

DAFTAR PUSTAKA

- Aggarwal, S. (2023). *Munich Personal RePEc Archive Machine Learning algorithms, perspectives, and real-world application: Empirical evidence from United States trade data*.
- Ahmed, G., Er, M. J., Fareed, M. M. S., Zikria, S., Mahmood, S., He, J., Asad, M., Jilani, S. F., & Aslam, M. (2022). DAD-Net: Classification of Alzheimer's Disease Using ADASYN Oversampling Technique and Optimized Neural Network. *Molecules*, 27(20), 7085. <https://doi.org/10.3390/molecules27207085>
- Alfakih, A. Y. (2018). *Euclidean Distance Matrices and Their Applications in Rigidity Theory*. Springer International Publishing. <https://doi.org/10.1007/978-3-319-97846-8>
- Alghamdi, B., & Alharby, F. (2019). An Intelligent Model for Online Recruitment Fraud Detection. *Journal of Information Security*, 10(03), 155–176. <https://doi.org/10.4236/jis.2019.103009>
- Ali, A., Ralescu, A., Shamsuddin, S. M., & Ralescu, A. L. (2013). Classification with class imbalance problem: A review. *Classification Int. J. Advance Soft Compu. Appl*, 5(3). <https://www.researchgate.net/publication/288228469>
- Biau, G., & Scornet, E. (2016). A Random Forest guided tour. *TEST*, 25(2), 197–227. <https://doi.org/10.1007/s11749-016-0481-7>
- Blockeel, H., Devos, L., Frénay, B., Nanfack, G., & Nijssen, S. (2023). Decision trees: from efficient prediction to responsible AI. *Frontiers in Artificial Intelligence*, 6. <https://doi.org/10.3389/frai.2023.1124553>
- Breiman, L. (2001). *Random Forests*. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Buttarelo, M. (2016). Laboratory diagnosis of anemia: are the old and new red cell parameters useful in classification and treatment, how? *International Journal of Laboratory Hematology*, 38(S1), 123–132. <https://doi.org/10.1111/ijlh.12500>
- Chen, W., Wang, H., Fei, M., Du, D., & Rakić, A. (2022). An Intrusion Detection Method Using ADASYN and Bayesian Optimized LightGBM. *Proceedings of the 34th Chinese Control and Decision Conference, CCDC 2022*, 4622–4627. <https://doi.org/10.1109/CCDC55256.2022.10033879>
- Chen, Z., Yu, W., & Zhou, L. (n.d.). *ADASYN-Random Forest Based Intrusion Detection Model*.
- Damarta, R., Hidayat, A., & Abdullah, A. S. (2021). The application of k-nearest neighbors classifier for sentiment analysis of PT PLN (Persero) twitter account service quality. *Journal of Physics: Conference Series*, 1722(1), 012002. <https://doi.org/10.1088/1742-6596/1722/1/012002>
- Decaro, C., Montanari, G. B., Bianconi, M., & Bellanca, G. (2021). Prediction of hematocrit through imbalanced dataset of blood spectra. *Healthcare Technology Letters*, 8(2), 37–44. <https://doi.org/10.1049/htl2.12006>
- Dewi, C., Chen, R.-C., Christanto, H. J., & Cauteruccio, F. (2023). Multinomial Naïve Bayes Classifier for Sentiment Analysis of Internet Movie Database. *Vietnam Journal of Computer Science*, 10(04), 485–498. <https://doi.org/10.1142/S2196888823500100>
- Dey, D., Haque, Md. S., Islam, Md. M., Aishi, U. I., Shammy, S. S., Mayen, Md. S. A., Noor, S. T. A., & Uddin, Md. J. (2025). The proper application of

- Logistic Regression* model in complex survey data: a systematic review. *BMC Medical Research Methodology*, 25(1), 15. <https://doi.org/10.1186/s12874-024-02454-5>
- Ditia Yosmita Praptiwi, Anang Kurnia, Anwar Fitrianto, & Fitrah Ernawati. (2024). *Random Forest* and CatBoost with Handling Imbalanced Class for Detection of Risk Factors Anemia in Children (5-12 Years). *International Journal of Scientific Research in Science, Engineering and Technology*, 11(3), 302–312. <https://doi.org/10.32628/IJSRSET24113134>
- Du, K.-L., Jiang, B., Lu, J., Hua, J., & Swamy, M. N. S. (2024). Exploring Kernel Machines and *Support Vector Machines*: Principles, Techniques, and Future Directions. *Mathematics*, 12(24), 3935. <https://doi.org/10.3390/math12243935>
- Fahmi, M. (2023). *PENANGANAN DATA TIDAK SEIMBANG MENGGUNAKAN SMOTE PADA MODEL KLASIFIKASI UNTUK DIAGNOSA ANEMIA*. Universitas Gadjah Mada.
- Gohari, K., Kazemnejad, A., Mohammadi, M., Eskandari, F., Saberi, S., Esmaili, M., & Sheidaei, A. (2023). A Bayesian latent class extension of naive Bayesian classifier and its application to the classification of gastric cancer patients. *BMC Medical Research Methodology*, 23(1), 190. <https://doi.org/10.1186/s12874-023-02013-4>
- Haibo He, Yang Bai, Garcia, E. A., & Shutao Li. (2008). ADASYN: Adaptive synthetic *sampling* approach for imbalanced learning. *2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, 1322–1328. <https://doi.org/10.1109/IJCNN.2008.4633969>
- Halabaku, E., & Bytyçi, E. (2024). Overfitting in Machine Learning: A Comparative Analysis of *Decision Trees* and *Random Forests*. *Intelligent Automation & Soft Computing*, 39(6), 987–1006. <https://doi.org/10.32604/iasc.2024.059429>
- Halder, R. K., Uddin, M. N., Uddin, Md. A., Aryal, S., & Khraisat, A. (2024). Enhancing K-nearest neighbor algorithm: a comprehensive review and performance analysis of modifications. *Journal of Big Data*, 11(1), 113. <https://doi.org/10.1186/s40537-024-00973-y>
- Harimoorthy, K., & Thangavelu, M. (2021). RETRACTED ARTICLE: Multi-disease prediction model using improved SVM-radial bias technique in healthcare monitoring system. *Journal of Ambient Intelligence and Humanized Computing*, 12(3), 3715–3723. <https://doi.org/10.1007/s12652-019-01652-0>
- Hu, Y.-J., Ku, T.-H., Jan, R.-H., Wang, K., Tseng, Y.-C., & Yang, S.-F. (2012). Decision tree-based learning to predict patient controlled analgesia consumption and readjustment. *BMC Medical Informatics and Decision Making*, 12(1), 131. <https://doi.org/10.1186/1472-6947-12-131>
- Jaiswal, M., Srivastava, A., & Siddiqui, T. J. (2019). *Machine Learning Algorithms for Anemia Disease Prediction* (pp. 463–469). https://doi.org/10.1007/978-981-13-2685-1_44
- Khan, J. R., Chowdhury, S., Islam, H., & Raheem, E. (2021). Machine Learning Algorithms To Predict The Childhood Anemia In Bangladesh. *Journal of Data Science*, 17(1), 195–218. [https://doi.org/10.6339/JDS.201901_17\(1\).0009](https://doi.org/10.6339/JDS.201901_17(1).0009)



- Kurniawati, Y. E., & Prabowo, Y. D. (2022). Model optimisation of class imbalanced learning using ensemble classifier on over-sampling data. *IAES International Journal of Artificial Intelligence (IJ-AI)*, 11(1), 276. <https://doi.org/10.11591/ijai.v11.i1.pp276-283>
- Lango, M., & Stefanowski, J. (2022). What makes multi-class imbalanced problems difficult? An experimental study. *Expert Systems with Applications*, 199, 116962. <https://doi.org/10.1016/j.eswa.2022.116962>
- Leboeuf, J.-S., LeBlanc, F., & Marchand, M. (2022). *Generalization Properties of Decision Trees on Real-valued and Categorical Features*. <http://arxiv.org/abs/2210.10781>
- Lee, D. T., & Plesa, M. L. (2022). Anemia. In *Family Medicine* (pp. 1815–1829). Springer International Publishing. https://doi.org/10.1007/978-3-030-54441-6_132
- Li, Y., Xu, W., Li, W., Li, A., & Liu, Z. (2021). Research on hybrid intrusion detection method based on the ADASYN and ID3 algorithms. *Mathematical Biosciences and Engineering*, 19(2), 2030–2042. <https://doi.org/10.3934/MBE.2022095>
- Matemilola, A. S., & Aliyu, S. (2024). Development of an enhanced naive bayes algorithm for fake news classification. *Science World Journal*, 19(2), 512–517. <https://doi.org/10.4314/swj.v19i2.28>
- Meng, L., Bai, B., Zhang, W., Liu, L., & Zhang, C. (2023). Research on a Decision Tree Classification Algorithm Based on Granular Matrices. *Electronics*, 12(21), 4470. <https://doi.org/10.3390/electronics12214470>
- Milovanovic, T., Dragasevic, S., Nikolic, A. N., Markovic, A. P., Lalošević, M. S., Popovic, D. D., & Krstic, M. N. (2022). Anemia as a Problem: GP Approach. *Digestive Diseases*, 40(3), 370–375. <https://doi.org/10.1159/000517579>
- Mittal, M., Kobielnik, M., Gupta, S., Cheng, X., & Wozniak, M. (2022). An efficient quality of services based wireless sensor network for anomaly detection using soft computing approaches. *Journal of Cloud Computing*, 11(1), 70. <https://doi.org/10.1186/s13677-022-00344-z>
- Munshi, R. M. (2024). Novel ensemble learning approach with SVM-imputed ADASYN features for enhanced cervical cancer prediction. *PLOS ONE*, 19(1), e0296107. <https://doi.org/10.1371/journal.pone.0296107>
- Nadiyah, N., Sitoayu, L., & Dewanti, L. P. (2022). REMAJA PUTRI PEDESAAN DI INDONESIA BERISIKO ANEMIA DUA KALI LEBIH TINGGI. *GIZI INDONESIA*, 45(1), 35–46. <https://doi.org/10.36457/gizindo.v45i1.614>
- Nie, F., Hao, Z., & Wang, R. (2024). *Multi-Class Support Vector Machine with Maximizing Minimum Margin*. www.aaai.org
- Ozturk Kiyak, E., Ghasemkhani, B., & Birant, D. (2023). High-Level K-Nearest Neighbors (HLKNN): A Supervised Machine Learning Model for Classification Analysis. *Electronics*, 12(18), 3828. <https://doi.org/10.3390/electronics12183828>
- Pal, A. (2021). Logistic Regression. *Cancer Research, Statistics, and Treatment*, 4(3), 551–554. https://doi.org/10.4103/crst.crst_164_21
- Peretz, O., Koren, M., & Koren, O. (2024). Naive Bayes classifier – An ensemble procedure for Recall and Precision enrichment. *Engineering Applications of Artificial Intelligence*, 136, 108972. <https://doi.org/10.1016/j.engappai.2024.108972>



- Rajaguru, H., & S R, S. C. (2019). Analysis of *Decision Tree* and K-Nearest Neighbor Algorithm in the Classification of Breast Cancer. *Asian Pacific Journal of Cancer Prevention*, 20(12), 3777–3781. <https://doi.org/10.31557/APJCP.2019.20.12.3777>
- Romano, M., Zammarchi, G., & Conversano, C. (2024). Iterative threshold-based Naïve bayes classifier. *Statistical Methods & Applications*, 33(1), 235–265. <https://doi.org/10.1007/s10260-023-00721-1>
- Sáez, J. A., Krawczyk, B., & Woźniak, M. (2016). Analyzing the *oversampling* of different classes and types of examples in multi-class imbalanced datasets. *Pattern Recognition*, 57, 164–178. <https://doi.org/10.1016/j.patcog.2016.03.012>
- Sandhya, I., & Muhasin, T. (2014). Study of RBC histogram in various anemias. *Journal of Evolution of Medical and Dental Sciences*, 3(74), 15521.
- Sasirega, D., & Krishapriya, V. (2024). Hybrid Feature Selection with Chaotic Rat Swarm Optimization-Based Convolutional Neural Network for Heart Disease Prediction from Imbalanced Datasets. *Data and Metadata*, 3, 262. <https://doi.org/10.56294/dm2024262>
- Satapathy, S. K., Mishra, S., Mallick, P. K., & Chae, G.-S. (2023). ADASYN and ABC-optimized RBF convergence network for classification of electroencephalograph signal. *Personal and Ubiquitous Computing*, 27(3), 1161–1177. <https://doi.org/10.1007/s00779-021-01533-4>
- Schipper, A., Rutten, M., van Gammeren, A., Hartevelde, C. L., Urrechaga, E., Weerkamp, F., den Besten, G., Krabbe, J., Slomp, J., Schoonen, L., Broeren, M., van Wijnen, M., Huijskens, M. J. A. J., Koopmann, T., van Ginneken, B., Kusters, R., & Kurstjens, S. (2024). Machine Learning-Based Prediction of Hemoglobinopathies Using Complete Blood Count Data. *Clinical Chemistry*. <https://doi.org/10.1093/clinchem/hvae081>
- Schweta, N., & Pande, S. D. (2023). Prediction of Anemia using various Ensemble Learning and *Boosting* Techniques. *EAI Endorsed Transactions on Pervasive Health and Technology*, 9. <https://doi.org/10.4108/eetpht.9.4197>
- Siswanto, T., & Ernawati, M. (2023). Parity and anemia among pregnant women at a public health center in Bojonegoro, Indonesia. *Majalah Obstetri & Ginekologi*, 31(3), 110–116. <https://doi.org/10.20473/mog.V31I32023.110-116>
- Sleeman IV, W. C., & Krawczyk, B. (2021). Multi-class imbalanced big data classification on Spark. *Knowledge-Based Systems*, 212, 106598. <https://doi.org/10.1016/j.knosys.2020.106598>
- Suárez, C. A., Castro, M., Leon, M., Martin-Barreiro, C., & Liut, M. (2025). Improving SVM performance through data reduction and misclassification analysis with linear programming. *Complex & Intelligent Systems*, 11(8), 356. <https://doi.org/10.1007/s40747-025-01989-4>
- Sunardi, D., Bardosono, S., Basrowi, R. W., Wasito, E., & Vandenplas, Y. (2021). Dietary Determinants of Anemia in Children Aged 6–36 Months: A Cross-Sectional Study in Indonesia. *Nutrients*, 13(7), 2397. <https://doi.org/10.3390/nu13072397>
- Tahirovic, E., & Krivic, S. (2023). Interpretability and Explain ability of *Logistic Regression* Model for Breast Cancer Detection. *International Conference on*

- Agents and Artificial Intelligence*, 3, 161–168.
<https://doi.org/10.5220/0011627600003393>
- Tesfaye, S. H., Seboka, B. T., & Sisay, D. (2024). Application of *Machine Learning* methods for predicting childhood anaemia: Analysis of Ethiopian Demographic Health Survey of 2016. *PLOS ONE*, 19(4), e0300172.
<https://doi.org/10.1371/journal.pone.0300172>
- van den Goorbergh, R., van Smeden, M., Timmerman, D., & Van Calster, B. (2022). The harm of class imbalance corrections for risk prediction models: illustration and simulation using *Logistic Regression*. *Journal of the American Medical Informatics Association*, 29(9), 1525–1534.
<https://doi.org/10.1093/jamia/ocac093>
- Vos, T., Flaxman, A. D., Naghavi, M., Lozano, R., Michaud, C., Ezzati, M., Shibuya, K., Salomon, J. A., Abdalla, S., Aboyans, V., Abraham, J., Ackerman, I., Aggarwal, R., Ahn, S. Y., Ali, M. K., AlMazroa, M. A., Alvarado, M., Anderson, H. R., Anderson, L. M., ... Murray, C. J. (2012). Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, 380(9859), 2163–2196.
[https://doi.org/10.1016/S0140-6736\(12\)61729-2](https://doi.org/10.1016/S0140-6736(12)61729-2)
- Walker HK, Hall WD, & Hurst JW. (1990). Chapter 152 Red Cell Indices. In Walker HK, Hall WD, & Hurst JW (Eds.), *Clinical Methods: The History, Physical, and Laboratory Examinations*. 3rd edition. Butterworths.
- Wang, S., Ren, J., Lian, X., Bai, R., & Jiang, X. (2022). *Boosting the Discriminant Power of Naive Bayes*. <https://doi.org/10.1109/ICPR56361.2022.9956358>
- Wang, X., Wang, S., Du, Y., & Huang, Z. (2023). Improved large margin classifier via bounding hyperellipsoid. *Information Sciences*, 648, 119597.
<https://doi.org/10.1016/j.ins.2023.119597>
- Wang, Y., Rosli, M. M., Musa, N., & Li, F. (2024a). Multi-Class Imbalanced Data Classification: A Systematic Mapping Study. *Engineering, Technology & Applied Science Research*, 14(3), 14183–14190.
<https://doi.org/10.48084/etasr.7206>
- Wang, Y., Rosli, M. M., Musa, N., & Li, F. (2024b). Multi-Class Imbalanced Data Classification: A Systematic Mapping Study. *Engineering, Technology & Applied Science Research*, 14(3), 14183–14190.
<https://doi.org/10.48084/etasr.7206>
- WHO. (2017). *WHO I Nutritional anaemias: tools for effective prevention and control NUTRITIONAL ANAEMIAS: TOOLS FOR EFFECTIVE PREVENTION AND CONTROL*.
- Yang, X., Kuang, Q., Zhang, W., & Zhang, G. (2018). AMDO: An Over-Sampling Technique for Multi-Class Imbalanced Problems. *IEEE Transactions on Knowledge and Data Engineering*, 30(9), 1672–1685.
<https://doi.org/10.1109/TKDE.2017.2761347>
- Zhang, S., Li, J., & Li, Y. (2023). Reachable Distance Function for KNN Classification. *IEEE Transactions on Knowledge and Data Engineering*, 35(7), 7382–7396. <https://doi.org/10.1109/TKDE.2022.3185149>
- Zhu, B., Jing, X., Qiu, L., & Li, R. (2024). An Imbalanced Data Classification Method Based on Hybrid Resampling and Fine Cost Sensitive Support Vector



UNIVERSITAS
GADJAH MADA

Machine. Computers, Materials & Continua, 79(3), 3977–3999.
<https://doi.org/10.32604/cmc.2024.048062>



UNIVERSITAS
GADJAH MADA

ADASYN Termodifikasi (Modified) Untuk Multiclass Pada Kasus Klasifikasi Penyakit Anemia
Velssy, Afiahayati, S.Kom., M.Cs., Ph.D ; Dr. Sri Mulyana., M.Kom
Universitas Gadjah Mada, 2026 | Diunduh dari <http://etd.repository.ugm.ac.id/>

LAMPIRAN