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: 69.463287.

P-value computed by Friedman Test: 9.257938859974502E-11.

Algorithm	Ranking
best-precision	7.9808
best-recall	4.0577
balanced	4.5385
promethee-precision	7.9808
promethee-recall	4.0577
bac	5.2885
precision	9.0962
recall	6.6346
f1	5.8269
auc	5.6346
gmean	4.9038

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	best-recall vs. precision	5.477385	0
54	promethee-recall vs. precision	5.477385	0
53	balanced vs. precision	4.954734	0.000001
52	precision vs. gmean	4.557519	0.000005
51	best-precision vs. best-recall	4.264834	0.00002
50	best-precision vs. promethee-recall	4.264834	0.00002
49	best-recall vs. promethee-precision	4.264834	0.00002
48	promethee-precision vs. promethee-recall	4.264834	0.00002
47	bac vs. precision	4.139398	0.000035
46	precision vs. auc	3.763089	0.000168
45	best-precision vs. balanced	3.742183	0.000182
44	balanced vs. promethee-precision	3.742183	0.000182
43	precision vs. f1	3.554029	0.000379
42	best-precision vs. gmean	3.344968	0.000823
41	promethee-precision vs. gmean	3.344968	0.000823
40	best-precision vs. bac	2.926847	0.003424
39	promethee-precision vs. bac	2.926847	0.003424
38	best-recall vs. recall	2.801411	0.005088
37	promethee-recall vs. recall	2.801411	0.005088
36	precision vs. recall	2.675974	0.007451
35	best-precision vs. auc	2.550538	0.010756
34	promethee-precision vs. auc	2.550538	0.010756
33	best-precision vs. f1	2.341478	0.019208
32	promethee-precision vs. f1	2.341478	0.019208
31	balanced vs. recall	2.278759	0.022681
30	best-recall vs. f1	1.923357	0.054435
29	promethee-recall vs. f1	1.923357	0.054435
28	recall vs. gmean	1.881545	0.059898
27	best-recall vs. auc	1.714296	0.086474
26	promethee-recall vs. auc	1.714296	0.086474
25	bac vs. recall	1.463424	0.143352
24	best-precision vs. recall	1.463424	0.143352
23	promethee-precision vs. recall	1.463424	0.143352
22	balanced vs. f1	1.400705	0.161302
21	best-recall vs. bac	1.337987	0.180901
20	promethee-recall vs. bac	1.337987	0.180901
19	best-precision vs. precision	1.212551	0.225302
18	promethee-precision vs. precision	1.212551	0.225302
17	balanced vs. auc	1.191645	0.233401
16	recall vs. auc	1.087115	0.276986
15	f1 vs. gmean	1.00349	0.315624
14	best-recall vs. gmean	0.919866	0.357643
13	promethee-recall vs. gmean	0.919866	0.357643
12	recall vs. f1	0.878054	0.379914
11	balanced vs. bac	0.815336	0.41488
10	auc vs. gmean	0.79443	0.426945
9	bac vs. f1	0.585369	0.558299
8	best-recall vs. balanced	0.522651	0.601217
7	balanced vs. promethee-recall	0.522651	0.601217
6	bac vs. gmean	0.418121	0.675859
5	balanced vs. gmean	0.397215	0.691209
4	bac vs. auc	0.376309	0.706687
3	f1 vs. auc	0.209061	0.834401
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	best-recall vs. precision	5.477385	0
54	promethee-recall vs. precision	5.477385	0
53	balanced vs. precision	4.954734	0.000001
52	precision vs. gmean	4.557519	0.000005
51	best-precision vs. best-recall	4.264834	0.00002
50	best-precision vs. promethee-recall	4.264834	0.00002
49	best-recall vs. promethee-precision	4.264834	0.00002
48	promethee-precision vs. promethee-recall	4.264834	0.00002
47	bac vs. precision	4.139398	0.000035
46	precision vs. auc	3.763089	0.000168
45	best-precision vs. balanced	3.742183	0.000182
44	balanced vs. promethee-precision	3.742183	0.000182
43	precision vs. f1	3.554029	0.000379
42	best-precision vs. gmean	3.344968	0.000823
41	promethee-precision vs. gmean	3.344968	0.000823
40	best-precision vs. bac	2.926847	0.003424
39	promethee-precision vs. bac	2.926847	0.003424
38	best-recall vs. recall	2.801411	0.005088
37	promethee-recall vs. recall	2.801411	0.005088
36	precision vs. recall	2.675974	0.007451
35	best-precision vs. auc	2.550538	0.010756
34	promethee-precision vs. auc	2.550538	0.010756
33	best-precision vs. f1	2.341478	0.019208
32	promethee-precision vs. f1	2.341478	0.019208
31	balanced vs. recall	2.278759	0.022681
30	best-recall vs. f1	1.923357	0.054435
29	promethee-recall vs. f1	1.923357	0.054435
28	recall vs. gmean	1.881545	0.059898
27	best-recall vs. auc	1.714296	0.086474
26	promethee-recall vs. auc	1.714296	0.086474
25	bac vs. recall	1.463424	0.143352
24	best-precision vs. recall	1.463424	0.143352
23	promethee-precision vs. recall	1.463424	0.143352
22	balanced vs. f1	1.400705	0.161302
21	best-recall vs. bac	1.337987	0.180901
20	promethee-recall vs. bac	1.337987	0.180901
19	best-precision vs. precision	1.212551	0.225302
18	promethee-precision vs. precision	1.212551	0.225302
17	balanced vs. auc	1.191645	0.233401
16	recall vs. auc	1.087115	0.276986
15	f1 vs. gmean	1.00349	0.315624
14	best-recall vs. gmean	0.919866	0.357643
13	promethee-recall vs. gmean	0.919866	0.357643
12	recall vs. f1	0.878054	0.379914
11	balanced vs. bac	0.815336	0.41488
10	auc vs. gmean	0.79443	0.426945
9	bac vs. f1	0.585369	0.558299
8	best-recall vs. balanced	0.522651	0.601217
7	balanced vs. promethee-recall	0.522651	0.601217
6	bac vs. gmean	0.418121	0.675859
5	balanced vs. gmean	0.397215	0.691209
4	bac vs. auc	0.376309	0.706687
3	f1 vs. auc	0.209061	0.834401
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	best-recall vs . precision	0	0.000002
2	promethee-recall vs . precision	0	0.000002
3	balanced vs . precision	0.000001	0.00004
4	precision vs . gmean	0.000005	0.000285
5	best-precision vs . best-recall	0.00002	0.0011
6	best-precision vs . promethee-recall	0.00002	0.0011
7	best-recall vs . promethee-precision	0.00002	0.0011
8	promethee-precision vs . promethee-recall	0.00002	0.0011
9	bac vs . precision	0.000035	0.001915
10	precision vs . auc	0.000168	0.009231
11	best-precision vs . balanced	0.000182	0.010034
12	balanced vs . promethee-precision	0.000182	0.010034
13	precision vs . f1	0.000379	0.020866
14	best-precision vs . gmean	0.000823	0.045261
15	promethee-precision vs . gmean	0.000823	0.045261
16	best-precision vs . bac	0.003424	0.188329
17	promethee-precision vs . bac	0.003424	0.188329
18	best-recall vs . recall	0.005088	0.279838
19	promethee-recall vs . recall	0.005088	0.279838
20	precision vs . recall	0.007451	0.409818
21	best-precision vs . auc	0.010756	0.591562
22	promethee-precision vs . auc	0.010756	0.591562
23	best-precision vs . f1	0.019208	1.056417
24	promethee-precision vs . f1	0.019208	1.056417
25	balanced vs . recall	0.022681	1.247475
26	best-recall vs . f1	0.054435	2.99394
27	promethee-recall vs . f1	0.054435	2.99394
28	recall vs . gmean	0.059898	3.294384
29	best-recall vs . auc	0.086474	4.75609
30	promethee-recall vs . auc	0.086474	4.75609
31	bac vs . recall	0.143352	7.884334
32	best-precision vs . recall	0.143352	7.884334
33	promethee-precision vs . recall	0.143352	7.884334
34	balanced vs . f1	0.161302	8.871621
35	best-recall vs . bac	0.180901	9.949533
36	promethee-recall vs . bac	0.180901	9.949533
37	best-precision vs . precision	0.225302	12.391586
38	promethee-precision vs . precision	0.225302	12.391586
39	balanced vs . auc	0.233401	12.837029
40	recall vs . auc	0.276986	15.234239
41	f1 vs . gmean	0.315624	17.359337
42	best-recall vs . gmean	0.357643	19.670347
43	promethee-recall vs . gmean	0.357643	19.670347
44	recall vs . f1	0.379914	20.89529
45	balanced vs . bac	0.41488	22.818402
46	auc vs . gmean	0.426945	23.481986
47	bac vs . f1	0.558299	30.706465
48	best-recall vs . balanced	0.601217	33.066933
49	balanced vs . promethee-recall	0.601217	33.066933
50	bac vs . gmean	0.675859	37.172226
51	balanced vs . gmean	0.691209	38.016493
52	bac vs . auc	0.706687	38.867799
53	f1 vs . auc	0.834401	45.892055
54	best-precision vs . promethee-precision	1	55
55	best-recall vs . promethee-recall	1	55

Table 4: Adjusted p-values