REFERENCES

1. Fayyad, U.M., Piatetsky-Shapiro G., and Smyth, P. From data mining to knowledge discovery: an overview. In Usama, M.F. et al.(ed), Advances in knowledge discovery and data mining. pp.1-34. American Association for Artificial Intelligence, 1996.
2. Cerrito, P.B. Introduction to Data Mining Using SAS® Enterprise Miner™. Cary, NC: SAS Institute Inc., 2006.
3. Berry, M.J., Linoff, G. Data Mining Techniques: For Marketing, Sales, and Customer Support. New York, NY, USA: John Wiley & Sons, Inc., 1997.
4. Kumar, V. Data Mining for Scientific and Engineering Applications. Kluwer Norwell, MA, USA: Academic Publishers, 2001.
5. Tomar, D. and Agarwal, S. A survey on Data Mining approaches for Healthcare. International Journal of Bio-Science & Bio-Technology. pp. 241-266. **5**(5), 2013.
6. Larrañaga, P. et al, Machine learning in bioinformatics. Brief Bioinformatics. pp. 86-112. 7(1), March 2006.
7. Quinlan, J.R. Induction of Decision Trees. Machine Learning. pp. 81-106. **1**(1), 1986.
8. Zhang, H. Exploring Conditions For The optimality of naive Bayes. International Journal of Pattern Recognition and Artificial Intelligence. pp. 183-198. **19**, 2005.
9. Tetko, I.V., Livingstone, D.J., and Luik, A.I. Neural-Network Studies .1. Comparison of Overfitting and Overtraining. Journal of Chemical Information and Computer Sciences. pp. 826-833. **35**(5), 1995.
10. Boser, B.E., Guyon I.M., and Vapnik, V.N. A training algorithm for optimal margin classifiers, In Proceedings of the fifth annual workshop on Computational learning theory. pp. 144-152. Pittsburgh, Pennsylvania, USA: ACM, 1992.
11. Cover, T. and Hart, P. Nearest neighbor pattern classification. IEEE Transactions on Information Theory. pp. 21-27. **13**(1), 1967.
12. MacQueen, J. Some methods for classification and analysis of multivariate observations. In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. Volume 1: Statistics. pp. 281-297, University of California Press, 1967.

13. Kaufman, L. and Rousseeuw, P.J. Finding Groups in Data An Introduction to Cluster Analysis. New York: Wiley Interscience, 1990.

14. Ester, M., et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96). pp. 226-231, 1996.

15. Hair, J., et al. Multivariate Data Analysis (7th Edition). Prentice Hall, 2009.

16. Hipp, J., Guntzer, U., and Nakhaeizadeh, G. Algorithms for association rule mining - a general survey and comparison. In SIGKDD Exploration Newsletter. pp. 58-64. 2(1), 2000.

17. Agrawal, R. and Srikant, R. Fast Algorithms for Mining Association Rules in Large Databases, In Proceedings of the 20th International Conference on Very Large Data Bases. pp. 487-499. Morgan Kaufmann Publishers Inc., 1994.

18. Witten, I.H. and Frank, E. Data Mining: Practical Machine Learning Tools and Techniques. Second Edition (Morgan Kaufmann Series in Data Management Systems). Morgan Kaufmann Publishers Inc., 2005.

19. He, H.B. and Garcia, E.A. Learning from Imbalanced Data. IEEE Transactions on Knowledge and Data Engineering. pp. 1263-1284. 21(9), 2009.

20. Kubat, M., Holte, R.C. and Matwin, S. Machine Learning for the Detection of Oil Spills in Satellite Radar Images. Machine Learning. pp. 195-215. 30(2-3), 1998.

21. Fawcett, T.E. and Provost, F. Handbook of data mining and knowledge discovery. Klösgen, W. and Zytkow, J.M. (ed). pp. 726-731. New York, NY, USA: Oxford University Press, Inc, 2002.

22. Japkowicz, N. and Stephen, S. The class imbalance problem: A systematic study. Intelligent Data Analysis. pp. 429-449 6(5), 2002.

23. Lewis, D.D. and Catlett, J. Heterogeneous Uncertainty Sampling for Supervised Learning, In Proceedings of the Eleventh International Conference on Machine Learning. pp. 148-156. Morgan Kaufmann, 1994.

24. Kousarrizi, M., Seiti, F. and Teshnehlab, M. An Experimental Comparative Study on Thyroid Disease Diagnosis Based on Feature Subset Selection and classification. International Journal of Electrical & Computer Sciences. pp. 13-19. 12(1), 2012.

25. Chawla, N.V., et al., SMOTE: Synthetic minority over-sampling technique. Journal of Artificial Intelligence Research. pp. 321-357. 16, 2002.

26. He, H., et al., ADASYN: Adaptive synthetic sampling approach for imbalanced learning, in Neural Networks, IEEE International Joint Conference. pp. 1322-1328, 2008.

27. Han, H., Wang, W.Y. and Mao, B.H. Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning, in Proceedings of the 2005 international conference on Advances in Intelligent Computing. pp. 878-887, Springer-Verlag: Berlin, Heidelberg, 2005.
28. Bunkhumpornpat, C., Sinapiromsaran, K. and Lursinsap, C. Safe-level-SMOTE: Safe-level-Synthetic Minority Over-Sampling TEchnique for Handling the Class Imbalanced Problem, In Proceedings of the 13th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining. pp. 475-482. Springer-Verlag: Berlin, Heidelberg, 2009.
29. Bunkhumpornpat, C., Sinapiromsaran, K. and Lursinsap, C. DBSMOTE: Density-Based Synthetic Minority Over-sampling TEchnique. Applied Intelligence, pp. 664-684. **36**(3), 2012.
30. Wilcoxon, F. Individual comparisons by ranking methods. Biometrics Bulletin, p. 80–83. **1**(6), 1945.
31. Batuwita, R. and V. Palade, A New Performance Measure for Class Imbalance Learning. Application to Bioinformatics Problems, in Proceedings of the 2009 International Conference on Machine Learning and Applications. pp. 545-550. IEEE Computer Society, 2009.
32. Quinlan, J.R., C4.5: programs for machine learning. Morgan Kaufmann Publishers Inc., 1993
33. R Core Team R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/, 2013.
34. Berthold, M.R., et al., KNIME - the Konstanz information miner. ACM SIGKDD Explorations Newsletter **11**(1), 2009.
35. Han, J., Kamber, M., and Pei, J. Data mining: concepts and techniques. Morgan kaufmann, 2006.
36. Murthy, S.K., Automatic Construction of Decision Trees from Data: A Multi-Disciplinary Survey. Data Mining and Knowledge Discovery. pp. 345-389. **2**(4), 1998.
37. Tan, P.N., Steinbach, M., and Kumar, V. Introduction to Data Mining. (First Edition). Addison-Wesley Longman Publishing Co., Inc., 2005.
38. Breiman, L., et al., Classification and Regression Trees. Chapman and Hall, 1984.

39. Quinlan, J.R., Improved use of continuous attributes in C4.5. Journal of Artificial Intelligence Research. pp. 77-90. 4(1), 1996.

40. Goodrich, M.T., et al., Decision tree conctruction in fixed dimensions: Being global is hard but local greed is good., In Technical Report. pp. 1-12. Johns Hopkins Univ., Department of Computer Science, 1995.

41. Hyafil, L. and Rivest, R.L. Constructing optimal binary decision trees is NP-complete. Information Processing Letters. pp. 15-17. 5(1), 1976.

42. Hastie, T., Tibshirani, R. and Friedman, J. The Elements of Statistical Learning : Data Mining, Inference and Prediction. Springer, 2001.

43. Yang Y., Webb, G.I., A Comparative Study of Discretization Methods for Naive-Bayes Classifiers, In Proceedings of PKAW 2002: The 2002 Pacific Rim Knowledge Acquisition Workshop. pp. 159-173, 2002.

44. Burges, C.J.C. A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery. pp. 121-167. 2(2), 1998.

45. Cortes, C. and Vapnik, V. Support Vector Networks. Machine Learning. pp. 273-297. 20(3), 1995.

46. Buhmann, M.D. Radial Basis Functions. Cambridge University Press, 2003.

47. Rumelhart, D.E., Hinton, G.E., and Williams, R.J. Learning internal representations by error propagation In David, E.R., James, L.M. and C.P.R. Group, (ed). Parallel distributed processing: explorations in the microstructure of cognition. pp. 318-362. vol. 1, MIT Press, 1986.

48. Rosenblatt, F. Principles of neurodynamics: Perceptrons and the theory of brain mechanisms. Spartan Books, 1961.

49. Haykin, S. Neural Networks: A Comprehensive Foundation. Prentice Hall PTR, 1998.

50. Bailey, T., Jain A.K. A Note on Distance-Weighted k-Nearest Neighbor Rules. IEEE Transactions on systems, Man and Cybernetics. p. 311-313. 8(4), 1978.

51. Sproull, R.F., Refinements to nearest-neighbor searching in *k*-dimensional trees. Algorithmica. pp. 579-589. 6(1-6), 1991.

52. Beygelzimer, A., Kakade, S. and Langford, J. Cover trees for nearest neighbor, In Proceedings of the 23rd international conference on Machine learning. pp. 97-104, New York, NY, USA: ACM, 2006.

53. Guo, H. and Viktor, H.L. Learning from imbalanced data sets with boosting and data generation: the DataBoost-IM approach. SIGKDD Explorations Newsletter. pp. 30-39. 6(1), pp. 30-39.

54. Woods, K.S., et al., Comparative evaluation of pattern recognition techniques for detection of microcalcifications. In Biomedical Image Processing and Biomedical Visualization. pp. 841-852. San Jose, CA, USA: SPIE, 1993.
55. Weiss, G.M. and Provost, F. Learning when training data are costly: The effect of class distribution on tree induction. Journal of Artificial Intelligence Research. pp. 315-354. 19, 2003.
56. Chawla, N.V., C4.5 and imbalanced data sets: investigating the effect of sampling method, probabilistic estimate, and decision tree structure. In the ICML'03 Workshop on Class Imbalances. 2003.
57. Weiss, G.M. and Provost, F. The effect of class distribution on classifier learning: an empirical study. Technical Report MLTR-43. 2001.
58. Laurikkala, J., Improving Identification of Difficult Small Classes by Balancing Class Distribution. In Proceedings of the 8th Conference on AI in Medicine in Europe: Artificial Intelligence Medicine. pp. 63-66. London, UK: Springer-Verlag, 2001.
59. Estabrooks, A., Jo, T.H. and Japkowicz, N. A multiple resampling method for learning from imbalanced data sets. Computational Intelligence. pp. 18-36. 20(1), 2004.
60. Holte, R.C., Acker, L. and Porter, B.W. Concept Learning and the Problem of Small Disjuncts, In Sridharan, N.S. (ed). International Joint Conferences on Artificial Intelligence. pp. 813-818. Morgan Kaufmann, 1989.
61. Mease, D., Wyner, A.J. and Buja, A. Boosted Classification Trees and Class Probability/Quantile Estimation. Journal of Machine Learning Research. pp. 409-439. 8, 2007.
62. Drummond, C. and Holte, R. C4.5, Class Imbalance, and Cost Sensitivity: Why Under-sampling beats Over-sampling, pp 1-8, 2003; Available from: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.132.9672.
63. Liu, X.Y., Wu, J. and Zhou, Z.H. Exploratory Under-Sampling for Class-Imbalance Learning. Sixth International Conference on Data Mining (ICDM'06). pp. 965-969, 2006.
64. Cohen, W., Fast Effective Rule Induction, In Proceedings of the Twelfth International Conference on Machine Learning. pp. 115-123, 1995.
65. Ling, C.X. and Sheng, V.S. Cost-sensitive learning. Encyclopedia of Machine Learning. pp.231-235 Sammut C. (ed) Springer, New York, 2008.

66. Zhou, Z.H. and Liu, X.Y. Training cost-sensitive neural networks with methods addressing the class imbalance problem. IEEE Transactions on Knowledge and Data Engineering. pp. 63-77. 18(1), 2006.

67. McCarthy, K., Zabar, B. and Weiss, G. Does cost-sensitive learning beat sampling for classifying rare classes?. In Proceedings of the 1st international workshop on Utility-based data mining. pp. 69-77. ACM: Chicago, Illinois, 2005.

68. Elkan, C. The Foundations of Cost-Sensitive Learning. In Proceedings of the Seventeenth International Joint Conference of Artificial Intelligence. pp. 973-978. Morgan Kaufmann: Seattle, Washington, 2001.

69. Sheng, V.S. and Ling, C.X. Thresholding for Making Classifiers Cost-sensitive. In Proceedings of the 21st National Conference on Artificial Intelligence. pp. 476-481. Boston, Massachusetts, 2006.

70. Drummond, C. and Holte, R. Exploiting the cost (in)sensitivity of decision tree splitting criteria. In Proceedings of the 17th International Conference on Machine Learning. pp.239-246, 2000.

71. Turney, P.D. Cost-Sensitive Classification: Empirical Evaluation of a Hybrid Genetic Decision Tree Induction Algorithm. Journal of Artificial Intelligence Research. pp. 369-409. 2, 1995.

72. Ling, C.X., Yang, Q., Wang, J., and Zhang, S. Decision Trees with Minimal Costs. In Proceedings of 2004 International Conference on Machine Learning (ICML'2004). pp. 69-76, 2004.

73. Sun, Y., et al., Cost-sensitive boosting for classification of imbalanced data. Pattern Recognition. p. 3358-3378. 40(12), 2007.

74. Fan, W., et al., AdaCost: Misclassification Cost-Sensitive Boosting. In Proceedings of the Sixteenth International Conference on Machine Learning. p. 97-105. Morgan Kaufmann Publishers Inc, 1999.

75. Domingos, P. MetaCost: A general method for making classifiers cost-sensitive. In Proceedings of the Fifth International Conference on Knowledge Discovery and Data Mining. pp. 155-164, ACM Press, 1999.

76. Witten, I.H., and Frank, E. Data Mining – Practical Machine Learning Tools and Techniques with Java Implementations. Morgan Kaufmann Publishers, 2005.

77. Chai, X., Deng, L., Yang, Q., and Ling, C.X. Test-Cost Sensitive Naïve Bayesian Classification. In Proceedings of the Fourth IEEE International Conference on Data Mining. IEEE Computer Society Press, Brighton, UK , 2004.

78. Zadrozny, B., Langford, J., and Abe, N. Cost-sensitive learning by Cost-Proportionate instance Weighting. In Proceedings of the 3th International Conference on Data Mining. 2003.

79. Hong, X., Chen, S. and Harris, C.J. A Kernel-Based Two-Class Classifier for Imbalanced Data Sets. IEEE Transactions on Neural Networks and Learning System. pp. 28-41. 18(1), 2007.

80. Wu, G. and Chang, E. Class-Boundary Alignment for Imbalanced Data Set Learning. In Proceeding of international conference of Data Mining (ICDM '03). workshop Learning from Imbalanced Data Sets II. 2003.

81. Wu, G. and Chang, E.Y. KBA: Kernel Boundary Alignment Considering Imbalanced Data Distribution. IEEE Transactions on Knowledge and Data Engineering. pp. 786-795. 17(6), 2005.

82. Wu, G. and Chang, E. Adaptive Feature-Space Conformal Transformation for Imbalanced Data Learning. In Proceedings of the Twentieth International Conference on Machine Learning. pp. 816-823, 2003.

83. Hung, L.Y. and Ting, C.Y. Total margin based adaptive fuzzy support vector machines for multiview face recognition. In 2005 IEEE International Conference on Systems, Man and Cybernetics. pp. 1704-1711, 2005.

84. Fung, G. and Mangasarian, O.L. Multicategory Proximal Support Vector Machine Classifiers. Machine Learning. p. 77-97. 59(1-2), 2005.

85. Ertekin, S., et al., Learning on the border: active learning in imbalanced data classification. In Proceedings of the sixteenth ACM conference on Conference on information and knowledge management. pp. 127-136. Lisbon, Portugal : ACM, 2007.

86. Bordes, A., et al., Fast Kernel Classifiers with Online and Active Learning. Journal of Machine Learning Research. pp. 1579-1619. 6, 2005.

87. Sokolova, M. and Lapalme, G. A systematic analysis of performance measures for classification tasks. Information Processing & Management. pp. 427-437. 45(4), 2009.

88. Kubat, M. and Matwin, S. Addressing the curse of imbalanced training sets: one-sided selection. In Proceedings of the Fourteenth International Conference on Machine Learning. pp. 179-186. Morgan Kaufmann, 1997.

89. Arya, S., et al., An optimal algorithm for approximate nearest neighbor searching fixed dimensions. Journal of the ACM. pp. 891-923. 45(6), 1998.

90. Newman, D.J., et al., UCI Repository of machine learning databases. 1998.

91. Boetticher, G., Menzies, T. and Ostrand, T. {PROMISE} Repository of empirical software engineering data. 2007.

APPENDIX

Appendix A : Experimental result

Table 22: The average percentage of minority outcasts in positive instances in each dataset when the number of *c* is varied

c	cm1	ecoli	glass	haberman	jm1	kc1	kc2
1	56.816%	17.700%	22.895%	47.704%	44.813%	42.975%	33.626%
2	44.612%	11.400%	13.711%	34.173%	33.679%	34.110%	23.364%
3	38.449%	9.800%	7.474%	25.556%	26.759%	27.675%	17.682%
4	33.510%	8.600%	4.658%	18.963%	21.949%	23.718%	14.224%
5	29.184%	6.500%	3.263%	14.543%	18.257%	20.479%	12.093%
6	25.714%	4.900%	2.763%	11.407%	15.239%	18.031%	10.935%
7	22.000%	3.700%	2.184%	8.815%	12.914%	16.313%	9.832%
8	18.531%	2.000%	1.895%	7.012%	11.124%	14.982%	8.879%
9	16.122%	1.300%	1.658%	5.630%	9.670%	13.626%	8.318%
10	14.612%	1.100%	1.526%	4.519%	8.476%	12.380%	8.000%
11	13.306%	0.900%	1.395%	3.679%	7.486%	11.166%	7.664%
12	12.327%	0.700%	1.211%	2.889%	6.607%	10.031%	7.383%
13	11.673%	0.600%	1.026%	2.321%	5.855%	9.086%	7.103%
14	11.265%	0.500%	0.895%	1.506%	5.213%	8.160%	6.879%
15	10.776%	0.300%	0.711%	1.136%	4.645%	7.350%	6.617%
16	10.408%	0.000%	0.605%	0.840%	4.172%	6.669%	6.336%
17	10.000%	0.000%	0.500%	0.667%	3.702%	6.080%	6.187%
18	9.469%	0.000%	0.342%	0.519%	3.286%	5.442%	5.925%
19	8.898%	0.000%	0.237%	0.321%	2.935%	4.816%	5.776%
20	8.163%	0.000%	0.158%	0.247%	2.641%	4.288%	5.589%

c	letter	optdigits	pc1	satimage	segment	vehicle	yeast	Average
1	8.586%	0.159%	43.195%	20.431%	7.533%	5.321%	22.724%	26.748%
2	4.817%	0.152%	36.468%	12.198%	4.382%	1.725%	13.890%	19.191%
3	3.297%	0.126%	31.169%	8.534%	3.303%	0.679%	10.233%	15.053%
4	2.460%	0.105%	27.013%	6.907%	2.624%	0.431%	8.663%	12.416%
5	1.877%	0.072%	24.753%	5.712%	2.030%	0.147%	7.926%	10.488%
6	1.455%	0.018%	22.675%	4.776%	1.576%	0.073%	6.982%	9.039%
7	1.139%	0.014%	20.156%	4.006%	1.309%	0.028%	6.012%	7.744%
8	0.896%	0.014%	17.818%	3.447%	1.055%	0.018%	5.252%	6.637%
9	0.668%	0.011%	16.052%	3.067%	0.879%	0.009%	4.589%	5.828%
10	0.499%	0.011%	14.779%	2.716%	0.715%	0.009%	4.098%	5.246%
11	0.371%	0.007%	13.688%	2.463%	0.588%	0.000%	3.632%	4.739%
12	0.278%	0.004%	12.805%	2.195%	0.479%	0.000%	3.252%	4.297%
13	0.218%	0.000%	11.870%	1.978%	0.424%	0.000%	2.822%	3.927%
14	0.174%	0.000%	11.351%	1.843%	0.358%	0.000%	2.601%	3.625%
15	0.131%	0.000%	10.571%	1.728%	0.333%	0.000%	2.380%	3.334%
16	0.117%	0.000%	9.974%	1.585%	0.309%	0.000%	2.221%	3.088%
17	0.104%	0.000%	9.532%	1.466%	0.291%	0.000%	2.135%	2.905%
18	0.098%	0.000%	9.065%	1.307%	0.291%	0.000%	2.000%	2.696%
19	0.093%	0.000%	8.675%	1.211%	0.267%	0.000%	1.951%	2.513%
20	0.084%	0.000%	8.286%	1.137%	0.242%	0.000%	1.840%	2.334%

In this section, the average F-measure values from original imbalanced dataset, SMOTE and Triangular minority oversampling technique with a specified classifier are shown. The result of comparison among these techniques is reported by highlighting the best value in each case of dataset and classifier with the dark gray shading and the second best value with the light gray shading.

Table 23: The comparison of Triangular minority oversampling technique with using original imbalanced dataset (ORIG) and SMOTE under F-measure

Classifier	Dataset	ORIG	SMOTE	TMOT
DT	CM1	0.7102	0.7615	0.7137
	ecoli	0.6266	0.6498	0.6342
	glass	0.2426	0.4881	0.5029
	haberman	0.2943	0.3837	0.3546
	jm1	0.3534	0.4210	0.4015
	kc1	0.4971	0.5452	0.5437
	kc2	0.8058	0.7542	0.7557
	letter	0.9670	0.9644	0.9664
	optdigits	0.2735	0.3395	0.3427
	pc1	0.5637	0.5734	0.5732
	satimage	0.8553	0.8798	0.8724
	segment	0.9135	0.9142	0.9206
	vehicle	0.7712	0.7677	0.7675
	yeast	0.0883	0.2582	0.2609

Classifier	Dataset	ORIG	SMOTE	TMOT	Classifier	Dataset	ORIG	SMOTE	TMOT
NB	CM1	0.6227	0.7190	0.6635	SVM	CM1	0.8661	0.7936	0.8032
	ecoli	0.3430	0.5572	0.5572		ecoli	0.4982	0.6220	0.6223
	glass	0.3383	0.4108	0.4188		glass	0.1077	0.4490	0.4533
	haberman	0.4155	0.3929	0.3937		haberman	0.0422	0.3804	0.3812
	jm1	0.4501	0.4565	0.4568		jm1	0.2168	0.4183	0.4167
	kc1	0.5690	0.5724	0.5718		kc1	0.4576	0.5487	0.5532
	kc2	0.2679	0.3274	0.3244		kc2	0.4021	0.4675	0.4793
	letter	0.7701	0.6885	0.6471		letter	0.9930	0.9930	0.9931
	optdigits	0.2197	0.2122	0.2131		optdigits	0.0307	0.2816	0.2850
	pc1	0.0584	0.2980	0.2887		pc1	0.5555	0.5753	0.5782
	satimage	0.3247	0.3503	0.3478		satimage	0.7297	0.7198	0.7226
	segment	0.5567	0.5959	0.5927		segment	0.9465	0.9422	0.9405
	vehicle	0.4931	0.7241	0.7087		vehicle	0.7659	0.7345	0.7379
	yeast	0.2707	0.2888	0.2895		yeast	0.0260	0.3711	0.3691
MLP	CM1	0.8297	0.7632	0.7610	KNN	CM1	0.8432	0.7647	0.7774
	ecoli	0.5807	0.6136	0.6152		ecoli	0.6685	0.6986	0.7009
	glass	0.3719	0.5091	0.5086		glass	0.2860	0.4145	0.4081
	haberman	0.2295	0.4212	0.4174		haberman	0.3367	0.4021	0.4036
	jm1	0.3372	0.4033	0.3983		jm1	0.3724	0.4425	0.4424
	kc1	0.4798	0.5491	0.5450		kc1	0.5362	0.5143	0.5088
	kc2	0.5302	0.2508	0.2446		kc2	0.8808	0.8064	0.8023
	letter	0.9860	0.9856	0.9861		letter	0.9977	0.9971	0.9971
	optdigits	0.2216	0.3035	0.3028		optdigits	0.2624	0.3386	0.3385
	pc1	0.4452	0.4435	0.4524		pc1	0.6994	0.6230	0.6072
	satimage	0.7597	0.7067	0.7139		satimage	0.8757	0.8702	0.8702
	segment	0.9301	0.9184	0.9232		segment	0.9091	0.8774	0.8797
	vehicle	0.7632	0.6879	0.6959		vehicle	0.7348	0.6955	0.7000
	yeast	0.2098	0.3657	0.3533		yeast	0.1506	0.2901	0.2906

In this section, the average values of three performance measures from each oversampling technique and RSLS-ANS1-ANS2 with a specified classifier are shown. The result of comparison among these techniques are displayed by highlighting the best value in each case of dataset and classifier with the dark gray shading and the top three with the light gray shading.

Table 24: The comparison with relocating safe-level SMOTE under F-measure

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	RSLS
DT	cm1	0.7102	0.7615	0.6653	0.7504	0.7140	0.7548
	ecoli	0.6266	0.6498	0.6605	0.6468	0.6377	0.6697
	glass	0.2426	0.4881	0.5056	0.4746	0.4576	0.4751
	haberman	0.2943	0.3837	0.3791	0.3903	0.3514	0.4158
	jm1	0.3534	0.4210	0.4166	0.4176	0.3742	0.4528
	kc1	0.4971	0.5452	0.5422	0.5532	0.5239	0.5487
	kc2	0.8058	0.7542	0.7584	0.7530	0.7126	0.7581
	letter	0.9670	0.9644	0.9661	0.9678	0.9668	0.9665
	optdigits	0.2735	0.3395	0.3532	0.3346	0.2830	0.4082
	pc1	0.5637	0.5734	0.5641	0.5798	0.5664	0.5989
	satimage	0.8553	0.8798	0.8704	0.8801	0.8831	0.8787
	segment	0.9135	0.9142	0.9336	0.9206	0.9211	0.9184
	vehicle	0.7712	0.7677	0.7602	0.7725	0.7683	0.7759
	yeast	0.0883	0.2582	0.2801	0.1947	0.2242	0.2336

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	RSLS
NB	cm1	0.6227	0.7190	0.6241	0.6802	0.5371	0.6555
	ecoli	0.3430	0.5572	0.4236	0.5352	0.4423	0.5617
	glass	0.3383	0.4108	0.4657	0.4449	0.3677	0.4622
	haberman	0.4155	0.3929	0.3981	0.4023	0.3812	0.4137
	jm1	0.4501	0.4565	0.4543	0.4508	0.4331	0.4730
	kc1	0.5690	0.5724	0.5550	0.5739	0.5756	0.5849
	kc2	0.2679	0.3274	0.1597	0.3294	0.2515	0.3581
	letter	0.7701	0.6885	0.1047	0.6881	0.2167	0.6905
	optdigits	0.2197	0.2122	0.2154	0.2040	0.2296	0.2338
	pc1	0.0584	0.2980	0.4024	0.2466	0.0510	0.2971
	satimage	0.3247	0.3503	0.3471	0.3501	0.3147	0.3596
	segment	0.5567	0.5959	0.4011	0.5866	0.4775	0.5929
	vehicle	0.4931	0.7241	0.7404	0.6901	0.4170	0.7192
	yeast	0.2707	0.2888	0.2793	0.2870	0.1493	0.3166
MLP	cm1	0.8297	0.7632	0.7243	0.8148	0.7626	0.8310
	ecoli	0.5807	0.6136	0.6110	0.6095	0.6234	0.6244
	glass	0.3719	0.5091	0.5107	0.4822	0.4902	0.4997
	haberman	0.2295	0.4212	0.4199	0.4350	0.3868	0.4488
	jm1	0.3372	0.4033	0.3987	0.4344	0.4051	0.4506
	kc1	0.4798	0.5491	0.5341	0.5433	0.5402	0.5659
	kc2	0.5302	0.2508	0.2023	0.2586	0.2979	0.2417
	letter	0.9860	0.9856	0.9849	0.9856	0.9822	0.9863
	optdigits	0.2216	0.3035	0.3106	0.2925	0.2571	0.3282
	pc1	0.4452	0.4435	0.4048	0.4856	0.4942	0.4794
	satimage	0.7597	0.7067	0.6904	0.7126	0.7349	0.7054
	segment	0.9301	0.9184	0.9087	0.9245	0.9163	0.9202
	vehicle	0.7632	0.6879	0.6422	0.7104	0.7227	0.7172
	yeast	0.2098	0.3657	0.3661	0.3338	0.3026	0.3748

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	RSLS
SVM	cm1	0.8661	0.7936	0.7608	0.8208	0.8075	0.8200
	ecoli	0.4982	0.6220	0.6357	0.6354	0.6529	0.6522
	glass	0.1077	0.4490	0.4861	0.4724	0.4512	0.5058
	haberman	0.0422	0.3804	0.4092	0.4144	0.3798	0.4476
	jm1	0.2168	0.4183	0.4157	0.4423	0.4113	0.4550
	kc1	0.4576	0.5487	0.5443	0.5639	0.5625	0.5763
	kc2	0.4021	0.4675	0.4057	0.4808	0.5344	0.4807
	letter	0.9930	0.9930	0.9930	0.9929	0.9930	0.9933
	optdigits	0.0307	0.2816	0.2891	0.2643	0.2579	0.3133
	pc1	0.5555	0.5753	0.5421	0.6027	0.6329	0.6077
	satimage	0.7297	0.7198	0.7088	0.7266	0.7543	0.7284
	segment	0.9465	0.9422	0.9390	0.9426	0.9456	0.9433
	vehicle	0.7659	0.7345	0.6906	0.7567	0.7601	0.7587
	yeast	0.0260	0.3711	0.3689	0.3503	0.3604	0.3886
KNN	cm1	0.8432	0.7647	0.7516	0.8381	0.7919	0.8317
	ecoli	0.6685	0.6986	0.6978	0.7084	0.6951	0.7137
	glass	0.2860	0.4145	0.3980	0.3563	0.3839	0.4021
	haberman	0.3367	0.4021	0.3992	0.4010	0.3860	0.4347
	jm1	0.3724	0.4425	0.4388	0.4381	0.4262	0.4775
	kc1	0.5362	0.5143	0.5045	0.5372	0.5370	0.5522
	kc2	0.8808	0.8064	0.8016	0.8435	0.8376	0.8015
	letter	0.9977	0.9971	0.9977	0.9971	0.9974	0.9969
	optdigits	0.2624	0.3386	0.3340	0.3399	0.3279	0.4162
	pc1	0.6994	0.6230	0.6075	0.6501	0.6516	0.6280
	satimage	0.8757	0.8702	0.8634	0.8737	0.8767	0.8595
	segment	0.9091	0.8774	0.8736	0.8841	0.8817	0.8611
	vehicle	0.7348	0.6955	0.6773	0.7478	0.7255	0.7537
	yeast	0.1506	0.2901	0.2827	0.2626	0.2801	0.3059

Table 25: The comparison with adaptive neighbors SMOTE without minority outcast handling under F-measure

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	ANS1
DT	cm1	0.7102	0.7615	0.6653	0.7504	0.7140	0.7531
	ecoli	0.6266	0.6498	0.6605	0.6468	0.6377	0.6689
	glass	0.2426	0.4881	0.5056	0.4746	0.4576	0.5016
	haberman	0.2943	0.3837	0.3791	0.3903	0.3514	0.3593
	jm1	0.3534	0.4210	0.4166	0.4176	0.3742	0.4410
	kc1	0.4971	0.5452	0.5422	0.5532	0.5239	0.5615
	kc2	0.8058	0.7542	0.7584	0.7530	0.7126	0.7450
	letter	0.9670	0.9644	0.9661	0.9678	0.9668	0.9667
	optdigits	0.2735	0.3395	0.3532	0.3346	0.2830	0.3233
	pc1	0.5637	0.5734	0.5641	0.5798	0.5664	0.5783
	satimage	0.8553	0.8798	0.8704	0.8801	0.8831	0.8821
	segment	0.9135	0.9142	0.9336	0.9206	0.9211	0.9183
	vehicle	0.7712	0.7677	0.7602	0.7725	0.7683	0.7755
	yeast	0.0883	0.2582	0.2801	0.1947	0.2242	0.2663
NB	cm1	0.6227	0.7190	0.6241	0.6802	0.5371	0.7012
	ecoli	0.3430	0.5572	0.4236	0.5352	0.4423	0.5445
	glass	0.3383	0.4108	0.4657	0.4449	0.3677	0.4528
	haberman	0.4155	0.3929	0.3981	0.4023	0.3812	0.4075
	jm1	0.4501	0.4565	0.4543	0.4508	0.4331	0.4500
	kc1	0.5690	0.5724	0.5550	0.5739	0.5756	0.5814
	kc2	0.2679	0.3274	0.1597	0.3294	0.2515	0.3186
	letter	0.7701	0.6885	0.1047	0.6881	0.2167	0.6029
	optdigits	0.2197	0.2122	0.2154	0.2040	0.2296	0.2164
	pc1	0.0584	0.2980	0.4024	0.2466	0.0510	0.1574
	satimage	0.3247	0.3503	0.3471	0.3501	0.3147	0.3421
	segment	0.5567	0.5959	0.4011	0.5866	0.4775	0.5739
	vehicle	0.4931	0.7241	0.7404	0.6901	0.4170	0.6635
	yeast	0.2707	0.2888	0.2793	0.2870	0.1493	0.2796

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	ANS1
MLP	cm1	0.8297	0.7632	0.7243	0.8148	0.7626	0.7853
	ecoli	0.5807	0.6136	0.6110	0.6095	0.6234	0.6068
	glass	0.3719	0.5091	0.5107	0.4822	0.4902	0.5025
	haberman	0.2295	0.4212	0.4199	0.4350	0.3868	0.4264
	jm1	0.3372	0.4033	0.3987	0.4344	0.4051	0.4400
	kc1	0.4798	0.5491	0.5341	0.5433	0.5402	0.5612
	kc2	0.5302	0.2508	0.2023	0.2586	0.2979	0.2473
	letter	0.9860	0.9856	0.9849	0.9856	0.9822	0.9840
	optdigits	0.2216	0.3035	0.3106	0.2925	0.2571	0.3075
	pc1	0.4452	0.4435	0.4048	0.4856	0.4942	0.4884
	satimage	0.7597	0.7067	0.6904	0.7126	0.7349	0.7047
	segment	0.9301	0.9184	0.9087	0.9245	0.9163	0.9204
	vehicle	0.7632	0.6879	0.6422	0.7104	0.7227	0.7177
	yeast	0.2098	0.3657	0.3661	0.3338	0.3026	0.3616
SVM	cm1	0.8661	0.7936	0.7608	0.8208	0.8075	0.7990
	ecoli	0.4982	0.6220	0.6357	0.6354	0.6529	0.6159
	glass	0.1077	0.4490	0.4861	0.4724	0.4512	0.4869
	haberman	0.0422	0.3804	0.4092	0.4144	0.3798	0.4266
	jm1	0.2168	0.4183	0.4157	0.4423	0.4113	0.4354
	kc1	0.4576	0.5487	0.5443	0.5639	0.5625	0.5701
	kc2	0.4021	0.4675	0.4057	0.4808	0.5344	0.4822
	letter	0.9930	0.9930	0.9930	0.9929	0.9930	0.9930
	optdigits	0.0307	0.2816	0.2891	0.2643	0.2579	0.2814
	pc1	0.5555	0.5753	0.5421	0.6027	0.6329	0.6150
	satimage	0.7297	0.7198	0.7088	0.7266	0.7543	0.7287
	segment	0.9465	0.9422	0.9390	0.9426	0.9456	0.9388
	vehicle	0.7659	0.7345	0.6906	0.7567	0.7601	0.7562
	yeast	0.0260	0.3711	0.3689	0.3503	0.3604	0.3802

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	ANS1
KNN	cm1	0.8432	0.7647	0.7516	0.8381	0.7919	0.8096
	ecoli	0.6685	0.6986	0.6978	0.7084	0.6951	0.6951
	glass	0.2860	0.4145	0.3980	0.3563	0.3839	0.4146
	haberman	0.3367	0.4021	0.3992	0.4010	0.3860	0.4522
	jm1	0.3724	0.4425	0.4388	0.4381	0.4262	0.4733
	kc1	0.5362	0.5143	0.5045	0.5372	0.5370	0.5821
	kc2	0.8808	0.8064	0.8016	0.8435	0.8376	0.7778
	letter	0.9977	0.9971	0.9977	0.9971	0.9974	0.9971
	optdigits	0.2624	0.3386	0.3340	0.3399	0.3279	0.3534
	pc1	0.6994	0.6230	0.6075	0.6501	0.6516	0.6098
	satimage	0.8757	0.8702	0.8634	0.8737	0.8767	0.8696
	segment	0.9091	0.8774	0.8736	0.8841	0.8817	0.8613
	vehicle	0.7348	0.6955	0.6773	0.7478	0.7255	0.7505
	yeast	0.1506	0.2901	0.2827	0.2626	0.2801	0.3140

Table 26: The comparison with adaptive neighbors SMOTE with minority outcast handling under F-measure

classifier	dataset	ORIG	SMOTE	ADASYN	SLSORIG	DBSMOTE	ANS2
DT	cm1	0.7102	0.7615	0.6653	0.7504	0.7140	0.7549
	ecoli	0.6266	0.6498	0.6605	0.6468	0.6377	0.6793
	glass	0.2426	0.4881	0.5056	0.4746	0.4576	0.5197
	haberman	0.2943	0.3837	0.3791	0.3903	0.3514	0.3985
	jm1	0.3534	0.4210	0.4166	0.4176	0.3742	0.4673
	kc1	0.4971	0.5452	0.5422	0.5532	0.5239	0.5737
	kc2	0.8058	0.7542	0.7584	0.7530	0.7126	0.7526
	letter	0.9670	0.9644	0.9661	0.9678	0.9668	0.9670
	optdigits	0.2735	0.3395	0.3532	0.3346	0.2830	0.3804
	pc1	0.5637	0.5734	0.5641	0.5798	0.5664	0.5918
	satimage	0.8553	0.8798	0.8704	0.8801	0.8831	0.8833
	segment	0.9135	0.9142	0.9336	0.9206	0.9211	0.9184
	vehicle	0.7712	0.7677	0.7602	0.7725	0.7683	0.7806
	yeast	0.0883	0.2582	0.2801	0.1947	0.2242	0.3019
NB	cm1	0.6227	0.7190	0.6241	0.6802	0.5371	0.7106
	ecoli	0.3430	0.5572	0.4236	0.5352	0.4423	0.5652
	glass	0.3383	0.4108	0.4657	0.4449	0.3677	0.4758
	haberman	0.4155	0.3929	0.3981	0.4023	0.3812	0.4186
	jm1	0.4501	0.4565	0.4543	0.4508	0.4331	0.4721
	kc1	0.5690	0.5724	0.5550	0.5739	0.5756	0.5855
	kc2	0.2679	0.3274	0.1597	0.3294	0.2515	0.3459
	letter	0.7701	0.6885	0.1047	0.6881	0.2167	0.6083
	optdigits	0.2197	0.2122	0.2154	0.2040	0.2296	0.2464
	pc1	0.0584	0.2980	0.4024	0.2466	0.0510	0.2133
	satimage	0.3247	0.3503	0.3471	0.3501	0.3147	0.3509
	segment	0.5567	0.5959	0.4011	0.5866	0.4775	0.5807
	vehicle	0.4931	0.7241	0.7404	0.6901	0.4170	0.6909
	yeast	0.2707	0.2888	0.2793	0.2870	0.1493	0.3147

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	ANS2
MLP	cm1	0.8297	0.7632	0.7243	0.8148	0.7626	0.7967
	ecoli	0.5807	0.6136	0.6110	0.6095	0.6234	0.6205
	glass	0.3719	0.5091	0.5107	0.4822	0.4902	0.5182
	haberman	0.2295	0.4212	0.4199	0.4350	0.3868	0.4455
	jm1	0.3372	0.4033	0.3987	0.4344	0.4051	0.4663
	kc1	0.4798	0.5491	0.5341	0.5433	0.5402	0.5734
	kc2	0.5302	0.2508	0.2023	0.2586	0.2979	0.2585
	letter	0.9860	0.9856	0.9849	0.9856	0.9822	0.9880
	optdigits	0.2216	0.3035	0.3106	0.2925	0.2571	0.3558
	pc1	0.4452	0.4435	0.4048	0.4856	0.4942	0.4709
	satimage	0.7597	0.7067	0.6904	0.7126	0.7349	0.7157
	segment	0.9301	0.9184	0.9087	0.9245	0.9163	0.9243
	vehicle	0.7632	0.6879	0.6422	0.7104	0.7227	0.7183
	yeast	0.2098	0.3657	0.3661	0.3338	0.3026	0.3880
SVM	cm1	0.8661	0.7936	0.7608	0.8208	0.8075	0.7990
	ecoli	0.4982	0.6220	0.6357	0.6354	0.6529	0.6240
	glass	0.1077	0.4490	0.4861	0.4724	0.4512	0.5070
	haberman	0.0422	0.3804	0.4092	0.4144	0.3798	0.4475
	jm1	0.2168	0.4183	0.4157	0.4423	0.4113	0.4523
	kc1	0.4576	0.5487	0.5443	0.5639	0.5625	0.5751
	kc2	0.4021	0.4675	0.4057	0.4808	0.5344	0.4842
	letter	0.9930	0.9930	0.9930	0.9929	0.9930	0.9930
	optdigits	0.0307	0.2816	0.2891	0.2643	0.2579	0.3150
	pc1	0.5555	0.5753	0.5421	0.6027	0.6329	0.6206
	satimage	0.7297	0.7198	0.7088	0.7266	0.7543	0.7309
	segment	0.9465	0.9422	0.9390	0.9426	0.9456	0.9389
	vehicle	0.7659	0.7345	0.6906	0.7567	0.7601	0.7595
	yeast	0.0260	0.3711	0.3689	0.3503	0.3604	0.4038

classifier	dataset	Original	SMOTE	ADASYN	SLS	DBSMOTE	ANS2
KNN	cm1	0.8432	0.7647	0.7516	0.8381	0.7919	0.8067
	ecoli	0.6685	0.6986	0.6978	0.7084	0.6951	0.7063
	glass	0.2860	0.4145	0.3980	0.3563	0.3839	0.4371
	haberman	0.3367	0.4021	0.3992	0.4010	0.3860	0.4508
	jm1	0.3724	0.4425	0.4388	0.4381	0.4262	0.4891
	kc1	0.5362	0.5143	0.5045	0.5372	0.5370	0.5[illegible]21
	kc2	0.8808	0.8064	0.8016	0.8435	0.8376	0.7317
	letter	0.9977	0.9971	0.9977	0.9971	0.9974	0.9960
	optdigits	0.2624	0.3386	0.3340	0.3399	0.3279	0.3853
	pc1	0.6994	0.6230	0.6075	0.6501	0.6516	0.5978
	satimage	0.8757	0.8702	0.8634	0.8737	0.8767	0.8575
	segment	0.9091	0.8774	0.8736	0.8841	0.8817	0.8399
	vehicle	0.7348	0.6955	0.6773	0.7478	0.7255	0.7515
	yeast	0.1506	0.2901	0.2827	0.2626	0.2801	0.3257

Table 27: The comparison of F-measure from SMOTE with the default setting *k* as 5 and ANS.

classifier	dataset	SMOTEO-1	ANS1	SMOTEO-2	ANS2
DT	cm1	0.7394	0.7531	0.7412	0.7549
	ecoli	0.6601	0.6689	0.6788	0.6793
	glass	0.4764	0.5016	0.4984	0.5197
	haberman	0.3858	0.3593	0.4216	0.3985
	jm1	0.4218	0.4410	0.4575	0.4673
	kc1	0.5541	0.5615	0.5652	0.5737
	kc2	0.7512	0.7450	0.7590	0.7526
	letter	0.9671	0.9667	0.9674	0.9670
	optdigits	0.3258	0.3233	0.4111	0.3804
	pc1	0.5838	0.5783	0.5994	0.5918
	satimage	0.8772	0.8821	0.8795	0.8833
	segment	0.9184	0.9183	0.9190	0.9184
	vehicle	0.7822	0.7755	0.7869	0.7806
	yeast	0.2155	0.2663	0.2606	0.3019
NB	cm1	0.6712	0.7012	0.6801	0.7106
	ecoli	0.5619	0.5445	0.5789	0.5652
	glass	0.4396	0.4528	0.4625	0.4758
	haberman	0.4014	0.4075	0.4115	0.4186
	jm1	0.4503	0.4500	0.4721	0.4721
	kc1	0.5827	0.5814	0.5878	0.5855
	kc2	0.3283	0.3186	0.3556	0.3459
	letter	0.6864	0.6029	0.6911	0.6083
	optdigits	0.2179	0.2164	0.2464	0.2464
	pc1	0.2405	0.1574	0.2911	0.2133
	satimage	0.3525	0.3421	0.3602	0.3509
	segment	0.5937	0.5739	0.6002	0.5807
	vehicle	0.6785	0.6635	0.7046	0.6909
	yeast	0.2707	0.2796	0.3069	0.3147

classifier	dataset	SMOTEO-1	ANS1	SMOTEO-2	ANS2
MLP	cm1	0.8170	0.7853	0.8027	0.7967
	ecoli	0.6138	0.6068	0.6239	0.6205
	glass	0.4900	0.5025	0.5067	0.5182
	haberman	0.4288	0.4264	0.4414	0.4455
	jm1	0.4308	0.4400	0.4522	0.4663
	kc1	0.5471	0.5612	0.5651	0.5739
	kc2	0.2498	0.2473	0.2529	0.2585
	letter	0.9856	0.9840	0.9874	0.9880
	optdigits	0.2972	0.3075	0.3347	0.3558
	pc1	0.4457	0.4884	0.4598	0.4709
	satimage	0.6986	0.7047	0.6967	0.7157
	segment	0.9219	0.9204	0.9198	0.9243
	vehicle	0.7114	0.7177	0.7136	0.7183
	yeast	0.3484	0.3616	0.3993	0.3880
SVM	cm1	0.7939	0.7990	0.7939	0.7990
	ecoli	0.6268	0.6159	0.6385	0.6240
	glass	0.4859	0.4869	0.5083	0.5070
	haberman	0.4240	0.4266	0.4395	0.4475
	jm1	0.4375	0.4354	0.4532	0.4523
	kc1	0.5657	0.5701	0.5710	0.5751
	kc2	0.4752	0.4822	0.4771	0.4842
	letter	0.9931	0.9930	0.9934	0.9930
	optdigits	0.2857	0.2814	0.3138	0.3150
	pc1	0.6001	0.6150	0.6051	0.6206
	satimage	0.7234	0.7287	0.7257	0.7309
	segment	0.9400	0.9388	0.9401	0.9389
	vehicle	0.7523	0.7562	0.7557	0.7595
	yeast	0.3697	0.3802	0.3926	0.4038

classifier	dataset	SMOTEO-1	ANS1	SMOTEO-2	ANS2
KNN	cm1	0.8045	0.8096	0.8011	0.8067
	ecoli	0.7024	0.6951	0.7100	0.7063
	glass	0.3924	0.4146	0.4157	0.4371
	haberman	0.4080	0.4322	0.4376	0.4508
	jm1	0.4540	0.4733	0.4837	0.4891
	kc1	0.5511	0.5821	0.5616	0.5921
	kc2	0.8123	0.7778	0.7646	0.7317
	letter	0.9972	0.9971	0.9963	0.9960
	optdigits	0.3428	0.3534	0.4002	0.3853
	pc1	0.6325	0.6098	0.6127	0.5978
	satimage	0.8679	0.8696	0.8545	0.8575
	segment	0.8737	0.8613	0.8462	0.8399
	vehicle	0.7500	0.7505	0.7463	0.7515
	yeast	0.2979	0.3140	0.3186	0.3257

Table 28: The comparison with relocating safe-level SMOTE under geometric mean

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	RSLS
DT	cm1	0.7951	0.8732	0.8154	0.8396	0.8365	0.8354
	ecoli	0.7042	0.7253	0.7359	0.7240	0.7163	0.7436
	glass	0.3338	0.6379	0.6526	0.6269	0.6084	0.6300
	haberman	0.4529	0.5764	0.5722	0.5825	0.5277	0.5984
	jm1	0.5146	0.6270	0.6229	0.6079	0.5614	0.6324
	kc1	0.6426	0.7242	0.7278	0.7120	0.6925	0.7106
	kc2	0.8735	0.8710	0.8762	0.8615	0.8218	0.8647
	letter	0.9804	0.9828	0.9821	0.9831	0.9790	0.9829
	optdigits	0.4263	0.6291	0.6449	0.5570	0.5076	0.6162
	pc1	0.7184	0.7731	0.7684	0.7576	0.7466	0.7718
	satimage	0.9108	0.9459	0.9450	0.9387	0.9389	0.9398
	segment	0.9422	0.9475	0.9591	0.9507	0.9523	0.9494
	vehicle	0.8736	0.9145	0.9206	0.8912	0.8930	0.8869
	yeast	0.1757	0.5489	0.5582	0.4033	0.4602	0.4406
NB	cm1	0.7189	0.7917	0.7534	0.7284	0.6015	0.7242
	ecoli	0.4737	0.6196	0.5368	0.6134	0.5560	0.6321
	glass	0.4724	0.5329	0.6038	0.5919	0.5059	0.6112
	haberman	0.6408	0.6192	0.6279	0.6361	0.5855	0.6470
	jm1	0.6797	0.7011	0.7077	0.6792	0.6729	0.6995
	kc1	0.7639	0.7715	0.7616	0.7671	0.7724	0.7755
	kc2	0.4214	0.5158	0.3560	0.5164	0.4164	0.5441
	letter	0.7915	0.7333	0.2009	0.7326	0.3513	0.7349
	optdigits	0.5941	0.6375	0.6314	0.6059	0.6108	0.6598
	pc1	0.1705	0.4223	0.5292	0.3767	0.1549	0.4200
	satimage	0.5099	0.5793	0.5571	0.5744	0.4917	0.5833
	segment	0.6495	0.6966	0.5007	0.6856	0.5665	0.6899
	vehicle	0.5838	0.7895	0.8218	0.7578	0.5169	0.7822
	yeast	0.5597	0.6145	0.6046	0.5850	0.3032	0.6186

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	RSLS
MLP	cm1	0.8865	0.8911	0.8863	0.8789	0.8923	0.8868
	ecoli	0.6679	0.6637	0.6788	0.6792	0.7018	0.6899
	glass	0.5192	0.6559	0.6609	0.6377	0.6420	0.6516
	haberman	0.3735	0.6547	0.6537	0.6570	0.6020	[illegible]
	jm1	0.4801	0.6892	0.6842	0.6767	0.6482	0.7023
	kc1	0.6235	0.7448	0.7347	0.7215	0.7215	0.7445
	kc2	0.6201	0.8465	0.7833	0.8432	0.8197	0.8401
	letter	0.9909	0.9909	0.9897	0.9898	0.9878	0.9900
	optdigits	0.3747	0.7234	0.7040	0.6461	0.5637	0.7190
	pc1	0.5885	0.8255	0.8085	0.8356	0.7953	0.8354
	satimage	0.8315	0.9113	0.9029	0.9060	0.8723	0.9055
	segment	0.9565	0.9560	0.9479	0.9569	0.9489	0.9547
	vehicle	0.8538	0.9131	0.9107	0.9035	0.8990	0.9066
	yeast	0.3877	0.6998	0.6947	0.5978	0.5735	0.6525
SVM	cm1	0.8984	0.9226	0.9156	0.9097	0.9222	0.9083
	ecoli	0.6002	0.6427	0.6922	0.6851	0.7165	0.7013
	glass	0.2007	0.5477	0.6320	0.6209	0.5975	0.6548
	haberman	0.1458	0.5815	0.6358	0.6042	0.6066	0.6478
	jm1	0.3559	0.7081	0.7062	0.6897	0.6877	0.7185
	kc1	0.5677	0.7531	0.7516	0.7500	0.7538	0.7656
	kc2	0.5016	0.9327	0.9081	0.9280	0.8996	0.9306
	letter	0.9945	0.9945	0.9946	0.9945	0.9945	0.9948
	optdigits	0.0694	0.7314	0.7207	0.5653	0.6364	0.7284
	pc1	0.6735	0.8824	0.8780	0.8790	0.8588	0.8844
	satimage	0.7804	0.9265	0.9231	0.9255	0.8949	0.9269
	segment	0.9623	0.9706	0.9712	0.9694	0.9704	0.9695
	vehicle	0.8434	0.9209	0.9282	0.9100	0.9068	0.9125
	yeast	0.0584	0.7273	0.7230	0.6199	0.6661	0.7091

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	RSLS
KNN	cm1	0.8769	0.8942	0.8940	0.8995	0.8971	0.9120
	ecoli	0.7402	0.7679	0.7665	0.7775	0.7654	0.7803
	glass	0.4468	0.5743	0.5624	0.5264	0.5511	0.5664
	haberman	0.5022	0.6307	0.6280	0.6021	0.6067	0.6497
	jm1	0.5238	0.6860	0.6803	0.6318	0.6622	0.6942
	kc1	0.6767	0.7155	0.7079	0.7105	0.7277	0.7381
	kc2	0.9119	0.9760	0.9767	0.9615	0.9551	0.9709
	letter	0.9978	0.9985	0.9980	0.9985	0.9978	0.9987
	optdigits	0.4356	0.7087	0.7081	0.5749	0.6659	0.6942
	pc1	0.8210	0.9005	0.8993	0.8790	0.8860	0.8890
	satimage	0.9230	0.9536	0.9539	0.9440	0.9408	0.9512
	segment	0.9444	0.9465	0.9449	0.9454	0.9418	0.9374
	vehicle	0.8228	0.8824	0.8812	0.8688	0.8733	0.8933
	yeast	0.2979	0.6431	0.6349	0.4983	0.6045	0.5834

Table 29: The number of cases each technique achieves the average geometric mean in the ranking 1st -3rd

# of cases as	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	RSLS
1st	1	21	15	3	3	27
2nd	4	26	13	10	5	12
3rd	0	15	16	13	7	19
Total in 1st -3rd	5	62	44	26	15	58

Table 30: The comparison with adaptive neighbors SMOTE without minority outcast handling under geometric mean

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	ANS1
DT	cm1	0.7951	0.8732	0.8154	0.8396	0.8365	0.8504
	ecoli	0.7042	0.7253	0.7359	0.7240	0.7163	0.7414
	glass	0.3338	0.6379	0.6526	0.6269	0.6084	0.6508
	haberman	0.4529	0.5764	0.5722	0.5825	0.5277	0.5328
	jm1	0.5146	0.6270	0.6229	0.6079	0.5614	0.6430
	kc1	0.6426	0.7242	0.7278	0.7120	0.6925	0.7307
	kc2	0.8735	0.8710	0.8762	0.8615	0.8218	0.8703
	letter	0.9804	0.9828	0.9821	0.9831	0.9790	0.9833
	optdigits	0.4263	0.6291	0.6449	0.5570	0.5076	0.6068
	pc1	0.7184	0.7731	0.7684	0.7576	0.7466	0.7791
	satimage	0.9108	0.9459	0.9450	0.9387	0.9389	0.9476
	segment	0.9422	0.9475	0.9591	0.9507	0.9523	0.9539
	vehicle	0.8736	0.9145	0.9206	0.8912	0.8930	0.9012
	yeast	0.1757	0.5489	0.5582	0.4033	0.4602	0.5367
NB	cm1	0.7189	0.7917	0.7534	0.7284	0.6015	0.7702
	ecoli	0.4737	0.6196	0.5368	0.6134	0.5560	0.6164
	glass	0.4724	0.5329	0.6038	0.5919	0.5059	0.6046
	haberman	0.6408	0.6192	0.6279	0.6361	0.5855	0.6406
	jm1	0.6797	0.7011	0.7077	0.6792	0.6729	0.6761
	kc1	0.7639	0.7715	0.7616	0.7671	0.7724	0.7721
	kc2	0.4214	0.5158	0.3560	0.5164	0.4164	0.5009
	letter	0.7915	0.7333	0.2009	0.7326	0.3513	0.6648
	optdigits	0.5941	0.6375	0.6314	0.6059	0.6108	0.6158
	pc1	0.1705	0.4223	0.5292	0.3767	0.1549	0.2910
	satimage	0.5099	0.5793	0.5571	0.5744	0.4917	0.5576
	segment	0.6495	0.6966	0.5007	0.6856	0.5665	0.6611
	vehicle	0.5838	0.7895	0.8218	0.7578	0.5169	0.7326
	yeast	0.5597	0.6145	0.6046	0.5850	0.3032	0.5701

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	ANS1
MLP	cm1	0.8865	0.8911	0.8863	0.8789	0.8923	0.8780
	ecoli	0.6679	0.6637	0.6788	0.6792	0.7018	0.6668
	glass	0.5192	0.6559	0.6609	0.6377	0.6420	0.6557
	haberman	0.3735	0.6547	0.6537	0.6570	0.6020	0.6455
	jm1	0.4801	0.6892	0.6842	0.6767	0.6482	0.6725
	kc1	0.6235	0.7448	0.7347	0.7215	0.7215	0.7384
	kc2	0.6201	0.8465	0.7833	0.8432	0.8197	0.8406
	letter	0.9909	0.9909	0.9897	0.9898	0.9878	0.9900
	optdigits	0.3747	0.7234	0.7040	0.6461	0.5637	0.6621
	pc1	0.5885	0.8255	0.8085	0.8356	0.7953	0.8303
	satimage	0.8315	0.9113	0.9029	0.9060	0.8723	0.9035
	segment	0.9565	0.9560	0.9479	0.9569	0.9489	0.9563
	vehicle	0.8538	0.9131	0.9107	0.9035	0.8990	0.9016
	yeast	0.3877	0.6998	0.6947	0.5978	0.5735	0.6379
SVM	cm1	0.8984	0.9226	0.9156	0.9097	0.9222	0.9072
	ecoli	0.6002	0.6427	0.6922	0.6851	0.7165	0.6257
	glass	0.2007	0.5477	0.6320	0.6209	0.5975	0.6304
	haberman	0.1458	0.5815	0.6358	0.6042	0.6066	0.6394
	jm1	0.3559	0.7081	0.7062	0.6897	0.6877	0.7001
	kc1	0.5677	0.7531	0.7516	0.7500	0.7538	0.7598
	kc2	0.5016	0.9327	0.9081	0.9280	0.8996	0.9309
	letter	0.9945	0.9945	0.9946	0.9945	0.9945	0.9945
	optdigits	0.0694	0.7314	0.7207	0.5653	0.6364	0.6791
	pc1	0.6735	0.8824	0.8780	0.8790	0.8588	0.8775
	satimage	0.7804	0.9265	0.9231	0.9255	0.8949	0.9256
	segment	0.9623	0.9706	0.9712	0.9694	0.9704	0.9688
	vehicle	0.8434	0.9209	0.9282	0.9100	0.9068	0.9095
	yeast	0.0584	0.7273	0.7230	0.6199	0.6661	0.6938

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	ANS1
KNN	cm1	0.8769	0.8942	0.8940	0.8995	0.8971	0.8963
	ecoli	0.7402	0.7679	0.7665	0.7775	0.7654	0.7624
	glass	0.4468	0.5743	0.5624	0.5264	0.5511	0.5765
	haberman	0.5022	0.6307	0.6280	0.6021	0.6067	0.6464
	jm1	0.5238	0.6860	0.6803	0.6318	0.6622	0.7021
	kc1	0.6767	0.7155	0.7079	0.7105	0.7277	0.7589
	kc2	0.9119	0.9760	0.9767	0.9615	0.9551	0.9743
	letter	0.9978	0.9985	0.9980	0.9985	0.9978	0.9985
	optdigits	0.4356	0.7087	0.7081	0.5749	0.6659	0.6959
	pc1	0.8210	0.9005	0.8993	0.8790	0.8860	0.8936
	satimage	0.9230	0.9536	0.9539	0.9440	0.9408	0.9526
	segment	0.9444	0.9465	0.9449	0.9454	0.9418	0.9401
	vehicle	0.8228	0.8824	0.8812	0.8688	0.8733	0.8951
	yeast	0.2979	0.6431	0.6349	0.4983	0.6045	0.6321

Table 31: The number of cases each technique achieves the average geometric mean in the ranking 1st -3rd

# of cases as	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	ANS1
1st	2	26	16	8	4	14
2nd	3	25	17	8	4	13
3rd	2	10	16	15	3	24
Total in 1st -3rd	7	61	49	31	11	51

Table 32: The comparison with adaptive neighbors SMOTE with minority outcast handling under geometric mean

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	ANS2
DT	cm1	0.7951	0.8732	0.8154	0.8396	0.8365	0.8520
	ecoli	0.7042	0.7253	0.7359	0.7240	0.7163	0.7500
	glass	0.3338	0.6379	0.6526	0.6269	0.6084	0.6663
	haberman	0.4529	0.5764	0.5722	0.5825	0.5277	0.5678
	jm1	0.5146	0.6270	0.6229	0.6079	0.5614	0.6674
	kc1	0.6426	0.7242	0.7278	0.7120	0.6925	0.7418
	kc2	0.8735	0.8710	0.8762	0.8615	0.8218	0.8774
	letter	0.9804	0.9828	0.9821	0.9831	0.9790	0.9836
	optdigits	0.4263	0.6291	0.6449	0.5570	0.5076	0.6697
	pc1	0.7184	0.7731	0.7684	0.7576	0.7466	0.7919
	satimage	0.9108	0.9459	0.9450	0.9387	0.9389	0.9488
	segment	0.9422	0.9475	0.9591	0.9507	0.9523	0.9540
	vehicle	0.8736	0.9145	0.9206	0.8912	0.8930	0.9061
	yeast	0.1757	0.5489	0.5582	0.4033	0.4602	0.5801
NB	cm1	0.7189	0.7917	0.7534	0.7284	0.6015	0.7781
	ecoli	0.4737	0.6196	0.5368	0.6134	0.5560	0.6326
	glass	0.4724	0.5329	0.6038	0.5919	0.5059	0.6240
	haberman	0.6408	0.6192	0.6279	0.6361	0.5855	0.6516
	jm1	0.6797	0.7011	0.7077	0.6792	0.6729	0.6974
	kc1	0.7639	0.7715	0.7616	0.7671	0.7724	0.7761
	kc2	0.4214	0.5158	0.3560	0.5164	0.4164	0.5263
	letter	0.7911	0.7333	0.2009	0.7326	0.3513	0.6691
	optdigits	0.5941	0.6375	0.6314	0.6059	0.6108	0.6637
	pc1	0.1705	0.4223	0.5292	0.3767	0.1549	0.3454
	satimage	0.5099	0.5793	0.5571	0.5744	0.4917	0.5662
	segment	0.6495	0.6966	0.5007	0.6856	0.5665	0.6666
	vehicle	0.5838	0.7895	0.8218	0.7578	0.5169	0.7554
	yeast	0.5597	0.6145	0.6046	0.5850	0.3032	0.6111

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	ANS2
MLP	cm1	0.8865	0.8911	0.8863	0.8789	0.8923	0.8796
	ecoli	0.6679	0.6637	0.6788	0.6792	0.7018	0.6766
	glass	0.5192	0.6559	0.6609	0.6377	0.6420	0.6690
	haberman	0.3735	0.6547	0.6537	0.6570	0.6020	0.6644
	jm1	0.4801	0.6892	0.6842	0.6767	0.6482	0.6992
	kc1	0.6235	0.7448	0.7347	0.7215	0.7215	0.7504
	kc2	0.6201	0.8465	0.7833	0.8432	0.8197	0.8488
	letter	0.9909	0.9909	0.9897	0.9898	0.9878	0.9915
	optdigits	0.3747	0.7234	0.7040	0.6461	0.5637	0.7226
	pc1	0.5885	0.8255	0.8085	0.8356	0.7953	0.8282
	satimage	0.8315	0.9113	0.9029	0.9060	0.8723	0.9102
	segment	0.9565	0.9560	0.9479	0.9569	0.9489	0.9589
	vehicle	0.8538	0.9131	0.9107	0.9035	0.8990	0.9063
	yeast	0.3877	0.6998	0.6947	0.5978	0.5735	0.6703
SVM	cm1	0.8984	0.9226	0.9156	0.9097	0.9222	0.9072
	ecoli	0.6002	0.6427	0.6922	0.6851	0.7165	0.6317
	glass	0.2007	0.5477	0.6320	0.6209	0.5975	0.6474
	haberman	0.1458	0.5815	0.6358	0.6042	0.6066	0.6592
	jm1	0.3559	0.7081	0.7062	0.6897	0.6877	0.7113
	kc1	0.5677	0.7531	0.7516	0.7500	0.7538	0.7644
	kc2	0.5016	0.9327	0.9081	0.9280	0.8996	0.9334
	letter	0.9945	0.9945	0.9946	0.9945	0.9945	0.9945
	optdigits	0.0694	0.7314	0.7207	0.5653	0.6364	0.7255
	pc1	0.6735	0.8824	0.8780	0.8790	0.8588	0.8832
	satimage	0.7804	0.9265	0.9231	0.9255	0.8949	0.9278
	segment	0.9623	0.9706	0.9712	0.9694	0.9704	0.9690
	vehicle	0.8434	0.9209	0.9282	0.9100	0.9068	0.9127
	yeast	0.0584	0.7273	0.7230	0.6199	0.6661	0.7196

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	ANS2
KNN	cm1	0.8769	0.8942	0.8940	0.8995	0.8971	0.9101
	ecoli	0.7402	0.7679	0.7665	0.7775	0.7654	0.7676
	glass	0.4468	0.5743	0.5624	0.5264	0.5511	0.5944
	haberman	0.5022	0.6307	0.6280	0.6021	0.6067	0.6655
	jm1	0.5238	0.6860	0.6803	0.6318	0.6622	0.7230
	kc1	0.6767	0.7155	0.7079	0.7105	0.7277	0.7739
	kc2	0.9119	0.9760	0.9767	0.9615	0.9551	0.9770
	letter	0.9978	0.9985	0.9980	0.9985	0.9978	0.9984
	optdigits	0.4356	0.7087	0.7081	0.5749	0.6659	0.7595
	pc1	0.8210	0.9005	0.8993	0.8790	0.8860	0.8958
	satimage	0.9230	0.9536	0.9539	0.9440	0.9408	0.9570
	segment	0.9444	0.9465	0.9449	0.9454	0.9418	0.9294
	vehicle	0.8228	0.8824	0.8812	0.8688	0.8733	0.9093
	yeast	0.2979	0.6431	0.6349	0.4983	0.6045	0.6648

Table 33: The number of cases each technique achieves the average geometric mean in the ranking 1st -3rd

# of cases as	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	ANS2
1^{st}	1	14	8	4	3	40
2^{nd}	1	31	17	9	4	8
3^{rd}	4	15	21	14	4	12
Total in 1^{st} -3^{rd}	6	60	46	27	11	60

Table 34: The comparison with relocating safe-level SMOTE under adjusted geometric mean

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	RSLS
DT	cm1	0.8870	0.9256	0.8918	0.9099	0.9057	0.9081
	ecoli	0.7694	0.7414	0.7564	0.7573	0.7480	0.7650
	glass	0.3986	0.6437	0.6626	0.6601	0.6697	0.6594
	haberman	0.6572	0.6891	0.6865	0.6927	0.6801	0.7094
	jm1	0.7001	0.7335	0.7308	0.7331	0.7104	0.7512
	kc1	0.7603	0.7664	0.7630	0.7806	0.7642	0.7778
	kc2	0.9252	0.9217	0.9246	0.9166	0.8940	0.9185
	letter	0.9884	0.9890	0.9890	0.9896	0.9879	0.9893
	optdigits	0.6804	0.7656	0.7740	0.7440	0.7153	0.7778
	pc1	0.8341	0.8508	0.8472	0.8483	0.8423	0.8565
	satimage	0.9424	0.9579	0.9553	0.9559	0.9567	0.9559
	segment	0.9549	0.9547	0.9652	0.9585	0.9584	0.9572
	vehicle	0.9112	0.9284	0.9295	0.9192	0.9193	0.9178
	yeast	0.3091	0.6894	0.6906	0.6267	0.6579	0.6518
NB	cm1	0.8467	0.8864	0.8618	0.8558	0.7702	0.8521
	ecoli	0.6509	0.5847	0.6244	0.6078	0.6468	0.6136
	glass	0.6540	0.5140	0.6862	0.6787	0.6609	0.6478
	haberman	0.6718	0.5789	0.5939	0.6194	0.5263	0.6237
	jm1	0.7462	0.7461	0.7414	0.7469	0.7343	0.7592
	kc1	0.7675	0.7653	0.7465	0.7710	0.7689	0.7786
	kc2	0.6654	0.7159	0.6223	0.7164	0.6615	0.7322
	letter	0.8938	0.8628	0.5134	0.8624	0.6681	0.8636
	optdigits	0.6981	0.6737	0.6843	0.6728	0.7035	0.6903
	pc1	0.5692	0.6987	0.7487	0.6755	0.5418	0.6979
	satimage	0.7065	0.7123	0.7146	0.7132	0.7019	0.7183
	segment	0.7856	0.7945	0.7258	0.7921	0.7602	0.7952
	vehicle	0.7458	0.8673	0.8839	0.8486	0.7065	0.8634
	yeast	0.6965	0.7033	0.6862	0.7064	0.4744	0.7250

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	RSLS
MLP	cm1	0.9364	0.9325	0.9269	0.9317	0.9336	0.9370
	ecoli	0.7261	0.6176	0.6575	0.6622	0.7165	0.6682
	glass	0.6554	0.6423	0.6585	0.6684	0.6772	0.6740
	haberman	0.6284	0.6296	0.6357	0.6898	0.6726	0.6891
	jm1	0.6907	0.6755	0.6720	0.7342	0.7169	0.7394
	kc1	0.7533	0.7580	0.7476	0.7672	0.7655	0.7778
	kc2	0.7809	0.8201	0.7741	0.8239	0.8366	0.8137
	letter	0.9948	0.9947	0.9942	0.9943	0.9931	0.9945
	optdigits	0.6333	0.7623	0.7661	0.7490	0.7204	0.7778
	pc1	0.7627	0.7865	0.7513	0.8195	0.8222	0.8121
	satimage	0.8989	0.8944	0.8844	0.8963	0.9003	0.8934
	segment	0.9634	0.9561	0.9516	0.9600	0.9556	0.9574
	vehicle	0.9013	0.9110	0.8976	0.9122	0.9132	0.9149
	yeast	0.6257	0.7549	0.7540	0.7359	0.7166	0.7589
SVM	cm1	0.9444	0.9495	0.9433	0.9460	0.9505	0.9452
	ecoli	0.7332	0.5565	0.6528	0.6284	0.6841	0.6433
	glass	0.3784	0.5093	0.6942	0.6906	0.6289	0.6917
	haberman	0.5071	0.5170	0.6298	0.7059	0.5839	0.7168
	jm1	0.6297	0.6765	0.6729	0.7368	0.6985	0.7369
	kc1	0.7490	0.7464	0.7393	0.7714	0.7672	0.7753
	kc2	0.7139	0.9274	0.9053	0.9273	0.9207	0.9285
	letter	0.9970	0.9970	0.9970	0.9970	0.9970	0.9972
	optdigits	0.1777	0.7426	0.7520	0.7222	0.7263	0.7705
	pc1	0.8195	0.8735	0.8582	0.8830	0.8877	0.8858
	satimage	0.8789	0.9024	0.8972	0.9053	0.9109	0.9062
	segment	0.9722	0.9695	0.9676	0.9697	0.9714	0.9702
	vehicle	0.8971	0.9246	0.9172	0.9244	0.9237	0.9258
	yeast	0.1330	0.7570	0.7554	0.7458	0.7411	0.7697

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	RSLS
KNN	cm1	0.9334	0.9346	0.9333	0.9431	0.9384	0.9479
	ecoli	0.7953	0.7533	0.7593	0.7868	0.7732	0.7719
	glass	0.6066	0.5863	0.5958	0.6018	0.6033	0.6019
	haberman	0.6764	0.6611	0.6590	0.6932	0.6668	0.6988
	jm1	0.7092	0.7386	0.7370	0.7437	0.7313	0.7630
	kc1	0.7791	0.7376	0.7298	0.7682	0.7583	0.7672
	kc2	0.9485	0.9792	0.9793	0.9733	0.9697	0.9763
	letter	0.9989	0.9991	0.9990	0.9991	0.9988	0.9991
	optdigits	0.6803	0.7816	0.7791	0.7499	0.7688	0.8046
	pc1	0.8914	0.8942	0.8886	0.8974	0.8996	0.8934
	satimage	0.9504	0.9576	0.9560	0.9559	0.9557	0.9542
	segment	0.9522	0.9301	0.9277	0.9355	0.9345	0.9196
	vehicle	0.8836	0.9015	0.8970	0.9055	0.9037	0.9169
	yeast	0.5481	0.6970	0.6913	0.6923	0.6977	0.7183

Table 35: The number of cases each technique achieves the average adjusted geometric mean in the ranking 1st -3rd

# of cases as	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	RSLS
1^{st}	14	4	8	3	10	31
2^{nd}	3	15	9	25	8	10
3^{rd}	2	16	6	18	15	13
Total in 1^{st} -3^{rd}	19	35	23	46	33	54

Table 36: The comparison with adaptive neighbors SMOTE without minority outcast handling under adjusted geometric mean

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	ANS1
DT	cm1	0.8870	0.9256	0.8918	0.9099	0.9057	0.9148
	ecoli	0.7694	0.7414	0.7564	0.7573	0.7480	0.7387
	glass	0.3986	0.6437	0.6626	0.6601	0.6697	0.6799
	haberman	0.6572	0.6891	0.6865	0.6927	0.6801	0.6850
	jm1	0.7001	0.7335	0.7308	0.7331	0.7104	0.7445
	kc1	0.7603	0.7664	0.7630	0.7806	0.7642	0.7808
	kc2	0.9252	0.9217	0.9246	0.9166	0.8940	0.9209
	letter	0.9884	0.9890	0.9890	0.9896	0.9879	0.9895
	optdigits	0.6804	0.7656	0.7740	0.7440	0.7153	0.7550
	pc1	0.8341	0.8508	0.8472	0.8483	0.8423	0.8536
	satimage	0.9424	0.9579	0.9553	0.9559	0.9567	0.9589
	segment	0.9549	0.9547	0.9652	0.9585	0.9584	0.9569
	vehicle	0.9112	0.9284	0.9295	0.9192	0.9193	0.9240
	yeast	0.3091	0.6894	0.6906	0.6267	0.6579	0.6840
NB	cm1	0.8467	0.8864	0.8618	0.8558	0.7702	0.8757
	ecoli	0.6509	0.5847	0.6244	0.6078	0.6468	0.5968
	glass	0.6540	0.5140	0.6862	0.6787	0.6609	0.6487
	haberman	0.6718	0.5789	0.5939	0.6194	0.5263	0.6277
	jm1	0.7462	0.7461	0.7414	0.7469	0.7343	0.7468
	kc1	0.7675	0.7653	0.7465	0.7710	0.7689	0.7766
	kc2	0.6654	0.7159	0.6223	0.7164	0.6615	0.7080
	letter	0.8938	0.8628	0.5134	0.8624	0.6681	0.8278
	optdigits	0.6981	0.6737	0.6843	0.6728	0.7035	0.6902
	pc1	0.5692	0.6987	0.7487	0.6755	0.5418	0.6314
	satimage	0.7065	0.7123	0.7146	0.7132	0.7019	0.7109
	segment	0.7856	0.7945	0.7258	0.7921	0.7602	0.7935
	vehicle	0.7458	0.8673	0.8839	0.8486	0.7065	0.8338
	yeast	0.6965	0.7033	0.6862	0.7064	0.4744	0.7031

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	ANS1
MLP	cm1	0.9364	0.9325	0.9269	0.9317	0.9336	0.9288
	ecoli	0.7261	0.6176	0.6575	0.6622	0.7165	0.6277
	glass	0.6554	0.6423	0.6585	0.6684	0.6772	0.6655
	haberman	0.6284	0.6296	0.6357	0.6898	0.6726	0.6893
	jm1	0.6907	0.6755	0.6720	0.7342	0.7169	0.7397
	kc1	0.7533	0.7580	0.7476	0.7672	0.7655	0.7774
	kc2	0.7809	0.8201	0.7741	0.8239	0.8366	0.8166
	letter	0.9948	0.9947	0.9942	0.9943	0.9931	0.9942
	optdigits	0.6333	0.7623	0.7661	0.7490	0.7204	0.7589
	pc1	0.7627	0.7865	0.7513	0.8195	0.8222	0.8212
	satimage	0.8989	0.8944	0.8844	0.8963	0.9003	0.8928
	segment	0.9634	0.9561	0.9516	0.9600	0.9556	0.9576
	vehicle	0.9013	0.9110	0.8976	0.9122	0.9132	0.9132
	yeast	0.6257	0.7549	0.7540	0.7359	0.7166	0.7525
SVM	cm1	0.9444	0.9495	0.9433	0.9460	0.9505	0.9430
	ecoli	0.7332	0.5565	0.6528	0.6284	0.6841	0.5347
	glass	0.3784	0.5093	0.6942	0.6906	0.6289	0.6490
	haberman	0.5071	0.5170	0.6298	0.7059	0.5839	0.6964
	jm1	0.6297	0.6765	0.6729	0.7368	0.6985	0.7246
	kc1	0.7490	0.7464	0.7393	0.7714	0.7672	0.7730
	kc2	0.7139	0.9274	0.9053	0.9273	0.9207	0.9288
	letter	0.9970	0.9970	0.9970	0.9970	0.9970	0.9970
	optdigits	0.1777	0.7426	0.7520	0.7222	0.7263	0.7468
	pc1	0.8195	0.8735	0.8582	0.8830	0.8877	0.8868
	satimage	0.8789	0.9024	0.8972	0.9053	0.9109	0.9063
	segment	0.9722	0.9695	0.9676	0.9697	0.9714	0.9678
	vehicle	0.8971	0.9246	0.9172	0.9244	0.9237	0.9241
	yeast	0.1330	0.7570	0.7554	0.7458	0.7411	0.7646

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	ANS1
KNN	cm1	0.9334	0.9346	0.9333	0.9431	0.9384	0.9395
	ecoli	0.7953	0.7533	0.7593	0.7868	0.7732	0.7374
	glass	0.5066	0.5863	0.5958	0.6018	0.6033	0.6014
	haberman	0.6764	0.6611	0.6590	0.6932	0.6668	0.6982
	jm1	0.7092	0.7386	0.7370	0.7437	0.7313	0.7592
	kc1	0.7791	0.7376	0.7298	0.7682	0.7583	0.7879
	kc2	0.9485	0.9792	0.9793	0.9733	0.9697	0.9768
	letter	0.9989	0.9991	0.9990	0.9991	0.9988	0.9991
	optdigits	0.6803	0.7816	0.7791	0.7499	0.7688	0.7862
	pc1	0.8914	0.8942	0.8886	0.8974	0.8996	0.8883
	satimage	0.9504	0.9576	0.9560	0.9559	0.9557	0.9572
	segment	0.9522	0.9301	0.9277	0.9355	0.9345	0.9192
	vehicle	0.8836	0.9015	0.8970	0.9055	0.9037	0.9172
	yeast	0.5481	0.6970	0.6913	0.6923	0.6977	0.7216

Table 37: The number of cases each technique achieves the average adjusted geometric mean in the ranking 1st -3rd

# of cases as	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	ANS1
1^{st}	14	6	13	10	9	18
2^{nd}	3	20	2	20	9	16
3^{rd}	2	11	10	20	13	14
Total in 1^{st} -3^{rd}	19	37	25	50	31	48

Table 38: The comparison with adaptive neighbors SMOTE with minority outcast handling under adjusted geometric mean

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	ANS2
DT	cm1	0.8870	0.9256	0.8918	0.9099	0.9057	0.9157
	ecoli	0.7694	0.7414	0.7564	0.7573	0.7480	0.7432
	glass	0.3986	0.6437	0.6626	0.6601	0.6697	0.6893
	haberman	0.6572	0.6891	0.6865	0.6927	0.6801	0.7052
	jm1	0.7001	0.7335	0.7308	0.7331	0.7104	0.7584
	kc1	0.7603	0.7664	0.7630	0.7806	0.7642	0.7869
	kc2	0.9252	0.9217	0.9246	0.9166	0.8940	0.9250
	letter	0.9884	0.9890	0.9890	0.9896	0.9879	0.9897
	optdigits	0.6804	0.7656	0.7740	0.7440	0.7153	0.7883
	pc1	0.8341	0.8508	0.8472	0.8483	0.8423	0.8602
	satimage	0.9424	0.9579	0.9553	0.9559	0.9567	0.9595
	segment	0.9549	0.9547	0.9652	0.9585	0.9584	0.9570
	vehicle	0.9112	0.9284	0.9295	0.9192	0.9193	0.9270
	yeast	0.3091	0.6894	0.6906	0.6267	0.6579	0.7159
NB	cm1	0.8467	0.8864	0.8618	0.8558	0.7702	0.8798
	ecoli	0.6509	0.5847	0.6244	0.6078	0.6468	0.6052
	glass	0.6540	0.5140	0.6862	0.6787	0.6609	0.6605
	haberman	0.6718	0.5789	0.5939	0.6194	0.5263	0.6340
	jm1	0.7462	0.7461	0.7414	0.7469	0.7343	0.7589
	kc1	0.7675	0.7653	0.7465	0.7710	0.7689	0.7788
	kc2	0.6654	0.7159	0.6223	0.7164	0.6615	0.7226
	letter	0.8938	0.8628	0.5134	0.8624	0.6681	0.8300
	optdigits	0.6981	0.6737	0.6843	0.6728	0.7035	0.7155
	pc1	0.5692	0.6987	0.7487	0.6755	0.5418	0.6597
	satimage	0.7065	0.7123	0.7146	0.7132	0.7019	0.7154
	segment	0.7856	0.7945	0.7258	0.7921	0.7602	0.7965
	vehicle	0.7458	0.8673	0.8839	0.8486	0.7065	0.8476
	yeast	0.6965	0.7033	0.6862	0.7064	0.4744	0.7248

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	ANS2
MLP	cm1	0.9364	0.9325	0.9269	0.9317	0.9336	0.9304
	ecoli	0.7261	0.6176	0.6575	0.6622	0.7165	0.6309
	glass	0.6554	0.6423	0.6585	0.6684	0.6772	0.6684
	haberman	0.6284	0.6296	0.6357	0.6898	0.6726	0.6991
	jm1	0.6907	0.6755	0.6720	0.7342	0.7169	0.7539
	kc1	0.7533	0.7580	0.7476	0.7672	0.7655	0.7837
	kc2	0.7809	0.8201	0.7741	0.8239	0.8366	0.8262
	letter	0.9948	0.9947	0.9942	0.9943	0.9931	0.9953
	optdigits	0.6333	0.7623	0.7661	0.7490	0.7204	0.7923
	pc1	0.7627	0.7865	0.7513	0.8195	0.8222	0.8081
	satimage	0.8989	0.8944	0.8844	0.8963	0.9003	0.8980
	segment	0.9634	0.9561	0.9516	0.9600	0.9556	0.9598
	vehicle	0.9013	0.9110	0.8976	0.9122	0.9132	0.9151
	yeast	0.6257	0.7549	0.7540	0.7359	0.7166	0.7666
SVM	cm1	0.9444	0.9495	0.9433	0.9460	0.9505	0.9430
	ecoli	0.7332	0.5565	0.6528	0.6284	0.6841	0.5378
	glass	0.3784	0.5093	0.6942	0.6906	0.6289	0.6593
	haberman	0.5071	0.5170	0.6298	0.7059	0.5839	0.7078
	jm1	0.6297	0.6765	0.6729	0.7368	0.6985	0.7344
	kc1	0.7490	0.7464	0.7393	0.7714	0.7672	0.7756
	kc2	0.7139	0.9274	0.9053	0.9273	0.9207	0.9303
	letter	0.9970	0.9970	0.9970	0.9970	0.9970	0.9970
	optdigits	0.1777	0.7426	0.7520	0.7222	0.7263	0.7713
	pc1	0.8195	0.8735	0.8582	0.8830	0.8877	0.8898
	satimage	0.8789	0.9024	0.8972	0.9053	0.9109	0.9074
	segment	0.9722	0.9695	0.9676	0.9697	0.9714	0.9678
	vehicle	0.8971	0.9246	0.9172	0.9244	0.9237	0.9261
	yeast	0.1330	0.7570	0.7554	0.7458	0.7411	0.7782

classifier	dataset	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	ANS2
KNN	cm1	0.9334	0.9346	0.9333	0.9431	0.9384	0.9450
	ecoli	0.7953	0.7533	0.7593	0.7868	0.7732	0.7234
	glass	0.6066	0.5863	0.5958	0.6018	0.6033	0.6009
	haberman	0.6764	0.6611	0.6590	0.6932	0.6668	0.7069
	jm1	0.7092	0.7386	0.7370	0.7437	0.7313	0.7661
	kc1	0.7791	0.7376	0.7298	0.7682	0.7583	0.7898
	kc2	0.9485	0.9792	0.9793	0.9733	0.9697	0.9755
	letter	0.9989	0.9991	0.9990	0.9991	0.9988	0.9989
	optdigits	0.6803	0.7816	0.7791	0.7499	0.7688	0.8113
	pc1	0.8914	0.8942	0.8886	0.8974	0.8996	0.8844
	satimage	0.9504	0.9576	0.9560	0.9559	0.9557	0.9552
	segment	0.9522	0.9301	0.9277	0.9355	0.9345	0.9042
	vehicle	0.8836	0.9015	0.8970	0.9055	0.9037	0.9233
	yeast	0.5481	0.6970	0.6913	0.6923	0.6977	0.7262

Table 39: The number of cases each technique achieves the average adjusted geometric mean in the ranking 1st -3rd

# of cases as	ORIG	SMOTE	ADASYN	SLS	DBSMOTE	ANS2
1^{st}	13	3	8	2	7	37
2^{nd}	3	17	6	25	11	8
3^{rd}	3	14	11	22	13	7
Total in 1^{st} -3^{rd}	19	34	25	49	31	52

VITA

Wacharasak Siriseriwan was born on December 18th, 1983 in Bangkok, Thailand. He received a bachelor degree in Mathematics from Department of Mathematics, Faculty of Science, Chulalongkorn University in 2005. He also received a master degree in Computational Science from Department of Mathematics, Faculty of Science, Chulalongkorn University in 2009 which is changed later to Department of Mathematics and Computer Science. He has been financial supported by Development and Promotion of Science and Technology (DPST), Institute of the Promotion of Teaching Science and Technology (IPST).

ห้องสมุดคณะวิทยาศาสตร์ จุฬาฯ
FACULTY OF SCIENCE LIBRARY
CHULALONGKORN UNIVERSITY