE-1	IntImp	IPD	IIE
0 to 0.95	3629	5214	4019
0. 95 to 0.99	1011	145	1072
0.99 to 1	762	43	311
Total	5402	5402	5402

Solutions :

- weighting by a discriminating index (IIE, [GKCG01])
- contextual functional transformation (IPD, [LA03])

$\hookrightarrow$ These two solutions have been generalized to be applied whatever the reference threshold is for the confidence [LVL07].

How to fix a threshold ? [LMVL08]

Easiness to fix a threshold

Even if references situations have been taken into account, it is still difficult to decide on the best threshold value that separates interesting from uninteresting rules.

To establish this property, we can provide a sense of the strength of the evidence against the null hypothesis H_0 (absence of a link between `A` and `B`), that is the p-value [LT04] i.e. the threshold σ_μ should be the value exceeded 5 times out of 100 by the measure μ in case of independence.

Does the semantics of the measure make sense ? [LMVL08]

Intelligibility criteria

- $\text{TIIE} = \{[(1 - H^*(\mathtt{B}/\mathtt{A})^2)(1 - H^*(\bar{\mathtt{A}}/\bar{\mathtt{B}})^2)]^{1/4}\textsc{IntImp}\}^{1/2}$ where $H^*(X/Y) = 1$, if $p_{x/y} > \max\{0.5;\ p_x\}$, $H^*(X/Y) = -p_{x/y}\log_2 p_{x/y} - (1 - p_{x/y})\log_2(1 - p_{x/y})$ otherwise
- $\textsc{Conf} = p_{b/a} = \frac{p_{ab}}{p_a}$

What does it means when the measure values 0.7, when it changes from 0.7 to 0.72 ?

- with $\textsc{Conf}$ the user may understand that the rule is now true in 72% of cases (or that the rule is not true for 28 % of cases) and that he gains 2 %
- but with TIIE ?

↪ Importance of intelligibility/comprehensibility depends on the user/application/domain/context.

Does the measure help recognizing robust rules ? [LLV06]

$(\mathtt{r}_1)$ A\B	0	1	total
0	0.2	0.2	0.4
1	0.2	0.4	0.6
total	0.4	0.6	1

$(\mathtt{r}_2)$ A\B	0	1	total
0	0.18	0.32	0.5
1	0.1	0.4	0.5
total	0.28	0.72	1

	Sup	Conf	Lift
$\mathtt{r}_1$	0.40	0.66	1.11
$\mathtt{r}_2$	0.40	0.80	1.11

Is $\mathtt{r}_2$ better than $\mathtt{r}_1$?

Does the measure help recognizing robust rules ?

(r_1) A\B	0	1	total
0	0.2	0.2	0.4
1	0.2	0.4	0.6
total	0.4	0.6	1

	SUP	CONF	LIFT
r_1	0.40	0.66	1.11
r_2	0.40	0.80	1.11

(r_2) A\B	0	1	total
0	0.18	0.32	0.5
1	0.1	0.4	0.5
total	0.28	0.72	1

Robustness criteria [LLV06]

- although having a lower confidence, r_1 may lose 25% of its examples while r_2 may lose only 20% of them, and still have a lift value above 1.0.
- r_1 is more robust when being evaluated in a post-processing step with the lift and a lift threshold of 1.0

$\hookrightarrow$ To be above the threshold is not sufficient. See also for example works to discover false positive [LPT04] and to filter random noise in transaction data [HH07].

Sensitivity to noise, to threshold, etc. ?

Robustness

Four strategies have been proposed :

- experimental approach, using simulation [AK02] [Cad05]
- statistical approach, using statistical tests [LPT04], [RM08]
- formal approach, by studying the derivative of the measures [LLV06], [GDGB07]
- algebraic definition of the robustness by considering the distance between the considered rule and the nearest rule corresponding to the threshold [LBMLL10]

Does the measure help to efficiently find good rules ? [AS94]

Pruning property

How to find nuggets (rules with a very small support and a high confidence) [Li06], or rare rules [SVN10] ?

For FIM, a solution is to avoid the use of the support and to use a measure with a pruning property, see for example [Li06] and its generalization in [LBLL09].

Some criteria [LMVL08]

Property	Semantic	Modalities
g_1	asymmetric processing of `A` and `B`	`asym`, `sym`
g_2	decrease with p_b	`dec`(p_b), `no-dec`(p_b)
g_3	reference situations : independence	`cst`, `var`
g_4	reference situations : logical rule	`cst`, `var`
g_5	linearity with $p_{a\bar{b}}$ around 0^+	`conv`, `lin`, `conc`
g_6	sensitivity to n	`desc`, `stat`
g_7	easiness to fix a threshold	`easy`, `hard`
g_8	intelligibility	`a`, `b`, `c`

$\hookrightarrow$ Study of qualities and drawbacks of interestingness measures (decision aid, classification).
List of criteria could be extended (robustness, discriminant power, etc.).

Evaluation matrix on 20 measures [LMVL08]

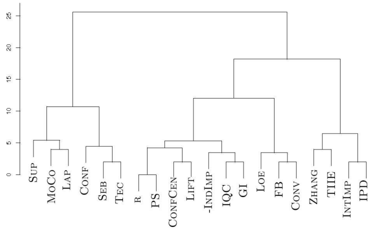
	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8
BF	asym	dec(n_b)	cst	cst	conv	desc	easy	a
ConfCen	asym	dec(n_b)	cst	var	lin	desc	easy	a
Conf	asym	no-dec(n_b)	var	cst	lin	desc	easy	a
Conv	asym	dec(n_b)	cst	cst	conv	desc	easy	b
Tec	asym	no-dec(n_b)	var	cst	conc	desc	easy	b
GI	sym	dec(n_b)	cst	var	conc	desc	easy	c
- IndImp	asym	dec(n_b)	cst	var	lin	stat	easy	c
IntImp	asym	dec(n_b)	cst	var	conc	stat	easy	c
IQC	sym	dec(n_b)	cst	var	lin	desc	easy	c
Lap	asym	no-dec(n_b)	var	var	lin	desc	easy	c
MoCo	asym	dec(n_b)	var	var	lin	desc	easy	b
Lift	sym	dec(n_b)	cst	var	lin	desc	easy	a
Loe	asym	dec(n_b)	cst	cst	lin	desc	easy	b
IPD	asym	dec(n_b)	cst	var	conc	stat	easy	c
PS	sym	dec(n_b)	cst	var	lin	stat	easy	b
R	sym	dec(n_b)	cst	var	lin	desc	easy	b
Seb	asym	no-dec(n_b)	var	cst	conv	desc	easy	b
Sup	sym	no-dec(n_b)	var	var	lin	desc	easy	a
TIIE	asym	dec(n_b)	cst	var	conc	stat	hard	c
Zhang	asym	dec(n_b)	cst	cst	conc	desc	hard	c

Outline

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Comparison of two classifications [VLL04]

[measures×propriety]



Pre-orders on rules sets

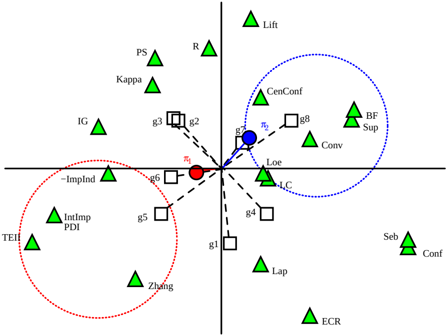
	e_1	e_2	e_3	e_5	e_5
f_1	{LAP}		{SUP, MOCO}		
f_2	{TEC, SEB, CONF}				
f_3				{CONFCEN, LIFT, GI}	{-INDIMP, R, PS, IQC}
f_4		{LOE, CONV, BF}			
f_5		{ZHANG, TIIE}			{INTIMP, IPD}

Decision Aiding [LMVL08]

Selection of the good measures : use of [measures×propriety] and decision maker preferences and a multicriteria decision aid tool.

normative criterion	order
g_1 asymmetric ...	asym $\succ$ sym
g_2 decrease with p_b	dec(n_b) $\succ$ no-dec(n_b)
g_3 independence	cst $\succ$ var
g_4 logical rule	cst $\succ$ var
g_7 easiness to fix σ_μ	easy $\succ$ hard

subjective criterion	order
g_5 linearity with $p_{a\bar{b}}$	conc $\succ$ lin $\succ$ conv
	(tolerance for c-ex)
	conv $\succ$ lin $\succ$ conc
	(no tolerance for c-ex)
g_6 sensitivity to n	stat $\succ$ desc
g_8 intelligibility	a $\succ$ b $\succ$ c



Outline

Conclusion

Some key points presented today

- interest of a rule is context dependent
- there are a lot of interestingness measures with very different behaviours
- the user has to select the good ones in order to select the good rules
- proposition of a systematic/characterizing approach
- formal/experimental clustering and analysis of the measures
- applying MCDA methods for measures selection

$\hookrightarrow$ Theoretical and practical framework leading to useful characterizations and operational applications of interestingness measures.

Readings . . .

- Tan, P-N., Kumar, V. and Srivastava, J., Selecting the Right Objective Measure for Association Analysis, Information Systems, 29 :(4), pp. 293-313, 2004.
- McGarry, K., A survey of Interestingness Measures for Knowledge Discovery, Knowledge Engineering Review Journal, 20 :(1), pp. 39-61, 2005.
- Geng, L., and Hamilton, H.J., Interestingness Measures for Data Mining : A Survey, ACM Computing Surveys, 38(3), Article 9, 2006.
- Lenca P., Meyer P., Vaillant B. and Lallich S., On selecting interestingness measures for association rules : user oriented description and multiple criteria decision aid, Eur. J. of Operational Research, 184(2), pp. 610-626, 2008.
- Suzuki E., Pitfalls for Categorizations of Objective Interestingness Measures for Rule Discovery, Statistical Implicative Analysis : Theory and Applications, Springer-Verlag, SCI (127), pp. 383-395, 2008.

Outline

References I

[AK02] J. Azé and Y. Kodratoff.
Evaluation de la résistance au bruit de quelques mesures d'extraction de règles d'association.
In *EGC*, pages 143–154, 2002.

[AS94] R. Agrawal and R. Srikant.
Fast algorithms for mining association rules.
In *VLDB*, pages 487–499, 1994.

[BFSO84] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen.
Classification and Regression Trees.
Wadsworth Press, 1984.

[BGB03] J. Blanchard, F. Guillet, and H. Briand.
Exploratory visualization for association rule rummaging.
In *Proceedings of the KDD'2003 Workshop on Multimedia Data Mining MDM'03*, pages 107–114, 2003.

[BGBG05] J. Blanchard, F. Guillet, H. Briand, and R. Gras.
Assessing the interestingness of rules with a probabilistic measure of deviation from equilibrium.
In *The XIth International Symposium on Applied Stochastic Models and Data Analysis*, pages 191–200, 2005.

References II

[BMUT97] S. Brin, R. Motwani, J. D. Ullman, and S. Tsur.
Dynamic itemset counting and implication rules for market basket data.
In J. Peckham, editor, *SIGMOD 1997, Proceedings ACM SIGMOD International Conference on Management of Data, Tucson, Arizona, USA*, pages 255–264. ACM Press, 1997.

[Cad05] M. Cadot.
A simulation technique for extracting robust association rules.
In *CSDA*, pages 143–154, 2005.

[Cao10] L. Cao.
Domain driven data mining : challenges and prospects.
IEEE Trans. on Knowledge and Data Engineering, 22(6) :755–769, 2010.

[DP03] T-N. Do and F. Poulet.
Interactive visualization tools for visual data mining.
In R. Bisdorff, editor, *2nd Human Centered Processes Conference*, pages 299–304, Luxembourg, 2003.

[FPSSU96] U.M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, editors.
Advances in Knowledge Discovery and Data Mining.
AAAI/MIT Press, 1996.

References III

[Fre99] A. Freitas.
On rule interestingness measures.
Knowledge-Based Systems journal, pages 309–315, 1999.

[GDGB07] R. Gras, J. David, F. Guillet, and H. Briand.
Stabilité en A.S.I. de l'intensité d'implication et comparaisons avec d'autres indices de qualité de règles d'association.
In *Qualité des Données et des Connaissances*, pages 35–43, 2007.

[GH06] L. Geng and H. J. Hamilton.
Interestingness measures for data mining : A survey.
ACM Computing Surveys, 38(3), 2006.

[GKCG01] R. Gras, P. Kuntz, R. Couturier, and F. Guillet.
Une version entropique de l'intensité d'implication pour les corpus volumineux.
Extraction des connaissances et apprentissage (Extraction et Gestion des Connaissances 2001), 1(1-2) :69–80, 2001.

[HH00] R. J. Hilderman and H. J. Hamilton.
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)*, pages 432–439. Springer-Verlag, 2000.

References IV

[HH07] M. Hahsler and K. Hornik.
New probabilistic interest measures for association rules.
Intelligent Data Analysis, 11(5) :437–455, 2007.

[KNZ01] Y. Kodratoff, A. Napoli, and D. Zighed.
Bulletin de l'association française d'intelligence artificielle, Extraction de connaissances dans des bases de données, 2001.

[LA03] I.C. Lerman and J. Azé.
Une mesure probabiliste contextuelle discriminante de qualité des règles d'association.
In M.-S. Hacid, Y. Kodratoff, and D. Boulanger, editors, *Extraction et gestion des connaissances*, volume 17 of *RSTI-RIA*, pages 247–262. Lavoisier, 2003.

[Lal02] S. Lallich.
Mesure et validation en extraction des connaissances à partir des données.
Habilitation à Diriger des Recherches – Université Lyon 2, 2002.

[LBLL09] Y. Le Bras, P. Lenca, and S. Lallich.
On optimal rules mining : a framework and a necessary and sufficient condition for optimality.
In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, volume 5476 of *Lecture Notes in Computer Science*, pages 705–712. Springer-Verlag Berlin Heidelberg, 2009.

References V

[LBMLL10] Y. Le Bras, P. Meyer, P. Lenca, and S. Lallich.
A robustness measure of association rules.
In *ECML/PKDD*, volume 6322 of *Lecture Notes in Computer Science*, pages 227–242. Springer-Verlag Berlin Heidelberg, 2010.

[LGB98] R. Lehn, F. Guillet, and H. Briand.
Eliminating redundancy in a rule system.
In *Eurepean meeting on Cybernetics and System Research*, volume 2, pages 793–798, 1998.

[LHM98] B. Liu, W. Hsu, and Y. Ma.
Integrating classification and association rule mining.
In *Knowledge Discovery and Data Mining*, pages 80–86, 1998.

[Li06] J. Li.
On optimal rule discovery.
IEEE Transformation on Knowledge and Data Engineering, 18(4) :460–471, 2006.

[LLV06] P. Lenca, S. Lallich, and B. Vaillant.
On the robustness of association rules.
In *The IEEE International Conference on Cybernetics and Intelligent Systems*, pages 596–601, June 7-9 2006.

References VI

[LMVL08] P. Lenca, P. Meyer, B. Vaillant, and S. Lallich.
On selecting interestingness measures for association rules : user oriented description and multiple criteria decision aid.
European Journal of Operational Research, 184(2) :610–626, 2008.

[LPT04] S. Lallich, E. Prudhomme, and O. Teytaud.
Contrôle du risque multiple en sélection de règles d'association significatives.
RNTI-E-2 (EGC 2004), 2 :305–316, 2004.

[LT04] S. Lallich and O. Teytaud.
Évaluation et validation de l'intérêt des règles d'association.
Revue des Nouvelles Technologies de l'Information (Mesures de Qualité pour la Fouille de Données), (RNTI-E-1) :193–218, 2004.

[LVL07] S. Lallich, B. Vaillant, and P. Lenca.
A probabilistic framework towards the parameterization of association rule interestingness measures.
Methodology and Computing in Applied Probability, 9(3) :447–463, 2007.

[MBY10] A. Mouakher and S. Ben Yahia.
Anthropocentric visualisation of optimal cover of association rules.
In *Concept Lattices and Their Applications*, pages 211–222, 2010.

References VII

[PCKW89] K. Parsaye, M. Chignell, S. Khoshafian, and H. Wong.
Intelligent Databases ; Object-Oriented, Deductive Hypermedia Technologies.
John Wiley & Sons, 1989.

[PS91] G. Piatetsky-Shapiro.
Discovery, analysis and presentation of strong rules.
In G. Piatetsky-Shapiro and W.J. Frawley, editors, *Knowledge Discovery in Databases*, pages 229–248. AAAI/MIT Press, 1991.

[RM08] R. Rakotomalala and A. Morineau.
Statistical Implicative Analysis, Theory and Applications, chapter The TVpercent principle for the counterexamples statistic, pages 449–462.
Springer, 2008.

[ST95] A. Silberschatz and A. Tuzhilin.
On subjective measures of interestingness in knowledge discovery.
In *Knowledge Discovery and Data Mining*, pages 275–281, 1995.

[SVN10] Laszlo Szathmary, Petko Valtchev, and Amedeo Napoli.
Finding minimal rare itemsets and rare association rules.
In *Knowledge Science, Engineering and Management*, volume 6291 of *Lecture Notes in Computer Science*, pages 16–27. Springer, 2010.

References VIII

[Tsu00] S. Tsumoto.
Clinical knowledge discovery in hospital information systems : Two case studies.
In D. A. Zighed, H. J. Komorowski, and J. M. Zytkow, editors, *4th European Conference on Principles of Data Mining and Knowledge Discovery, Lyon, France*, pages 652–656. Springer, 2000.

[VLL04] B. Vaillant, P. Lenca, and S. Lallich.
A clustering of interestingness measures.
In *Discovery Science*, volume 3245 of *Lecture Notes in Artificial Intelligence*, pages 290–297. Springer-Verlag, 2004.

[WL00] M. L. Wong and K. S. Leung.
Data mining using grammar based genetic programming and applications.
Kluwer Academic Publishers, 2000.