Uczenie się preferencji w problemie rankingu z wykorzystaniem dominacyjnej teorii zbiorów przybliżonych (ang. DRSA)

Marcin Szeląg

Institute of Computing Science, Poznań University of Technology

13.11.2019

Outline

Introduction

Problem Setting

Multicriteria Decision Aiding (MCDA)

- *multicriteria decision problems*: *classification*, *ranking*, and *choice*
- *objects* (variants, alternatives, options, candidates)
- *evaluation criteria* with explicit monotonic preference scales
- *consistent set of criteria* – conditions of completeness, monotonicity, and non-redundancy
- *information table*, *decision table*
- *decision maker* (user), DM
- *dominance relation*
- *preference information*
- indirect preference information – *decision examples*
- (*in*)*consistency* of decision examples

Problem Setting

Table: Exemplary decision table with evaluations of students

Student	Mathematics	Physics	Literature	Overall Evaluation
$S1$	**good**	medium	bad	**bad**
$S2$	**medium**	medium	bad	**medium**
$S3$	medium	medium	medium	medium
$S4$	good	good	medium	good
$S5$	good	medium	good	good
$S6$	good	good	good	good
$S7$	bad	bad	bad	bad
$S8$	bad	bad	medium	bad

Problem Setting

Multicriteria Decision Aiding (MCDA)

- *preference model* – value function, outranking relation, set of if-then decision rules
- *induction* of preference model from decision examples
- application of preference model $\rightarrow$ *preference structure* on a set of objects
- *exploitation* of preference structure $\rightarrow$ *recommendation*
- importance of *readability* of a preference model
- attractiveness of *rule preference model*
- *Dominance-based Rough Set Approach* (DRSA) $\rightarrow$ *structuring* of decision examples into *lower* and *upper approximations* + induction and application of decision rules
- *Variable Consistency DRSA* (VC-DRSA) $\rightarrow$ *object consistency measures* (e.g., ϵ), *monotonicity properties* (m1)-(m4)

Problem Setting

Machine Learning (ML)

- learning on training objects, testing on unseen (test) objects
- stochastic process generating the observed data (the "ground truth")
- monotonic preference scales converting elementary features to criteria are (usually) neither used nor revealed explicitly

Preference Learning (PL)

- emerging as an important subfield of ML
- "learning to rank" (recommender systems, information retrieval)
- minimization of a loss function

Problem Setting

Beyond the frame of MCDA

- DRSA can also handle monotonic relationships observed for problems where preference are not considered, e.g.,
  - "the colder the weather, the higher the energy consumption",
  - "the more a tomato is red, the more it is ripe",
  - "the larger the mass and the smaller the distance, the larger the gravity".

Multicriteria Ranking Problem

Multicriteria ranking problem is a decision problem in which a finite set of objects A described by a set of criteria $G = \{g_1, \ldots, g_n\}$ has to be ordered, either completely (total preorder, also called weak order) or partially (partial preorder).

Each criterion $g_i \in G$ is modeled as a real-valued function

$$g_i : A \to \mathbb{R},$$

with

- cardinal scale (i.e., interval scale or ratio scale) or
- ordinal scale (given a priori or resulting from an order-preserving number-coding of non-numerical ordinal evaluations).

Multicriteria Ranking Problem

- Cardinal criterion = criterion with cardinal scale.
  - One can measure the intensity of preference (positive or negative) of object a over object b, taking into account evaluations $g_i(a)$, $g_i(b)$, $a, b \in A$, using any function

    $$k_i : \mathbb{R}^2 \to \mathbb{R}$$

    non-decreasing w.r.t. the first evaluation, and non-increasing w.r.t. the second evaluation.
    Greco S, Matarazzo B, Słowiński R, Rough sets theory for multicriteria decision analysis, European J. Operational Research 129(1), 2001, pp. 1–47.
  - For the sake of simplicity, it is assumed that

    $$k_i(g_i(a), g_i(b)) = \Delta_i(a, b) = g_i(a) - g_i(b).$$

- Ordinal criterion = criterion with ordinal scale.
  - Differences of evaluations are not meaningful.
  - One can only establish an order of evaluations $g_i(a)$, $a \in A$.

Multicriteria Ranking Problem – Example

Car ranking problem

Order a given set of 14 cars from the best to the worst (with possible ties), taking into account the following criteria:

1. maximum speed in km/h (to be maximized),
2. comfort: low $\prec$ medium $\prec$ high (to be maximized),
3. price in EUR (to be minimized),
4. fuel consumption per 100 km (to be minimized).

Multicriteria Ranking Problem

Existing MCDA approaches

- Multiple Attribute Utility Theory (MAUT) $\rightarrow$ UTA, GRIP, AHP, PAPRIKA, ...
- outranking methods $\rightarrow$ ELECTRE III and IV, ELECTRE$^{\mathsf{GKMS}}$, PROMETHEE I and II, PROMETHEE$^{\mathsf{GKS}}$, . . .
- previous decision rule-based approaches (α, β, γ, and δ)

Existing PL approaches

- SVM^{rank}
- RankBoost
- Ensembles of decision rules
- . . .

General Motivations

- Practical importance of the ranking problem.
- Many methods applied to solve this problem:
  - are hard to use (i.e., require too much cognitive effort on the part of a DM),
  - are not always appropriate (e.g., in case of ordinal attributes),
  - produce preference/classification models that are not meaningful to a DM.
- Main difficulty consists in aggregation of different criteria; usually such aggregation is performed arbitrary, using weights or aggregation operators like sum, average or distance metrics.
- Need for multicriteria modeling method that allows to:
  - include domain knowledge,
  - handle possible inconsistencies w.r.t. dominance relation,
  - avoid using aggregation operators.

Motivations for Application of DRSA

Dominance-based Rough Set Approach (DRSA), introduced by Greco, Matarazzo and Słowiński in 90's:

- handles inconsistencies in decision examples, resulting, e.g., from imprecise of incomplete information,
- takes into account domain knowledge:
  - domains of attributes, i.e., sets of values that an attribute may take while being meaningful for user's perception,
  - division of attributes into condition and decision attributes,
  - preference order in the domains of attributes and monotonic relationships between attributes,
- works with heterogeneous attributes – nominal, ordinal (without conversion) and cardinal (no need of discretization),
- enables to infer decision rule model (inductive learning).

Motivations for Using Decision Rule Model

Advantages of decision rules:

- comprehensible form of knowledge representation,
- combination of elementary conditions instead of (arbitrary) aggregation of criteria/features,
- can represent any function (more general than utility functions or binary relations),
- give account of most complex interactions among criteria,
- accept ordinal evaluation scales,
- exploit only ordinal properties of criteria/marginal similarity functions,
- support "backtracking",
- can explain past decisions and predict future decisions,
- "resistant" to irrelevant attributes.

Application of VC-DRSA to Multicriteria Ranking Problem

How Multicriteria Ranking Fits MCDA Framework?

Summary of main features

- recommendation = ranking,
- decision examples = pairwise comparisons of reference objects,
- dominance relation on pairs of objects,
- consistency of pairs of objects,
- pairwise comparison table (PCT),
- PCT-oriented adaptation of (VC-)DRSA,
- decision rules concern pairs of objects.

The only objective information concerning set A of objects is the dominance relation D over A:

$$aDb \Leftrightarrow g_i(a) \succeq g_i(b) \text{ for all } g_i \in G.$$

However, usually this relation leaves many objects incomparable.

In order to make the objects more comparable, the DM has to supply some preference information.

Sources of preference information:

- pairwise comparisons (or ranking, or ordinal classification) of some reference objects (set A^R), i.e., objects relatively well known to the DM,
- ranking of reference objects, called reference ranking,
- ordinal classification.

Methodology for Multicriteria Ranking – VC-DRSA^{rank}

Preference information is used to induce a preference model in terms of a set of "*if* ... *then* ..." decision rules.

After acceptance by the DM, this model can be used to build a ranking (complete or partial) of all objects from set A.

Methodology for Multicriteria Ranking – VC-DRSA^{rank}

Two problem settings are considered:

(1) set G is a consistent set of criteria, i.e., G satisfies the properties of:

- *completeness* (all relevant criteria are considered),
- *monotonicity* (the better the evaluation of an object on considered criteria, the more it is preferable to another object),
- *non-redundancy* (there is no criterion which could be removed without violating one of the previous two properties),

(2) set G is a not necessarily consistent set of criteria.

Setting (1) $\rightarrow s_{MCDA}$; typical for Multiple Criteria Decision Aiding.
Setting (2) $\rightarrow s_{ML}$; typical for Machine Learning.

Pairwise Comparison Table (PCT)

- Defined by pairwise comparisons of reference objects.
- $B \subseteq A^R \times A^R$ = set of pairs of compared reference objects.
- Given objects $a, b \in A^R$, $a \neq b$, the DM can declare that:
  - "a is at least as good as b" (a outranks b, denoted by aSb) or
  - "a is NOT at least as good as b" (a does not outrank b, denoted by aS^cb)

  or (s)he can abstain from any judgment.
- aSa is fixed for every $a \in A^R$.
- For s_{MCDA}, aSb is fixed for $a, b \in A^R$ such that aDb.
- When comparing objects $a, b \in A^R$ on a cardinal criterion, one puts in the corresponding column of PCT value $k_i(g_i(a), g_i(b)) = \Delta_i(a, b)$, i.e., difference of evaluations.
- When comparing objects $a, b \in A^R$ on an ordinal criterion, one puts in the corresponding column of PCT ordered pair of evaluations $(g_i(a), g_i(b))$.

Pairwise Comparison Table (PCT)

Exemplary PCT, where g_1 – cardinal criterion, g_n – ordinal criterion:

Pair of ref. objects	Evaluations of pair on criteria			Preference information
	g_1	$\ldots$	g_n	
(a, b)	$\Delta_1(a, b)$	$\ldots$	$(g_n(a), g_n(b))$	aSb
(b, a)	$\Delta_1(b, a)$	$\ldots$	$(g_n(b), g_n(a))$	bS^ca
(b, c)	$\Delta_1(b, c)$	$\ldots$	$(g_n(b), g_n(c))$	bSc
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
(d, e)	$\Delta_1(d, e)$	$\ldots$	$(g_n(d), g_n(e))$	dS^ce

Dominance Relation for Pairs of Objects

Dominance principle - monotonic relationship expected to hold

"If a is preferred to b at least as much as c is preferred to d with respect to each $g_i \in G$, then the comprehensive preference of a over b is not weaker than the comprehensive preference of c over d".

Dominance Relation for Pairs of Objects

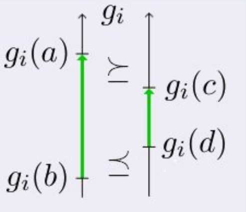
Marginal dominance relation D_2^i for pairs $(a,b),(c,d) \in B$

For cardinal criterion $g_i \in G$:

$$(a,b) D_2^i (c,d) \Leftrightarrow \Delta_i(a,b) \succeq \Delta_i(c,d)$$

For ordinal criterion $g_i \in G$:

$$(a,b) D_2^i (c,d) \Leftrightarrow g_i(a) \succeq g_i(c) \text{ and } g_i(d) \succeq g_i(b)$$



Dominance Relation for Pairs of Objects

Dominance relation D_2 for pairs $(a, b), (c, d) \in B$

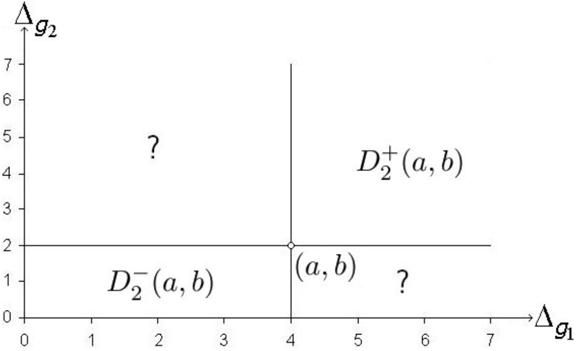
$$(a, b) D_2 (c, d) \text{ if } (a, b) D_2^i (c, d) \text{ for all } g_i \in G, \text{ i.e.,}$$

if a is preferred to b at least as much as c is preferred to d for all $g_i \in G$.

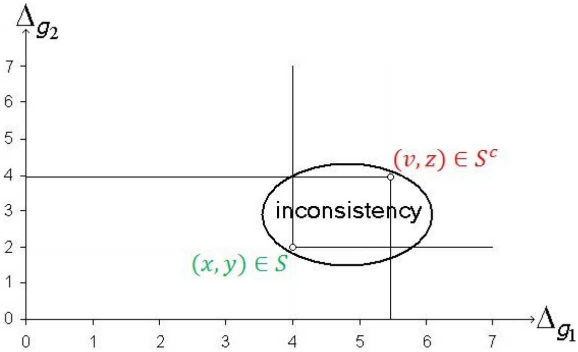
Dominance Cones

For a pair of objects $(a, b) \in B$:

positive dominance cone $D_2^+(a,b) = \{(c,d) \in B : (c,d) D_2 (a,b)\}$,
negative dominance cone $D_2^-(a,b) = \{(c,d) \in B : (a,b) D_2 (c,d)\}$.



Example of Inconsistent Preference Information



Inconsistency of the Preference Information

Preference information (pairwise comparisons of reference objects) may be inconsistent w.r.t. dominance relation D_2 due to:

- uncertainty of information – hesitation of the DM, unstable preferences,
- incomplete determination of the set G of criteria,
- granularity of information.

The inconsistency is handled using a dominance-based rough set approach. Before learning of a rule-based preference model of the DM, pairs of objects contained in a PCT are structured by calculation of lower approximations of S and S^c.

In this way, one restricts a priori the set of pairs of objects on which the preference model is build to a subset of sufficiently consistent pairs of objects. The goal is to obtain a reliable preference model.

Adaptation of DRSA – approximation of S and S^c

Lower approximations of S and S^c

$$\underline{S} = \{(a,b) \in B : D_2^+(a,b) \subseteq S\},$$
$$\underline{S^c} = \{(a,b) \in B : D_2^-(a,b) \subseteq S^c\}.$$

Upper approximations of S and S^c

$$\overline{S} = \bigcup_{(a,b)\in S} D_2^+(a,b),$$
$$\overline{S^c} = \bigcup_{(a,b)\in S^c} D_2^-(a,b).$$

Boundaries of S and S^c

$$Bn(S) = \overline{S} - \underline{S},$$
$$Bn(S^c) = \overline{S^c} - \underline{S^c}.$$

Adaptation of ϵ-VC-DRSA – approximation of S and S^c

Błaszczyński J, Greco S, Słowiński R, Szeląg M, Monotonic Variable Consistency Rough Set Approaches, International J. of Approximate Reasoning, 50(7), 2009, pp. 979–999.

Consistency is quantified using cost-type consistency measures $\epsilon_S, \epsilon_{S^c} : B \to [0,1]$, defined as:

$$\epsilon_S(a,b) = \frac{|D_2^+(a,b) \cap S^c|}{|S^c|}, \qquad \epsilon_{S^c}(a,b) = \frac{|D_2^-(a,b) \cap S|}{|S|}.$$

Parameterized lower approximations of S and S^c

$$\underline{S} = \{(a,b) \in S : \epsilon_S(a,b) \leq \theta_S\},$$
$$\underline{S^c} = \{(a,b) \in S^c : \epsilon_{S^c}(a,b) \leq \theta_{S^c}\},$$

where cost-type consistency thresholds $\theta_S, \theta_{S^c} \in [0,1)$.

Adaptation of ϵ-VC-DRSA – Positive Regions of S and S^c

Positive regions of relations S and S^c:

$$POS(S) = \bigcup_{(a,b) \in \underline{S}} D_2^+(a,b),$$

$$POS(S^c) = \bigcup_{(a,b) \in \underline{S^c}} D_2^-(a,b).$$

Positive regions defined above contain sufficiently consistent pairs of objects, i.e., pairs belonging to lower approximations of relation S or S^c, and can also contain some inconsistent pairs of objects which fall into dominance cones $D_2^+(\cdot,\cdot)$ or $D_2^-(\cdot,\cdot)$ originating in pairs of objects from lower approximations of relation S or S^c, respectively.

Adaptation of ϵ-VC-DRSA – Quality of Approximation

The coefficient

$$\gamma(S, S^c) = \frac{|\underline{S} \cup \underline{S^c}|}{|B|}$$

defines quality of approximation of S and S^c by set G.

$\gamma(S, S^c) \in [0, 1]$, and $\gamma(S, S^c) = 1$ indicates that the lower approximations of S and S^c contain all the pairs of objects from relations S and S^c, respectively.

Decision Rules

- Decision rules are induced in order to generalize description of sufficiently consistent pairs of objects from S_{PCT} (i.e., pairs of objects from parameterized lower approximations of S and S^c).
- Only minimal decision rules are considered. A decision rule suggesting assignment to S (S^c) is minimal, if there is no other rule suggesting assignment to S (resp. S^c), which has not stronger conditions and not worse consistency.
- Each rule is supported by at least one object from respective lower approximation and is allowed to cover only objects from respective positive region.

Decision rules constitute a preference model of the DM who gave the pairwise comparisons of reference objects.

Decision Rules

Decision rules are induced using VC-DomLEM[a] sequential covering algorithm, which generates minimal set of decision rules.

[a] Błaszczyński J, Słowiński R, Szeląg M, Sequential Covering Rule Induction Algorithm for Variable Consistency Rough Set Approaches, Information Sciences, 181, 2011, 987-1002.

Rule consistency is measured by cost-type rule consistency measure $\widehat{\epsilon}_T : R_T \rightarrow [0,1]$ defined as:

$$\widehat{\epsilon}_T(r_T) = \frac{\big|\|r_T\| \cap \neg T\big|}{|\neg T|},$$

where $T \in \{S, S^c\}$, R_T = set of rules suggesting assignment to relation T, $r_T \in R_T$, $\|r_T\|$ = the set of pairs of objects covered by r_T, $\neg T = B \setminus T$. For each $r_T \in R_T$, $\widehat{\epsilon}_T(r_T) \leq \theta_T$ has to hold.

Decision Rules

Exemplary S-decision rule (induced from $\underline{S}$):

If $\Delta_{maxSpeed}(a, b) \geq 25 \wedge$
$comfort(a) \geq 3 \wedge comfort(b) \leq 2$
then aSb.
"If car a has max speed at least 25 km/h greater than car b (cardinal criterion) and car a has comfort at least 3 while car b has comfort at most 2 (ordinal criterion), then car a is at least as good as car b".

Exemplary S^c-decision rule, (induced from $\underline{S^c}$):

If $\Delta_{maxSpeed}(a, b) \leq 20 \wedge$
$comfort(a) \leq 2 \wedge comfort(y) \geq 1$
then aS^cb.

As it can be seen from above, decision rules make use of ordinal properties of criteria only.

Application of Decision Rules

- Application of induced decision rules on set A of objects to be ranked yields a preference structure on A.
- Each pair of objects $(a, b) \in A \times A$ can be covered by some decision rules suggesting assignment to relation S and/or to relation S^c. It can also be not covered by any rule. In order to address these possibilities, two relations over set A, denoted by $\mathbb{S}$ and $\mathbb{S}^c$, are defined.
- Relations $\mathbb{S}$ and $\mathbb{S}^c$:
  - depend on adopted problem setting (s_{MCDA} or s_{ML}),
  - can be defined as crisp or valued relations,
  - can be defined differently when are valued relations.

Application of Decision Rules – Crisp Relations

s_{MCDA}

$$\mathbb{S} = \{(a,b) \in A \times A : (\exists\, r_S \in R_S : r_S \text{ covers } (a,b)) \textbf{ or } \ \mathsf{\textbf{a D b}}\},$$
$$\mathbb{S}^c = \{(a,b) \in A \times A : (\exists\, r_{S^c} \in R_{S^c} : r_{S^c} \text{ covers } (a,b)) \textbf{ and not } \ \mathsf{\textbf{a D b}}\}.$$

s_{ML}

$$\mathbb{S} = \{(a,b) \in A \times A : (\exists\, r_S \in R_S : r_S \text{ covers } (a,b)) \textbf{ or } \ \mathsf{\textbf{a = b}}\},$$
$$\mathbb{S}^c = \{(a,b) \in A \times A : (\exists\, r_{S^c} \in R_{S^c} : r_{S^c} \text{ covers } (a,b)) \textbf{ and not } \ \mathsf{\textbf{a = b}}\}.$$

Relation $\mathbb{S}$ is reflexive and relation $\mathbb{S}^c$ is irreflexive. Moreover, relations $\mathbb{S}$ and $\mathbb{S}^c$ are, in general, neither transitive nor complete.

Application of Decision Rules – Valued Relations

1. Each rule r_T covering pair (a, b) is treated as an argument (piece of evidence) for assignment of this pair to relation T.
2. Strength σ of each argument (rule r_T) defined as:
   (σ_1) $\sigma(r_T) = \big(1 - \widehat{\epsilon}_T(r_T)\big)$ ("credibility"), or
   (σ_2) $\sigma(r_T) = \big(1 - \widehat{\epsilon}_T(r_T)\big)cf(r_T)$ (product of "credibility" and coverage factor),

   where $cf(r_T)$ denotes coverage factor of rule r_T, defined as the ratio of the number of pairs of objects supporting r_T and the cardinality of relation T.
3. Aggregated strength of the arguments supporting assignment of pair (a, b) to relation T is calculated as maximum strength of these arguments.

Application of Decision Rules – Valued Relations

s_{MCDA}

$$\mathbb{S}(a,b) = \begin{cases} \max\{\sigma(r_S) : r_S \in R_S, r_S \text{ covers } (a,b)\}, & \textbf{if not} \quad \textbf{a D b} \\ 1, \ \textbf{if} \quad \textbf{a D b} \end{cases}$$

$$\mathbb{S}^c(a,b) = \begin{cases} \max\{\sigma(r_{S^c}) : r_{S^c} \in R_{S^c}, r_{S^c} \text{ covers } (a,b)\}, & \textbf{if not} \quad \textbf{a D b} \\ 0, \ \textbf{if} \quad \textbf{a D b} \end{cases}$$

s_{ML}

$$\mathbb{S}(a,b) = \begin{cases} \max\{\sigma(r_S) : r_S \in R_S, r_S \text{ covers } (a,b)\}, & \textbf{if not} \quad \textbf{a = b} \\ 1, \ \textbf{if} \quad \textbf{a = b} \end{cases}$$

$$\mathbb{S}^c(a,b) = \begin{cases} \max\{\sigma(r_{S^c}) : r_{S^c} \in R_{S^c}, r_{S^c} \text{ covers } (a,b)\}, & \textbf{if not} \quad \textbf{a = b} \\ 0, \ \textbf{if} \quad \textbf{a = b} \end{cases}$$

Relation $\mathbb{S}$ is reflexive and relation $\mathbb{S}^c$ is irreflexive.

Application of Decision Rules – $\mathbb{S}/\mathbb{S}^c$ summary

Six versions of VC-DRSA^{rank}

- VC-DRSA$^{rank}_{c\,0|1}$ – s_{MCDA}, $\mathbb{S}$ and $\mathbb{S}^c$ crisp,
- VC-DRSA$^{rank}_{c\,0\text{-}1_{cr}}$ – s_{MCDA}, $\mathbb{S}$ and $\mathbb{S}^c$ valued,
  value $\rightarrow$ max "credibility",
- VC-DRSA$^{rank}_{c\,0\text{-}1_{\times}}$ – s_{MCDA}, $\mathbb{S}$ and $\mathbb{S}^c$ valued,
  value $\rightarrow$ max "credibility" $\times$ coverage factor,

- VC-DRSA$^{rank}_{nc\,0|1}$ – s_{ML}, $\mathbb{S}$ and $\mathbb{S}^c$ crisp,
- VC-DRSA$^{rank}_{nc\,0\text{-}1_{cr}}$ – s_{ML}, $\mathbb{S}$ and $\mathbb{S}^c$ valued,
  value $\rightarrow$ max "credibility",,
- VC-DRSA$^{rank}_{nc\,0\text{-}1_{\times}}$ – s_{ML}, $\mathbb{S}$ and $\mathbb{S}^c$ valued,
  value $\rightarrow$ max "credibility" $\times$ coverage factor.

Application of Decision Rules

Relations $\mathbb{S}$ and $\mathbb{S}^c$ can be jointly represented by a directed multigraph $\mathcal{G}$ called preference graph. Each vertex (node) v_a of $\mathcal{G}$ corresponds to exactly one object $a \in A$. $\mathcal{G}$ contains two types of arcs: $\mathbb{S}$-arcs and $\mathbb{S}^c$-arcs.

In case of crisp relations, an $\mathbb{S}$-arc ($\mathbb{S}^c$-arc) from vertex v_a to vertex v_b indicates that $a\mathbb{S}b$ (resp. $a\mathbb{S}^c b$).

In case of valued relations, each $\mathbb{S}$-arc ($\mathbb{S}^c$-arc) from vertex v_a to vertex v_b is assigned the weight equal to $\mathbb{S}(a,b)$ (resp. $\mathbb{S}^c(a,b)$).

A final recommendation for the multicriteria ranking problem at hand, in terms of a total/partial preorder over set A, can be obtained upon a suitable exploitation of the preference graph.

Exploitation of Preference Graph

Two ways of exploitation of preference graph $\mathcal{G}$:

1. direct exploitation of relations $\mathbb{S}$ and $\mathbb{S}^c$ by the Net Flow Score (NFS) procedure that induces a total preorder over A by employing scoring function $S^{NF} : A \to \mathbb{R}$ defined as:

$$S^{NF}(a) = \sum_{b \in A \setminus \{a\}} \mathbb{S}(a,b) - \mathbb{S}(b,a) - \mathbb{S}^c(a,b) + \mathbb{S}^c(b,a)$$

2. transformation of preference graph $\mathcal{G}$ to another graph $\mathcal{G}'$ representing single valued relation $\mathcal{R}$ over set A, then exploitation of this relation using a ranking method (RM) $\succeq$, i.e., a function assigning a total or partial preorder $\succeq (A, R)$ over A to any finite set A and any valued relation R over A.

Exploitation of Preference Graph

Valued relation $\mathcal{R}$ is defined as:

$$\mathcal{R}(a,b) = \frac{\mathbb{S}(a,b) + (1 - \mathbb{S}^c(a,b))}{2},$$

where $a, b \in A$.

- Scoring function S^{NF} can be expressed in terms of $\mathcal{R}$ as:

$$S^{NF}(a) = 2\Big[\sum_{b \in A \setminus \{a\}} \mathcal{R}(a,b) - \mathcal{R}(b,a)\Big].$$

- Relation $\mathcal{R}$ is reflexive.
- If relations $\mathbb{S}$ and $\mathbb{S}^c$ are crisp, then $\mathcal{R}(a,b) \in \{0, \frac{1}{2}, 1\}$, for any $(a,b) \in A \times A$ – three-valued relation.

Literature Review of Ranking Methods

- Net Flow Rule (NFR) – yields a weak order using scoring function $SD : A \to \mathbb{R}$ defined as:
  $SD(a) = \sum_{b \in A \setminus \{a\}} \mathcal{R}(a,b) - \mathcal{R}(b,a)$.
- Iterative Net Flow Rule (It.NFR) – yields a weak order by iterative application of scoring function SD.
- Min in Favor (MiF) – yields a weak order using scoring function $mF : A \to \mathbb{R}$ defined as: $mF(a) = \min_{b \in A \setminus \{a\}} \mathcal{R}(a,b)$.
- Iterative Min in Favor (It.MiF) – yields a weak order by iterative application of scoring function mF.
- Leaving and Entering Flows (L/E) – yields a partial preorder being the intersection of two weak orders obtained using scoring functions SF and $-SA$, defined as:
  $SF(a) = \sum_{b \in A \setminus \{a\}} \mathcal{R}(a,b), \qquad -SA(a) = -\sum_{b \in A \setminus \{a\}} \mathcal{R}(b,a)$.

Desirable Properties of Ranking Methods

three-valued relation $\mathcal{R}$	**general relation $\mathcal{R}$**
neutrality (N)	neutrality (N)
monotonicity (M)	monotonicity (M)
covering compatibility (CC)	covering compatibility (CC)
discrimination (D)	independence of non-discriminating objects $(INDO)$
faithfulness (F)	independence of circuits (IC)
data-preservation (DP)	ordinality (O)
independence of non-discriminating objects $(INDO)$	continuity (C)
independence of circuits (IC)	faithfulness (F)
ordinality (O)	data-preservation (DP)
greatest-faithfulness (GF)	greatest-faithfulness (GF)

Given priority orders reflect relative importance of the properties.

http://www.cs.put.poznan.pl/mszelag/Research/rankingINS2014.pdf, str. 17-20

Desirable Properties of Ranking Methods

- (N) – a ranking method does not discriminate between objects just because of their labels (or, in other words, their order in the considered set A),
- (M) – improving an object cannot decrease its position in the ranking and, moreover, deteriorating an object cannot improve its position in the ranking,
- (CC) – when a "covers" b, b should not be ranked before a; in case of exploitation of valued relation $\mathcal{R}$, property CC of applied RM guaranties that the final ranking produced by this method respects dominance relation D over set A,
- (D) – for each set of objects A there exists at least one valued relation R over A such that the ranking obtained by a considered RM is a total order over set A,
- (F) – a RM applied to a weak order preserves it,

Desirable Properties of Ranking Methods

- (DP) – when it is possible to obtain a partial preorder on the basis of given transitive crisp relation without deleting information contained in this relation, a RM does so,
- $(INDO)$ – when there is a subset of objects that compare in the same way to all other objects, the ranking of the other objects is not affected by the presence of this subset,
- (IC) – the ranking is not affected by adding the same positive or negative value to the weights of all arcs in any cycle of $\mathcal{G}'$,
- (O) – ordinality implies that a RM should not make use of the "cardinal" properties of exploited valued relation,
- (C) – "small" changes in an exploited valued relation should not lead to radical changes in the final ranking produced by a RM,
- (GF) – if there are some greatest elements of a given set A, then the top-ranked objects should be chosen among them.

Desirable Properties of Ranking Methods – 3-valued $\mathcal{R}$

Property / RM	NFR	$It.NFR$	MiF	$It.MiF$	L/E
N	**T**	**T**	**T**	*T*	*T*
M	*T*	*F*	*T*	*F*	**T**
CC	**T**	**T**	*T*	*T*	*T*
D	*T*	*T*	*F*	*T*	*T*
F	**T**	**T**	**F**	**T**	*T*
DP	**T**	**T**	*T*	*T*	*T*
$INDO$	**T**	**T**	*F*	*F*	*T*
IC	**T**	*F*	**F**	*F*	*F*
O	**F**	*F*	**T**	**T**	*F*
GF	**F**	*F*	**T**	**T**	*T*

where:

T = presence of given property, F = lack of given property,
bold – proof in the literature

All considered ranking methods yield final ranking that respects the dominance relation on set A (since they have property CC).

Desirable Properties of Ranking Methods – arbitrary $\mathcal{R}$

Property / RM	NFR	$It.NFR$	MiF	$It.MiF$	L/E
N	**T**	**T**	**T**	*T*	*T*
M	*T*	*F*	*T*	*F*	**T**
CC	**T**	**T**	*T*	*T*	*T*
$INDO$	**T**	**T**	*F*	*F*	*T*
IC	**T**	*F*	**F**	*F*	*F*
O	**F**	*F*	**T**	**T**	*F*
C	**T**	*F*	**T**	**F**	*T*
F	**T**	**T**	**F**	**T**	*T*
DP	**T**	**T**	*T*	*T*	*T*
GF	**F**	*F*	**T**	**T**	*T*

where:

T = presence of given property, F = lack of given property,
bold – proof in the literature

All considered ranking methods yield final ranking that respects the dominance relation on set A (since they have property CC).

Desirable Properties of Ranking Methods – Conclusions

In view of the considered list of desirable properties, the best ranking method for exploitation of valued relation $\mathcal{R}$ is the Net Flow Rule method. This is because it satisfies most (eight out of ten) of the properties (which is, however, true also for the L/E ranking method) and, moreover, satisfies the first eight/five properties.

NFR ranking method is attractive also because it represents an intuitive way of reasoning about relative worth of objects in set A, as it takes into account both positive and negative arguments concerning each object (i.e., strength and weakness of each object).

Exploitation of relation $\mathcal{R}$ using NFR ranking method yields the same ranking (weak order) as direct exploitation of relations $\mathbb{S}$ and $\mathbb{S}^c$ using scoring function S^{NF}.

Analysis of the Final Ranking

Kendall rank correlation coefficient $\tau \in [-1, 1]$

$\tau(\succeq_{A^R}, \succeq_A)$ – measures rank correlation between 2 total preorders.

Modified Kendall rank correlation coefficient $\tau^{\neg I} \in [-1, 1]$

$\tau^{\neg I}(\succeq_{A^R}, \succeq_A)$ – measures rank correlation between two total preorders but does not take into account the pairs of objects $(a, b) \in A^R \times A^R$ such that a and b are considered indifferent according to the input preference information on A^R.

New concordance measure τ'[a] (generalizing τ)

[a] M. Szeląg, Application of the Dominance-based Rough Set Approach to Ranking and Similarity-based Classification Problems, Ph.D. th., 2015

$\tau'(S, S^c, \succeq_A)$ – measures concordance between pairwise comparisons in terms of S and S^c and final ranking being a partial preorder.

Illustrative Example

Notebooks

Experimental Verification of VC-DRSA^{rank}

Experimental Setup

- Comparisons of six variants of VC-DRSA^{rank} and SVM^{rank} method.
- In VC-DRSA^{rank}, exploitation of preference structure using NFR ranking method.
- Comparison on 14 ordinal classification problems of different data set consistency; results of SVM^{rank} could not be obtained for 3 data sets (marked by '(-)').
- To limit computational time, larger data sets were shrinked (preserving class distribution) to have at most around 350 objects (data sets marked in the table by suffix '$*$').
- Remark: 317 training objects results in around 100,000 pairwise comparisons!

Experimental Setup – Data Sets

- 10-fold stratified cross-validation (repeated 3 times).
- In each fold, preference information concerning training part A^R of each data set was obtained from ordinal classification, i.e., if class of a is not worse than class of b then aSb, otherwise aS^cb.
- For s_{MCDA}, if ordinal classification implied aS^cb but aDb, then the preference information was "**corrected**" by assuming aSb.
- In each fold, performance on test part A of each data set was measured in terms of $\tau(\succeq_A^i, \succeq_A^f)$ and $\tau^{\neg I}(\succeq_A^i, \succeq_A^f)$.
- Tested consistency thresholds:
  $\theta_S = \theta_{S^c}, \qquad \theta_S, \theta_{S^c} \in \{0, 0.01, 0.05, 0.1, 0.15\}$.
- Tested values of SVM^{rank}'s C parameter (trade-off between training error and margin):
  $C \in \{0.001, 0.01, 0.1, 1, 10\}$.

Experimental Setup – Data Sets

Table: Characteristics of data sets and average values of measure $\gamma(S, S^c)$ for $\theta_S = \theta_{S^c} = 0$ and not necessarily consistent set of criteria

Id	Data set	#Obj.	#Crit.	#Class.	$\overline{\gamma}(S, S^c)$
1	(-) car	324*	6	4	0.9732
2	housing	253*	13	4	0.9703
3	cpu	209	6	4	0.7545
4	denbosch	119	8	2	0.7291
5	bank-g	353*	16	2	0.7210
6	fame	332*	10	5	0.6454
7	(-) windsor	273*	10	4	0.6084
8	breast-w	350*	9	2	0.6048
9	balance-scale	313*	4	3	0.4886
10	ESL	244*	4	9	0.3360
11	(-) breast-c	286	7	2	0.2494
12	SWD	334*	10	4	0.1844
13	LEV	334*	4	5	0.1219
14	ERA	334*	4	9	0.0087

Experimental Results

Table: Performance in terms of **measure** τ

Data set	$V^{rank}_{c\,0\|1}$	$V^{rank}_{c\,0\text{-}1_{cr}}$	$V^{rank}_{c\,0\text{-}1_{\times}}$	$V^{rank}_{nc\,0\|1}$	$V^{rank}_{nc\,0\text{-}1_{cr}}$	$V^{rank}_{nc\,0\text{-}1_{\times}}$	SVM^{rank}
housing	**0.6727**$^{(2.5)}$ ±0.0433	**0.6727**$^{(2.5)}$ ±0.0433	0.6562$^{(6)}$ ±0.0560	**0.6727**$^{(2.5)}$ ±0.0433	**0.6727**$^{(2.5)}$ ±0.0433	0.6607$^{(5)}$ ±0.0567	0.6534$^{(7)}$ ±0.0523
cpu	**0.7873**$^{(1.5)}$ ±0.0155	0.7786$^{(6)}$ ±0.0147	0.7735$^{(7)}$ ±0.0154	**0.7873**$^{(1.5)}$ ±0.0155	0.7788$^{(5)}$ ±0.0147	0.7796$^{(4)}$ ±0.0114	0.7858$^{(3)}$ ±0.0061
denbosch	**0.5125**$^{(1.5)}$ ±0.1102	0.4774$^{(4)}$ ±0.0937	0.4570$^{(7)}$ ±0.0861	**0.5125**$^{(1.5)}$ ±0.1100	0.4792$^{(3)}$ ±0.0915	0.4754$^{(5)}$ ±0.0925	0.4747$^{(6)}$ ±0.0843
bank-g	**0.2696**$^{(1)}$ ±0.0344	0.2543$^{(4)}$ ±0.0286	0.2500$^{(6)}$ ±0.0293	0.2691$^{(2)}$ ±0.0342	0.2494$^{(7)}$ ±0.0318	0.2505$^{(5)}$ ±0.0289	0.2688$^{(3)}$ ±0.0191
fame	0.7097$^{(4)}$ ±0.0306	0.7070$^{(6)}$ ±0.0315	0.7030$^{(7)}$ ±0.0286	0.7097$^{(3)}$ ±0.0307	0.7072$^{(5)}$ ±0.0312	**0.7132**$^{(1)}$ ±0.0270	0.7131$^{(2)}$ ±0.0317
breast-w	**0.5387**$^{(1)}$ ±0.0458	*0.4839*$^{(4)}$ ±0.0097	*0.4696*$^{(6)}$ ±0.0062	0.5385$^{(2)}$ ±0.0458	0.5078$^{(3)}$ ±0.0219	*0.4819*$^{(5)}$ ±0.0178	*0.4678*$^{(7)}$ ±0.0078
balance-scale	**0.5787**$^{(1.5)}$ ±0.0210	0.5772$^{(3.5)}$ ±0.0224	0.5659$^{(7)}$ ±0.0206	**0.5787**$^{(1.5)}$ ±0.0210	0.5772$^{(3.5)}$ ±0.0224	0.5665$^{(6)}$ ±0.0200	0.5670$^{(5)}$ ±0.0226
ESL	**0.7650**$^{(1)}$ ±0.0446	0.7607$^{(3)}$ ±0.0416	0.7556$^{(7)}$ ±0.0351	0.7648$^{(2)}$ ±0.0370	0.7599$^{(4)}$ ±0.0374	0.7592$^{(5)}$ ±0.0374	0.7574$^{(6)}$ ±0.0403
SWD	0.4074$^{(3)}$ ±0.0934	0.4045$^{(6)}$ ±0.0938	0.4132$^{(2)}$ ±0.0965	0.4054$^{(4)}$ ±0.0954	0.4020$^{(7)}$ ±0.0945	**0.4157**$^{(1)}$ ±0.0967	0.4046$^{(5)}$ ±0.0986
LEV	0.5452$^{(5)}$ ±0.0717	0.5424$^{(7)}$ ±0.0713	0.5573$^{(3)}$ ±0.0734	0.5474$^{(4)}$ ±0.0719	0.5424$^{(6)}$ ±0.0751	**0.5634**$^{(1)}$ ±0.0789	0.5615$^{(2)}$ ±0.0753
ERA	0.3658$^{(6)}$ ±0.0946	0.3656$^{(7)}$ ±0.0936	0.3837$^{(3)}$ ±0.0901	0.3685$^{(4)}$ ±0.0919	0.3671$^{(5)}$ ±0.0934	0.3876$^{(2)}$ ±0.0892	**0.3976**$^{(1)}$ ±0.0871
avg rank (14)	2.57 (2^{nd})	4.68 (5^{th})	5.64 (6^{th})	**2.25** (1^{st})	4.21 (4^{th})	3.79 (3^{rd})	–
avg rank (11)	**2.55** (1^{st})	4.82 (5^{th})	5.55 (6^{th})	**2.55** (1^{st})	4.64 (4^{th})	3.64 (2^{nd})	4.27 (3^{rd})

Experimental Results

Table: Performance in terms of **measure** $\tau^{\neg \mathbf{I}}$

Data set	$V^{rank}_{c\,0\|1}$	$V^{rank}_{c\,0\text{-}1_{cr}}$	$V^{rank}_{c\,0\text{-}1_{\times}}$	$V^{rank}_{nc\,0\|1}$	$V^{rank}_{nc\,0\text{-}1_{cr}}$	$V^{rank}_{nc\,0\text{-}1_{\times}}$	SVM^{rank}
housing	$\mathbf{0.8566}^{(2.5)}$ ±0.0538	$\mathbf{0.8566}^{(2.5)}$ ±0.0538	$0.8418^{(6)}$ ±0.0721	$\mathbf{0.8566}^{(2.5)}$ ±0.0538	$\mathbf{0.8566}^{(2.5)}$ ±0.0538	$0.8475^{(5)}$ ±0.0729	$0.8382^{(7)}$ ±0.0673
cpu	$\mathit{0.9866}^{(5.5)}$ ±0.0211	$\mathit{0.9888}^{(3.5)}$ ±0.0184	$\mathit{0.9823}^{(7)}$ ±0.0187	$\mathit{0.9866}^{(5.5)}$ ±0.0211	$\mathit{0.9888}^{(3.5)}$ ±0.0184	$\mathit{0.9897}^{(2)}$ ±0.0139	$\mathbf{0.9980}^{(1)}$ ±0.0064
denbosch	$0.8485^{(6)}$ ±0.1701	$0.8533^{(3)}$ ±0.1262	$0.8378^{(7)}$ ±0.1579	$0.8494^{(5)}$ ±0.1687	$0.8500^{(4)}$ ±0.1695	$\mathbf{0.8715}^{(1)}$ ±0.1697	$0.8704^{(2)}$ ±0.1546
bank-g	$\mathit{0.9064}^{(4)}$ ±0.0989	$\mathit{0.9055}^{(5.5)}$ ±0.0986	$\mathit{0.9256}^{(3)}$ ±0.0908	$\mathit{0.9047}^{(7)}$ ±0.1042	$\mathit{0.9055}^{(5.5)}$ ±0.1015	$\mathit{0.9272}^{(2)}$ ±0.0893	$\mathbf{0.9970}^{(1)}$ ±0.0142
fame	$0.8769^{(6)}$ ±0.0381	$0.8778^{(4)}$ ±0.0392	$0.8728^{(7)}$ ±0.0362	$0.8772^{(5)}$ ±0.0382	$0.8780^{(3)}$ ±0.0388	$\mathbf{0.8855}^{(1)}$ ±0.0338	$0.8850^{(2)}$ ±0.0394
breast-w	$0.9952^{(4.5)}$ ±0.0095	$0.9952^{(4.5)}$ ±0.0096	$\mathbf{0.9957}^{(1)}$ ±0.0090	$0.9952^{(4.5)}$ ±0.0095	$0.9952^{(4.5)}$ ±0.0094	$0.9954^{(2)}$ ±0.0086	$0.9923^{(7)}$ ±0.0141
balance-scale	$\mathbf{0.9637}^{(1.5)}$ ±0.0319	$0.9635^{(3)}$ ±0.0313	$0.9614^{(7)}$ ±0.0318	$\mathbf{0.9637}^{(1.5)}$ ±0.0319	$0.9631^{(4)}$ ±0.0318	$0.9624^{(6)}$ ±0.0304	$0.9630^{(5)}$ ±0.0299
ESL	$0.9089^{(3)}$ ±0.0446	$\mathbf{0.9101}^{(1)}$ ±0.0443	$0.9041^{(7)}$ ±0.0366	$0.9086^{(4)}$ ±0.0447	$0.9093^{(2)}$ ±0.0398	$0.9085^{(5)}$ ±0.0396	$0.9062^{(6)}$ ±0.0436
SWD	$0.5805^{(5)}$ ±0.1359	$0.5807^{(4)}$ ±0.1359	$0.5933^{(2)}$ ±0.1397	$0.5770^{(7)}$ ±0.1367	$0.5772^{(6)}$ ±0.1369	$\mathbf{0.5970}^{(1)}$ ±0.1400	$0.5810^{(3)}$ ±0.1426
LEV	$0.7317^{(6)}$ ±0.0951	$0.7322^{(5)}$ ±0.0955	$0.7526^{(3)}$ ±0.0983	$0.7289^{(7)}$ ±0.0952	$0.7323^{(4)}$ ±0.1009	$\mathbf{0.7609}^{(1)}$ ±0.1059	$0.7583^{(2)}$ ±0.1011
ERA	$0.4075^{(7)}$ ±0.1057	$0.4084^{(6)}$ ±0.1046	$0.4288^{(3)}$ ±0.1005	$0.4108^{(4)}$ ±0.1030	$0.4101^{(5)}$ ±0.1045	$0.4332^{(2)}$ ±0.1000	$\mathbf{0.4445}^{(1)}$ ±0.0969
avg rank (14)	4.43 (4^{th})	3.61 (2^{nd})	4.71 (5^{th})	4.43 (4^{th})	3.82 (3^{rd})	**2.86** (1^{st})	–
avg rank (11)	4.64 (5^{th})	3.77 (3^{rd})	4.82 (6^{th})	4.82 (6^{th})	4.05 (4^{th})	**2.55** (1^{st})	3.36 (2^{nd})

Table: Best parameter values for the six versions of VC-DRSA^{rank} (in short V^{rank}) and for SVM^{rank} – performance measured using τ

Data set	$V^{rank}_{c\,0\|1}$	$V^{rank}_{c\,0\text{-}1_{cr}}$	$V^{rank}_{c\,0\text{-}1_{\times}}$	$V^{rank}_{nc\,0\|1}$	$V^{rank}_{nc\,0\text{-}1_{cr}}$	$V^{rank}_{nc\,0\text{-}1_{\times}}$	SVM^{rank}
(-) car	0.1	0	0.1	0.1	0	0.1	–
housing	0	0	0.01	0	0	0.01	0.1
cpu	0.05	0.05	0.05	0.05	0.05	0.01	0.1
denbosch	0.01	0	0.05	0.01	0	0.01	0.01
bank-g	0.01	0	0.01	0.01	0	0.01	0.001
fame	0.01	0.01	0.01	0.01	0.01	0.01	0.001
(-) windsor	0.01	0	0.05	0.01	0.01	0.01	–
breast-w	0.01	0	0.1	0.01	0	0	0.001
balance-scale	0.05	0	0.15	0.05	0	0	1
ESL	0.01	0.01	0.15	0.15	0.15	0.15	1
(-) breast-c	0.1	0	0.15	0	0	0	–
SWD	0.01	0.01	0.1	0.01	0.01	0.01	0.001
LEV	0.01	0.01	0.1	0.15	0.15	0.1	10
ERA	0.01	0.01	0.1	0.01	0.01	0.1	0.01

Table: Best parameter values for the six versions of VC-DRSA^{rank} (in short V^{rank}) and for SVM^{rank} – performance measured using $\tau^{\neg\mathbf{I}}$

Data set	$V^{rank}_{c\,0\vert 1}$	$V^{rank}_{c\,0\text{-}1_{cr}}$	$V^{rank}_{c\,0\text{-}1_{\times}}$	$V^{rank}_{nc\,0\vert 1}$	$V^{rank}_{nc\,0\text{-}1_{cr}}$	$V^{rank}_{nc\,0\text{-}1_{\times}}$	SVM^{rank}
(-) car	0.01	0.01	0.1	0.01	0.01	0.01	–
housing	0	0	0.01	0	0	0.01	0.1
cpu	0.05	0.05	0.05	0.05	0.05	0.01	0.1
denbosch	0.01	0.05	0.05	0.01	0.01	0.01	0.01
bank-g	0.05	0.05	0.01	0.01	0.01	0.01	0.1
fame	0.01	0.01	0.01	0.01	0.01	0.01	0.001
(-) windsor	0.01	0.01	0.05	0.01	0.01	0.01	–
breast-w	0	0	0.1	0.1	0.1	0.1	0.001
balance-scale	0.05	0.1	0.15	0.05	0.1	0	1
ESL	0.01	0.01	0.15	0.01	0.15	0.15	1
(-) breast-c	0.1	0.1	0.15	0.15	0.15	0.15	–
SWD	0.01	0.01	0.1	0.01	0.01	0.01	0.001
LEV	0.01	0.01	0.1	0.01	0.15	0.1	10
ERA	0.01	0.01	0.1	0.01	0.01	0.05	0.01

Experimental Results – Most Important Conclusions

- VC-DRSA^{rank} is highly competitive to SVM^{rank}. Considering its wider applicability (all 14 data sets), and interpretability of decision rules, it seems to be more attractive for a DM.
- Arguably, values of $\tau^{\neg I}$, directly addressing correct prediction of preference and inverse preference relations, should be considered more important than values of τ.
- The choice of the best version of VC-DRSA^{rank} depends on the chosen performance measure:
    - $\tau \rightarrow$ "crisp" versions VC-DRSA$^{rank}_{c\,0|1}$ and VC-DRSA$^{rank}_{nc\,0|1}$,
    - $\tau^{\neg I} \rightarrow$ "valued" version VC-DRSA$^{rank}_{nc\,0\text{-}1_\times}$,
- The version VC-DRSA$^{rank}_{c\,0\text{-}1_\times}$ is systematically (i.e., for both performance measures) the worst version of VC-DRSA^{rank} $\rightarrow$ not recommended.
- Employing ϵ-VC-DRSA improves performance, especially in terms of $\tau^{\neg I}$ – in most of the cases the largest avg. value was obtained for $\theta_S = \theta_{S^c} > 0$.

Summary and Conclusions

Summary and Conclusions

- VC-DRSA is a flexible modeling method that allows to include domain knowledge and handles inconsistencies in data.
- VC-DRSA allows to work with heterogeneous attributes – nominal, ordinal, and cardinal (no need of discretization).
- Preference information in terms of pairwise comparisons of some reference objects is relatively easy to elicit from the DM,
- Presented methodology involves non-statistical processing of preference information and induction of decision rules from decision examples (pairwise comparisons of reference objects).
- Applied rule preference model has many advantages, e.g., comprehensibility, generality, lack of aggregation operators.
- Net Flow Rule appears to be the best ranking method for exploitation of a valued relation over a set of objects.
- Concordance with the current trend in MCDA which consists in induction of preference model from decision examples.

Summary and Conclusions (2)

- Presented approach to preference learning in multicriteria ranking is competitive to state-of-the-art SVM^{rank}.
- By adaptation of ϵ-VC-DRSA, it was possible to obtain better average values of applied performance measures than in case of adapting classical DRSA.
- According to measure τ, the "crisp" versions of VC-DRSA^{rank}, i.e., VC-DRSA$^{rank}_{c\,0|1}$ and VC-DRSA$^{rank}_{nc\,0|1}$, obtained in the experiment the best (i.e., the lowest) average ranks over 11 data sets.
- "Valued" version VC-DRSA$^{rank}_{nc\,0\text{-}1_{\times}}$ obtained the lowest average rank with respect to measure $\tau^{\neg I}$.

References

- Bouyssou D, Vincke Ph, Ranking alternatives on the basis of preference relations: A progress report with special emphasis on outranking relations, Journ. of MCDA, 6, 1997, pp. 77–85.
- Greco S, Matarazzo B, Słowiński R, Tsoukiàs A, Exploitation of a Rough Approximation of the Outranking Relation in Multicriteria Choice and Ranking, Lecture Notes in Economics and Mathematical Systems, 465, 1998, pp. 45–60.
- Fortemps Ph, Greco S, Słowiński R, Multicriteria decision support using rules that represent rough-graded preference relations, EJOR, 188 (2008) 206-223.
- Szeląg M, Greco S, Słowiński R, Variable Consistency Dominance-Based Rough Set Approach to Preference Learning in Multicriteria Ranking, INS, 277, 2014, pp. 525-552.
- M. Szeląg, Application of the Dominance-based Rough Set Approach to Ranking and Similarity-based Classification Problems, Ph.D. thesis, PUT, 2015.

Questions and Discussion

Thank you for your attention.