

1 kMeans

1.1 Scores

Dataset	FeatureSet	adjusted_rand	v_measure
Arcene	1-distance	-0.006465	0.004965
	all-distance	-0.006760	0.006098
	cardinality	-0.011700	0.000675
	impurities	-0.006465	0.004965
	induced	-0.006465	0.004965
	original	-0.008967	0.004114
BloodTransfusion	1-distance	-0.003782	0.002600
	all-distance	0.006607	0.000192
	cardinality	-0.019734	0.013180
	impurities	-0.003782	0.002600
	induced	-0.003782	0.002600
	original	-0.006249	0.000271
BreastCancerDiag	1-distance	0.701895	0.614174
	all-distance	0.320927	0.284364
	cardinality	0.742996	0.625092
	impurities	0.701895	0.614174
	induced	0.701895	0.614174
	original	0.730175	0.623086
ConnectionistBenchSonar	1-distance	-0.004429	0.001501
	all-distance	-0.001420	0.008729
	cardinality	-0.004702	0.000198
	impurities	-0.004429	0.001501
	induced	-0.004429	0.001501
	original	0.006409	0.008560
Ecoli	1-distance	0.389873	0.583985
	all-distance	0.441016	0.548142
	cardinality	0.472276	0.556490
	impurities	0.372906	0.550758
	induced	0.376303	0.551535
	original	0.454764	0.598052
GlassIdentification	1-distance	0.230023	0.381088
	all-distance	0.193896	0.311725
	cardinality	0.153630	0.351215
	impurities	0.233918	0.394491
	induced	0.221757	0.372039
	original	0.164356	0.324170
HabermansSurvival	1-distance	-0.003979	0.000741
	all-distance	-0.003979	0.000741
	cardinality	-0.007615	0.003101
	impurities	-0.003979	0.000741
	induced	-0.003979	0.000741
	original	-0.003979	0.000741
IndianLiverPatients	1-distance	0.032686	0.005730
	all-distance	0.032686	0.005730
	cardinality	0.032686	0.005730
	impurities	0.032686	0.005730
	induced	0.032686	0.005730
	original	0.032686	0.005730
Ionosphere	1-distance	0.097159	0.098255
	all-distance	0.161769	0.096197
	cardinality	0.182808	0.147080
	impurities	0.100870	0.100503
	induced	0.104641	0.102781
	original	0.177607	0.134857

1.2 Wins

1.2.1 v_measure

	original	induced	impurities	cardinality	1-distance	all-disctance
original	0	5	5	4	5	8
induced	6	0	3	6	2	9
impurities	6	1	0	6	3	9
cardinality	7	6	6	0	6	9
1-distance	6	3	2	6	0	9
all-disctance	3	2	2	3	2	0

1.2.2 adjusted_rand

	original	induced	impurities	cardinality	1-distance	all-disctance
original	0	5	5	6	5	7
induced	6	0	3	8	2	7
impurities	6	1	0	8	3	7
cardinality	5	4	4	0	4	6
1-distance	6	3	2	8	0	7
all-disctance	4	4	4	6	4	0

2 1-NN

2.1 Scores

Dataset	FeatureSet	accuracy	auroc
Arcene	1-distance	0.600000	0.591667
	all-distance	0.550000	0.541667
	cardinality	0.590000	0.583333
	impurities	0.600000	0.591667
	induced	0.600000	0.591667
	original	0.640000	0.620833
BloodTransfusion	1-distance	0.549892	0.453225
	all-distance	0.641964	0.530693
	cardinality	0.565946	0.483084
	impurities	0.549892	0.453225
	induced	0.549892	0.453225
	original	0.568559	0.474845
BreastCancerDiag	1-distance	0.950897	0.950422
	all-distance	0.927997	0.922338
	cardinality	0.954376	0.951255
	impurities	0.950897	0.950422
	induced	0.950897	0.950422
	original	0.954194	0.952807
ConnectionistBenchSonar	1-distance	0.630065	0.627626
	all-distance	0.645498	0.639747
	cardinality	0.630974	0.627475
	impurities	0.639827	0.636717
	induced	0.639827	0.636717
	original	0.693420	0.688460
Ecoli	1-distance	0.801999	NaN
	all-distance	0.799348	NaN
	cardinality	0.785118	NaN
	impurities	0.801999	NaN
	induced	0.801999	NaN
	original	0.803067	NaN
GlassIdentification	1-distance	0.693175	NaN
	all-distance	0.580233	NaN
	cardinality	0.682306	NaN
	impurities	0.693175	NaN
	induced	0.693175	NaN
	original	0.656928	NaN
HabermansSurvival	1-distance	0.644664	0.538560
	all-distance	0.634335	0.539995
	cardinality	0.650793	0.539427
	impurities	0.644664	0.538560
	induced	0.638105	0.530037
	original	0.653925	0.541203
IndianLiverPatients	1-distance	0.655885	0.619912
	all-distance	0.640511	0.571446
	cardinality	0.590886	0.543170
	impurities	0.657580	0.621102
	induced	0.657580	0.621102
	original	0.643397	0.582193
Ionosphere	1-distance	0.923861	0.908274
	all-distance	0.915364	0.908342
	cardinality	0.912838	0.894981
	impurities	0.918310	0.899008
	induced	0.915369	0.894841
	original	0.861120	0.817429

2.2 Wins

2.2.1 accuracy

	original	induced	impurities	cardinality	1-distance	all-distance
original	0	7	7	7	7	10
induced	6	0	2	6	2	10
impurities	6	3	0	6	2	10
cardinality	6	7	7	0	8	7
1-distance	6	5	5	5	0	10
all-distance	3	3	3	6	3	0

2.2.2 auroc

	original	induced	impurities	cardinality	1-distance	all-distance
original	0	5	5	5	5	6
induced	3	0	1	3	2	4
impurities	3	3	0	4	3	4
cardinality	3	5	4	0	4	3
1-distance	3	2	2	4	0	4
all-distance	2	4	4	5	4	0

3 3-NN

3.1 Scores

Dataset	FeatureSet	accuracy	auroc
Arcene	1-distance	0.610000	0.616667
	all-distance	0.530000	0.533333
	cardinality	0.620000	0.610417
	impurities	0.620000	0.616667
	induced	0.620000	0.616667
	original	0.550000	0.558333
BloodTransfusion	1-distance	0.634000	0.481567
	all-distance	0.650054	0.495975
	cardinality	0.639387	0.496394
	impurities	0.631315	0.490881
	induced	0.631315	0.477981
	original	0.631297	0.485612
BreastCancerDiag	1-distance	0.956133	0.980746
	all-distance	0.943973	0.963918
	cardinality	0.959580	0.979641
	impurities	0.956133	0.980813
	induced	0.956133	0.980813
	original	0.970201	0.982742
ConnectionistBenchSonar	1-distance	0.619827	0.675328
	all-distance	0.623874	0.614356
	cardinality	0.610801	0.682184
	impurities	0.600758	0.672904
	induced	0.610281	0.676843
	original	0.643182	0.719987
Ecoli	1-distance	0.840310	NaN
	all-distance	0.827841	NaN
	cardinality	0.824508	NaN
	impurities	0.837084	NaN
	induced	0.837084	NaN
	original	0.826728	NaN
GlassIdentification	1-distance	0.678241	NaN
	all-distance	0.621519	NaN
	cardinality	0.688912	NaN
	impurities	0.678241	NaN
	induced	0.678241	NaN
	original	0.655450	NaN
HabermansSurvival	1-distance	0.683488	0.602778
	all-distance	0.676815	0.587562
	cardinality	0.686499	0.580892
	impurities	0.696606	0.618632
	induced	0.693273	0.617212
	original	0.696183	0.610437
IndianLiverPatients	1-distance	0.638578	0.632687
	all-distance	0.626596	0.590480
	cardinality	0.631949	0.610246
	impurities	0.640332	0.631848
	induced	0.640273	0.631837
	original	0.645089	0.613579
Ionosphere	1-distance	0.924108	0.950241
	all-distance	0.918226	0.951987
	cardinality	0.909897	0.941419
	impurities	0.923866	0.950320
	induced	0.918231	0.945972
	original	0.855733	0.891185

3.2 Wins

3.2.1 accuracy

	original	induced	impurities	cardinality	1-distance	all-distance
original	0	5	4	7	5	8
induced	8	0	3	5	4	9
impurities	9	3	0	5	4	9
cardinality	6	6	6	0	8	8
1-distance	8	6	6	5	0	9
all-distance	5	4	4	5	4	0

3.2.2 auROC

	original	induced	impurities	cardinality	1-distance	all-distance
original	0	3	2	4	4	5
induced	5	0	2	5	4	5
impurities	6	4	0	5	5	5
cardinality	4	3	3	0	3	6
1-distance	3	3	2	5	0	5
all-distance	3	3	3	2	3	0

4 5-NN

4.1 Scores

Dataset	FeatureSet	accuracy	auroc
Arcene	1-distance	0.590000	0.595833
	all-distance	0.540000	0.516667
	cardinality	0.610000	0.620833
	impurities	0.590000	0.606250
	induced	0.590000	0.606250
	original	0.590000	0.497917
BloodTransfusion	1-distance	0.704685	0.594846
	all-distance	0.648685	0.511447
	cardinality	0.667423	0.547558
	impurities	0.658018	0.559787
	induced	0.667351	0.582293
	original	0.656685	0.513966
BreastCancerDiag	1-distance	0.956164	0.984007
	all-distance	0.945698	0.977121
	cardinality	0.964907	0.988065
	impurities	0.956164	0.985901
	induced	0.956164	0.986090
	original	0.966693	0.986791
ConnectionistBenchSonar	1-distance	0.610996	0.683851
	all-distance	0.599589	0.585783
	cardinality	0.654156	0.716263
	impurities	0.606234	0.673775
	induced	0.610996	0.675189
	original	0.604805	0.695455
Ecoli	1-distance	0.854622	NaN
	all-distance	0.815693	NaN
	cardinality	0.833900	NaN
	impurities	0.851990	NaN
	induced	0.851990	NaN
	original	0.838367	NaN
GlassIdentification	1-distance	0.710256	NaN
	all-distance	0.669743	NaN
	cardinality	0.688554	NaN
	impurities	0.696817	NaN
	induced	0.691817	NaN
	original	0.666401	NaN
HabermansSurvival	1-distance	0.686808	0.609450
	all-distance	0.673159	0.560476
	cardinality	0.709194	0.619805
	impurities	0.693367	0.599380
	induced	0.693259	0.603555
	original	0.705853	0.603936
IndianLiverPatients	1-distance	0.671018	0.653922
	all-distance	0.650680	0.609481
	cardinality	0.633287	0.628298
	impurities	0.672654	0.652558
	induced	0.667511	0.653019
	original	0.634771	0.617694
Ionosphere	1-distance	0.932358	0.957336
	all-distance	0.909972	0.954331
	cardinality	0.909650	0.962510
	impurities	0.923866	0.950647
	induced	0.918147	0.949455
	original	0.841438	0.920457

4.2 Wins

4.2.1 accuracy

	original	induced	impurities	cardinality	1-distance	all-distance
original	0	4	4	5	4	8
induced	9	0	3	5	3	12
impurities	9	6	0	5	3	12
cardinality	8	8	7	0	7	11
1-distance	9	6	6	5	0	12
all-distance	5	1	1	2	1	0

4.2.2 auROC

	original	induced	impurities	cardinality	1-distance	all-distance
original	0	3	3	1	2	5
induced	5	0	6	3	3	6
impurities	5	1	0	3	3	6
cardinality	7	5	5	0	5	7
1-distance	6	5	5	3	0	7
all-distance	3	2	2	1	1	0

5 7-NN

5.1 Scores

Dataset	FeatureSet	accuracy	auroc
Arcene	1-distance	0.570000	0.568750
	all-distance	0.520000	0.491667
	cardinality	0.550000	0.545833
	impurities	0.570000	0.575000
	induced	0.570000	0.575000
	original	0.580000	0.541667
BloodTransfusion	1-distance	0.706072	0.611716
	all-distance	0.674054	0.554925
	cardinality	0.675405	0.588829
	impurities	0.655387	0.570110
	induced	0.664721	0.599923
	original	0.667351	0.537650
BreastCancerDiag	1-distance	0.963213	0.988076
	all-distance	0.942188	0.981802
	cardinality	0.964845	0.990340
	impurities	0.963213	0.988194
	induced	0.963213	0.988194
	original	0.966724	0.990176
ConnectionistBenchSonar	1-distance	0.644567	0.664924
	all-distance	0.564589	0.579331
	cardinality	0.653658	0.701376
	impurities	0.620043	0.659217
	induced	0.620043	0.651351
	original	0.598442	0.664747
Ecoli	1-distance	0.851002	NaN
	all-distance	0.813472	NaN
	cardinality	0.830039	NaN
	impurities	0.851086	NaN
	induced	0.854211	NaN
	original	0.852555	NaN
GlassIdentification	1-distance	0.700453	NaN
	all-distance	0.621896	NaN
	cardinality	0.678792	NaN
	impurities	0.690710	NaN
	induced	0.690710	NaN
	original	0.632209	NaN
HabermansSurvival	1-distance	0.709496	0.628497
	all-distance	0.695961	0.555842
	cardinality	0.725316	0.653029
	impurities	0.712937	0.624115
	induced	0.712937	0.627535
	original	0.722628	0.606540
IndianLiverPatients	1-distance ¹⁰	0.681189	0.659789
	all-distance	0.665812	0.625731
	cardinality	0.677504	0.654864
	impurities	0.679434	0.655123
	induced	0.679434	0.655342
	original	0.667154	0.631605
Ionosphere	1-distance	0.940854	0.966924
	all-distance	0.898618	0.950314
	cardinality	0.912512	0.959823
	impurities	0.929585	0.957659
	induced	0.929585	0.955819
	original	0.844211	0.924426

5.2 Wins

5.2.1 accuracy

	original	induced	impurities	cardinality	1-distance	all-distance
original	0	6	7	5	6	9
induced	7	0	2	6	6	11
impurities	6	1	0	6	6	11
cardinality	8	7	7	0	6	13
1-distance	7	5	5	6	0	12
all-distance	4	2	2	0	1	0

5.2.2 auroc

	original	induced	impurities	cardinality	1-distance	all-distance
original	0	2	2	0	1	5
induced	6	0	4	3	3	8
impurities	6	2	0	2	3	8
cardinality	8	4	6	0	4	8
1-distance	7	5	5	4	0	8
all-distance	3	0	0	0	0	0

6 9-NN

6.1 Scores

Dataset	FeatureSet	accuracy	auroc
Arcene	1-distance	0.570000	0.518750
	all-distance	0.600000	0.570833
	cardinality	0.530000	0.525000
	impurities	0.570000	0.518750
	induced	0.570000	0.514583
	original	0.520000	0.545833
BloodTransfusion	1-distance	0.698054	0.611275
	all-distance	0.674054	0.601322
	cardinality	0.688739	0.598042
	impurities	0.671387	0.590801
	induced	0.695387	0.620408
	original	0.674018	0.556029
BreastCancerDiag	1-distance	0.959643	0.988995
	all-distance	0.940434	0.982581
	cardinality	0.963091	0.989679
	impurities	0.961428	0.989256
	induced	0.959643	0.989054
	original	0.966785	0.990249
ConnectionistBenchSonar	1-distance	0.639329	0.629798
	all-distance	0.545996	0.551881
	cardinality	0.634827	0.687891
	impurities	0.601688	0.630846
	induced	0.601472	0.648018
	original	0.594805	0.647689
Ecoli	1-distance	0.860562	NaN
	all-distance	0.791770	NaN
	cardinality	0.829570	NaN
	impurities	0.860378	NaN
	induced	0.857153	NaN
	original	0.859256	NaN
GlassIdentification	1-distance	0.676499	NaN
	all-distance	0.606167	NaN
	cardinality	0.668333	NaN
	impurities	0.681045	NaN
	induced	0.686045	NaN
	original	0.590231	NaN
HabermansSurvival	1-distance	0.728864	0.653121
	all-distance	0.712843	0.527807
	cardinality	0.754886	0.662467
	impurities	0.728757	0.654404
	induced	0.728757	0.655088
	original	0.725423	0.615648
IndianLiverPatients	1-distance ¹²	0.691210	0.670251
	all-distance	0.670987	0.633310
	cardinality	0.675956	0.667490
	impurities	0.687850	0.668738
	induced	0.687850	0.667576
	original	0.674140	0.642371
Ionosphere	1-distance	0.941018	0.964497
	all-distance	0.895840	0.956237
	cardinality	0.906629	0.961051
	impurities	0.929664	0.957744
	induced	0.929500	0.956072
	original	0.838413	0.922041

6.2 Wins

6.2.1 accuracy

	original	induced	impurities	cardinality	1-distance	all-distance
original	0	3	3	3	2	7
induced	10	0	2	7	2	11
impurities	10	5	0	6	3	11
cardinality	10	6	7	0	5	12
1-distance	11	9	9	8	0	12
all-distance	6	2	2	1	1	0

6.2.2 auroc

	original	induced	impurities	cardinality	1-distance	all-distance
original	0	2	3	2	3	4
induced	6	0	4	3	5	5
impurities	5	4	0	2	4	5
cardinality	6	5	6	0	4	5
1-distance	5	3	3	4	0	6
all-distance	4	3	3	3	2	0

7 Parameters

number_of_trees: 200

execution_time: 3863.796885728836

percentage_of_removal: 0