

Test accuracy comparison of KNN (noise rate 5%)

Datasets	KNN	CAGBS-KNN	Best purity	Accuracy improvement
diabetes	0.6095	0.6537	0.90	0.0442
sonar	0.4861	0.5382	0.93	0.0521
creditApproval	0.6478	0.6899	0.91	0.0421
votes	0.8920	0.9195	0.96	0.0275
fourclass	0.6900	0.7714	0.67	0.0814
breastcancer	0.8278	0.8331	0.90	0.0053
isolet5	0.9756	1	0.93	0.0244
mushrooms	0.9930	0.9957	0.97	0.0027
svmguide1	0.4799	0.5225	0.78	0.0426
splice	0.6900	0.7030	0.65	0.0130
clean1	0.9937	1	0.80	0.0063
svmguide3	0.6783	0.7437	0.98	0.0654
codrna	0.6849	0.6528	0.78	-0.0321

Test accuracy comparison of KNN(noise rate 10%)

Datasets	KNN	CAGBS-KNN	Best purity	Accuracy improvement
diabetes	0.5831	0.6576	0.81	0.0745
sonar	0.4979	0.5869	0.71	0.0890
creditApproval	0.6826	0.6913	0.71	0.0087
votes	0.8391	0.8805	0.97	0.0414
fourclass	0.6600	0.7656	0.64	0.1056
breastcancer	0.7266	0.8375	0.83	0.1109
isolet5	0.9480	1	0.80	0.0520
mushrooms	0.9751	0.9935	0.96	0.0184
svmguide1	0.4579	0.5386	0.77	0.0807
splice	0.6880	0.6940	0.61	0.0060
clean1	0.9895	1	0.62	0.0105
svmguide3	0.6970	0.7398	0.87	0.0428
codrna	0.6425	0.6507	0.84	0.0082

Test accuracy comparison of KNN(noise rate 15%)

Datasets	KNN	CAGBS-KNN	Best purity	Accuracy improvement
diabetes	0.5166	0.6537	0.81	0.1371
sonar	0.4858	0.5334	0.91	0.0476
creditApproval	0.6029	0.6826	0.91	0.0797
votes	0.8230	0.8782	0.93	0.0552
fourclass	0.6277	0.7146	0.78	0.0869
breastcancer	0.7214	0.8258	0.83	0.1044
isolet5	0.9282	1	0.77	0.0718
mushrooms	0.9420	0.9913	0.92	0.0493
svmguide1	0.4712	0.5413	0.71	0.0701
splice	0.6190	0.6720	0.62	0.0530
clean1	0.9832	1	0.70	0.0168
svmguide3	0.6604	0.7305	0.92	0.0701
codrna	0.6145	0.6393	0.82	0.0248

Test accuracy comparison of KNN(noise rate 20%)

Datasets	KNN	CAGBS-KNN	Best purity	Accuracy improvement
diabetes	0.5235	0.6589	0.96	0.1354
sonar	0.4566	0.5386	0.74	0.0820
creditApproval	0.6493	0.6705	0.80	0.0212
votes	0.7655	0.8782	0.89	0.1127
fourclass	0.6299	0.7100	0.74	0.0801
breastcancer	0.7339	0.8551	0.87	0.1212
isolet5	0.8916	1	0.90	0.1084
mushrooms	0.8871	0.9846	0.98	0.0975
svmguide1	0.4518	0.5849	0.67	0.1331
splice	0.6350	0.6550	0.64	0.0200
clean1	0.9790	1	0.74	0.0210
svmguide3	0.6296	0.7227	0.90	0.0931
codrna	0.6058	0.6122	0.74	0.0064

Test accuracy comparison of KNN(noise rate 30%)

Datasets	KNN	CAGBS-KNN	Best purity	Accuracy improvement
diabetes	0.5004	0.6082	0.67	0.1078
sonar	0.2401	0.7315	0.74	0.4914
creditApproval	0.5710	0.6594	0.96	0.0884
votes	0.7310	0.8713	0.98	0.1403
fourclass	0.5427	0.6763	0.78	0.1336
breastcancer	0.6090	0.8565	0.74	0.2475
isolet5	0.8897	1	0.62	0.1103
mushrooms	0.7719	0.9307	0.86	0.1588
svmguide1	0.4282	0.5875	0.61	0.1593
splice	0.5720	0.6460	0.61	0.0740
clean1	0.9811	1	0.96	0.0189
svmguide3	0.5755	0.7095	0.81	0.1340
codrna	0.5695	0.6128	0.79	0.0433

Test accuracy comparison of KNN(noise rate 40%)

Datasets	KNN	CAGBS-KNN	Best purity	Accuracy improvement
diabetes	0.5224	0.6719	0.78	0.1495
sonar	0.4037	0.4564	0.88	0.0527
creditApproval	0.5522	0.6493	0.98	0.0971
votes	0.554	0.7632	0.92	0.2092
fourclass	0.5753	0.6914	0.85	0.1161
breastcancer	0.5620	0.6914	0.84	0.1294
isolet5	0.8230	0.9731	0.72	0.1501
mushrooms	0.6640	0.8090	0.81	0.1450
svmguide1	0.4280	0.4709	0.69	0.0429
splice	0.5090	0.5330	0.65	0.0240
clean1	0.9790	0.9874	0.80	0.0084
svmguide3	0.5140	0.6300	0.95	0.1160
codrna	0.5556	0.5952	0.75	0.0396

Test accuracy comparison of DT(noise rate 5%)

Datasets	DT	CAGBS-DT	Best purity	Accuracy improvement
diabetes	0.5012	0.6745	0.85	0.1733
sonar	0.4767	0.5822	0.73	0.1055
creditApproval	0.4594	0.5551	0.99	0.0957
votes	0.6345	0.8552	0.94	0.2207
fourclass	0.5186	0.7262	1	0.2076
breastcancer	0.8828	0.8725	0.97	-0.0103
isolet5	0.8691	1	0.61	0.1309
mushrooms	0.9132	0.9948	0.98	0.0816
svmguide1	0.4590	0.5729	0.77	0.1139
splice	0.4890	0.6530	0.78	0.1640
clean1	0.8952	1	0.78	0.1048
svmguide3	0.4828	0.7041	0.82	0.2213
codrna	0.5110	0.5721	0.68	0.0611

Test accuracy comparison of DT (noise rate 10%)

Datasets	DT	CAGBS-DT	Best purity	Accuracy improvement
diabetes	0.4267	0.6602	0.70	0.2335
sonar	0.4330	0.5382	0.78	0.1052
creditApproval	0.4014	0.5551	0.98	0.1537
votes	0.6322	0.8506	0.77	0.2184
fourclass	0.5751	0.7204	1	0.1453
breastcancer	0.8345	0.9017	0.92	0.0672
isolet5	0.8210	1	0.62	0.1790
mushrooms	0.8587	0.9876	0.97	0.1289
svmguide1	0.5372	0.5720	0.68	0.0348
splice	0.4840	0.5770	0.92	0.0930
clean1	0.8717	1	0.64	0.1283
svmguide3	0.4796	0.7196	0.92	0.2400
codrna	0.4921	0.5624	0.82	0.0703

Test accuracy comparison of DT (noise rate 15%)

Datasets	DT	CAGBS-DT	Best purity	Accuracy improvement
diabetes	0.4269	0.6368	0.77	0.2099
sonar	0.4175	0.5530	0.76	0.1355
creditApproval	0.4826	0.5580	0.88	0.0754
votes	0.4483	0.8092	0.96	0.3609
fourclass	0.5950	0.7135	0.97	0.1185
breastcancer	0.5106	0.8622	0.89	0.3516
isolet5	0.7972	1	0.97	0.2028
mushrooms	0.8039	0.9894	0.96	0.1855
svmguide1	0.4531	0.5631	0.64	0.1100
splice	0.4370	0.6170	0.79	0.1800
clean1	0.8465	1	0.69	0.1535
svmguide3	0.4410	0.6900	0.88	0.2490
codrna	0.5305	0.5943	0.80	0.0638

Test accuracy comparison of DT (noise rate 20%)

Datasets	DT	CAGBS-DT	Best purity	Accuracy improvement
diabetes	0.3359	0.6939	0.95	0.3580
sonar	0.4560	0.5668	0.85	0.1108
creditApproval	0.4768	0.5551	0.79	0.0783
votes	0.4621	0.8046	0.95	0.3425
fourclass	0.5348	0.7204	0.98	0.1856
breastcancer	0.2211	0.8652	0.88	0.6441
isolet5	0.8377	1	0.67	0.1623
mushrooms	0.7594	0.9844	0.88	0.2250
svmguide1	0.4177	0.4792	0.91	0.0615
splice	0.4470	0.5580	0.91	0.1110
clean1	0.7859	1	0.75	0.2141
svmguide3	0.2937	0.7194	0.82	0.4257
codrna	0.5189	0.5762	0.91	0.0573

Test accuracy comparison of DT (noise rate 30%)

Datasets	DT	CAGBS-DT	Best purity	Accuracy improvement
diabetes	0.4559	0.6081	0.66	0.1522
sonar	0.5056	0.5339	0.96	0.0283
creditApproval	0.4739	0.6116	0.84	0.1377
votes	0.4713	0.8437	0.97	0.3724
fourclass	0.4629	0.6428	0.85	0.1799
breastcancer	0.4614	0.8331	0.86	0.3717
isolet5	0.7780	1	0.91	0.2220
mushrooms	0.6620	0.9366	0.78	0.2746
svmguide1	0.4885	0.5514	0.69	0.0629
splice	0.4840	0.5910	0.79	0.1070
clean1	0.8130	1	0.82	0.1870
svmguide3	0.3629	0.6737	0.75	0.3108
codrna	0.5110	0.6012	0.70	0.0902

Test accuracy comparison of DT (noise rate 40%)

Datasets	DT	CAGBS-DT	Best purity	Accuracy improvement
diabetes	0.3008	0.4868	0.74	0.1860
sonar	0.4667	0.5287	0.96	0.0620
creditApproval	0.5159	0.5957	0.74	0.0798
votes	0.4092	0.6598	0.90	0.2506
fourclass	0.5289	0.6891	0.76	0.1602
breastcancer	0.5621	0.8375	0.85	0.2754
isolet5	0.8288	1	0.98	0.1712
mushrooms	0.5825	0.7394	0.96	0.1569
svmguide1	0.5444	0.5572	0.74	0.0128
splice	0.4580	0.5230	0.67	0.0650
clean1	0.8048	1	0.87	0.1952
svmguide3	0.3848	0.5880	0.81	0.2032
codrna	0.4643	0.5104	0.86	0.0461

Test accuracy comparison of XGBoost (noise rate 5%)

Datasets	XGBoost	CAGBS-XGBoost	Best purity	Accuracy improvement
diabetes	0.6556	0.6546	0.96	-0.0010
sonar	0.4791	0.4843	0.97	0.0052
creditApproval	0.5445	0.5551	0.93	0.0106
votes	0.9076	0.9287	0.69	0.0211
fourclass	0.7160	0.7309	0.88	0.0149
breastcancer	0.8713	0.8535	0.83	-0.0178
isolet5	0.9716	1	1	0.0284
mushrooms	0.9996	0.8984	0.66	-0.1012
svmguide1	0.4736	0.5406	0.79	0.0670
splice	0.5144	0.5807	0.87	0.0663
clean1	0.9958	1	0.89	0.0042
svmguide3	0.4165	0.7417	0.99	0.3252
codrna	0.5297	0.6617	0.66	0.1320

Test accuracy comparison of XGBoost (noise rate 10%)

Datasets	XGBoost	CAGBS-XGBoost	Best purity	Accuracy improvement
diabetes	0.6610	0.6552	0.92	-0.0058
sonar	0.4800	0.4843	0.65	0.0043
creditApproval	0.5393	0.5542	1	0.0149
votes	0.8549	0.8906	0.62	0.0357
fourclass	0.7120	0.7767	0.64	0.0647
breastcancer	0.8551	0.8613	0.78	0.0062
isolet5	0.9276	1	0.61	0.0724
mushrooms	0.9995	0.8966	0.61	-0.1029
svmguide1	0.4733	0.5616	0.77	0.0883
splice	0.5210	0.5833	0.81	0.0623
clean1	0.9975	1	0.98	0.0025
svmguide3	0.4209	0.7395	0.90	0.3186
codrna	0.5703	0.6279	0.67	0.0576

Test accuracy comparison of XGBoost (noise rate 15%)

Datasets	XGBoost	CAGBS-XGBoost	Best purity	Accuracy improvement
diabetes	0.6448	0.6552	0.96	0.0104
sonar	0.4868	0.4886	0.77	0.0018
creditApproval	0.5479	0.5596	0.85	0.0117
votes	0.794	0.9062	0.62	0.1122
fourclass	0.7023	0.7654	0.62	0.0631
breastcancer	0.8471	0.8595	0.70	0.0124
isolet5	0.9017	1	0.68	0.0983
mushrooms	0.9984	0.8952	0.61	-0.1032
svmguide1	0.4453	0.5516	0.73	0.1063
splice	0.5230	0.5992	0.69	0.0762
clean1	0.9985	1	0.81	0.0015
svmguide3	0.4799	0.7360	0.93	0.2561
codrna	0.5925	0.6647	0.65	0.0722

Test accuracy comparison of XGBoost (noise rate 20%)

Datasets	XGBoost	CAGBS-XGBoost	Best purity	Accuracy improvement
diabetes	0.6209	0.6446	0.96	0.0237
sonar	0.4918	0.4830	0.62	-0.0088
creditApproval	0.5470	0.5568	0.9	0.0098
votes	0.6837	0.7903	0.81	0.1066
fourclass	0.6821	0.7124	0.89	0.0303
breastcancer	0.8030	0.8629	0.74	0.0599
isolet5	0.8814	1	0.84	0.1186
mushrooms	0.9973	0.9200	0.69	-0.0773
svmguide1	0.4871	0.5895	0.71	0.1024
splice	0.5434	0.5769	0.85	0.0335
clean1	0.9992	1	0.78	0.0008
svmguide3	0.4585	0.7338	0.88	0.2753
codrna	0.6057	0.6273	0.64	0.0216

Test accuracy comparison of XGBoost (noise rate 30%)

Datasets	XGBoost	CAGBS-XGBoost	Best purity	Accuracy improvement
diabetes	0.5699	0.6035	0.85	0.0336
sonar	0.4899	0.4875	0.88	-0.0024
creditApproval	0.5107	0.5554	0.96	0.0447
votes	0.5975	0.7064	0.86	0.1089
fourclass	0.6337	0.6832	0.72	0.0495
breastcancer	0.7764	0.8485	0.67	0.0721
isolet5	0.8500	1	0.65	0.1500
mushrooms	0.9950	0.8991	0.61	-0.0959
svmguide1	0.4613	0.4886	0.98	0.0273
splice	0.5455	0.5837	0.63	0.0382
clean1	1	0.9811	0.63	-0.0189
svmguide3	0.4570	0.7031	0.80	0.2461
codrna	0.6207	0.6476	0.72	0.0269

Test accuracy comparison of XGBoost (noise rate 40%)

Datasets	XGBoost	CAGBS-XGBoost	Best purity	Accuracy improvement
diabetes	0.5107	0.6219	0.65	0.1112
sonar	0.4851	0.4879	0.86	0.0028
creditApproval	0.4986	0.5317	0.99	0.0331
votes	0.4805	0.6228	0.87	0.1423
fourclass	0.4954	0.5425	0.64	0.0471
breastcancer	0.5659	0.7603	0.64	0.1944
isolet5	0.7827	1	0.79	0.2173
mushrooms	0.9697	0.7933	0.61	-0.1764
svmguide1	0.4546	0.5528	0.66	0.0982
splice	0.5490	0.5312	0.71	-0.0178
clean1	1	0.9159	0.71	-0.0841
svmguide3	0.4549	0.6418	0.96	0.1869
codrna	0.5673	0.6667	0.90	0.0994

Test accuracy comparison of lightGBM (noise rate 5%)

Datasets	lightGBM	CAGBS-lightGBM	Best purity	Accuracy improvement
diabetes	0.6598	0.6514	0.62	-0.0084
sonar	0.4777	0.4813	0.91	0.0036
creditApproval	0.5130	0.5568	0.93	0.0438
votes	0.9209	0.9223	0.67	0.0014
fourclass	0.7171	0.7708	0.70	0.0537
breastcancer	0.8688	0.8894	0.74	0.0206
isolet5	0.9962	1	0.71	0.0038
mushrooms	1	0.8956	0.64	-0.1044
svmguide1	0.4696	0.5678	0.78	0.0982
splice	0.5170	0.5896	0.92	0.0726
clean1	0.9957	1	0.70	0.0043
svmguide3	0.4560	0.7404	0.93	0.2844
codrna	0.6614	0.6653	0.65	0.0039

Test accuracy comparison of lightGBM (noise rate 10%)

Datasets	lightGBM	CAGBS-lightGBM	Best purity	Accuracy improvement
diabetes	0.6236	0.6529	0.93	0.0293
sonar	0.4811	0.4752	0.63	-0.0059
creditApproval	0.5258	0.5516	0.97	0.0258
votes	0.9133	0.9064	0.66	-0.0069
fourclass	0.6986	0.7301	0.84	0.0315
breastcancer	0.8288	0.8834	0.73	0.0546
isolet5	0.9936	1	0.95	0.0064
mushrooms	0.9999	0.9122	0.75	-0.0877
svmguide1	0.4743	0.5777	0.76	0.1034
splice	0.5259	0.5984	0.75	0.0725
clean1	0.9998	1	0.64	0.0002
svmguide3	0.4500	0.7225	0.96	0.2725
codrna	0.6403	0.6701	0.92	0.0298

Test accuracy comparison of lightGBM (noise rate 15%)

Datasets	lightGBM	CAGBS-lightGBM	Best purity	Accuracy improvement
diabetes	0.6401	0.6526	0.92	0.0125
sonar	0.4776	0.4877	0.98	0.0101
creditApproval	0.5126	0.5584	0.94	0.0458
votes	0.8936	0.9034	0.62	0.0098
fourclass	0.7059	0.7648	0.62	0.0589
breastcancer	0.8403	0.8922	0.64	0.0519
isolet5	0.9987	1	0.94	0.0013
mushrooms	0.9988	0.9032	0.68	-0.0956
svmguide1	0.4741	0.5716	0.71	0.0975
splice	0.5518	0.5923	0.97	0.0405
clean1	0.9992	1	0.98	0.0008
svmguide3	0.4685	0.7281	0.88	0.2596
codrna	0.6355	0.7021	0.87	0.0666

Test accuracy comparison of lightGBM (noise rate 20%)

Datasets	lightGBM	CAGBS-lightGBM	Best purity	Accuracy improvement
diabetes	0.6308	0.6457	0.81	0.0149
sonar	0.4810	0.4833	0.68	0.0023
creditApproval	0.5010	0.5564	0.99	0.0554
votes	0.8441	0.8805	0.62	0.0364
fourclass	0.6767	0.7191	0.79	0.0424
breastcancer	0.8356	0.8883	0.72	0.0527
isolet5	0.9917	1	0.75	0.0083
mushrooms	0.9968	0.9012	0.62	-0.0956
svmguide1	0.4645	0.5872	0.73	0.1227
splice	0.5322	0.6012	0.64	0.069
clean1	0.9989	1	0.73	0.0011
svmguide3	0.4705	0.7306	0.97	0.2601
codrna	0.6233	0.6855	0.62	0.0622

Test accuracy comparison of lightGBM (noise rate 30%)

Datasets	lightGBM	CAGBS-lightGBM	Best purity	Accuracy improvement
diabetes	0.5975	0.6457	0.74	0.0482
sonar	0.4887	0.4905	0.72	0.0018
creditApproval	0.5158	0.5667	0.99	0.0509
votes	0.7407	0.868	0.61	0.1273
fourclass	0.6487	0.7375	0.63	0.0888
breastcancer	0.7646	0.8859	0.65	0.1213
isolet5	0.9692	1	0.69	0.0308
mushrooms	0.9714	0.9072	0.61	-0.0642
svmguide1	0.4429	0.4840	0.92	0.0411
splice	0.5653	0.5922	0.61	0.0269
clean1	0.9998	0.9799	0.78	-0.0199
svmguide3	0.4371	0.7006	0.97	0.2635
codrna	0.5569	0.6727	0.80	0.1158

Test accuracy comparison of lightGBM (noise rate 40%)

Datasets	lightGBM	CAGBS-lightGBM	Best purity	Accuracy improvement
diabetes	0.4837	0.5578	0.61	0.0741
sonar	0.4840	0.4998	0.76	0.0158
creditApproval	0.5029	0.5442	0.87	0.0413
votes	0.6513	0.7393	0.61	0.0880
fourclass	0.5817	0.6214	0.80	0.0397
breastcancer	0.5923	0.6996	0.72	0.1073
isolet5	0.9416	1	0.77	0.0584
mushrooms	0.8572	0.7454	0.61	-0.1118
svmguide1	0.4026	0.4776	0.94	0.0750
splice	0.5387	0.5409	0.61	0.0022
clean1	1	0.9261	0.61	-0.0739
svmguide3	0.4739	0.6182	0.96	0.1443
codrna	0.5569	0.6101	0.63	0.0532