Appendix: Multi-criteria approach for selecting an explanation from the set of counterfactuals produced by an ensemble of explainers

Ignacy Stępka, Mateusz Lango, Jerzy Stefanowski

Poznan University of Technology

This appendix contains the following:

1. Extended walkthrough of the toy example presented in Section 3.5, including all the data supporting the case
2. Radar plot visualizations of the scores of explanation quality measures obtained by different explanation algorithms
3. The detailed specification of the trained black-box models
4. The results supporting the ablation study presented in Section 4.4 (the analysis of the impact of the elements in our approach)

1 Extension of the toy example walkthrough

In the paper, we provide a detailed architectural overview to aid readers in comprehending the approach we have introduced. This appendix contains all the data supporting the case presented in the paper.

Recall that we have chosen the following real-world example from the Adult dataset[1]:

age	edu.num	cpt.gain	cpt.loss	hrs/week	workclass	marital.status	occupation	race	sex	nat.cntry	income
24	10	0	0	30	Self-emp-not-inc	Never-married	Prof-specialty	Asian-Pac-Islander	Male	USA	>50K

All 82 counterfactual explanations obtained from the ensemble of explainers for the above example are presented in Table 1. Each row is color-coded to represent the stage at which each alternative was filtered out: *red* for invalid, *gray* for non-actionable, *blue* for dominated, *orange* for nondominated, and *green* for the selected explanation. Table 2 provides the scores of three selected quality measures (proximity, discriminative power, feasibility) assigned to these counterfactuals.

Finally, we present a 3D visualization illustrating the selection process from the Pareto front in Figure 1, followed by a two-dimensional projection in Figure 2.

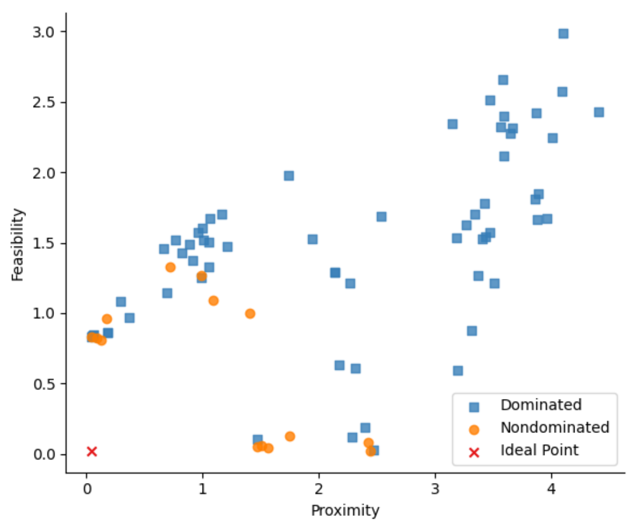

Fig. 2. This plot depicts the same data as the one above, except with the *DiscriminativePower* axis collapsed, which accounts for the apparent discrepancy in the Pareto front observed in the two-criteria plot.

[1] For the full specification of the dataset refer to: https://archive.ics.uci.edu/dataset/2/adult

ID	explainer	age	edu.num	capital.gain	capital.loss	hrs/week	workclass	marital.status	occupation	race	sex	nat.cntry	income
1	dice	24.0	10.0	5193.0	0.0	30.0	Self-emp-not-inc	Married-AF-spouse	Prof-specialty	As-Pa-Is	Male	USA	<=50K
2	dice	24.0	10.0	0.0	3851.0	30.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
3	dice	24.0	10.0	0.0	3139.0	30.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
4	dice	24.0	6.0	49751.0	0.0	30.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
5	dice	24.0	10.0	17327.0	0.0	30.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
6	dice	24.0	10.0	26115.0	3511.0	30.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
7	dice	24.0	10.0	94281.0	0.0	30.0	Self-emp-not-inc	Never-married	Sales	As-Pa-Is	Male	USA	<=50K
8	dice	24.0	10.0	0.0	402.0	30.0	Self-emp-not-inc	Married-civ-spouse	Prof-specialty	As-Pa-Is	Male	USA	<=50K
9	dice	78.0	10.0	0.0	0.0	30.0	Self-emp-not-inc	Married-AF-spouse	Prof-specialty	As-Pa-Is	Male	USA	<=50K
10	dice	24.0	10.0	29264.0	0.0	30.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
11	dice	24.0	10.0	0.0	0.0	51.0	Self-emp-not-inc	Widowed	Prof-specialty	As-Pa-Is	Male	USA	<=50K
12	dice	24.0	10.0	4349.0	0.0	30.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
13	dice	70.0	10.0	53380.0	0.0	30.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
14	dice	54.0	10.0	0.0	0.0	30.0	Self-emp-not-inc	Married-civ-spouse	Prof-specialty	As-Pa-Is	Male	USA	<=50K
15	dice	24.0	14.0	10377.0	0.0	30.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
16	dice	39.0	10.0	0.0	3097.0	30.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
17	dice	24.0	10.0	66733.0	0.0	30.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
18	dice	24.0	10.0	0.0	4354.0	30.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
19	dice	24.0	10.0	0.0	3568.0	30.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
20	dice	24.0	10.0	0.0	4207.0	30.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
21	cadex	24.0	10.0	4999.0	0.0	30.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
22	cadex	27.0	10.0	4999.0	0.0	30.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
23	cadex	27.0	10.0	4999.0	0.0	34.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
24	cadex	27.0	10.0	4999.0	217.0	34.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
25	cadex	27.0	10.0	4999.0	217.0	34.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
26	fimap	30.0	11.0	10317.0	102.0	34.0	Private	Married-civ-spouse	Prof-specialty	As-Pa-Is	Male	USA	<=50K
27	fimap	30.0	10.0	5907.0	43.0	32.0	Private	Married-civ-spouse	Prof-specialty	As-Pa-Is	Male	USA	<=50K
28	fimap	31.0	10.0	6444.0	49.0	32.0	Private	Married-civ-spouse	Machine-op-inspct	As-Pa-Is	Male	USA	<=50K
29	fimap	31.0	11.0	9949.0	102.0	33.0	Private	Married-civ-spouse	Handlers-cleaners	As-Pa-Is	Male	USA	<=50K
30	fimap	31.0	11.0	7127.0	61.0	32.0	Private	Divorced	Prof-specialty	As-Pa-Is	Male	USA	<=50K
31	fimap	31.0	10.0	6595.0	49.0	32.0	Private	Divorced	?	As-Pa-Is	Male	USA	>50K
32	wachter	72.0	10.0	3885.0	0.0	65.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
33	wachter	70.0	9.0	3640.0	0.0	55.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
34	wachter	53.0	9.0	2918.0	0.0	33.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	>50K
35	wachter	24.0	10.0	4402.0	0.0	29.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
36	wachter	31.0	9.0	2614.0	1.0	30.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	>50K
37	wachter	24.0	10.0	4469.0	0.0	30.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
38	wachter	23.0	9.0	4474.0	0.0	29.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	>50K
39	wachter	70.0	10.0	3712.0	0.0	62.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
40	wachter	58.0	8.0	4964.0	0.0	34.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
41	wachter	23.0	10.0	4405.0	0.0	30.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	>50K
42	cem	24.0	10.0	5939.0	0.0	30.0	Self-emp-not-inc	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
43	cfproto	24.0	10.0	15154.0	0.0	30.0	Self-emp-inc	Never-married	Protective-serv	As-Pa-Is	Male	France	<=50K
44	cfproto	28.0	10.0	35794.0	0.0	30.0	?	Never-married	Protective-serv	White	Male	Outlying-US(	<=50K
45	cfproto	24.0	10.0	13763.0	0.0	30.0	Self-emp-not-inc	Never-married	Prof-specialty	White	Male	Philippines	<=50K
46	cfproto	24.0	10.0	18256.0	0.0	30.0	Self-emp-inc	Never-married	Prof-specialty	White	Male	China	<=50K
47	cfproto	24.0	10.0	34508.0	0.0	30.0	Self-emp-inc	Never-married	Prof-specialty	White	Male	Vietnam	<=50K
48	cfproto	24.0	10.0	40939.0	0.0	30.0	State-gov	Never-married	Prof-specialty	White	Male	India	<=50K
49	cfproto	24.0	10.0	26873.0	0.0	30.0	Self-emp-inc	Never-married	Prof-specialty	White	Male	China	<=50K
50	cfproto	23.0	10.0	45951.0	0.0	30.0	Without-pay	Never-married	Prof-specialty	White	Male	Greece	<=50K
51	cfproto	24.0	10.0	13763.0	0.0	30.0	Self-emp-not-inc	Never-married	Prof-specialty	White	Male	Philippines	<=50K
52	grow-sph	45.0	9.0	26255.0	703.0	22.0	Federal-gov	Married-civ-spouse	?	As-Pa-Is	Male	USA	<=50K
53	grow-sph	17.0	14.0	42417.0	86.0	36.0	Self-emp-inc	Divorced	?	As-Pa-Is	Male	USA	<=50K
54	grow-sph	34.0	10.0	26651.0	333.0	27.0	?	Divorced	?	As-Pa-Is	Male	USA	<=50K
55	grow-sph	30.0	13.0	31719.0	1324.0	24.0	?	Divorced	?	As-Pa-Is	Male	USA	<=50K
56	grow-sph	17.0	9.0	29154.0	47.0	43.0	Private	Married-AF-spouse	Armed-Forces	As-Pa-Is	Male	USA	<=50K
57	grow-sph	46.0	12.0	17359.0	0.0	36.0	?	Married-AF-spouse	Craft-repair	As-Pa-Is	Male	USA	<=50K
58	grow-sph	17.0	9.0	19830.0	0.0	36.0	Federal-gov	Married-AF-spouse	Adm-clerical	As-Pa-Is	Male	USA	<=50K
59	grow-sph	24.0	6.0	22589.0	115.0	24.0	?	Married-AF-spouse	Craft-repair	As-Pa-Is	Male	USA	<=50K
60	grow-sph	33.0	8.0	25089.0	552.0	29.0	Federal-gov	Divorced	Armed-Forces	As-Pa-Is	Male	USA	<=50K
61	grow-sph	29.0	10.0	31798.0	215.0	30.0	Federal-gov	Divorced	Adm-clerical	As-Pa-Is	Male	USA	<=50K
62	grow-sph	17.0	8.0	49887.0	641.0	52.0	?	Married-AF-spouse	Craft-repair	As-Pa-Is	Male	USA	<=50K
63	grow-sph	30.0	10.0	35880.0	0.0	33.0	Federal-gov	Divorced	Craft-repair	As-Pa-Is	Male	USA	<=50K
64	grow-sph	21.0	10.0	27720.0	0.0	35.0	?	Divorced	?	As-Pa-Is	Male	USA	<=50K
65	grow-sph	31.0	7.0	47001.0	198.0	22.0	Private	Married-AF-spouse	Adm-clerical	As-Pa-Is	Male	USA	<=50K
66	grow-sph	25.0	11.0	24283.0	518.0	15.0	?	Divorced	Craft-repair	As-Pa-Is	Male	USA	<=50K
67	grow-sph	40.0	7.0	36015.0	275.0	5.0	Federal-gov	Divorced	Armed-Forces	As-Pa-Is	Male	USA	<=50K
68	grow-sph	32.0	9.0	29964.0	0.0	36.0	Self-emp-not-inc	Divorced	Machine-op-inspct	As-Pa-Is	Male	USA	<=50K
69	grow-sph	39.0	11.0	20938.0	0.0	38.0	Federal-gov	Married-AF-spouse	Farming-fishing	As-Pa-Is	Male	USA	<=50K
70	grow-sph	35.0	10.0	34992.0	0.0	67.0	Private	Married-AF-spouse	Machine-op-inspct	As-Pa-Is	Male	USA	<=50K
71	grow-sph	17.0	3.0	30575.0	430.0	26.0	Private	Married-AF-spouse	Exec-managerial	As-Pa-Is	Male	USA	<=50K
72	act-rec	24.0	10.0	0.0	0.0	29.0	Self-emp-not-inc	Married-AF-spouse	Prof-specialty	As-Pa-Is	Male	USA	<=50K
73	face	32.0	14.0	0.0	0.0	39.0	Self-emp-not-inc	Never-married	Prof-specialty	White	Male	USA	<=50K
74	face	39.0	16.0	4787.0	0.0	39.0	Self-emp-not-inc	Never-married	Prof-specialty	White	Male	USA	<=50K
75	face	27.0	13.0	13549.0	0.0	39.0	Private	Never-married	Prof-specialty	As-Pa-Is	Male	USA	<=50K
76	face	37.0	13.0	0.0	0.0	39.0	Self-emp-not-inc	Married-civ-spouse	Prof-specialty	White	Male	USA	<=50K
77	face	38.0	10.0	3137.0	0.0	50.0	Self-emp-not-inc	Married-civ-spouse	Prof-specialty	White	Male	USA	<=50K
78	face	35.0	9.0	8614.0	0.0	39.0	Self-emp-not-inc	Never-married	Machine-op-inspct	White	Male	USA	<=50K
79	face	30.0	15.0	0.0	0.0	39.0	Self-emp-not-inc	Never-married	Prof-specialty	White	Male	USA	<=50K
80	face	35.0	13.0	0.0	0.0	39.0	Self-emp-not-inc	Married-civ-spouse	Prof-specialty	White	Male	USA	<=50K
81	face	39.0	14.0	0.0	0.0	39.0	Self-emp-not-inc	Never-married	Prof-specialty	White	Male	USA	<=50K
82	face	37.0	9.0	3137.0	0.0	29.0	Self-emp-not-inc	Married-civ-spouse	Prof-specialty	White	Male	USA	<=50K

Table 1. Full list of generated counterfactual explanations for the toy example.

ID	Proximity	K_Feasibility(3)	DiscriminativePower(9)	explainer
1	1.052	1.506	0.222	dice
2	0.884	1.491	0.222	dice
3	0.721	1.328	0.222	dice
4	0.764	1.517	0.000	dice
5	0.173	0.962	0.111	dice
6	1.067	1.674	0.222	dice
7	1.943	1.529	0.111	dice
8	1.092	1.093	0.444	dice
9	1.740	1.975	0.667	dice
10	0.293	1.082	0.111	dice
11	1.214	1.475	0.444	dice
12	0.043	0.832	0.000	dice
13	1.164	1.706	0.111	dice
14	1.411	1.004	0.556	dice
15	0.370	0.966	0.111	dice
16	0.916	1.372	0.222	dice
17	0.667	1.456	0.111	dice
18	1.000	1.607	0.222	dice
19	0.819	1.426	0.222	dice
20	0.966	1.573	0.222	dice
21	0.050	0.839	0.000	cadex
22	0.091	0.825	0.000	cadex
23	0.132	0.812	0.000	cadex
24	0.182	0.861	0.000	cadex
25	0.182	0.861	0.000	cadex
26	2.316	0.608	1.000	fimap
27	2.172	0.633	0.889	fimap
28	3.192	0.595	0.667	fimap
29	3.316	0.879	0.333	fimap
30	2.268	1.212	0.000	fimap
31	3.193	1.522	0.667	fimap
32	1.054	1.331	0.222	wachter
33	0.988	1.255	0.111	wachter
34	0.524	1.002	1.000	wachter
35	0.054	0.843	0.000	wachter
36	0.189	0.883	1.000	wachter
37	0.045	0.834	0.000	wachter
38	0.135	0.882	1.000	wachter
39	0.994	1.271	0.222	wachter
40	0.690	1.148	0.000	wachter
41	0.058	0.846	1.000	wachter
42	0.059	0.848	0.000	cem
43	3.152	2.348	0.000	cfproto
44	4.413	2.433	0.000	cfproto
45	2.138	1.287	0.000	cfproto
46	3.183	1.537	0.333	cfproto
47	3.345	1.700	0.333	cfproto
48	3.409	1.528	0.000	cfproto
49	3.269	1.624	0.333	cfproto
50	3.473	2.513	0.111	cfproto
51	2.138	1.287	0.000	cfproto
52	3.860	1.812	0.778	growing-spheres
53	3.868	2.419	0.333	growing-spheres
54	3.511	1.218	0.000	growing-spheres
55	3.965	1.671	0.000	growing-spheres
56	3.598	2.397	0.222	growing-spheres
57	3.670	2.317	0.778	growing-spheres
58	3.422	1.778	0.111	growing-spheres
59	3.580	2.658	0.556	growing-spheres
60	3.644	2.278	0.111	growing-spheres
61	3.436	1.545	0.000	growing-spheres
62	4.100	2.984	0.778	growing-spheres
63	3.472	1.571	0.222	growing-spheres
64	3.369	1.270	0.000	growing-spheres
65	3.893	1.846	0.222	growing-spheres
66	3.595	2.119	0.333	growing-spheres
67	4.098	2.573	0.111	growing-spheres
68	2.537	1.689	0.222	growing-spheres
69	3.563	2.320	0.556	growing-spheres
70	3.878	1.665	0.556	growing-spheres
71	4.008	2.249	0.444	growing-spheres
72	1.010	1.519	0.222	actionable-recourse
73	1.468	0.056	0.667	face
74	1.745	0.128	1.000	face
75	1.468	0.110	0.111	face
76	2.470	0.028	1.000	face
77	2.427	0.084	1.000	face
78	2.395	0.191	0.111	face
79	1.507	0.058	0.778	face
80	2.443	0.019	1.000	face
81	1.564	0.047	0.667	face
82	2.286	0.121	0.667	face

Table 2. The values of selected quality measures computed for counterfactual explanations in the toy example. IDs correspond to counterfactuals presented in Tab. 1

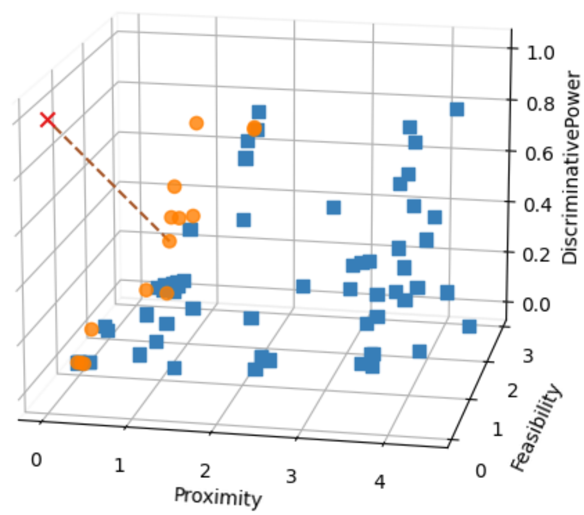

Fig. 1. This plot showcases the generated counterfactuals produced by an ensemble of explainers, visualized over three criteria employed in our proposed approach for multi-criteria analysis. The orange circles represent non-dominated counterfactuals that remain after applying the dominance relation, while the red cross denotes the calculated Ideal Point that our method uses to select the closest solution.

2 Radar plots visualising performance on all metrics

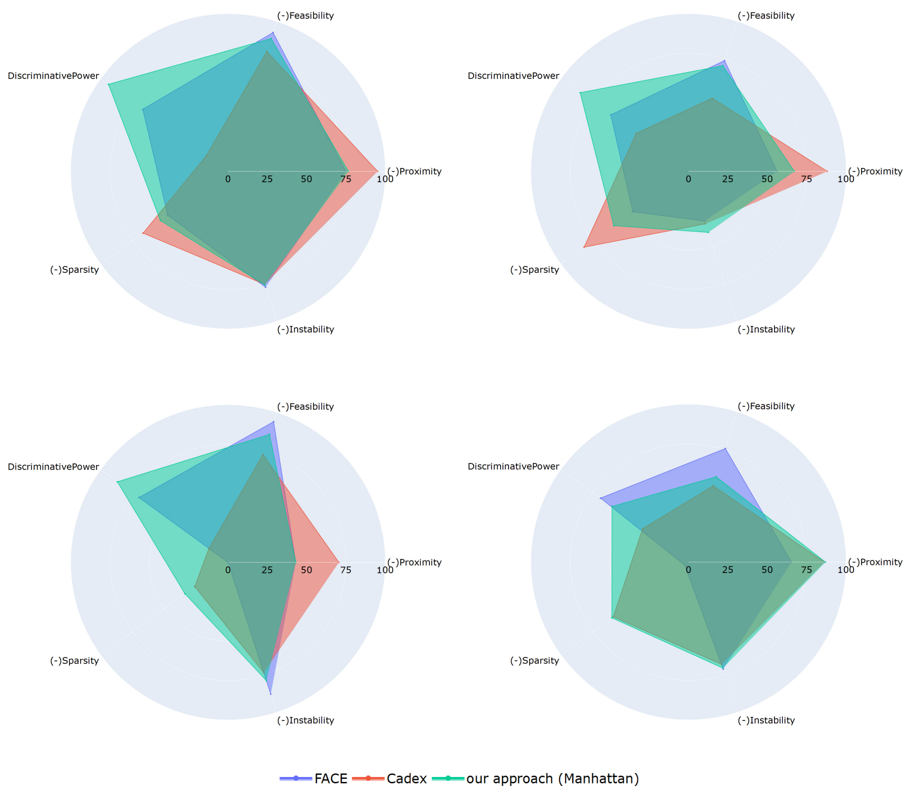

Fig. 3. Radar plots comparing three methods, two well performing explainers and our approach with Manhattan distance for Ideal Point selection. Metrics denoted with (-) normally are the lower the better, but to visualise them on this plot they were inverted as 100%-metric value. Each plot represents distinct dataset: *upper-left* Adult, *upper-right* German, *lower-left* Compas, *lower-right* Fico.

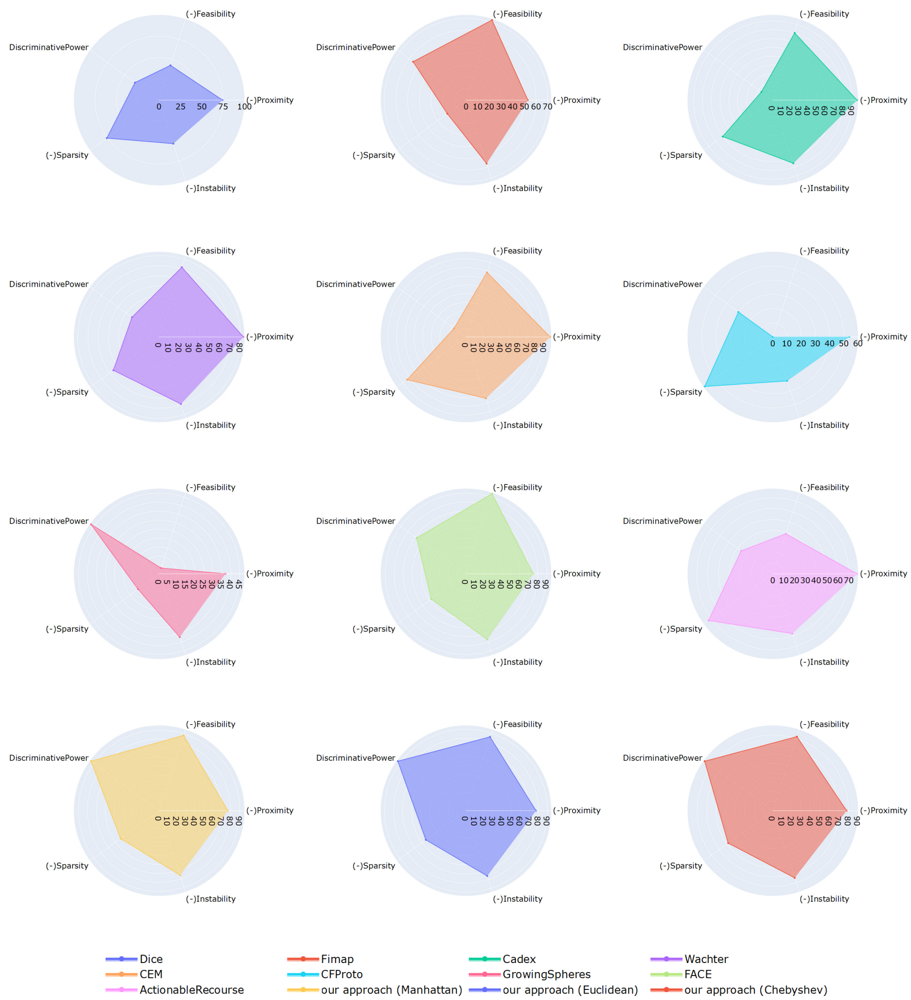

Fig. 4. Side by side performance comparison of all considered explainers in our study on five metrics for Adult dataset. Metrics denoted with (-) normally are the lower the better, but to visualise them on this plot they were inverted as 100%-metric value. Note, that three plots in the last row represent our approach.

3 Detailed specification of black-box models

As a black-box classification method, we employed an artificial neural network consisting of two hidden layers with 16-128 neurons in each layer. The number of neurons was optimized for each dataset separately to maximize the accuracy on the validation set. The training and validation sets were obtained by performing a 90-10 split, where the latter set had balanced classes. After each layer of neurons, except the last one, we used ReLU activation function and a dropout of 0.2 rate. The last layer in all models consisted of two neurons with softmax activation function.

Dataset name	Target feature	Layers
german	class	128-64
adult	income	128-64
fico	RiskPerformance	32-16
compas	two_year_recid	64-32

Table 3. Black-box specification for each dataset

During training, the following values of hyperparameters were used:

- seed: 44
- max epochs: 50
- batch size: 32
- early stopping patience: 10
- loss function: Categorical Crossentropy
- optimizer: Adam (lr: $1e^{-3}$, β_1: 0.9, β_2: 0.999, ϵ: $1e^{-7}$)
- encoding of categorical features: one-hot
- dropout layer drop rate: 0.2
- class weights: (1 / {pos or neg count}) * ({total count} / 2.0)
- input transformation: min-max normalization

4 Analyzing the impact of the elements in our approach - experiment results

In this section, we present detailed results from experiments analysing the importance of particular steps of our approach.

Abbreviations of the method names:

- *s-op* - unweighted sum of criteria, performed only on counterfactuals from the Pareto front
- *s-bd* - unweighted sum of criteria, performed on all counterfactuals that were not removed by the filtering step
- *ip-{manhattan,euclidean,chebyshev}* - Ideal Point method selection with the distance metric specified after the dash
- *rc-op,bd* - random choice, with -op and -bd having the same meaning as for the unweighted sum

Contents:

- Tables 4-7: Results after removing the filtering step
- Tables 8-11: Results after removing the dominance relation enforcement
- Tables 12-15: Nadir variation - results after removing the dominance relation enforcement
- Tables 16-19: Nadir variation - overall results with all steps included

Table 4. Dataset: German. **Setting**: Ensemble - Dominance Relation - Ideal Point Selection. **Removed element**: Filtering

Method	proximity	feasibility	discriminative_power	sparsity	instability	coverage	validity	actionable	rank
dice	1.93	4.02	0.43	2.18	4.34	1	1	1	4.12
cadex	1.34	3.71	0.45	2.73	3.84	1	0.81	1	4
fimap	6.58	2.96	0.76	9.78	3.65	1	0.57	1	2.88
wachter	0.68	3.5	0.61	3.29	3.67	1	0.2	1	3.38
cem	0.39	3.55	0.65	1.61	3.75	1	0.13	1	3.12
cfproto	4.18	4.55	0.48	5.65	4.62	1	0.97	0.87	5.5
growing-spheres	7.38	5.76	0.57	10.43	5.29	1	0.65	1	3.5
actionable-recourse	0.67	3.38	0.46	0.94	3.4	0.36	0.87	0.36	7.62
face	5.05	1.87	0.62	8.19	3.75	1	0.99	0.98	4.5
s-op	3.19	2.12	0.9	5.77	3.41	1	0.64	0.98	3.75
s-bd	2.2	2.63	0.92	4.51	3.58	1	0.48	0.98	3.5
ip-manhattan	2.2	2.63	0.92	0.92	3.58	1	0.48	0.98	3.5
ip-euclidean	1.65	2.83	0.87	3.57	3.66	1	0.53	0.99	3.62
ip-chebyshev	1.76	2.86	0.82	3.48	3.65	1	0.57	0.99	3.75
rc-op	2.86	2.64	0.68	4.95	3.43	1	0.7	1	3
rc-bd	3.9	3.86	0.54	5.99	4.26	1	0.75	1	3.62

Table 5. Dataset: Adult. **Setting**: Ensemble - Dominance Relation - Ideal Point Selection. **Removed element**: Filtering

Method	proximity	feasibility	discriminative_power	sparsity	instability	coverage	validity	actionable	rank
dice	1.14	0.88	0.35	1.74	1.23	1	1	1	3.88
fimap	2.06	0.39	0.67	5.74	1.14	1	0.79	1	3.12
cadex	0.21	0.3	0.19	2.38	0.64	0.99	0.98	0.99	6.12
wachter	0.69	0.43	0.64	3.32	0.83	0.96	0.54	0.96	6
cem	0.13	0.31	0.39	1.16	0.67	1	0.66	1	3.75
cfproto	2.09	1.51	0.31	3.02	1.82	0.99	0.99	0.1	7.38
growing-spheres	2.7	1.38	0.47	6.02	1.59	1	0.79	1	3.62
face	1.16	0.09	0.69	3.8	0.57	1	0.96	0.82	4.5
actionable-recourse	0.66	0.66	0.69	1.72	0.91	0.4	0.7	0.4	6.38
s-op	0.66	0.1	0.94	2.82	0.54	1	0.53	0.9	4.12
s-bd	0.45	0.16	0.99	2.61	0.61	1	0.27	0.94	3.25
ip-manhattan	0.45	0.16	0.99	2.61	0.61	1	0.27	0.94	3.25
ip-euclidean	0.36	0.21	0.97	2.5	0.62	1	0.24	0.97	3.12
ip-chebyshev	0.36	0.22	0.96	2.53	0.62	1	0.24	0.97	3.25
rc-op	0.71	0.2	0.67	2.67	0.6	1	0.73	0.93	4.38
rc-bd	1.51	0.76	0.5	3.73	1.12	1	0.86	0.88	5

Table 6. Dataset: Fico. **Setting**: Ensemble - Dominance Relation - Ideal Point Selection. **Removed element**: Filtering

Method	proximity	feasibility	discriminative_power	sparsity	instability	coverage	validity	actionable	rank
dice	1.14	2.18	0.36	1.99	2.54	1	1	1	4
cadex	0.86	1.7	0.46	7.26	2.04	1	0.85	1	3.75
fimap	1.39	1.74	0.6	15.62	1.85	1	0.85	0.62	4.38
wachter	0.91	1.67	0.47	13.72	1.88	1	0.69	1	3.62
cem	1.12	2.08	0.5	5.8	2.47	1	1	1	3.38
cfproto	0.79	1.55	0.46	10.63	1.82	1	0.7	0.49	5.25
growing-spheres	1.38	1.88	0.53	16.26	2.4	1	0.54	1	3.25
face	2.34	0.83	0.69	17.98	1.82	1	0.99	0.02	4.62
s-op	1.42	1.13	0.93	11.98	1.7	1	0.58	0.48	3.88
s-bd	0.88	1.43	0.9	8	1.86	1	0.36	0.78	3.38
ip-manhattan	0.88	1.43	0.9	8	1.86	1	0.36	0.78	3.38
ip-euclidean	0.79	1.47	0.86	9.04	1.85	1	0.35	0.86	3.38
ip-chebyshev	0.88	1.46	0.8	10.33	1.83	1	0.4	0.89	3.38
rc-op	1.32	1.26	0.68	11.49	1.85	1	0.68	0.56	4.38
rc-bd	1.32	1.77	0.49	11.14	2.22	1	0.78	0.84	4.38

Table 7. Dataset: Compas. **Setting**: Ensemble - Dominance Relation - Ideal Point Selection. **Removed element**: Filtering

Method	proximity	feasibility	discriminative_power	sparsity	instability	coverage	validity	actionable	rank
dice	0.96	0.84	0.36	1.7	1.3	1	1	1	3.88
fimap	0.5	0.11	0.7	3.48	0.18	1	0.92	0.59	4.38
cadex	0.26	0.22	0.32	2.66	0.33	1	0.81	0.99	4.5
wachter	0.24	0.17	0.46	2.26	0.22	1	0.62	1	3.62
cem	0.33	0.32	0.29	1.57	0.38	1	1	1	4.12
cfproto	0.75	0.19	0.51	2.92	0.35	1	0.6	0.35	5.12
growing-spheres	0.26	0.18	0.67	2.96	0.3	1	0.42	1	3.25
actionable-recourse	0.01	0.08	0.44	0.14	0.15	0.03	0.98	0.03	7.5
face	0.54	0.05	0.69	3.54	0.17	1	1	0.04	4.88
s-op	0.29	0.07	0.96	2.68	0.19	1	0.38	0.65	4
s-bd	0.22	0.1	0.99	2.38	0.22	1	0.22	0.83	2.75
ip-manhattan	0.22	0.1	0.99	2.38	0.22	1	0.22	0.83	2.75
ip-euclidean	0.22	0.1	0.99	2.48	0.22	1	0.2	0.83	2.75
ip-chebyshev	0.22	0.1	0.98	2.5	0.22	1	0.2	0.82	3.5
rc-op	0.39	0.07	0.73	2.98	0.16	1	0.75	0.44	4.38
rc-bd	0.5	0.3	0.54	2.63	0.5	1	0.77	0.79	4.38

Table 8. Dataset: German. **Setting**: Ensemble - Filtering - Ideal Point Selection. **Removed element**: Dominance Relation

Method	**proximity**	**feasibility**	**discriminative_power**	**sparsity**	**instability**	**coverage**	**actionable**	**rank**
dice	1.92	4.11	0.46	2.18	4.3	1	1	3.38
cadex	1.35	3.76	0.41	2.61	3.93	0.97	0.97	5.62
fimap	6.75	3.07	0.55	9.79	3.79	0.97	0.97	5
wachter	1.27	3.73	0.34	3.97	3.98	0.3	0.3	6.5
cem	0.62	4.18	0.31	2.15	3.99	0.13	0.13	7.12
cfproto	4.05	4.54	0.49	5.43	4.74	0.93	0.93	5.75
growing-spheres	7.51	5.82	0.59	10.56	5.6	1	1	2.88
actionable-recourse	1.01	3.55	0.44	1.39	3.6	0.23	0.23	6.5
face	5.11	1.89	0.62	8.16	3.79	0.99	0.99	4.5
ip-manhattan	3.83	2.15	0.85	6.06	3.5	1	1	2.25
ip-euclidean	3.21	2.46	0.8	4.99	3.68	1	1	2.5
ip-chebyshev	2.9	2.7	0.74	4.38	3.71	1	1	2.62
s-bd	3.83	2.15	0.85	6.06	3.5	1	1	2.25
rc-bd	4.51	4.21	0.5	6.38	4.57	1	1	3.12

Table 9. Dataset: Adult. **Setting**: Ensemble - Filtering - Ideal Point Selection. **Removed element**: Dominance Relation

Method	**proximity**	**feasibility**	**discriminative_power**	**sparsity**	**instability**	**coverage**	**actionable**	**rank**
dice	1.14	0.86	0.34	1.7	1.2	1	1	3.5
fimap	2.06	0.38	0.57	5.72	1.15	0.99	0.99	4.38
cadex	0.2	0.3	0.17	2.29	0.64	0.99	0.99	5.25
wachter	0.72	0.41	0.32	3.16	0.81	0.81	0.81	6.12
cem	0.13	0.32	0.17	1.16	0.67	0.66	0.66	6.5
cfproto	1.12	0.66	0.36	2.09	1.17	0.39	0.39	6.25
growing-spheres	2.76	1.34	0.45	6.08	1.63	0.99	0.99	4.5
face	1.02	0.11	0.68	3.67	0.63	0.93	0.93	5
actionable-recourse	0.98	0.9	0.36	1.92	1.1	0.1	0.1	6.5
ip-manhattan	1.02	0.17	0.94	3.34	0.63	1	1	2.25
ip-euclidean	0.99	0.21	0.93	3.22	0.63	1	1	2.5
ip-chebyshev	0.96	0.26	0.89	2.97	0.66	1	1	2.62
s-bd	1.02	0.17	0.94	3.34	0.63	1	1	2.25
rc-bd	1.54	0.78	0.41	3.74	1.15	1	1	3.12

Table 10. Dataset: Fico. **Setting**: Ensemble - Filtering - Ideal Point Selection. **Removed element**: Dominance Relation

Method	proximity	feasibility	discriminative_power	sparsity	instability	coverage	actionable	rank
dice	1.15	2.17	0.36	2.1	2.54	1	1	3.5
cadex	0.94	1.71	0.38	7.82	2.03	1	1	3.38
fimap	1.63	1.82	0.62	15.97	1.91	0.62	0.62	5.12
wachter	0.91	1.66	0.33	13.66	1.88	0.99	0.99	5.75
cem	1.12	2.08	0.5	5.8	2.47	1	1	3
cfproto	0.66	1.54	0.42	9.32	1.77	0.58	0.58	5.88
growing-spheres	1.43	1.91	0.34	16.33	2.41	0.97	0.97	5.88
face	2.08	0.84	0.66	16.54	1.82	0.24	0.24	5.25
ip-manhattan	0.87	1.51	0.6	7.33	1.89	1	1	2.75
ip-euclidean	0.9	1.59	0.63	7.48	1.93	1	1	2.38
ip-chebyshev	0.99	1.66	0.63	7.11	2.03	1	1	2.38
s-bd	0.87	1.51	0.6	7.33	1.89	1	1	2.75
rc-bd	1.12	1.92	0.39	7.96	2.28	1	1	3.25

Table 11. Dataset: Compas. **Setting**: Ensemble - Filtering - Ideal Point Selection. **Removed element**: Dominance Relation

Method	proximity	feasibility	discriminative_power	sparsity	instability	coverage	actionable	rank
dice	0.91	0.76	0.37	1.67	1.21	1	1	3.12
fimap	0.49	0.1	0.68	3.12	0.18	0.58	0.58	5.5
cadex	0.29	0.24	0.15	2.78	0.36	0.97	0.97	5.62
wachter	0.28	0.19	0.22	2.44	0.25	0.77	0.77	5.88
cem	0.33	0.32	0.29	1.57	0.38	1	1	3.38
cfproto	0.16	0.14	0.24	1.92	0.26	0.21	0.21	6.5
growing-spheres	0.3	0.19	0.18	3.15	0.33	0.8	0.8	5.75
actionable-recourse	0.07	0.32	0.89	1	0.66	0	0	5.5
face	0.38	0.03	0.71	2.75	0.13	0.27	0.27	5.62
ip-manhattan	0.54	0.12	0.87	2.4	0.28	1	1	2.38
ip-euclidean	0.55	0.14	0.86	2.45	0.28	1	1	2.62
ip-chebyshev	0.55	0.17	0.84	2.44	0.3	1	1	2.75
s-bd	0.54	0.12	0.87	2.4	0.28	1	1	2.38
rc-bd	0.62	0.48	0.31	2.36	0.76	1	1	3.25

Table 12. Dataset: German. **Setting**: Ensemble - Filtering - Nadir Selection. **Removed element**: Dominance Relation

Method	**proximity**	**feasibility**	**discriminative_power**	**sparsity**	**instability**	**coverage**	**actionable**	**rank**
dice	1.91	4.13	0.4	2.16	4.32	1	1	3.71
cadex	1.42	3.79	0.42	2.75	3.92	0.97	0.97	6.43
fimap	6.93	3.12	0.6	9.97	3.86	0.97	0.97	5.57
wachter	1.26	3.78	0.3	3.93	3.96	0.3	0.3	7.71
cem	0.62	4.18	0.31	2.15	3.99	0.13	0.13	8.14
cfproto	4.25	4.46	0.49	5.89	4.72	0.93	0.93	6.43
growing-spheres	7.43	5.61	0.57	10.58	5.47	1	1	2.86
actionable-recourse	1.01	3.55	0.44	1.39	3.6	0.23	0.23	7.43
face	5.15	1.86	0.61	8.22	3.85	0.99	0.99	5.14
ip-manhattan	3.83	2.15	0.85	6.06	3.5	1	1	1.86
ip-euclidean	3.32	2.47	0.83	5.21	3.62	1	1	2.14
ip-chebyshev	3.08	2.67	0.79	4.73	3.67	1	1	2.29
s-op	1.85	3.99	0.45	2.2	4.34	1	1	3.29
s-bd	3.83	2.15	0.85	6.06	3.5	1	1	1.86
rc-op	3.49	2.76	0.64	5.39	3.56	1	1	2.43
rc-bd	3.68	3.99	0.47	5.33	4.34	1	1	3.14

Table 13. Dataset: Adult. **Setting**: Ensemble - Filtering - Nadir Selection. **Removed element**: Dominance Relation

Method	**proximity**	**feasibility**	**discriminative_power**	**sparsity**	**instability**	**coverage**	**actionable**	**rank**
dice	1.17	0.86	0.41	1.74	1.24	1	1	3
fimap	2.14	0.4	0.61	5.82	1.21	0.99	0.99	4.86
cadex	0.2	0.29	0.16	2.15	0.64	0.99	0.99	6.29
wachter	0.74	0.42	0.3	3.11	0.85	0.81	0.81	7.14
cem	0.13	0.32	0.17	1.16	0.67	0.66	0.66	7.57
cfproto	1.15	0.67	0.35	2.09	1.22	0.39	0.39	7.57
growing-spheres	2.72	1.37	0.44	6.1	1.64	0.99	0.99	5.14
face	1.03	0.11	0.71	3.66	0.62	0.93	0.93	5.57
actionable-recourse	0.98	0.9	0.36	1.92	1.1	0.1	0.1	7.71
ip-manhattan	1.02	0.17	0.94	3.34	0.63	1	1	2.14
ip-euclidean	1.03	0.21	0.95	3.32	0.65	1	1	1.86
ip-chebyshev	1.03	0.25	0.95	3.26	0.68	1	1	1.86
s-op	1.15	0.91	0.38	1.77	1.23	1	1	3.29
s-bd	1.02	0.17	0.94	3.34	0.63	1	1	2.14
rc-op	0.73	0.27	0.52	2.73	0.72	1	1	2.71
rc-bd	1.6	0.84	0.41	3.58	1.23	1	1	3

Table 14. Dataset: Fico. **Setting**: Ensemble - Filtering - Nadir Selection. **Removed element**: Dominance Relation

Method	proximity	feasibility	discriminative_power	sparsity	instability	coverage	actionable	rank
dice	1.2	2.23	0.39	2.08	2.62	1	1	3.14
cadex	0.94	1.71	0.38	8.03	2.04	1	1	3.43
fimap	1.72	1.84	0.65	16.02	1.92	0.62	0.62	5.43
wachter	0.9	1.66	0.34	13.63	1.87	0.99	0.99	6.57
cem	1.12	2.08	0.5	5.8	2.47	1	1	2.86
cfproto	0.67	1.54	0.41	9.39	1.77	0.58	0.58	6.71
growing-spheres	1.41	1.89	0.34	16.29	2.4	0.97	0.97	6.86
face	2.1	0.84	0.65	16.54	1.78	0.24	0.24	6
ip-manhattan	0.87	1.51	0.6	7.33	1.89	1	1	2.43
ip-euclidean	0.91	1.59	0.64	7.18	1.93	1	1	2.29
ip-chebyshev	0.98	1.66	0.66	7.13	1.97	1	1	1.86
s-op	1.12	2.15	0.37	2	2.48	1	1	3.57
s-bd	0.87	1.51	0.6	7.33	1.89	1	1	2.43
rc-op	1.03	1.66	0.52	8.1	1.99	1	1	2.71
rc-bd	1.17	1.99	0.39	7.89	2.4	1	1	3.14

Table 15. Dataset: Compas. **Setting**: Ensemble - Filtering - Nadir Selection. **Removed element**: Dominance Relation

Method	proximity	feasibility	discriminative_power	sparsity	instability	coverage	actionable	rank
dice	0.97	0.83	0.36	1.74	1.35	1	1	3.14
fimap	0.49	0.1	0.69	3.14	0.17	0.58	0.58	6.14
cadex	0.29	0.23	0.14	2.65	0.35	0.97	0.97	6.57
wachter	0.27	0.2	0.22	2.42	0.24	0.77	0.77	6.86
cem	0.33	0.32	0.29	1.57	0.38	1	1	3.43
cfproto	0.16	0.14	0.24	1.92	0.25	0.21	0.21	7.57
growing-spheres	0.3	0.19	0.18	3.19	0.33	0.8	0.8	6.71
actionable-recourse	0.07	0.32	0.89	1	0.66	0	0	6.14
face	0.39	0.03	0.72	2.78	0.13	0.27	0.27	6.29
ip-manhattan	0.54	0.12	0.87	2.4	0.28	1	1	2.29
ip-euclidean	0.57	0.14	0.88	2.39	0.29	1	1	2.14
ip-chebyshev	0.58	0.16	0.89	2.4	0.31	1	1	1.86
s-op	0.92	0.8	0.39	1.7	1.17	1	1	3
s-bd	0.54	0.12	0.87	2.4	0.28	1	1	2.29
rc-op	0.42	0.2	0.5	2.21	0.32	1	1	2.86
rc-bd	0.66	0.51	0.36	2.36	0.82	1	1	3.14

Table 16. Dataset: German. **Setting**: Ensemble - Filtering - Dominance Relation - Nadir Selection. **Removed element**: None

Method	proximity	feasibility	discriminative_power	sparsity	instability	coverage	actionable	rank
dice	1.92	4.11	0.46	2.18	4.3	1	1	3.14
cadex	1.35	3.76	0.41	2.61	3.93	0.97	0.97	5.71
fimap	6.75	3.07	0.55	9.79	3.79	0.97	0.97	5
wachter	1.27	3.73	0.34	3.97	3.98	0.3	0.3	6.71
cem	0.62	4.18	0.31	2.15	3.99	0.13	0.13	7.43
cfproto	4.05	4.54	0.49	5.43	4.74	0.93	0.93	5.86
growing-spheres	7.51	5.82	0.59	10.56	5.6	1	1	2.57
actionable-recourse	1.01	3.55	0.44	1.39	3.6	0.23	0.23	6.71
face	5.11	1.89	0.62	8.16	3.79	0.99	0.99	4.43
ip-manhattan	3.83	2.15	0.85	6.06	3.5	1	1	1.86
ip-euclidean	3.34	2.38	0.82	5.27	3.57	1	1	2.14
ip-chebyshev	2.87	2.65	0.76	4.43	3.74	1	1	2.29
s-bd	3.83	2.15	0.85	6.06	3.5	1	1	1.86
rc-bd	4.51	4.21	0.5	6.38	4.57	1	1	2.86

Table 17. Dataset: Adult. **Setting**: Ensemble - Filtering - Dominance Relation - Nadir Selection. **Removed element**: None

Method	proximity	feasibility	discriminative_power	sparsity	instability	coverage	actionable	rank
dice	1.14	0.86	0.34	1.7	1.2	1	1	3.29
fimap	2.06	0.38	0.57	5.72	1.15	0.99	0.99	4.29
cadex	0.2	0.3	0.17	2.29	0.64	0.99	0.99	5.29
wachter	0.72	0.41	0.32	3.16	0.81	0.81	0.81	6.29
cem	0.13	0.32	0.17	1.16	0.67	0.66	0.66	6.71
cfproto	1.12	0.66	0.36	2.09	1.17	0.39	0.39	6.43
growing-spheres	2.76	1.34	0.45	6.08	1.63	0.99	0.99	4.43
face	1.02	0.11	0.68	3.67	0.63	0.93	0.93	5
actionable-recourse	0.98	0.9	0.36	1.92	1.1	0.1	0.1	6.71
ip-manhattan	1.02	0.17	0.94	3.34	0.63	1	1	1.86
ip-euclidean	0.99	0.19	0.94	3.23	0.63	1	1	1.86
ip-chebyshev	0.98	0.24	0.91	3.14	0.65	1	1	2.29
s-bd	1.02	0.17	0.94	3.34	0.63	1	1	1.86
rc-bd	1.54	0.78	0.41	3.74	1.15	1	1	2.86

Table 18. Dataset: Fico. **Setting**: Ensemble - Filtering - Dominance Relation - Nadir Selection. **Removed element**: None

Method	proximity	feasibility	discriminative_power	sparsity	instability	coverage	actionable	rank
dice	1.15	2.17	0.36	2.1	2.54	1	1	3.29
cadex	0.94	1.71	0.38	7.82	2.03	1	1	3.14
fimap	1.63	1.82	0.62	15.97	1.91	0.62	0.62	5.14
wachter	0.91	1.66	0.33	13.66	1.88	0.99	0.99	5.86
cem	1.12	2.08	0.5	5.8	2.47	1	1	2.71
cfproto	0.66	1.54	0.42	9.32	1.77	0.58	0.58	6
growing-spheres	1.43	1.91	0.34	16.33	2.41	0.97	0.97	6
face	2.08	0.84	0.66	16.54	1.82	0.24	0.24	5.29
ip-manhattan	0.87	1.51	0.6	7.33	1.89	1	1	2.43
ip-euclidean	0.88	1.57	0.63	7.28	1.92	1	1	2
ip-chebyshev	0.94	1.6	0.63	7.28	1.93	1	1	2
s-bd	0.87	1.51	0.6	7.33	1.89	1	1	2.43
rc-bd	1.12	1.92	0.39	7.96	2.28	1	1	3

Table 19. Dataset: Compas. **Setting**: Ensemble - Filtering - Dominance Relation - Nadir Selection. **Removed element**: None

Method	proximity	feasibility	discriminative_power	sparsity	instability	coverage	actionable	rank
dice	0.91	0.76	0.37	1.67	1.21	1	1	2.86
fimap	0.49	0.1	0.68	3.12	0.18	0.58	0.58	5.57
cadex	0.29	0.24	0.15	2.78	0.36	0.97	0.97	5.71
wachter	0.28	0.19	0.22	2.44	0.25	0.77	0.77	6
cem	0.33	0.32	0.29	1.57	0.38	1	1	3.14
cfproto	0.16	0.14	0.24	1.92	0.26	0.21	0.21	6.71
growing-spheres	0.3	0.19	0.18	3.15	0.33	0.8	0.8	5.86
actionable-recourse	0.07	0.32	0.89	1	0.66	0	0	5.57
face	0.38	0.03	0.71	2.75	0.13	0.27	0.27	5.71
ip-manhattan	0.54	0.12	0.87	2.4	0.28	1	1	2
ip-euclidean	0.54	0.13	0.86	2.43	0.28	1	1	2.29
ip-chebyshev	0.55	0.15	0.86	2.48	0.28	1	1	2.29
s-bd	0.54	0.12	0.87	2.4	0.28	1	1	2
rc-bd	0.62	0.48	0.31	2.36	0.76	1	1	3