

DAFTAR PUSTAKA

- Adawiyah, R., Badriyah, T., Syarif, I., Elektronika Negeri Surabaya, P., & Kampus PENS Jl Raya, I. (2021). *Hospital Length of Stay Prediction Based on Patient Examination Using Neural Network*. 08(1).
- Adawiyah, R., Badriyah, T., & Syarif, I. (2021). *Hospital Length of Stay Prediction based on Patient Examination Using General Features*. EMITTER International Journal of Engineering Technology, 9(1), 169–181. <https://doi.org/10.24003/emitter.v9i1.609>.
- AIS Universitas Siliwangi. (2020). Roadmap Artificial Intelligence Research Group Universitas Siliwangi tahun 2020-2030.
- Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). *Optuna: A Next-generation Hyperparameter Optimization Framework*. <http://arxiv.org/abs/1907.10902>
- Alsinglawi, B. S., Alnajjar, F., Alorjani, M. S., Al-Shari, O. M., Munoz, M. N., & Mubin, O. (2024). Predicting Hospital Stay Length Using Explainable Machine Learning. *IEEE Access*, 12, 90571–90585. <https://doi.org/10.1109/ACCESS.2024.3421295>
- Anglin, K., Liu, Q., & Wong, V. C. (2024). A primer on the validity typology and threats to validity in education research. *Asia Pacific Education Review*, 25(3), 557–574. <https://doi.org/10.1007/s12564-024-09955-4>
- Arat, M. M. (2022). Learning From High-Cardinality Categorical Features in Deep Neural Networks. *Journal of Advanced Research in Natural and Applied Sciences*, 8(2), 222–236. <https://doi.org/10.28979/jarnas.1014469>
- Asal, B., & Yalciner, B. (2026). Benchmarking TabNet, NODE, and FT-Transformer for Software Defect Prediction: An Empirical Comparison and Explainability Analysis. *IEEE Access*, 14, 11660–11681. <https://doi.org/10.1109/ACCESS.2026.3656247>
- Balik, M. Y. (2024). *Comparing Federated Stochastic Gradient Descent and Federated Averaging for Predicting Hospital Length of Stay*. <http://arxiv.org/abs/2407.12741>
- Brenndoerfer, M. (2025). *Gradient Clipping: Preventing Exploding Gradients in Deep Learning*. <https://mbrenndoerfer.com/writing/gradient-clipping-deep-learning>
- Butt, M. U., Hotho, A., & Schlör, D. (2026). *Evaluating Tabular Representation Learning for Network Intrusion Detection*. <http://arxiv.org/abs/2605.02519>

- Chen, K.-Y., Shyu, C.-R., Tsai, Y.-Y., Baskett, W. I., Chang, C.-Y., Chou, C.-Y., Tsai, J. J. P., & Shae, Z.-Y. (2025). Effective Non-IID Degree Estimation for Robust Federated Learning in Healthcare Datasets. *Journal of Healthcare Informatics Research*, 9(3), 437–464. <https://doi.org/10.1007/s41666-025-00195-8>
- Chen, Y., Ding, W., Huang, J., Zhang, W., & Zhou, T. (2025). Multigranularity Fuzzy Autoencoder for Discriminative Feature Selection in High-Dimensional Data. *IEEE Transactions on Neural Networks and Learning Systems*, 36(9), 17433–17447. <https://doi.org/10.1109/TNNLS.2025.3569893>
- Choi, J.-W., Yang, M., Kim, J.-W., Shin, Y. M., Shin, Y.-G., & Park, S. (2024). Prognostic prediction of sepsis patient using transformer with skip connected token for tabular data. *Artificial Intelligence in Medicine*, 149, 102804. <https://doi.org/10.1016/j.artmed.2024.102804>
- Ciecińska, B. (2023). *Identification of Defects Causes: Ishikawa Diagram and 5 Whys in Theoretical and Practical Terms*. DOI: 10.5772/intechopen.113990. Diambil www.intechopen.com
- Digitale, J., Franzon, D., Pletcher, M. J., McCulloch, C. E., & Gennatas, E. D. (2025). Methods for Addressing Missingness in Electronic Health Record Data for Clinical Prediction Models: Comparative Evaluation. *JMIR Medical Informatics*, 13, e79307–e79307. <https://doi.org/10.2196/79307>
- Fa, R., Stoessel, D., Artemova, S., von Schenck, U., Rostami, H. N., Madiot, P. E., Landelle, C., & Bosson, J. L. (2025). Prediction of prolonged length of stay from first 2 days of hospitalization data: a SHAP value-based variable selection method for a simplified model. *International Journal for Quality in Health Care*, 37(4). <https://doi.org/10.1093/intqhc/mzaf116>
- Fani, T., Kun Saptorini, K., & Perguruan Tinggi, A. (2024). Factors Associated with Length of Stay of Inpatients: A Cross-sectional Study at RSD KRMT Wongsonegoro, Semarang, 2024. Dalam *Procedia of Engineering and Life Science* (Vol. 6).
- Ferrão, J. C., Oliveira, M. D., Gartner, D., Janela, F., & Martins, H. M. G. (2021). Leveraging electronic health record data to inform hospital resource management. *Health Care Management Science*, 24(4), 716–741. <https://doi.org/10.1007/s10729-021-09554-4>
- Gaib, F. A., Gusti Ngurah Bagus Ferry Mahayudha, I., Karta Wijaya, A., Andini, N., & Zamhariroyan, A. (2025). *SNESTIK Seminar Nasional Teknik Elektro, Sistem Informasi, dan Teknik Informatika Advanced Sleep Disorder Classification: An ML-Based Study with Optuna for Model Optimization*. <https://doi.org/10.31284/p.snestik.2025.7380>

- Gebreyesus, Y., Dalton, D., De Chiara, D., Chinnici, M., & Chinnici, A. (2024). AI for Automating Data Center Operations: Model Explainability in the Data Centre Context Using Shapley Additive Explanations (SHAP). *Electronics (Switzerland)*, 13(9). <https://doi.org/10.3390/electronics13091628>
- Gorishniy, Y., Rubachev, I., Khrulkov, V., & Babenko, A. (2021). *Revisiting Deep Learning Models for Tabular Data*. Diambil <https://github.com/yandex-research/rtdl>.
- Grinsztajn, L., Oyallon, E., & Varoquaux, G. (2022). *Why do tree-based models still outperform deep learning on typical tabular data?*, 36th Conference on Neural Information Processing Systems (NeurIPS 2022) Track on Datasets and Benchmarks.
- Gurupur, V., Hooshmand, S., Prabhu, D. F., Trader, E., & Salvi, S. (2025). Incompleteness of Electronic Health Records: An Impending Process Problem Within Healthcare. *Healthcare*, 13(22), 2900. <https://doi.org/10.3390/healthcare13222900>
- Health at a Glance 2025*. (2025). OECD Publishing. <https://doi.org/10.1787/8f9e3f98-en>
- Hidayaturrohman, Q. A., & Hanada, E. (2024). Impact of Data Pre-Processing Techniques on XGBoost Model Performance for Predicting All-Cause Readmission and Mortality Among Patients with Heart Failure. *BioMedInformatics*, 4(4), 2201–2212. <https://doi.org/10.3390/biomedinformatics4040118>
- Hirani, R., Podder, D., Stala, O., Mohebpour, R., Tiwari, R. K., & Etienne, M. (2025). Strategies to Reduce Hospital Length of Stay: Evidence and Challenges. Dalam *Medicina (Lithuania)* (Vol. 61, Nomor 5). Multidisciplinary Digital Publishing Institute (MDPI). <https://doi.org/10.3390/medicina61050922>
- Hu, J., & Szymczak, S. (2023). A review on longitudinal data analysis with random forest. Dalam *Briefings in Bioinformatics* (Vol. 24, Nomor 2). Oxford University Press. <https://doi.org/10.1093/bib/bbad002>
- Huber, P. J. (1964). Robust Estimation of a Location Parameter. *The Annals of Mathematical Statistics*, 35(1), 73–101. <https://doi.org/10.1214/aoms/1177703732>
- Jadon, A., Patil, A., & Jadon, S. (2022). *A Comprehensive Survey of Regression Based Loss Functions for Time Series Forecasting*. <http://arxiv.org/abs/2211.02989>

- Jain, R., Singh, M., Rao, A. R., & Garg, R. (2024). Predicting hospital length of stay using machine learning on a large open health dataset. *BMC Health Services Research*, 24(1). <https://doi.org/10.1186/s12913-024-11238-y>
- Kee, T., & Ho, W. K. O. (2025). Optimizing Machine Learning Models for Urban Sciences: A Comparative Analysis of Hyperparameter Tuning Methods. *Urban Science*, 9(9). <https://doi.org/10.3390/urbansci9090348>
- Koukaras, P., & Tjortjis, C. (2025). Data Preprocessing and Feature Engineering for Data Mining: Techniques, Tools, and Best Practices. *AI*, 6(10), 257. <https://doi.org/10.3390/ai6100257>
- Kumah, A., Nwogu, C. N., Issah, A.-R., Obot, E., Kanamitie, D. T., Sifa, J. S., & Aidoo, L. A. (2024). Cause-and-Effect (Fishbone) Diagram: A Tool for Generating and Organizing Quality Improvement Ideas. *Global Journal on Quality and Safety in Healthcare*, 7(2), 85–87. <https://doi.org/10.36401/JQSH-23-42>
- Lago, P., Runeson, P., Song, Q., & Verdecchia, R. (2024). Threats to Validity in Software Engineering – hypocritical paper section or essential analysis? *Proceedings of the 18th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*, 314–324. <https://doi.org/10.1145/3674805.3686691>
- Lamane, H., Mouhir, L., Moussadek, R., Baghdad, B., Kisi, O., & El Bilali, A. (2025). Interpreting machine learning models based on SHAP values in predicting suspended sediment concentration. *International Journal of Sediment Research*, 40(1), 91–107. <https://doi.org/10.1016/j.ijsrc.2024.10.002>
- Lee, J., & Park, J. (2025). Relationships Among Comorbidities, Disease Severity, and Hospitalization Duration in the United States Using the Healthcare Cost and Utilization Project (HCUP) Database. *Journal of Clinical Medicine*, 14(3), 680. <https://doi.org/10.3390/jcm14030680>
- Lindauer, M., Eggensperger, K., Feurer, M., Biedenkapp, A., Marben, J., Müller, P., & Hutter, F. (2019). *BOAH: A Tool Suite for Multi-Fidelity Bayesian Optimization & Analysis of Hyperparameters*. <http://arxiv.org/abs/1908.06756>
- Loshchilov, I., & Hutter, F. (2017). *SGDR: Stochastic Gradient Descent With Warm Restarts*. Diambil <https://github.com/loshchil/SGDR>
- Lundberg, S., & Lee, S.-I. (2017). *A Unified Approach to Interpreting Model Predictions*. <http://arxiv.org/abs/1705.07874>
- Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S.-I. (2020). From local explanations

- to global understanding with explainable AI for trees. *Nature Machine Intelligence*, 2(1), 56–67. <https://doi.org/10.1038/s42256-019-0138-9>
- Max Kuhn, & Kjell Johnson. (2026). *Feature Engineering and Selection: A Practical Approach for Predictive Models*.
- Mekhaldi, R. N., Caulier, P., Chaabane, S., Chraibi, A., & Piechowiak, S. (2021). A comparative study of machine learning models for predicting length of stay in hospitals. *Journal of Information Science and Engineering*, 37(5), 1025–1038. [https://doi.org/10.6688/JISE.202109_37\(5\).0003](https://doi.org/10.6688/JISE.202109_37(5).0003)
- Mengistu, A. K., Getinet, K., Alemayehu, T., Yeneakal, K. A., Gedefaw, A. E., Teym, A., Endalew, B., & Assaye, B. T. (2026). Machine learning predicts prolonged patient length of stay in a resource constrained Ethiopian hospital. *Discover Artificial Intelligence*. <https://doi.org/10.1007/s44163-025-00794-9>
- Mermer, O., Liu, Y., Jennissen, C. A., Sonka, M., & Demir, I. (2026). Agricultural Injury Severity Prediction Using Integrated Data-Driven Analysis: Global Versus Local Explainability Using SHAP. *Safety*, 12(1), 6. <https://doi.org/10.3390/safety12010006>
- Miller, C., Portlock, T., Nyaga, D. M., & O’Sullivan, J. M. (2024). A review of model evaluation metrics for machine learning in genetics and genomics. *Frontiers in Bioinformatics*, 4. <https://doi.org/10.3389/fbinf.2024.1457619>
- Munshi, R. M., Munshi, L. R., Himdi, H., Qashlan, A., Munshi, R., Alyahyawy, O. Y., & Khayyat, M. M. (2025). Optimising hyperparameters with a tree structured Parzen estimator to improve diabetes prediction. *Scientific Reports*, 15(1). <https://doi.org/10.1038/s41598-025-19295-x>
- Mustafa, A., Naseem, U., & Rahimi Azghadi, M. (2025). Large language models vs human for classifying clinical documents. *International Journal of Medical Informatics*, 195, 105800. <https://doi.org/10.1016/j.ijmedinf.2025.105800>
- Mustafa, A., Wei, C., Khan, S., Rizvi, T., Grovu, R., Khan, D., Dilluvio, B., Bjorklund, J., El-Sayegh, S., & Weinberg, M. (2024). Predictors of complications and extended length of stay following percutaneous transluminal renal artery angioplasty. *Medicine*, 103(52), e41017. <https://doi.org/10.1097/MD.00000000000041017>
- Nasiatin, T. (2025.). The Relationship Between Length of Stay (LOS) and Patient Satisfaction With Hospitality Services at Haji General Hospital Medan, *Contagion: Scientific Periodical of Public Health and Coastal Health*, 7(1).
- Nijman, S. W. J., Leeuwenberg, A. M., Beekers, I., Verkouter, I., Jacobs, J. J. L., Bots, M. L., Asselbergs, F. W., Moons, K. G. M., & Debray, T. P. A. (2022). Missing data is poorly handled and reported in prediction model studies using

- machine learning: a literature review. Dalam *Journal of Clinical Epidemiology* (Vol. 142, hlm. 218–229). Elsevier Inc. <https://doi.org/10.1016/j.jclinepi.2021.11.023>
- Pargent, F., Pfisterer, F., Thomas, J., & Bischl, B. (2022). Regularized target encoding outperforms traditional methods in supervised machine learning with high cardinality features. *Computational Statistics*, 37(5), 2671–2692. <https://doi.org/10.1007/s00180-022-01207-6>
- Pascanu, R., Mikolov, T., & Bengio, Y. (2013). On the difficulty of training recurrent neural networks, *ICML'13: Proceedings of the 30th International Conference on International Conference on Machine Learning*.
- Peng, B., & Gao, S. (2025). Prediction of hospital length of stay: leveraging ensemble tree models and intelligent feature selection. *Journal of Hospital Management and Health Policy*, 9. <https://doi.org/10.21037/jhmhp-24-129>
- Pinheiro, J. M. H., Oliveira, S. V. B. de, Silva, T. H. S., Saraiva, P. A. R., Souza, E. F. de, Godoy, R. V., Ambrosio, L. A., & Becker, M. (2025). The Impact of Feature Scaling in Machine Learning: Effects on Regression and Classification Tasks. *IEEE Access*, 13, 199903–199931. <https://doi.org/10.1109/ACCESS.2025.3635541>
- Ponce-Bobadilla, A. V., Schmitt, V., Maier, C. S., Mensing, S., & Stodtmann, S. (2024). Practical guide to <scp>SHAP</scp> analysis: Explaining supervised machine learning model predictions in drug development. *Clinical and Translational Science*, 17(11). <https://doi.org/10.1111/cts.70056>
- Puspita, D., Lumandi, W., & Rachman, A. (2023). Hospital Length of Stay Prediction with Ensemble Learning Methode. Dalam *Hspital Length of Stay Prediction with Ensemble Learning Methode* (Vol. 29).
- Qin, L., Zhu, Y., Liu, S., Zhang, X., & Zhao, Y. (2025). The Shapley Value in Data Science: Advances in Computation, Extensions, and Applications. *Mathematics*, 13(10), 1581. <https://doi.org/10.3390/math13101581>
- Rahman, M. M., Kundu, D., Suha, S. A., Siddiqi, U. R., & Dey, S. K. (2022). Hospital patients' length of stay prediction: A federated learning approach. *Journal of King Saud University - Computer and Information Sciences*, 34(10), 7874–7884. <https://doi.org/10.1016/j.jksuci.2022.07.006>
- Rajeashwari, S., & Arunesh, K. (2025). *An Enhanced Learning Model Based on an Improved Random Forest Classifier and an Integrated Attribute Selector for Healthcare Datasets* (hlm. 234–249). https://doi.org/10.1007/978-3-031-73068-9_19

- Raschka, S. (2020). *Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning*. <http://arxiv.org/abs/1811.12808>
- Ren, Y., Zeng, L., Lou, J., Xiong, L., Ho, J. C., Jiang, X., & Bhavani, S. V. (2022). *Unraveling Complex Temporal Patterns in EHRs via Robust Irregular Tensor Factorization*.
- Roder, J., Maguire, L., Georgantas, R., & Roder, H. (2021). Explaining multivariate molecular diagnostic tests via Shapley values. *BMC Medical Informatics and Decision Making*, 21(1), 211. <https://doi.org/10.1186/s12911-021-01569-9>
- Salih, A. M., Raisi-Estabragh, Z., Galazzo, I. B., Radeva, P., Petersen, S. E., Lekadir, K., & Menegaz, G. (2025). A Perspective on Explainable Artificial Intelligence Methods: SHAP and LIME. *Advanced Intelligent Systems*, 7(1). <https://doi.org/10.1002/aisy.202400304>
- Sendak, M. P., Gao, M., Brajer, N., & Balu, S. (2020). Presenting machine learning model information to clinical end users with model facts labels. Dalam *npj Digital Medicine* (Vol. 3, Nomor 1). Nature Research. <https://doi.org/10.1038/s41746-020-0253-3>
- Shap.KernelExplainer — SHAP latest documentation*. (2018). Diambil 11 Juni 2026, dari <https://shap.readthedocs.io/en/latest/generated/shap.KernelExplainer.html>
- Siddique, S. M., Tipton, K., Leas, B., Greysen, S. R., Mull, N. K., Lane-Fall, M., McShea, K., & Tsou, A. Y. (2021). Interventions to Reduce Hospital Length of Stay in High-risk Populations. *JAMA Network Open*, 4(9), e2125846. <https://doi.org/10.1001/jamanetworkopen.2021.25846>
- Sivakumar, M., Parthasarathy, S., & Padmapriya, T. (2024). Trade-off between training and testing ratio in machine learning for medical image processing. *PeerJ Computer Science*, 10, e2245. <https://doi.org/10.7717/peerj-cs.2245>
- Stahl-Toyota, S., Dönnhoff, I., Nagy, E., Hochlehnert, A., Unger, I., Szendrödi, J., Frey, N., Michl, P., Müller-Tidow, C., Jäger, D., Friederich, H.-C., & Nikendei, C. (2025). Machine learning models incorporating somatic and mental comorbidities for prolonged length-of-stay prediction in a maximum care university hospital. *BMC Medical Informatics and Decision Making*, 25(1), 436. <https://doi.org/10.1186/s12911-025-03290-3>
- Sunardi, S., & Pradiarti, P. A. (2025). Relationship between quality of care and length of stay in the emergency department: A systematic literature review. *Malahayati International Journal of Nursing and Health Science*, 8(3), 333–344. <https://doi.org/10.33024/minh.v8i3.669>

- Tyralis, H., Papacharalampous, G., Dogulu, N., & Chun, K. P. (2025). Deep Huber quantile regression networks. *Neural Networks*, 187, 107364. <https://doi.org/10.1016/j.neunet.2025.107364>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2023). *Attention Is All You Need*. <http://arxiv.org/abs/1706.03762>
- Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., & Wesslén, A. (2012). *Experimentation in Software Engineering*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-29044-2>
- Wongoutong, C. (2024). The impact of neglecting feature scaling in k-means clustering. *PLoS ONE*, 19(12). <https://doi.org/10.1371/journal.pone.0310839>
- World Health Organization European Region. (2024). *Average length of stay, all hospitals*. https://gateway.euro.who.int/en/indicators/hfa_540-6100-average-length-of-stay-all-hospitals/#id=19635.
- Yang, A., Zhang, M., Chen, H., Li, T., Liu, S., & Xu, X. (2025). tanh As a robust feature scaling method in training deep learning models with imbalanced data. *Pattern Recognition*, 167, 111746. <https://doi.org/10.1016/j.patcog.2025.111746>
- Yuan, X., & Ren, S. (2024). *Transfer Learning for High Dimensional Robust Regression*. <http://arxiv.org/abs/2406.17567>
- Zhang, C., Shao, Y., Sun, H., Xing, L., Zhao, Q., & Zhang, L. (2023). The WuC-Adam algorithm based on joint improvement of Warmup and cosine annealing algorithms. *Mathematical Biosciences and Engineering*, 21(1), 1270–1285. <https://doi.org/10.3934/mbe.2024054>
- Zhang, J., He, T., Sra, S., & Jadbabaie, A. (2017). *Why Gradient Clipping Accelerates Training: A Theoretical Justification For Adaptivity*.
- Zhou, P., Xie, X., Lin, Z., & Yan, S. (2024). Towards Understanding Convergence and Generalization of AdamW. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(9), 6486–6493. <https://doi.org/10.1109/TPAMI.2024.3382294>
- Zhou, Y., Aryal, S., & Bouadjenek, M. R. (2024). *Review for Handling Missing Data with special missing mechanism*. <http://arxiv.org/abs/2404.04905>