

DAFTAR PUSTAKA

- [1] M. Cannistrà, C. Masci, F. Ieva, T. Agasisti, and A. M. Paganoni, "Early-predicting dropout of university students: an application of innovative multilevel machine learning and statistical techniques," *Stud. High. Educ.*, vol. 47, no. 9, pp. 1935–1956, 2022, doi: 10.1080/03075079.2021.2018415.
- [2] V. Realinho, J. Machado, L. Baptista, and M. V. Martins, "Predicting Student Dropout and Academic Success," *Data*, vol. 7, no. 11, p. 146, 2022, doi: 10.3390/data7110146.
- [3] S. C. Matz, C. S. Bukow, H. Peters, C. Deacons, and C. Stachl, "Using machine learning to predict student retention from socio-demographic characteristics and app-based engagement metrics," *Sci. Rep.*, vol. 13, no. 1, 2023, doi: 10.1038/s41598-023-32484-w.
- [4] J. Mangundu, "Higher education institutional innovation: An analysis of challenges to e-academic advising during emergency remote teaching," *SA J. Inf. Manag.*, vol. 24, no. 1, pp. 1–9, 2022, doi: 10.4102/sajim.v24i1.1569.
- [5] PDDikti Kemendikbud, "Statistik Pendidikan Tinggi 2021 (Final)," Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi. [Online]. Available: <https://pddikti.kemdikbud.go.id/>
- [6] Unesco, "Unesco Global Education Monitoring report." [Online]. Available: <https://www.unesco.org/>
- [7] Badan Pusat Statistik, "Statistik Pendidikan Tinggi 2022. Jakarta: BPS RI," <https://www.bps.go.id>. [Online]. Available: <https://pddikti.kemdikbud.go.id/>
- [8] B. J. Mathew and O. A. Ibrahim, "Effective advising: How does academic advising influence student learning outcomes in higher education institutions in Oman?," *Cogent Educ.*, vol. 10, no. 1, pp. 1–19, 2023, doi: 10.1080/2331186X.2023.2197663.
- [9] D. Pembelajaran and D. Tahun, "Direktorat Pembelajaran dan Kemahasiswaan DIKTI Tahun 2021." [Online]. Available: <https://dikti.kemdikbud.go.id/>
- [10] M. Roy, I. Prasetyawan, and R. I. Desanti, "Predicting Student Academic Success Using XGBoost and Optuna Tuning Based on Student Attendance," *J. Appl. Data Sci.*, vol. 5, no. 3, pp. 1–15, 2024.
- [11] A. F. A. Naibaho and A. Zahra, "Prediksi Kelulusan Siswa Sekolah Menengah Pertama Menggunakan Machine Learning," *J. Inform. dan Tek. Elektro Terap.*, vol. 11, no. 3, 2023, doi: 10.23960/jitet.v11i3.3056.
- [12] J. Jović, E. Kisić, M. R. Milić, D. Domazet, and K. Chandra, "Prediction of student academic performance using machine learning algorithms," *CEUR Workshop Proc.*, vol. 3454, pp. 31–39, 2022.

- [13] Y. Liu, T. Yang, L. Tian, B. Huang, J. Yang, and Z. Zeng, "Ada-XG-CatBoost: A Combined Forecasting Model for Gross Ecosystem Product (GEP) Prediction," 2024. doi: 10.3390/su16167203.
- [14] N. Nguyen and D. Ngo, "Comparative analysis of boosting algorithms for predicting personal default," *Cogent Econ. Financ.*, vol. 13, no. 1, p., 2025, doi: 10.1080/23322039.2025.2465971.
- [15] T. Wongvorachan, S. He, and O. Bulut, "A Comparison of Undersampling, Oversampling, and SMOTE Methods for Dealing with Imbalanced Classification in Educational Data Mining," 2023. doi: 10.3390/info14010054.
- [16] G. Ramaswami, T. Susnjak, and A. Mathrani, "On Developing Generic Models for Predicting Student Outcomes in Educational Data Mining," *Big Data Cogn. Comput.*, vol. 6, no. 1, 2022, doi: 10.3390/bdcc6010006.
- [17] S. D. A. Bujang *et al.*, "Multiclass Prediction Model for Student Grade Prediction Using Machine Learning," *IEEE Access*, vol. 9, pp. 95608–95621, 2021, doi: 10.1109/ACCESS.2021.3093563.
- [18] L. Vives *et al.*, "Prediction of Students' Academic Performance in the Programming Fundamentals Course Using Long Short-Term Memory Neural Networks," *IEEE Access*, vol. 12, pp. 5882–5898, 2024, doi: 10.1109/ACCESS.2024.3350169.
- [19] S. Albahli, "Efficient hyperparameter tuning for predicting student performance with Bayesian optimization," *Multimed. Tools Appl.*, vol. 83, no. 17, pp. 52711–52735, 2024, doi: 10.1007/s11042-023-17525-w.
- [20] J. Lu *et al.*, "Machine learning analysis of factors affecting college students' academic performance," *Front. Psychol.*, vol. 15, 2024, doi: 10.3389/fpsyg.2024.1447825.
- [21] S. Poudyal, M. J. Mohammadi-Aragh, and J. E. Ball, "Prediction of Student Academic Performance Using a Hybrid 2D CNN Model," 2022. doi: 10.3390/electronics11071005.
- [22] D. Wang, D. Lian, Y. Xing, S. Dong, X. Sun, and J. Yu, "Analysis and prediction of influencing factors of college student achievement based on machine learning," *Front. Psychol.*, vol. 13, p. 881859, 2022.
- [23] R. M. Sagala, "Prediksi Kelulusan Mahasiswa Menggunakan Data Mining Algoritma K-Means," *TeIka*, vol. 11, no. 2, pp. 131–142, 2021, doi: 10.36342/teika.v11i2.2610.
- [24] H. Sakız, F. Özdaş, I. Göksu, and A. Ekinci, "A Longitudinal Analysis of Academic Achievement and Its Correlates in Higher Education," *SAGE Open*, vol. 11, 2021, doi: 10.1177/21582440211003085.
- [25] Z. Cai and Q. Meng, "Academic resilience and academic performance of university students: the mediating role of teacher support," *Front. Psychol.*,

vol. 16, 2025, doi: 10.3389/fpsyg.2025.1463643.

- [26] R. Gilar-Corbí, T. Pozo-Rico, J. Castejón, T. Sanchez, I. Sandoval-Palis, and J. Vidal, "Academic Achievement and Failure in University Studies: Motivational and Emotional Factors," *Sustainability*, 2020, doi: 10.3390/su12239798.
- [27] R. Abdelrahman, "Metacognitive awareness and academic motivation and their impact on academic achievement of Ajman University students," *Heliyon*, vol. 6, 2020, doi: 10.1016/j.heliyon.2020.e04192.
- [28] E. Acosta-Gonzaga and A. Ramírez-Arellano, "The Influence of Motivation, Emotions, Cognition, and Metacognition on Students' Learning Performance: A Comparative Study in Higher Education in Blended and Traditional Contexts," *SAGE Open*, vol. 11, 2021, doi: 10.1177/21582440211027561.
- [29] X. Gu, N. C. Hassan, and T. Sulaiman, "The Relationship between Family Factors and Academic Achievement of Junior High School Students in Rural China: Mediation Effect of Parental Involvement," *Behav. Sci. (Basel)*, vol. 14, 2024, doi: 10.3390/bs14030221.
- [30] M. R. Siddiky and I. E. Haque, "Factors Affecting Students' Academic Performance Mediated by their Motivation for Learning," *Asian J. Univ. Educ.*, vol. 20, no. 1, pp. 15–27, 2024, doi: 10.24191/ajue.v20i1.25696.
- [31] D. Urhahne and L. Wijnia, "Theories of Motivation in Education: an Integrative Framework," *Educ. Psychol. Rev.*, vol. 35, pp. 1–35, 2023, doi: 10.1007/s10648-023-09767-9.
- [32] J. L. Howard, J. S. Bureau, F. Guay, J. X. Y. Chong, and R. M. Ryan, "Student Motivation and Associated Outcomes: A Meta-Analysis From Self-Determination Theory," *Perspect. Psychol. Sci.*, vol. 16, pp. 1300–1323, 2021, doi: 10.1177/1745691620966789.
- [33] W. Hong, R. Zhen, R. Liu, M.-T. Wang, Y. Ding, and J. Wang, "The longitudinal linkages among Chinese children's behavioural, cognitive, and emotional engagement within a mathematics context," *Educ. Psychol.*, vol. 40, pp. 666–680, 2020, doi: 10.1080/01443410.2020.1719981.
- [34] J. Kittur, J. Bekki, and S. R. Brunhaver, "Development of a student engagement score for online undergraduate engineering courses using learning management system interaction data," *Comput. Appl. Eng. Educ.*, vol. 30, pp. 661–677, 2021, doi: 10.1002/cae.22479.
- [35] A. Behr, M. Giese, H. T. K., and K. Theune, "Early Prediction of University Dropouts – A Random Forest Approach," *Jahrb. Natl. Okon. Stat.*, vol. 240, pp. 743–789, 2020, doi: 10.1515/jbnst-2019-0006.
- [36] C. Romero and S. Ventura, "Educational data mining and learning analytics: An updated survey," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 10, no. 3, 2020, doi: 10.1002/widm.1355.

- [37] S. Batool, J. Rashid, M. W. Nisar, J. Kim, H. Y. Kwon, and A. Hussain, "Educational data mining to predict students' academic performance: A survey study," *Educ. Inf. Technol.*, vol. 28, no. 1, pp. 905–971, 2023, doi: 10.1007/s10639-022-11152-y.
- [38] Y. Lin, H. Chen, W. Xia, F. Lin, Z. Wang, and Y. Liu, "A Comprehensive Survey on Deep Learning Techniques in Educational Data Mining," *Data Sci. Eng.*, vol. abs/2309.0, 2025, doi: 10.1007/s41019-025-00303-z.
- [39] A. Kord, A. Aboelfetouh, and S. M. Shohieb, "Academic course planning recommendation and students' performance prediction multi-modal based on educational data mining techniques," *J. Comput. High. Educ.*, 2025, doi: 10.1007/s12528-024-09426-0.
- [40] C. Fischer *et al.*, "Mining Big Data in Education: Affordances and Challenges," *Rev. Res. Educ.*, vol. 44, no. 1, pp. 130–160, 2020, doi: 10.3102/0091732X20903304.
- [41] D. S. PM, "CRISP-DM is Still the Most Popular Framework for Executing Data Science Projects," Data Science PM.
- [42] C. Schröer, F. Kruse, and J. M. Gómez, "A systematic literature review on applying CRISP-DM process model," *Procedia Comput. Sci.*, vol. 181, pp. 526–534, 2021, doi: 10.1016/j.procs.2021.01.199.
- [43] K. Okoye, J. T. Nganji, J. Escamilla, and S. Hosseini, "Machine learning model (RG-DMML) and ensemble algorithm for prediction of students' retention and graduation in education," *Comput. Educ. Artif. Intell.*, vol. 6, 2024, doi: 10.1016/j.caeai.2024.100205.
- [44] F. Martinez-Plumed *et al.*, "CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories," *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 8, pp. 3048–3061, 2021, doi: 10.1109/TKDE.2019.2962680.
- [45] V. Plotnikova, M. Dumas, and F. Milani, "Adaptations of data mining methodologies: A systematic literature review," *PeerJ Comput. Sci.*, vol. 6, pp. 1–43, 2020, doi: 10.7717/PEERJ-CS.267.
- [46] J. S. Saltz, "CRISP-DM for Data Science: Strengths, Weaknesses and Potential Next Steps," 2021. doi: 10.1109/BigData52589.2021.9671634.
- [47] W. Y. Ayele, "Adapting CRISP-DM for idea mining a data mining process for generating ideas using a textual dataset," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 6, pp. 20–32, 2020, doi: 10.14569/IJACSA.2020.0110603.
- [48] M. Akhtar, M. Tanveer, and M. Arshad, "Advancing Supervised Learning with the Wave Loss Function: A Robust and Smooth Approach," *Pattern Recognit.*, vol. 155, 2024, doi: 10.1016/j.patcog.2024.110637.
- [49] A. J. Vanhove, B. Z. Graham, T. Titareva, and A. Udomvisawakul, "Classification Performance of Supervised Machine Learning to Predict

- Human Resource Management Outcomes: A Meta-Analysis Using Cross-Classified Multilevel Modeling,” *Hum. Resour. Manage.*, 2025, doi: 10.1002/hrm.70012.
- [50] V. Muthukumar, A. Narang, V. Subramanian, M. Belkin, D. Hsu, and A. Sahai, “Classification vs regression in overparameterized regimes: Does the loss function matter?,” *J. Mach. Learn. Res.*, vol. 22, pp. 222:1-222:69, 2021.
 - [51] R. Graf, M. Zeldovich, and S. Friedrich, “Comparing linear discriminant analysis and supervised learning algorithms for binary classification—A method comparison study,” *Biometrical J.*, vol. 66, no. 1, 2024, doi: 10.1002/bimj.202200098.
 - [52] M. Akhtar, M. Tanveer, and M. Arshad, “RoBoSS: A Robust, Bounded, Sparse, and Smooth Loss Function for Supervised Learning,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 47, no. 1, pp. 149–160, 2025, doi: 10.1109/TPAMI.2024.3465535.
 - [53] H. Zhao, C. Zhang, B. Xue, M. Zhang, and B. Zhang, “A Two-Stage Differential Evolutionary Algorithm for Deep Ensemble Model Generation,” *IEEE Trans. Evol. Comput.*, vol. 28, no. 3, pp. 810–824, 2024, doi: 10.1109/TEVC.2022.3231387.
 - [54] M. Arifuzzaman, M. R. Hasan, T. J. Toma, S. B. Hassan, and A. K. Paul, “An Advanced Decision Tree-Based Deep Neural Network in Nonlinear Data Classification,” *Technologies*, vol. 11, no. 1, 2023, doi: 10.3390/technologies11010024.
 - [55] M. V. C. Aragão *et al.*, “A practical evaluation of AutoML tools for binary, multiclass, and multilabel classification,” *Sci. Rep.*, vol. 15, no. 1, 2025, doi: 10.1038/s41598-025-02149-x.
 - [56] J. Pachouly, S. Ahirrao, K. Kotecha, A. Kulkarni, and S. Alfarhood, “Multilabel classification for defect prediction in software engineering,” *Sci. Rep.*, vol. 15, no. 1, 2025, doi: 10.1038/s41598-025-93242-8.
 - [57] S. Farhadpour, T. A. Warner, and A. E. Maxwell, “Selecting and Interpreting Multiclass Loss and Accuracy Assessment Metrics for Classifications with Class Imbalance: Guidance and Best Practices,” *Remote Sens.*, vol. 16, no. 3, p. 533, 2024, doi: 10.3390/rs16030533.
 - [58] O. Rainio, J. Teuho, and R. Klén, “Evaluation metrics and statistical tests for machine learning,” *Sci. Rep.*, vol. 14, no. 1, 2024, doi: 10.1038/s41598-024-56706-x.
 - [59] S. Malik *et al.*, “Advancing educational data mining for enhanced student performance prediction: a fusion of feature selection algorithms and classification techniques with dynamic feature ensemble evolution,” *Sci. Rep.*, vol. 15, no. 1, 2025, doi: 10.1038/s41598-025-92324-x.
 - [60] I. D. Mienye and N. Jere, “A Survey of Decision Trees: Concepts,

- Algorithms, and Applications,” *IEEE Access*, vol. 12, pp. 86716–86727, 2024, doi: 10.1109/ACCESS.2024.3416838.
- [61] E. Agliari, F. Alemanno, M. Aquaro, and A. Fachechi, “Regularization, early-stopping and dreaming: A Hopfield-like setup to address generalization and overfitting,” *Neural Networks*, vol. 177, p. 106389, 2024, doi: 10.1016/j.neunet.2024.106389.
 - [62] M. Mujahid *et al.*, “Data oversampling and imbalanced datasets: an investigation of performance for machine learning and feature engineering,” *J. Big Data*, vol. 11, no. 1, pp. 1–32, 2024, doi: 10.1186/s40537-024-00943-4.
 - [63] D. Chicco and G. Jurman, “The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification,” *BioData Min.*, vol. 16, no. 1, 2023, doi: 10.1186/s13040-023-00322-4.
 - [64] J. Li, “Area under the ROC Curve has the most consistent evaluation for binary classification,” *PLoS One*, vol. 19, no. 12 December, 2024, doi: 10.1371/journal.pone.0316019.
 - [65] G. M. Foody, “Challenges in the real world use of classification accuracy metrics: From recall and precision to the Matthews correlation coefficient,” *PLoS One*, vol. 18, no. 10 October, 2023, doi: 10.1371/journal.pone.0291908.
 - [66] A. Abdullah, W. S. Wong, and D. Albashish, “EB-CNN: Ensemble of branch convolutional neural network for image classification,” *Pattern Recognit. Lett.*, vol. 189, pp. 1–7, 2025, doi: 10.1016/j.patrec.2024.12.017.
 - [67] I. D. Mienye and Y. Sun, “A Survey of Ensemble Learning: Concepts, Algorithms, Applications, and Prospects,” *IEEE Access*, vol. 10, pp. 99129–99149, 2022, doi: 10.1109/ACCESS.2022.3207287.
 - [68] B. Amirshahi and S. Lahmiri, “Bankruptcy prediction using optimal ensemble models under balanced and imbalanced data,” *Expert Syst.*, vol. 41, no. 8, 2024, doi: 10.1111/exsy.13599.
 - [69] C. Li, “A gentle intro to gradient boosting,” vol. 23, no. 5, pp. 566–569, 2015.
 - [70] W. Chang, X. Wang, J. Yang, and T. Qin, “An Improved CatBoost-Based Classification Model for Ecological Suitability of Blueberries,” *Sensors*, vol. 23, no. 4, 2023, doi: 10.3390/s23041811.
 - [71] M. P. A. Ariawan, I. B. A. Peling, and G. B. Subiksa, “Prediksi Nilai Akhir Matakuliah Mahasiswa Menggunakan Metode K-Means Clustering (Studi Kasus : Matakuliah Pemrograman Dasar),” *J. Nas. Teknol. dan Sist. Inf.*, vol. 9, no. 2, pp. 122–131, 2023, doi: 10.25077/teknosi.v9i2.2023.122-131.
 - [72] J. T. Hancock and T. M. Khoshgoftaar, “CatBoost for big data: an

- interdisciplinary review,” *J. Big Data*, vol. 7, no. 1, 2020, doi: 10.1186/s40537-020-00369-8.
- [73] Y. Cai, Y. Yuan, and A. Zhou, “Predictive slope stability early warning model based on CatBoost,” *Sci. Rep.*, vol. 14, no. 1, 2024, doi: 10.1038/s41598-024-77058-6.
 - [74] M. Luo *et al.*, “Combination of feature selection and catboost for prediction: The first application to the estimation of aboveground biomass,” *Forests*, vol. 12, no. 2, pp. 1–22, 2021, doi: 10.3390/f12020216.
 - [75] M. Saber *et al.*, “Examining LightGBM and CatBoost models for wadi flash flood susceptibility prediction,” *Geocarto Int.*, vol. 37, no. 25, pp. 7462–7487, 2022, doi: 10.1080/10106049.2021.1974959.
 - [76] B. So and E. A. Valdez, “Zero-inflated Tweedie boosted trees with CatBoost for Insurance loss analytics,” *Appl. Soft Comput.*, vol. 177, p. 113226, 2025, doi: 10.1016/j.asoc.2025.113226.
 - [77] C. Li, N. Guan, and H. Zhang, “Anticancer drug synergy prediction based on CatBoost,” *PeerJ Comput. Sci.*, vol. 11, 2025, doi: 10.7717/peerj-cs.2829.
 - [78] T. M. Nguyen, H. H. P. Vo, and M. Yoo, “Enhancing Intrusion Detection in Wireless Sensor Networks Using a GSWO-CatBoost Approach,” *Sensors*, vol. 24, no. 11, 2024, doi: 10.3390/s24113339.
 - [79] G. Douzas, F. Bacao, and F. Last, “Improving imbalanced learning through a heuristic oversampling method based on k-means and SMOTE,” *Inf. Sci. (Ny)*, vol. 465, pp. 1–20, 2018, doi: 10.1016/j.ins.2018.06.056.
 - [80] J. H. Joloudari, A. Marefat, M. A. Nematollahi, S. S. Oyelere, and S. Hussain, “Effective Class-Imbalance Learning Based on SMOTE and Convolutional Neural Networks,” *Appl. Sci.*, vol. 13, no. 6, 2023, doi: 10.3390/app13064006.
 - [81] G. Husain *et al.*, “SMOTE vs. SMOTEENN: A Study on the Performance of Resampling Algorithms for Addressing Class Imbalance in Regression Models,” *Algorithms*, vol. 18, no. 1, 2025, doi: 10.3390/a18010037.
 - [82] M. Ilham, A. Winarno, M. Lutfi, and A. Indrasetyaningsih, “Handling Imbalanced Fraudulent Transaction Data Using SMOTE-Tomek and Random Forest: A Classification Approach,” *BEST J. Appl. Electr. Sci. Technol.*, vol. 7, no. 1, pp. 35–38, 2025, doi: 10.36456/best.vol7.no1.10335.
 - [83] J. Hemmatian, R. Hajizadeh, and F. Nazari, “Addressing imbalanced data classification with Cluster-Based Reduced Noise SMOTE,” *PLoS One*, vol. 20, no. 2 February, 2025, doi: 10.1371/journal.pone.0317396.
 - [84] R. Setiawan, E. Nursasongko, A. Syukur, F. Budiman, and D. Kurniadi, “Handling Class Imbalance in Student Success Prediction Using Machine

- Learning: A Comparison of SMOTE and SMOTETomek,” *2025 Int. Conf. Smart Comput. IoT Mach. Learn. SIML 2025*, pp. 1–6, 2025, doi: 10.1109/SIML65326.2025.11081128.
- [85] S. F. Taskiran, B. Turkoglu, E. Kaya, and T. Asuroglu, “A comprehensive evaluation of oversampling techniques for enhancing text classification performance,” *Sci. Rep.*, vol. 15, no. 1, 2025, doi: 10.1038/s41598-025-05791-7.
 - [86] D. Elreedy, A. F. Atiya, and F. Kamalov, “A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning,” *Mach. Learn.*, vol. 113, no. 7, pp. 4903–4923, 2024, doi: 10.1007/s10994-022-06296-4.
 - [87] J. Fonseca and F. Bacao, “Geometric SMOTE for imbalanced datasets with nominal and continuous features,” *Expert Syst. Appl.*, vol. 234, p. 121053, 2023, doi: 10.1016/j.eswa.2023.121053.
 - [88] G. Almeida and F. Bacao, “UMAP-SMOTENC: A simple, efficient, and consistent alternative for privacy-aware synthetic data generation,” *Knowledge-Based Syst.*, vol. 300, p. 112174, 2024, doi: 10.1016/j.knosys.2024.112174.
 - [89] S. D. A. Bujang *et al.*, “Multiclass Prediction Model for Student Grade Prediction Using Machine Learning,” *IEEE Access*, vol. 9, pp. 95608–95621, 2021, doi: 10.1109/ACCESS.2021.3093563.
 - [90] H. Sahlaoui, E. A. A. Alaoui, S. Agoujl, and A. Nayyar, “An empirical assessment of smote variants techniques and interpretation methods in improving the accuracy and the interpretability of student performance models,” *Educ. Inf. Technol.*, vol. 29, no. 5, pp. 5447–5483, 2024, doi: 10.1007/s10639-023-12007-w.
 - [91] A. Stuke, P. Rinke, and M. Todorovic, “Efficient hyperparameter tuning for kernel ridge regression with Bayesian optimization,” *Mach. Learn. Sci. Technol.*, vol. 2, no. 3, 2021, doi: 10.1088/2632-2153/abee59.
 - [92] L. I. Souadda, A. R. Halitim, B. Benilles, J. M. Oliveira, and P. Ramos, “Optimizing Credit Risk Prediction for Peer-to-Peer Lending Using Machine Learning,” *Forecasting*, vol. 7, no. 3, 2025, doi: 10.3390/forecast7030035.
 - [93] E. Elgeldawi, A. Sayed, A. R. Galal, and A. M. Zaki, “Hyperparameter tuning for machine learning algorithms used for arabic sentiment analysis,” *Informatics*, vol. 8, no. 4, p. 79, 2021, doi: 10.3390/informatics8040079.
 - [94] M. Wojciuk, Z. Swiderska-Chadaj, K. Siwek, and A. Gertych, “Improving classification accuracy of fine-tuned CNN models: Impact of hyperparameter optimization,” *Heliyon*, vol. 10, no. 5, 2024, doi: 10.1016/j.heliyon.2024.e26586.
 - [95] N. H. Tiep *et al.*, “A New Hyperparameter Tuning Framework for

- Regression Tasks in Deep Neural Network: Combined-Sampling Algorithm to Search the Optimized Hyperparameters,” *Mathematics*, vol. 12, no. 24, 2024, doi: 10.3390/math12243892.
- [96] M. Liang *et al.*, “Improving Genomic Prediction with Machine Learning Incorporating TPE for Hyperparameters Optimization,” *Biology (Basel)*, vol. 11, no. 11, 2022, doi: 10.3390/biology11111647.
- [97] L. Wang, C. Zhang, W. Wang, T. Deng, T. Ma, and P. Shuai, “Slope Stability Assessment Using an Optuna-TPE-Optimized CatBoost Model,” *Eng*, vol. 6, no. 8, 2025, doi: 10.3390/eng6080185.
- [98] S. Hanifi, A. Cammarono, and H. Zare-Behtash, “Advanced hyperparameter optimization of deep learning models for wind power prediction,” *Renew. Energy*, vol. 221, 2024, doi: 10.1016/j.renene.2023.119700.
- [99] A. C. R. Klaar, S. F. Stefenon, L. O. Seman, V. C. Mariani, and L. dos S. Coelho, “Optimized EWT-Seq2Seq-LSTM with Attention Mechanism to Insulators Fault Prediction,” *Sensors*, vol. 23, no. 6, 2023, doi: 10.3390/s23063202.
- [100] Y. Zhang and Z. Wang, “Feature Engineering and Model Optimization Based Classification Method for Network Intrusion Detection,” *Appl. Sci.*, vol. 13, no. 16, 2023, doi: 10.3390/app13169363.
- [101] M. Imani and H. R. Arabnia, “Hyperparameter Optimization and Combined Data Sampling Techniques in Machine Learning for Customer Churn Prediction: A Comparative Analysis,” 2023. doi: 10.3390/technologies11060167.
- [102] S. Ali, F. Akhlaq, A. S. Imran, Z. Kastrati, S. M. Daudpota, and M. Moosa, “The enlightening role of explainable artificial intelligence in medical & healthcare domains: A systematic literature review,” *Comput. Biol. Med.*, vol. 166, p. 107555, 2023, doi: 10.1016/j.compbio.2023.107555.
- [103] A. Nambiar, S. Harikrishnaa, and S. Sharanprasath, “Model-agnostic explainable artificial intelligence tools for severity prediction and symptom analysis on Indian COVID-19 data,” *Front. Artif. Intell.*, vol. 6, 2023, doi: 10.3389/frai.2023.1272506.
- [104] V. Vimbi, N. Shaffi, and M. Mahmud, “Interpreting artificial intelligence models: a systematic review on the application of LIME and SHAP in Alzheimer’s disease detection,” *Brain Informatics*, vol. 11, no. 1, 2024, doi: 10.1186/s40708-024-00222-1.
- [105] Z. Li, “Extracting spatial effects from machine learning model using local interpretation method: An example of SHAP and XGBoost,” *Comput. Environ. Urban Syst.*, vol. 96, p. 101845, 2022, doi: 10.1016/j.compenvurbsys.2022.101845.
- [106] R. Alkhanbouli, H. Matar Abdulla Almadhaani, F. Alhosani, and M. C. E.

- Simsekler, "The role of explainable artificial intelligence in disease prediction: a systematic literature review and future research directions," *BMC Med. Inform. Decis. Mak.*, vol. 25, no. 1, 2025, doi: 10.1186/s12911-025-02944-6.
- [107] A. V. Ponce-Bobadilla, V. Schmitt, C. S. Maier, S. Mensing, and S. Stodtmann, "Practical guide to SHAP analysis: Explaining supervised machine learning model predictions in drug development," *Clin. Transl. Sci.*, vol. 17, no. 11, 2024, doi: 10.1111/cts.70056.
- [108] P. Hermosilla, S. Berrios, and H. Allende-Cid, "Explainable AI for Forensic Analysis: A Comparative Study of SHAP and LIME in Intrusion Detection Models," *Appl. Sci.*, vol. 15, no. 13, 2025, doi: 10.3390/app15137329.
- [109] O. O. Bifarin, "Interpretable machine learning with treebased shapley additive explanations: Application to metabolomics datasets for binary classification," *PLoS One*, vol. 18, no. 5 May, 2023, doi: 10.1371/journal.pone.0284315.
- [110] Y. Nohara, K. Matsumoto, H. Soejima, and N. Nakashima, "Explanation of machine learning models using shapley additive explanation and application for real data in hospital," *Comput. Methods Programs Biomed.*, vol. 214, p. 106584, 2022, doi: 10.1016/j.cmpb.2021.106584.
- [111] R. Rodríguez-Pérez and J. Bajorath, "Interpretation of machine learning models using shapley values: application to compound potency and multi-target activity predictions," *J. Comput. Aided. Mol. Des.*, vol. 34, no. 10, pp. 1013–1026, 2020, doi: 10.1007/s10822-020-00314-0.
- [112] D. Wang, S. Thunell, U. Lindberg, L. Jiang, J. Trygg, and M. Tysklind, "Towards better process management in wastewater treatment plants: Process analytics based on SHAP values for tree-based machine learning methods," *J. Environ. Manage.*, vol. 301, p. 113941, 2022, doi: 10.1016/j.jenvman.2021.113941.
- [113] R. Mitchell, E. Frank, and G. Holmes, "GPUShap: Massively parallel exact calculation of SHAP scores for tree ensembles," *PeerJ Comput. Sci.*, vol. 8, 2022, doi: 10.7717/PEERJ-CS.880.
- [114] M. Heydarian, T. E. Doyle, and R. Samavi, "MLCM: Multi-label confusion matrix," *IEEE Access*, 2022.
- [115] Y. Wang, Y. Jia, Y. Tian, and J. Xiao, "Deep reinforcement learning with the confusion-matrix-based dynamic reward function for customer credit scoring," *Expert Syst. Appl.*, vol. 200, 2022, doi: 10.1016/j.eswa.2022.117013.
- [116] W. Lee, *Python® Machine Learning*. books.google.com, 2019. doi: 10.1002/9781119557500.
- [117] N. Ketkar and J. Moolayil, *Deep Learning with Python*. books.google.com,

2021. doi: 10.1007/978-1-4842-5364-9.

- [118] S. V. Halyal, “Running Google Colaboratory as a server – transferring dynamic data in and out of colabs,” 2019, *mecs-press.org*. doi: 10.5815/ijeme.2019.06.04.
- [119] G. Vos, L. van Eijk, Z. Sarnyai, and M. Rahimi Azghadi, “Stabilizing machine learning for reproducible and explainable results: A novel validation approach to subject-specific insights,” *Comput. Methods Programs Biomed.*, vol. 269, p. 108899, 2025, doi: 10.1016/j.cmpb.2025.108899.
- [120] T. El Moudden and M. Amnai, “Building an efficient convolution neural network from scratch: A case study on detecting and localizing slums,” *Sci. African*, vol. 20, 2023, doi: 10.1016/j.sciaf.2023.e01612.
- [121] Y. Li, Y. Yang, P. Song, L. Duan, and R. Ren, “An improved SMOTE algorithm for enhanced imbalanced data classification by expanding sample generation space,” *Sci. Rep.*, vol. 15, no. 1, pp. 1–21, 2025, doi: 10.1038/s41598-025-09506-w.
- [122] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, “Catboost: Unbiased boosting with categorical features,” *Adv. Neural Inf. Process. Syst.*, vol. 2018-Decem, no. Section 4, pp. 6638–6648, 2018.
- [123] P. Virtanen *et al.*, “SciPy 1.0: fundamental algorithms for scientific computing in Python,” *Nat. Methods*, vol. 17, no. 3, pp. 261–272, 2020, doi: 10.1038/s41592-019-0686-2.
- [124] K. Gao, G. Mei, F. Piccialli, S. Cuomo, J. Tu, and Z. Huo, “Julia language in machine learning: Algorithms, applications, and open issues,” 2020. doi: 10.1016/j.cosrev.2020.100254.
- [125] F. M. Giorgi, C. Ceraolo, and D. Mercatelli, “The R Language: An Engine for Bioinformatics and Data Science,” 2022. doi: 10.3390/life12050648.

UNIVERSITAS
MULTIMEDIA
NUSANTARA