

DAFTAR PUSTAKA

- Ade, M., & Sherifdeen, K. (2024). Leveraging XGBoost, Random Forest, and Linear Regression for Accurate Home Price Forecasting. *ResearchGate, September*, 1–41. <https://www.researchgate.net/publication/387463583>
- Agung, A., Daniswara, A., Kadek, I., & Nuryana, D. (2023). Data Preprocessing Pola Pada Penilaian Mahasiswa Program Profesi Guru. *Journal of Informatics and Computer Science*, 05, 97–100.
- Akyap, B. (2023). Machine learning and feature selection: Applications in economics and climate change. *Cambridge University Press*, 2, 1–23. <https://doi.org/10.1017/eds.2023.36>
- Aqsha, D. (2025). Perbandingan Kinerja Algoritma Extreme Gradient Boosting Dan Random Forest Untuk Prediksi Harga Rumah Di Jabodetabek. *Jurnal Ilmu Komputer Dan Sistem Informasi*, 13(1), 1–7. <https://doi.org/10.24912/jiksi.v13i1.32863>
- Arifin, M., Widowati, W., & Farikhin, F. (2023). Ingénierie des Systèmes d'Information Optimization of Hyperparameters in Machine Learning for Enhancing Predictions of Student Academic Performance. *IIETA*, 28(3), 575–582. <https://doi.org/10.18280/isi.280305>
- Aswin, R., Kaderi, S., Anny, Y., Fajar, K., Yaman, R., Shazrin, A., Azmi, M., & Zuhaira, F. (2022). Determining the factors influencing residential property price: A comparative study between Indonesia and Malaysia. *GrowingScience*, 11, 485–496. <https://doi.org/10.5267/dsl.2022.6.002>
- Aufa, S., Santoso, R., & Suparti, S. (2022). Pemodelan Indeks Harga Properti Residensial Di Indonesia Menggunakan Metode Generalized Space Time Autoregressive. *Jurnal Gaussian*, 11(1), 31–44. <https://doi.org/10.14710/j.gauss.v11i1.34001>
- B, J. Y. (2023). *Prediction on Housing Price Based on the Data on Kaggle*. Atlantis Press International BV. <https://doi.org/10.2991/978-94-6463-005-3>

- Çılgin, C., Gökşen, Y., & Gökçen, H. (2023). *The Effect of Outlier Detection Methods in Real Estate Valuation with Machine Learning* *. 5(1), 9–20. <https://doi.org/10.47899/ijss.1270433>
- Deng, L. (2025). *Real estate valuation with multi-source image fusion and enhanced machine learning pipeline.* 1–32. <https://doi.org/10.1371/journal.pone.0321951>
- Dhany, H. W., Izhari, F., & Sutarman. (2023). Journal of Intelligent Decision Support System (IDSS) Exploratory Data Analysis (EDA) methods for healthcare classification. *Journal of Intelligent Decision Support System (IDSS)*, 6(4), 209–215.
- Diewert, W. E., & Shimizu, C. (2022). Residential property price indexes: spatial coordinates versus neighborhood dummy variables. *Review of Income and Wealth*, 3, 770–796. <https://doi.org/10.1111/roiw.12534>
- Emmanuel, T., Maupong, T., Mpoeleng, D., Semong, T., & Mphago, B. (2021). A survey on missing data in machine learning. In *Journal of Big Data*. Springer International Publishing. <https://doi.org/10.1186/s40537-021-00516-9>
- Erdebilli, B., & Devrim-İçtenbaş, B. (2022). Ensemble Voting Regression Based on Machine Learning for Predicting Medical Waste: A Case from Turkey. *Mathematics*, 10(14), 7–9. <https://doi.org/10.3390/math10142466>
- Fauzi, A., Maulidah, N., Supriyadi, R., Nalatissifa, H., & Diantika, S. (2025). Prediksi Harga Properti Di Indonesia Menggunakan Algoritma Random Forest. *Journal of Artificial Intelligence and Digital Business (RIGGS)*, 4(1), 43–49.
- Fitri, E. (2023). Analisis Perbandingan Metode Regresi Linier, Random Forest Regression dan Gradient Boosted Trees Regression Method untuk Prediksi Harga Rumah. *Journal of Applied Computer Science and Technology*, 4(1), 58–64. <https://doi.org/10.52158/jacost.v4i1.491>
- Gamal, A. (2023). Urban clusters and land price variation in Jakarta , Indonesia. *Journal of Property Research*, 00(00), 1–23.

<https://doi.org/10.1080/09599916.2023.2269956>

- Gheisari, M., Hamidpour, H., Liu, Y., Saedi, P., Raza, A., Jalili, A., Rokhsati, H., & Amin, R. (2023). Data Mining Techniques for Web Mining: A Survey. *Artificial Intelligence and Applications*, 1(1), 3–10. <https://doi.org/10.47852/bonviewAIA2202290>
- Hadi, N., & Benedict, J. (2024). Implementasi Machine Learning Untuk Prediksi Harga Rumah Menggunakan Algoritma Random Forest. *Computatio : Journal of Computer Science and Information Systems*, 8(1), 50–61. <https://doi.org/10.24912/computatio.v8i1.15173>
- Haryanto, C., Rahaningsih, N., & Basysyar, F. M. (2023). Komparasi Algoritma Machine Learning Dalam Memprediksi Harga Rumah. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 7(1), 533–539.
- Hibatulloh, M. N., Prakoso, G. D., Putri Yunus, A. D., & Putra, T. D. (2025). Prediksi Harga Rumah di Bandung 2024 Menggunakan Ensemble Learning: Analisis Komparatif dan Interpretabilitas. *Jurnal Informatika: Jurnal Pengembangan IT*, 10(2), 484–493. <https://doi.org/10.30591/jpit.v10i2.8200>
- Igamo, A. M., Saputra, A., & Ismail, R. G. (2023). Monetary Policy and Demographics: Empirical Evidence for Housing Prices in Indonesia. *SRIWIJAYA INTERNATIONAL JOURNAL OF DYNAMIC ECONOMICS AND BUSINESS*, 6(4), 371–384. <https://doi.org/https://doi.org/10.29259/sijdeb.v6i4.371-384>
- Janssen, K. M. J., Mulder, P., & Yudhistira, M. H. (2022). Spatial Sorting of Rich Versus Poor People in Jakarta. *Bulletin of Indonesian Economic Studies*, 58, 167–194. <https://doi.org/10.1080/00074918.2021.1876209>
- Khan, M. K., Khan, A. U., Ahmed, S., & Ali, W. (2026). A Robust Stacked Ensemble Approach for House Price Prediction. *Proceedings of the Pakistan Academy of Sciences: A Physical and Computational Sciences*, 2(63), 123–132. [https://doi.org/http://doi.org/10.53560/PPASA\(63-2\)708](https://doi.org/http://doi.org/10.53560/PPASA(63-2)708)
- Li, H. (2023). House price prediction based on machine learning. *Applied and*

Computational Engineering, 4, 623–628. <https://doi.org/10.54254/2755-2721/4/2023362>

Liang, M., Chang, T., An, B., Duan, X., Du, L., Wang, X., Miao, J., Xu, L., Gao, X., Zhang, L., Li, J., & Gao, H. (2021). A Stacking Ensemble Learning Framework for Genomic Prediction. *Frontiers in Genetics*, 12(March), 1–9. <https://doi.org/10.3389/fgene.2021.600040>

Liao, J. (2023). House price prediction using machine learning : A case study. *Applied and Computational Engineering*, 32, 259–269. <https://doi.org/10.54254/2755-2721/32/20230222>

Lumumba, V. W., Kiprotich, D., Makena, N. G., Kavita, M. D., & Mpaine, M. L. (2024). Comparative Analysis of Cross-Validation Techniques : LOOCV , K-folds Cross-Validation , and Repeated K-folds Cross-Validation in Machine Learning Models. *Science Publishing Group*, 13(5), 127–137. <https://doi.org/https://doi.org/10.11648/j.ajtas.20241305.13>

Malakouti, S. M., Menhaj, M. B., & Suratgar, A. A. (2023). The usage of 10-fold cross-validation and grid search to enhance ML methods performance in solar farm power generation prediction. *Cleaner Engineering and Technology*, 15(July), 100664. <https://doi.org/10.1016/j.clet.2023.100664>

Mienye, I. D., Sun, Y., & Member, S. (2022). A Survey of Ensemble Learning : Concepts , Algorithms , Applications , and Prospects. *IEEE Access*, 10(August), 99129–99149. <https://doi.org/10.1109/ACCESS.2022.3207287>

Moscovich, A., & Rosset, S. (2022). On the cross-validation bias due to unsupervised preprocessing. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 84(4), 1474–1502. <https://doi.org/10.1111/rssb.12537>

Nascimento, M., Carolina, A., Nascimento, C., Azevedo, C. F., Carlos, A., Oliveira, B. De, Caixeta, E. T., & Jarquin, D. (2024). Enhancing genomic prediction with Stacking Ensemble Learning in Arabica Coffee. *Frontiers in Plant Science*, 15, 1–14. <https://doi.org/10.3389/fpls.2024.1373318>

- Nguyen, Q. H., Ly, H. B., Ho, L. S., Al-Ansari, N., Van Le, H., Tran, V. Q., Prakash, I., & Pham, B. T. (2021). Influence of data splitting on performance of machine learning models in prediction of shear strength of soil. *Mathematical Problems in Engineering*, 4832864, 15. <https://doi.org/10.1155/2021/4832864>
- Nti, I. K., Nyarko-Boateng, O., & Aning, J. (2021). Performance of Machine Learning Algorithms with Different K Values in K-fold CrossValidation. *International Journal of Information Technology and Computer Science*, 13(6), 61–71. <https://doi.org/10.5815/ijitcs.2021.06.05>
- Nuris, N. (2024). Analisis Prediksi Harga Rumah Pada Machine Learning Metode Regresi Linear. *Explore*, 14(2), 108–112. <https://doi.org/10.35200/ex.v14i2.123>
- 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>
- Park, U., Kang, Y., Lee, H., & Yun, S. (2022). A Stacking Heterogeneous Ensemble Learning Method for the Prediction of Building Construction Project Costs. *Appl. Sci*, 12(9729), 11.
- Prusty, S., Patnaik, S., & Dash, S. K. (2022). SKCV : Strati fi ed K-fold cross-validation on ML classi fi ers for predicting cervical cancer. *Frontiers in Nanotechnology*, 4(972421), 1–12. <https://doi.org/10.3389/fnano.2022.972421>
- Putri, N. H., Azizah, R. N., Prasetya, R. P. D., Simanjuntak, G. A. D. P., & Zuhri, S. (2024). Analisis Dampak Warga Negara Asing dalam Membuka Usaha Ilegal di Bali. *Indonesian Culture and Religion Issues*, 1(2), 14. <https://doi.org/10.47134/diksima.v1i2.17>
- Qolbi, M. B. S., Puteh, T. N., Rivandi, R., & Rozikin, C. (2025). Prediksi Harga Rumah Di Jakarta Pusat Menggunakan Algoritma Machine Learning. *Jurnal Ilmu Komputer Dan Bisnis*, 16(1), 16–24.

<https://doi.org/10.47927/jikb.v16i1.840>

Rayadin, M. A., Musaruddin, M., Saputra, R. A., & Isnawaty, I. (2024). Implementasi Ensemble Learning Metode XGBoost dan Random Forest untuk Prediksi Waktu Penggantian Baterai Aki. *BIOS : Jurnal Teknologi Informasi Dan Rekayasa Komputer*, 5(2), 111–119. <https://doi.org/10.37148/bios.v5i2.128>

Saiful, A., Andryana, S., & Gunaryati, A. (2021). Prediksi Harga Rumah Menggunakan Web Scrapping dan Machine Learning Dengan Algoritma Linear Regression. *JATISI (Jurnal Teknik Informatika Dan Sistem Informasi)*, 8(1), 41–50. <https://doi.org/10.35957/jatisi.v8i1.701>

Saputra, R. A., & Pratama, A. (2025). Implementasi Decision Tree Untuk Prediksi Harga Rumah Di Daerah Tebet. *Journal of Information System Management (JOISM)*, 6(2), 164–170. <https://doi.org/10.24076/joism.2025v6i2.1928>

Sembiring, M. A., & Sembiring, F. W. (2024). Analisa Kinerja Model Regresi Dalam Machine Learning Untuk Memprediksi Harga Beras. *Journal Of Information Systems And Informatics Engineering*, 8(1), 144–152. <https://doi.org/10.35145/joisie.v8i1.3902>

Sharma, H., Harsora, H., & Ogunleye, B. (2024). An Optimal House Price Prediction Algorithm: XGBoost. *Analytics*, 3(1), 30–45. <https://doi.org/10.3390/analytics3010003>

Soares, T. G., Tonggiroh, M., Erkamim, M., & Widarti, E. (2025). Enhancing Liver Disease Classification Using Support Vector Machine with IQR-Based Outlier Handling. *JURNAL ILMIAH FIFO*, 17(1), 91–101. <https://doi.org/10.22441/fifo.2025.v17i1.010>

Suparmadi, & Ramadhani, A. (2022). Sistem Estimasi Pencapaian Target Profit Menggunakan Model Regresi Berbasis Machine Learning. *Journal of Science and Social Research*, 5(3), 703–708. <https://doi.org/10.54314/jssr.v5i3.1042>

Tegar, A., Satria, B., Hasanah, H., & Oktaviani, I. (2025). Predicting Social Media Addiction Using Machine Learning and Interactive Visualization with

Streamlit. *Komunitas Dosen Indonesia*, 8(1), 778–788.
<https://doi.org/10.32877/bt.v8i1.2715>

Thakker, Z. L., & Buch, S. H. (2024). Effect of Feature Scaling Pre-processing Techniques on Machine Learning Algorithms to Predict Particulate Matter Concentration for Gandhinagar, Gujarat, India. *International Journal of Scientific Research in Science and Technology*, 11(1), 410–419.
<https://doi.org/10.32628/ijrst52411150>

Verdonck, T., Baesens, B., & Óskarsdóttir, M. (2024). Special issue on feature engineering editorial. *Machine Learning*, 113(7), 3917–3928.
<https://doi.org/10.1007/s10994-021-06042-2>

Wardani, A. P., Irawan, H. A., Syah, M. P., Akmal, M. A., Nariswari, N. U., & Hindrayani, K. M. (2024). Analisis dan Prediksi Harga Properti Rumah di Kota Surabaya dengan Algoritma Random Forest. *Prosiding Seminar Nasional Sains Data*, 4(1), 885–894.
<https://doi.org/10.33005/senada.v4i1.375>

Wu, W. T., Li, Y. J., Feng, A. Z., Li, L., Huang, T., Xu, A. D., & Lyu, J. (2021). Data mining in clinical big data: the frequently used databases, steps, and methodological models. *Military Medical Research*, 8(1), 1–12.
<https://doi.org/10.1186/s40779-021-00338-z>

Yu, J., Hamdan, S., Sasse, L., Morrison, A., & Patil, K. R. (2024). Empirical Comparison between Cross-Validation and Mutation-Validation in Model Selection. *ArXiv, Cv*, 1–12. <https://doi.org/10.48550/arXiv.2311.14079>

Zhang, Y., Ma, J., Liang, S., Li, X., & Liu, J. (2022). A stacking ensemble algorithm for improving the biases of forest aboveground biomass estimations from multiple remotely sensed datasets. *GIScience and Remote Sensing*, 59(1), 234–249. <https://doi.org/10.1080/15481603.2021.2023842>

Zhao, Y., Zhang, W., & Liu, X. (2024). Grid search with a weighted error function : Hyper-parameter optimization for financial time series forecasting. *Applied Soft Computing*, 154(111362), 20.
<https://doi.org/https://doi.org/10.1016/j.asoc.2024.111362>