Multicriteria attractiveness evaluation of association and decision rules

Izabela Szczęch

Poznań University of Technology

Wielokryterialna ocena atrakcyjności reguł asocjacyjnych i decyzyjnych

Poznań, 15.05.2007

Motivations

- Rule induction
- Rule evaluation
- Attractiveness measures and their properties
- Relationships and dependencies between different attractiveness measures
- Potential efficiency gains

Presentation plan

- Rule induction
- Rule evaluation and attractiveness measures
- Desirable properties of attractiveness measures
- Relationships between measures f, s, any measure with property M, and support-confidence Pareto border
- Support and anti-support rule evaluation space
- Conclusions

Introduction – rule induction

- Patterns in form of rules are induced from a data table
- $S=\langle U, A\rangle$ – *data table*, where U and A are finite, non-empty sets
  U – *universe*; A – set of *attributes*
- $S=\langle U, C, D\rangle$ – *decision table*, where C – set of *condition attributes*,
  D – set of *decision attributes*, $C\cap D=\emptyset$
- *Decision rule* or *association rule* induced from S
  is a *consequence relation*: $\phi\rightarrow\psi$ *read as* ***if*** ϕ ***then*** ψ
  where ϕ and ψ are condition and conclusion formulas
  built from attribute-value pairs (q,v)
- If the division into independent and dependent attributes is fixed, then rules are regarded as decision rules, otherwise as association rules.

Introduction – rule induction

Characterization of nationalities

U	*Height*	*Hair*	*Eyes*	*Nationality*	Support
1	*tall*	*blond*	*blue*	*Swede*	270
2	*medium*	*dark*	*hazel*	*German*	90
3	*medium*	*blond*	*blue*	*Swede*	90
4	*tall*	*blond*	*blue*	*German*	360
5	*short*	*red*	*blue*	*German*	45
6	*medium*	*dark*	*hazel*	*Swede*	45

C *D*

- E.g. decision rules induced from „characterization of nationalities”:
  1) If (*Height=tall*), then (*Nationality=Swede*)
  2) If (*Height=medium*) & (*Hair=dark*), then (*Nationality=German*)

Introduction – attractiveness measures

- To measure the relevance and utility of rules, quantitative measures called attractiveness or interestingness measures, have been proposed

  (e.g. support, confidence, lift, gain, conviction, Piatetsky-Shapiro,…)
- Unfortunately, there is no evidence which measure(s) is (are) the best
- Notation:
  - $sup(\circ)$ is the number of all objects from *U*, having property °

    e.g. $sup(\phi)$, $sup(\psi)$

Basic quantitative characteristics of rules

- Basic quantitative characteristics of rules
  - *Support* of rule $\phi \rightarrow \psi$ in *S*:

$$sup(\phi \rightarrow \psi) = sup(\phi \wedge \psi)$$

  - *Confidence* (called also *certainty factor*) of rule $\phi \rightarrow \psi$ in *S*:

$$conf(\phi \rightarrow \psi) = \frac{sup(\phi \rightarrow \psi)}{sup(\phi)}$$

  - *Anti-support* of rule $\phi \rightarrow \psi$ in *S*:

$$anti\text{-}sup(\phi \rightarrow \psi) = sup(\phi \wedge \neg\psi)$$

Confirmation measure *f* and *s*

- Confirmation measure *f* (Good 1984, Heckerman 1988, Pearl 1988, Fitelson 2001)

$$f(\phi \to \psi) = \frac{conf(\psi \to \phi) - conf(\neg\psi \to \phi)}{conf(\psi \to \phi) + conf(\neg\psi \to \phi)}$$

- Confirmation measure *s* (Christensen 1999)

$$s(\phi \to \psi) = conf(\phi \to \psi) - conf(\neg\phi \to \psi)$$

- Gain measure (Fukuda et al. 1996)
- Rule Interest Function (Piatetsky-Shapiro 1991)
- Dependency Factor (Pawlak 2002)
- Conviction (Brin et al. 1997)
- ...

Bayesian confirmation property

- An attractiveness *c* measure has the property of confirmation if is satisfies the following condition:

$$c(\phi,\psi)\begin{cases} >0 \;\; if \;\; Pr(\psi|\phi) > Pr(\psi) \\ =0 \;\; if \;\; Pr(\psi|\phi) = Pr(\psi) \\ <0 \;\; if \;\; Pr(\psi|\phi) < Pr(\psi) \end{cases}$$

- Measures of confirmation quantify the strength of confirmation that premise ϕ gives to conclusion ψ
- „ψ is verified more often, when ϕ is verified, rather than when ϕ is not verified”

Bayesian confirmation property - interpretation

- $c(\phi, \psi)>0$ means that property ψ is satisfied more frequently when ϕ is satisfied (then, this frequency is $conf(\phi, \psi)$), rather than generically in S (where the frequency is $Pr(\psi)$),

- $c(\phi, \psi)=0$ means that property ψ is satisfied with the same frequency whether ϕ is satisfied or not

- $c(\phi, \psi)<0$ means that property ψ is satisfied less frequently when ϕ is satisfied, rather than generically

Property M

- Property M (Greco, Pawlak, Słowiński 2004)

- An attractiveness measure $I(a, b, c, d)$ has the property M if it is a function non-decreasing with respect to *a* and *d* and non-increasing with respect to *b* and *c*

  where:

  $a = sup(\phi \rightarrow \psi)$

  the number of objects in *U* for which ϕ and ψ hold together

  $b = sup(\neg\phi \rightarrow \psi)$,

  $c = sup\ (\phi \rightarrow \neg\psi)$,

  $d = sup(\neg\phi \rightarrow \neg\psi)$

Interpretation of the property M

- E.g. (Hempel) consider rule $\phi \rightarrow \psi$:

  if x is a raven then x is black

- ϕ is the property *to be a raven*, ψ is the property *to be black*
  - *a* – the number of objects in *U* which are black ravens

    //the more black ravens we observe, the more credible becomes the rule
  - *b* – the no. of objects in *U* which are black non-ravens
  - *c* – the no. of objects in *U* which are non-black ravens
  - *d* – the no. of objects in *U* which are non-black non-ravens

Attractiveness measures with the property M

- Open question:
  which attractiveness measures possess the property M?

- *Theorem* [2004]*:*
  Confirmation measures f, s have the property M

- *Theorem:*
  Rule support, Confidence, Rule Interest Function, Gain measure
  have the property M

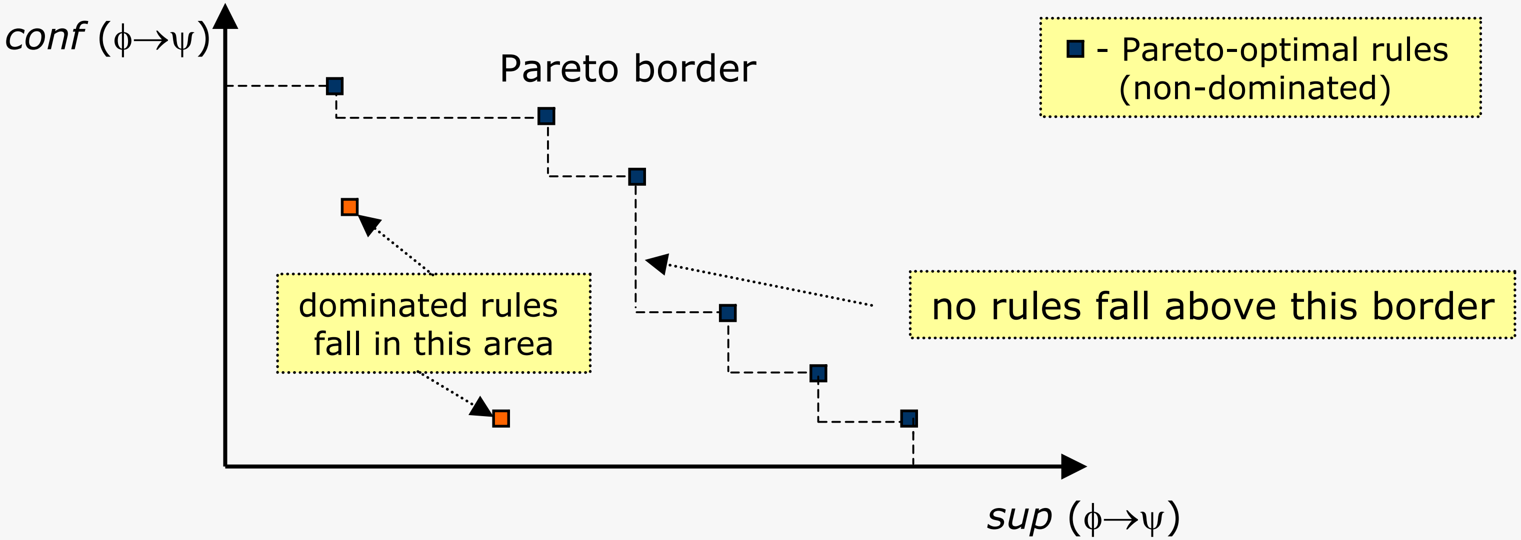
- *Theorem:*
  Dependency factor does not have the property M

Greco S., Słowiński R., Szczęch I.: Analysis of monotonicity properties of some rule interestingness measures (manuscript, 2007)

Support-confidence Pareto border

Support-confidence Pareto border

- Support-confidence Pareto border is the set of non-dominated, Pareto-optimal rules with respect to both rule support and confidence



- Mining the border identifies rules optimal with respect to measures such as: *lift*, *gain*, *conviction*, *Piatetsky-Shapiro*,…

Support-confidence Pareto border

- The following conditions are sufficient for verifying whether rules optimal according to a measure $g(x)$ are included on the support-confidence Pareto border:

  1. $g(x)$ is monotone in support over rules with the same confidence

  and

  2. $g(x)$ is monotone in confidence over rules with the same support

- A function $g(x)$ is understood to be *monotone* in x, if $x_1 \prec x_2$ implies that $g(x_1) \leq g(x_2)$

Support-f Pareto border

Monotonicty of f in support and confidence

- Is measure f included in the support-confidence Pareto border?

- <u>*Theorem:*</u>

  Confirmation measure f is independent of support, and, therefore, monotone in support, when the value of confidence is held fixed.
- <u>*Theorem:*</u>

  Confirmation measure f is increasing, and, therefore, monotone in confidence

- Conclusion:

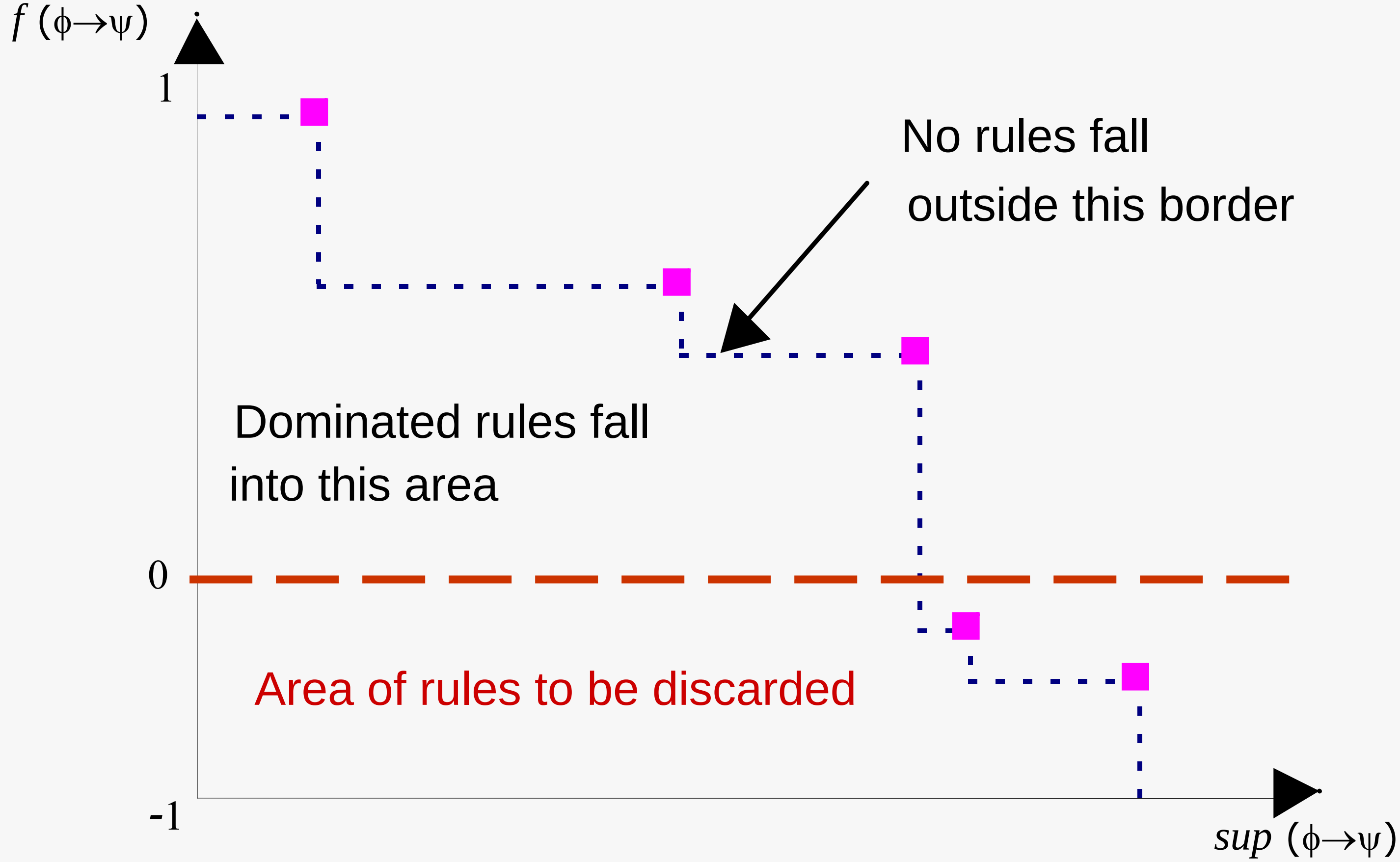
  Rules maximizing f lie on the support-confidence Pareto border

Support-confidence vs. support-*f* Pareto border

- The utility of confirmation measure *f* outranks utility of confidence
- **Claim:** Substitute the $conf(\phi \to \psi)$ dimension for $f(\phi \to \psi)$
- *Theorem:*
  The set of rules located on the support-confidence Pareto border is exactly the same as on the support-*f* Pareto border

Brzezińska I., Słowiński R.: Monotonicity of a Bayesian confirmation measure in rule support and confidence. (AI-METH Gliwice, 2005)

Support-*f* Pareto border is more meaningful



Confirmation perspective on support-confidence space

- Is there a curve separating rules with negative value of any measure with the confirmation property in the support-confidence space?

- <u>*Theorem:*</u>

  Rules lying above a constant:

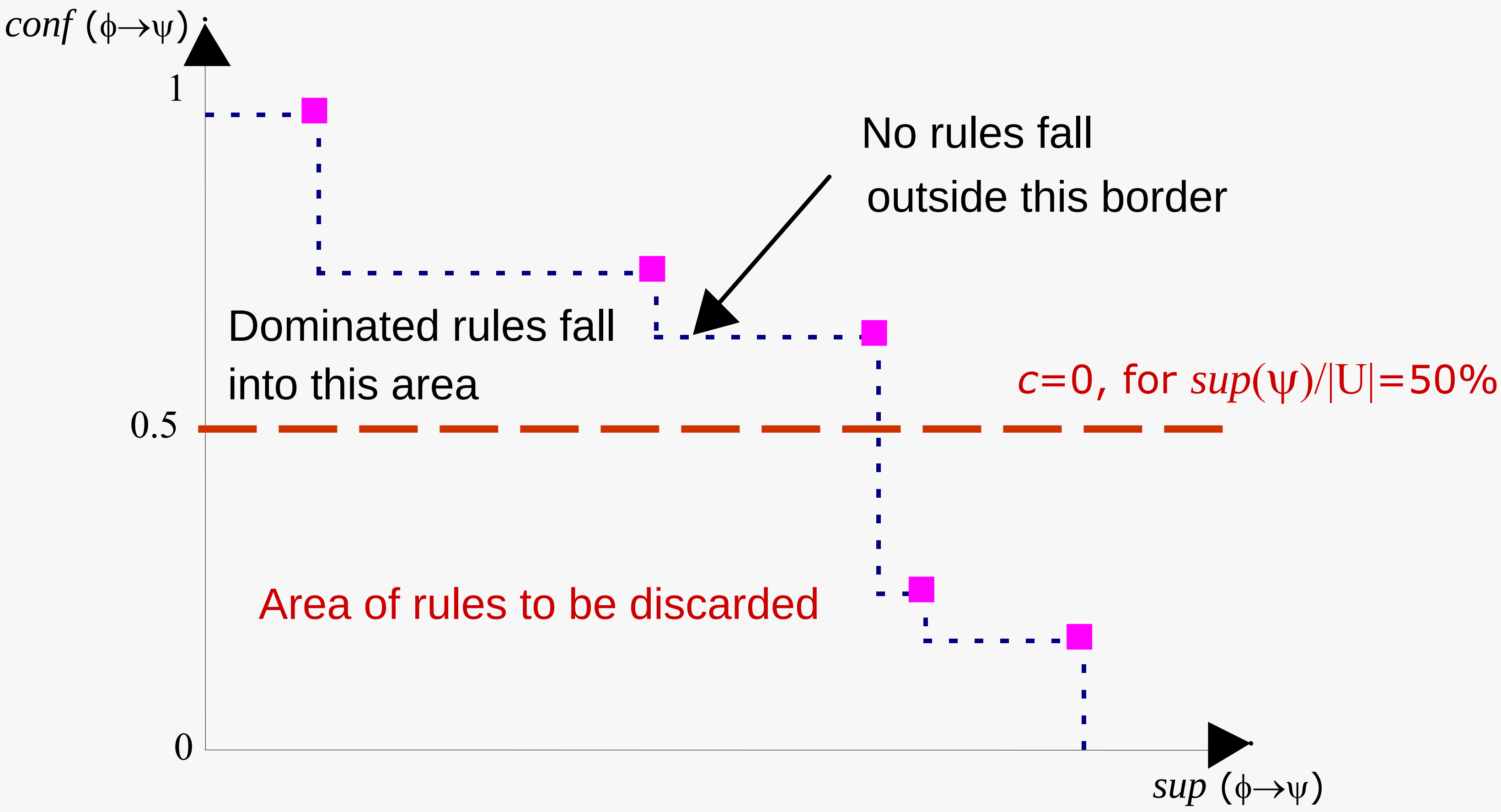
$$sup(\psi)/|U|$$

  have a negative value of any confirmation measure.

  For those rules, the premise only disconfirms the conclusion!

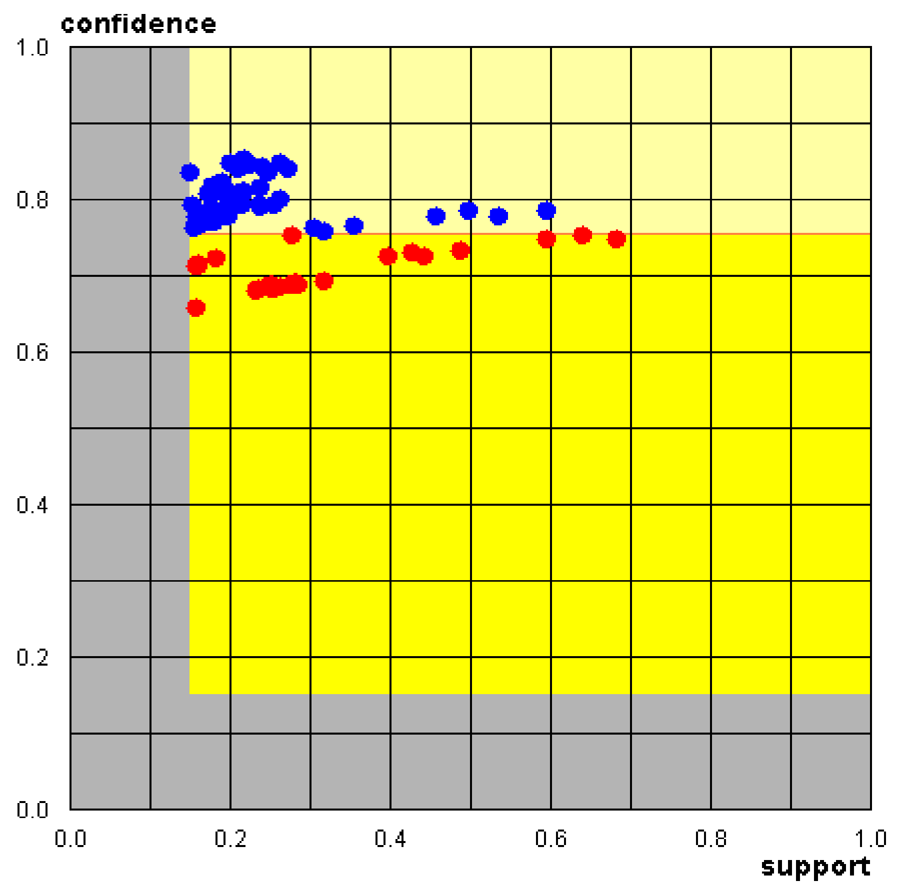
Słowiński R., Szczęch I., Greco S.: Mining Association Ruleswith respect to Support and Anti–support - experimental results. (RSEISP LNAI Warszawa, to appear)

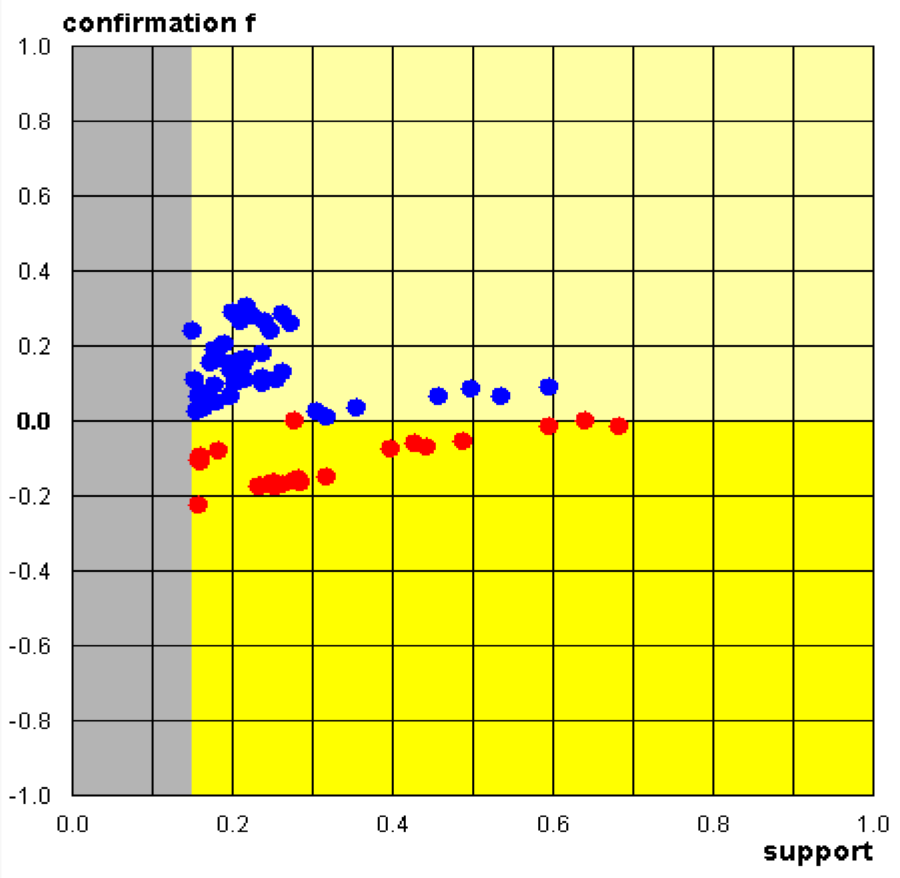
Confirmation perspective on support-confidence space



For rules lying below the curve for which *c*=0
the premise only disconfirms the conclusion

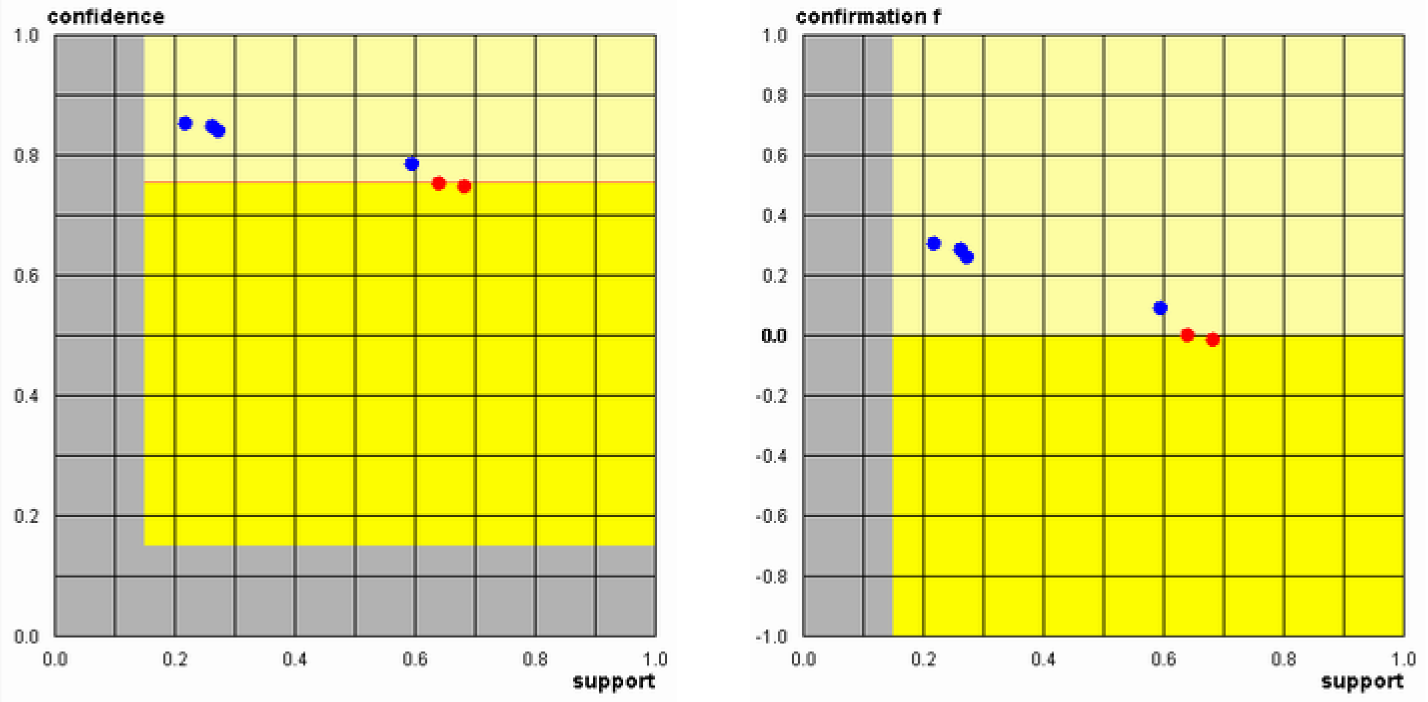
Support-confidence Pareto border vs. support-f





- ● indicates rules with negative confirmation
- the class constitutes over 70% of the whole dataset
- rules with high confidence can be disconfirming
- even some rules from the Pareto border need to be discarded

Support-confidence Pareto border vs. support-f



premise	conclusion	supp	conf	s	f	a-supp
marital-status **is** Never-married **and** **race** **is** White **and** **class** **is** =50K	**workclass** **is** Private	0.22	0.85	0.13	0.30	0.04
marital-status **is** Never-married **and** **class** **is** =50K	**workclass** **is** Private	0.26	0.85	0.13	0.28	0.05
marital-status **is** Never-married	**workclass** **is** Private	0.27	0.84	0.13	0.26	0.05
race **is** White	**workclass** **is** Private	0.64	0.75	-0.01	-0.00	0.21
native-country **is** United-States	**workclass** **is** Private	0.68	0.75	-0.07	-0.02	0.23
class **is** =50K	**workclass** **is** Private	0.60	0.78	0.13	0.09	0.16

- both Pareto borders contain the same rules

Support-s Pareto border

Monotonicty of *s* in support and confidence

- Is measure s on rule support-confidence Pareto border?

- Theorem:
  Confirmation measure s is increasing, and, therefore, monotone in confidence when the value of support is held fixed

- Theorem:
  For a fixed value of confidence, confirmation measure s is:
  - increasing in $sup(\phi\rightarrow\psi) \Leftrightarrow s(\phi\rightarrow\psi)>0$
  - constant in $sup(\phi\rightarrow\psi) \Leftrightarrow s(\phi\rightarrow\psi)=0$
  - decreasing in $sup(\phi\rightarrow\psi) \Leftrightarrow s(\phi\rightarrow\psi)<0$

- The above theorem states the monotone relationship just in the non-negative range of the value of *s* (i.e. the only interesting)

Support-confidence vs. support-*s* Pareto border

- Theorem:

  If a rule resides on the support-s Pareto border

  (in case of positive value of s),

  then it also resides on the support-confidence Pareto border,

  while one can have rules being on the support-confidence Pareto border which are not on the support-s Pareto border.

- Conclusion:

  The support-confidence Pareto border is, in general, larger than the support-s Pareto border

Measures with the property M in support-confidence space

- What are the conditions for rules maximizing

  any measure with the property M

  to be included in the rule support-confidence Pareto border?

- Reminder of the property M:

  $a=sup(\phi\rightarrow\psi)$, $b=sup(\neg\phi\rightarrow\psi)$, $c=sup\ (\phi\rightarrow\neg\psi)$, $d=sup(\neg\phi\rightarrow\neg\psi)$

  $I(a,b,c,d)$ is a function non-decreasing with respect to a and d,

  and non-increasing with respect to b and c

Measures with the property M in support-confidence space

- Theorem:

  When the value of support is held fixed, then $I(a, b, c, d)$ is monotone in confidence.

- Theorem:

  When the value of confidence is held fixed, then $I(a, b, c, d)$ admitting derivative with respect to all its variables a, b, c and d, is monotone in support if:

$$\frac{\partial I}{\partial c} = \frac{\partial I}{\partial d} = 0 \quad or \quad \frac{\dfrac{\partial I}{\partial a} - \dfrac{\partial I}{\partial b}}{\dfrac{\partial I}{\partial d} - \dfrac{\partial I}{\partial c}} \geq \frac{1}{conf(\phi \rightarrow \psi)} - 1$$

Słowiński R., Brzezińska I., Greco S.: Application of Bayesian confirmation measures for mining rules from support-confidence Pareto-optimal set. (ICAISC LNAI Zakopane, 2006)

Measures with the property M in support-confidence space

- Conclusions:
  - *F*or a set of rules with the same conclusion, any interestingness measure with property M is always non-decreasing with respect to confidence when the value of support is kept fixed
  - All those interestingness measures that are independent of $c=sup(\phi\to\neg\psi)$ and $d=sup(\neg\phi\to\neg\psi)$ are always monotone in support when the value of confidence remains unchanged
  - There are some measure with property M whose optimal rules will not be on the support-confidence Pareto border.

Support-anti-support Pareto border

Support - anti-support Pareto border

- How to find rules optimal according to **any** measure with the property M?

- Theorem:
  When the value of support is held fixed, then $I(a, b, c, d)$ is anti-monotone (non-increasing) in anti-support

- Theorem:
  When the value of anti-support is held fixed, then $I(a, b, c, d)$ is monotone (non-decreasing) in support

Support - anti-support Pareto border

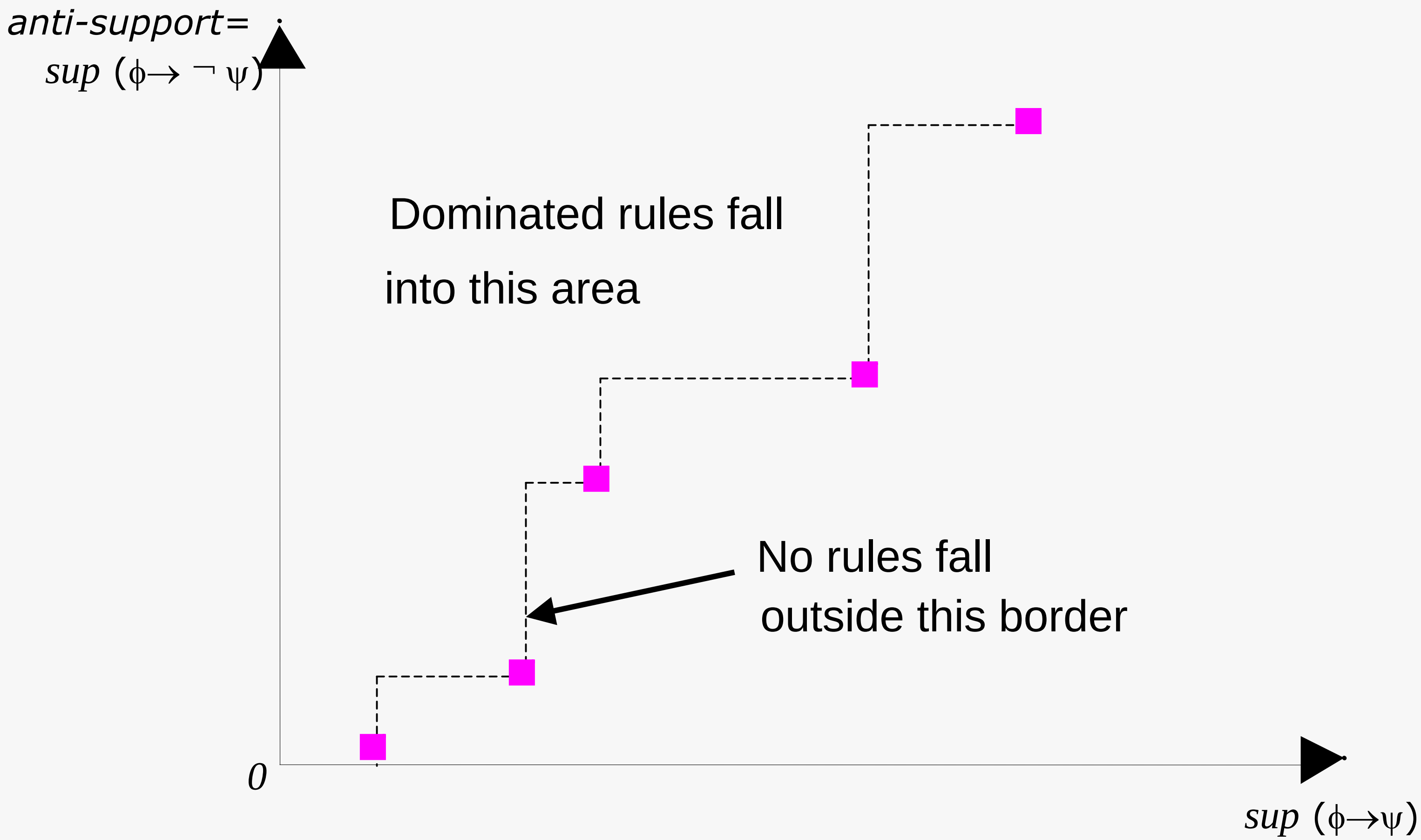
- <u>Theorem:</u>
  For rules with the same conclusion,
  the best rules according to any measure with the property M
  must reside on the support-anti-support Pareto border

- The support-anti-support Pareto border is the set of rules such that there is no other rule having greater support and smaller anti-support

- <u>Theorem:</u>
  The support - anti-support Pareto border is, in general, not smaller than the support-confidence Pareto border

Brzezińska I., Greco S., Słowiński R.: Mining Pareto-Optimal Rules with Respect to Support and Confirmation or Support and Anti-Support (EAAI Journal, 2007)

Support - anti-support Pareto border



The best rules according to any measure with the property M must reside on the support - anti-support Pareto border

Confirmation perspective on the support - anti-support Pareto border

Confirmation perspective on support - anti-support border

- Is there a curve separating rules with negative value of any confirmation measure in the support-anti-support space?

- *Theorem:*
  Rules lying above a linear function:

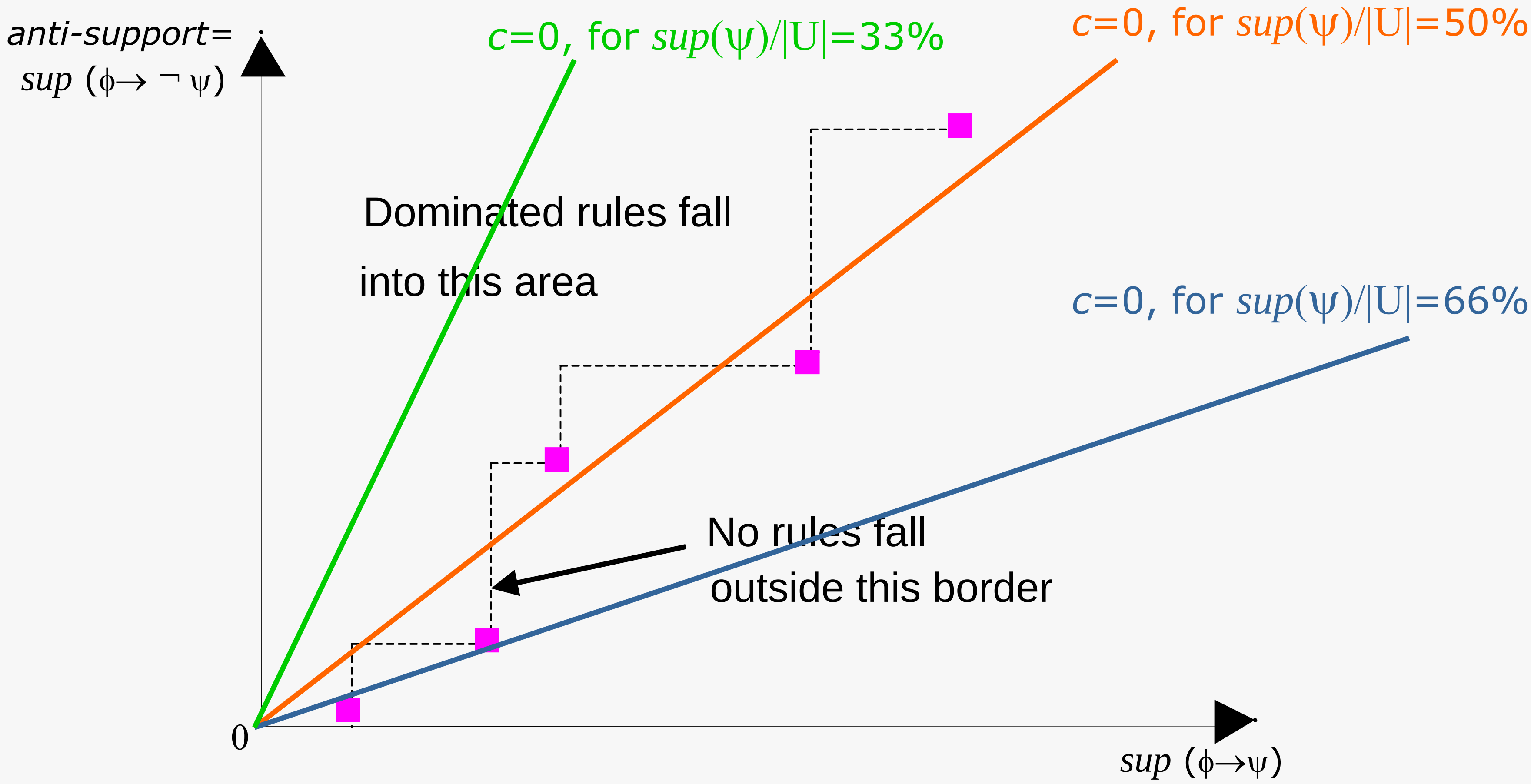
$$sup(\phi\rightarrow \psi)[|U|/sup(\psi)\text{-}1]$$

  have a negative value of any confirmation measure.

  For those rules, the premise only disconfirms the conclusion!

Słowiński R., Szczęch I., Greco S.: Mining Association Ruleswith respect to Support and Anti–support - experimental results. (RSEISP LNAI Warszawa, to appear)

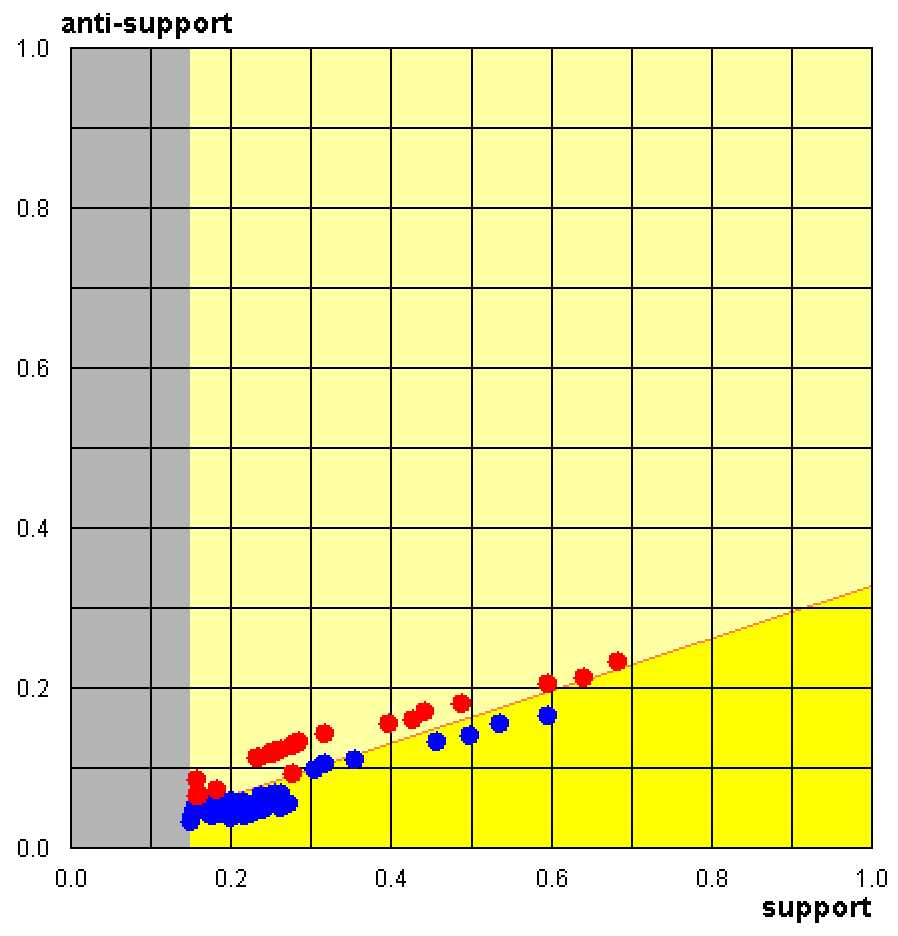
Confirmation perspective on support - anti-support border



For rules lying above the curve for which *c*=0

the premise only disconfirms the conclusion

Support - anti-support (workclass=Private)



- ● indicates rules with negative confirmation
- even some rules from the Pareto border need to be discarded

Inner monotonicity
in support - anti-support space

The gist of the algorithm for support - anti-support rules

- Traditional Apriori approach to generation of association rules (Agrawal et al) proceeds in a two step framework:
  - find frequent itemsets (i.e. sets of items which occur more frequently than the minimum support threshold),
  - generate rules from frequent itemsets and filter out those that do not exceed the minimum confidence threshold

- Generation of association rules regarding support and anti-support, in general, requires only the substitution of the parameter calculated in step 2. Confidence -> anti-support

The gist of the algorithm for support - anti-support rules

- **Claim**: calculation of anti-support (instead of confidence) does not introduce any more computational overhead to the algorithm

- Let us observe that: $\textit{anti-sup}(\phi \rightarrow \psi) = sup(\phi \rightarrow \neg\, \psi) = sup(\phi) - sup(\phi \rightarrow \psi)$.
- All the data required to calculate anti-support are also gathered in step 1 of Apriori
- The data needed to calculate anti-support is the same as to calculate confidence

The gist of the algorithm for support - anti-support rules

- **Claim**: When generating association rules from a frequent set it is advisable to first generate rules with few conclusion elements (for optimisation reasons)
- Let us observe three different rules constructed from the same frequent itemset $\{x, y, z, v\}$:
  - $r_1: x \rightarrow yzv$ $\quad$ $anti\text{-}sup(r_1) = sup(x) - sup(xyzv)$
  - $r_2: xy \rightarrow zv$ $\quad$ $anti\text{-}sup(r_2) = sup(xy) - sup(xyzv)$
  - $r_3: xyz \rightarrow v$ $\quad$ $anti\text{-}sup(r_3) = sup(xyz) - sup(xyzv)$
- $anti\text{-}sup(r_1) \geq anti\text{-}sup(r_2) \geq anti\text{-}sup(r_3)$
- **Conclusion**: $anti\text{-}sup(r_3)$ *> max_acceptable anti-support =>*

  $anti\text{-}sup(r_2)$ *> max_acceptable anti-support*

  Generate and verify $\boldsymbol{r_3}$ first!

Summary

Summary

- Attractiveness measures with desirable property M were considered
  - we have verified property M for measures: gain, rule interest function, dependency factor
- Relationships between measures were analysed

  We have analytically shown:
  - relationships between measures f, s, any measure with property M, and support-confidence Pareto border
  - enclosure relationship between considered Pareto-optimal borders
  - a way to impose the confirmation perspective on the support-confidence space

Summary

Moreover, we have proved that:

- Pareto border w.r.t. support and anti-support includes rules maximizing any measures with the property M
- support - anti-support Pareto border includes support-confidence border
- a linear function narrows the area of rules only to rules for which the premise confirms the conclusion
- there is a monotonic relationship between support and anti-support

Verification of those relationships results in potential efficiency improvement as it allows to limit the space of the analysed rules and concentrate on mining one Pareto set while having rules optimal wrt many different measures.

Further lines of investigation

- Verification of the property M for other attractiveness measures

- Development of algorithm for finding in support - anti-support space a set of rules (both dominated and non-dominated) that covers the objects in a certain percentage

Thank you!

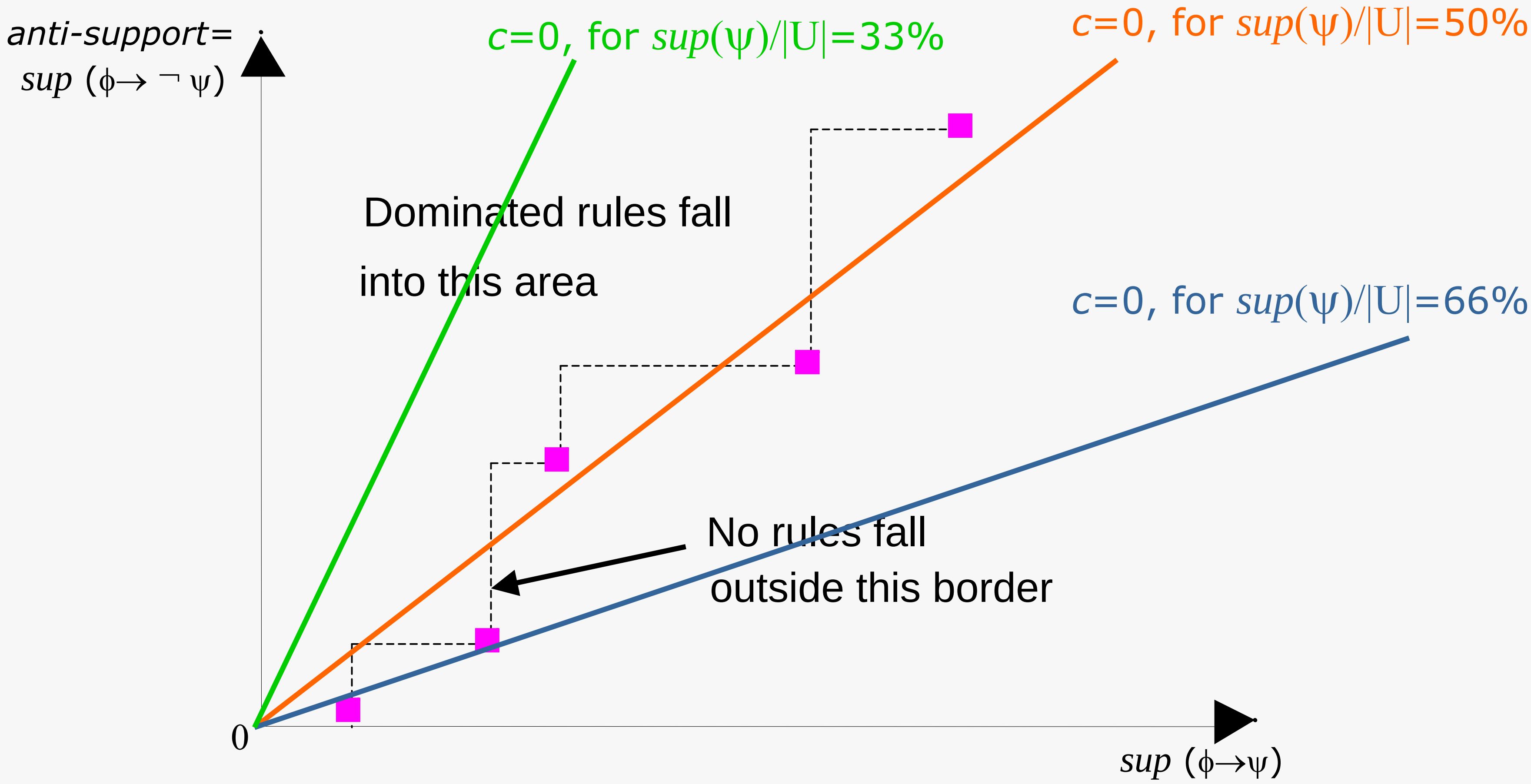
References

- Brzezińska, I., Słowiński, R.: Monotonicity of a Bayesian confirmation measure in rule support and confidence, in: Recent Developments in Artificial Intelligence Methods, AI-METH Series, Gliwice (2005), 39-42.
- Brzezińska, I, Greco, S., Słowiński, R.: Mining Pareto–optimal rules with respect to support and anti–support. [In]: *Journal of Engineering Applications of Artificial Intelligence* (2006).
- Słowiński R., Brzezińska I., Greco S.: Application of Bayesian confirmation measures for mining rules from support-confidence Pareto-optimal set. (ICAISC LNAI Zakopane, 2006)
- Słowiński R., Szczęch I., Greco S.: Mining Association Ruleswith respect to Support and Anti–support - experimental results. (RSEISP LNAI Warszawa, to appear)

General info about the dataset

- Dataset ***adult***, created in ’96 by B. Becker/R. Kohavi from census database
- **32 561 instances**
- 9 nominal attributes
  - workclass: Private, Local-gov, etc.;
  - education: Bachelors, Some-college, etc.;
  - marital-status: Married, Divorced, Never-married, et.;
  - occupation: Tech-support, Craft-repair, etc.;
  - relationship: Wife, Own-child, Husband, etc.;
  - race: White, Asian-Pac-Islander, etc.;
  - sex: Female, Male;
  - native-country: United-States, Cambodia, England, etc.;
  - salary: >50K, <=50K
- **throughout the experiment, $sup(\phi \rightarrow \psi)$ is denoted as „support” and expressed as a relative rule support [0-1]**

Confirmation perspective on support - anti-support border



For rules lying above the curve for which c=0

the premise only disconfirms the conclusion

Few rules describing class: workclass=Private

premise	conclusion	supp	conf	s	f	a-supp
education is HS-grad **and race is** White **and native-country is** United-States	**workclass is** Private	0.20	0.79	0.05	0.10	0.06
education is HS-grad **and sex is** Male **and native-country is** United-States	**workclass is** Private	0.16	0.76	0.01	0.02	0.05
education is HS-grad **and native-country is** United-States	**workclass is** Private	0.24	0.79	0.05	0.10	0.06
education is Some-college **and native-country is** United-States	**workclass is** Private	0.16	0.77	0.02	0.03	0.05
marital-status is Married-civ-spouse **and relationship is** Husband **and race is** White	**workclass is** Private	0.25	0.69	-0.11	-0.17	0.12
relationship is Husband **and race is** White **and sex is** Male	**workclass is** Private	0.25	0.69	-0.11	-0.17	0.12
relationship is Husband **and sex is** Male **and native-country is** United-States	**workclass is** Private	0.25	0.68	-0.12	-0.18	0.12
race is White **and sex is** Male	**workclass is** Private	0.43	0.73	-0.06	-0.06	0.16
sex is Male **and native-country is** United-States	**workclass is** Private	0.44	0.72	-0.07	-0.07	0.17

- the table contains few examples of rules with the conclusion workclass=Private

Confirmation perspective on support-anti-support border