

DAFTAR PUSTAKA

- [1] D. Tomic, J. E. Shaw, and D. J. Magliano, "The burden and risks of emerging complications of diabetes mellitus," *Nat. Rev. Endocrinol.*, vol. 18, no. 9, pp. 525–539, 2022, doi: 10.1038/s41574-022-00690-7.
- [2] R. Pahlevi, "Proyeksi Jumlah Penderita Diabetes di Indonesia," *Databoks*, 2021. <https://databoks.katadata.co.id/layanan-konsumen-kesehatan/statistik/8380d72a010b4f0/jumlah-penderita-diabetes-di-indonesia-diproyeksikan-capai-2857-juta-pada-2045> (accessed Jun. 01, 2025).
- [3] World Health Organization, "Diabetes," *World Health Organization*, 2023. <https://www.who.int/news-room/fact-sheets/detail/diabetes> (accessed Jun. 11, 2024).
- [4] R. Unnikrishnan, V. N. Shah, and V. Mohan, "Challenges in diagnosis and management of diabetes in the young," *Clin. Diabetes Endocrinol.*, vol. 2, no. 1, pp. 1–9, 2016, doi: 10.1186/s40842-016-0036-6.
- [5] K. Anissa, H. Rumahorbo, and S. Wahyuni, "Development of Instruments Test to Detect Diabetes Mellitus in Pregnancy," *J. Kebidanan Unimus*, vol. 12, no. 1, pp. 27–36, 2023, doi: 10.26714/jk.12.1.2023.27-36.
- [6] K. Alnowaiser, "Improving Healthcare Prediction of Diabetic Patients Using KNN Imputed Features and Tri-Ensemble Model," *IEEE Access*, vol. 12, no. February, pp. 16783–16793, 2024, doi: 10.1109/ACCESS.2024.3359760.
- [7] A. Kumar, M. Boehm, and J. Yang, "Data management in machine learning: Challenges, techniques, and systems," *Proc. ACM SIGMOD Int. Conf. Manag. Data*, vol. Part F1277, no. Section 4, pp. 1717–1722, 2017, doi: 10.1145/3035918.3054775.
- [8] K. Psychogyios, L. Ilias, C. Ntanos, and D. Askounis, "Missing Value Imputation Methods for Electronic Health Records," *IEEE Access*, vol. 11, no. February, pp. 21562–21574, 2023, doi: 10.1109/ACCESS.2023.3251919.
- [9] P. Ranganathan and S. Hunsberger, "Handling missing data in research," *Perspect. Clin. Res.*, vol. 15, no. 2, pp. 99–101, 2024, doi: 10.4103/picr.picr_38_24.
- [10] W. Sudrajat and I. Cholid, "K-Nearest Neighbor (K-Nn) Untuk Penanganan Missing Value Pada Data Umkm," *J. Rekayasa Sist. Inf. dan Teknol.*, vol. 1, no. 2, pp. 54–63, 2023, doi: 10.59407/jrsit.v1i2.77.
- [11] M. V.K., "Dealing with Missing Values in a Relation Dataset Using the DROPNA Function in Python," in *Mathematics and Computer Science Volume 1*, vol. 1, and S. Das S. Ghosh, M. Niranjnamurthy, K. Deyasi, B.B. Mallik, Ed. 2023. doi: 10.1002/9781119879831.ch25.

- [12] A. Juna *et al.*, “Water Quality Prediction Using KNN Imputer and Multilayer Perceptron,” *Water (Switzerland)*, vol. 14, no. 17, pp. 1–19, 2022, doi: 10.3390/w14172592.
- [13] W. M. Hameed and N. A. Ali, “Comparison of Seventeen Missing Value Imputation Techniques,” *J. Hunan Univ. Nat. Sci.*, vol. 49, no. 7, pp. 26–36, 2022, doi: 10.55463/issn.1674-2974.49.7.4.
- [14] A. Fadlil, Herman, and D. Praseptian M, “K Nearest Neighbor Imputation Performance on Missing Value Data Graduate User Satisfaction,” *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 4, pp. 570–576, 2022, doi: 10.29207/resti.v6i4.4173.
- [15] V. Chang, J. Bailey, Q. A. Xu, and Z. Sun, “Pima Indians diabetes mellitus classification based on machine learning (ML) algorithms,” *Neural Comput. Appl.*, vol. 35, no. 22, pp. 16157–16173, 2023, doi: 10.1007/s00521-022-07049-z.
- [16] H. Karamti *et al.*, “Improving Prediction of Cervical Cancer Using KNN Imputed SMOTE Features and Multi-Model Ensemble Learning Approach,” *Cancers (Basel)*, vol. 15, no. 17, 2023, doi: 10.3390/cancers15174412.
- [17] S. Jäger, A. Allhorn, and F. Bießmann, “A Benchmark for Data Imputation Methods,” *Front. Big Data*, vol. 4, no. July, pp. 1–16, 2021, doi: 10.3389/fdata.2021.693674.
- [18] J. Mantik, A. Susanto, E. Wijaya, and D. Djohan, “Analysis of the random forest and grid search algorithms in early detection of diabetes mellitus disease,” *J. Mantik*, vol. 7, no. 2, pp. 2685–4236, 2023.
- [19] F. I. Tushar *et al.*, “Virtual imaging trials improved the transparency and reliability of AI systems in COVID-19 imaging,” pp. 1–17, 2023, [Online]. Available: <http://arxiv.org/abs/2308.09730>
- [20] J. Schwerter, K. Gurtskaia, A. Romero, B. Zeyer-Gliozzo, and M. Pauly, “Evaluating tree-based imputation methods as an alternative to MICE PMM for drawing inference in empirical studies,” *arXiv Prepr.*, pp. 1–47, 2024.
- [21] S. V. Mbona, P. Ndlovu, H. Mwambi, and S. Ramroop, “Multiple imputation using chained equations for missing data in survival models: applied to multidrug-resistant tuberculosis and HIV data,” *J. Public Health Africa*, vol. 14, no. 8, pp. 12–18, 2023, doi: 10.4081/jphia.2023.2388.
- [22] I. Nirmala, H. Wijayanto, and K. A. Notodiputro, “Prediction of Undergraduate Student’s Study Completion Status Using MissForest Imputation in Random Forest and XGBoost Models,” *ComTech Comput. Math. Eng. Appl.*, vol. 13, no. 1, pp. 53–62, 2022, doi: 10.21512/comtech.v13i1.7388.

- [23] A. Finocchi *et al.*, “Multiple imputation integrated to machine learning: predicting post-stroke recovery of ambulation after intensive inpatient rehabilitation,” *Sci. Rep.*, vol. 14, no. 1, p. 25188, 2024, doi: 10.1038/s41598-024-74537-8.
- [24] J. H. Li *et al.*, “Comparison of the effects of imputation methods for missing data in predictive modelling of cohort study datasets,” *BMC Med. Res. Methodol.*, vol. 24, no. 1, pp. 1–9, 2024, doi: 10.1186/s12874-024-02173-x.
- [25] M. H. Nadimi-Shahraki, S. Mohammadi, H. Zamani, M. Gandomi, and A. H. Gandomi, “A hybrid imputation method for multi-pattern missing data: A case study on type ii diabetes diagnosis,” *Electron.*, vol. 10, no. 24, 2021, doi: 10.3390/electronics10243167.
- [26] D. J. Magliano and E. J. Boyko, *IDF DIABETES ATLAS*, 10th Editi. Brussels, 2021. [Online]. Available: <https://www.ncbi.nlm.nih.gov/books/NBK581934/>
- [27] J. Loscalzo, A. Fauci, D. Kasper, S. Hauser, D. Longo, and J. L. Jameson, *Harrison’s Principles of Internal Medicine*, 21st Editi. New York, NY: McGraw-Hill Education, 2022. [Online]. Available: accessmedicine.mhmedical.com/content.aspx?aid=1197729784
- [28] J. Abadie, “Normal Laboratory Values,” *MSD Manual*, 2024. <https://www.msdmanuals.com/professional/resources/normal-laboratory-values/normal-laboratory-values> (accessed Mar. 18, 2025).
- [29] C. J. Harrison and C. J. Sidey-Gibbons, “Machine learning in medicine: a practical introduction to natural language processing,” *BMC Med. Res. Methodol.*, vol. 21, no. 1, pp. 1–18, 2021, doi: 10.1186/s12874-021-01347-1.
- [30] R. Miotto, F. Wang, S. Wang, X. Jiang, and J. T. Dudley, “Deep learning for healthcare: Review, opportunities and challenges,” *Brief. Bioinform.*, vol. 19, no. 6, pp. 1236–1246, 2017, doi: 10.1093/bib/bbx044.
- [31] R. Little and D. Rubin, *Statistical Analysis with Missing Data*, Third Edition. Wiley, 2019. doi: 10.1002/9781119482260.
- [32] A. R. Ismail, N. Z. Abidin, and M. K. Maen, “Systematic Review on Missing Data Imputation Techniques with Machine Learning Algorithms for Healthcare,” *J. Robot. Control*, vol. 3, no. 2, pp. 143–152, Feb. 2022, doi: 10.18196/jrc.v3i2.13133.
- [33] M. Alabadla *et al.*, “Systematic Review of Using Machine Learning in Imputing Missing Values,” *IEEE Access*, vol. 10, pp. 44483–44502, 2022, doi: 10.1109/ACCESS.2022.3160841.
- [34] L. Beretta and A. Santaniello, “Nearest neighbor imputation algorithms: A critical evaluation,” *BMC Med. Inform. Decis. Mak.*, vol. 16, no. Suppl 3,

2016, doi: 10.1186/s12911-016-0318-z.

- [35] M. J. Azur, E. A. Stuart, C. Frangakis, and P. J. Leaf, “Multiple imputation by chained equations: what is it and how does it work?,” *Int. J. Methods Psychiatr. Res.*, vol. 20, no. 1, pp. 40–49, Mar. 2011, doi: 10.1002/mpr.329.
- [36] D. J. Stekhoven and P. Bühlmann, “MissForest--non-parametric missing value imputation for mixed-type data,” *Bioinformatics*, vol. 28, no. 1, pp. 112–118, 2012, doi: 10.1093/bioinformatics/btr597.
- [37] R. Mohammed, J. Rawashdeh, and M. Abdullah, “Machine Learning with Oversampling and Undersampling Techniques: Overview Study and Experimental Results,” *2020 11th Int. Conf. Inf. Commun. Syst. ICICS 2020*, pp. 243–248, 2020, doi: 10.1109/ICICS49469.2020.239556.
- [38] L. Breiman, “Random Forests,” *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.
- [39] J. Bergstra, J. B. Ca, and Y. B. Ca, “Random Search for Hyper-Parameter Optimization,” *J. Mach. Learn. Res.*, vol. 13, no. 10, pp. 281–305, 2012, Accessed: Jun. 18, 2025. [Online]. Available: <http://jmlr.org/papers/v13/bergstra12a.html>
- [40] J. Liang, “Confusion Matrix: Machine Learning,” *PAC*, vol. 3, no. 4, 2022, [Online]. Available: <https://pac.pogil.org/index.php/pac/article/view/304>
- [41] N. Japkowicz and Z. Boukouvalas, “Machine learning evaluation : towards reliable and responsible AI,” 2025.
- [42] G. W. Corder and D. I. Foreman, *NONPARAMETRIC STATISTICS A Step-by-Step Approach*, 2nd ed., vol. 11, no. 1. 2014. [Online]. Available: http://scioteca.caf.com/bitstream/handle/123456789/1091/RED2017-Eng-8ene.pdf?sequence=12&isAllowed=y%0Ahttp://dx.doi.org/10.1016/j.regsciurbeco.2008.06.005%0Ahttps://www.researchgate.net/publication/305320484_SISTEM_PEMBETUNGAN_TERPUSAT_STRATEGI_MELESTARI