

DAFTAR PUSTAKA

- Altmann, A., Toloşi, L., Sander, O., & Lengauer, T. (2010). Permutation importance: A corrected feature importance measure. *Bioinformatics*, 26(10), 1340–1347. <https://doi.org/10.1093/bioinformatics/btq134>
- Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys*, 4, 40–79. <https://doi.org/10.1214/09-SS054>
- Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 7, 1–24. <https://doi.org/10.7717/PEERJ-CS.623>
- Hancock, J. T., & Khoshgoftaar, T. M. (2020). Survey on categorical data for neural networks. *Journal of Big Data*, 7(1). <https://doi.org/10.1186/s40537-020-00305-w>
- Hyndman, R. J., & Koehler, A. B. (2006). *Another look at measures of forecast accuracy*. International Journal of Forecasting. <https://doi.org/10.1016/j.ijforecast.2006.03.001>
- Jin, Z., Shang, J., Zhu, Q., Ling, C., Xie, W., & Qiang, B. (2020). RFRSF: Employee Turnover Prediction Based on Random Forests and Survival Analysis. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12343 LNCS, 503–515. https://doi.org/10.1007/978-3-030-62008-0_35
- Kaufman, S., Rosset, S., Perlich, C., & Stitelman, O. (2012). Leakage in data

- mining: Formulation, detection, and avoidance. *ACM Transactions on Knowledge Discovery from Data*, 6(4), 1–21.
<https://doi.org/10.1145/2382577.2382579>
- KURGAN, L. A., And, & MUSILEK, P. (2006). *A survey of Knowledge Discovery and Data Mining process models*. Cambridge Core.
<https://www.cambridge.org/core/journals/knowledge-engineering-review/article/abs/survey-of-knowledge-discovery-and-data-mining-process-models/368D6AFE435EB5E30378398D34D61C17>
- Mitchell, M. I. J. and T. M. (2015). *Machine learning: Trends, perspectives, and prospects*. Science. <https://www.science.org/doi/10.1126/science.aaa8415>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Probst, P., Wright, M. N., & Boulesteix, A. L. (2019). *Hyperparameters and tuning strategies for random forest*. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. <https://doi.org/10.1002/widm.1301>
- Sembiring, A., & Santoso, H. (2026). Comparative Academic Performance Prediction in Primary Schools Using Linear Regression and Random Forest. *Sinkron*, 10(2), 1104–1113. <https://doi.org/10.33395/sinkron.v10i2.15953>
- Soraya, N. S., & Hendry, H. (2023). Komparasi linear regression, random forest regression, dan multilayer perceptron regression untuk prediksi tren musik

TikTok. *Aiti*, 20(2), 191–205. <https://doi.org/10.24246/aiti.v20i2.191-205>

Trevor Hastie , Robert Tibshirani, J. F. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. Springer.
<https://link.springer.com/book/10.1007/978-0-387-84858-7>

U. Fayyad, G. Piatetsky-Shapiro, P. S. (1996). *From Data Mining to Knowledge Discovery in Databases*. Semantic Scholar.
<https://www.semanticscholar.org/paper/From-Data-Mining-to-Knowledge-Discovery-in-Fayyad-Piatetsky-Shapiro/a1874aafa8730bdd4b28f29d025141c13ee28b58>