

DAFTAR PUSTAKA

- [1] S. Kamwaga, D. M. M. Mulungu, and P. Valimba, “Assessment of empirical and regression methods for infilling missing streamflow data in Little Ruaha catchment Tanzania,” *Phys. Chem. Earth*, vol. 106, no. May 2016, pp. 17–28, 2018, doi: 10.1016/j.pce.2018.05.008.
- [2] R. J. Abrahart *et al.*, “Two decades of anarchy? Emerging themes and outstanding challenges for neural network river forecasting,” *Prog. Phys. Geogr.*, vol. 36, no. 4, pp. 480–513, 2012, doi: 10.1177/0309133312444943.
- [3] E. Acu, “Classification, Clustering, and Data Mining Applications,” *Classif. Clust. Data Min. Appl.*, no. January 2004, 2004, doi: 10.1007/978-3-642-17103-1.
- [4] J. Luengo, S. García, and F. Herrera, “A study on the use of imputation methods for experimentation with Radial Basis Function Network classifiers handling missing attribute values: The good synergy between RBFNs and EventCovering method,” *Neural Networks*, vol. 23, no. 3, pp. 406–418, Apr. 2010, doi: 10.1016/J.NEUNET.2009.11.014.
- [5] L. Sunitha, M. Balraju, and J. Sasikiran, “Data Mining: Estimation of Missing Values Using Lagrange Interpolation Technique,” *Int. J. Adv. Res. Comput. Eng. Technol.*, vol. 2, no. 4, pp. 1579–1582, 2013, [Online]. Available: <http://ijarcet.org/wp-content/uploads/IJARCET-VOL-2-ISSUE-4-1579-1582.pdf>.
- [6] A. Jadhav, D. Pramod, and K. Ramanathan, “Comparison of Performance

- of Data Imputation Methods for Numeric Dataset,” *Appl. Artif. Intell.*, vol. 33, no. 10, pp. 913–933, 2019, doi: 10.1080/08839514.2019.1637138.
- [7] C. Curley, R. M. Krause, R. Feiock, and C. V. Hawkins, “Dealing with Missing Data: A Comparative Exploration of Approaches Using the Integrated City Sustainability Database,” *Urban Aff. Rev.*, vol. 55, no. 2, pp. 591–615, 2019, doi: 10.1177/1078087417726394.
- [8] R. J. A. Little, “Missing-data adjustments in large surveys,” *J. Bus. Econ. Stat.*, vol. 6, no. 3, pp. 287–296, 1988, doi: 10.1080/07350015.1988.10509663.
- [9] D. B. Rubin, “Inference and missing data,” *Biometrika*, vol. 63, no. 3, pp. 581–592, 1976, doi: 10.1093/biomet/63.3.581.
- [10] J. L. Schafer and J. W. Graham, “Missing data: Our view of the state of the art,” *Psychol. Methods*, vol. 7, no. 2, pp. 147–177, 2002, doi: 10.1037/1082-989X.7.2.147.
- [11] D. Ferguson and W. Winkler, “Glossary of Terms on Statistical Data Editing,” *Conf. Eur. Stat. Methodol. Mater. Gloss.*, 2000, [Online]. Available: https://ec.europa.eu/eurostat/ramon/statmanuals/files/UN_editing_glossary_2000.pdf.
- [12] W. M. Champion and D. B. Rubin, “Multiple Imputation for Nonresponse in Surveys,” *J. Mark. Res.*, vol. 26, no. 4, p. 485, 1989, doi: 10.2307/3172772.
- [13] Statistics Canada, “Catalogue no. 12-539,” *Stat. Canada Qual. Guidel.*, no.

- 12, pp. 21–24, 2003.
- [14] S. P. Mandel J, “A Comparison of Six Methods for Missing Data Imputation,” *J. Biom. Biostat.*, vol. 06, no. 01, pp. 1–6, 2015, doi: 10.4172/2155-6180.1000224.
- [15] G. Chhabra, V. Vashisht, and J. Ranjan, “A review on missing data value estimation using imputation algorithm,” *J. Adv. Res. Dyn. Control Syst.*, vol. 11, no. 7 Special Issue, pp. 312–318, 2019.
- [16] A. Kowarik and M. Templ, “Imputation with the R package VIM,” *J. Stat. Softw.*, vol. 74, no. 7, 2016, doi: 10.18637/jss.v074.i07.
- [17] Doreswamy, I. Gad, and B. R. Manjunatha, “Performance evaluation of predictive models for missing data imputation in weather data,” *2017 Int. Conf. Adv. Comput. Commun. Informatics, ICACCI 2017*, vol. 2017-Janua, pp. 1327–1334, 2017, doi: 10.1109/ICACCI.2017.8126025.
- [18] E. C. Blessie, E. Karthikeyan, and V. Thavavel, “An Extended RELIEF-DISC for Handling of Incomplete Data to Improve the Classifier Performance,” 2012.
- [19] S. Jain and K. Jain, “Estimation of Missing Attribute Value in Time Series Database in Data Mining,” *Glob. J. Comput. Sci. Technol.*, 2017.
- [20] D. M. P. Murti, U. Pujianto, A. P. Wibawa, and M. I. Akbar, “K-Nearest Neighbor (K-NN) based Missing Data Imputation,” *Proceeding - 2019 5th Int. Conf. Sci. Inf. Technol. Embrac. Ind. 4.0 Towar. Innov. Cyber Phys. Syst. ICSITech 2019*, pp. 83–88, 2019, doi:

10.1109/ICSITech46713.2019.8987530.

- [21] G. Chhabra, V. Vashisht, and J. Ranjan, “A Comparison of Multiple Imputation Methods for Data with Missing Values,” *Indian J. Sci. Technol.*, vol. 10, no. 19, pp. 1–7, 2017, doi: 10.17485/ijst/2017/v10i19/110646.
- [22] L. D. Jackel, “Limits in Learning Machine Accuracy Imposed by Data Quality,” *KDD-95 Proc.*, vol. 1, no. KDD-95, pp. 57–62, 1994.
- [23] J. M. Engels and P. Diehr, “Imputation of missing longitudinal data: A comparison of methods,” *J. Clin. Epidemiol.*, vol. 56, no. 10, pp. 968–976, 2003, doi: 10.1016/S0895-4356(03)00170-7.
- [24] J. Maillo, S. Ramírez, I. Triguero, and F. Herrera, “kNN-IS: An Iterative Spark-based design of the k-Nearest Neighbors classifier for big data,” *Knowledge-Based Syst.*, vol. 117, pp. 3–15, 2017, doi: 10.1016/j.knosys.2016.06.012.
- [25] G. E. A. P. A. Batista and M. C. Monard, “A study of k-nearest neighbour as an imputation method,” *Front. Artif. Intell. Appl.*, vol. 87, no. January, pp. 251–260, 2002.
- [26] B. Suthar, H. Patel, and A. Goswami, “A Survey: Classification of Imputation Methods in Data Mining,” *Int. J. Emerg. Technol. Adv. Eng.*, vol. 2, no. 1, pp. 309–312, 2012, [Online]. Available: www.ijetae.com.
- [27] D. Priya, R. Sivaraj, R. Assistant, and S. Gr, “a Review of Missing Data Handling Methods,” *Int. J. Eng. Technol. Sci. – IJETSTM ISSN*, vol. 2, no. 2, pp. 2349–3968, 2015.

- [28] E. Fix and J. L. Hodges, “Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties,” *Int. Stat. Rev. / Rev. Int. Stat.*, vol. 57, no. 3, pp. 238–247, Oct. 1989, doi: 10.2307/1403797.
- [29] T. Cover and P. Hart, “Nearest neighbor pattern classification,” *IEEE Trans. Inf. Theory*, vol. 13, no. 1, pp. 21–27, 1967, doi: 10.1109/TIT.1967.1053964.
- [30] G. L. Pritalia, “Klasifikasi Gaya Belajar Global Dan Sekuensial Pada Multimedia Learning Menggunakan Pendekatan Eye-Tracking Dan Machine Learning Pada Multimedia Learning Menggunakan Pendekatan Eye-Tracking Dan Machine Learning,” Yogyakarta, 2020.
- [31] Q. Wang and J. N. K. Rao, “Empirical likelihood-based inference under imputation for missing response data,” *Ann. Stat.*, vol. 30, no. 3, pp. 896–924, 2002, doi: 10.1214/aos/1028674845.
- [32] Y. Sun, A. K. C. Wong, and M. S. Kamel, “Classification of imbalanced data: A review,” *Int. J. Pattern Recognit. Artif. Intell.*, vol. 23, no. 4, pp. 687–719, 2009, doi: 10.1142/S0218001409007326.
- [33] A. MELISSA J, “Multiple imputation by chained equations: what is it and how does it work?,” *Int. J. Methods Psychiatr. Res.*, vol. 17 Suppl 1, no. 1, pp. S78–S82, 2008, doi: 10.1002/mpr.
- [34] P. Hayati Rezvan, K. J. Lee, and J. A. Simpson, “The rise of multiple imputation: A review of the reporting and implementation of the method in medical research Data collection, quality, and reporting,” *BMC Med. Res. Methodol.*, vol. 15, no. 1, pp. 1–14, 2015, doi: 10.1186/s12874-015-0022-1.

- [35] S. van Buuren and K. Groothuis-Oudshoorn, “mice: Multivariate imputation by chained equations in R,” *J. Stat. Softw.*, vol. 45, no. 3, pp. 1–67, 2011, doi: 10.18637/jss.v045.i03.
- [36] U. Azmi, Z. N. Hadi, and S. Soraya, “ARDL METHOD: Forecasting Data Curah Hujan Harian NTB,” *J. Varian*, vol. 3, no. 2, pp. 73–82, 2020, doi: 10.30812/varian.v3i2.627.
- [37] D. F. Hamilton, M. Ghert, and A. H. R. W. Simpson, “Interpreting regression models in clinical outcome studies,” *Bone Jt. Res.*, vol. 4, no. 9, pp. 152–153, 2015, doi: 10.1302/2046-3758.49.2000571.
- [38] T. Chai and R. R. Draxler, “Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature,” *Geosci. Model Dev.*, vol. 7, no. 3, pp. 1247–1250, 2014, doi: 10.5194/gmd-7-1247-2014.
- [39] A. A. Suryanto, “Penerapan Metode Mean Absolute Error (Mea) Dalam Algoritma Regresi Linear Untuk Prediksi Produksi Padi,” *Saintekbu*, vol. 11, no. 1, pp. 78–83, 2019, doi: 10.32764/saintekbu.v11i1.298.
- [40] W. Wang and Y. Lu, “Analysis of the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) in Assessing Rounding Model,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 324, no. 1, 2018, doi: 10.1088/1757-899X/324/1/012049.
- [41] M. J. Hartmann and G. Carleo, “Neural-Network Approach to Dissipative Quantum Many-Body Dynamics,” *Phys. Rev. Lett.*, vol. 122, no. 25, p. 250502, 2019, doi: 10.1103/PhysRevLett.122.250502.

- [42] K. Crammer, “On the algorithmic implementation of multiclass kernel-based vector machines,” *J. Mach. Learn. Res. - JMLR*, vol. 2, no. 2, pp. 265–292, 2002.
- [43] X. Wu *et al.*, *Top 10 algorithms in data mining*, vol. 14, no. 1. 2008.
- [44] J. Shao, “Linear model selection by cross-validation,” *J. Am. Stat. Assoc.*, vol. 88, no. 422, pp. 486–494, 1993, doi: 10.1080/01621459.1993.10476299.
- [45] Y. Sarvina, “Pemanfaatan Software Open Source ‘R’ Untuk Penelitian Agroklimat,” *Inform. Pertan.*, vol. 26, no. 1, p. 23, 2017, doi: 10.21082/ip.v26n1.2017.p23-30.
- [46] R-project, “What is R?,” *CRAN*. <https://www.r-project.org/about.html>.
- [47] A. F. Zuur, E. N. Ieno, and E. Meesters, “Introduction,” in *A Beginner’s Guide to R*, New York, NY: Springer New York, 2009, pp. 1–27.
- [48] P. L. S. D. Air, “Katalog Basis Data 2014 Sumber Daya Air,” *Badan Penelit. dan Pengemb. Kementrian Pekerj. Umum*, vol. 1, no. 1, pp. 341–353, 2014.
- [49] S. L. Dingman, *Physical Hydrology: Third Edition*, Third Edit. Illinois: Waveland Press, 2015.
- [50] S. J. Goldman, T. A. Bursztynsky, and K. Jackson, *Erosion and sediment control handbook*, 1st ed. New York: McGraw-Hill, 1986.
- [51] M. Rivki and A. M. Bachtiar, “IMPLEMENTASI ALGORITMA K-NEAREST NEIGHBOR DALAM PENGKLASIFIKASIAN FOLLOWER TWITTER YANG MENGGUNAKAN BAHASA INDONESIA,” *J. Sist.*

- Inf.*, vol. 13, no. 1, pp. 31–37, May 2017, doi: 10.21609/jsi.v13i1.500.
- [52] J. Han, M. Kamber, and J. Pei, “3 - Data Preprocessing,” in *Data Mining (Third Edition)*, Third Edit., J. Han, M. Kamber, and J. Pei, Eds. Boston: Morgan Kaufmann, 2012, pp. 83–124.
- [53] A. Khairi, “IMPLEMENTASI K-NEAREST NEIGHBOR (KNN) UNTUK KLASIFIKASI MASYARAKAT PRA SEJAHTERA DESA SAPIKEREK KECAMATAN SUKAPURA,” *TRILOGI J. Ilmu Teknol. Kesehatan, dan Hum.*, vol. 2, no. 3, pp. 319–323, 2021.
- [54] T. Hartono, “IMPUTATION USING FUZZY K-MEANS FOR MISSING DATA CASE STUDY LARGE MANUFACTURING ESTABLISHMENT DATA OF EAST JAVA PROVINCE 2008,” Institut Teknologi Sepuluh Nopember Surabaya, 2011.
- [55] M. Nishom and M. Y. Fathoni, “Implementasi Pendekatan Rule-Of-Thumb untuk Optimasi Algoritma K-Means Clustering,” *J. Inform. J. Pengemb. IT*, vol. 3, no. 2, pp. 237–241, 2018, doi: 10.30591/jpit.v3i2.909.