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: 74.057904.

P-value computed by Friedman Test: 1.1299439162115732E-10.

Algorithm	Ranking
best-precision	9.3654
best-recall	4.75
balanced	5.1154
promethee-precision	9.3654
promethee-recall	4.75
bac	5.9038
precision	10.4423
recall	7.7885
f1	6.3654
auc	6.3269
gmean	5.5192
AdaBoost	8.2692
Bagging	7.0385

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	best-recall vs. precision	5.270053	0
77	promethee-recall vs. precision	5.270053	0
76	balanced vs. precision	4.931772	0.000001
75	precision vs. gmean	4.557884	0.000005
74	best-precision vs. best-recall	4.273016	0.000019
73	best-precision vs. promethee-recall	4.273016	0.000019
72	best-recall vs. promethee-precision	4.273016	0.000019
71	promethee-precision vs. promethee-recall	4.273016	0.000019
70	bac vs. precision	4.201799	0.000026
69	best-precision vs. balanced	3.934735	0.000083
68	balanced vs. promethee-precision	3.934735	0.000083
67	precision vs. auc	3.810106	0.000139
66	precision vs. f1	3.774497	0.00016
65	best-precision vs. gmean	3.560847	0.00037
64	promethee-precision vs. gmean	3.560847	0.00037
63	best-recall vs. AdaBoost	3.258175	0.001121
62	promethee-recall vs. AdaBoost	3.258175	0.001121
61	best-precision vs. bac	3.204762	0.001352
60	promethee-precision vs. bac	3.204762	0.001352
59	precision vs. Bagging	3.151349	0.001625
58	balanced vs. AdaBoost	2.919894	0.003502
57	best-recall vs. recall	2.813069	0.004907
56	promethee-recall vs. recall	2.813069	0.004907
55	best-precision vs. auc	2.813069	0.004907
54	promethee-precision vs. auc	2.813069	0.004907
53	best-precision vs. f1	2.77746	0.005479
52	promethee-precision vs. f1	2.77746	0.005479
51	gmean vs. AdaBoost	2.546005	0.010896
50	balanced vs. recall	2.474788	0.013332
49	precision vs. recall	2.456984	0.014011
48	bac vs. AdaBoost	2.189921	0.02853
47	best-precision vs. Bagging	2.154312	0.031216
46	promethee-precision vs. Bagging	2.154312	0.031216
45	best-recall vs. Bagging	2.118704	0.034116
44	promethee-recall vs. Bagging	2.118704	0.034116
43	recall vs. gmean	2.100899	0.03565
42	precision vs. AdaBoost	2.011878	0.044233
41	auc vs. AdaBoost	1.798228	0.072141
40	balanced vs. Bagging	1.780423	0.075007
39	f1 vs. AdaBoost	1.762619	0.077965
38	bac vs. recall	1.744815	0.081017
37	best-recall vs. f1	1.495556	0.13477
36	promethee-recall vs. f1	1.495556	0.13477
35	best-recall vs. auc	1.459947	0.144305
34	promethee-recall vs. auc	1.459947	0.144305
33	best-precision vs. recall	1.459947	0.144305
32	promethee-precision vs. recall	1.459947	0.144305
31	gmean vs. Bagging	1.406534	0.159566
30	recall vs. auc	1.353122	0.176017
29	recall vs. f1	1.317513	0.187667
28	balanced vs. f1	1.157275	0.24716
27	AdaBoost vs. Bagging	1.139471	0.254507
26	balanced vs. auc	1.121667	0.262004
25	best-recall vs. bac	1.068254	0.285406
24	promethee-recall vs. bac	1.068254	0.285406
23	bac vs. Bagging	1.05045	0.293511
22	best-precision vs. AdaBoost	1.014841	0.310182
21	promethee-precision vs. AdaBoost	1.014841	0.310182
20	best-precision vs. precision	0.997037	0.318747
19	promethee-precision vs. precision	0.997037	0.318747
18	f1 vs. gmean	0.783386	0.4334
17	auc vs. gmean	0.747778	0.454594
16	balanced vs. bac	0.729974	0.465406
15	best-recall vs. gmean	0.712169	0.47636
14	promethee-recall vs. gmean	0.712169	0.47636
13	recall vs. Bagging	0.694365	0.487453
12	auc vs. Bagging	0.658757	0.510052
11	f1 vs. Bagging	0.623148	0.533187
10	recall vs. AdaBoost	0.445106	0.656243
9	bac vs. f1	0.427302	0.66916
8	bac vs. auc	0.391693	0.695285
7	balanced vs. gmean	0.373889	0.708487
6	bac vs. gmean	0.356085	0.721777
5	best-recall vs. balanced	0.33828	0.735152
4	balanced vs. promethee-recall	0.33828	0.735152
3	f1 vs. auc	0.035608	0.971595
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	best-recall vs. precision	5.270053	0
77	promethee-recall vs. precision	5.270053	0
76	balanced vs. precision	4.931772	0.000001
75	precision vs. gmean	4.557884	0.000005
74	best-precision vs. best-recall	4.273016	0.000019
73	best-precision vs. promethee-recall	4.273016	0.000019
72	best-recall vs. promethee-precision	4.273016	0.000019
71	promethee-precision vs. promethee-recall	4.273016	0.000019
70	bac vs. precision	4.201799	0.000026
69	best-precision vs. balanced	3.934735	0.000083
68	balanced vs. promethee-precision	3.934735	0.000083
67	precision vs. auc	3.810106	0.000139
66	precision vs. f1	3.774497	0.00016
65	best-precision vs. gmean	3.560847	0.00037
64	promethee-precision vs. gmean	3.560847	0.00037
63	best-recall vs. AdaBoost	3.258175	0.001121
62	promethee-recall vs. AdaBoost	3.258175	0.001121
61	best-precision vs. bac	3.204762	0.001352
60	promethee-precision vs. bac	3.204762	0.001352
59	precision vs. Bagging	3.151349	0.001625
58	balanced vs. AdaBoost	2.919894	0.003502
57	best-recall vs. recall	2.813069	0.004907
56	promethee-recall vs. recall	2.813069	0.004907
55	best-precision vs. auc	2.813069	0.004907
54	promethee-precision vs. auc	2.813069	0.004907
53	best-precision vs. f1	2.77746	0.005479
52	promethee-precision vs. f1	2.77746	0.005479
51	gmean vs. AdaBoost	2.546005	0.010896
50	balanced vs. recall	2.474788	0.013332
49	precision vs. recall	2.456984	0.014011
48	bac vs. AdaBoost	2.189921	0.02853
47	best-precision vs. Bagging	2.154312	0.031216
46	promethee-precision vs. Bagging	2.154312	0.031216
45	best-recall vs. Bagging	2.118704	0.034116
44	promethee-recall vs. Bagging	2.118704	0.034116
43	recall vs. gmean	2.100899	0.03565
42	precision vs. AdaBoost	2.011878	0.044233
41	auc vs. AdaBoost	1.798228	0.072141
40	balanced vs. Bagging	1.780423	0.075007
39	f1 vs. AdaBoost	1.762619	0.077965
38	bac vs. recall	1.744815	0.081017
37	best-recall vs. f1	1.495556	0.13477
36	promethee-recall vs. f1	1.495556	0.13477
35	best-recall vs. auc	1.459947	0.144305
34	promethee-recall vs. auc	1.459947	0.144305
33	best-precision vs. recall	1.459947	0.144305
32	promethee-precision vs. recall	1.459947	0.144305
31	gmean vs. Bagging	1.406534	0.159566
30	recall vs. auc	1.353122	0.176017
29	recall vs. f1	1.317513	0.187667
28	balanced vs. f1	1.157275	0.24716
27	AdaBoost vs. Bagging	1.139471	0.254507
26	balanced vs. auc	1.121667	0.262004
25	best-recall vs. bac	1.068254	0.285406
24	promethee-recall vs. bac	1.068254	0.285406
23	bac vs. Bagging	1.05045	0.293511
22	best-precision vs. AdaBoost	1.014841	0.310182
21	promethee-precision vs. AdaBoost	1.014841	0.310182
20	best-precision vs. precision	0.997037	0.318747
19	promethee-precision vs. precision	0.997037	0.318747
18	f1 vs. gmean	0.783386	0.4334
17	auc vs. gmean	0.747778	0.454594
16	balanced vs. bac	0.729974	0.465406
15	best-recall vs. gmean	0.712169	0.47636
14	promethee-recall vs. gmean	0.712169	0.47636
13	recall vs. Bagging	0.694365	0.487453
12	auc vs. Bagging	0.658757	0.510052
11	f1 vs. Bagging	0.623148	0.533187
10	recall vs. AdaBoost	0.445106	0.656243
9	bac vs. f1	0.427302	0.66916
8	bac vs. auc	0.391693	0.695285
7	balanced vs. gmean	0.373889	0.708487
6	bac vs. gmean	0.356085	0.721777
5	best-recall vs. balanced	0.33828	0.735152
4	balanced vs. promethee-recall	0.33828	0.735152
3	f1 vs. auc	0.035608	0.971595
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	0.000011
2	promethee-recall vs . precision	0	0.000011
3	balanced vs . precision	0.000001	0.000064
4	precision vs . gmean	0.000005	0.000403
5	best-precision vs . best-recall	0.000019	0.001504
6	best-precision vs . promethee-recall	0.000019	0.001504
7	best-recall vs . promethee-precision	0.000019	0.001504
8	promethee-precision vs . promethee-recall	0.000019	0.001504
9	bac vs . precision	0.000026	0.002065
10	best-precision vs . balanced	0.000083	0.006496
11	balanced vs . promethee-precision	0.000083	0.006496
12	precision vs . auc	0.000139	0.010835
13	precision vs . f1	0.00016	0.012506
14	best-precision vs . gmean	0.00037	0.028834
15	promethee-precision vs . gmean	0.00037	0.028834
16	best-recall vs .AdaBoost	0.001121	0.087462
17	promethee-recall vs .AdaBoost	0.001121	0.087462
18	best-precision vs . bac	0.001352	0.105436
19	promethee-precision vs . bac	0.001352	0.105436
20	precision vs .Bagging	0.001625	0.126764
21	balanced vs .AdaBoost	0.003502	0.273117
22	best-recall vs . recall	0.004907	0.382755
23	promethee-recall vs . recall	0.004907	0.382755
24	best-precision vs . auc	0.004907	0.382755
25	promethee-precision vs . auc	0.004907	0.382755
26	best-precision vs . f1	0.005479	0.427327
27	promethee-precision vs . f1	0.005479	0.427327
28	gmean vs .AdaBoost	0.010896	0.849916
29	balanced vs . recall	0.013332	1.039858
30	precision vs . recall	0.014011	1.092849
31	bac vs .AdaBoost	0.02853	2.22534
32	best-precision vs .Bagging	0.031216	2.434825
33	promethee-precision vs .Bagging	0.031216	2.434825
34	best-recall vs .Bagging	0.034116	2.66101
35	promethee-recall vs .Bagging	0.034116	2.66101
36	recall vs . gmean	0.03565	2.780684
37	precision vs .AdaBoost	0.044233	3.450156
38	auc vs .AdaBoost	0.072141	5.626995
39	balanced vs .Bagging	0.075007	5.850524
40	f1 vs .AdaBoost	0.077965	6.081252
41	bac vs . recall	0.081017	6.319335
42	best-recall vs . f1	0.13477	10.512022
43	promethee-recall vs . f1	0.13477	10.512022
44	best-recall vs . auc	0.144305	11.25576
45	promethee-recall vs . auc	0.144305	11.25576
46	best-precision vs . recall	0.144305	11.25576
47	promethee-precision vs . recall	0.144305	11.25576
48	gmean vs .Bagging	0.159566	12.446109
49	recall vs . auc	0.176017	13.729307
50	recall vs . f1	0.187667	14.637999
51	balanced vs . f1	0.24716	19.278478
52	AdaBoost vs .Bagging	0.254507	19.851531
53	balanced vs . auc	0.262004	20.436327
54	best-recall vs . bac	0.285406	22.261666
55	promethee-recall vs . bac	0.285406	22.261666
56	bac vs .Bagging	0.293511	22.893889
57	best-precision vs .AdaBoost	0.310182	24.194157
58	promethee-precision vs .AdaBoost	0.310182	24.194157
59	best-precision vs . precision	0.318747	24.86223
60	promethee-precision vs . precision	0.318747	24.86223
61	f1 vs . gmean	0.4334	33.805226
62	auc vs . gmean	0.454594	35.458348
63	balanced vs . bac	0.465406	36.301696
64	best-recall vs . gmean	0.47636	37.156075
65	promethee-recall vs . gmean	0.47636	37.156075
66	recall vs .Bagging	0.487453	38.021357
67	auc vs .Bagging	0.510052	39.784062
68	f1 vs .Bagging	0.533187	41.5886
69	recall vs .AdaBoost	0.656243	51.186976
70	bac vs . f1	0.66916	52.194455
71	bac vs . auc	0.695285	54.232228
72	balanced vs . gmean	0.708487	55.261985
73	bac vs . gmean	0.721777	56.298619
74	best-recall vs . balanced	0.735152	57.341846
75	balanced vs . promethee-recall	0.735152	57.341846
76	f1 vs . auc	0.971595	75.784376
77	best-precision vs . promethee-precision	1	78
78	best-recall vs . promethee-recall	1	78

Table 4: Adjusted p-values