

DAFTAR PUSTAKA

- [1] A. F. S. Borges, F. J. B. Laurindo, M. M. Spínola, R. F. Gonçalves, and C. A. Mattos, “The strategic use of artificial intelligence in the digital era: Systematic literature review and future research directions,” *Int J Inf Manage*, vol. 57, p. 102225, Apr. 2021, doi: 10.1016/J.IJINFOMGT.2020.102225.
- [2] S. L. Wamba-Taguimdje, S. Fosso Wamba, J. R. Kala Kamdjoug, and C. E. Tchatchouang Wanko, “Influence of artificial intelligence (AI) on firm performance: the business value of AI-based transformation projects,” *Business Process Management Journal*, vol. 26, no. 7, pp. 1893–1924, 2020, doi: 10.1108/BPMJ-10-2019-0411/FULL/XML.
- [3] A. Suprun, T. Petrishina, and I. Vasylichuk, “Competition and cooperation between fintech companies and traditional financial institutions,” *E3S Web of Conferences*, vol. 166, no. 13028, Apr. 2020, doi: 10.1051/E3SCONF/202016613028.
- [4] Otoritas Jasa Keuangan, “Statistik P2P Lending Periode September 2023,” Otoritas Jasa Keuangan. Accessed: Jan. 07, 2024. [Online]. Available: <https://www.ojk.go.id/id/kanal/iknb/data-dan-statistik/fintech/Pages/Statistik-P2P-Lending-Periode-September-2023.aspx>
- [5] H. Purwanto, D. Yandri, and M. P. Yoga, “Perkembangan dan Dampak Financial Technology (Fintech) Terhadap Perilaku Manajemen Keuangan di Masyarakat,” *Kompleksitas: Jurnal Ilmiah Manajemen, Organisasi Dan Bisnis*, vol. 11, no. 1, pp. 80–91, Jun. 2022, doi: 10.56486/KOMPLEKSITAS.VOL11NO1.220.
- [6] Y. Guo, “Credit risk assessment of P2P lending platform towards big data based on BP neural network,” *J Vis Commun Image Represent*, vol. 71, p. 102730, Aug. 2020, doi: 10.1016/J.JVCIR.2019.102730.
- [7] G. R. Murodovich and N. Jahongir, “Improving the Loaning Process in Commercial Banks,” *CENTRAL ASIAN JOURNAL OF ARTS AND DESIGN*, vol. 3, no. 5, pp. 10–15, 2022, doi: 10.17605/OSF.IO/2JFPM.
- [8] V. Moscato, A. Picariello, and G. Sperlí, “A benchmark of machine learning approaches for credit score prediction,” *Expert Syst Appl*, vol. 165, p. 113986, Mar. 2021, doi: 10.1016/J.ESWA.2020.113986.
- [9] L. Rodriguez, “All Data Is Not Credit Data: Closing the Gap Between the Fair Housing Act and Algorithmic Decisionmaking in the Lending Industry,” *Columbia Law Review*, vol. 120, no. 7, pp. 1843–1884, Nov. 2020, Accessed: May 30, 2024. [Online]. Available: <https://www.jstor.org/stable/26958733>

- [10] A. B. M. F. Shahriar, M. Z. Moon, H. Mahmud, and K. Hasan, "Online product recommendation system by using eye gaze data," in *Proceedings of the International Conference on Computing Advancements*, Dhaka: Association for Computing Machinery, Jan. 2020, pp. 1–7. doi: 10.1145/3377049.3377108.
- [11] U. Javed, K. Shaukat, I. A. Hameed, F. Iqbal, T. M. Alam, and S. Luo, "A Review of Content-Based and Context-Based Recommendation Systems," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 16, no. 3, pp. 274–306, Feb. 2021.
- [12] J. M. Amigo, "Data Mining, Machine Learning, Deep Learning, Chemometrics Definitions, Common Points and Trends (Spoiler Alert: VALIDATE your models!)," *Brazilian Journal of Analytical Chemistry*, vol. 8, no. 32, pp. 45–61, 2021, doi: 10.30744/brjac.2179-3425.AR-38-2021.
- [13] D. Papakyriakou and I. S. Barbounakis, "Data Mining Methods: A Review," *Int J Comput Appl*, vol. 183, no. 48, pp. 5–19, Jan. 2022, doi: 10.5120/ijca2022921884.
- [14] L. Yang and A. Shami, "On hyperparameter optimization of machine learning algorithms: Theory and practice," *Neurocomputing*, vol. 415, pp. 295–316, Nov. 2020, doi: 10.1016/J.NEUCOM.2020.07.061.
- [15] L. Hertel, J. Collado, P. Sadowski, J. Ott, and P. Baldi, "Sherpa: Robust hyperparameter optimization for machine learning," *SoftwareX*, vol. 12, p. 100591, Jul. 2020, doi: 10.1016/J.SOFTX.2020.100591.
- [16] S. K. Trivedi, "A study on credit scoring modeling with different feature selection and machine learning approaches," *Technol Soc*, vol. 63, p. 101413, Nov. 2020, doi: 10.1016/J.TECHSOC.2020.101413.
- [17] J. Xu, Z. Lu, and Y. Xie, "Loan default prediction of Chinese P2P market: a machine learning methodology," *Sci Rep*, vol. 11, no. 18759, Sep. 2021, doi: 10.1038/s41598-021-98361-6.
- [18] Y. Li and W. Chen, "A Comparative Performance Assessment of Ensemble Learning for Credit Scoring," *Mathematics*, vol. 8, no. 10, p. 1756, Oct. 2020, doi: 10.3390/MATH8101756.
- [19] S. G. K. Patro *et al.*, "A Hybrid Action-Related K-Nearest Neighbour (HAR-KNN) Approach for Recommendation Systems," *IEEE Access*, vol. 8, pp. 90978–90991, May 2020, doi: 10.1109/ACCESS.2020.2994056.
- [20] K. K. Jena *et al.*, "E-Learning Course Recommender System Using Collaborative Filtering Models," *Electronics (Basel)*, vol. 12, no. 1, p. 157, Dec. 2023, doi: 10.3390/ELECTRONICS12010157.

- [21] M. Y. Shams, A. M. Elshewey, E. S. M. El-kenawy, A. Ibrahim, F. M. Talaat, and Z. Tarek, "Water quality prediction using machine learning models based on grid search method," *Multimed Tools Appl*, Sep. 2023, doi: 10.1007/S11042-023-16737-4/FIGURES/15.
- [22] T. Wang *et al.*, "Random Forest-Bayesian Optimization for Product Quality Prediction with Large-Scale Dimensions in Process Industrial Cyber-Physical Systems," *IEEE Internet Things J*, vol. 7, no. 9, pp. 8641–8653, Sep. 2020, doi: 10.1109/JIOT.2020.2992811.
- [23] S. Lee *et al.*, "Estimation of rainfall erosivity factor in Italy and Switzerland using Bayesian optimization based machine learning models," *Catena (Amst)*, vol. 211, p. 105957, Apr. 2022, doi: 10.1016/J.CATENA.2021.105957.
- [24] G. Rong *et al.*, "Rainfall Induced Landslide Susceptibility Mapping Based on Bayesian Optimized Random Forest and Gradient Boosting Decision Tree Models—A Case Study of Shuicheng County, China," *Water (Basel)*, vol. 12, no. 11, p. 3066, Nov. 2020, doi: 10.3390/W12113066.
- [25] S. Jaggia, A. Kelly, K. Lertwachara, and L. Chen, "Applying the CRISP-DM Framework for Teaching Business Analytics," *Decision Sciences Journal of Innovative Education*, vol. 18, no. 4, pp. 612–634, Oct. 2020, doi: 10.1111/DSJI.12222.
- [26] W. S. Fana, R. Soviab, R. Permana, and M. A. Islam, "Data Warehouse Design with ETL Method (Extract, Transform, And Load) for Company Information Centre," *International Journal of Artificial Intelligence Research*, vol. 5, no. 2, pp. 132–137, Dec. 2021, doi: 10.29099/ijair.v5i2.215.
- [27] R. Moro-Visconti, S. C. Rambaud, and J. L. Pascual, "Sustainability in FinTechs: An Explanation through Business Model Scalability and Market Valuation," *Sustainability*, vol. 12, no. 24, p. 10316, Dec. 2020, doi: 10.3390/SU122410316.
- [28] P. Senyo and E. L. C. Osabutey, "Unearthing antecedents to financial inclusion through FinTech innovations," *Technovation*, vol. 98, p. 102155, Dec. 2020, doi: 10.1016/J.TECHNOVATION.2020.102155.
- [29] M. Sebira Hermawan, A. Syamil, and P. Heriyati, *Perspektif Industri Financial Technology di Indonesia*. Qiara Media, 2020.
- [30] H. Wang, G. Kou, and Y. Peng, "Multi-class misclassification cost matrix for credit ratings in peer-to-peer lending," *Journal of the Operational Research Society*, vol. 72, no. 4, pp. 923–934, 2021, doi: 10.1080/01605682.2019.1705193.

- [31] Z. Liu, J. Shang, S. Wu, and P. Chen, "Social collateral, soft information and online peer-to-peer lending: A theoretical model," *Eur J Oper Res*, vol. 281, no. 2, pp. 428–438, Mar. 2020, doi: 10.1016/J.EJOR.2019.08.038.
- [32] J. P. Barddal, L. Loezer, F. Enembreck, and R. Lanzuolo, "Lessons learned from data stream classification applied to credit scoring," *Expert Syst Appl*, vol. 162, p. 113899, Dec. 2020, doi: 10.1016/J.ESWA.2020.113899.
- [33] Y. Xia, J. Zhao, L. He, Y. Li, and M. Niu, "A novel tree-based dynamic heterogeneous ensemble method for credit scoring," *Expert Syst Appl*, vol. 159, p. 113615, Nov. 2020, doi: 10.1016/J.ESWA.2020.113615.
- [34] P. Pławiak, M. Abdar, J. Pławiak, V. Makarenkov, and U. R. Acharya, "DGHNL: A new deep genetic hierarchical network of learners for prediction of credit scoring," *Inf Sci (N Y)*, vol. 516, pp. 401–418, Apr. 2020, doi: 10.1016/J.INS.2019.12.045.
- [35] A. Kumar, S. Sharma, and M. Mahdavi, "Machine Learning (ML) Technologies for Digital Credit Scoring in Rural Finance: A Literature Review," *Risks*, vol. 9, no. 11, p. 192, Oct. 2021, doi: 10.3390/RISKS9110192.
- [36] S. S. Khanal, P. W. C. Prasad, A. Alsadoon, and A. Maag, "A systematic review: machine learning based recommendation systems for e-learning," *Educ Inf Technol (Dordr)*, vol. 25, pp. 2635–2664, 2020, doi: 10.1007/S10639-019-10063-9/METRICS.
- [37] N. Hug, "Surprise: A Python library for recommender systems," *J Open Source Softw*, vol. 5, no. 52, p. 2174, 2020, doi: 10.21105/joss.02174.
- [38] A. Beheshti, S. Yakhchi, S. Mousaeirad, S. M. Ghafari, S. R. Goluguri, and M. A. Edrisi, "Towards Cognitive Recommender Systems," *Algorithms*, vol. 13, no. 8, p. 176, Jul. 2020, doi: 10.3390/A13080176.
- [39] S. H. Nallamala, U. R. Bajjuri, S. Anandarao, Dr. D. DurgaPrasad, and Dr. P. Mishra, "A Brief Analysis of Collaborative and Content Based Filtering Algorithms used in Recommender Systems," *IOP Conf Ser Mater Sci Eng*, vol. 981, no. 2, p. 022008, Dec. 2020.
- [40] R. Galici, L. Ordile, M. Marchesi, A. Pinna, and R. Tonelli, "Applying the ETL Process to Blockchain Data. Prospect and Findings," *Information*, vol. 11, no. 4, p. 204, Apr. 2020, doi: 10.3390/INFO11040204.
- [41] X. Wang and J. Li, "Design of Data Quality Control System Based on ETL," *J Phys Conf Ser*, vol. 2476, p. 012083, 2023, doi: 10.1088/1742-6596/2476/1/012083.
- [42] R. Kimball and M. Ross, *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*, 3rd ed. John Wiley & Sons, 2013.

- [43] G. M. Vaidya and Dr. M. M. Kshirsagar, “ETL - A Solution to Challenging Issues in Big Data Analytics,” *Ann Rom Soc Cell Biol*, vol. 25, no. 2, pp. 947–952, 2021.
- [44] I. Priyadarshini and C. Cotton, “A novel LSTM–CNN–grid search-based deep neural network for sentiment analysis,” *Journal of Supercomputing*, vol. 77, pp. 13911–13932, Dec. 2021, doi: 10.1007/S11227-021-03838-W/TABLES/5.
- [45] P. Mikkola, M. Todorović, J. Järvi, P. Rinke, and S. Kaski, “Projective Preferential Bayesian Optimization,” in *Proceedings of the 37th International Conference on Machine Learning*, PMLR, 2020, pp. 6884–6892.
- [46] A. H. Victoria and G. Maragatham, “Automatic tuning of hyperparameters using Bayesian optimization,” *Evolving Systems*, vol. 12, pp. 217–223, 2021, doi: 10.1007/S12530-020-09345-2/METRICS.
- [47] R. Moriconi, M. P. Deisenroth, and K. S. S. Kumar, “High-dimensional Bayesian optimization using low-dimensional feature spaces,” *Mach Learn*, vol. 109, pp. 1925–1943, Sep. 2020, doi: 10.1007/S10994-020-05899-Z.
- [48] P. I. Frazier, “Bayesian Optimization,” in *Recent Advances in Optimization and Modeling of Contemporary Problems*, INFORMS, 2018, pp. 255–278. doi: 10.1287/EDUC.2018.0188.
- [49] X. Dastile, T. Celik, and M. Potsane, “Statistical and machine learning models in credit scoring: A systematic literature survey,” *Appl Soft Comput*, vol. 91, p. 106263, Jun. 2020, doi: 10.1016/J.ASOC.2020.106263.
- [50] Z. Shahbazi, D. Hazra, S. Park, and Y. C. Byun, “Toward Improving the Prediction Accuracy of Product Recommendation System Using Extreme Gradient Boosting and Encoding Approaches,” *Symmetry (Basel)*, vol. 12, no. 9, p. 1566, Sep. 2020, doi: 10.3390/SYM12091566.
- [51] M. Tsagris and N. Pandis, “Normality test: Is it really necessary?,” *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 159, no. 4, pp. 548–549, Apr. 2021, doi: 10.1016/J.AJODO.2021.01.003.
- [52] N. S. Turhan, “Karl Pearson’s Chi-Square Tests,” *Educational Research and Reviews*, vol. 16, no. 9, pp. 575–580, Sep. 2020, doi: 10.5897/ERR2019.3817.
- [53] W. Q. Meeker, L. A. Escobar, and F. G. Pascual, *Statistical Methods for Reliability Data*, 2nd ed. John Wiley & Sons, 2022.
- [54] J. Liu and Y. Xu, “T-Friedman Test: A New Statistical Test for Multiple Comparison with an Adjustable Conservativeness Measure,” *International*

Journal of Computational Intelligence Systems, vol. 15, no. 1, p. 29, Apr. 2022, doi: 10.1007/S44196-022-00083-8/FIGURES/13.

- [55] N. Chukhrova and A. Johannssen, “Fuzzy hypothesis testing: Systematic review and bibliography,” *Appl Soft Comput*, vol. 106, p. 107331, Mar. 2021, doi: 10.1016/J.ASOC.2021.107331.
- [56] B. Gerald and T. F. Patson, “Parametric and Nonparametric Tests: A Brief Review,” *International Journal of Statistical Distributions and Applications*, vol. 7, no. 3, pp. 78–82, Oct. 2021, doi: 10.11648/J.IJSD.20210703.12.
- [57] S. Yang and G. Berdine, “Normality tests,” *The Southwest Respiratory and Critical Care Chronicles*, vol. 9, no. 37, pp. 87–90, 2021, doi: 10.12746/swrccc.v9i37.805.
- [58] S. S. Shapiro and M. B. Wilk, “An analysis of variance test for normality (complete samples),” *Biometrika*, vol. 52, no. 3–4, pp. 591–611, Dec. 1965, doi: 10.1093/BIOMET/52.3-4.591.
- [59] A. Adhikari, S. C. Gorain, P. Gayen, I. Pal, and Dr. S. Sen, “Studying the Differences: A Review on t-Test,” *International Research Journal of Education and Technology*, vol. 5, no. 5, pp. 338–349, 2023.
- [60] C. S. de Brito, D. E. W. Silva, L. G. de A. Aguiar, R. Abrãao, and N. A. de S. Sampaio, “Using the paired T-Test to compare suppliers,” *Revista de Gestão e Secretariado*, vol. 14, no. 10, pp. 18565–18575, 2023, doi: 10.7769/GESEC.V14I10.3068.
- [61] H. M. Park, “Comparing Group Means: T-tests and One-way ANOVA Using Stata, SAS, R, and SPSS,” 2009. Accessed: Apr. 27, 2024. [Online]. Available: <http://www.indiana.edu/~statmath>
- [62] C. Schröer, F. Kruse, and J. M. Gómez, “A Systematic Literature Review on Applying CRISP-DM Process Model,” *Procedia Comput Sci*, vol. 181, pp. 526–534, 2021, doi: 10.1016/J.PROCS.2021.01.199.
- [63] O. Firas, “A combination of SEMMA & CRISP-DM models for effectively handling big data using formal concept analysis based knowledge discovery: A data mining approach,” *World Journal of Advanced Engineering Technology and Sciences*, vol. 8, no. 1, pp. 9–14, 2023, doi: 10.30574/WJAETS.2023.8.1.0147.
- [64] F. Martínez-Plumed *et al.*, “CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories,” *IEEE Trans Knowl Data Eng*, vol. 33, no. 8, pp. 3048–3061, Aug. 2021, doi: 10.1109/TKDE.2019.2962680.
- [65] Z. Tang, Z. Mei, W. Liu, and Y. Xia, “Identification of the key factors affecting Chinese carbon intensity and their historical trends using random

- forest algorithm,” *Journal of Geographical Sciences*, vol. 30, no. 5, pp. 743–756, 2020, doi: 10.1007/S11442-020-1753-4/METRICS.
- [66] A. Segnini and J. J. T. Motchoffo, “Random Forests and Text Mining,” 2021.
- [67] H. Wu, A. Lin, X. Xing, D. Song, and Y. Li, “Identifying core driving factors of urban land use change from global land cover products and POI data using the random forest method,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 103, p. 102475, Dec. 2021, doi: 10.1016/J.JAG.2021.102475.
- [68] L. Breiman, “Random forests,” *Mach Learn*, vol. 45, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.
- [69] C. Aroef, Y. Rivan, and Z. Rustam, “Comparing random forest and support vector machines for breast cancer classification,” *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 18, no. 2, pp. 815–821, Apr. 2020, doi: 10.12928/TELKOMNIKA.V18I2.14785.
- [70] G. E. Atteia, H. Mengash, and N. Abdhelsamee, “Evaluation of using Parametric and Non-parametric Machine Learning Algorithms for Covid-19 Forecasting,” *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 10, pp. 647–657, Jan. 2021, doi: 10.14569/IJACSA.2021.0121071.
- [71] I. Dănilă-Cernat, “Nearest neighbor estimates of Kaniadakis entropy,” *Analele științifice ale Universității “Ovidius” Constanța. Seria Matematică*, vol. 30, no. 1, pp. 171–189, Feb. 2022, doi: 10.2478/auom-2022-0010.
- [72] E. Indrayuni, A. Nurhadi, and D. A. Kristiyanti, “Implementasi Algoritma Naive Bayes, Support Vector Machine, dan K-Nearest Neighbors untuk Analisa Sentimen Aplikasi Halodoc,” *Faktor Exacta*, vol. 14, no. 2, pp. 64–71, Jun. 2021, doi: 10.30998/FAKTOREXACTA.V14I2.9697.
- [73] S. Zhang, J. Li, and Y. Li, “Reachable Distance Function for KNN Classification,” *IEEE Trans Knowl Data Eng*, vol. 35, no. 7, pp. 7382–7397, 2022, doi: 10.1109/TKDE.2022.3185149.
- [74] D. Jiang, J. Zhang, Z. Wang, C. Feng, K. Jiao, and R. Xu, “A Prediction Model of Blast Furnace Slag Viscosity Based on Principal Component Analysis and K-Nearest Neighbor Regression,” *JOM*, vol. 72, pp. 3908–3916, 2020, doi: 10.1007/S11837-020-04360-9/METRICS.
- [75] The PostgreSQL Global Development Group, “PostgreSQL: About,” The PostgreSQL Global Development Group. Accessed: Jan. 16, 2024. [Online]. Available: <https://www.postgresql.org/about/>

- [76] L. Ferrari and E. Pirozzi, *Learn PostgreSQL: Build and manage high-performance database solutions using PostgreSQL 12 and 13*. Packt Publishing Ltd, 2020.
- [77] The PostgreSQL Global Development Group, “PostgreSQL: Documentation: 16: 43.1. Overview,” The PostgreSQL Global Development Group. Accessed: Apr. 27, 2024. [Online]. Available: <https://www.postgresql.org/docs/current/plpgsql-overview.html#PLPGSQL-ADVANTAGES>
- [78] DBeaver, “About | DBeaver Community,” DBeaver. Accessed: Jan. 16, 2024. [Online]. Available: <https://dbeaver.io/about/>
- [79] V. Anbumani, V. Geetha, V. P. Kumar, D. Sabaree, and K. Sivanantham, “Development of Cloud-Based Agriculture Marketing System with Intellectual Weigh Machine,” *IOP Conf Ser Mater Sci Eng*, vol. 1055, no. 1, p. 12016, Dec. 2021, doi: 10.1088/1757-899X/1055/1/012016.
- [80] DBeaver, “DBeaver Documentation,” DBeaver. Accessed: Jan. 16, 2024. [Online]. Available: <https://dbeaver.com/docs/dbeaver/>
- [81] Python Software Foundation, “The Python Tutorial — Python 3.12.1 documentation,” Python Software Foundation. Accessed: Jan. 16, 2024. [Online]. Available: <https://docs.python.org/3/tutorial/index.html>
- [82] Prof. T. Arivanantham, K. Swami, K. Swami, S. Ugalmugale, and D. Giri, “VIRTUAL MOUSE AND KEYBOARD BASED ON HUMAN GESTURES USING MACHINE LEARNING,” *International Research Journal of Modernization in Engineering Technology and Science*, vol. 5, no. 11, Nov. 2023, doi: 10.56726/IRJMETS4624.
- [83] A. Sharma, F. Khan, D. Sharma, and S. Gupta, “Python: The Programming Language of Future,” *International Journal of Innovative Research in Technology*, vol. 6, no. 12, pp. 115–118, May 2020.
- [84] S. Raj and M. Paliwal, “Why Python is Most Famous,” *International Journal of Innovative Research in Computer Science & Technology*, vol. 9, no. 6, pp. 234–238, Nov. 2021, doi: 10.55524/ijircst.2021.9.6.52.
- [85] J. Wang, L. Li, and A. Zeller, “Better Code, Better Sharing: On the Need of Analyzing Jupyter Notebooks,” in *2020 IEEE/ACM 42nd International Conference on Software Engineering: New Ideas and Emerging Results (ICSE-NIER)*, Seoul: Institute of Electrical and Electronics Engineers Inc., Oct. 2020, pp. 53–56. doi: 10.1145/3377816.3381724.
- [86] J. Wang, L. Li, and A. Zeller, “Restoring Execution Environments of Jupyter Notebooks,” in *2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE)*, Madrid: Institute of Electrical and Electronics

- Engineers Inc., May 2021, pp. 1622–1633. doi: 10.1109/ICSE43902.2021.00144.
- [87] The R Foundation, “R: What is R?,” The R Foundation. Accessed: Feb. 06, 2024. [Online]. Available: <https://www.r-project.org/about.html>
- [88] Posit, “Download RStudio | The Popular Open-Source IDE from Posit,” Posit Software, PBC (formerly RStudio, PBC). Accessed: Feb. 06, 2024. [Online]. Available: <https://posit.co/products/open-source/rstudio/>
- [89] F. Kronthaler and S. Zöllner, *Data Analysis with R Studio: An Easygoing Introduction*, 1st ed. Berlin: Springer Spektrum, 2021. doi: 10.1007/978-3-662-62518-7/COVER.
- [90] Rukminingsih, G. Adnan, and M. A. Latief, *Metode Penelitian Pendidikan: Penelitian Kuantitatif, Penelitian Kualitatif, Penelitian Tindakan Kelas*. Sleman: Erhaka Utama, 2020.
- [91] J. Sreemathy, I. J. V., S. Nisha, C. P. I., and G. P. R.M., “Data Integration in ETL Using TALEND,” in *6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore: Institute of Electrical and Electronics Engineers Inc., 2020, pp. 1444–1448. doi: 10.1109/ICACCS48705.2020.9074186.
- [92] D. A. Kristiyanti, S. A. Sanjaya, V. C. Tjokro, and J. Suhali, “Dealing imbalance dataset problem in sentiment analysis of recession in Indonesia,” *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 13, no. 2, pp. 2060–2072, Jun. 2024, doi: 10.11591/IJAI.V13.I2.PP2060-2072.
- [93] R. Wazirali, “An Improved Intrusion Detection System Based on KNN Hyperparameter Tuning and Cross-Validation,” *Arab J Sci Eng*, vol. 45, pp. 10859–10873, Dec. 2020, doi: 10.1007/S13369-020-04907-7/METRICS.
- [94] O. A. M. López, A. M. López, and J. Crossa, “Overfitting, Model Tuning, and Evaluation of Prediction Performance,” in *Multivariate Statistical Machine Learning Methods for Genomic Prediction*, Springer, Cham, 2022, pp. 109–139. doi: 10.1007/978-3-030-89010-0_4.
- [95] D. F. Oliveira and M. A. Brito, “Position Paper: Quality Assurance in Deep Learning Systems,” in *Proceedings of the 11th International Conference on Data Science, Technology and Applications (DATA 2022)*, Setúbal: SciTePress, 2022, pp. 203–210. doi: 10.5220/0011107100003269.
- [96] S. Devi, P. Maurya, M. Saran, R. K. Yadav, and U. N. Tripathi, “Using Different Methodologies of Data Science to Find Comparison Between Them for Cyber-Security,” *International Journal of Innovative Research in Engineering & Management*, vol. 9, no. 5, pp. 53–58, Oct. 2022, doi: 10.55524/ijirem.2022.9.5.7.

- [97] A. Rotondo and F. Quilligan, "Evolution Paths for Knowledge Discovery and Data Mining Process Models," *SN Comput Sci*, vol. 1, no. 109, Mar. 2020, doi: 10.1007/S42979-020-0117-6/METRICS.
- [98] P. Butka, P. Bednár, and J. Ivančáková, "Methodologies for Knowledge Discovery Processes in Context of AstroGeoInformatics," in *Knowledge Discovery in Big Data from Astronomy and Earth Observation*, P. Škoda and F. Adam, Eds., Elsevier, 2020, pp. 1–20. doi: 10.1016/B978-0-12-819154-5.00010-2.
- [99] V. Plotnikova, M. Dumas, and F. Milani, "Adaptations of data mining methodologies: A systematic literature review," *PeerJ Comput Sci*, vol. 6, p. e267, May 2020, doi: 10.7717/PEERJ-CS.267.
- [100] P. Pandey and M. M. Pandey, *Research methodology tools and techniques*. Buzau: Bridge Center, 2021.
- [101] A. E. Berndt, "Sampling Methods," *Journal of Human Lactation*, vol. 36, no. 2, pp. 224–226, Mar. 2020, doi: 10.1177/0890334420906850.
- [102] E. I. OBILOR, "Convenience and Purposive Sampling Techniques: Are they the Same?," *International Journal of Innovative Social & Science Education Research*, vol. 11, no. 1, pp. 1–7, 2023, Accessed: Mar. 13, 2024. [Online]. Available: www.seahipaj.org
- [103] R. Ulfa, "Variabel Penelitian Dalam Penelitian Pendidikan," *Jurnal AL-Fathonah*, vol. 1, no. 1, pp. 342–351, Apr. 2021.

