

Results

May 8, 2024

1 Tables of Friedman, Bonferroni-Dunn, Holm, Hochberg and Hommel Tests

Table 1: Average Rankings of the algorithms

Algorithm	Ranking
LWRK	2.15
MS	3.36
SPT	1.35
WINQ	3.71
—Flexible-PPO—Complete-model-3	7.18
—Flexible-PPO—MachineCompressed-model-4	7.11
—Flexible-PPO—MachineCompressed-model-5	5.72
—Flexible-PPO—MachineCompressed-model-6	8.19
—Flexible-PPO—MachineCompressed-model-7	6.23

Friedman statistic considering reduction performance (distributed according to chi-square with 8 degrees of freedom: 1884.398222222203.
P-value computed by Friedman Test: 0.0.

Iman and Davenport statistic considering reduction performance (distributed according to F-distribution with 8 and 2392 degrees of freedom: 1092.77177218589.
P-value computed by Iman and Daveport Test: 0.0.

Table 2: Holm / Hochberg Table for $\alpha = 0.05$

i	algorithm	$z = (R_0 - R_i) / SE$	p	Holm/Hochberg/Hommel
8	—Flexible-PPO—MachineCompressed-model-6	30.55959569249716	4.216315290413759E-205	0.00625
7	—Flexible-PPO—Complete-model-3	26.04273837794752	1.6259111677652102E-149	0.0071428571428571435
6	—Flexible-PPO—MachineCompressed-model-4	25.729688861097568	5.441405399684134E-146	0.008333333333333333
5	—Flexible-PPO—MachineCompressed-model-7	21.824023460397857	1.3723757925740571E-105	0.01
4	—Flexible-PPO—MachineCompressed-model-5	19.513419883648094	8.44396628051309E-85	0.0125
3	WINQ	10.539333733948945	5.690057749259103E-26	0.016666666666666666
2	MS	8.988993269549105	2.495053098565265E-19	0.025
1	LWRK	3.5777087639996576	3.4661935113463444E-4	0.05

Bonferroni-Dunn's procedure rejects those hypotheses that have a p-value ≤ 0.00625 .
Hochberg's procedure rejects those hypotheses that have a p-value ≤ 0.05 .
Hommel's procedure rejects all hypotheses.
Bonferroni-Dunn's procedure rejects those hypotheses that have a p-value ≤ 0.0125 .
Hochberg's procedure rejects those hypotheses that have a p-value ≤ 0.1 .
Hommel's procedure rejects all hypotheses.
Nemenyi's procedure rejects those hypotheses that have a p-value $\leq 0.001388888888888889$.

Table 3: Holm / Hochberg Table for $\alpha = 0.10$

i	algorithm	$z = (R_0 - R_i)/SE$	p	Holm/Hochberg/Hommel
8	—Flexible-PPO—MachineCompressed-model-6	30.55959569249716	4.216315290413759E-205	0.0125
7	—Flexible-PPO—Complete-model-3	26.04273837794752	1.6259111677652102E-149	0.014285714285714287
6	—Flexible-PPO—MachineCompressed-model-4	25.729688861097568	5.441405399684134E-146	0.016666666666666666
5	—Flexible-PPO—MachineCompressed-model-7	21.824023460397857	1.3723757925740571E-105	0.02
4	—Flexible-PPO—MachineCompressed-model-5	19.513419883648094	8.44396628051309E-85	0.025
3	WINQ	10.539333733948945	5.690057749259103E-26	0.03333333333333333
2	MS	8.988993269549105	2.495053098565265E-19	0.05
1	LWRK	3.5777087639996576	3.4661935113463444E-4	0.1

Table 4: Adjusted p -values

i	algorithm	unadjusted p	p_{Bonf}	p_{Holm}	p_{Hoch}	p_{Hommel}
1	—Flexible-PPO—MachineCompressed-model-6	4.216315290413759E-205	3.373052232331007E-204	3.373052232331007E-204	3.373052232331007E-204	3.373052232331007E-204
2	—Flexible-PPO—Complete-model-3	1.6259111677652102E-149	1.3007289342121681E-148	1.1381378174356471E-148	1.1381378174356471E-148	1.1381378174356471E-148
3	—Flexible-PPO—MachineCompressed-model-4	5.441405399684134E-146	4.353124319747307E-145	3.26484323981048E-145	3.26484323981048E-145	3.26484323981048E-145
4	—Flexible-PPO—MachineCompressed-model-7	1.3723757925740571E-105	1.0979006340592457E-104	6.861878962870286E-105	6.861878962870286E-105	6.861878962870286E-105
5	—Flexible-PPO—MachineCompressed-model-5	8.44396628051309E-85	6.755173024410472E-84	3.377586512205236E-84	3.377586512205236E-84	3.377586512205236E-84
6	WINQ	5.690057749259103E-26	4.552046199407282E-25	1.7070173247777307E-25	1.7070173247777307E-25	1.7070173247777307E-25
7	MS	2.495053098565265E-19	1.996042478852212E-18	4.99010619713053E-19	4.99010619713053E-19	4.99010619713053E-19
8	LWRK	3.4661935113463444E-4	0.0027729548090770755	3.4661935113463444E-4	3.4661935113463444E-4	3.4661935113463444E-4

Table 5: Holm / Shaffer Table for $\alpha = 0.05$

i	algorithms	$z = (R_0 - R_i)/SE$	p	Holm	Shaffer
36	SPT vs. —Flexible-PPO—MachineCompressed-model-6	30.55959569249716	4.216315290413759E-205	0.0013888888888888889	0.0013888888888888889
35	LWRK vs. —Flexible-PPO—MachineCompressed-model-6	26.981886928497502	2.4113282328489952E-160	0.0014285714285714286	0.0017857142857142859
34	SPT vs. —Flexible-PPO—Complete-model-3	26.04273837794752	1.6259111677652102E-149	0.0014705882352941176	0.0017857142857142859
33	SPT vs. —Flexible-PPO—MachineCompressed-model-4	25.729688861097568	5.441405399684134E-146	0.0015151515151515152	0.0017857142857142859
32	LWRK vs. —Flexible-PPO—Complete-model-3	22.465029613947863	9.125107514809185E-112	0.0015625	0.0017857142857142859
31	LWRK vs. —Flexible-PPO—MachineCompressed-model-4	22.15198009709791	9.98334497432293E-109	0.0016129032258064516	0.0017857142857142859
30	SPT vs. —Flexible-PPO—MachineCompressed-model-7	21.824023460397857	1.3723757925740571E-105	0.0016666666666666668	0.0017857142857142859
29	MS vs. —Flexible-PPO—MachineCompressed-model-6	21.570602422948056	3.3924746025762414E-103	0.001724137931034483	0.0017857142857142859
28	WINQ vs. —Flexible-PPO—MachineCompressed-model-6	20.020621958548218	3.6678771158744114E-89	0.0017857142857142859	0.0017857142857142859
27	SPT vs. —Flexible-PPO—MachineCompressed-model-5	19.513419883648094	8.44396628051309E-85	0.001851851851851852	0.002272727272727273
26	LWRK vs. —Flexible-PPO—MachineCompressed-model-7	18.246314696398198	2.2134690237171673E-74	0.0019230769230769232	0.002272727272727273
25	MS vs. —Flexible-PPO—Complete-model-3	17.053745108398417	3.278379948241815E-65	0.002	0.002272727272727273
24	MS vs. —Flexible-PPO—MachineCompressed-model-4	16.740695591548462	6.620993918955883E-63	0.0020833333333333333	0.002272727272727273
23	LWRK vs. —Flexible-PPO—MachineCompressed-model-5	15.935711119648436	3.5810359595092633E-57	0.002173913043478261	0.002272727272727273
22	WINQ vs. —Flexible-PPO—Complete-model-3	15.503404643998575	3.289870313537462E-54	0.002272727272727273	0.002272727272727273
21	WINQ vs. —Flexible-PPO—MachineCompressed-model-4	15.190355127148623	4.096940885813027E-52	0.002380952380952381	0.002380952380952381
20	MS vs. —Flexible-PPO—MachineCompressed-model-7	12.835030190848752	1.0436228304569693E-37	0.0025	0.0027777777777777778
19	WINQ vs. —Flexible-PPO—MachineCompressed-model-7	11.284689726448912	1.5619172745632465E-29	0.002631578947368421	0.0027777777777777778
18	—Flexible-PPO—MachineCompressed-model-5 vs. —Flexible-PPO—MachineCompressed-model-6	11.046175808849068	2.2875307196442617E-28	0.0027777777777777778	0.0027777777777777778
17	SPT vs. WINQ	10.539333733948945	5.690057749259103E-26	0.0029411764705882353	0.003125
16	MS vs. —Flexible-PPO—MachineCompressed-model-5	10.524426614098989	6.66609326927244E-26	0.003125	0.003125
15	MS vs. SPT	8.988993269549105	2.495053098565265E-19	0.0033333333333333335	0.0033333333333333335
14	WINQ vs. —Flexible-PPO—MachineCompressed-model-5	8.974086149699149	2.8571363354688545E-19	0.0035714285714285718	0.0038461538461538464
13	—Flexible-PPO—MachineCompressed-model-6 vs. —Flexible-PPO—MachineCompressed-model-7	8.735572232099305	2.4242573585673707E-18	0.0038461538461538464	0.0038461538461538464
12	LWRK vs. WINQ	6.961624969949288	3.3636997930586434E-12	0.0041666666666666667	0.0041666666666666667
11	—Flexible-PPO—Complete-model-3 vs. —Flexible-PPO—MachineCompressed-model-5	6.529318494299425	6.606962320187035E-11	0.004545454545454546	0.004545454545454546
10	—Flexible-PPO—MachineCompressed-model-4 vs. —Flexible-PPO—MachineCompressed-model-5	6.216268977449474	5.091150334584427E-10	0.005	0.005
9	LWRK vs. MS	5.411284505549447	6.257426372258213E-8	0.005555555555555556	0.005555555555555556
8	—Flexible-PPO—MachineCompressed-model-4 vs. —Flexible-PPO—MachineCompressed-model-6	4.829906831399594	1.3659695054902853E-6	0.00625	0.00625
7	—Flexible-PPO—Complete-model-3 vs. —Flexible-PPO—MachineCompressed-model-6	4.516857314549642	6.276416664578976E-6	0.0071428571428571435	0.0071428571428571435
6	—Flexible-PPO—Complete-model-3 vs. —Flexible-PPO—MachineCompressed-model-7	4.218714917549663	2.4569870018019462E-5	0.0083333333333333333	0.0083333333333333333
5	—Flexible-PPO—MachineCompressed-model-4 vs. —Flexible-PPO—MachineCompressed-model-7	3.905665400699711	9.396645921747932E-5	0.01	0.01
4	LWRK vs. SPT	3.5777087639996576	3.4661935113463444E-4	0.0125	0.0125
3	—Flexible-PPO—MachineCompressed-model-5 vs. —Flexible-PPO—MachineCompressed-model-7	2.3106035767497626	0.02085476150971584	0.016666666666666666	0.016666666666666666
2	MS vs. WINQ	1.5503404643998404	0.12105981997910276	0.025	0.025
1	—Flexible-PPO—Complete-model-3 vs. —Flexible-PPO—MachineCompressed-model-4	0.31304951684995197	0.7542430423910516	0.05	0.05

Holm’s procedure rejects those hypotheses that have a p-value $\leq 0.016666666666666666$.
Shaffer’s procedure rejects those hypotheses that have a p-value $\leq 0.001388888888888889$.

Table 6: Holm / Shaffer Table for $\alpha = 0.10$

i	algorithms	$z = (R_0 - R_i) / SE$	p	Holm	Shaffer
36	SPT vs. —Flexible-PPO—MachineCompressed-model-6	30.55959569249716	4.216315290413759E-205	0.002777777777777778	0.002777777777777778
35	LWRK vs. —Flexible-PPO—MachineCompressed-model-6	26.981886928497502	2.4113282328489952E-160	0.002857142857142857	0.0035714285714285718
34	SPT vs. —Flexible-PPO—Complete-model-3	26.04273837794752	1.6259111677652102E-149	0.0029411764705882353	0.0035714285714285718
33	SPT vs. —Flexible-PPO—MachineCompressed-model-4	25.729688861097568	5.4411405399684134E-146	0.0030303030303030303	0.0035714285714285718
32	LWRK vs. —Flexible-PPO—Complete-model-3	22.465029613947863	9.125107514809185E-112	0.003125	0.0035714285714285718
31	LWRK vs. —Flexible-PPO—MachineCompressed-model-4	22.15198009709791	9.98334497432293E-109	0.0032258064516129032	0.0035714285714285718
30	SPT vs. —Flexible-PPO—MachineCompressed-model-7	21.824023460397857	1.3723757925740571E-105	0.0033333333333333335	0.0035714285714285718
29	MS vs. —Flexible-PPO—MachineCompressed-model-6	21.570602422948056	3.3924746025762414E-103	0.003448275862068966	0.0035714285714285718
28	WINQ vs. —Flexible-PPO—MachineCompressed-model-6	20.020261958548218	3.6678771158744114E-89	0.0035714285714285718	0.0035714285714285718
27	SPT vs. —Flexible-PPO—MachineCompressed-model-5	19.513419883648094	8.44396628051309E-85	0.003703703703703704	0.004545454545454546
26	LWRK vs. —Flexible-PPO—MachineCompressed-model-7	18.246314696398198	2.2134690237171673E-74	0.0038461538461538464	0.004545454545454546
25	MS vs. —Flexible-PPO—Complete-model-3	17.053745108398417	3.278379948241815E-65	0.004	0.004545454545454546
24	MS vs. —Flexible-PPO—MachineCompressed-model-4	16.740695591548462	6.620993918955883E-63	0.004166666666666667	0.004545454545454546
23	LWRK vs. —Flexible-PPO—MachineCompressed-model-5	15.935711119648436	3.5810359595092633E-57	0.004347826086956522	0.004545454545454546
22	WINQ vs. —Flexible-PPO—Complete-model-3	15.503404643998575	3.289870313537462E-54	0.004545454545454546	0.004545454545454546
21	WINQ vs. —Flexible-PPO—MachineCompressed-model-4	15.190355127148623	4.096940885813027E-52	0.004761904761904762	0.004761904761904762
20	MS vs. —Flexible-PPO—MachineCompressed-model-7	12.835030190848752	1.0436228304569693E-37	0.005	0.005555555555555556
19	WINQ vs. —Flexible-PPO—MachineCompressed-model-7	11.284689726448912	1.5619172745632465E-29	0.005263157894736842	0.005555555555555556
18	—Flexible-PPO—MachineCompressed-model-5 vs. —Flexible-PPO—MachineCompressed-model-6	11.046175808849068	2.2875307196442617E-28	0.005555555555555556	0.005555555555555556
17	SPT vs. WINQ	10.53933733948945	5.690057749259103E-26	0.0058823529411764705	0.00625
16	MS vs. —Flexible-PPO—MachineCompressed-model-5	10.524426614098989	6.666609326927244E-26	0.00625	0.00625
15	MS vs. SPT	8.988993269549105	2.495053098565265E-19	0.006666666666666667	0.006666666666666667
14	WINQ vs. —Flexible-PPO—MachineCompressed-model-5	8.974086149699149	2.8571363354688545E-19	0.0071428571428571435	0.007692307692307693
13	—Flexible-PPO—MachineCompressed-model-6 vs. —Flexible-PPO—MachineCompressed-model-7	8.735572232099305	2.4242573585673707E-18	0.007692307692307693	0.007692307692307693
12	LWRK vs. WINQ	6.961624969949288	3.3636997930586434E-12	0.008333333333333333	0.008333333333333333
11	—Flexible-PPO—Complete-model-3 vs. —Flexible-PPO—MachineCompressed-model-5	6.529318494299425	6.606962320187035E-11	0.009090909090909092	0.009090909090909092
10	—Flexible-PPO—MachineCompressed-model-4 vs. —Flexible-PPO—MachineCompressed-model-5	6.216268977449474	5.091150334584427E-10	0.01	0.01
9	LWRK vs. MS	5.411284505549447	6.257426372258213E-8	0.011111111111111112	0.011111111111111112
8	—Flexible-PPO—MachineCompressed-model-4 vs. —Flexible-PPO—MachineCompressed-model-6	4.829906831399594	1.3659695054902853E-6	0.0125	0.0125
7	—Flexible-PPO—Complete-model-3 vs. —Flexible-PPO—MachineCompressed-model-6	4.516857314549642	6.276416664578976E-6	0.014285714285714287	0.014285714285714287
6	—Flexible-PPO—Complete-model-3 vs. —Flexible-PPO—MachineCompressed-model-7	4.218714917549663	2.4569870018019462E-5	0.016666666666666666	0.016666666666666666
5	—Flexible-PPO—MachineCompressed-model-4 vs. —Flexible-PPO—MachineCompressed-model-7	3.905665400699711	9.396645921747932E-5	0.02	0.02
4	LWRK vs. SPT	3.5777087639996576	3.4661935113463444E-4	0.025	0.025
3	—Flexible-PPO—MachineCompressed-model-5 vs. —Flexible-PPO—MachineCompressed-model-7	2.3106035767497626	0.02085476150971584	0.03333333333333333	0.03333333333333333
2	MS vs. WINQ	1.5503404643998404	0.12105981997910276	0.05	0.05
1	—Flexible-PPO—Complete-model-3 vs. —Flexible-PPO—MachineCompressed-model-4	0.31304951684995197	0.7542430423910516	0.1	0.1

Nemenyi’s procedure rejects those hypotheses that have a p-value $\leq 0.002777777777777778$.
Holm’s procedure rejects those hypotheses that have a p-value ≤ 0.05 .
Shaffer’s procedure rejects those hypotheses that have a p-value $\leq 0.002777777777777778$.

Table 7: Adjusted p -values

i	hypothesis	unadjusted p	P_{Neme}	P_{Holm}	P_{Shaf}	P_{Berg}
1	SPT vs .—Flexible-PPO—MachineCompressed-model-6	4.216315290413759E-205	1.5178735045489532E-203	1.5178735045489532E-203	1.5178735045489532E-203	0.0
2	LWRK vs .—Flexible-PPO—MachineCompressed-model-6	2.4113282328489952E-160	8.680781638256383E-159	8.439648814971483E-159	6.751719051977186E-159	0.0
3	SPT vs .—Flexible-PPO—Complete-model-3	1.6259111677652102E-149	5.853280203954757E-148	5.528097970401714E-148	4.5525512697425884E-148	0.0
4	SPT vs .—Flexible-PPO—MachineCompressed-model-4	5.441405399684134E-146	1.9589059438862883E-144	1.795663781895764E-144	1.5235935119115573E-144	0.0
5	LWRK vs .—Flexible-PPO—Complete-model-3	9.125107514809185E-112	3.2850387053313065E-110	2.920034404738939E-110	2.5550301041465717E-110	0.0
6	LWRK vs .—Flexible-PPO—MachineCompressed-model-4	9.98334497432293E-109	3.5940041907562544E-107	3.094836942040108E-107	2.7953365928104205E-107	0.0
7	SPT vs .—Flexible-PPO—MachineCompressed-model-7	1.3723757925740571E-105	4.9405528532666056E-104	4.117127377722171E-104	3.84265221920736E-104	0.0
8	MS vs .—Flexible-PPO—MachineCompressed-model-6	3.3924746025762414E-103	1.221290856927447E-101	9.8381763474711E-102	9.498928887213475E-102	0.0
9	WINQ vs .—Flexible-PPO—MachineCompressed-model-6	3.6678771158744114E-89	1.320435761714788E-87	1.0270055924448352E-87	1.0270055924448352E-87	0.0
10	SPT vs .—Flexible-PPO—MachineCompressed-model-5	8.44396628051309E-85	3.0398278609847123E-83	2.2798708957385344E-83	1.85767258171288E-83	0.0
11	LWRK vs .—Flexible-PPO—MachineCompressed-model-7	2.2134690237171673E-74	7.968488485381803E-73	5.755019461664635E-73	4.869631852177768E-73	0.0
12	MS vs .—Flexible-PPO—Complete-model-3	3.278379948241815E-65	1.1802167813670533E-63	8.195949870604538E-64	7.212435886131993E-64	0.0
13	MS vs .—Flexible-PPO—MachineCompressed-model-4	6.620993918955883E-63	2.383557810824118E-61	1.589038540549412E-61	1.4566186621702942E-61	0.0
14	LWRK vs .—Flexible-PPO—MachineCompressed-model-5	3.5810359595092633E-57	1.289172945423347E-55	8.236382706871306E-56	7.878279110920379E-56	0.0
15	WINQ vs .—Flexible-PPO—Complete-model-3	3.289870313537462E-54	1.1843533128734863E-52	7.237714689782416E-53	7.237714689782416E-53	0.0
16	WINQ vs .—Flexible-PPO—MachineCompressed-model-4	4.096940885813027E-52	1.4748987188926897E-50	8.603575860207356E-51	8.603575860207356E-51	0.0
17	MS vs .—Flexible-PPO—MachineCompressed-model-7	1.0436228304569693E-37	3.7570421896450896E-36	2.0872456609139387E-36	1.8785210948225448E-36	0.0
18	WINQ vs .—Flexible-PPO—MachineCompressed-model-7	1.5619172745632465E-29	5.622902188427687E-28	2.9676428216701685E-28	2.8114510942138436E-28	0.0
19	—Flexible-PPO—MachineCompressed-model-5 vs .—Flexible-PPO—MachineCompressed-model-6	2.2875307196442617E-28	8.235110590719342E-27	4.117555295359671E-27	4.117555295359671E-27	0.0
20	SPT vs .WINQ	5.690057749259103E-26	2.048420789733277E-24	9.673098173740474E-25	9.104092398814564E-25	0.0
21	MS vs .—Flexible-PPO—MachineCompressed-model-5	6.66609326927244E-26	2.3999793576938077E-24	1.066657492308359E-24	1.066657492308359E-24	0.0
22	MS vs .SPT	2.495053098565265E-19	8.982191154834954E-18	3.7425796478478976E-18	3.7425796478478976E-18	0.0
23	WINQ vs .—Flexible-PPO—MachineCompressed-model-5	2.8571363354688545E-19	1.0285690807687876E-17	3.999990869656396E-18	3.7425796478478976E-18	0.0
24	—Flexible-PPO—MachineCompressed-model-6 vs .—Flexible-PPO—MachineCompressed-model-7	2.4242573585673707E-18	8.727326490842534E-17	3.151534566137582E-17	3.151534566137582E-17	0.0
25	LWRK vs .WINQ	3.3636997930586434E-12	1.2109319255011117E-10	4.036439751670372E-11	4.036439751670372E-11	0.0
26	—Flexible-PPO—Complete-model-3 vs .—Flexible-PPO—MachineCompressed-model-5	6.606962320187035E-11	2.3785064352673326E-9	7.267658552205738E-10	7.267658552205738E-10	0.0
27	—Flexible-PPO—MachineCompressed-model-4 vs .—Flexible-PPO—MachineCompressed-model-5	5.091150334584427E-10	1.8328141204503938E-8	5.091150334584427E-9	5.091150334584427E-9	0.0
28	LWRK vs .MS	6.257426372258213E-8	2.252673494012957E-6	5.631683735032392E-7	5.631683735032392E-7	0.0
29	—Flexible-PPO—MachineCompressed-model-4 vs .—Flexible-PPO—MachineCompressed-model-6	1.3659695054902853E-6	4.917490219765027E-5	1.0927756043922282E-5	1.0927756043922282E-5	0.0
30	—Flexible-PPO—Complete-model-3 vs .—Flexible-PPO—MachineCompressed-model-6	6.276416664578976E-6	2.2595099992484314E-4	4.393491665205283E-5	4.393491665205283E-5	0.0
31	—Flexible-PPO—Complete-model-3 vs .—Flexible-PPO—MachineCompressed-model-7	2.4569870018019462E-5	8.845153206487006E-4	1.4741922010811677E-4	1.4741922010811677E-4	0.0
32	—Flexible-PPO—MachineCompressed-model-4 vs .—Flexible-PPO—MachineCompressed-model-7	9.396645921747932E-5	0.003382792531829256	4.698322960873966E-4	4.698322960873966E-4	0.0
33	LWRK vs .SPT	3.4661935113463444E-4	0.01247829664084684	0.0013864774045385378	0.0013864774045385378	0.0
34	—Flexible-PPO—MachineCompressed-model-5 vs .—Flexible-PPO—MachineCompressed-model-7	0.02085476150971584	0.7507714143497701	0.06256428452914752	0.06256428452914752	0.0
35	MS vs .WINQ	0.12105981997910276	4.358153519247699	0.2421196399582055	0.2421196399582055	0.0
36	—Flexible-PPO—Complete-model-3 vs .—Flexible-PPO—MachineCompressed-model-4	0.7542430423910516	27.152749526077855	0.7542430423910516	0.7542430423910516	0.0