DAFTAR PUSTAKA

Alamgir, Z., Akram, H., Karim, S., & Wali, A. (2023). Enhancing Student Performance Prediction via Educational Data Mining on Academic data. *Informatics in Education*. https://doi.org/10.15388/infedu.2024.04

Aldania, A. N. A., Soleh, A. M., & Notodiputro, K. A. (2023). A Comparative Study of CatBoost and Double Random Forest for Multi-class Classification. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, *7*(1), 129–137. https://doi.org/10.29207/resti.v7i1.4766

Ayulani, I. D., Yunawan, A. M., Prihutaminingsih, T., Sarwinda, D., Ardaneswari, G., & Handari, B. D. (2023a). Tree-Based Ensemble Methods and Their Applications for Predicting Studentsâ€™ Academic Performance. *International Journal on Advanced Science, Engineering and Information Technology*, *13*(3), 919–927. https://doi.org/10.18517/ijaseit.13.3.16880

Ayulani, I. D., Yunawan, A. M., Prihutaminingsih, T., Sarwinda, D., Ardaneswari, G., & Handari, B. D. (2023b). Tree-Based Ensemble Methods and Their Applications for Predicting Studentsâ€™ Academic Performance. *International Journal on Advanced Science, Engineering and Information Technology*, *13*(3), 919–927. https://doi.org/10.18517/ijaseit.13.3.16880

Banerjee, A. K. (2010). A multidimensional Gini index. *Mathematical Social Sciences*, *60*(2), 87–93. https://doi.org/10.1016/j.mathsocsci.2010.06.001

Beckham, N. R., Akeh, L. J., Mitaart, G. N. P., & Moniaga, J. V. (2023). Determining factors that affect student performance using various machine learning methods. *Procedia Computer Science*, *216*, 597–603.

https://doi.org/10.1016/j.procs.2022.12.174

Bentéjac, C., Csörgő, A., & Martínez-Muñoz, G. (2021a). A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review*, *54*(3), 1937–1967. https://doi.org/10.1007/s10462-020-09896-5

Bentéjac, C., Csörgő, A., & Martínez-Muñoz, G. (2021b). A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review*, *54*(3), 1937–1967. https://doi.org/10.1007/s10462-020-09896-5

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, *16*, 321–357. https://doi.org/10.1613/jair.953

Dietterich, T. G. (2000). Ensemble Methods in Machine Learning. Dalam G. Goos, J. Hartmanis, & J. Van Leeuwen (Eds.), *Multiple Classifier Systems* (Vol. 1857, hlm. 1–15). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-45014-9_1

Dorogush, A. V., Ershov, V., & Gulin, A. (2018). *CatBoost: Gradient boosting with categorical features support* (Versi 1). arXiv. https://doi.org/10.48550/ARXIV.1810.11363

E. Ben George. (2025). Explainable AI Methods for Predicting Student Grades and Improving Academic Success. *Journal of Information Systems Engineering and Management*, *10*(23s), 117–126. https://doi.org/10.52783/jisem.v10i23s.3680

García-Vicente, C., Chushig-Muzo, D., Mora-Jiménez, I., Fabelo, H., Gram, I. T.,

Løchen, M.-L., Granja, C., & Soguero-Ruiz, C. (2023). Evaluation of Synthetic Categorical Data Generation Techniques for Predicting Cardiovascular Diseases and Post-Hoc Interpretability of the Risk Factors. *Applied Sciences*, *13*(7), 4119. https://doi.org/10.3390/app13074119

Genuer, R., Poggi, J.-M., & Tuleau-Malot, C. (2010). Variable selection using random forests. *Pattern Recognition Letters*, *31*(14), 2225–2236. https://doi.org/10.1016/j.patrec.2010.03.014

Guan, X., Xue, R., He, Z., Chen, S., & Chen, X. (2025). CatBoost-Optimized Hyperspectral Modeling for Accurate Prediction of Wood Dyeing Formulations. *Forests*, *16*(8), 1279. https://doi.org/10.3390/f16081279

Hastie, T., Tibshirani, R., & Friedman, J. H. (2017). *The elements of statistical learning: Data mining, inference, and prediction* (Second edition). Springer.

Haymet, A. D. J. (1994). Calculation of entropy. *Computational Materials Science*, *2*(1), 85–92. https://doi.org/10.1016/0927-0256(94)90050-7

Johora, F. T., Hasan, M. N., Rajbongshi, A., Ashrafuzzaman, M., & Akter, F. (2025). An explainable AI-based approach for predicting undergraduate students academic performance. *Array*, *26*, 100384. https://doi.org/10.1016/j.array.2025.100384

Kesgin, K., Kiraz, S., Kosunalp, S., & Stoycheva, B. (2025). Beyond Performance: Explaining and Ensuring Fairness in Student Academic Performance Prediction with Machine Learning. *Applied Sciences*, *15*(15), 8409. https://doi.org/10.3390/app15158409

Kharis, S. A. A., & Zili, A. H. A. (2022). Learning Analytics dan Educational Data Mining pada Data Pendidikan. *JURNAL RISET PEMBELAJARAN MATEMATIKA SEKOLAH*, *6*(1), 12–20. https://doi.org/10.21009/jrpms.061.02

Kumar, M., Singh, N., Wadhwa, J., Singh, P., Kumar, G., & Qtaishat, A. (2024). Utilizing Random Forest and XGBoost DataMining Algorithms for Anticipating Students'Academic Performance. *International Journal of Modern Education and Computer Science*, *16*(2), 29–44. https://doi.org/10.5815/ijmecs.2024.02.03

Lundberg, S., & Lee, S.-I. (2017). *A Unified Approach to Interpreting Model Predictions* (Versi 2). arXiv. https://doi.org/10.48550/ARXIV.1705.07874

Matthay, E. C., & Glymour, M. M. (2020). A Graphical Catalog of Threats to Validity: Linking Social Science with Epidemiology. *Epidemiology*, *31*(3), 376–384. https://doi.org/10.1097/EDE.0000000000001161

Mingyu, Z., Sutong, W., Yanzhang, W., & Dujuan, W. (2022). An interpretable prediction method for university student academic crisis warning. *Complex & Intelligent Systems*, *8*(1), 323–336. https://doi.org/10.1007/s40747-021-00383-0

Najwa Laila Anggraini, Basuki Rahmat, & Achmad Junaidi. (2025). Performance Analysis of SMOTE and SMOTEN Techniques for Daily Rainfall Classification using XGBoost. *Journal of Artificial Intelligence and Engineering Applications (JAIEA)*, *4*(3), 1987–1993. https://doi.org/10.59934/jaiea.v4i3.1066

Ninov Royana, A., Vita Via, Y., & Aji Putra, C. (2025). EVALUASI KINERJA LIGHTGBM DAN CATBOOST UNTUK PREDIKSI CHURN BERDASARKAN DATASET PELANGGAN LAYANAN STREAMING MUSIK. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *9*(4), 6977–6982. https://doi.org/10.36040/jati.v9i4.14358

Nugraha, W., & Syarif, M. (2023). Teknik Weighting untuk Mengatasi Ketidakseimbangan Kelas Pada Prediksi Churn Menggunakan XGBoost, LightGBM, dan CatBoost. *Techno.Com*, *22*(1), 97–108. https://doi.org/10.33633/tc.v22i1.7191

Ovirianti, N. H., Zarlis, M., & Mawengkang, H. (2022). Support Vector Machine Using A Classification Algorithm. *SinkrOn*, *7*(3), 2103–2107. https://doi.org/10.33395/sinkron.v7i3.11597

Patle, A., & Chouhan, D. S. (2013). SVM kernel functions for classification. *2013 International Conference on Advances in Technology and Engineering (ICATE)*, 1–9. https://doi.org/10.1109/ICAdTE.2013.6524743

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2019). *CatBoost: Unbiased boosting with categorical features* (Versi 5). arXiv. https://doi.org/10.48550/ARXIV.1706.09516

Ptr, A. F. L., Siregar, M. M., & Daniel, I. (2024). Analysis of Gradient Boosting, XGBoost, and CatBoost on Mobile Phone Classification. *Journal of Computer Networks, Architecture and High Performance Computing*, *6*(2), 661–670. https://doi.org/10.47709/cnahpc.v6i2.3790

Ramaswami, G., Susnjak, T., & Mathrani, A. (2022). On Developing Generic

Models for Predicting Student Outcomes in Educational Data Mining. *Big Data and Cognitive Computing*, *6*(1), 6. https://doi.org/10.3390/bdcc6010006

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why Should I Trust You?": Explaining the Predictions of Any Classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144. https://doi.org/10.1145/2939672.2939778

Sahlaoui, H., Alaoui, E. A. A., Nayyar, A., Agoujil, S., & Jaber, M. M. (2021). Predicting and Interpreting Student Performance Using Ensemble Models and Shapley Additive Explanations. *IEEE Access*, *9*, 152688–152703. https://doi.org/10.1109/ACCESS.2021.3124270

Salih, A., Raisi-Estabragh, Z., Galazzo, I. B., Radeva, P., Petersen, S. E., Menegaz, G., & Lekadir, K. (2023). *A Perspective on Explainable Artificial Intelligence Methods: SHAP and LIME*. https://doi.org/10.48550/ARXIV.2305.02012

Samosir P, P., & Salamah, U. (2025). Perbandingan Performa Algoritma XGBoost, CatBoost Dan GBM Dalam Prediksi Penyakit Kardiovaskular. *JSAI (Journal Scientific and Applied Informatics)*, *8*(1), 268–273. https://doi.org/10.36085/jsai.v8i1.7552

Shoukath, T. K., & Chakkaravarthy, M. (2025). Predictive analytics in education: Machine learning approaches and performance metrics for student success – a systematic literature review. *Data and Metadata*, *4*, 730. https://doi.org/10.56294/dm2025730

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, *45*(4), 427–437. https://doi.org/10.1016/j.ipm.2009.03.002

Szeghalmy, S., & Fazekas, A. (2023). A Comparative Study of the Use of Stratified Cross-Validation and Distribution-Balanced Stratified Cross-Validation in Imbalanced Learning. *Sensors*, *23*(4), 2333. https://doi.org/10.3390/s23042333

Tirumanadham, N. S. K. M. K., S, T., & M, S. (2024). Evaluating Boosting Algorithms for Academic Performance Prediction in E-Learning Environments. *2024 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE)*, 1–8. https://doi.org/10.1109/IITCEE59897.2024.10467968

Verma, S., Yadav, R. K., & Kholiya, K. (2022). Prediction of Academic Performance of Engineering Students by Using Data Mining Techniques. *International Journal of Information and Education Technology*, *12*(11), 1164–1171. https://doi.org/10.18178/ijiet.2022.12.11.1734

Wang, S., & Luo, B. (2024). Academic achievement prediction in higher education through interpretable modeling. *PLOS ONE*, *19*(9), e0309838. https://doi.org/10.1371/journal.pone.0309838

Widiyatmoko, C. B., Gernowo, R., & Warsito, B. (2025). Unstacking the Stack: Synthesis of Optimization Strategies for Stacked Ensemble Models in Multi-Domain Contexts. *Jurnal Sisfokom (Sistem Informasi dan Komputer)*, *15*(01), 88–94. https://doi.org/10.32736/sisfokom.v15i01.2545

Wolpert, D. H. (1992). Stacked generalization. *Neural Networks*, *5*(2), 241–259. https://doi.org/10.1016/S0893-6080(05)80023-1

Yağcı, M. (2022). Educational data mining: Prediction of students' academic performance using machine learning algorithms. *Smart Learning Environments*, *9*(1), 11. https://doi.org/10.1186/s40561-022-00192-z

Zang, F., & Zhang, J. (2011). Softmax Discriminant Classifier. *2011 Third International Conference on Multimedia Information Networking and Security*, 16–19. https://doi.org/10.1109/MINES.2011.123