Output tables for the test of Multiple comparisons.

January 10, 2022

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 12 degrees of freedom: 65.212595.

P-value computed by Friedman Test: 2.5550600524226 75E-9.

Algorithm	Ranking
best-precision	9.4423
best-recall	5.1731
balanced	5.1538
promethee-precision	9.4423
promethee-recall	5.1731
bac	6.0577
precision	10.5577
recall	7.0577
f1	6.2885
auc	6.3269
gmean	6.2885
AdaBoost	7.7308
Bagging	6.3077

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 ≤ 0.000641.

i	algorithms	$z = (R_0 - R_i)/SE$	p
78	balanced vs. precision	5.002989	0.000001
77	best-recall vs. precision	4.985185	0.000001
76	promethee-recall vs. precision	4.985185	0.000001
75	bac vs. precision	4.16619	0.000031
74	best-precision vs. balanced	3.970344	0.000072
73	balanced vs. promethee-precision	3.970344	0.000072
72	precision vs. gmean	3.95254	0.000077
71	precision vs. f1	3.95254	0.000077
70	best-precision vs. best-recall	3.95254	0.000077
69	best-precision vs. promethee-recall	3.95254	0.000077
68	best-recall vs. promethee-precision	3.95254	0.000077
67	promethee-precision vs. promethee-recall	3.95254	0.000077
66	precision vs. Bagging	3.934735	0.000083
65	precision vs. auc	3.916931	0.00009
64	precision vs. recall	3.24037	0.001194
63	best-precision vs. bac	3.133545	0.001727
62	promethee-precision vs. bac	3.133545	0.001727
61	best-precision vs. gmean	2.919894	0.003502
60	promethee-precision vs. gmean	2.919894	0.003502
59	best-precision vs. f1	2.919894	0.003502
58	promethee-precision vs. f1	2.919894	0.003502
57	best-precision vs. Bagging	2.90209	0.003707
56	promethee-precision vs. Bagging	2.90209	0.003707
55	best-precision vs. auc	2.884286	0.003923
54	promethee-precision vs. auc	2.884286	0.003923
53	precision vs. AdaBoost	2.617222	0.008865
52	balanced vs. AdaBoost	2.385767	0.017044
51	best-recall vs. AdaBoost	2.367963	0.017886
50	promethee-recall vs. AdaBoost	2.367963	0.017886
49	best-precision vs. recall	2.207725	0.027263
48	promethee-precision vs. recall	2.207725	0.027263
47	balanced vs. recall	1.762619	0.077965
46	best-recall vs. recall	1.744815	0.081017
45	promethee-recall vs. recall	1.744815	0.081017
44	best-precision vs. AdaBoost	1.584577	0.113063
43	promethee-precision vs. AdaBoost	1.584577	0.113063
42	bac vs. AdaBoost	1.548968	0.121389
41	gmean vs. AdaBoost	1.335317	0.181772
40	f1 vs. AdaBoost	1.335317	0.181772
39	AdaBoost vs. Bagging	1.317513	0.187667
38	auc vs. AdaBoost	1.299709	0.193701
37	balanced vs. auc	1.086058	0.277453
36	best-recall vs. auc	1.068254	0.285406
35	promethee-recall vs. auc	1.068254	0.285406
34	balanced vs. Bagging	1.068254	0.285406
33	best-recall vs. Bagging	1.05045	0.293511
32	promethee-recall vs. Bagging	1.05045	0.293511
31	balanced vs. f1	1.05045	0.293511
30	balanced vs. gmean	1.05045	0.293511
29	best-precision vs. precision	1.032645	0.30177
28	promethee-precision vs. precision	1.032645	0.30177
27	best-recall vs. f1	1.032645	0.30177
26	promethee-recall vs. f1	1.032645	0.30177
25	best-recall vs. gmean	1.032645	0.30177
24	promethee-recall vs. gmean	1.032645	0.30177
23	bac vs. recall	0.92582	0.354539
22	balanced vs. bac	0.836799	0.402706
21	best-recall vs. bac	0.818995	0.412789
20	promethee-recall vs. bac	0.818995	0.412789
19	recall vs. gmean	0.712169	0.47636
18	recall vs. f1	0.712169	0.47636
17	recall vs. Bagging	0.694365	0.487453
16	recall vs. auc	0.676561	0.498685
15	recall vs. AdaBoost	0.623148	0.533187
14	bac vs. auc	0.249259	0.80316
13	bac vs. Bagging	0.231455	0.816961
12	bac vs. f1	0.213651	0.830819
11	bac vs. gmean	0.213651	0.830819
10	auc vs. gmean	0.035608	0.971595
9	f1 vs. auc	0.035608	0.971595
8	gmean vs. Bagging	0.017804	0.985795
7	f1 vs. Bagging	0.017804	0.985795
6	auc vs. Bagging	0.017804	0.985795
5	best-recall vs. balanced	0.017804	0.985795
4	balanced vs. promethee-recall	0.017804	0.985795
3	f1 vs. gmean	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 ≤ 0.001282.

i	algorithms	$z = (R_0 - R_i)/SE$	p
78	balanced vs. precision	5.002989	0.000001
77	best-recall vs. precision	4.985185	0.000001
76	promethee-recall vs. precision	4.985185	0.000001
75	bac vs. precision	4.16619	0.000031
74	best-precision vs. balanced	3.970344	0.000072
73	balanced vs. promethee-precision	3.970344	0.000072
72	precision vs. gmean	3.95254	0.000077
71	precision vs. f1	3.95254	0.000077
70	best-precision vs. best-recall	3.95254	0.000077
69	best-precision vs. promethee-recall	3.95254	0.000077
68	best-recall vs. promethee-precision	3.95254	0.000077
67	promethee-precision vs. promethee-recall	3.95254	0.000077
66	precision vs. Bagging	3.934735	0.000083
65	precision vs. auc	3.916931	0.00009
64	precision vs. recall	3.24037	0.001194
63	best-precision vs. bac	3.133545	0.001727
62	promethee-precision vs. bac	3.133545	0.001727
61	best-precision vs. gmean	2.919894	0.003502
60	promethee-precision vs. gmean	2.919894	0.003502
59	best-precision vs. f1	2.919894	0.003502
58	promethee-precision vs. f1	2.919894	0.003502
57	best-precision vs. Bagging	2.90209	0.003707
56	promethee-precision vs. Bagging	2.90209	0.003707
55	best-precision vs. auc	2.884286	0.003923
54	promethee-precision vs. auc	2.884286	0.003923
53	precision vs. AdaBoost	2.617222	0.008865
52	balanced vs. AdaBoost	2.385767	0.017044
51	best-recall vs. AdaBoost	2.367963	0.017886
50	promethee-recall vs. AdaBoost	2.367963	0.017886
49	best-precision vs. recall	2.207725	0.027263
48	promethee-precision vs. recall	2.207725	0.027263
47	balanced vs. recall	1.762619	0.077965
46	best-recall vs. recall	1.744815	0.081017
45	promethee-recall vs. recall	1.744815	0.081017
44	best-precision vs. AdaBoost	1.584577	0.113063
43	promethee-precision vs. AdaBoost	1.584577	0.113063
42	bac vs. AdaBoost	1.548968	0.121389
41	gmean vs. AdaBoost	1.335317	0.181772
40	f1 vs. AdaBoost	1.335317	0.181772
39	AdaBoost vs. Bagging	1.317513	0.187667
38	auc vs. AdaBoost	1.299709	0.193701
37	balanced vs. auc	1.086058	0.277453
36	best-recall vs. auc	1.068254	0.285406
35	promethee-recall vs. auc	1.068254	0.285406
34	balanced vs. Bagging	1.068254	0.285406
33	best-recall vs. Bagging	1.05045	0.293511
32	promethee-recall vs. Bagging	1.05045	0.293511
31	balanced vs. f1	1.05045	0.293511
30	balanced vs. gmean	1.05045	0.293511
29	best-precision vs. precision	1.032645	0.30177
28	promethee-precision vs. precision	1.032645	0.30177
27	best-recall vs. f1	1.032645	0.30177
26	promethee-recall vs. f1	1.032645	0.30177
25	best-recall vs. gmean	1.032645	0.30177
24	promethee-recall vs. gmean	1.032645	0.30177
23	bac vs. recall	0.92582	0.354539
22	balanced vs. bac	0.836799	0.402706
21	best-recall vs. bac	0.818995	0.412789
20	promethee-recall vs. bac	0.818995	0.412789
19	recall vs. gmean	0.712169	0.47636
18	recall vs. f1	0.712169	0.47636
17	recall vs. Bagging	0.694365	0.487453
16	recall vs. auc	0.676561	0.498685
15	recall vs. AdaBoost	0.623148	0.533187
14	bac vs. auc	0.249259	0.80316
13	bac vs. Bagging	0.231455	0.816961
12	bac vs. f1	0.213651	0.830819
11	bac vs. gmean	0.213651	0.830819
10	auc vs. gmean	0.035608	0.971595
9	f1 vs. auc	0.035608	0.971595
8	gmean vs. Bagging	0.017804	0.985795
7	f1 vs. Bagging	0.017804	0.985795
6	auc vs. Bagging	0.017804	0.985795
5	best-recall vs. balanced	0.017804	0.985795
4	balanced vs. promethee-recall	0.017804	0.985795
3	f1 vs. gmean	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	balanced vs . precision	0.000001	0.000044
2	best-recall vs . precision	0.000001	0.000048
3	promethee-recall vs . precision	0.000001	0.000048
4	bac vs . precision	0.000031	0.002416
5	best-precision vs . balanced	0.000072	0.005598
6	balanced vs . promethee-precision	0.000072	0.005598
7	precision vs . gmean	0.000077	0.006031
8	precision vs . f1	0.000077	0.006031
9	best-precision vs . best-recall	0.000077	0.006031
10	best-precision vs . promethee-recall	0.000077	0.006031
11	best-recall vs . promethee-precision	0.000077	0.006031
12	promethee-precision vs . promethee-recall	0.000077	0.006031
13	precision vs .Bagging	0.000083	0.006496
14	precision vs . auc	0.00009	0.006995
15	precision vs . recall	0.001194	0.093112
16	best-precision vs . bac	0.001727	0.134713
17	promethee-precision vs . bac	0.001727	0.134713
18	best-precision vs . gmean	0.003502	0.273117
19	promethee-precision vs . gmean	0.003502	0.273117
20	best-precision vs . f1	0.003502	0.273117
21	promethee-precision vs . f1	0.003502	0.273117
22	best-precision vs .Bagging	0.003707	0.289132
23	promethee-precision vs .Bagging	0.003707	0.289132
24	best-precision vs . auc	0.003923	0.305996
25	promethee-precision vs . auc	0.003923	0.305996
26	precision vs .AdaBoost	0.008865	0.691459
27	balanced vs .AdaBoost	0.017044	1.329396
28	best-recall vs .AdaBoost	0.017886	1.395134
29	promethee-recall vs .AdaBoost	0.017886	1.395134
30	best-precision vs . recall	0.027263	2.12655
31	promethee-precision vs . recall	0.027263	2.12655
32	balanced vs . recall	0.077965	6.081252
33	best-recall vs . recall	0.081017	6.319335
34	promethee-recall vs . recall	0.081017	6.319335
35	best-precision vs .AdaBoost	0.113063	8.818878
36	promethee-precision vs .AdaBoost	0.113063	8.818878
37	bac vs .AdaBoost	0.121389	9.46837
38	gmean vs .AdaBoost	0.181772	14.178253
39	f1 vs .AdaBoost	0.181772	14.178253
40	AdaBoost vs .Bagging	0.187667	14.637999
41	auc vs .AdaBoost	0.193701	15.108657
42	balanced vs . auc	0.277453	21.641353
43	best-recall vs . auc	0.285406	22.261666
44	promethee-recall vs . auc	0.285406	22.261666
45	balanced vs .Bagging	0.285406	22.261666
46	best-recall vs .Bagging	0.293511	22.893889
47	promethee-recall vs .Bagging	0.293511	22.893889
48	balanced vs . f1	0.293511	22.893889
49	balanced vs . gmean	0.293511	22.893889
50	best-precision vs . precision	0.30177	23.538047
51	promethee-precision vs . precision	0.30177	23.538047
52	best-recall vs . f1	0.30177	23.538047
53	promethee-recall vs . f1	0.30177	23.538047
54	best-recall vs . gmean	0.30177	23.538047
55	promethee-recall vs . gmean	0.30177	23.538047
56	bac vs . recall	0.354539	27.654079
57	balanced vs . bac	0.402706	31.411036
58	best-recall vs . bac	0.412789	32.197576
59	promethee-recall vs . bac	0.412789	32.197576
60	recall vs . gmean	0.47636	37.156075
61	recall vs . f1	0.47636	37.156075
62	recall vs .Bagging	0.487453	38.021357
63	recall vs . auc	0.498685	38.897401
64	recall vs .AdaBoost	0.533187	41.5886
65	bac vs . auc	0.80316	62.646499
66	bac vs .Bagging	0.816961	63.722983
67	bac vs . f1	0.830819	64.803912
68	bac vs . gmean	0.830819	64.803912
69	auc vs . gmean	0.971595	75.784376
70	f1 vs . auc	0.971595	75.784376
71	gmean vs .Bagging	0.985795	76.892012
72	f1 vs .Bagging	0.985795	76.892012
73	auc vs .Bagging	0.985795	76.892012
74	best-recall vs . balanced	0.985795	76.892012
75	balanced vs . promethee-recall	0.985795	76.892012
76	f1 vs . gmean	1	78
77	best-precision vs . promethee-precision	1	78
78	best-recall vs . promethee-recall	1	78

Table 4: Adjusted p-values