

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: 60.407343.

P-value computed by Friedman Test: 3.0664201178254302E-9.

Algorithm	Ranking
best-precision	7.9038
best-recall	4.3269
balanced	4.4615
promethee-precision	7.9038
promethee-recall	4.3269
bac	5.3654
precision	9.0577
recall	5.9808
f1	5.5962
auc	5.5577
gmean	5.5192

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	best-recall vs. precision	5.142888	0
54	promethee-recall vs. precision	5.142888	0
53	balanced vs. precision	4.996546	0.000001
52	bac vs. precision	4.013962	0.00006
51	best-precision vs. best-recall	3.888525	0.000101
50	best-precision vs. promethee-recall	3.888525	0.000101
49	best-recall vs. promethee-precision	3.888525	0.000101
48	promethee-precision vs. promethee-recall	3.888525	0.000101
47	precision vs. gmean	3.846713	0.00012
46	precision vs. auc	3.804901	0.000142
45	precision vs. f1	3.763089	0.000168
44	best-precision vs. balanced	3.742183	0.000182
43	balanced vs. promethee-precision	3.742183	0.000182
42	precision vs. recall	3.344968	0.000823
41	best-precision vs. bac	2.759599	0.005787
40	promethee-precision vs. bac	2.759599	0.005787
39	best-precision vs. gmean	2.59235	0.009532
38	promethee-precision vs. gmean	2.59235	0.009532
37	best-precision vs. auc	2.550538	0.010756
36	promethee-precision vs. auc	2.550538	0.010756
35	best-precision vs. f1	2.508726	0.012117
34	promethee-precision vs. f1	2.508726	0.012117
33	best-precision vs. recall	2.090605	0.036563
32	promethee-precision vs. recall	2.090605	0.036563
31	best-recall vs. recall	1.79792	0.07219
30	promethee-recall vs. recall	1.79792	0.07219
29	balanced vs. recall	1.651578	0.098621
28	best-recall vs. f1	1.379799	0.167648
27	promethee-recall vs. f1	1.379799	0.167648
26	best-recall vs. auc	1.337987	0.180901
25	promethee-recall vs. auc	1.337987	0.180901
24	best-recall vs. gmean	1.296175	0.194915
23	promethee-recall vs. gmean	1.296175	0.194915
22	best-precision vs. precision	1.254363	0.20971
21	promethee-precision vs. precision	1.254363	0.20971
20	balanced vs. f1	1.233457	0.217405
19	balanced vs. auc	1.191645	0.233401
18	balanced vs. gmean	1.149833	0.250213
17	best-recall vs. bac	1.128927	0.258929
16	promethee-recall vs. bac	1.128927	0.258929
15	balanced vs. bac	0.982584	0.325812
14	bac vs. recall	0.668994	0.5035
13	recall vs. gmean	0.501745	0.615847
12	recall vs. auc	0.459933	0.645564
11	recall vs. f1	0.418121	0.675859
10	bac vs. f1	0.250873	0.801913
9	bac vs. auc	0.209061	0.834401
8	bac vs. gmean	0.167248	0.867175
7	best-recall vs. balanced	0.146342	0.883651
6	balanced vs. promethee-recall	0.146342	0.883651
5	f1 vs. gmean	0.083624	0.933355
4	auc vs. gmean	0.041812	0.966648
3	f1 vs. auc	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	best-recall vs. precision	5.142888	0
54	promethee-recall vs. precision	5.142888	0
53	balanced vs. precision	4.996546	0.000001
52	bac vs. precision	4.013962	0.00006
51	best-precision vs. best-recall	3.888525	0.000101
50	best-precision vs. promethee-recall	3.888525	0.000101
49	best-recall vs. promethee-precision	3.888525	0.000101
48	promethee-precision vs. promethee-recall	3.888525	0.000101
47	precision vs. gmean	3.846713	0.00012
46	precision vs. auc	3.804901	0.000142
45	precision vs. f1	3.763089	0.000168
44	best-precision vs. balanced	3.742183	0.000182
43	balanced vs. promethee-precision	3.742183	0.000182
42	precision vs. recall	3.344968	0.000823
41	best-precision vs. bac	2.759599	0.005787
40	promethee-precision vs. bac	2.759599	0.005787
39	best-precision vs. gmean	2.59235	0.009532
38	promethee-precision vs. gmean	2.59235	0.009532
37	best-precision vs. auc	2.550538	0.010756
36	promethee-precision vs. auc	2.550538	0.010756
35	best-precision vs. f1	2.508726	0.012117
34	promethee-precision vs. f1	2.508726	0.012117
33	best-precision vs. recall	2.090605	0.036563
32	promethee-precision vs. recall	2.090605	0.036563
31	best-recall vs. recall	1.79792	0.07219
30	promethee-recall vs. recall	1.79792	0.07219
29	balanced vs. recall	1.651578	0.098621
28	best-recall vs. f1	1.379799	0.167648
27	promethee-recall vs. f1	1.379799	0.167648
26	best-recall vs. auc	1.337987	0.180901
25	promethee-recall vs. auc	1.337987	0.180901
24	best-recall vs. gmean	1.296175	0.194915
23	promethee-recall vs. gmean	1.296175	0.194915
22	best-precision vs. precision	1.254363	0.20971
21	promethee-precision vs. precision	1.254363	0.20971
20	balanced vs. f1	1.233457	0.217405
19	balanced vs. auc	1.191645	0.233401
18	balanced vs. gmean	1.149833	0.250213
17	best-recall vs. bac	1.128927	0.258929
16	promethee-recall vs. bac	1.128927	0.258929
15	balanced vs. bac	0.982584	0.325812
14	bac vs. recall	0.668994	0.5035
13	recall vs. gmean	0.501745	0.615847
12	recall vs. auc	0.459933	0.645564
11	recall vs. f1	0.418121	0.675859
10	bac vs. f1	0.250873	0.801913
9	bac vs. auc	0.209061	0.834401
8	bac vs. gmean	0.167248	0.867175
7	best-recall vs. balanced	0.146342	0.883651
6	balanced vs. promethee-recall	0.146342	0.883651
5	f1 vs. gmean	0.083624	0.933355
4	auc vs. gmean	0.041812	0.966648
3	f1 vs. auc	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	best-recall vs . precision	0	0.000015
2	promethee-recall vs . precision	0	0.000015
3	balanced vs . precision	0.000001	0.000032
4	bac vs . precision	0.000006	0.003284
5	best-precision vs . best-recall	0.000101	0.005547
6	best-precision vs . promethee-recall	0.000101	0.005547
7	best-recall vs . promethee-precision	0.000101	0.005547
8	promethee-precision vs . promethee-recall	0.000101	0.005547
9	precision vs . gmean	0.00012	0.006584
10	precision vs . auc	0.000142	0.007802
11	precision vs . f1	0.000168	0.009231
12	best-precision vs . balanced	0.000182	0.010034
13	balanced vs . promethee-precision	0.000182	0.010034
14	precision vs . recall	0.000823	0.045261
15	best-precision vs . bac	0.005787	0.318298
16	promethee-precision vs . bac	0.005787	0.318298
17	best-precision vs . gmean	0.009532	0.524275
18	promethee-precision vs . gmean	0.009532	0.524275
19	best-precision vs . auc	0.010756	0.591562
20	promethee-precision vs . auc	0.010756	0.591562
21	best-precision vs . f1	0.012117	0.666421
22	promethee-precision vs . f1	0.012117	0.666421
23	best-precision vs . recall	0.036563	2.010992
24	promethee-precision vs . recall	0.036563	2.010992
25	best-recall vs . recall	0.07219	3.97043
26	promethee-recall vs . recall	0.07219	3.97043
27	balanced vs . recall	0.098621	5.424134
28	best-recall vs . f1	0.167648	9.220664
29	promethee-recall vs . f1	0.167648	9.220664
30	best-recall vs . auc	0.180901	9.949533
31	promethee-recall vs . auc	0.180901	9.949533
32	best-recall vs . gmean	0.194915	10.720334
33	promethee-recall vs . gmean	0.194915	10.720334
34	best-precision vs . precision	0.20971	11.534055
35	promethee-precision vs . precision	0.20971	11.534055
36	balanced vs . f1	0.217405	11.957293
37	balanced vs . auc	0.233401	12.837029
38	balanced vs . gmean	0.250213	13.761702
39	best-recall vs . bac	0.258929	14.241081
40	promethee-recall vs . bac	0.258929	14.241081
41	balanced vs . bac	0.325812	17.919662
42	bac vs . recall	0.5035	27.692476
43	recall vs . gmean	0.615847	33.871572
44	recall vs . auc	0.645564	35.506033
45	recall vs . f1	0.675859	37.172226
46	bac vs . f1	0.801913	44.105193
47	bac vs . auc	0.834401	45.892055
48	bac vs . gmean	0.867175	47.694603
49	best-recall vs . balanced	0.883651	48.600812
50	balanced vs . promethee-recall	0.883651	48.600812
51	f1 vs . gmean	0.933355	51.334537
52	auc vs . gmean	0.966648	53.165667
53	f1 vs . auc	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