Multi-objective learning of accurate and comprehensible classifiers – a case study

Rok Piltaver, Mitja Luštrek and Matjaž Gams

Jožef Stefan Institute, Ljubljana, Slovenia

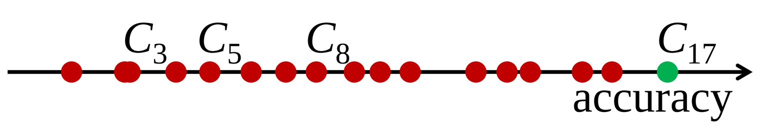
Multi-objective learning

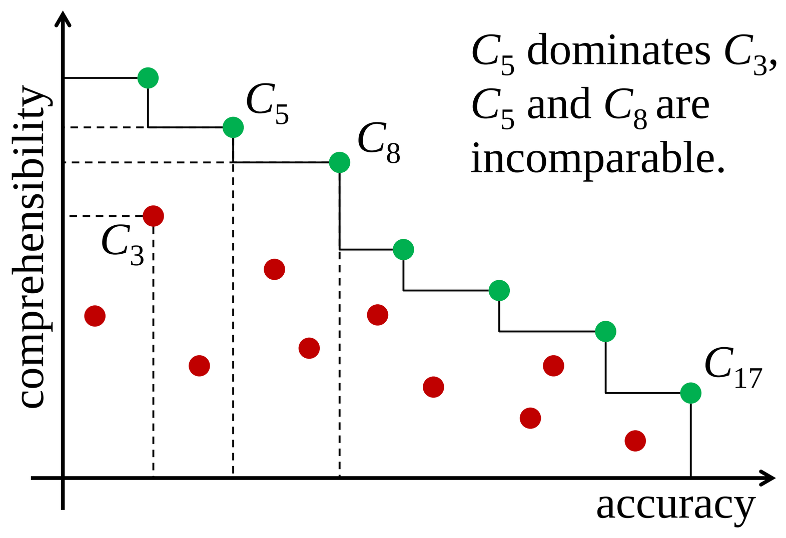
Considering only accuracy:

- classifiers compared with greater then relation
- single best solution

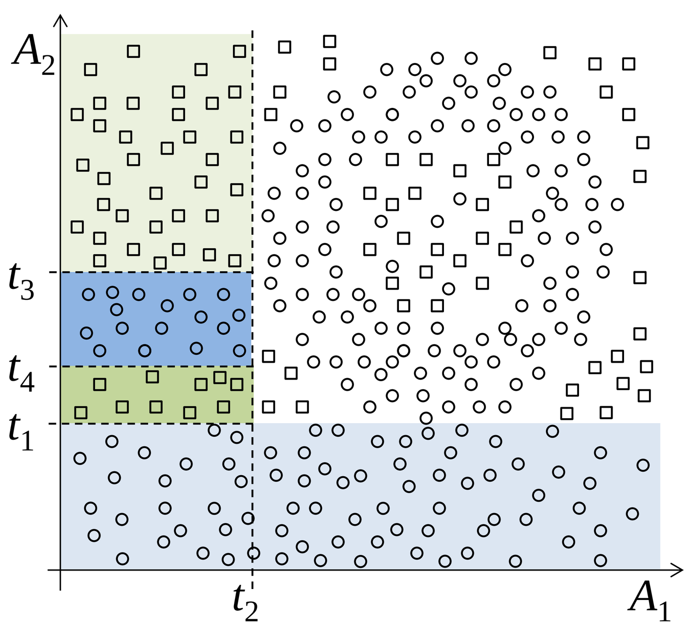
Considering accuracy and comprehensibility:

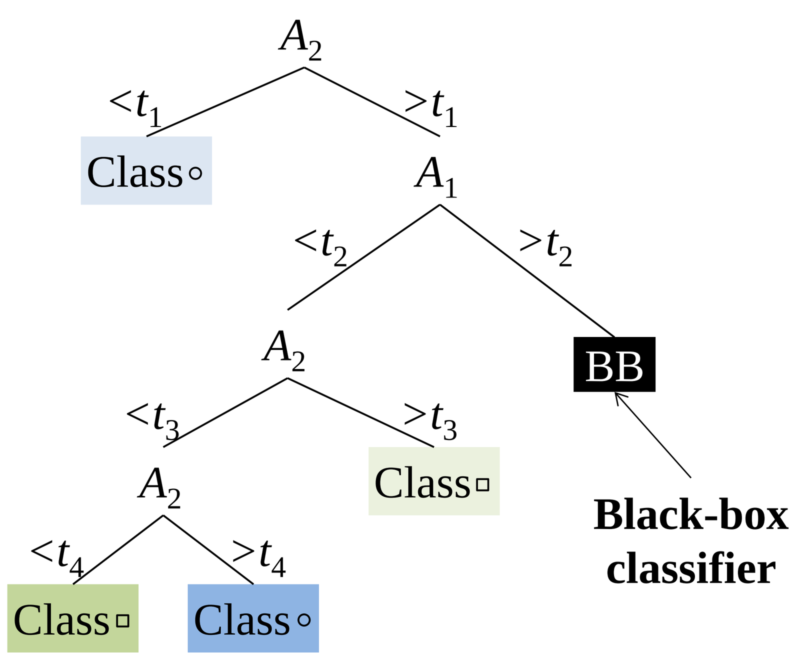
- classifiers compared with Pareto dominates relation
- a set of non-dominated solutions





Hybrid trees with black-box leaves





Domain experts can **explain some parts of the domain with simple rules** but fail to formulize the rest of their knowledge. Hybrid tree should therefore be used as the classification model: if accuracy of a regular leaf is high enough use it; if not **use a more accurate black box** (BB) **classifier instead** to classify the instances belonging to the leaf.

Algorithm

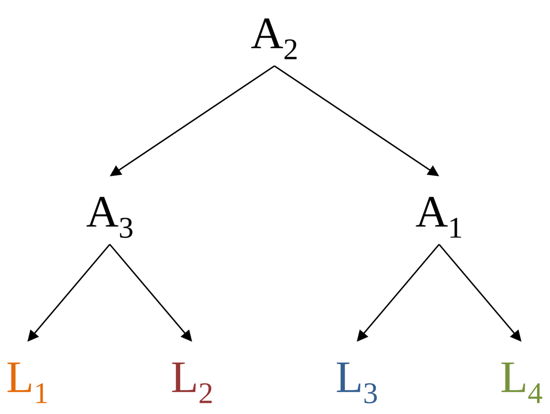
Find hybrid trees $T_i = (t_{i,1}, t_{i,2}, \ldots, t_{i,n}) \in \{0, 1\}^n$ that are non-dominated according to the two objectives $(a_i, c_i) \in [0, 1] \times [0, 1]$:

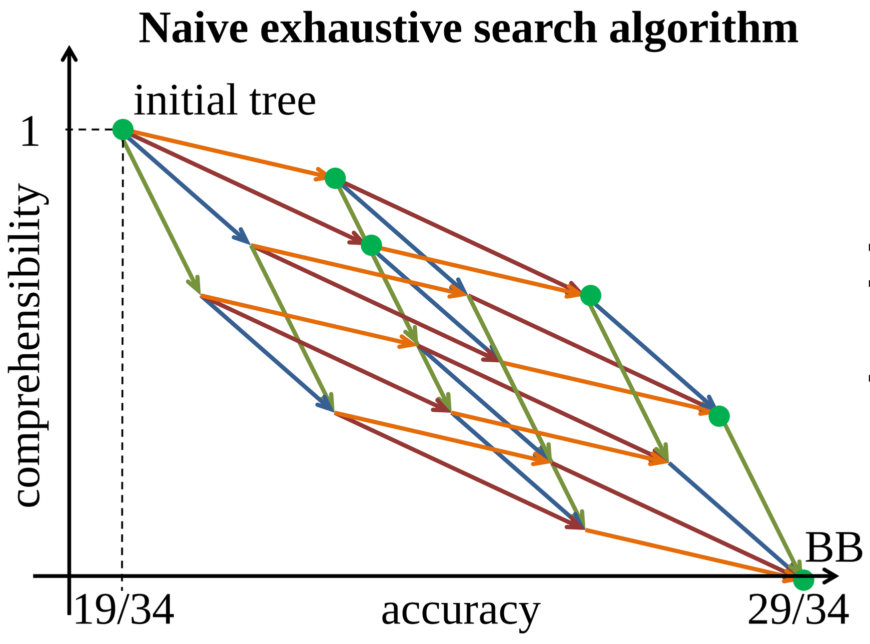
- accuracy: $a_i = (\sum_{j \in \text{regular leaf}} N_{j,t} + \sum_{j \in \text{hybrid leaf}} N_{j,bb}) / N$
- comprehensibility: $c_i = (\sum_{j \in \text{regular leaf}} N_j) / N$

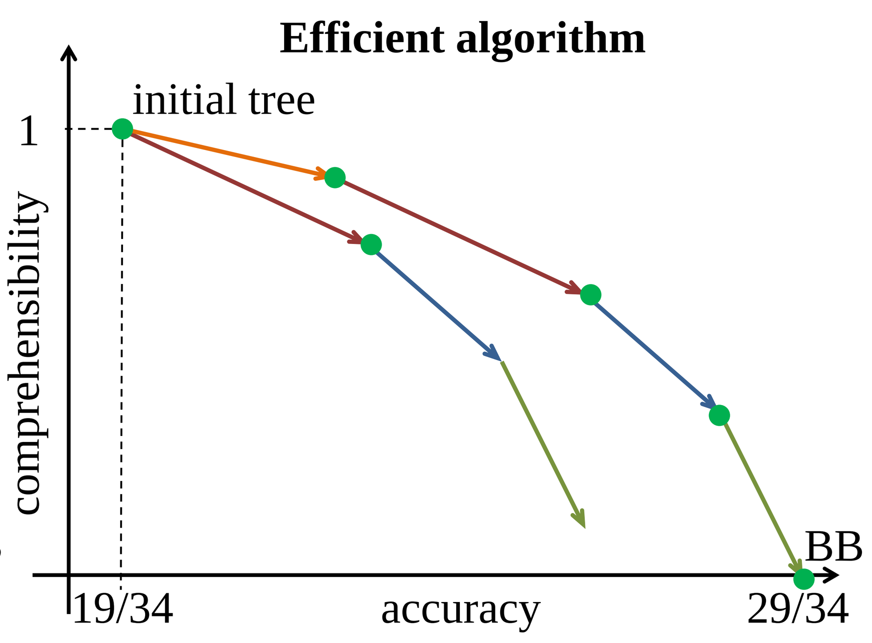
Use dynamic programing to compute the quality of tree $T_j = (a_i + \delta_{l,a}, c_i + \delta_{l,c})$ from the quality of tree $T_i = (a_i, c_i)$ where:

- T_j is obtained by replacing a single leaf l in T_i for the BB $t_{j,k} = \begin{cases} t_{j,l} = 1 \\ t_{j,k} = t_{i,k}, k \neq l \end{cases}$
- relative difference in accuracy is: $\delta_{l,a} = (N_{l,bb} - N_{l,t}) / N > 0$
- relative difference in comprehensibility is: $\delta_{l,c} = -N_l / N < 0$

leaf i	1	2	3	4	∑
#instances in the leaf N_i	8	9	7	10	**34**
#correctly classified using the tree $N_{i,t}$	4	5	4	6	**19**
#correctly classified using the BB $N_{i,BB}$	8	7	7	7	**29**
relative difference in accuracy $\delta_{i,a}$	**+4**	**+2**	**+3**	**+1**	10
relative diff. in comprehensibility $\delta_{i,c}$	**-8**	**-9**	**-7**	**-10**	34





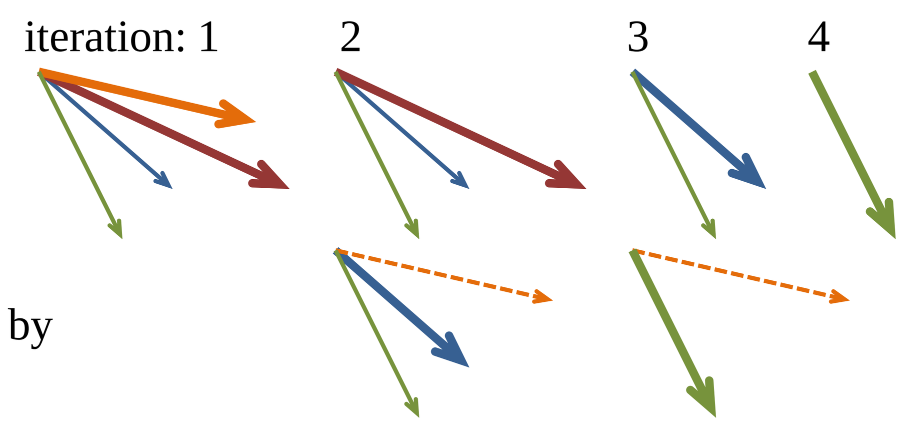


Naive algorithm generates:

- most hybrid trees more than once,
- many dominated hybrid trees.

The efficient algorithm:

- generates each hybrid trees only once,
- avoids generating dominated hybrid trees by replacing only the non-dominated leaves

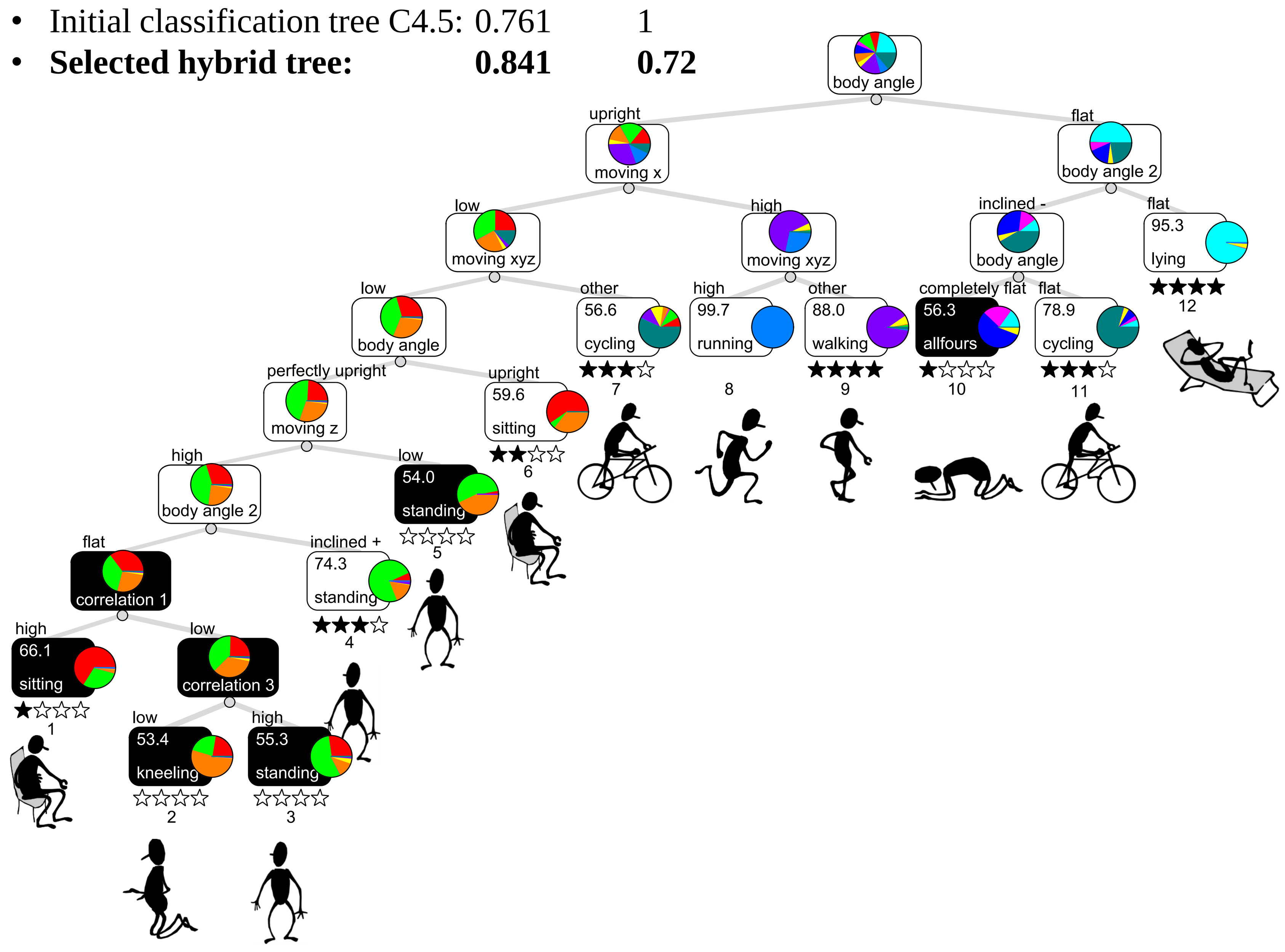


Case study: activity recognition

Activity recognition domain:

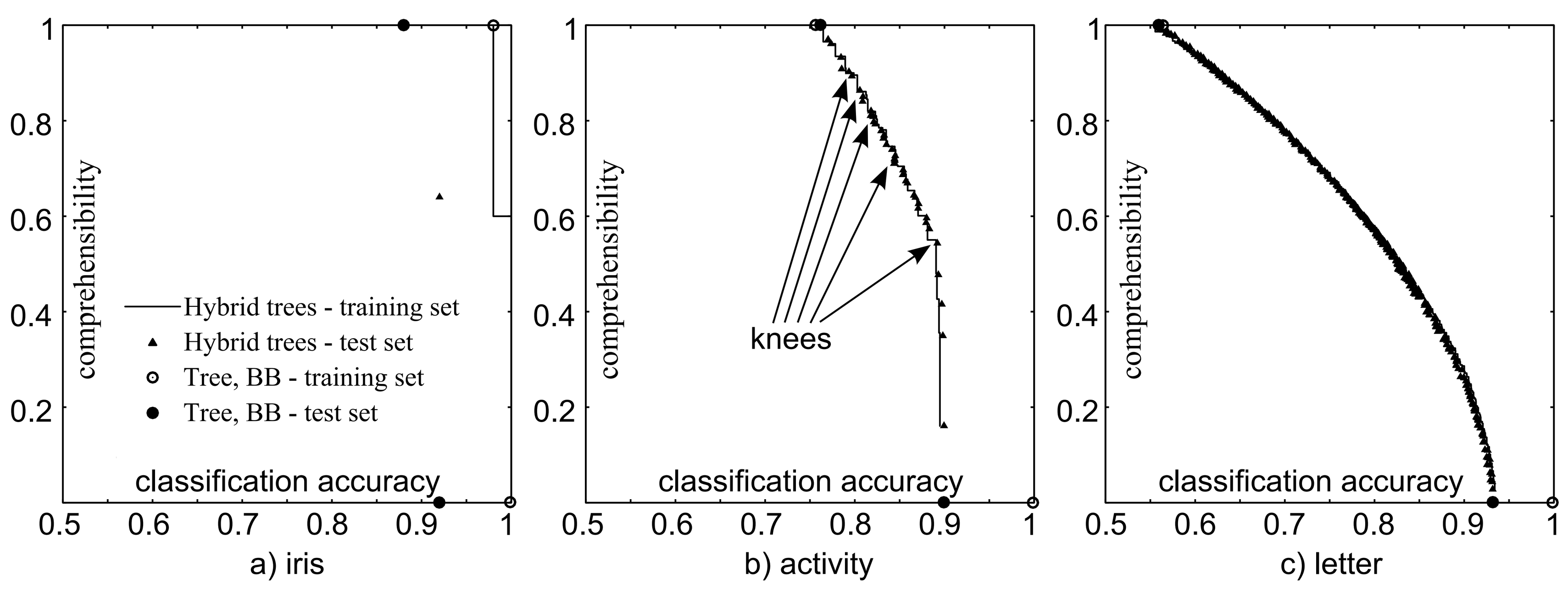
- **10 classes**: lying (22 %), walking (17 %), cycling (14 %), standing (10 %), sitting (8 %), kneeling (8 %), on all fours (7 %), running (7 %), transitions (4 %), leaning (3 %)
- **61 attributes** computed from 2 s windows of 3-axis accelerometer data (on chest)
- **48.000 instances**, 9 persons (1.5 h each)

Classifier	**accuracy**	**comprehensibility**
BB classifier random forest:	0.906	0
Initial classification tree C4.5:	0.761	1
Selected hybrid tree:	**0.841**	**0.72**

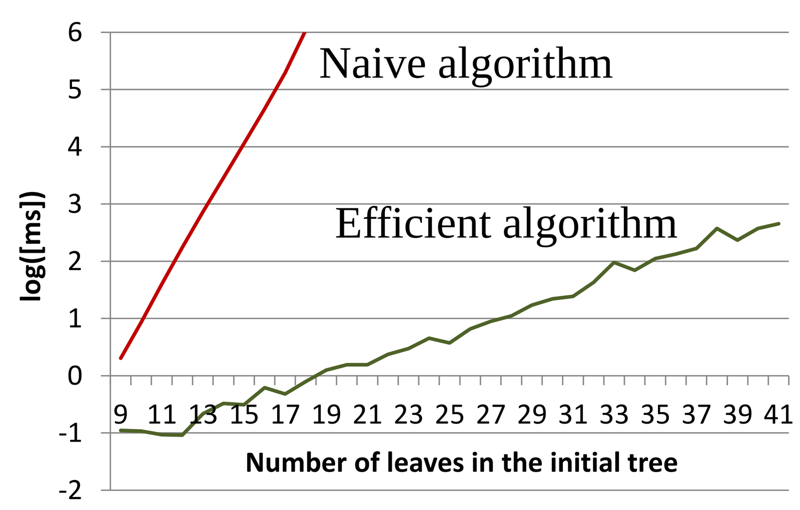


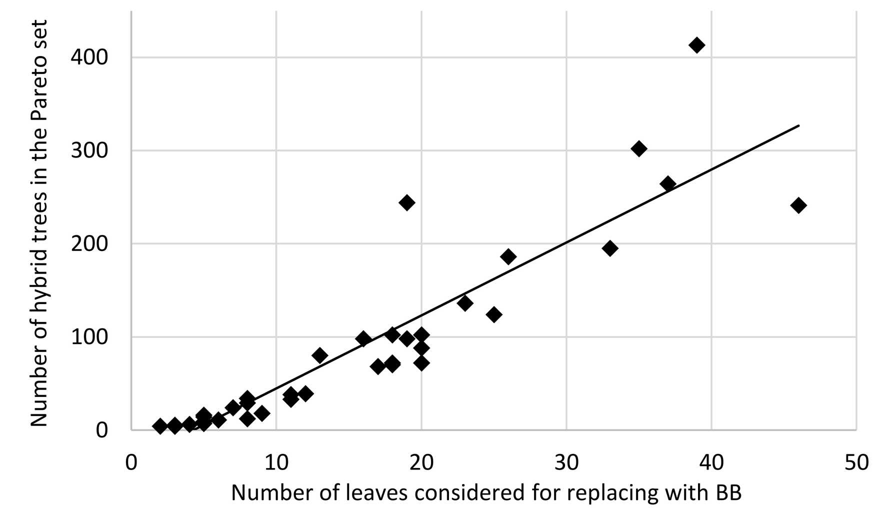
Using the Pareto front

Steepness of the Pareto front corresponds to the difference in accuracy between the initial tree and black-box classifier. Choose comprehensible classifier if Pareto front is steep.
Number of hybrid trees increases with the number of leaves in the initial tree.
Presence of knees on the Pareto front reduces the number of hybrid trees to be analyzed.
The **leaves** that are never replaced with the black-box classifiers are **validated.**
Relative quality of leaves corresponds to the number of hybrid trees with the leaf replaced.
Errors in the estimated comprehensibility and accuracy decrease with increasing number of learning instances.



Evaluation





Naive algorithm:

- 17 leaves, 3.26 min
- 18 leaves, 18.08 min

Efficient algorithm:

- 17 leaves, 0.48 ms
- 40 leaves, ~0.5 s

The suggested algorithm:

- is guaranteed to find the entire set of non-dominated solutions,
- outperforms the baseline solution set in terms of hyper-volume,
- outperforms state-of-the-art multi objective optimization algorithm NSGA-II in terms of execution time.

rok.piltaver@ijs.si