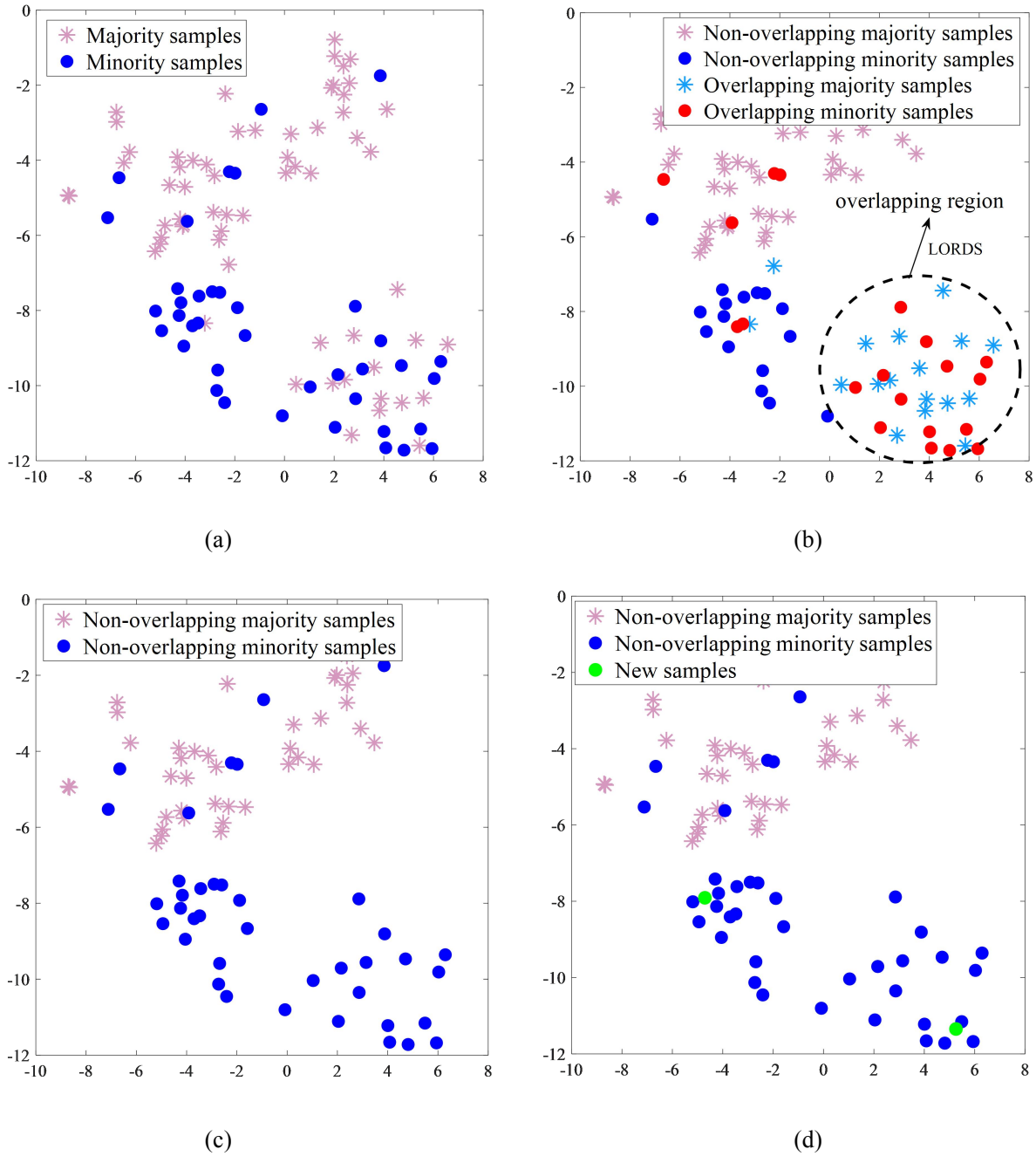


Fig. 6 The schematic sketch of the effect of SHS: (a) represents the original feature distribution of the subset, (b) describes the identification results of LORDS for the overlapping samples.(c) is the subset after undersampling of overlapping majority samples, and (d) depicts the final balanced subset with low overlap after oversampling of minority samples.

Table 3 Classification performance of different algorithms using SVM on the datasets with low and high IR. The superscript m1-m4 represents the methods of oversampling, undersampling, hybrid sampling and ensemble learning (see section4.1.1)

Dataset	Ionosphere				glass1				Monk			
Measure	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC
SMOTE^{m1}	0.7006	0.74710	0.7939	0.8029	0.9093	0.5603	0.4869	0.5875	**0.8471**	0.5772	0.6063	0.6422
KNSMOTE^{m1}	0.66412	0.74112	0.78412	0.79712	0.23511	0.28112	0.40512	0.54112	0.52811	0.42410	0.51110	0.51411
ADPCHFO^{m1}	0.65713	0.7479	0.78611	0.80110	0.13014	0.19014	0.32314	0.52813	0.6449	0.4869	0.5559	0.5659
ENN^{m2}	0.64214	0.73513	0.77713	0.79313	0.22513	0.29811	0.42611	0.54811	0.02414	0.04614	0.15414	0.51212
CBIS^{m2}	0.7622	0.7545	0.8063	0.8094	0.9502	0.5604	0.43310	0.5788	0.6697	0.5247	0.6006	0.6076
AdaOBU^{m2}	0.7404	0.63714	0.70414	0.70714	0.23212	0.23413	0.36313	0.42514	0.54010	0.41511	0.49211	0.50114
SMTL^{m3}	0.7175	0.7537	0.8004	0.8075	0.9054	0.5605	0.4878	0.5876	0.8333	0.5663	0.5957	0.6323
CUSS^{m3}	0.6957	0.7546	0.7967	0.8077	0.60710	0.48110	0.5034	0.5923	0.40412	0.26012	0.24113	0.53710
RFMSE^{m3}	0.68810	0.7593	0.7995	0.8103	0.8327	0.5408	0.5035	0.57210	0.6568	0.4918	0.5588	0.5718
SBE^{m4}	0.67611	0.74210	0.78610	0.79811	0.6109	0.4839	0.5353	0.5779	0.20113	0.19613	0.30812	0.50613
DTE-SBD^{m4}	0.6909	0.7538	0.7968	0.8068	0.8925	0.5576	0.4917	0.5857	0.7804	0.5565	0.6035	0.6255
REMDD^{m4}	0.6918	0.7554	0.7976	0.8076	0.8786	0.5554	0.5026	0.5884	0.7695	0.5556	0.6062	0.6264
EASE^{m4}	**0.8321**	0.7612	0.8132	0.8152	0.7498	0.6542	0.6272	0.6372	0.6746	0.5624	0.6044	0.6057
DCSHS	0.7423	**0.8101**	**0.8341**	**0.8401**	**0.9571**	**0.6111**	**0.5721**	**0.6491**	0.8392	**0.6261**	**0.6841**	**0.7011**
Dataset	haberman				Vehicle1				Vehicle2			
Measure	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC
SMOTE^{m1}	0.4128	0.4323	0.5743	**0.6201**	0.9377	0.8476	0.9192	0.9202	0.7863	**0.6131**	**0.7571**	0.7582
KNSMOTE^{m1}	0.4945	0.4006	0.5596	0.5789	0.82312	0.8544	0.8909	0.8939	0.60513	0.54911	0.69111	0.69811
ADPCHFO^{m1}	0.219	0.309	0.443	0.579	0.916	0.857	0.919	0.919	0.679	0.584	0.724	0.727

	12	12	12	8	8	2	4	4	10	6	7	7
ENN^{m2}	0.140	0.218	0.354	0.550	0.819	0.855	0.889	0.892	0.416	0.507	0.618	0.669
	13	13	13	12	13	3	11	11	14	13	14	12
CBIS^{m2}	0.067	0.119	0.254	0.526	0.911	0.820	0.901	0.901	0.638	0.583	0.717	0.723
	14	14	14	14	9	9	7	7	11	7	9	8
AdaOBU^{m2}	0.585	0.394	0.538	0.545	0.835	0.715	0.829	0.829	0.624	0.497	0.656	0.658
	3	8	9	13	11	13	13	13	12	14	13	14
SMTL^{m3}	0.435	0.426	0.573	0.611	0.937	0.846	0.919	0.920	0.797	0.603	0.749	0.752
	7	4	4	3	6	7	3	3	2	3	3	3
CUSS^{m3}	0.313	0.394	0.525	0.612	0.861	0.849	0.899	0.901	0.687	0.592	0.729	0.734
	11	9	10	2	10	5	8	8	8	4	5	5
RFMSE^{m3}	0.385	0.362	0.512	0.566	0.944	0.772	0.881	0.883	0.682	0.510	0.668	0.669
	9	11	11	10	2	12	12	12	9	12	12	13
SBE^{m4}	0.507	0.437	0.586	0.611	**0.950**	0.806	0.901	0.903	0.728	0.571	0.720	0.721
	4	2	2	4	**1**	10	6	6	6	9	8	9
DTE-SBD^{m4}	0.467	0.414	0.570	0.593	0.941	0.826	0.911	0.911	0.738	0.582	0.729	0.730
	6	5	5	7	4	8	5	5	4	8	6	6
REMDD^{m4}	0.375	0.398	0.541	0.602	0.942	0.803	0.898	0.900	0.693	0.554	0.705	0.706
	10	7	8	5	3	11	9	9	7	10	10	10
EASE^{m4}	0.599	0.388	0.548	0.556	0.761	0.599	0.741	0.742	0.738	0.592	0.737	0.738
	2	10	7	11	14	14	14	14	5	5	4	4
DCSHS	**0.602**	**0.442**	**0.594**	0.601	0.939	**0.878**	**0.935**	**0.935**	**0.892**	0.607	0.752	**0.763**
	1	**1**	**1**	6	5	**1**	**1**	**1**	**1**	2	2	**1**
Dataset	Vehicle3				ecoli1				yeast2			
Measure	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC
SMOTE^{m1}	0.791	0.594	0.747	0.749	0.932	0.767	0.889	0.891	0.706	0.550	0.764	0.768
	3	2	2	2	3	3	4	4	7	4	2	2
KNSMOTE^{m1}	0.613	0.514	0.673	0.678	0.837	0.753	0.858	0.861	0.678	0.525	0.743	0.752
	12	11	11	11	12	8	12	11	10	10	9	9
ADPCHFO^{m1}	0.621	0.557	0.701	0.710	0.820	0.739	0.848	0.850	0.622	**0.578**	0.745	0.759
	10	5	6	6	13	11	13	13	11	**1**	8	7
ENN^{m2}	0.092	0.151	0.278	0.537	0.793	0.733	0.837	0.841	0.362	0.476	0.590	0.666
	14	14	14	14	14	13	14	14	14	13	14	14
CBIS^{m2}	0.656	0.546	0.700	0.704	0.902	0.759	0.878	0.879	0.492	0.549	0.678	0.717
	9	7	7	7	9	6	7	7	12	5	12	11
AdaOBU^{m2}	0.616	0.489	0.654	0.657	0.926	0.733	0.869	0.872	0.728	0.492	0.741	0.742
	11	13	13	13	4	14	10	10	2	11	10	10
SMTL^{m3}	0.800	0.584	0.739	0.742	**0.945**	0.763	0.890	0.892	0.710	0.546	0.764	0.767
	2	3	3	3	**1**	5	3	3	5	7	3	3
CUSS^{m3}	0.561	0.535	0.674	0.693	0.900	0.752	0.873	0.876	0.482	0.544	0.672	0.713
	13	9	10	10	11	10	8	8	13	8	13	13

RFMSE[m3]	0.6698	0.50112	0.66612	0.66812	0.9138	0.73812	0.8709	0.8729	**0.7851**	0.49212	0.7527	0.7548
SBE[m4]	0.7167	0.53310	0.6949	0.6979	0.9177	0.7557	0.8805	0.8815	0.7096	0.5399	0.7605	0.7635
DTE-SBD[m4]	0.7604	0.5566	0.7155	0.7185	0.9442	0.7664	**0.8911**	0.8932	0.7028	0.5486	0.7624	0.7664
REMDD[m4]	0.7236	0.5358	0.6968	0.6998	0.9196	0.7529	0.8796	0.8816	0.6809	0.5513	0.7576	0.7626
EASE[m4]	0.7395	0.5744	0.7264	0.7274	0.90210	0.7682	0.86311	0.86612	0.7124	0.45414	0.71611	0.71712
DCSHS	**0.8901**	**0.6021**	**0.7541**	**0.7661**	0.9235	**0.7771**	0.8902	**0.8941**	0.7223	0.5772	**0.7801**	**0.7841**
Dataset	ecoli2				yeast3				ecoli3			
Measure	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC
SMOTE[m1]	0.9166	0.7206	0.8975	0.8995	0.8758	0.6756	0.8927	0.8937	0.8947	0.5775	0.8734	0.8765
KNSMOTE[m1]	0.76612	0.62413	0.78113	0.80313	0.81510	0.62013	0.85010	0.85710	0.76111	0.5228	0.78711	0.82211
ADPCHFO[m1]	0.86710	0.7322	0.8847	0.8876	0.80111	0.7462	0.8759	0.8799	0.8439	0.6792	0.8792	0.8832
ENN[m2]	0.61214	0.68012	0.76514	0.79014	0.56414	0.6747	0.74414	0.77614	0.37114	0.46113	0.58514	0.67513
CBIS[m2]	0.80011	0.7275	0.8599	0.8629	0.72913	0.7384	0.83913	0.84913	0.82610	**0.6801**	0.8725	0.8774
AdaOBU[m2]	0.9097	0.52714	0.80311	0.81112	0.8209	0.59114	0.84911	0.85012	**0.9711**	0.49611	0.8638	0.8708
SMTL[m3]	0.9203	0.7197	0.8994	0.9004	0.8807	0.6699	0.8936	0.8936	0.9116	0.5726	0.8773	0.8803
CUSS[m3]	0.69513	0.6919	0.80312	0.82011	0.73512	0.7443	0.84212	0.85311	0.65613	0.5894	0.76012	0.80012
RFMSE[m3]	0.9175	0.68111	0.8846	0.8867	0.9135	0.62512	0.8935	0.8945	0.9493	0.49012	0.8539	0.8609
SBE[m4]	**0.9231**	0.7324	0.9042	0.9052	0.9163	0.66210	0.9052	0.9052	0.9494	0.51310	0.8667	0.8717
DTE-SBD[m4]	0.8829	0.6968	0.8638	0.8668	0.9076	0.6698	0.9033	0.9033	0.68912	0.37714	0.60513	0.62714
REMDD[m4]	0.9202	0.7323	0.9033	0.9053	0.9144	0.65111	0.9014	0.9014	0.9572	0.5169	0.8696	0.8746
EASE[m4]	0.8978	0.68610	0.84910	0.85610	**0.9201**	0.6765	0.8798	0.8818	0.8818	0.5547	0.81810	0.82810

DCSHS	0.9204	**0.7411**	**0.9061**	**0.9071**	0.9182	**0.7441**	**0.9251**	**0.9251**	0.9115	0.5923	**0.8821**	**0.8851**
Dataset	yeast2vs4				ecoli067vs35				ecoli0234vs5			
Measure	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC
SMOTE[m1]	0.8337	0.6926	0.8803	0.8843	0.7649	0.6275	0.8317	0.8455	0.8509	0.7538	0.8946	0.9016
KNSMOTE[m1]	0.8099	0.7074	0.8737	0.8787	0.6951	0.5697	0.7791	0.8077	0.8508	0.7855	0.9004	0.9073
ADPCHFO[m1]	0.7410	0.7472	0.8452	0.8572	0.7420	**0.7661**	0.8442	0.8612	0.8450	**0.8401**	0.9052	0.9142
ENN[m2]	0.4934	0.6411	0.6955	0.7455	0.3665	0.5131	0.5911	0.6822	0.7044	0.7776	0.8222	0.8486
CBIS[m2]	0.6273	0.7253	0.7833	0.8093	0.6862	0.7343	0.8133	0.8342	0.7847	0.8032	0.8694	0.8844
AdaOBU[m2]	0.8258	0.6502	0.8689	0.8739	0.8513	0.4755	0.8209	0.8259	0.8507	0.5422	0.8477	0.8527
SMTL[m3]	0.8375	0.6847	0.8812	0.8842	0.7648	0.6374	0.8334	0.8464	0.8552	0.7529	0.8975	0.9035
CUSS[m3]	0.6412	0.7075	0.7851	0.8131	0.4913	0.5796	0.6766	0.7386	0.7541	0.7923	0.8507	0.8717
RFMSE[m3]	0.8434	0.6622	0.8784	0.8824	0.8592	0.4644	0.8162	0.8222	0.8506	0.6811	0.8811	0.8871
SBE[m4]	0.8443	0.6555	0.8785	0.8815	0.8285	0.5599	0.8423	0.8483	0.8554	0.7104	0.8897	0.8958
DTE-SBD[m4]	0.7251	0.5792	0.7433	0.7533	0.8027	0.5618	0.8316	0.8396	0.8603	0.7657	0.9013	0.9064
REMDD[m4]	0.8356	0.6749	0.8786	0.8816	0.8256	0.5377	0.8335	0.8395	0.8406	0.7436	0.8898	0.8958
EASE[m4]	**0.9341**	0.6808	0.8708	0.8788	**0.8631**	0.5311	0.8308	0.8378	0.8632	0.6032	0.8892	0.8949
DCSHS	0.8672	**0.7651**	**0.9111**	**0.9121**	0.8334	0.7142	**0.8901**	**0.8921**	**0.8951**	0.7914	**0.9211**	**0.9261**
Dataset	yeast0359vs78				yeast02579vs368				ecoli01vs235			
Measure	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC
SMOTE[m1]	0.7143	0.3693	0.7332	0.7392	0.8717	0.7228	0.9045	0.9065	0.8336	0.7197	0.8842	0.8902
KNSMOTE[m1]	0.5549	0.2921	0.6188	0.6547	0.8261	0.7566	0.8899	0.8939	0.7649	0.7416	0.8518	0.8657
ADPCHFO[m1]	0.3051	0.3476	0.5222	0.6309	0.8539	0.7974	0.9082	0.9112	0.7989	**0.8061**	0.8804	0.8893

ENN[m2]	0.227	0.344	0.462	0.610	0.777	**0.832**	0.876	0.884	0.642	0.767	0.792	0.821
	14	9	13	12	14	**1**	12	12	14	5	13	13
CBIS[m2]	0.227	0.344	0.462	0.610	0.788	0.808	0.879	0.885	0.678	0.788	0.814	0.838
	13	8	12	11	13	2	11	11	12	3	12	11
AdaOBU[m2]	0.239	0.322	0.460	0.600	0.894	0.602	0.886	0.887	0.829	0.554	0.845	0.850
	11	11	14	13	3	14	10	10	7	12	9	9
SMTL[m3]	0.704	0.364	0.727	0.734	0.870	0.734	0.906	0.907	0.833	0.708	0.882	0.888
	5	4	3	3	8	7	4	3	5	9	3	4
CUSS[m3]	0.227	0.344	0.462	0.610	0.813	0.806	0.890	0.895	0.703	0.768	0.824	0.846
	12	7	11	10	12	3	8	7	11	4	10	10
RFMSE[m3]	**0.802**	0.262	0.618	0.649	0.886	0.642	0.893	0.894	0.874	0.548	0.856	0.862
	1	13	7	8	5	12	7	8	2	13	7	8
SBE[m4]	0.696	0.335	0.700	0.710	0.893	0.684	0.906	0.907	0.837	0.669	0.875	0.881
	6	10	5	5	4	10	3	4	4	10	5	5
DTE-SBD[m4]	0.704	0.354	0.721	0.727	0.841	0.680	0.867	0.869	0.666	0.506	0.657	0.683
	4	5	4	4	10	11	13	13	13	14	14	14
REMDD[m4]	0.560	0.370	0.669	0.697	0.876	0.690	0.900	0.901	0.811	0.718	0.872	0.880
	8	2	6	6	6	9	6	6	8	8	6	6
EASE[m4]	0.592	0.220	0.576	0.593	0.903	0.614	0.862	0.865	**0.919**	0.610	0.822	0.832
	7	14	9	14	2	13	14	14	**1**	11	11	12
DCSHS	0.714	**0.377**	**0.751**	**0.752**	**0.938**	0.760	**0.937**	**0.937**	0.857	0.800	**0.912**	**0.913**
	2	**1**	**1**	**1**	**1**	5	**1**	**1**	3	2	**1**	**1**

Dataset	glass06vs5				ecoli01vs5				glass0146vs2			
Measure	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC
SMOTE[m1]	**1.000**	0.647	**0.938**	**0.941**	0.850	0.748	0.898	0.905	0.932	0.198	0.537	0.624
	2	3	**1**	**1**	7	9	3	4	5	10	3	4
KNSMOTE[m1]	0.663	0.551	0.761	0.796	0.770	0.788	0.862	0.877	0.844	0.200	0.535	0.612
	11	7	8	8	11	5	9	9	7	8	5	13
ADPCHFO[m1]	0.707	0.601	0.791	0.824	0.830	0.813	0.897	0.905	0.230	**0.373**	0.479	0.615
	8	5	7	6	9	2	5	3	14	**5**	14	11
ENN[m2]	0.377	0.540	0.609	0.688	0.685	0.790	0.817	0.842	0.230	**0.373**	0.479	0.615
	14	9	14	14	14	4	14	14	13	**4**	13	10
CBIS[m2]	0.607	0.487	0.716	0.753	0.720	**0.815**	0.839	0.859	0.230	**0.373**	0.479	0.615
	12	10	9	9	12	**1**	12	11	12	**3**	12	9
AdaOBU[m2]	0.673	0.327	0.699	0.719	0.810	0.645	0.861	0.870	0.230	**0.373**	0.479	0.615
	9	14	12	12	10	12	10	10	11	**2**	11	8
SMTL[m3]	**1.000**	0.646	0.937	0.940	0.850	0.740	0.897	0.903	0.937	0.199	0.536	0.625
	1	4	2	2	6	10	6	6	4	9	4	3
CUSS[m3]	0.387	0.547	0.615	0.693	0.715	0.782	0.831	0.854	0.230	**0.373**	0.479	0.615
	13	8	13	13	13	6	13	13	10	**1**	10	7

RFMSE[m3]	0.8336	0.42311	0.7966	0.8117	0.8505	0.61313	0.8728	0.8788	0.9373	0.18714	0.4909	0.60014
SBE[m4]	0.67010	0.37113	0.71510	0.73911	0.8504	0.72711	0.8957	0.9017	0.9076	0.19013	0.4967	0.61312
DTE-SBD[m4]	0.9434	0.5886	0.8954	0.9044	0.8503	0.7677	0.9012	0.9072	0.9632	0.19611	0.5146	0.6225
REMDD[m4]	0.7177	0.39112	0.71411	0.74810	0.8458	0.7648	0.8984	0.9045	**0.987**1	0.19412	0.4948	0.6206
EASE[m4]	0.9523	0.7172	0.8855	0.8915	**0.879**1	0.55414	0.84111	0.85912	0.6439	0.2277	0.5982	0.6262
DCSHS	0.9175	**0.803**1	0.9313	0.9393	0.8552	0.7983	**0.907**1	**0.914**1	0.7148	0.2946	**0.655**1	**0.657**1

Dataset	glass2				ecoli0146vs5				yeast1vs7			
Measure	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC
SMOTE[m1]	0.8506	0.18611	0.5472	0.61211	0.8407	0.65810	0.8846	0.8935	0.7502	0.3068	0.7582	0.7644
KNSMOTE[m1]	0.8427	0.18910	0.5413	0.6149	0.78011	0.7206	0.8609	0.8749	0.7276	0.3292	0.7563	0.7672
ADPCHFO[m1]	0.23014	**0.373**5	0.47914	0.6158	0.80010	**0.771**1	0.8787	0.8897	0.29411	0.3146	0.51711	0.62611
ENN[m2]	0.23013	**0.373**4	0.47913	0.6157	0.61914	0.7335	0.77514	0.80814	0.14313	0.25013	0.37814	0.57114
CBIS[m2]	0.23012	**0.373**3	0.47912	0.6156	0.66413	0.7672	0.80312	0.83112	0.14314	0.25012	0.37813	0.57113
AdaOBU[m2]	0.23011	**0.373**2	0.47911	0.6155	0.8159	0.51113	0.84510	0.85110	0.34510	**0.332**1	0.54710	0.65010
SMTL[m3]	0.8505	0.18413	0.5404	0.60814	0.8456	0.6747	0.8902	0.8962	**0.763**1	0.3097	**0.763**1	**0.769**1
CUSS[m3]	0.23010	**0.373**1	0.47910	0.6154	0.67912	0.7504	0.81011	0.83511	0.14312	0.25011	0.37812	0.57112
RFMSE[m3]	0.9373	0.18314	0.5069	0.61013	0.8503	0.56412	0.8708	0.8758	0.7404	0.27210	0.7298	0.7378
SBE[m4]	0.9952	0.1919	0.5128	0.6323	0.8502	0.62611	0.8855	0.8906	0.7109	0.3253	0.7495	0.7595
DTE-SBD[m4]	0.9084	0.18512	0.5236	0.61310	0.8455	0.6728	0.8893	0.8953	0.7403	0.3009	0.7417	0.7537
REMDD[m4]	**1.000**1	0.1928	0.5157	0.6352	0.8454	0.6649	0.8894	0.8944	0.7108	0.3205	0.7476	0.7566
EASE[m4]	0.6828	0.2527	**0.649**1	**0.664**1	0.8228	0.42514	0.78913	0.80913	0.7335	0.24014	0.6819	0.6959

DCSHS	0.3509	0.2646	0.5325	0.61112	**0.8501**	0.7523	**0.9021**	**0.9081**	0.7277	0.3234	0.7564	0.7653
Dataset	glass4				ecoli4				zoo3			
Measure	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC
SMOTE^{m1}	0.8775	0.5473	**0.8851**	**0.8961**	0.8758	0.7514	0.9175	0.9236	0.8008	0.7874	0.8762	0.8922
KNSMOTE^{m1}	0.8377	0.5454	0.8644	0.8793	0.70412	0.7247	0.81912	0.84612	0.7909	0.7807	0.8705	0.8875
ADPCHFO^{m1}	0.74211	0.5582	0.8228	0.8428	0.8659	0.8023	0.9176	0.9235	0.74013	0.7903	0.8458	0.8677
ENN^{m2}	0.33813	0.4888	0.57413	0.66813	0.45014	0.60112	0.65814	0.72514	0.73014	0.7599	0.83610	0.85810
CBIS^{m2}	0.31714	0.46910	0.55914	0.65714	0.83010	0.8512	0.9039	0.9129	0.77010	0.7836	0.8606	0.8796
AdaOBU^{m2}	0.74710	0.32114	0.76611	0.77911	0.80011	0.6808	0.87111	0.88211	**1.0001**	0.15814	0.64913	0.71413
SMTL^{m3}	0.8774	0.5415	0.8842	0.8952	0.8757	0.7476	0.9167	0.9227	0.8007	0.7658	0.8734	0.8894
CUSS^{m3}	0.39312	0.4849	0.60612	0.69112	0.47013	0.60311	0.66613	0.73513	0.74012	0.7932	0.8457	0.8677
RFMSE^{m3}	**0.9171**	0.38012	0.8566	0.8636	0.9504	0.53914	0.9138	0.9188	0.8006	0.75710	0.8733	0.8903
SBE^{m4}	0.9032	0.38311	0.8517	0.8597	**1.0002**	0.56513	0.9464	0.9484	0.8005	0.28912	0.74712	0.76312
DTE-SBD^{m4}	0.8576	0.5026	0.8643	0.8745	0.9456	0.7495	0.9503	0.9523	0.8004	**0.8001**	**0.8761**	**0.8931**
REMDD^{m4}	0.8109	0.35513	0.77010	0.78610	0.9903	0.61810	0.9512	0.9522	0.8202	0.38411	0.77511	0.79611
EASE^{m4}	0.8983	0.4937	0.7879	0.8269	0.9485	0.6429	0.89110	0.89910	0.8033	0.22713	0.53514	0.70214
DCSHS	0.8208	**0.5691**	0.8625	0.8754	**1.0001**	**0.8751**	**0.9891**	**0.9891**	0.74011	0.7875	0.8449	0.8669
Dataset	lymphography-normal-fibrosis				winequality-red4				poker9vs7			
Measure	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC
SMOTE^{m1}	0.8078	0.7964	0.8827	0.8987	0.6497	0.1347	0.6802	0.6872	0.5878	0.2388	0.6827	0.7156
KNSMOTE^{m1}	0.75011	0.7788	0.84810	0.87111	0.2129	0.05414	0.3899	0.49514	0.58310	0.2477	0.6856	0.7185
ADPCHFO^{m1}	0.75010	0.7709	0.84811	0.87210	0.10110	0.1575	0.31210	0.54810	0.5877	0.6352	0.7432	0.7842

DCSHS 0.3509 0.2646 0.5325 0.61112 **0.8501** 0.7523 **0.9021** **0.9081** 0.7277 0.3234 0.7564 0.7653

ENN[m2]	0.737 14	0.780 7	0.841 14	0.866 14	0.086 14	**0.159 4**	0.294 14	0.543 13	0.407 14	0.573 5	0.634 13	0.703 12
CBIS[m2]	0.740 13	0.783 6	0.843 13	0.868 13	0.086 13	**0.159 3**	0.294 13	0.543 12	0.407 13	0.573 4	0.634 12	0.703 11
AdaOBU[m2]	0.817 6	0.648 12	0.872 9	0.888 9	0.086 12	**0.159 2**	0.294 12	0.543 11	**0.747 1**	0.164 11	0.707 3	0.727 4
SMTL[m3]	0.873 4	**0.850 1**	0.921 3	0.933 2	0.650 6	0.133 8	0.679 3	0.686 3	0.583 9	0.231 9	0.678 9	0.711 8
CUSS[m3]	0.740 12	0.783 5	0.843 12	0.868 12	0.086 11	**0.159 1**	0.294 11	0.543 10	0.407 12	0.573 3	0.634 11	0.703 10
RFMSE[m3]	0.797 9	0.768 10	0.872 8	0.891 8	**0.715 1**	0.104 12	0.645 7	0.651 7	0.733 4	0.138 13	0.673 10	0.698 13
SBE[m4]	0.977 2	0.585 14	**0.950 1**	**0.954 1**	0.696 3	0.118 11	0.672 5	0.676 5	0.737 3	0.145 12	0.680 8	0.705 9
DTE-SBD[m4]	0.857 5	0.819 2	0.910 4	0.923 5	0.655 5	0.132 9	0.679 4	0.685 4	0.580 11	0.324 6	0.699 4	0.734 3
REMDD[m4]	0.917 3	0.605 13	0.922 2	0.930 3	0.692 4	0.119 10	0.671 6	0.674 6	0.713 5	0.203 10	0.686 5	0.712 7
EASE[m4]	**0.979 1**	0.747 11	0.887 5	0.929 4	0.713 2	0.103 13	0.589 8	0.613 8	0.628 6	0.071 14	0.505 14	0.604 14
DCSHS	0.810 7	0.810 3	0.883 6	0.901 6	0.637 8	0.141 6	**0.683 1**	**0.692 1**	0.743 2	**0.717 1**	**0.841 1**	**0.866 1**

Dataset	abalone3vs11				winequalitywhite9vs4				yeast6			
Measure	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC
SMOTE[m1]	**1.000 10**	**0.971 5**	**0.999 6**	**0.999 6**	**0.900 6**	0.635 7	0.922 3	0.930 3	0.854 7	0.305 5	0.879 4	0.881 4
KNSMOTE[m1]	**1.000 9**	**0.971 4**	**0.999 5**	**0.999 5**	0.550 10	0.519 9	0.722 10	0.755 10	0.846 9	0.329 3	0.879 3	0.882 3
ADPCHFO[m1]	**1.000 8**	**0.971 3**	**0.999 4**	**0.999 4**	**0.900 5**	0.709 2	**0.928 2**	**0.936 1**	0.729 10	**0.422 1**	0.829 9	0.843 9
ENN[m2]	0.933 14	0.954 10	0.960 14	0.967 14	0.500 14	0.667 6	0.707 14	0.750 14	0.125 14	0.222 13	0.354 14	0.563 14
CBIS[m2]	0.933 13	0.954 9	0.960 13	0.967 13	0.500 13	0.667 5	0.707 13	0.750 13	0.125 13	0.222 12	0.354 13	0.563 13
AdaOBU[m2]	**1.000 7**	0.666 14	0.983 11	0.983 11	0.500 12	0.667 4	0.707 12	0.750 12	**0.886 2**	0.270 6	0.885 2	0.886 2
SMTL[m3]	**1.000 6**	0.969 6	**0.999 3**	**0.999 3**	0.890 8	0.610 8	0.914 4	0.923 4	0.851 8	0.308 4	0.878 5	0.881 5
CUSS[m3]	0.933 12	0.954 8	0.960 12	0.967 12	0.500 11	0.667 3	0.707 11	0.750 11	0.143 12	0.228 10	0.366 12	0.571 12

RFMSE[m3]	**1.000**	0.959	0.998	0.998	**0.900**	0.480	0.904	0.912	0.877	0.224	0.864	0.865
	5	7	7	7	**4**	11	5	5	5	11	8	8
SBE[m4]	**1.000**	0.768	0.990	0.990	**0.900**	0.230	0.832	0.841	0.880	0.252	0.877	0.878
	4	13	10	10	**3**	14	8	8	4	8	7	7
DTE-SBD[m4]	**1.000**	0.947	0.998	0.998	0.890	0.515	0.901	0.909	0.708	0.215	0.694	0.709
	3	11	8	8	7	10	6	6	11	14	11	11
REMDD[m4]	**1.000**	0.773	0.990	0.990	**0.900**	0.265	0.842	0.852	0.883	0.253	0.878	0.879
	2	12	9	9	**2**	13	7	7	3	7	6	6
EASE[m4]	0.998	0.971	**0.999**	**0.999**	**0.785**	0.271	0.724	0.793	**0.886**	0.242	0.809	0.817
	11	2	**2**	**2**	**9**	12	9	9	**1**	9	10	10
DCSHS	**1.000**	**0.971**	**0.999**	**0.999**	**0.900**	**0.714**	**0.928**	0.935	0.867	0.413	**0.892**	**0.893**
	1	**1**	**1**	**1**	**1**	**1**	**1**	2	6	2	**1**	**1**

Dataset	winequalitywhite39vs5				shuttle2vs5				poker8vs6			
Measure	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC	Rec	F1-M	G-M	AUC
SMOTE[m1]	**0.568**	0.086	**0.663**	**0.685**	**1.000**	**1.000**	**1.000**	**1.000**	0.401	0.023	0.441	0.481
	1	9	**1**	**1**	**7**	**6**	**6**	**6**	6	12	11	12
KNSMOTE[m1]	0.316	0.056	0.492	0.568	0.986	0.992	0.993	0.993	0.346	0.027	0.474	0.518
	10	13	10	13	8	8	8	8	8	7	9	10
ADPCHFO[m1]	0.190	0.267	0.431	0.592	**1.000**	**1.000**	**1.000**	**1.000**	0.265	0.274	0.506	0.624
	11	4	11	9	**6**	**5**	**5**	**5**	10	6	3	2
ENN[m2]	0.167	0.286	0.408	0.583	0.740	0.839	0.857	0.870	0.230	0.373	0.479	0.615
	14	3	14	12	11	11	11	11	14	4	6	5
CBIS[m2]	0.167	0.286	0.408	0.583	0.738	0.838	0.856	0.869	0.230	0.373	0.479	0.615
	13	2	13	11	12	12	12	12	13	3	5	4
AdaOBU[m2]	0.365	0.083	0.544	0.607	0.734	0.824	0.853	0.867	0.230	0.323	0.474	0.605
	9	11	9	8	13	13	13	13	12	5	8	6
SMTL[m3]	0.560	0.085	0.658	0.681	**1.000**	**1.000**	**1.000**	**1.000**	0.412	0.023	0.446	0.484
	2	10	2	2	**5**	**4**	**4**	**4**	5	13	10	11
CUSS[m3]	0.167	**0.286**	0.408	0.583	0.740	0.839	0.857	0.870	0.230	0.373	0.479	0.615
	12	**1**	12	10	10	10	10	10	11	2	4	3
RFMSE[m3]	0.530	0.073	0.619	0.643	**1.000**	0.996	**1.000**	**1.000**	0.691	0.024	0.475	0.535
	5	12	6	7	**4**	7	**7**	**7**	2	10	7	7
SBE[m4]	0.519	0.092	0.642	0.676	0.756	0.845	0.867	0.878	0.687	0.021	0.355	0.481
	6	6	3	4	9	9	9	9	3	14	14	13
DTE-SBD[m4]	0.539	0.088	0.631	0.660	**1.000**	**1.000**	**1.000**	**1.000**	0.385	0.023	0.439	0.471
	4	8	5	6	**3**	**3**	**3**	**3**	7	11	12	14
REMDD[m4]	0.515	0.091	0.635	0.669	0.639	0.695	0.712	0.729	**0.733**	0.026	0.406	0.528
	7	7	4	5	14	14	14	14	**1**	8	13	9
EASE[m4]	0.555	0.042	0.553	0.566	**1.000**	**1.000**	**1.000**	**1.000**	0.462	0.025	0.514	0.533
	3	14	8	14	**2**	**2**	**2**	**2**	4	9	2	8

DCSHS	0.4398	0.1525	0.6167	0.6803	**1.0001**	**1.0001**	**1.0001**	**1.0001**	0.3169	**0.4461**	**0.5501**	**0.6581**

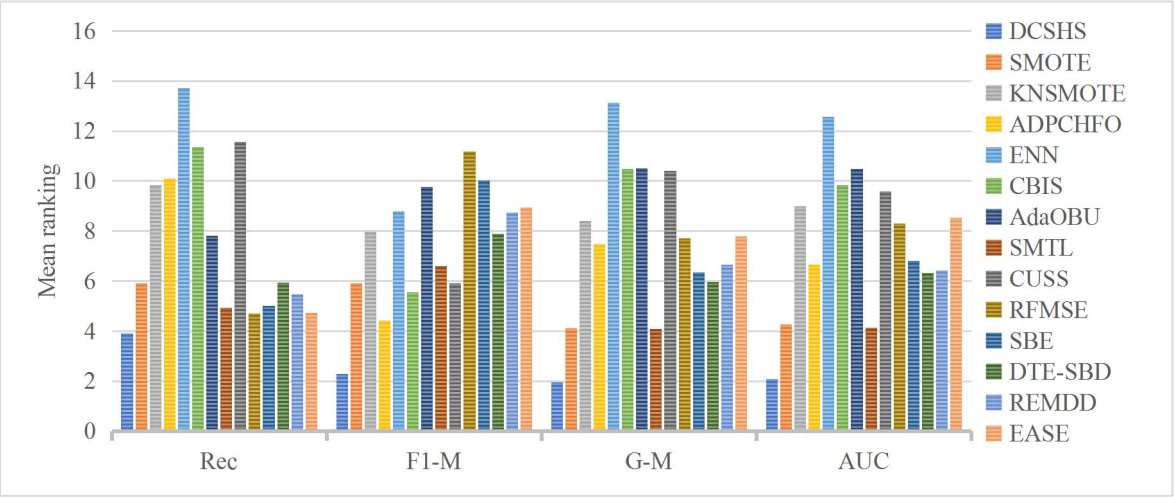


Fig. 7 Mean ranking of four evaluation criteria by the algorithms on all datasets

Table 4 The *p*-value of Holm test with DCSHS as control method.

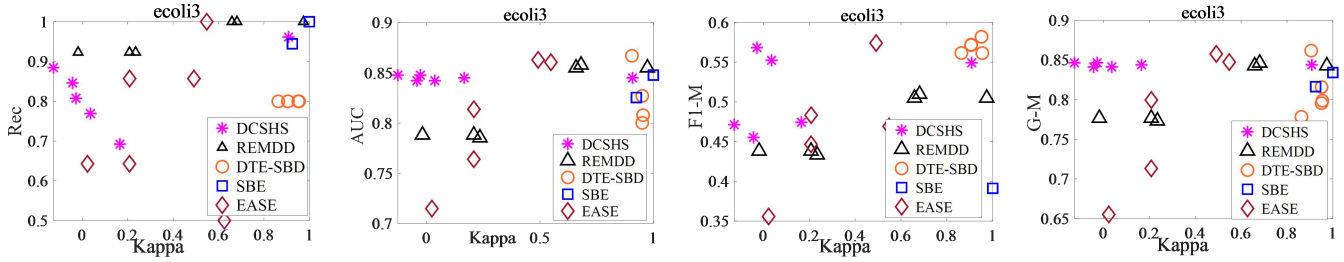
Algorithms	Rec	F1-M	G-M	AUC
SMOTE	1.42E-03	3.97E-06	1.30E-03	2.21E-03
KNSMOTE	9.78E-20	1.16E-12	2.53E-20	1.85E-20
ADPCHFO	2.49E-21	6.32E-02	1.47E-15	3.11E-10
ENN	1.91E-45	7.89E-16	6.35E-50	1.04E-40
CBIS	4.83E-29	3.13E-05	2.92E-32	8.69E-25
AdaOBU	9.47E-10	3.59E-20	1.96E-32	2.65E-28
SMTL	9.98E-02	4.53E-08	1.50E-03	4.12E-03
CUSS	2.59E-30	3.97E-06	6.48E-32	1.86E-23
RFMSE	2.13E-01	7.42E-27	1.98E-14	4.22E-17
SBE	5.42E-01	2.54E-21	1.13E-10	9.41E-11
DTE-SBD	1.22E-03	2.36E-12	3.59E-09	4.86E-09
REMDD	1.29E-02	1.03E-15	6.21E-12	2.48E-09
EASE	1.97E-01	1.59E-16	4.83E-17	3.78E-18

Table 5 Classification performance of different algorithms using SVM on the datasets with higher IR

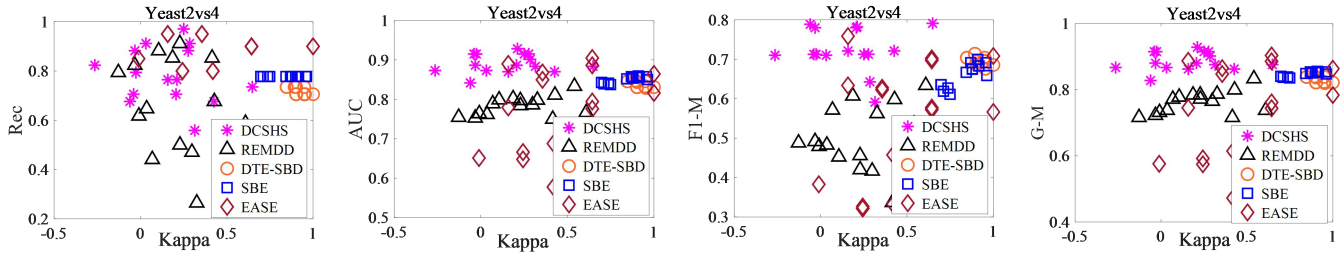
Dataset	Measure	SBE	DTE-SBD	REMDD	EASE	DCSHS
kddrootkitback	Rec	0.833	0.778	0.6	**1**	0.900
	F1-M	0.833	0.875	0.750	0.911	**0.947**
	G-M	0.912	0.882	0.775	0.920	**0.949**
	AUC	0.916	0.889	0.800	0.930	**0.950**
Abalone19	Rec	0.615	0.857	0.875	0.690	**0.909**
	F1-M	0.026	0.037	0.020	0.038	**0.058**
	G-M	0.565	0.802	0.634	0.657	**0.819**
	AUC	0.567	0.804	0.667	0.690	**0.824**
cod	Rec	**1**	**1**	**1**	0.995	**1**
	F1-M	0.057	0.467	0.099	0.341	**0.500**
	G-M	0.980	**0.999**	0.983	0.955	**0.999**
	AUC	0.981	**0.999**	0.983	0.958	**0.999**

Table 6 Diversity analysis of base classifier

Dataset	Indicators	DCSHS	SBE	DTE-SBD	REMDD	EASE
Ecoli3	*dis*	**0.290**	0.015	0.017	0.231	0.203
	ς	**0.157**	0.962	0.923	0.459	0.377
	Q-statistic	**0.177**	0.999	0.999	0.636	0.712
	κ	**0.152**	0.963	0.922	0.456	0.351
Yeast2vs4	*dis*	**0.203**	0.029	0.01	0.285	0.117
	ς	**0.177**	0.850	0.925	0.233	0.501
	Q-statistic	**0.359**	0.998	0.999	0.453	0.866
	κ	**0.159**	0.840	0.923	0.220	0.461



(a)



(b)

Fig.8 Diversity and performance analysis of base classifiers :(a) diversity and performance analysis for Ecoli3, (b) diversity and performance analysis for Yeast2vs4