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: 72.463287.

P-value computed by Friedman Test: 6.256795082038025E-11.

Algorithm	Ranking
best-precision	7.2115
best-recall	4.5962
balanced	3.1538
promethee-precision	7.2115
promethee-recall	4.5962
bac	5.9423
precision	8.2885
recall	8.8654
f1	4.6346
auc	5.4038
gmean	6.0962

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	balanced vs. recall	6.209097	0
54	balanced vs. precision	5.581915	0
53	best-recall vs. recall	4.641143	0.000003
52	promethee-recall vs. recall	4.641143	0.000003
51	recall vs. f1	4.599331	0.000004
50	best-precision vs. balanced	4.411177	0.00001
49	balanced vs. promethee-precision	4.411177	0.00001
48	best-recall vs. precision	4.013962	0.00006
47	promethee-recall vs. precision	4.013962	0.00006
46	precision vs. f1	3.97215	0.000071
45	recall vs. auc	3.763089	0.000168
44	balanced vs. gmean	3.198626	0.001381
43	bac vs. recall	3.17772	0.001484
42	precision vs. auc	3.135908	0.001713
41	balanced vs. bac	3.031377	0.002434
40	recall vs. gmean	3.010471	0.002608
39	best-precision vs. best-recall	2.843223	0.004466
38	best-precision vs. promethee-recall	2.843223	0.004466
37	best-recall vs. promethee-precision	2.843223	0.004466
36	promethee-precision vs. promethee-recall	2.843223	0.004466
35	best-precision vs. f1	2.801411	0.005088
34	promethee-precision vs. f1	2.801411	0.005088
33	bac vs. precision	2.550538	0.010756
32	balanced vs. auc	2.446008	0.014445
31	precision vs. gmean	2.38329	0.017159
30	best-precision vs. auc	1.965169	0.049395
29	promethee-precision vs. auc	1.965169	0.049395
28	best-precision vs. recall	1.79792	0.07219
27	promethee-precision vs. recall	1.79792	0.07219
26	best-recall vs. gmean	1.630672	0.10296
25	promethee-recall vs. gmean	1.630672	0.10296
24	balanced vs. f1	1.609766	0.107449
23	f1 vs. gmean	1.58886	0.112092
22	best-recall vs. balanced	1.567954	0.116892
21	balanced vs. promethee-recall	1.567954	0.116892
20	best-recall vs. bac	1.463424	0.143352
19	promethee-recall vs. bac	1.463424	0.143352
18	bac vs. f1	1.421611	0.155139
17	best-precision vs. bac	1.379799	0.167648
16	promethee-precision vs. bac	1.379799	0.167648
15	best-precision vs. gmean	1.212551	0.225302
14	promethee-precision vs. gmean	1.212551	0.225302
13	best-precision vs. precision	1.170739	0.241704
12	promethee-precision vs. precision	1.170739	0.241704
11	best-recall vs. auc	0.878054	0.379914
10	promethee-recall vs. auc	0.878054	0.379914
9	f1 vs. auc	0.836242	0.403019
8	auc vs. gmean	0.752618	0.45168
7	precision vs. recall	0.627182	0.53054
6	bac vs. auc	0.585369	0.558299
5	bac vs. gmean	0.167248	0.867175
4	best-recall vs. f1	0.041812	0.966648
3	promethee-recall vs. f1	0.041812	0.966648
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	balanced vs. recall	6.209097	0
54	balanced vs. precision	5.581915	0
53	best-recall vs. recall	4.641143	0.000003
52	promethee-recall vs. recall	4.641143	0.000003
51	recall vs. f1	4.599331	0.000004
50	best-precision vs. balanced	4.411177	0.00001
49	balanced vs. promethee-precision	4.411177	0.00001
48	best-recall vs. precision	4.013962	0.00006
47	promethee-recall vs. precision	4.013962	0.00006
46	precision vs. f1	3.97215	0.000071
45	recall vs. auc	3.763089	0.000168
44	balanced vs. gmean	3.198626	0.001381
43	bac vs. recall	3.17772	0.001484
42	precision vs. auc	3.135908	0.001713
41	balanced vs. bac	3.031377	0.002434
40	recall vs. gmean	3.010471	0.002608
39	best-precision vs. best-recall	2.843223	0.004466
38	best-precision vs. promethee-recall	2.843223	0.004466
37	best-recall vs. promethee-precision	2.843223	0.004466
36	promethee-precision vs. promethee-recall	2.843223	0.004466
35	best-precision vs. f1	2.801411	0.005088
34	promethee-precision vs. f1	2.801411	0.005088
33	bac vs. precision	2.550538	0.010756
32	balanced vs. auc	2.446008	0.014445
31	precision vs. gmean	2.38329	0.017159
30	best-precision vs. auc	1.965169	0.049395
29	promethee-precision vs. auc	1.965169	0.049395
28	best-precision vs. recall	1.79792	0.07219
27	promethee-precision vs. recall	1.79792	0.07219
26	best-recall vs. gmean	1.630672	0.10296
25	promethee-recall vs. gmean	1.630672	0.10296
24	balanced vs. f1	1.609766	0.107449
23	f1 vs. gmean	1.58886	0.112092
22	best-recall vs. balanced	1.567954	0.116892
21	balanced vs. promethee-recall	1.567954	0.116892
20	best-recall vs. bac	1.463424	0.143352
19	promethee-recall vs. bac	1.463424	0.143352
18	bac vs. f1	1.421611	0.155139
17	best-precision vs. bac	1.379799	0.167648
16	promethee-precision vs. bac	1.379799	0.167648
15	best-precision vs. gmean	1.212551	0.225302
14	promethee-precision vs. gmean	1.212551	0.225302
13	best-precision vs. precision	1.170739	0.241704
12	promethee-precision vs. precision	1.170739	0.241704
11	best-recall vs. auc	0.878054	0.379914
10	promethee-recall vs. auc	0.878054	0.379914
9	f1 vs. auc	0.836242	0.403019
8	auc vs. gmean	0.752618	0.45168
7	precision vs. recall	0.627182	0.53054
6	bac vs. auc	0.585369	0.558299
5	bac vs. gmean	0.167248	0.867175
4	best-recall vs. f1	0.041812	0.966648
3	promethee-recall vs. f1	0.041812	0.966648
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	balanced vs . recall	0	0
2	balanced vs . precision	0	0.000001
3	best-recall vs . recall	0.000003	0.000191
4	promethee-recall vs . recall	0.000003	0.000191
5	recall vs . f1	0.000004	0.000233
6	best-precision vs . balanced	0.00001	0.000565
7	balanced vs . promethee-precision	0.00001	0.000565
8	best-recall vs . precision	0.00006	0.003284
9	promethee-recall vs . precision	0.00006	0.003284
10	precision vs . f1	0.000071	0.003917
11	recall vs . auc	0.000168	0.009231
12	balanced vs . gmean	0.001381	0.075946
13	bac vs . recall	0.001484	0.081641
14	precision vs . auc	0.001713	0.094228
15	balanced vs . bac	0.002434	0.133892
16	recall vs . gmean	0.002608	0.143463
17	best-precision vs . best-recall	0.004466	0.245629
18	best-precision vs . promethee-recall	0.004466	0.245629
19	best-recall vs . promethee-precision	0.004466	0.245629
20	promethee-precision vs . promethee-recall	0.004466	0.245629
21	best-precision vs . f1	0.005088	0.279838
22	promethee-precision vs . f1	0.005088	0.279838
23	bac vs . precision	0.010756	0.591562
24	balanced vs . auc	0.014445	0.794463
25	precision vs . gmean	0.017159	0.943728
26	best-precision vs . auc	0.049395	2.716709
27	promethee-precision vs . auc	0.049395	2.716709
28	best-precision vs . recall	0.07219	3.97043
29	promethee-precision vs . recall	0.07219	3.97043
30	best-recall vs . gmean	0.10296	5.662776
31	promethee-recall vs . gmean	0.10296	5.662776
32	balanced vs . f1	0.107449	5.909694
33	f1 vs . gmean	0.112092	6.165062
34	best-recall vs . balanced	0.116892	6.429056
35	balanced vs . promethee-recall	0.116892	6.429056
36	best-recall vs . bac	0.143352	7.884334
37	promethee-recall vs . bac	0.143352	7.884334
38	bac vs . f1	0.155139	8.53265
39	best-precision vs . bac	0.167648	9.220664
40	promethee-precision vs . bac	0.167648	9.220664
41	best-precision vs . gmean	0.225302	12.391586
42	promethee-precision vs . gmean	0.225302	12.391586
43	best-precision vs . precision	0.241704	13.293708
44	promethee-precision vs . precision	0.241704	13.293708
45	best-recall vs . auc	0.379914	20.89529
46	promethee-recall vs . auc	0.379914	20.89529
47	f1 vs . auc	0.403019	22.166032
48	auc vs . gmean	0.45168	24.842379
49	precision vs . recall	0.53054	29.179714
50	bac vs . auc	0.558299	30.706465
51	bac vs . gmean	0.867175	47.694603
52	best-recall vs . f1	0.966648	53.165667
53	promethee-recall vs . f1	0.966648	53.165667
54	best-precision vs . promethee-precision	1	55
55	best-recall vs . promethee-recall	1	55

Table 4: Adjusted p-values