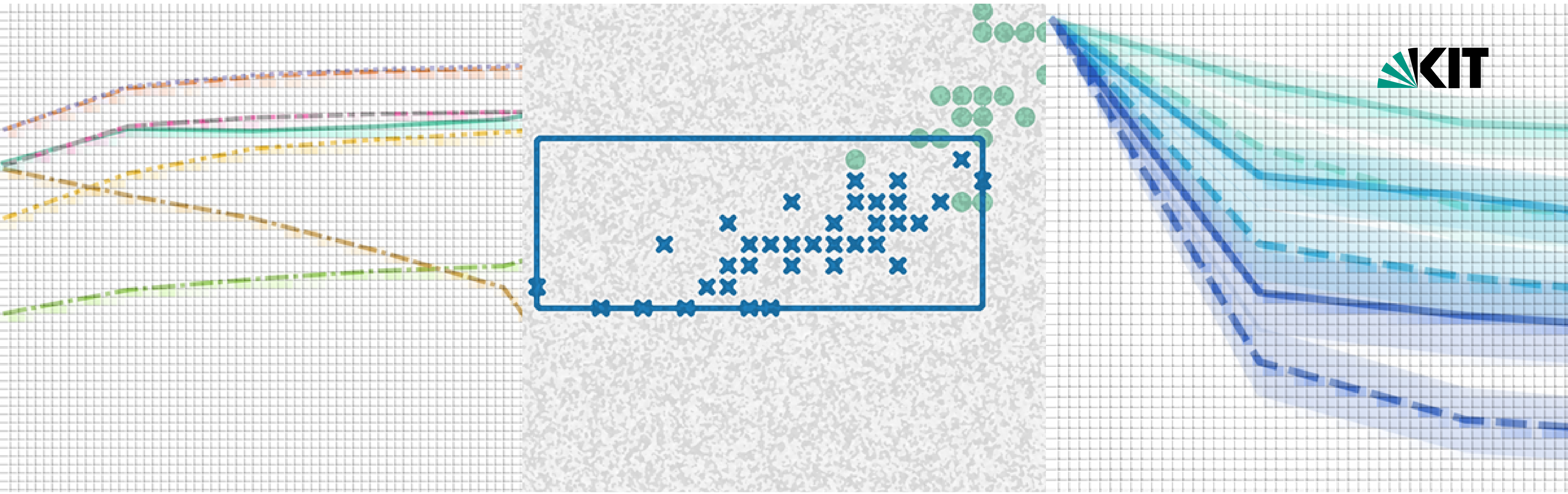


Subgroup Discovery with Small and Alternative Feature Sets

SIGMOD 2025 | Berlin

Jakob Bach | June 24, 2025

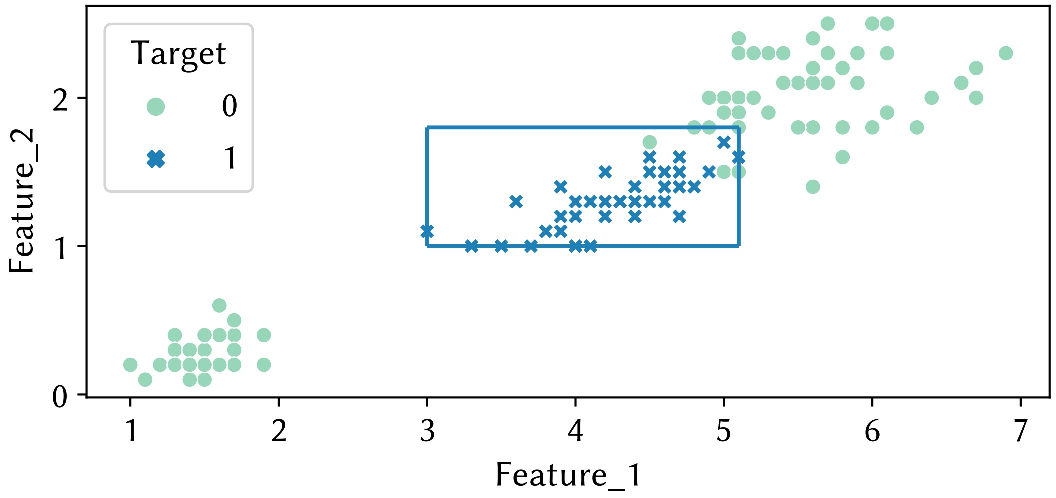
Subgroup Discovery

- **Subgroup discovery:** “Identifying descriptions of subsets of a dataset that show an interesting behavior” [1]
  - Language for descriptions typically simple
  - ‘Interesting’ according to some objective function

Subgroup Discovery

- **Subgroup discovery:** "Identifying descriptions of subsets of a dataset that show an interesting behavior" [1]
  - Language for descriptions typically simple
  - 'Interesting' according to some objective function

- **Our scope:** Binary classification with real-valued features
  - Tabular dataset $X \in \mathbb{R}^{m \times n}$ (data objects $\times$ features)
  - Prediction target $y \in \{0, 1\}^m$ ('interesting'/'positive' = 1)
  - Subgroup description: Hyperrectangle
  - Subgroup quality: Weighted Relative Accuracy
    - WRAcc $= \frac{m_b}{m} \cdot \left(\frac{m_b^+}{m_b} - \frac{m^+}{m} \right)$ [22]
    - $+ \leftrightarrow$ positive data object, $b \leftrightarrow$ in subgroup (box)



Contributions

- **Our focus**: Two constraint types to foster interpretability of subgroup descriptions
  - Limit number of features used in subgroup description
  - Find alternative subgroup descriptions

Contributions

- **Our focus**: Two constraint types to foster interpretability of subgroup descriptions
  - Limit number of features used in subgroup description
  - Find alternative subgroup descriptions

- **Contributions:**
  - Formalize subgroup discovery as an SMT (Satisfiability Modulo Theories) optimization problem
  - Formalize two constraint types:
    - Feature-cardinality constraints
    - Alternative subgroup descriptions
  - Analyze computational complexity and show $\mathcal{NP}$-hardness
  - Comprehensive experiments

Formalization – Basic Problem

$$lb, ub \in \{\mathbb{R} \cup \{-\infty, +\infty\}\}^n$$ (Variables: lower/upper bounds of subgroup)

Formalization – Basic Problem

$\forall j \in \{1, \dots, n\}$	$lb_j \leq ub_j$	(Constraint: relationship between bounds)
	$lb, ub \in \{\mathbb{R} \cup \{-\infty, +\infty\}\}^n$	(Variables: lower/upper bounds of subgroup)

Formalization – Basic Problem

$$\forall i \in \{1, \ldots, m\} \quad b_i \leftrightarrow \bigwedge_{j \in \{1,\ldots,n\}} \left((X_{ij} \geq lb_j) \land (X_{ij} \leq ub_j) \right) \quad \text{(i-th data object in subgroup?)}$$

$$\forall j \in \{1, \ldots, n\} \quad lb_j \leq ub_j \quad \text{(Constraint: relationship between bounds)}$$

$$b \in \{0, 1\}^m \quad \text{(Auxiliary variables: subgroup membership)}$$

$$lb, ub \in \{\mathbb{R} \cup \{-\infty, +\infty\}\}^n \quad \text{(Variables: lower/upper bounds of subgroup)}$$

Formalization – Basic Problem

$$
\begin{array}{llr}
\text{s.t.:} & m_b := \sum_{i=1}^{m} b_i \quad \text{and} \quad m_b^+ := \sum_{\substack{i \in \{1,\dots,m\} \\ y_i = 1}} b_i & \text{(Number of data objects in subgroup)} \\
\forall i \in \{1, \dots, m\} & b_i \leftrightarrow \bigwedge_{j \in \{1,\dots,n\}} \left((X_{ij} \geq lb_j) \land (X_{ij} \leq ub_j) \right) & \text{(i-th data object in subgroup?)} \\
\forall j \in \{1, \dots, n\} & lb_j \leq ub_j & \text{(Constraint: relationship between bounds)} \\
 & b \in \{0, 1\}^m & \text{(Auxiliary variables: subgroup membership)} \\
 & lb, ub \in \{\mathbb{R} \cup \{-\infty, +\infty\}\}^n & \text{(Variables: lower/upper bounds of subgroup)}
\end{array}
$$

Formalization – Basic Problem

$$
\begin{array}{llr}
\max & Q_{\text{WRAcc}} = \frac{m_b^+}{m} - \frac{m_b \cdot m^+}{m^2} & \text{(Objective: subgroup quality)} \\
\text{s.t.:} & m_b := \sum_{i=1}^{m} b_i \quad \text{and} \quad m_b^+ := \sum_{\substack{i \in \{1,\dots,m\} \\ y_i = 1}} b_i & \text{(Number of data objects in subgroup)} \\
\forall i \in \{1, \dots, m\} & b_i \leftrightarrow \bigwedge_{j \in \{1,\dots,n\}} \left((X_{ij} \geq lb_j) \land (X_{ij} \leq ub_j) \right) & \text{(i-th data object in subgroup?)} \\
\forall j \in \{1, \dots, n\} & lb_j \leq ub_j & \text{(Constraint: relationship between bounds)} \\
 & b \in \{0, 1\}^m & \text{(Auxiliary variables: subgroup membership)} \\
 & lb, ub \in \{\mathbb{R} \cup \{-\infty, +\infty\}\}^n & \text{(Variables: lower/upper bounds of subgroup)}
\end{array}
$$

Formalization – Cardinality Constraints

- **Concept:** Limit number of features used (= selected) in subgroup description to $k \in \mathbb{N}$
- **Formally:** Limit number of features whose bounds exclude at least one data object from subgroup

Formalization – Cardinality Constraints

- **Concept:** Limit number of features used (= selected) in subgroup description to $k \in \mathbb{N}$
- **Formally:** Limit number of features whose bounds exclude at least one data object from subgroup

$$s, s^{\text{lb}}, s^{\text{ub}} \in \{0, 1\}^n$$

Formalization – Cardinality Constraints

- **Concept:** Limit number of features used (= selected) in subgroup description to $k \in \mathbb{N}$
- **Formally:** Limit number of features whose bounds exclude at least one data object from subgroup

$$\forall j \in \{1, \ldots, n\}: \qquad s_j^{\text{lb}} \leftrightarrow \left(lb_j > \min_{i \in \{1,\ldots,m\}} X_{ij} \right)$$

$$\forall j \in \{1, \ldots, n\}: \qquad s_j^{\text{ub}} \leftrightarrow \left(ub_j < \max_{i \in \{1,\ldots,m\}} X_{ij} \right)$$

$$s, s^{\text{lb}}, s^{\text{ub}} \in \{0, 1\}^n$$

Formalization – Cardinality Constraints

- **Concept:** Limit number of features used (= selected) in subgroup description to $k \in \mathbb{N}$
- **Formally:** Limit number of features whose bounds exclude at least one data object from subgroup

$$\forall j \in \{1, \dots, n\}: \quad s_j^{\text{lb}} \leftrightarrow \left(lb_j > \min_{i \in \{1, \dots, m\}} X_{ij} \right)$$

$$\forall j \in \{1, \dots, n\}: \quad s_j^{\text{ub}} \leftrightarrow \left(ub_j < \max_{i \in \{1, \dots, m\}} X_{ij} \right)$$

$$\forall j \in \{1, \dots, n\}: \quad s_j \leftrightarrow \left(s_j^{\text{lb}} \lor s_j^{\text{ub}} \right)$$

$$s, s^{\text{lb}}, s^{\text{ub}} \in \{0, 1\}^n$$

Formalization – Cardinality Constraints

- **Concept:** Limit number of features used (= selected) in subgroup description to $k \in \mathbb{N}$
- **Formally:** Limit number of features whose bounds exclude at least one data object from subgroup

$$\begin{aligned}
\forall j \in \{1, \dots, n\}: \quad & s_j^{\text{lb}} \leftrightarrow \left(lb_j > \min_{i \in \{1, \dots, m\}} X_{ij} \right) \\
\forall j \in \{1, \dots, n\}: \quad & s_j^{\text{ub}} \leftrightarrow \left(ub_j < \max_{i \in \{1, \dots, m\}} X_{ij} \right) \\
\forall j \in \{1, \dots, n\}: \quad & s_j \leftrightarrow \left(s_j^{\text{lb}} \lor s_j^{\text{ub}} \right) \\
& \sum_{j=1}^{n} s_j \leq k \\
& s, s^{\text{lb}}, s^{\text{ub}} \in \{0, 1\}^n
\end{aligned}$$

Formalization – Alternative Descriptions

- **Concept:** Cover similar set of data objects as a given subgroup with different set of selected features
  - Repeat sequentially to get $a \in \mathbb{N}$ alternatives for dissimilarity threshold $\tau \in \mathbb{R}_{\geq 0}$

Formalization – Alternative Descriptions

- **Concept:** Cover similar set of data objects as a given subgroup with different set of selected features
  - Repeat sequentially to get $a \in \mathbb{N}$ alternatives for dissimilarity threshold $\tau \in \mathbb{R}_{\geq 0}$

- **Chosen optimization objective:** Maximize normalized Hamming similarity (= prediction accuracy)

$$\text{sim}_{\text{nHamm}}(b^{(a)}, b^{(0)}) = \frac{1}{m} \cdot \sum_{i=1}^{m} \left(b_i^{(a)} \leftrightarrow b_i^{(0)} \right) = \frac{1}{m} \cdot \left(\sum_{\substack{i \in \{1,\dots,m\} \\ b_i^{(0)}=1}} b_i^{(a)} + \sum_{\substack{i \in \{1,\dots,m\} \\ b_i^{(0)}=0}} \neg b_i^{(a)} \right)$$

Formalization – Alternative Descriptions

- **Concept:** Cover similar set of data objects as a given subgroup with different set of selected features
  - Repeat sequentially to get $a \in \mathbb{N}$ alternatives for dissimilarity threshold $\tau \in \mathbb{R}_{\geq 0}$

- **Chosen optimization objective:** Maximize normalized Hamming similarity (= prediction accuracy)

$$\text{sim}_{\text{nHamm}}(b^{(a)}, b^{(0)}) = \frac{1}{m} \cdot \sum_{i=1}^{m} \left(b_i^{(a)} \leftrightarrow b_i^{(0)} \right) = \frac{1}{m} \cdot \left(\sum_{\substack{i \in \{1,\dots,m\} \\ b_i^{(0)}=1}} b_i^{(a)} + \sum_{\substack{i \in \{1,\dots,m\} \\ b_i^{(0)}=0}} \neg b_i^{(a)} \right)$$

- **Chosen dissimilarity constraints:** From each existing feature set, deselect at least $\tau_{\text{abs}} \in \mathbb{N}$ (but $\leq k$) features

$$\forall l \in \{0, \dots, a-1\}: \ \text{dis}_{\text{des}}(s^{(a)}, s^{(l)}) = \sum_{\substack{j \in \{1,\dots,n\} \\ s_j^{(l)}=1}} \neg s_j^{(a)} \geq \min\left(\tau_{\text{abs}},\ k^{(l)}\right)$$

Experimental Design

- **27 binary-classification datasets** (with 106–9822 data objects and 20–168 features) from *PMLB* [31, 33]

Experimental Design

- **27 binary-classification datasets** (with 106–9822 data objects and 20–168 features) from *PMLB* [31, 33]

- **Eight subgroup-discovery methods:**
  - Solver-based (novel): *SMT* (using *Z3* [8, 13] as optimizer)
  - Exhaustive* algorithms (related work): *BSD* [24], *SD-Map* [2] (* both require discretized features)
  - Heuristics (related work): *Beam*, *BI* [28], *PRIM* [15]
  - Baselines (novel): *MORS* (Minimal Optimal Recall Subgroup), *Random*

Experimental Design

- **27 binary-classification datasets** (with 106–9822 data objects and 20–168 features) from *PMLB* [31, 33]

- **Eight subgroup-discovery methods:**
  - Solver-based (novel): *SMT* (using *Z3* [8, 13] as optimizer)
  - Exhaustive* algorithms (related work): *BSD* [24], *SD-Map* [2] (* both require discretized features)
  - Heuristics (related work): *Beam*, *BI* [28], *PRIM* [15]
  - Baselines (novel): *MORS* (Minimal Optimal Recall Subgroup), *Random*

- **Experimental scenarios:**
  - Solver timeouts: {1 s, 2 s, 4 s, . . . , 2048 s}
  - Feature-cardinality constraints: $k \in \{1, 2, 3, 4, 5\}$, 'no' (unconstrained)
  - Alternative subgroup descriptions: $k = 3$ features, $a = 5$ alternatives, and dissimilarity threshold $\tau_{\text{abs}} \in \{1, 2, 3\}$

Experimental Design

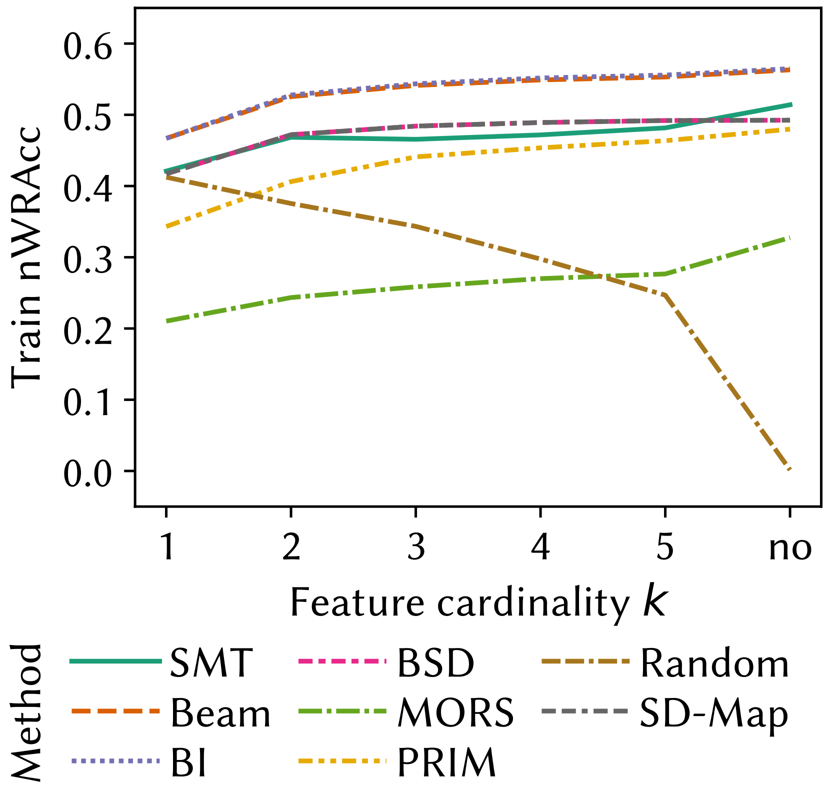
- **27 binary-classification datasets** (with 106–9822 data objects and 20–168 features) from *PMLB* [31, 33]

- **Eight subgroup-discovery methods:**
  - Solver-based (novel): *SMT* (using *Z3* [8, 13] as optimizer)
  - Exhaustive* algorithms (related work): *BSD* [24], *SD-Map* [2] (* both require discretized features)
  - Heuristics (related work): *Beam*, *BI* [28], *PRIM* [15]
  - Baselines (novel): *MORS* (Minimal Optimal Recall Subgroup), *Random*

- **Experimental scenarios:**
  - Solver timeouts: {1 s, 2 s, 4 s, . . . , 2048 s}
  - Feature-cardinality constraints: $k \in \{1, 2, 3, 4, 5\}$, 'no' (unconstrained)
  - Alternative subgroup descriptions: $k = 3$ features, $a = 5$ alternatives, and dissimilarity threshold $\tau_{\text{abs}} \in \{1, 2, 3\}$

- **Evaluation metrics:**
  - Subgroup quality (nWRAcc [22, 29])
  - Runtime
  - For alternatives: Similarity [12] (Normalized Hamming and Jaccard)

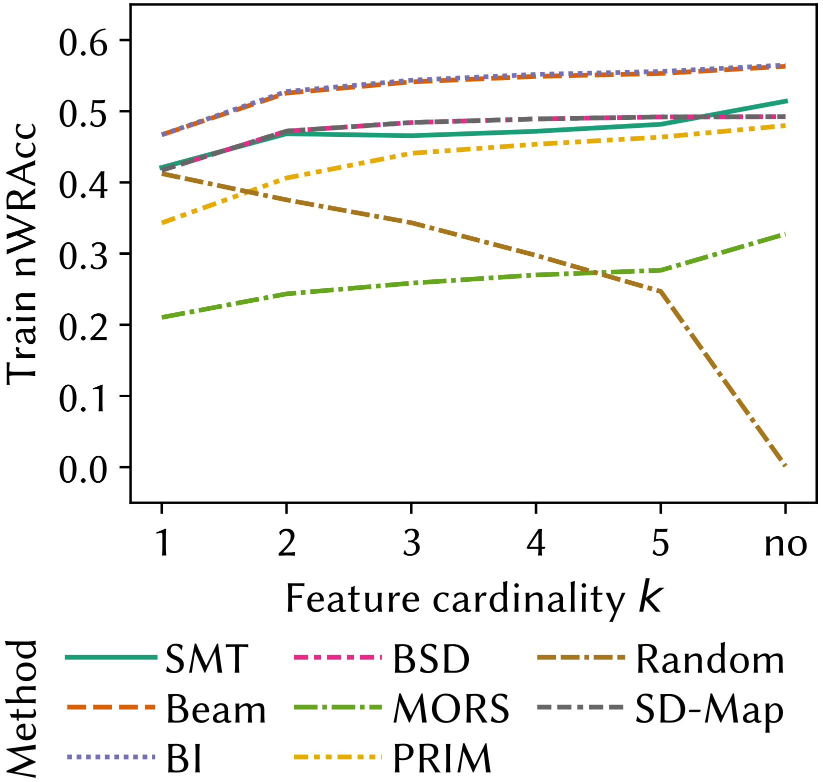
Results – Cardinality Constraints



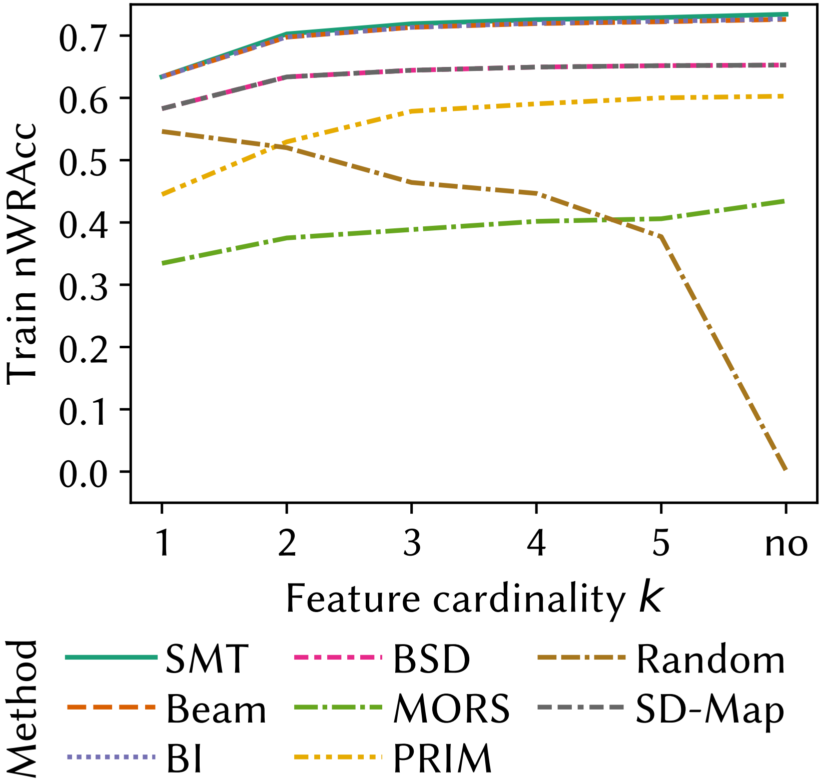
(a) All 27 datasets.

Subgroup quality of original subgroups.

Results – Cardinality Constraints



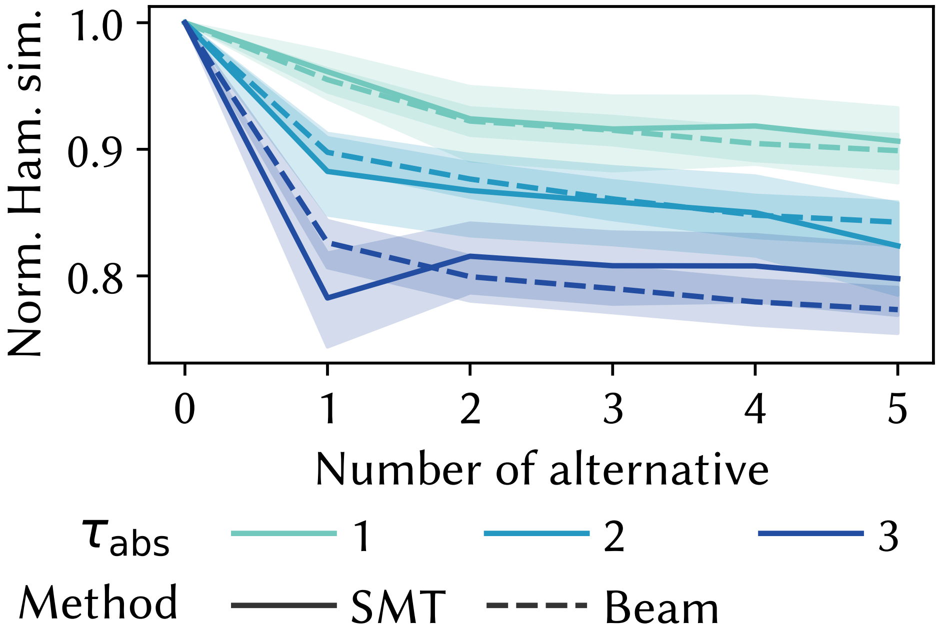
(a) All 27 datasets.



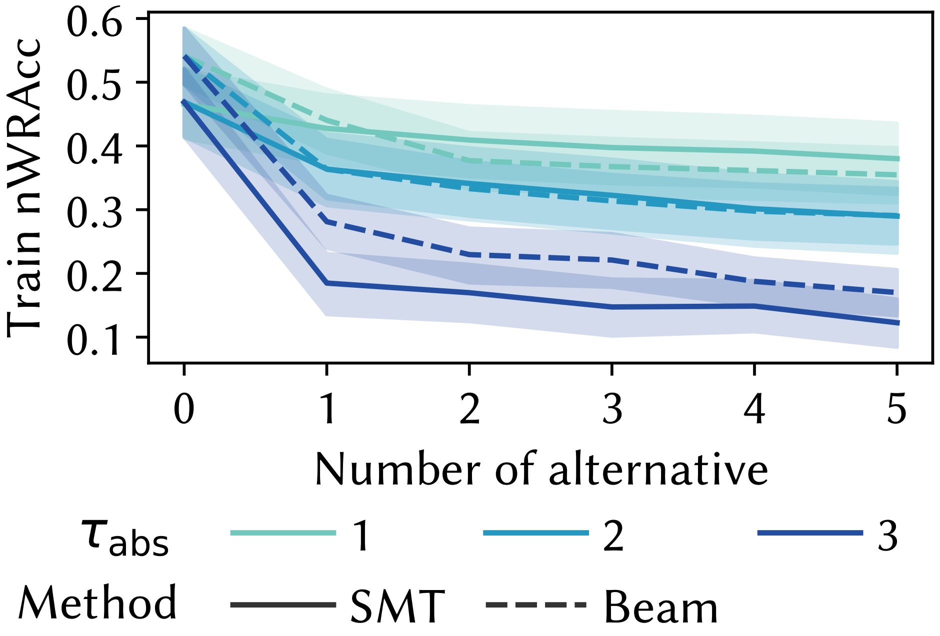
(b) 13 datasets without *SMT* timeouts.

Subgroup quality of original subgroups.

Results – Alternative Descriptions



(a) Subgroup similarity.



(b) Subgroup quality.

Similarity and quality of alternative subgroup descriptions.

Summary

- **Focus:** Two constraint types to foster interpretability of subgroup descriptions
  - Limit number of features used in subgroup description
  - Find alternative subgroup descriptions

Summary

- **Focus:** Two constraint types to foster interpretability of subgroup descriptions
  - Limit number of features used in subgroup description
  - Find alternative subgroup descriptions

- **Contributions:** Formalization (as SMT problem), complexity analyses, and experiments

Summary

- **Focus:** Two constraint types to foster interpretability of subgroup descriptions
  - Limit number of features used in subgroup description
  - Find alternative subgroup descriptions

- **Contributions:** Formalization (as SMT problem), complexity analyses, and experiments

- **Experimental results:**
  - Heuristic search methods yield close-to-optimal subgroup quality in short time
  - Using few features in subgroup descriptions suffices to reach close-to-optimal subgroup quality
  - Different subgroup descriptions (used features) may capture similar subgroup (set of data objects)

Summary

- **Focus:** Two constraint types to foster interpretability of subgroup descriptions
  - Limit number of features used in subgroup description
  - Find alternative subgroup descriptions

- **Contributions:** Formalization (as SMT problem), complexity analyses, and experiments

- **Experimental results:**
  - Heuristic search methods yield close-to-optimal subgroup quality in short time
  - Using few features in subgroup descriptions suffices to reach close-to-optimal subgroup quality
  - Different subgroup descriptions (used features) may capture similar subgroup (set of data objects)

- **Resources:**
  - Paper [5] (longer arXiv version [6])
  - Code [3] (including Python package `csd` with subgroup-discovery methods)
  - Experimental data [4]

Appendix

Related Work

- Existing subgroup-discovery methods [1, 18, 19, 34] typically algorithmic (exhaustive or heuristic)
- Few white-box formulations of (other) variants of subgroup discovery [10, 14, 17, 21, 26]
- Concept of feature cardinality well-established [19, 30], but empirical studies varying it [15, 24, 30, 32] are limited
- Concept of alternative subgroups well-established [1, 7, 11, 23, 27], but alternative descriptions [9, 16, 23, 25] less

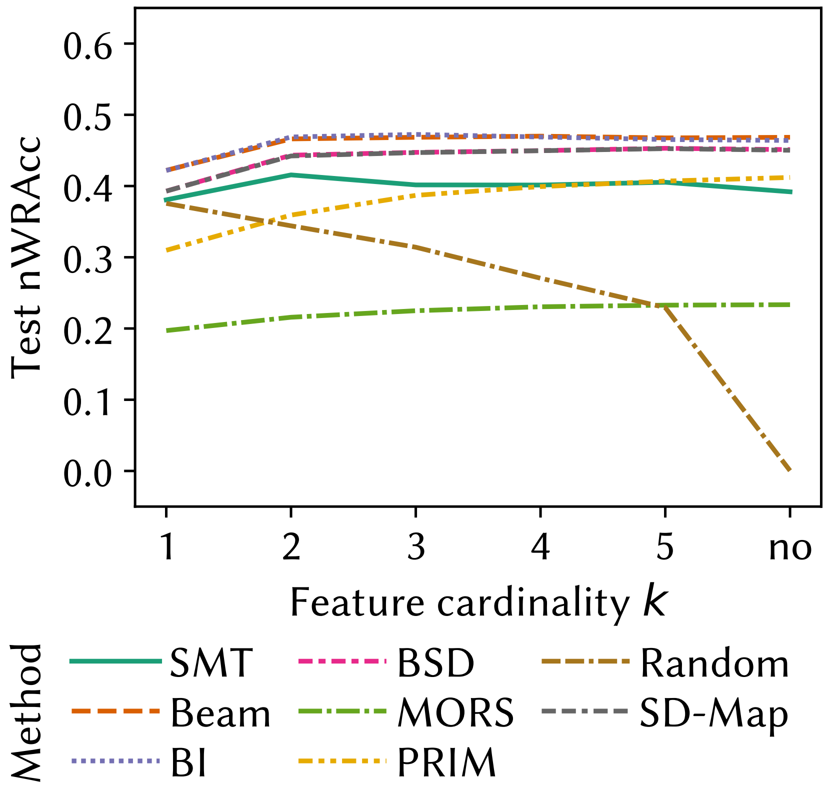
Computational Complexity

- Subgroup discovery with a feature-cardinality constraint is $\mathcal{NP}$-complete

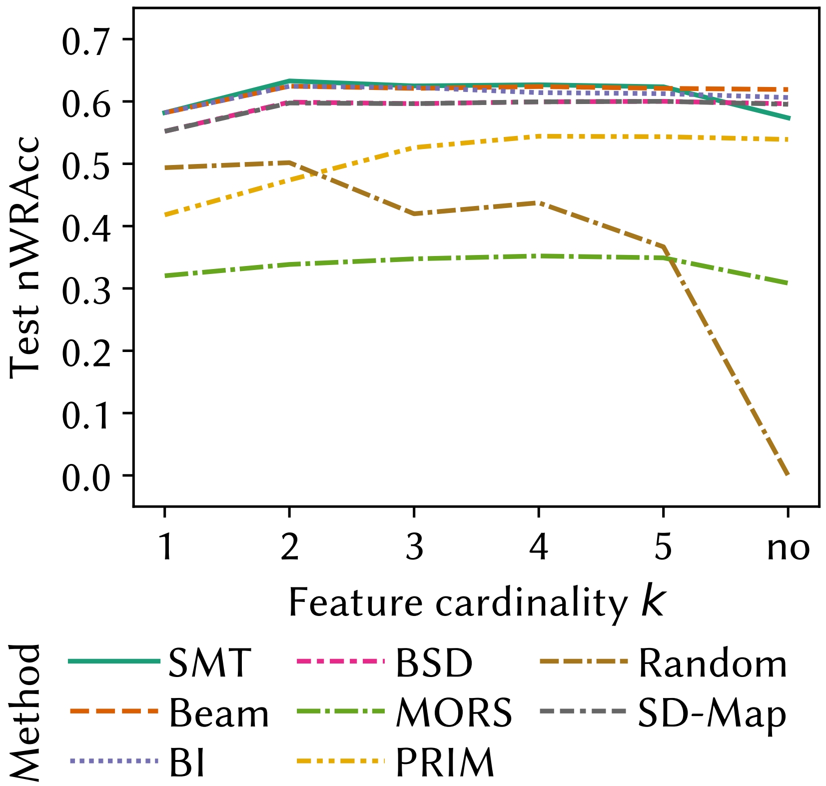
- **Proof:** Reduction from Set Covering [20]
  - Set-covering problem: Given set of elements $E = \{e_1, \dots, e_m\}$, set of sets $\mathbb{S} = \{S_1, \dots, S_n\}$ with $E = \bigcup_{S \in \mathbb{S}} S$, and a cardinality $k \in \mathbb{N}$, does subset $\mathbb{C} \subseteq \mathbb{S}$ with $|\mathbb{C}| \leq k$ and $E = \bigcup_{S \in \mathbb{C}} S$ exist?
  - Perfect-subgroup discovery: Find subgroup containing all positive data objects ($y_i = 1$) and zero negatives ($y_i = 0$)
  - Problem transformation:
    - Dataset $X \in \{0, 1\}^{(m+1) \times n}$ with $X_{ij} := (e_i \in S_j)$
    - Data object $m + 1$ represents an element not contained in any set, i.e., $X_{(m+1)j} = 0$ for all $j \in \{1, \dots, n\}$
    - Prediction target $y \in \{0, 1\}^{m+1}$ with $y_{m+1} = 1$ and $y_i = 0$ otherwise
  - Perfect subgroup only contains data object $m + 1$ and uses $lb_j = ub_j = 0$ conditions for selected features
  - Other data objects have value 1 for at least one selected feature $\rightarrow$ each element is in a selected set
  - I.e., algorithm for perfect-subgroup discovery also solves Set Covering
  - Optimizing subgroup quality typically at least as hard as finding perfect subgroup

- Finding alternative subgroup descriptions with a feature-cardinality constraint is $\mathcal{NP}$-complete as well

More Results – Feature-Cardinality Constraints



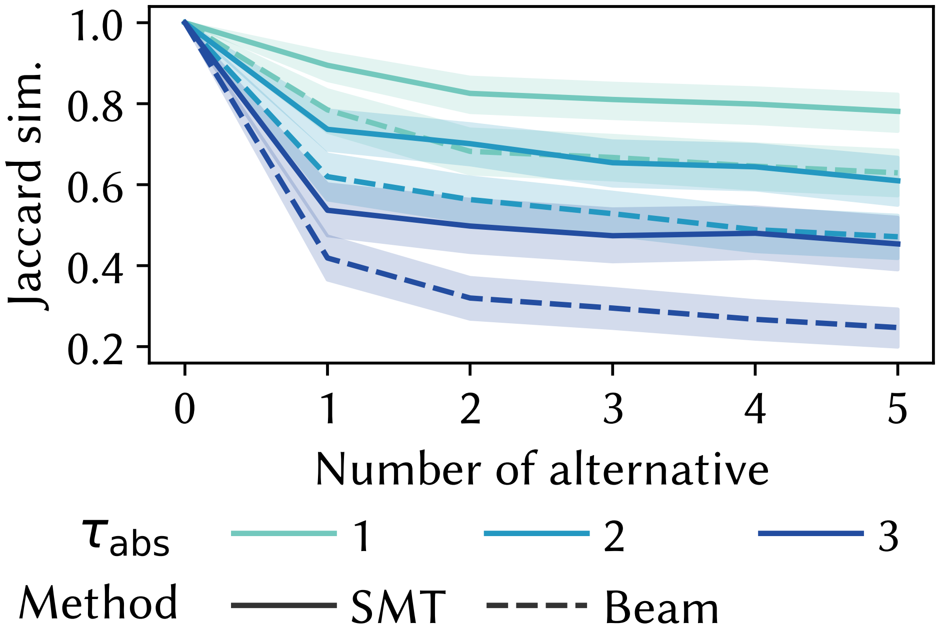
(a) All 27 datasets.



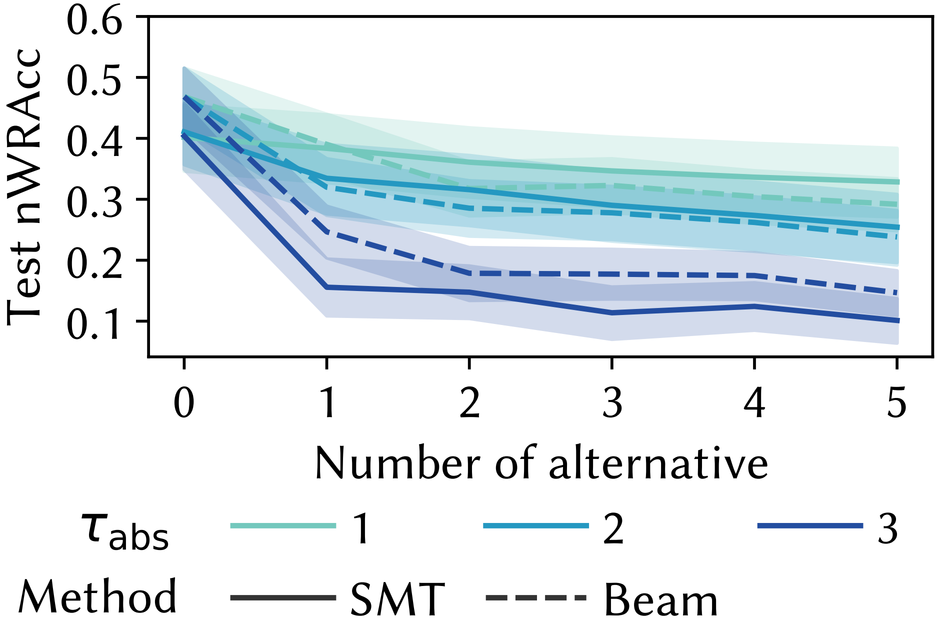
(b) 13 datasets without *SMT* timeouts.

Subgroup quality of original subgroups on the test set.

More Results – Alternative Subgroup Descriptions



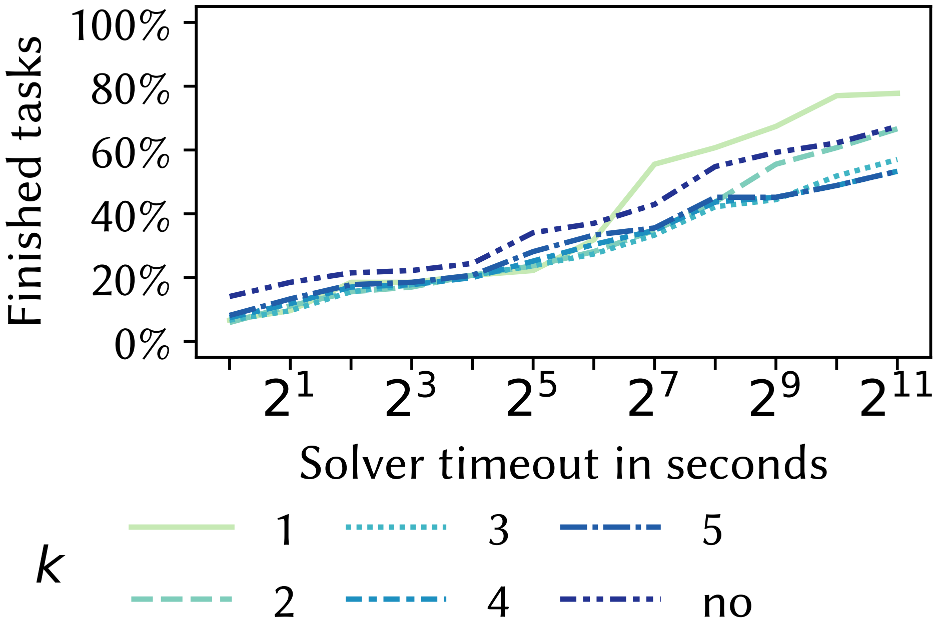
(a) Subgroup similarity (Jaccard).



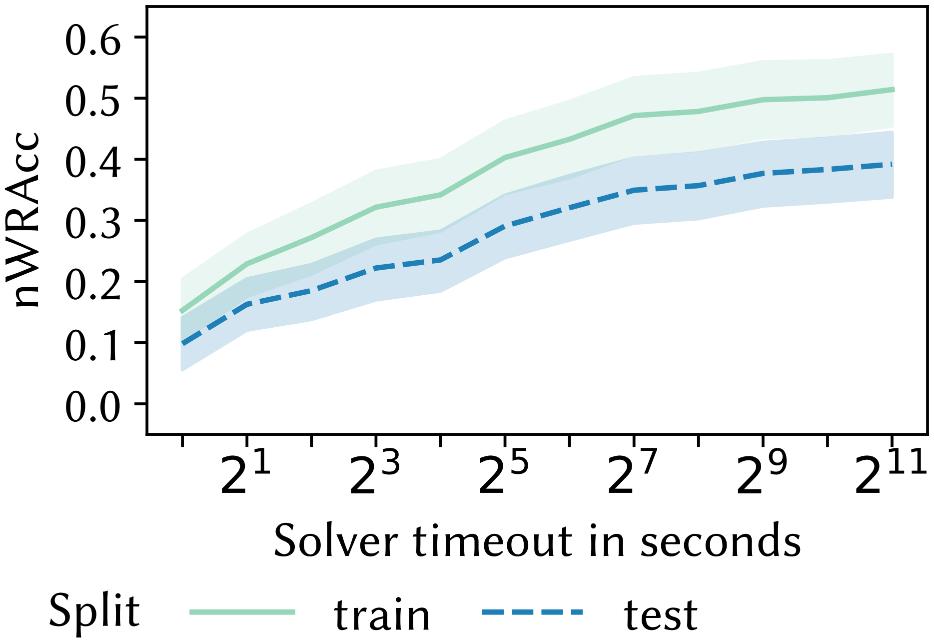
(b) Subgroup quality on the test set.

Similarity and quality of alternative subgroup descriptions.

More Results – Solver Timeouts



(a) Frequency of finished *SMT* tasks.



(b) Subgroup quality.

Impact of solver timeouts for *SMT* as the subgroup-discovery method.

More Results – Runtime

Mean runtime (in sec.) for searching original subgroups.

k	1	2	3	4	5	no
BI	7.8	11.7	14.2	16.7	18.7	35.0
BSD	0.9	0.9	0.9	2.7	29.5	55.7
Beam	6.8	10.1	12.8	14.6	16.1	30.5
MORS	0.0	0.0	0.0	0.0	0.0	0.0
PRIM	0.1	0.2	0.3	0.3	0.5	1.3
Random	0.6	0.6	0.6	0.7	0.7	0.9
SD-Map	2.3	3.3	9.6	54.0	345.2	367.4
SMT	648.2	911.3	1091.7	1113.4	1117.4	849.0

Mean runtime (in sec.) for searching alternative subgroup descriptions.

Method	τ_{abs}	Number of alternative					
		0	1	2	3	4	5
Beam	1	12.8	8.0	7.6	7.3	7.3	7.3
	2	12.8	7.7	7.4	7.2	7.0	6.8
	3	12.8	5.8	5.1	4.7	4.1	3.5
SMT	1	1091.7	166.0	221.5	239.6	258.1	277.9
	2	1105.2	377.5	463.5	537.5	599.4	658.3
	3	1107.4	869.1	670.8	597.6	588.1	557.6

References I

[1] Martin Atzmueller. “Subgroup discovery”. In: *WIREs Data Min. Knowl. Disc.* 5.1 (2015), pp. 35–49. DOI: 10.1002/widm.1144.

[2] Martin Atzmueller and Frank Puppe. “SD-Map – A Fast Algorithm for Exhaustive Subgroup Discovery”. In: *Proc. PKDD*. Berlin, Germany, 2006, pp. 6–17. DOI: 10.1007/11871637_6.

[3] Jakob Bach. *Constrained Subgroup Discovery*. GitHub repository. 2025. URL: https://github.com/Jakob-Bach/Constrained-Subgroup-Discovery.

[4] Jakob Bach. *Experimental Data for the Paper "Subgroup Discovery with Small and Alternative Feature Sets"*. RADAR4KIT. 2025. URL: https://doi.org/10.35097/nftgaf7w73hy2491.

[5] Jakob Bach. “Subgroup Discovery with Small and Alternative Feature Sets”. In: *Proc. SIGMOD*. Berlin, Germany, 2025. DOI: 10.1145/3725358.

[6] Jakob Bach. *Using Constraints to Discover Sparse and Alternative Subgroup Descriptions*. arXiv:2406.01411v2 [cs.LG]. 2025. DOI: 10.48550/arXiv.2406.01411.

References II

[7] Adnene Belfodil et al. “FSSD - A Fast and Efficient Algorithm for Subgroup Set Discovery”. In: *Proc. DSAA*. Washington, DC, USA, 2019, pp. 91–99. DOI: 10.1109/DSAA.2019.00023.

[8] Nikolaj Bjørner, Anh-Dung Phan, and Lars Fleckenstein. “νZ - An Optimizing SMT Solver”. In: *Proc. TACAS*. London, United Kingdom, 2015, pp. 194–199. DOI: 10.1007/978-3-662-46681-0_14.

[9] Mario Boley and Henrik Grosskreutz. “Non-redundant Subgroup Discovery Using a Closure System”. In: *Proc. ECML PKDD*. Bled, Slovenia, 2009, pp. 179–194. DOI: 10.1007/978-3-642-04180-8_29.

[10] Tibérius O. Bonates, Peter L. Hammer, and Alexander Kogan. “Maximum patterns in datasets”. In: *Discrete Appl. Math.* 156.6 (2008), pp. 846–861. DOI: 10.1016/j.dam.2007.06.004.

[11] Guillaume Bosc et al. “Anytime discovery of a diverse set of patterns with Monte Carlo tree search”. In: *Data Min. Knowl. Disc.* 32.3 (2018), pp. 604–650. DOI: 10.1007/s10618-017-0547-5.

[12] Seung-Seok Choi, Sung-Hyuk Cha, and Charles C. Tappert. “A Survey of Binary Similarity and Distance Measures”. In: *J. Syst. Cybern. Inf.* 8.1 (2010), pp. 43–48. URL: http://www.iiisci.org/Journal/pdv/sci/pdfs/GS315JG.pdf.

References III

[13] Leonardo De Moura and Nikolaj Bjørner. "Z3: An Efficient SMT Solver". In: *Proc. TACAS*. Budapest, Hungary, 2008, pp. 337–340. DOI: 10.1007/978-3-540-78800-3_24.

[14] Jonathan Eckstein et al. "The Maximum Box Problem and its Application to Data Analysis". In: *Comput. Optim. Appl.* 23.3 (2002), pp. 285–298. DOI: 10.1023/A:1020546910706.

[15] Jerome H. Friedman and Nicholas I. Fisher. "Bump hunting in high-dimensional data". In: *Stat. Comput.* 9.2 (1999), pp. 123–143. DOI: 10.1023/A:1008894516817.

[16] Esther Galbrun and Pauli Miettinen. *Redescription Mining*. 1st ed. Springer, 2017. DOI: 10.1007/978-3-319-72889-6.

[17] Tias Guns, Siegfried Nijssen, and Luc De Raedt. "Itemset mining: A constraint programming perspective". In: *Artif. Intell.* 175.12-13 (2011), pp. 1951–1983. DOI: 10.1016/j.artint.2011.05.002.

[18] Sumyea Helal. "Subgroup Discovery Algorithms: A Survey and Empirical Evaluation". In: *J. Comput. Sci. Technol.* 31.3 (2016), pp. 561–576. DOI: 10.1007/s11390-016-1647-1.

References IV

[19] Franciso Herrera et al. “An overview on subgroup discovery: foundations and applications”. In: *Knowl. Inf. Syst.* 29.3 (2011), pp. 495–525. DOI: 10.1007/s10115-010-0356-2.

[20] Richard M. Karp. “Reducibility among Combinatorial Problems”. In: *Complexity of Computer Computations*. 1st ed. Plenum Press, 1972. Chap. 9, pp. 85–103. DOI: 10.1007/978-1-4684-2001-2_9.

[21] Gökberk Koçak et al. “Exploiting Incomparability in Solution Dominance: Improving General Purpose Constraint-Based Mining”. In: *Proc. ECAI*. Santiago de Compostela, Spain, 2020, pp. 331–338. DOI: 10.3233/FAIA200110.

[22] Nada Lavrač, Peter Flach, and Blaz Zupan. “Rule Evaluation Measures: A Unifying View”. In: *Proc. ILP*. Bled, Slovenia, 1999, pp. 174–185. DOI: 10.1007/3-540-48751-4_17.

[23] Matthijs van Leeuwen and Arno Knobbe. “Diverse subgroup set discovery”. In: *Data Min. Knowl. Disc.* 25.2 (2012), pp. 208–242. DOI: 10.1007/s10618-012-0273-y.

[24] Florian Lemmerich, Mathias Rohlfs, and Martin Atzmueller. “Fast Discovery of Relevant Subgroup Patterns”. In: *Proc. FLAIRS*. Daytona Beach, FL, USA, 2010, pp. 428–433. URL: https://aaai.org/papers/flairs-2010-1262/.

References V

[25] Antonio Lopez-Martinez-Carrasco et al. "Discovering Diverse Top-K Characteristic Lists". In: *Proc. IDA*. Louvain-la-Neuve, Belgium, 2023, pp. 262–273. DOI: 10.1007/978-3-031-30047-9_21.

[26] Quentin Louveaux and Sébastien Mathieu. "A combinatorial branch-and-bound algorithm for box search". In: *Discrete Optim.* 13 (2014), pp. 36–48. DOI: 10.1016/j.disopt.2014.05.001.

[27] Tarcísio Lucas, Renato Vimieiro, and Teresa Ludermir. "SSDP+: A Diverse and More Informative Subgroup Discovery Approach for High Dimensional Data". In: *Proc. CEC*. Rio de Janeiro, Brazil, 2018. DOI: 10.1109/CEC.2018.8477855.

[28] Michael Mampaey et al. "Efficient Algorithms for Finding Richer Subgroup Descriptions in Numeric and Nominal Data". In: *Proc. ICDM*. Brussels, Belgium, 2012, pp. 499–508. DOI: 10.1109/ICDM.2012.117.

[29] Romain Mathonat et al. "Anytime Subgroup Discovery in High Dimensional Numerical Data". In: *Proc. DSAA*. Porto, Portugal, 2021. DOI: 10.1109/DSAA53316.2021.9564223.

[30] Marvin Meeng and Arno Knobbe. "For real: a thorough look at numeric attributes in subgroup discovery". In: *Data Min. Knowl. Disc.* 35.1 (2021), pp. 158–212. DOI: 10.1007/s10618-020-00703-x.

References VI

[31] Randal S. Olson et al. “PMLB: a large benchmark suite for machine learning evaluation and comparison”. In: *BioData Min.* 10 (2017). DOI: 10.1186/s13040-017-0154-4.

[32] Hugo M. Proença et al. “Robust subgroup discovery: Discovering subgroup lists using *MDL*”. In: *Data Min. Knowl. Disc.* 36.5 (2022), pp. 1885–1970. DOI: 10.1007/s10618-022-00856-x.

[33] Joseph D. Romano et al. *PMLB v1.0: An open source dataset collection for benchmarking machine learning methods*. arXiv:2012.00058v3 [cs.LG]. 2021. DOI: 10.48550/arXiv.2012.00058.

[34] Sebastián Ventura and José María Luna. “Subgroup Discovery”. In: *Supervised Descriptive Pattern Mining*. 1st ed. Springer, 2018. Chap. 4, pp. 71–98. DOI: 10.1007/978-3-319-98140-6_4.