Appendix A

Table A1 of Appendix A shows the summary of the rest of the datasets. The results given in Table A2 of Appendix A show that OTEC(oob) in terms of sensitivity give promising results ranging from 93.36% to 100% for a number of datasets. This shows the effectiveness of the proposed method. On the other hand, OTEC(sub) gives better results in terms of sensitivity for 14 datasets. Moreover, RF(smote) and RF(over) demonstrate high sensitivity on kc2 and Glass classification. Table A3 of Appendix A shows the results of the proposed method, i.e., OTEC(oob) and OTEC(sub) in terms of specificity, that consistently outperform other methods for the majority of the datasets. OTEC(oob) shows high values of specificity, ranging from 94.13% to 100%, and for OTEC(sub) 97.86% to 100%. Moreover, OTE, RF(smote), RF(over), k-NN, SVM, and Tree outperform the others in terms of specificity for one or 2 datasets. Tables A4-A5 of Appendix A provide insights into the proposed method, i.e., OTEC(oob) and OTEC(sub) in terms of recall and F1 score. The box plots reveal the best findings of the proposed method, i.e., OTEC(oob) and OTEC(sub) in terms of classification error rate for the rest of the data sets are given in Figures A1-A2 and for the rest of the methods in term of sensitivity, specificity recall, and F1 score are given in Figures A3-A14 of Appendix A.

Table A1. Summary of the datasets with class-imbalance problem.

	Dataset	Instances	Features	Class-based Distribution	Imbalance ratio n^1/n^0	Source
8	Madelon	1358	500	1300/58	(22.413:1)	https://openml.org/search?type=data&status=active&sort=runs&id=1485
9	Turing binary	6384	20	6260/124	(50.483:1)	https://www.openml.org/search?type=data&status=active&id=44269
10	KDD	2566	35	2515/51	(49.313:1)	https://openml.org/search?type=data&status=active&id=45075
11	Liver disorder	220	6	200/20	(10:1)	https://openml.org/search?type=data&status=active&id=8
12	Wine	143	13	106/37	(2.864:1)	https://archive.ics.uci.edu/dataset/109/wine
13	Soy bean	167	8	160/7	(22.857:1)	https://archive.ics.uci.edu/dataset/913/forty+soybean+cultivars+from+subsequent+harvests
14	Ionosphere	350	32	312/38	(8.210:1)	https://archive.ics.uci.edu/dataset/52/ionosphere
15	Room Occupancy	8407	14	8228/179	(45.966:1)	https://archive.ics.uci.edu/dataset/864/room+occupancy+estimation
16	Harth	7269	7	6771/498	(13.596:1)	https://archive.ics.uci.edu/dataset/779/harth
17	Rocket League	3015	6	2830/185	15.297:1	https://archive.ics.uci.edu/dataset/858/rocket+league+skillshots
18	Sirtuin6	54	6	50/4	(12.5:1)	https://archive.ics.uci.edu/dataset/748/sirtuin6+small+molecules-1
19	Toxicity	123	12	115/8	(14.375:1)	https://archive.ics.uci.edu/dataset/728/toxicity-2
20	Dry bean	4357	16	4246/129	(32.914:1)	https://archive.ics.uci.edu/dataset/602/dry+bean+dataset

Table A2. The proposed methods, i.e., OTEC(oob), OTEC(sub) and other state-of-the-art methods, in terms of sensitivity.

Dataset	**OTEC(*oob*)**	**OTEC(*sub*)**	**OTE**	**RF(*smote*)**	**RF(*over*)**	**RF(*under*)**	***k*-NN**	**SVM**	**ANN**	**Tree**
Breast Cancer	**1.0000**	0.9629	0.9364	0.9670	0.9795	0.9517	0.9696	0.9353	0.7200	0.9029
Credit Card	**0.9993**	**0.9992**	0.6503	0.9923	0.9241	0.9505	0.9860	0.9889	0.6123	0.7284
Drug Classification	**1.0000**	0.9378	0.8655	0.8277	0.9714	0.9361	0.6401	0.8328	0.2330	0.9031
Kc2	0.9267	0.7951	0.4658	**0.9572**	**0.9572**	0.6335	0.4017	0.1900	0.1407	0.5908
Eeg eye	**0.9929**	**1**	0.7578	0.9216	0.9576	0.9049	0.9327	0.0700	0.6646	0.7257
Glass Classification	0.9153	0.9022	0.8050	**0.9739**	0.8984	0.8166	0.6440	0.2400	0.2665	0.7204
Pc4	**0.9980**	**0.9971**	0.9755	0.9890	0.9885	0.9059	0.8811	0.9207	0	0.9287
Madelon	**0.9554**	**0.9563**	0.4063	0.6508	0.8614	0.6449	0.1750	0.0382	0	0.7171
Turing binary	**0.9796**	**0.9799**	0.0162	0.6648	0.9757	0.4800	0.2519	0.2176	0.0541	0.0140
KDD	**1**	**1**	0.9987	0.8649	0.9896	0.2506	0.9802	0.9794	0	0.9830
Liver disorder	**0.9860**	0.9830	0.9835	0.9713	0.9835	0.9134	0.9206	0.9135	0.1604	0.9365
Wine	**0.9584**	**0.9720**	0.8302	0.9111	0.8944	0.8921	0.7678	0.9177	0.0041	0.7702
Soy bean	**0.9888**	**0.9864**	0.4845	0.9309	0.6332	0.5975	0.0286	0.0288	0	0.3968
Ionosphere	**0.9733**	0.6763	0.9292	0.7239	0.9589	0.9209	0.8954	0.9138	0	0.9397
Room Occupancy	**1**	**1**	0.9971	0.9962	0.9971	0.9936	0.9943	1	0	0.9948
Harth	**0.9980**	**0.9975**	0.9750	0.7336	0.9828	0.9190	0.9800	0.9317	0	0.9749
Rocket League	**0.9336**	**0.9362**	0.0759	0.5431	0.3916	0.0890	0.1240	0	0	0
Sirtuin6	**0.9554**	**0.9589**	0.6373	0.7638	0.9509	0.7758	0.9200	0.9235	0	0.9182
Toxicity	**0.9323**	**0.9314**	0.0386	0.5925	0.4527	0.1049	0.0100	0	0	0
Dry bean	**0.9961**	**0.9952**	0.8161	0.9503	0.9639	0.8477	0.0353	0.9283	0	0.8880

Table A3. The proposed methods, i.e., OTEC(oob), OTEC(sub) and other state-of-the-art methods, in terms of specificity.

Dataset	**OTEC(*oob*)**	**OTEC(*sub*)**	**OTE**	**RF(*smote*)**	**RF(*over*)**	**RF(*under*)**	***k*-NN**	**SVM**	**ANN**	**Tree**
Breast Cancer	**0.9977**	0.9746	0.9711	0.9804	0.9913	0.9635	0.8768	0.9682	0.6215	0.9446
Credit Card	**0.9998**	***0.9998***	***0.9998***	0.9985	0.9996	0.9938	0.7885	0.9970	0.9989	0.9992
Drug Classification	**0.9941**	0.9696	0.9644	0.8058	0.9709	0.9375	0.6579	0.8220	0.6882	0.9443
Kc2	0.8823	0.9978	0.9191	**0.9996**	0.9572	0.8404	0.9345	0.7000	0.8177	0.8785
Eeg eye	**0.9999**	***1.0000***	0.9998	0.8915	0.9406	0.9334	0.9542	0.9700	0.7993	0.6966
Glass Classification	0.9036	0.9039	0.8982	0.9492	0.9718	0.8841	0.8656	**0.9800**	0.7396	0.8371
Pc4	**0.9669**	0.9604	0.4013	0.9645	0.9104	0.7624	0.2103	0.6461	0.0445	0.5508
Madelon	**1.0000**	***1.0000***	0.9197	0.6508	0.8596	0.6509	0.9591	0.6373	0.9565	0.8013
Turing binary	**1.0000**	***0.9999***	0.9962	0.7037	0.9490	0.8174	0.8721	0.8737	0.9815	0.8699
KDD	0.9803	0.9798	0.2393	0.8568	**0.9942**	0.9056	0.1289	0.1582	0.0196	0.6431
Liver disorder	0.9338	0.9264	0.2349	**0.9372**	0.8929	0.6402	0.3683	0.2322	0.1582	0.3683
Wine	**0.9966**	***0.9940***	0.9536	0.8912	0.9757	0.9665	0.8608	0.9704	0.7479	0.9216
Soy bean	**0.9922**	***0.9937***	0.8841	0.0706	0.9889	0.9364	0.9586	0.9580	0.9572	0.9765
Ionosphere	**0.9413**	***0.9988***	0.5066	0.2830	0.6602	0.5903	0.2519	0.8307	0.1087	0.5031
Room Occupancy	**0.9999**	***1***	**0.9999**	0.0034	0.9998	0.9992	***1***	**0.9999**	0.9788	**0.9999**
Harth	**0.9775**	***0.9786***	0.6987	0.2698	0.8552	0.7579	0.8854	0	0.0681	0.8709
Rocket League	**0.9995**	***0.9976***	0.9505	0.4525	0.9580	0.9407	0.9396	0.9383	0.9385	0.9386
Sirtuin6	**0.9428**	***0.9394***	0.0900	0.2680	0.6491	0.2142	*0*	0	0.3481	0
Toxicity	**0.9994**	***0.9986***	0.9074	0.3828	0.9502	0.8546	0.9369	0.9429	0.9305	0.9313
Dry bean	**0.9979**	***0.9977***	0.9952	0.0500	***0.9977***	0.9928	0.9702	0.9911	0.9704	0.9938

Table A4. The proposed methods, i.e., OTEC(oob), OTEC(sub) and other state-of-the-art methods, in terms of recall.

Dataset	**OTEC(*oob*)**	**OTEC(*sub*)**	**OTE**	**RF(*smote*)**	**RF(*over*)**	**RF(*under*)**	***k*-NN**	**SVM**	**ANN**	**Tree**
Breast Cancer	0.9946	0.9733	0.9498	0.9785	**0.9952**	0.9495	0.8483	0.9657	0.2915	0.9051
Credit Card	**0.9998**	**0.9998**	0.8391	0.9985	0.8329	0.8729	0.7281	0.9970	0.5809	0.6771
Drug Classification	**0.9898**	0.9681	0.8985	0.7866	0.9720	0.8850	0.6521	0.8072	0.3969	0.8406
Kc2	0.9063	0.8303	0.6016	**0.9996**	**0.9996**	0.5168	0.5637	0.0060	0.1115	0.4958
Eeg eye	**0.9999**	***1***	0.9911	0.8849	0.9251	0.9725	0.9420	0.0040	0.6507	0.5571
Glass Classification	0.9028	0.9038	0.7976	0.9410	**0.9750**	0.8640	0.7265	0.1224	0.3693	0.6458
Pc4	0.9621	0.9618	0.9216	0.9645	0.9453	0.8540	**0.9729**	0.9690	0	0.9453
Madelon	**1**	**1**	0.7463	0.6266	0.8983	0.6411	0.0121	0.0191	0.9565	0.6266
Turing binary	**1**	0.9999	0.3906	0.7223	0.8782	0.0423	0.0355	0.0704	0.0288	0.0009
KDD	0.9806	0.9802	0.9653	0.8587	**0.9989**	0.0973	1	0.9979	0	0.9984
Liver disorder	0.9379	0.9310	0.9264	0.9388	0.9397	**0.9587**	0.9003	0.7038	0.1615	0.9407
Wine	**0.9964**	**0.9938**	0.8617	0.8848	0.9018	0.9136	0.5640	0.9216	0.0100	0.7662
Soy bean	**0.9925**	**0.9934**	0.4762	0.9209	0.7300	0.6050	0.0680	0.0788	0	0.5167
Ionosphere	0.9438	0.9775	0.9399	0.7415	0.9236	0.8900	0.9914	**0.9931**	0	0.9387
Room Occupancy	**0.9999**	***1***	0.9961	0.9968	0.9944	0.9966	**1**	0.9992	0	0.9973
Harth	**0.9780**	***0.9790***	0.9778	0.7330	0.9721	0.9195	0.9930	0.9576	0	0.9929
Rocket League	**0.9995**	***0.9974***	0.0913	0.5412	0.4681	0.0965	0.0056	0	0	0
Sirtuin6	0.9422	0.9424	0.6201	0.7375	0.9060	0.7933	0.7300	0.9776	0	**0.9970**
Toxicity	**0.9994**	***0.9985***	0.0343	0.6200	0.5925	0.1600	0.0033	0	0	0
Dry bean	**0.9979**	***0.9977***	0.8415	0.9574	0.8384	0.8384	0.0015	0.7112	0	0.8020

Table A5. The proposed methods, i.e., OTEC(oob), OTEC(sub) and other state-of-the-art methods, in terms of F1 score.

Dataset	**OTEC(*oob*)**	**OTEC(*sub*)**	**OTE**	**RF(*smote*)**	**RF(*over*)**	**RF(*under*)**	***k*-NN**	**SVM**	**ANN**	**Tree**
Breast Cancer	**1**	0.9679	0.9197	0.9747	0.9974	0.9683	0.9043	0.9499	0.4035	0.9028
Credit Card	**0.9997**	***0.9995***	0.8571	0.9960	0.7692	0.8571	0.8376	0.9929	0.7196	0.6829
Drug Classification	**1**	0.9520	0.8667	0.7931	0.9778	0.8750	0.6425	0.8169	0.5472	0.8618
Kc2	0.9091	0.7556	0.5085	**0.9863**	**0.9863**	0.4839	0.4680	0.0122	0.1955	0.5331
Eeg eye	**0.9959**	***1***	0.8043	0.8970	0.9382	0.9732	0.9373	0.0076	0.6486	0.6284
Glass Classification	0.9184	0.9017	0.7179	**0.9565**	**0.9583**	0.9000	0.6776	0.2345	0.3052	0.6736
Pc4	**0.9807**	0.9790	0.9512	0.9773	0.9621	0.8750	0.9247	0.9442	0	0.9368
Madelon	**0.9824**	0.9776	0.5068	0.6000	0.8754	0.5631	0.0225	0.0288	0	0.6672
Turing binary	**0.9879**	**0.9898**	0.0432	0.6923	0.9297	0.0356	0.0621	0.1062	0.0352	0.0017
KDD	0.9902	0.9900	0.9807	0.8617	**0.9980**	0.0667	0.9899	0.9885	0	0.9906
Liver disorder	0.9508	***0.9559***	**0.9559**	0.9540	0.9244	0.8974	0.9543	0.8408	0.1608	0.9380
Wine	**1**	0.9824	0.8182	0.8955	0.7162	0.7273	0.6358	0.9414	0.0241	0.7539
Soy bean	**1**	0.9898	0.8900	0.9236	0.4000	**1**	0.0384	0.0771	0.0184	0.4123
Ionosphere	0.9604	0.9025	0.9327	0.7304	0.9513	0.8958	0.9406	0.9515	0	0.9385
Room Occupancy	**1**	***1***	**1**	0.9965	1	0.9622	0.9971	0.9996	0	0.9960
Harth	**0.9867**	***0.9882***	0.9797	0.7332	0.9807	0.9000	0.9864	0.9646	0	0.9838
Rocket League	**0.9633**	***0.9658***	0.1071	0.5419	0.4296	0.0909	0.0106	0	0	0
Sirtuin6	**0.9565**	***0.9488***	0	0.7356	0.8965	0.3333	0.9529	0.9485	0	0.9549
Toxicity	**0.9841**	***0.9633***	0	0.6000	0.3333	0.5000	0.0050	0	0	0
Dry bean	**0.9976**	***0.9964***	0.7500	0.9538	0.9781	0.8947	0.0028	0.8032	0	0.8397

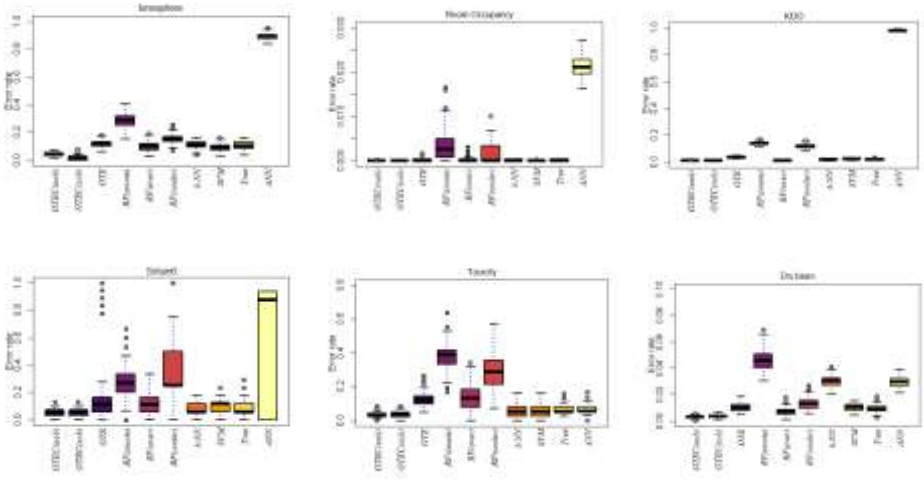

Figure A1. Box plots of OTEC(oob), and OTEC(sub) with other state-of-the-art methods in terms of classification error rate on various datasets.

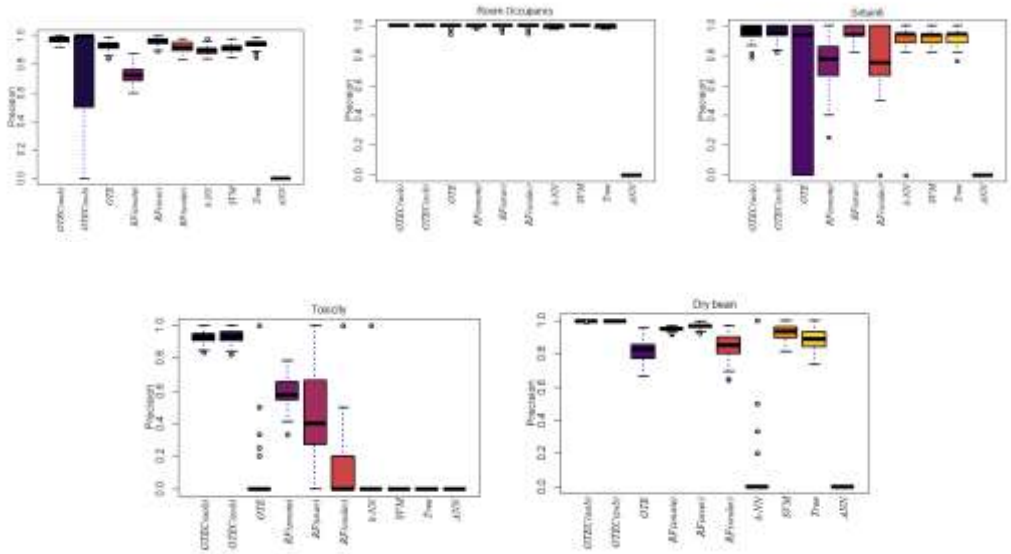

Figure A2. Box plots of OTEC(oob), and OTEC(sub) with other state-of-the-art methods in terms of precision on various datasets.

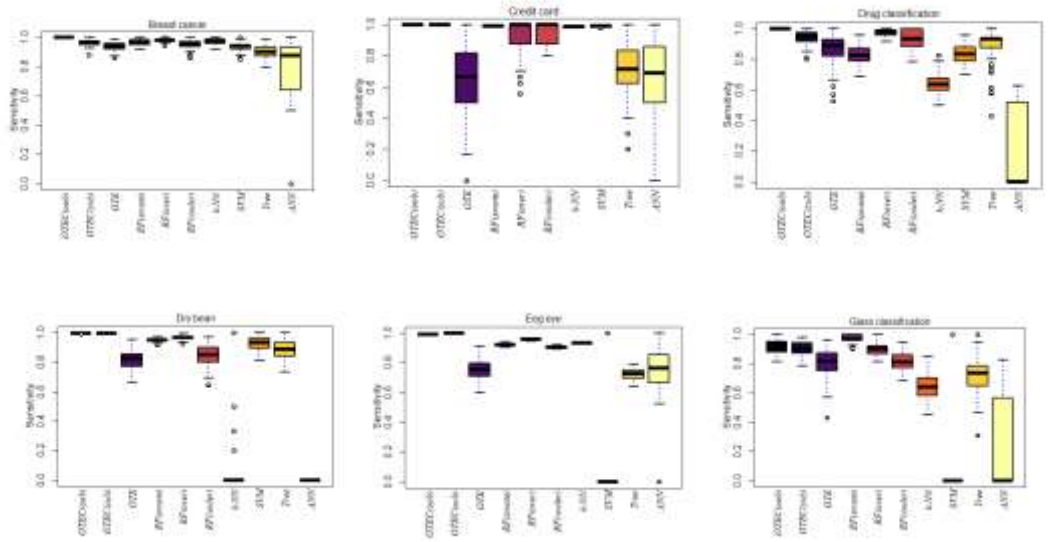

Figure A3. Box plots of OTEC(oob), and OTEC(sub) with other state-of-the-art methods in terms of sensitivity on various datasets.

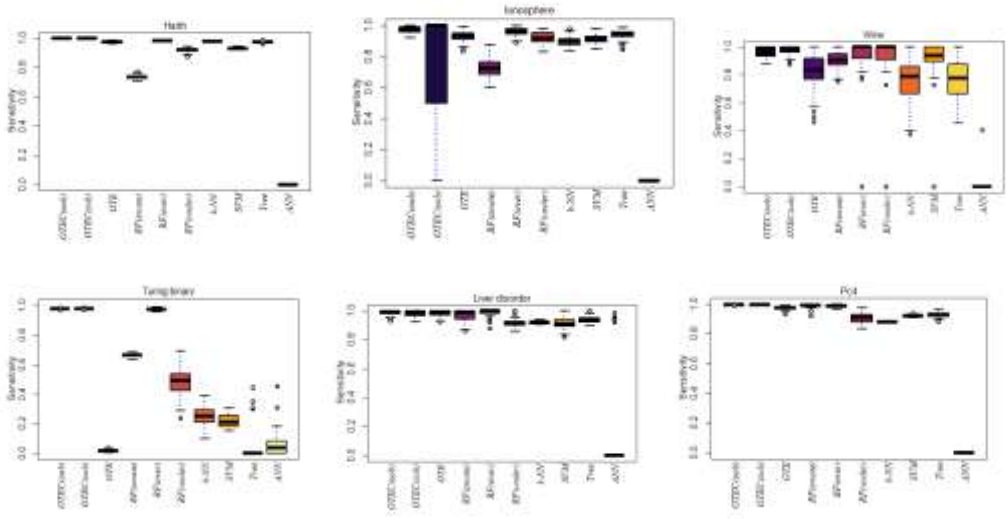

Figure A4. Box plots of OTEC(oob), and OTEC(sub) with other state-of-the-art methods in terms of sensitivity on various datasets.

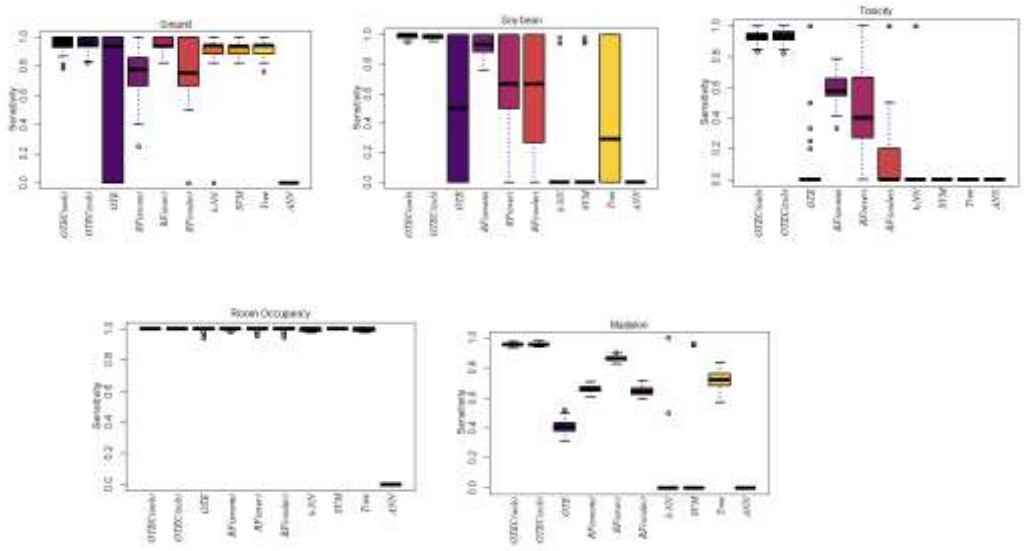

Figure A5. Box plots of OTEC(oob), and OTEC(sub) with other state-of-the-art methods in terms of sensitivity on various datasets.

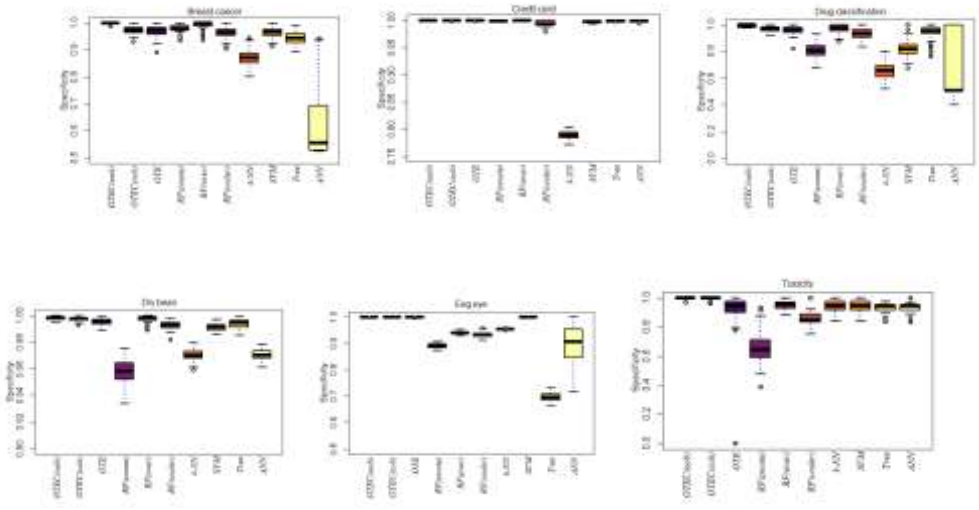

Figure A6. Box plots of OTEC(oob), and OTEC(sub) with other state-of-the-art methods in terms of specificity on various datasets.

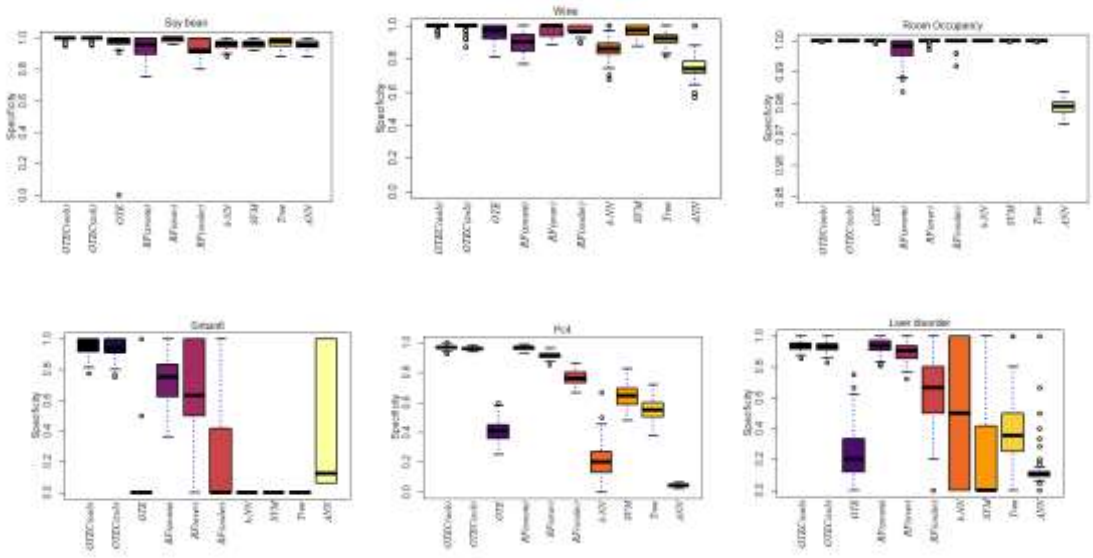

Figure A7. Box plots of OTEC(oob), and OTEC(sub) with other state-of-the-art methods in terms of specificity on various datasets.

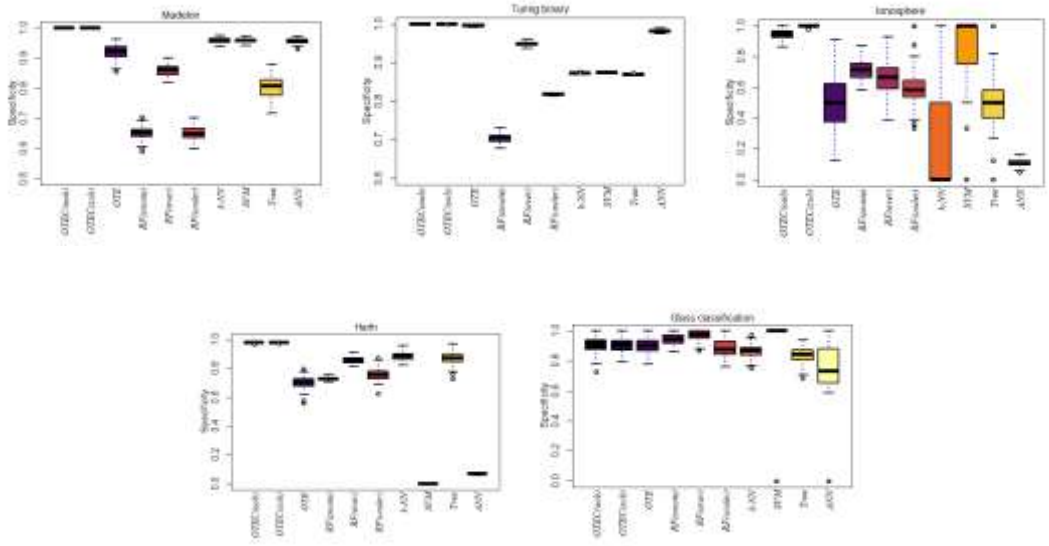

Figure A8. Box plots of OTEC(oob), OTEC(sub), and other state-of-the-art methods in terms of specificity on various datasets.

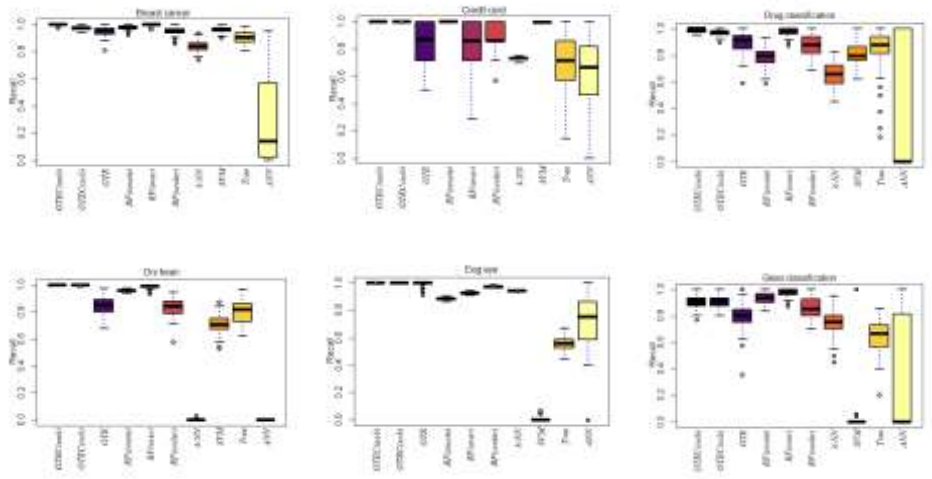

Figure A9. Box plots of OTEC(oob), OTEC(sub), and other state-of-the-art methods in terms of recall on various datasets.

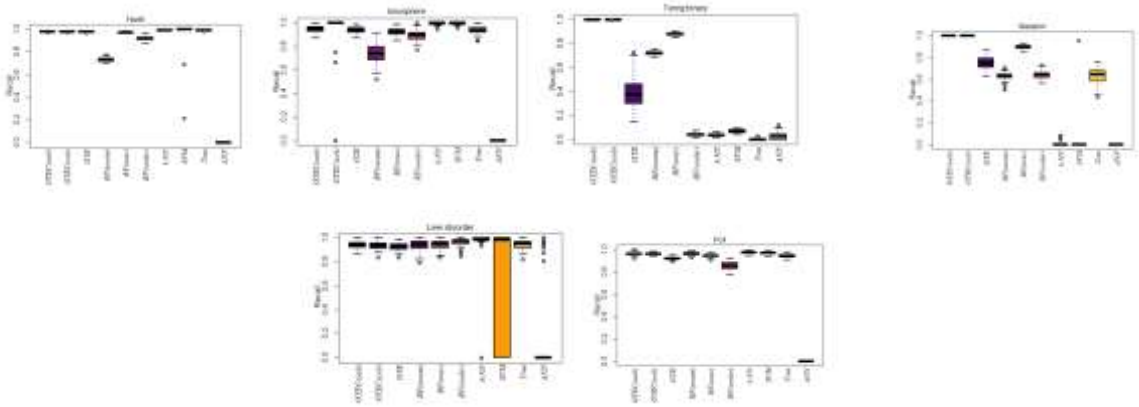

Figure A10. Box plots of OTEC(oob), OTEC(sub), and other state-of-the-art methods in terms of recall on various datasets.

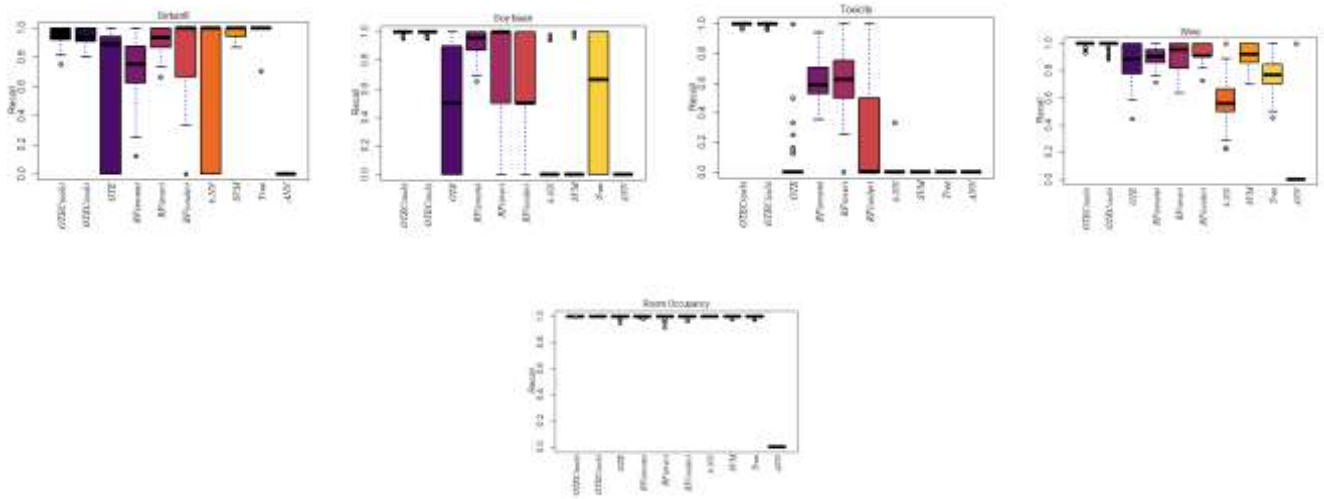

Figure A11. Box plots of OTEC(oob), OTEC(sub), and other state-of-the-art methods in terms of recall on various datasets.

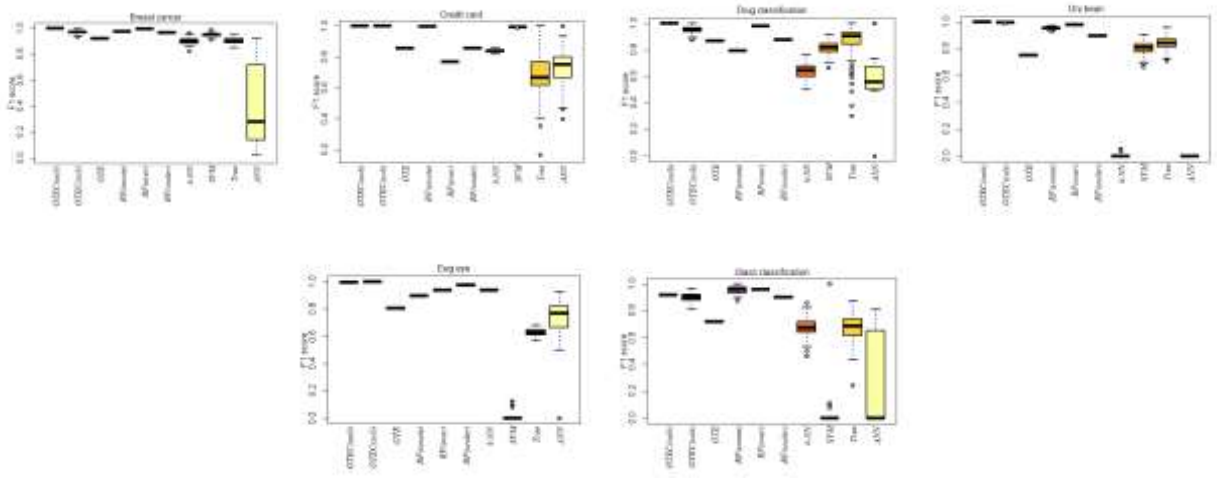

Figure A12. Box plots of OTEC(oob), OTEC(sub), and other state-of-the-art methods in terms of F1 score on various datasets.

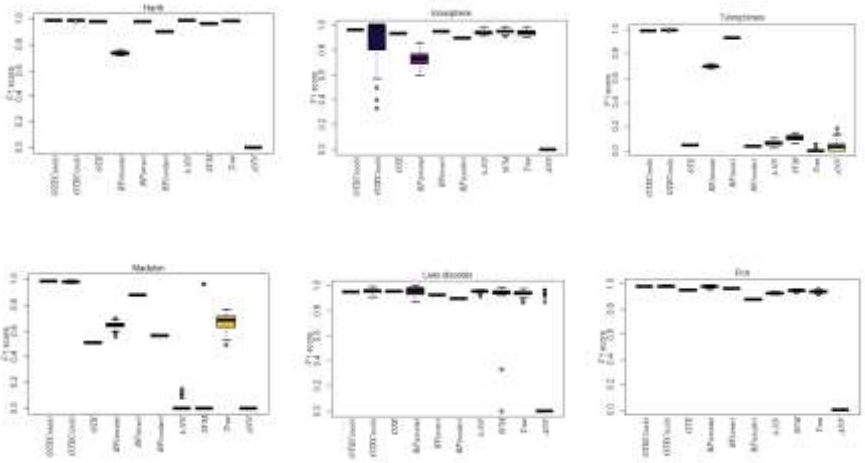

Figure A13. Box plots of OTEC(oob), OTEC(sub), and other state-of-the-art methods in terms of F1 score on the datasets.

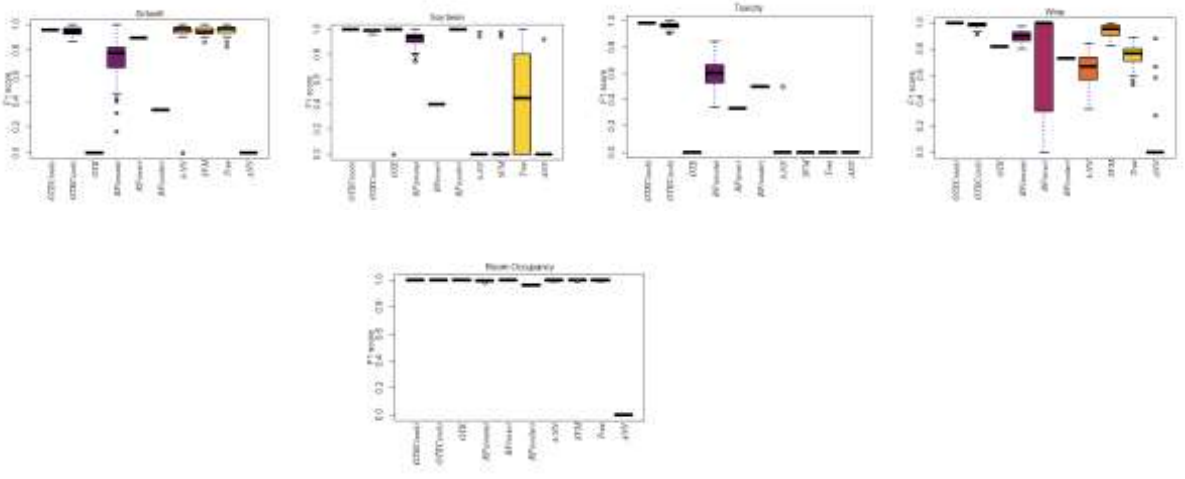

Figure A14. Box plots of OTEC(oob), OTEC(sub), and other state-of-the-art methods in terms of F1 score on various datasets.