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: 153.56993.

P-value computed by Friedman Test: 1.15912723863687E-10.

Algorithm	Ranking
best-precision	8.8462
best-recall	3.4615
balanced	5.2885
promethee-precision	8.8462
promethee-recall	3.4615
bac	6.0769
precision	9.8077
recall	1.5192
f1	6.4231
auc	6.6731
gmean	5.5962

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	9.010508	0
54	best-precision vs. recall	7.965205	0
53	promethee-precision vs. recall	7.965205	0
52	best-recall vs. precision	6.898997	0
51	promethee-recall vs. precision	6.898997	0
50	best-precision vs. best-recall	5.853694	0
49	best-precision vs. promethee-recall	5.853694	0
48	best-recall vs. promethee-precision	5.853694	0
47	promethee-precision vs. promethee-recall	5.853694	0
46	recall vs. auc	5.602821	0
45	recall vs. f1	5.331043	0
44	bac vs. recall	4.954734	0.000001
43	balanced vs. precision	4.912922	0.000001
42	precision vs. gmean	4.578425	0.000005
41	recall vs. gmean	4.432083	0.000009
40	balanced vs. recall	4.097586	0.000042
39	bac vs. precision	4.055774	0.00005
38	best-precision vs. balanced	3.867619	0.00011
37	balanced vs. promethee-precision	3.867619	0.00011
36	precision vs. f1	3.679465	0.000234
35	best-precision vs. gmean	3.533122	0.000411
34	promethee-precision vs. gmean	3.533122	0.000411
33	best-recall vs. auc	3.49131	0.000481
32	promethee-recall vs. auc	3.49131	0.000481
31	precision vs. auc	3.407686	0.000655
30	best-recall vs. f1	3.219532	0.001284
29	promethee-recall vs. f1	3.219532	0.001284
28	best-precision vs. bac	3.010471	0.002608
27	promethee-precision vs. bac	3.010471	0.002608
26	best-recall vs. bac	2.843223	0.004466
25	promethee-recall vs. bac	2.843223	0.004466
24	best-precision vs. f1	2.634162	0.008435
23	promethee-precision vs. f1	2.634162	0.008435
22	best-precision vs. auc	2.362384	0.018158
21	promethee-precision vs. auc	2.362384	0.018158
20	best-recall vs. gmean	2.320572	0.02031
19	promethee-recall vs. gmean	2.320572	0.02031
18	best-recall vs. recall	2.111511	0.034728
17	promethee-recall vs. recall	2.111511	0.034728
16	best-recall vs. balanced	1.986075	0.047025
15	balanced vs. promethee-recall	1.986075	0.047025
14	balanced vs. auc	1.505236	0.132264
13	balanced vs. f1	1.233457	0.217405
12	auc vs. gmean	1.170739	0.241704
11	best-precision vs. precision	1.045303	0.295883
10	promethee-precision vs. precision	1.045303	0.295883
9	f1 vs. gmean	0.89896	0.368674
8	balanced vs. bac	0.857148	0.391363
7	bac vs. auc	0.648088	0.516928
6	bac vs. gmean	0.522651	0.601217
5	bac vs. f1	0.376309	0.706687
4	balanced vs. gmean	0.334497	0.738005
3	f1 vs. auc	0.271779	0.785792
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	9.010508	0
54	best-precision vs. recall	7.965205	0
53	promethee-precision vs. recall	7.965205	0
52	best-recall vs. precision	6.898997	0
51	promethee-recall vs. precision	6.898997	0
50	best-precision vs. best-recall	5.853694	0
49	best-precision vs. promethee-recall	5.853694	0
48	best-recall vs. promethee-precision	5.853694	0
47	promethee-precision vs. promethee-recall	5.853694	0
46	recall vs. auc	5.602821	0
45	recall vs. f1	5.331043	0
44	bac vs. recall	4.954734	0.000001
43	balanced vs. precision	4.912922	0.000001
42	precision vs. gmean	4.578425	0.000005
41	recall vs. gmean	4.432083	0.000009
40	balanced vs. recall	4.097586	0.000042
39	bac vs. precision	4.055774	0.00005
38	best-precision vs. balanced	3.867619	0.00011
37	balanced vs. promethee-precision	3.867619	0.00011
36	precision vs. f1	3.679465	0.000234
35	best-precision vs. gmean	3.533122	0.000411
34	promethee-precision vs. gmean	3.533122	0.000411
33	best-recall vs. auc	3.49131	0.000481
32	promethee-recall vs. auc	3.49131	0.000481
31	precision vs. auc	3.407686	0.000655
30	best-recall vs. f1	3.219532	0.001284
29	promethee-recall vs. f1	3.219532	0.001284
28	best-precision vs. bac	3.010471	0.002608
27	promethee-precision vs. bac	3.010471	0.002608
26	best-recall vs. bac	2.843223	0.004466
25	promethee-recall vs. bac	2.843223	0.004466
24	best-precision vs. f1	2.634162	0.008435
23	promethee-precision vs. f1	2.634162	0.008435
22	best-precision vs. auc	2.362384	0.018158
21	promethee-precision vs. auc	2.362384	0.018158
20	best-recall vs. gmean	2.320572	0.02031
19	promethee-recall vs. gmean	2.320572	0.02031
18	best-recall vs. recall	2.111511	0.034728
17	promethee-recall vs. recall	2.111511	0.034728
16	best-recall vs. balanced	1.986075	0.047025
15	balanced vs. promethee-recall	1.986075	0.047025
14	balanced vs. auc	1.505236	0.132264
13	balanced vs. f1	1.233457	0.217405
12	auc vs. gmean	1.170739	0.241704
11	best-precision vs. precision	1.045303	0.295883
10	promethee-precision vs. precision	1.045303	0.295883
9	f1 vs. gmean	0.89896	0.368674
8	balanced vs. bac	0.857148	0.391363
7	bac vs. auc	0.648088	0.516928
6	bac vs. gmean	0.522651	0.601217
5	bac vs. f1	0.376309	0.706687
4	balanced vs. gmean	0.334497	0.738005
3	f1 vs. auc	0.271779	0.785792
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	best-recall vs . precision	0	0
5	promethee-recall vs . precision	0	0
6	best-precision vs . best-recall	0	0
7	best-precision vs . promethee-recall	0	0
8	best-recall vs . promethee-precision	0	0
9	promethee-precision vs . promethee-recall	0	0
10	recall vs . auc	0	0.000001
11	recall vs . f1	0	0.000005
12	bac vs . recall	0.000001	0.00004
13	balanced vs . precision	0.000001	0.000049
14	precision vs . gmean	0.000005	0.000258
15	recall vs . gmean	0.000009	0.000513
16	balanced vs . recall	0.000042	0.002296
17	bac vs . precision	0.00005	0.002748
18	best-precision vs . balanced	0.00011	0.006045
19	balanced vs . promethee-precision	0.00011	0.006045
20	precision vs . f1	0.000234	0.012855
21	best-precision vs . gmean	0.000411	0.022588
22	promethee-precision vs . gmean	0.000411	0.022588
23	best-recall vs . auc	0.000481	0.026436
24	promethee-recall vs . auc	0.000481	0.026436
25	precision vs . auc	0.000655	0.036034
26	best-recall vs . f1	0.001284	0.07062
27	promethee-recall vs . f1	0.001284	0.07062
28	best-precision vs . bac	0.002608	0.143463
29	promethee-precision vs . bac	0.002608	0.143463
30	best-recall vs . bac	0.004466	0.245629
31	promethee-recall vs . bac	0.004466	0.245629
32	best-precision vs . f1	0.008435	0.463898
33	promethee-precision vs . f1	0.008435	0.463898
34	best-precision vs . auc	0.018158	0.998681
35	promethee-precision vs . auc	0.018158	0.998681
36	best-recall vs . gmean	0.02031	1.117049
37	promethee-recall vs . gmean	0.02031	1.117049
38	best-recall vs . recall	0.034728	1.910062
39	promethee-recall vs . recall	0.034728	1.910062
40	best-recall vs . balanced	0.047025	2.586376
41	balanced vs . promethee-recall	0.047025	2.586376
42	balanced vs . auc	0.132264	7.274493
43	balanced vs . f1	0.217405	11.957293
44	auc vs . gmean	0.241704	13.293708
45	best-precision vs . precision	0.295883	16.273575
46	promethee-precision vs . precision	0.295883	16.273575
47	f1 vs . gmean	0.368674	20.277063
48	balanced vs . bac	0.391363	21.524968
49	bac vs . auc	0.516928	28.431058
50	bac vs . gmean	0.601217	33.066933
51	bac vs . f1	0.706687	38.867799
52	balanced vs . gmean	0.738005	40.590259
53	f1 vs . auc	0.785792	43.218572
54	best-precision vs . promethee-precision	1	55
55	best-recall vs . promethee-recall	1	55

Table 4: Adjusted p-values