
BIBLIOGRAPHY

- [1] A. K. Azlim Khan and N. H. Ahamed Hassain Malim. Comparative studies on re-sampling techniques in machine learning and deep learning models for drug-target interaction prediction. *Molecules*, 28(4), 2023. ISSN 1420-3049. doi: 10.3390/molecules28041663. URL <https://www.mdpi.com/1420-3049/28/4/1663>.
- [2] M. Bach and A. Werner. Improvement of random undersampling to avoid excessive removal of points from a given area of the majority class. In *International Conference on Computational Science*, pages 172–186. Springer, 2021.
- [3] M. Bagherian, E. Sabeti, K. Wang, M. A. Sartor, Z. Nikolovska-Coleska, and K. Najarian. Machine learning approaches and databases for prediction of drug–target interaction: a survey paper. *Briefings in bioinformatics*, 22(1):247–269, 2021.
- [4] M. R. Barkat, S. M. Moussa, and N. L. Badr. Drug-target interaction prediction using machine learning. In *2021 Tenth International Conference on Intelligent Computing and Information Systems (ICICIS)*, pages 480–485. IEEE, 2021.
- [5] N. Cahyana, S. Khomsah, and A. S. Aribowo. Improving imbalanced dataset classification using oversampling and gradient boosting. In *2019 5th International Conference on Science in Information Technology (ICSITech)*, pages 217–222. IEEE, 2019.
- [6] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357, 2002.
- [7] R. Chen, X. Liu, S. Jin, J. Lin, and J. Liu. Machine learning for drug-target interaction prediction. *Molecules*, 23(9):2208, 2018.
- [8] V. K. Chennuru and S. R. Timmappareddy. Simulated annealing based undersampling (saus): A hybrid multi-objective optimization method to tackle class imbalance. *Applied Intelligence*, 52(2):2092–2110, 2022.
- [9] D. Devi, S. K. Biswas, and B. Purkayastha. A review on solution to class imbalance problem: Undersampling approaches. In *2020 international conference on computational performance evaluation (ComPE)*, pages 626–631. IEEE, 2020.
- [10] A. Fernández, S. Garcia, F. Herrera, and N. V. Chawla. Smote for learning from imbalanced data: progress and challenges, marking the 15-year anniversary. *Journal of artificial intelligence research*, 61:863–905, 2018.
- [11] M. Hayaty, S. Muthmainah, and S. M. Ghufuran. Random and synthetic over-sampling approach to resolve data imbalance in classification. *International Journal of Artificial Intelligence Research*, 4(2):86–94, 2020.

-
- [12] H. He, Y. Bai, E. A. Garcia, and S. Li. Adasyn: Adaptive synthetic sampling approach for imbalanced learning. In *2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence)*, pages 1322–1328. Ieee, 2008.
- [13] H. He, Y. Bai, E. A. Garcia, and S. Li. Adasyn: Adaptive synthetic sampling approach for imbalanced learning. In *2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence)*, pages 1322–1328. Ieee, 2008.
- [14] R. Hendrawan, S. Al Faraby, et al. Multilabel classification of hate speech and abusive words on indonesian twitter social media. In *2020 International Conference on Data Science and Its Applications (ICoDSA)*, pages 1–7. IEEE, 2020.
- [15] Z. Jiang, T. Pan, C. Zhang, and J. Yang. A new oversampling method based on the classification contribution degree. *Symmetry*, 13(2):194, 2021.
- [16] P. Kaur and A. Gosain. Ff-smote: a metaheuristic approach to combat class imbalance in binary classification. *Applied Artificial Intelligence*, 33(5):420–439, 2019.
- [17] P. Kaur and A. Gosain. Ff-smote: a metaheuristic approach to combat class imbalance in binary classification. *Applied Artificial Intelligence*, 33(5):420–439, 2019.
- [18] H. Khojasteh, J. Pirgazi, and A. Ghanbari Sorkhi. Improving prediction of drug-target interactions based on fusing multiple features with data balancing and feature selection techniques. *Plos one*, 18(8):e0288173, 2023.
- [19] M. Kim and K.-B. Hwang. An empirical evaluation of sampling methods for the classification of imbalanced data. *PLoS one*, 17(7):e0271260, 2022.
- [20] S. H. Mahmud, W. Chen, H. Jahan, Y. Liu, N. I. Sujana, and S. Ahmed. idti-cssmoteb: identification of drug–target interaction based on drug chemical structure and protein sequence using xgboost with over-sampling technique smote. *IEEE Access*, 7:48699–48714, 2019.
- [21] S. H. Mahmud, W. Chen, H. Meng, H. Jahan, Y. Liu, and S. M. Hasan. Prediction of drug-target interaction based on protein features using undersampling and feature selection techniques with boosting. *Analytical biochemistry*, 589:113507, 2020.
- [22] R. Mohammed, J. Rawashdeh, and M. Abdullah. Machine learning with oversampling and undersampling techniques: overview study and experimental results. In *2020 11th international conference on information and communication systems (ICICS)*, pages 243–248. IEEE, 2020.

- [23] R. Mohammed, J. Rawashdeh, and M. Abdullah. Machine learning with oversampling and undersampling techniques: overview study and experimental results. In *2020 11th international conference on information and communication systems (ICICS)*, pages 243–248. IEEE, 2020.
- [24] H. D. Phan, K. Ellis, J. C. Barca, and A. Dorin. A survey of dynamic parameter setting methods for nature-inspired swarm intelligence algorithms. *Neural computing and applications*, 32(2):567–588, 2020.
- [25] S. Redkar, S. Mondal, A. Joseph, and K. Hareesha. A machine learning approach for drug-target interaction prediction using wrapper feature selection and class balancing. *Molecular informatics*, 39(5):1900062, 2020.
- [26] A. B. Shaik and S. Srinivasan. A brief survey on random forest ensembles in classification model. In *International Conference on Innovative Computing and Communications: Proceedings of ICICC 2018, Volume 2*, pages 253–260. Springer, 2019.
- [27] H. Shi, S. Liu, J. Chen, X. Li, Q. Ma, and B. Yu. Predicting drug-target interactions using lasso with random forest based on evolutionary information and chemical structure. *Genomics*, 111(6):1839–1852, 2019.
- [28] A. Tolios, J. De Las Rivas, E. Hovig, P. Trouillas, A. Scorilas, and T. Mohr. Computational approaches in cancer multidrug resistance research: Identification of potential biomarkers, drug targets and drug-target interactions. *Drug Resistance Updates*, 48: 100662, 2020.
- [29] T. Wongvorachan, S. He, and O. Bulut. A comparison of undersampling, oversampling, and smote methods for dealing with imbalanced classification in educational data mining. *Information*, 14(1):54, 2023.
- [30] T. Wongvorachan, S. He, and O. Bulut. A comparison of undersampling, oversampling, and smote methods for dealing with imbalanced classification in educational data mining. *Information*, 14(1):54, 2023.
- [31] Z. Xu, D. Shen, T. Nie, and Y. Kou. A hybrid sampling algorithm combining m-smote and enn based on random forest for medical imbalanced data. *Journal of Biomedical Informatics*, 107:103465, 2020.
- [32] Y. Yamanishi, M. Araki, A. Gutteridge, W. Honda, and M. Kanehisa. Prediction of drug–target interaction networks from the integration of chemical and genomic spaces. *Bioinformatics*, 24(13):i232–i240, 2008.
- [33] J. You, M. M. Islam, L. Grenier, Q. Kuang, R. D. McLeod, and P. Hu. Drug-target interaction network predictions for drug repurposing using lasso-based regularized linear

-
- classification model. In *Advances in Artificial Intelligence: 31st Canadian Conference on Artificial Intelligence, Canadian AI 2018, Toronto, ON, Canada, May 8–11, 2018, Proceedings 31*, pages 272–278. Springer, 2018.
- [34] H. Yu, G. Gu, H. Liu, J. Shen, and J. Zhao. A modified ant colony optimization algorithm for tumor marker gene selection. *Genomics, proteomics and bioinformatics*, 7(4):200–208, 2009.
- [35] H. Yu, J. Ni, and J. Zhao. Acosampling: An ant colony optimization-based under-sampling method for classifying imbalanced dna microarray data. *Neurocomputing*, 101:309–318, 2013.