Output tables for the test of Multiple comparisons.

November 30, 2021

1 Average rankings of Friedman test

Average ranks obtained by applying the Friedman procedure

Friedman statistic considering reduction performance (distributed according to chi-square with 10 degrees of freedom: 128.608392.

P-value computed by Friedman Test: 1.0023526453295517E-10.

Algorithm	Ranking
best-precision	3.0385
best-recall	6.7115
balanced	4.9231
promethee-precision	3.0385
promethee-recall	6.7115
bac	6.6731
precision	3.0192
recall	10.6346
f1	6.6731
auc	7.0577
gmean	7.5192

Table 1: Average Rankings of the algorithms

2 Post hoc comparisons

Results achieved on post hoc comparisons for $\alpha = 0.05$, $\alpha = 0.10$ and adjusted p-values.

2.1 P-values for $\alpha = 0.05$

Nemenyi's procedure rejects those hypotheses that have an unadjusted p-value ≤ 0.000909.

i	algorithms	$z = (R_0 - R_i)/SE$	p
55	precision vs. recall	8.278796	0
54	best-precision vs. recall	8.25789	0
53	promethee-precision vs. recall	8.25789	0
52	balanced vs. recall	6.209097	0
51	precision vs. gmean	4.892016	0.000001
50	best-precision vs. gmean	4.87111	0.000001
49	promethee-precision vs. gmean	4.87111	0.000001
48	precision vs. auc	4.390271	0.000011
47	best-precision vs. auc	4.369365	0.000012
46	promethee-precision vs. auc	4.369365	0.000012
45	bac vs. recall	4.306646	0.000017
44	recall vs. f1	4.306646	0.000017
43	best-recall vs. recall	4.264834	0.00002
42	promethee-recall vs. recall	4.264834	0.00002
41	best-recall vs. precision	4.013962	0.00006
40	promethee-recall vs. precision	4.013962	0.00006
39	best-precision vs. best-recall	3.993056	0.000065
38	best-precision vs. promethee-recall	3.993056	0.000065
37	best-recall vs. promethee-precision	3.993056	0.000065
36	promethee-precision vs. promethee-recall	3.993056	0.000065
35	precision vs. f1	3.97215	0.000071
34	bac vs. precision	3.97215	0.000071
33	best-precision vs. f1	3.951243	0.000078
32	promethee-precision vs. f1	3.951243	0.000078
31	best-precision vs. bac	3.951243	0.000078
30	promethee-precision vs. bac	3.951243	0.000078
29	recall vs. auc	3.888525	0.000101
28	recall vs. gmean	3.38678	0.000707
27	balanced vs. gmean	2.822317	0.004768
26	balanced vs. auc	2.320572	0.02031
25	balanced vs. precision	2.069699	0.038481
24	best-precision vs. balanced	2.048793	0.040482
23	balanced vs. promethee-precision	2.048793	0.040482
22	best-recall vs. balanced	1.944263	0.051864
21	balanced vs. promethee-recall	1.944263	0.051864
20	balanced vs. f1	1.902451	0.057112
19	balanced vs. bac	1.902451	0.057112
18	bac vs. gmean	0.919866	0.357643
17	f1 vs. gmean	0.919866	0.357643
16	best-recall vs. gmean	0.878054	0.379914
15	promethee-recall vs. gmean	0.878054	0.379914
14	auc vs. gmean	0.501745	0.615847
13	bac vs. auc	0.418121	0.675859
12	f1 vs. auc	0.418121	0.675859
11	best-recall vs. auc	0.376309	0.706687
10	promethee-recall vs. auc	0.376309	0.706687
9	best-recall vs. bac	0.041812	0.966648
8	promethee-recall vs. bac	0.041812	0.966648
7	best-recall vs. f1	0.041812	0.966648
6	promethee-recall vs. f1	0.041812	0.966648
5	best-precision vs. precision	0.020906	0.983321
4	promethee-precision vs. precision	0.020906	0.983321
3	bac vs. f1	0	1
2	best-precision vs. promethee-precision	0	1
1	best-recall vs. promethee-recall	0	1

Table 2: P-values Table for $\alpha = 0.05$

2.2 P-values for $\alpha = 0.10$

Nemenyi's procedure rejects those hypotheses that have an unadjusted p-value ≤ 0.001818.

i	algorithms	$z = (R_0 - R_i)/SE$	p
55	precision vs. recall	8.278796	0
54	best-precision vs. recall	8.25789	0
53	promethee-precision vs. recall	8.25789	0
52	balanced vs. recall	6.209097	0
51	precision vs. gmean	4.892016	0.000001
50	best-precision vs. gmean	4.87111	0.000001
49	promethee-precision vs. gmean	4.87111	0.000001
48	precision vs. auc	4.390271	0.000011
47	best-precision vs. auc	4.369365	0.000012
46	promethee-precision vs. auc	4.369365	0.000012
45	bac vs. recall	4.306646	0.000017
44	recall vs. f1	4.306646	0.000017
43	best-recall vs. recall	4.264834	0.00002
42	promethee-recall vs. recall	4.264834	0.00002
41	best-recall vs. precision	4.013962	0.00006
40	promethee-recall vs. precision	4.013962	0.00006
39	best-precision vs. best-recall	3.993056	0.000065
38	best-precision vs. promethee-recall	3.993056	0.000065
37	best-recall vs. promethee-precision	3.993056	0.000065
36	promethee-precision vs. promethee-recall	3.993056	0.000065
35	precision vs. f1	3.97215	0.000071
34	bac vs. precision	3.97215	0.000071
33	best-precision vs. f1	3.951243	0.000078
32	promethee-precision vs. f1	3.951243	0.000078
31	best-precision vs. bac	3.951243	0.000078
30	promethee-precision vs. bac	3.951243	0.000078
29	recall vs. auc	3.888525	0.000101
28	recall vs. gmean	3.38678	0.000707
27	balanced vs. gmean	2.822317	0.004768
26	balanced vs. auc	2.320572	0.02031
25	balanced vs. precision	2.069699	0.038481
24	best-precision vs. balanced	2.048793	0.040482
23	balanced vs. promethee-precision	2.048793	0.040482
22	best-recall vs. balanced	1.944263	0.051864
21	balanced vs. promethee-recall	1.944263	0.051864
20	balanced vs. f1	1.902451	0.057112
19	balanced vs. bac	1.902451	0.057112
18	bac vs. gmean	0.919866	0.357643
17	f1 vs. gmean	0.919866	0.357643
16	best-recall vs. gmean	0.878054	0.379914
15	promethee-recall vs. gmean	0.878054	0.379914
14	auc vs. gmean	0.501745	0.615847
13	bac vs. auc	0.418121	0.675859
12	f1 vs. auc	0.418121	0.675859
11	best-recall vs. auc	0.376309	0.706687
10	promethee-recall vs. auc	0.376309	0.706687
9	best-recall vs. bac	0.041812	0.966648
8	promethee-recall vs. bac	0.041812	0.966648
7	best-recall vs. f1	0.041812	0.966648
6	promethee-recall vs. f1	0.041812	0.966648
5	best-precision vs. precision	0.020906	0.983321
4	promethee-precision vs. precision	0.020906	0.983321
3	bac vs. f1	0	1
2	best-precision vs. promethee-precision	0	1
1	best-recall vs. promethee-recall	0	1

Table 3: P-values Table for $\alpha = 0.10$

2.3 Adjusted p-values

i	hypothesis	unadjusted p	p_{Neme}
1	precision vs . recall	0	0
2	best-precision vs . recall	0	0
3	promethee-precision vs . recall	0	0
4	balanced vs . recall	0	0
5	precision vs . gmean	0.000001	0.000055
6	best-precision vs . gmean	0.000001	0.000061
7	promethee-precision vs . gmean	0.000001	0.000061
8	precision vs . auc	0.000011	0.000623
9	best-precision vs . auc	0.000012	0.000685
10	promethee-precision vs . auc	0.000012	0.000685
11	bac vs . recall	0.000017	0.000912
12	recall vs . f1	0.000017	0.000912
13	best-recall vs . recall	0.00002	0.0011
14	promethee-recall vs . recall	0.00002	0.0011
15	best-recall vs . precision	0.00006	0.003284
16	promethee-recall vs . precision	0.00006	0.003284
17	best-precision vs . best-recall	0.000065	0.003587
18	best-precision vs . promethee-recall	0.000065	0.003587
19	best-recall vs . promethee-precision	0.000065	0.003587
20	promethee-precision vs . promethee-recall	0.000065	0.003587
21	precision vs . f1	0.000071	0.003917
22	bac vs . precision	0.000071	0.003917
23	best-precision vs . f1	0.000078	0.004276
24	promethee-precision vs . f1	0.000078	0.004276
25	best-precision vs . bac	0.000078	0.004276
26	promethee-precision vs . bac	0.000078	0.004276
27	recall vs . auc	0.000101	0.005547
28	recall vs . gmean	0.000707	0.038895
29	balanced vs . gmean	0.004768	0.262229
30	balanced vs . auc	0.02031	1.117049
31	balanced vs . precision	0.038481	2.11643
32	best-precision vs . balanced	0.040482	2.22653
33	balanced vs . promethee-precision	0.040482	2.22653
34	best-recall vs . balanced	0.051864	2.852508
35	balanced vs . promethee-recall	0.051864	2.852508
36	balanced vs . f1	0.057112	3.141175
37	balanced vs . bac	0.057112	3.141175
38	bac vs . gmean	0.357643	19.670347
39	f1 vs . gmean	0.357643	19.670347
40	best-recall vs . gmean	0.379914	20.89529
41	promethee-recall vs . gmean	0.379914	20.89529
42	auc vs . gmean	0.615847	33.871572
43	bac vs . auc	0.675859	37.172226
44	f1 vs . auc	0.675859	37.172226
45	best-recall vs . auc	0.706687	38.867799
46	promethee-recall vs . auc	0.706687	38.867799
47	best-recall vs . bac	0.966648	53.165667
48	promethee-recall vs . bac	0.966648	53.165667
49	best-recall vs . f1	0.966648	53.165667
50	promethee-recall vs . f1	0.966648	53.165667
51	best-precision vs . precision	0.983321	54.082633
52	promethee-precision vs . precision	0.983321	54.082633
53	bac vs . f1	1	55
54	best-precision vs . promethee-precision	1	55
55	best-recall vs . promethee-recall	1	55

Table 4: Adjusted p-values