

DAFTAR PUSTAKA

- Araf, I., Idri, A., Chairi, I. (2024). Cost-sensitive learning for imbalanced medical data: a review. *Artificial Intelligence Review*, 57, 80. <https://doi.org/10.1007/s10462-023-10652-8>
- Aggarwal, C. (2015). Data Classification: Advanced Concepts. In: *Data Mining*. Springer, Cham. https://doi.org/10.1007/978-3-319-14142-8_11
- Buda, M., Maki, A., Mazurowski, M.A. (2018). A systematic study of the class imbalance problem in Convolutional Neural Networks. *Neural Networks*, 106, 249–259. doi:10.1016/j.neunet.2018.07.011
- Cios, K., Swiniarski, R., Pedrycz, W., Kurgan, L. (2007). *A Knowledge Discovery Approach*. US: Springer.
- Collell, G., Prelec, D., Patil, K.R. (2018). A simple plug-in bagging ensemble based on threshold-moving for classifying binary and multiclass imbalanced data, *Neurocomputing* 275, 330-340. <https://doi.org/10.1016/j.neucom.2017.08.035>
- Ghaderi Zefrehi, H., Sheikhi, G., Altınçay, H. (2023). Threshold prediction for detecting rare positive samples using a meta-learner. *Pattern Analysis Applications*, 26, 289–306. <https://doi.org/10.1007/s10044-022-01103-1>
- Hofmann, H., Wickham, H., Kafadar, K. (2017). Letter-Value Plots: Boxplots for Large Data. *Journal of Computational and Graphical Statistics*, 26(3), 469–477. <https://doi.org/10.1080/10618600.2017.1305277>
- James, G., Witten, D., Hastie, T., Tibshirani, R. (2021). *An Introduction to Statistical Learning: with Applications in R. 2nd Edition*. US: Springer.
- Kriegel, HP. & Schubert, M. (2009). KDD Pipeline. In: LIU, L., ÖZSU, M.T. (eds) *Encyclopedia of Database Systems*. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-39940-9_1134

- Leevy, J.L., Johnson, J.M., Hancock, J., Khoshgoftarr, T.M. (2023). Threshold optimization and random undersampling for imbalanced credit card data. *Journal of Big Data*, 10, 58. <https://doi.org/10.1186/s40537-023-00738-z>
- Mulugeta, G., Zewotir, T., Tegegne, A.S., Juhar, L.H., Muleta, M.B. (2023). Classification of imbalanced data using machine learning algorithms to predict the risk of renal graft failures in Ethiopia. *BMC Medical Informatics and Decision Making*, 23, 98. <https://doi.org/10.1186/s12911-023-02185-5>
- Refaeilzadeh, P., Tang, L., Liu, H. (2009). Cross-Validation. In: LIU, L., ÖZSU, M.T. (eds) *Encyclopedia of Database Systems*. Springer, Boston, MA. https://doi.org/10.1007/978-0-387-39940-9_565
- Sarker, I.H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, 2, 160. <https://doi.org/10.1007/s42979-021-00592-x>
- Sheykhmousa, M., Mahdianpari, M., Ghanbari, H., Mohammadimanesh, F., Ghahmisi, P., Homayouni, S. (2020). Support vector machine versus random forest for remote sensing image classification: A meta-analysis and systematic review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 6308-6325.
- Tanha, J., Abdi, Y., Samadi, N., Razzaghi, N., Asadpour, M. (2020). Boosting methods for multi-class imbalanced data classification: an experimental review. *Journal of Big Data*, 7, 70. <https://doi.org/10.1186/s40537-020-00349-y>
- Tyralis, H. & Papacharalampous, G. (2021). Boosting algorithms in energy research: a systematic review. *Neural Computation & Applications* 33, 14101–14117. <https://doi.org/10.1007/s00521-021-05995-8>
- Yu, H., Mu, C., Sun, C., Yang, W., Yang, X., Zuo, X. (2015). Support vector machine-based optimized decision threshold adjustment strategy for classifying imbalanced data. *Knowledge-Based Systems*, 76, 67–78. <https://doi.org/10.1016/j.knosys.2014.12.007>

Zhang, H., & Singer, B. H. (2010). *Recursive partitioning and applications*. Springer Science & Business Media.

Zou, Q., Xie, S., Lin, Z., Wu, M., Ju, Y. (2016). Finding the Best Classification Threshold in Imbalanced Classification. *Big Data Research*, 5, 2–8. <https://doi.org/10.1016/j.bdr.2015.12.001>