

DAFTAR PUSTAKA

- [1] G. Florencia, “5 Efek Samping Diabetes yang Tidak Terduga,” *Halodoc.com*, 2020.
- [2] P. Wisanggeni, Satrio Pangarso Krisna, Albertus Rosalina, “Diabetes Makin Membebani Biaya Jaminan Kesehatan,” *KOMPAS*, 2023. <https://www.kompas.id/baca/investigasi/2023/04/13/diabetes-makin-membebani-biaya-jaminan-kesehatan>.
- [3] D. Sisodia and D. S. Sisodia, “Prediction of Diabetes using Classification Algorithms,” *Procedia Comput. Sci.*, vol. 132, no. Iccids, pp. 1578–1585, 2018, doi: 10.1016/j.procs.2018.05.122.
- [4] N. P. Tigga and S. Garg, “Prediction of Type 2 Diabetes using Machine Learning Classification Methods,” *Procedia Comput. Sci.*, vol. 167, no. 2019, pp. 706–716, 2020, doi: 10.1016/j.procs.2020.03.336.
- [5] A. Mujumdar and V. Vaidehi, “Diabetes Prediction using Machine Learning Algorithms,” *Procedia Comput. Sci.*, vol. 165, pp. 292–299, 2019, doi: 10.1016/j.procs.2020.01.047.
- [6] M. A. Sarwar, N. Kamal, W. Hamid, and M. A. Shah, “Prediction of diabetes using machine learning algorithms in healthcare,” *ICAC 2018 - 2018 24th IEEE Int. Conf. Autom. Comput. Improv. Product. through Autom. Comput.*, no. September, pp. 1–6, 2018, doi: `.
- [7] A. D. Association, “Diagnosis and classification of diabetes mellitus,” *Diabetes Care*, vol. 37, no. SUPPL.1, pp. 81–90, 2014, doi: 10.2337/dc14-S081.
- [8] K. Papatheodorou, M. Banach, E. Bekiari, M. Rizzo, and M. Edmonds, “Complications of Diabetes 2017,” *J. Diabetes Res.*, vol. 2018, pp. 10–13, 2018, doi: 10.1155/2018/3086167.
- [9] R. Pahlevi, “Jumlah Penderita Diabetes Indonesia Terbesar Kelima di Dunia,” *Katadata.co.id*, 2021.
- [10] S. H. Wahid, *A Tool Kit for Penulisan Karya Ilmiah (Skripsi dan Jurnal) Plus Analisis Data*. 2021.
- [11] B. Shamreen Ahamed, M. S. Arya, and A. O. Nancy, “Diabetes Mellitus Disease Prediction Using Machine Learning Classifiers and Techniques

- Using the Concept of Data Augmentation and Sampling,” in *Lecture Notes in Networks and Systems*, 2023, vol. 516, pp. 401–413, doi: 10.1007/978-981-19-5221-0_40.
- [12] S. S. Bhat, V. Selvam, G. A. Ansari, M. D. Ansari, and M. H. Rahman, “Prevalence and Early Prediction of Diabetes Using Machine Learning in North Kashmir: A Case Study of District Bandipora,” *Comput. Intell. Neurosci.*, vol. 2022, 2022, doi: 10.1155/2022/2789760.
- [13] R. Rastogi and M. Bansal, “Diabetes prediction model using data mining techniques,” *Meas. Sensors*, vol. 25, no. October 2022, p. 100605, 2023, doi: 10.1016/j.measen.2022.100605.
- [14] A. Hermawan and H. L. Yusran, *Penelitian Bisnis Pendekatan Kuantitatif*. Depok: Prenamedia Group, 2017.
- [15] S. Guritno, Sudaryono, and U. Rahardja, *Theory And Application Of IT Research : Metodologi Penelitian Teknologi Informasi*. Yogyakarta: Andi, 2011.
- [16] R. Munir, *Algoritma & Pemrograman dalam Bahasa pascal dan C*. Bandung: Informatika, 2011.
- [17] T. Cormen, *Introduction to Algorithms*. The MIT Press, 1989.
- [18] Q. Chang and J. Hu, “Research and Application of the Data Mining Technology in Economic Intelligence System,” *Comput. Intell. Neurosci.*, vol. 2022, 2022, doi: 10.1155/2022/6439315.
- [19] Yuli Mardi, “Data Mining : Klasifikasi Menggunakan Algoritma C4.5,” *J. Edik Inform.*, vol. 2, no. 2, pp. 213–219, 2019.
- [20] Mustika *et al.*, *Data Mining dan Aplikasinya*. 2021.
- [21] F. Marisa, “Educational Data Mining (Konsep dan Penerapan),” *J. Teknol. Inf.*, vol. 4, no. 2, pp. 91–93, 2013.
- [22] B. K. Pramono, *PENGANTAR RISET DIPLOMASI DIGITAL*. CV Literasi Nusantara Abadi, 2023.
- [23] R. B. Widodo, *Machine Learning Metode k- Nearest NeightBors Klasifikasi Angka Bahasa Isyarat*. 2022.
- [24] I. Izonin, R. Tkachenko, N. Shakhovska, B. Ilchyshyn, and K. K. Singh, “A Two-Step Data Normalization Approach for Improving Classification

- Accuracy in the Medical Diagnosis Domain,” *Mathematics*, vol. 10, no. 11, pp. 1–18, 2022, doi: 10.3390/math10111942.
- [25] I. Düntsch and G. Gediga, “Confusion Matrices and Rough Set Data Analysis,” in *Journal of Physics: Conference Series*, 2019, vol. 1229, no. 1, doi: 10.1088/1742-6596/1229/1/012055.
- [26] M. Grandini, E. Bagli, and G. Visani, “Metrics for Multi-Class Classification: an Overview,” pp. 1–17, 2020, [Online]. Available: <http://arxiv.org/abs/2008.05756>.
- [27] I. M. De Diego, A. R. Redondo, R. R. Fernández, J. Navarro, and J. M. Moguerza, “General Performance Score for classification problems,” *Appl. Intell.*, vol. 52, no. 10, pp. 12049–12063, 2022, doi: 10.1007/s10489-021-03041-7.
- [28] G. M. M. Aurup, “User Preference Extraction from Bio-Signals: An Experimental Study,” no. January, pp. 1–134, 2011.
- [29] I. Made, B. Adnyana, S. B. Jln, and R. Puputan, “Penerapan Feature Selection untuk Prediksi Lama Studi Mahasiswa,” *STIKOM Bali*, vol. 13, pp. 72–76, 2019.
- [30] Y. Bouchlaghem, Y. Akhiat, and S. Amjad, “Feature Selection: A Review and Comparative Study,” in *E3S Web of Conferences*, 2022, vol. 351, pp. 1–6, doi: 10.1051/e3sconf/202235101046.
- [31] A. Rahmansyah, O. Dewi, P. Andini, T. Hastuti, P. Ningrum, and M. E. Suryana, “Membandingkan Pengaruh Feature Selection Terhadap Algoritma Naïve Bayes dan Support Vector Machine,” *Semin. Nas. Apl. Teknol. Inf.*, pp. 1907–5022, 2018.
- [32] S. Yadav and S. Shukla, “Analysis of k-Fold Cross-Validation over Hold-Out Validation on Colossal Datasets for Quality Classification,” *Proc. - 6th Int. Adv. Comput. Conf. IACC 2016*, no. Cv, pp. 78–83, 2016, doi: 10.1109/IACC.2016.25.
- [33] T. T. Wong, “Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation,” *Pattern Recognit.*, vol. 48, no. 9, pp. 2839–2846, 2015, doi: 10.1016/j.patcog.2015.03.009.
- [34] R. Kohavi and S. Edu, “A study of cross-validation and bootstrap for

- accuracy estimation and model selection,” *Proc. 14th Int. Jt. Conf. Artif. Intell.*, vol. 2, pp. 1137–1143, 1993.
- [35] S. Raschka, “Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning,” 2018, [Online]. Available: <http://arxiv.org/abs/1811.12808>.
- [36] H. Henderi, “Comparison of Min-Max normalization and Z-Score Normalization in the K-nearest neighbor (kNN) Algorithm to Test the Accuracy of Types of Breast Cancer,” *IJIS Int. J. Informatics Inf. Syst.*, vol. 4, no. 1, pp. 13–20, 2021, doi: 10.47738/ijis.v4i1.73.
- [37] R. G. Whendasmoro and J. Joseph, “Analisis Penerapan Normalisasi Data Dengan Menggunakan Z-Score Pada Kinerja Algoritma K-NN,” *JURIKOM (Jurnal Ris. Komputer)*, vol. 9, no. 4, p. 872, 2022, doi: 10.30865/jurikom.v9i4.4526.
- [38] V. Aggarwal, V. Gupta, P. Singh, K. Sharma, and N. Sharma, “Detection of spatial outlier by using improved Z-score test,” *Proc. Int. Conf. Trends Electron. Informatics, ICOEI 2019*, vol. 2019-April, no. Icoei, pp. 788–790, 2019, doi: 10.1109/icoei.2019.8862582.
- [39] H. A. Prihanditya, “The Implementation of Z-Score Normalization and Boosting Techniques to Increase Accuracy of C4.5 Algorithm in Diagnosing Chronic Kidney Disease,” *J. Soft Comput. Explor.*, vol. 1, no. 1, pp. 63–69, 2020, doi: 10.52465/josce.v1i1.8.
- [40] S. Jameel Mukhyber and A. D. Majeed, “Effect Z-score Normalization on Accuracy of classification of liver disease,” *Turkish J. Comput. Math. Educ.*, vol. 12, no. 14, pp. 658–662, 2021.
- [41] G. Aksu, C. O. Güzeller, and M. T. Eser, “The Effect of the Normalization Method Used in Different Sample Sizes on the Success of Artificial Neural Network Model,” *Int. J. Assess. Tools Educ.*, vol. 6, no. 2, pp. 170–192, 2019, doi: 10.21449/ijate.479404.
- [42] J. J. Pangaribuan, “MENDIAGNOSIS PENYAKIT DIABETES MELITUS DENGAN MENGGUNAKAN METODE EXTREME LEARNING MACHINE,” *J. ISD*, vol. 2, no. 2, 2016, doi: 10.34288/jri.v2i2.121.
- [43] K. Bell *et al.*, “A position statement on screening and management of

- prediabetes in adults in primary care in Australia,” *Diabetes Res. Clin. Pract.*, vol. 164, p. 108188, 2020, doi: 10.1016/j.diabres.2020.108188.
- [44] R. K. Rezki and L. Septina, “Hubungan Prediabetes Terhadap Neuropati Diabetikum Di Klinik Aisyiyah Muhammadiyah Medan,” *Tjyybjb.Ac.Cn*, vol. 3, no. 2, pp. 58–66, 2019, [Online]. Available: <http://www.tjyybjb.ac.cn/CN/article/downloadArticleFile.do?attachType=PDF&id=9987>.
- [45] I. Budiastutik, M. I. Kartasurya, H. W. Subagio, and B. Widjanarko, “High Prevalence of Prediabetes and Associated Risk Factors in Urban Areas of Pontianak, Indonesia: A Cross-Sectional Study,” *J. Obes.*, vol. 2022, 2022, doi: 10.1155/2022/4851044.
- [46] D. Care and S. S. Suppl, “Classification and diagnosis of diabetes: Standards of medical care in diabetes-2021,” *Diabetes Care*, vol. 44, no. January, pp. S15–S33, 2021, doi: 10.2337/dc21-S002.
- [47] Lestari, Zulkarnain, and S. A. Sijid, “Diabetes Melitus: Review Etiologi, Patofisiologi, Gejala, Penyebab, Cara Pemeriksaan, Cara Pengobatan dan Cara Pencegahan,” *UIN Alauddin Makassar*, no. November, pp. 237–241, 2021, [Online]. Available: <http://journal.uin-alauddin.ac.id/index.php/psb>.
- [48] E. Kartinah, “Perlu Kebijakan Inovatif untuk Atasi Beban Ekonomi Diabetes,” *Media Indonesia*, Nov. 2019.
- [49] S. Barakat, M. Abujbara, R. Banimustafa, A. Batieha, and K. Ajlouni, “Sleep Quality in Patients With Type 2 Diabetes Mellitus,” *J. Clin. Med. Res.*, vol. 11, no. 4, pp. 261–266, 2019, doi: 10.14740/jocmr2947w.
- [50] S. Surani, “Effect of diabetes mellitus on sleep quality,” *World J. Diabetes*, vol. 6, no. 6, p. 868, 2015, doi: 10.4239/wjd.v6.i6.868.
- [51] L. Ismail, H. Materwala, and J. Al Kaabi, “Association of risk factors with type 2 diabetes: A systematic review,” *Comput. Struct. Biotechnol. J.*, vol. 19, pp. 1759–1785, 2021, doi: 10.1016/j.csbj.2021.03.003.
- [52] J. Maddatu, E. Anderson-Baucum, and C. Evans-Molina, “Smoking and the Risk of Type 2 Diabetes HHS Public Access,” *Physiol. Behav.*, vol. 176, no. 3, pp. 139–148, 2017, doi: 10.1016/j.trsl.2017.02.004.Smoking.
- [53] A. Przekaz, W. Bielka, and A. Pawlik, “Hypertension and Type 2 Diabetes—

- The Novel Treatment Possibilities,” *Int. J. Mol. Sci.*, vol. 23, no. 12, 2022, doi: 10.3390/ijms23126500.
- [54] I. H. De Boer *et al.*, “Diabetes and hypertension: A position statement by the American diabetes association,” *Diabetes Care*, vol. 40, no. 9, pp. 1273–1284, 2017, doi: 10.2337/dci17-0026.
- [55] M. Shi, X. Zhang, and H. Wang, “The Prevalence of Diabetes, Prediabetes and Associated Risk Factors in Hangzhou, Zhejiang Province: A Community-Based Cross-Sectional Study,” *Diabetes, Metab. Syndr. Obes.*, vol. 15, no. March, pp. 713–721, 2022, doi: 10.2147/DMSO.S351218.
- [56] P. W. Yoon, M. T. Scheuner, K. L. Peterson-Oehlke, M. Gwinn, A. Faucett, and M. J. Khoury, “Can family history be used as a tool for public health and preventive medicine?,” *Genetics in Medicine*, vol. 4, no. 4, pp. 304–310, 2002, doi: 10.1097/00125817-200207000-00009.
- [57] A. Parkkola *et al.*, “Family history of type 2 diabetes and characteristics of children with newly diagnosed type 1 diabetes,” *Diabetologia*, vol. 64, no. 3, pp. 581–590, 2021, doi: 10.1007/s00125-020-05342-x.
- [58] L. Penn *et al.*, “Importance of Weight Loss Maintenance and Risk Prediction in the Prevention of Type 2 Diabetes: Analysis of European Diabetes Prevention Study RCT,” *PLoS One*, vol. 8, no. 2, pp. 1–11, 2013, doi: 10.1371/journal.pone.0057143.
- [59] R. J. Sigal *et al.*, “Physical Activity and Diabetes,” *Can. J. Diabetes*, vol. 42, pp. S54–S63, 2018, doi: 10.1016/j.jcjd.2017.10.008.
- [60] K. Briggs Early and K. Stanley, “Position of the Academy of Nutrition and Dietetics: The Role of Medical Nutrition Therapy and Registered Dietitian Nutritionists in the Prevention and Treatment of Prediabetes and Type 2 Diabetes,” *J. Acad. Nutr. Diet.*, vol. 118, no. 2, pp. 343–353, 2018, doi: 10.1016/j.jand.2017.11.021.
- [61] A. Mir and S. N. Dhage, “Diabetes Disease Prediction Using Machine Learning on Big Data of Healthcare,” *Proc. - 2018 4th Int. Conf. Comput. Commun. Control Autom. ICCUBEA 2018*, pp. 1–6, 2018, doi: 10.1109/ICCUBEA.2018.8697439.
- [62] A. Singh, A. Dhillon, N. Kumar, M. S. Hossain, G. Muhammad, and M.

- Kumar, “eDiaPredict: An Ensemble-based Framework for Diabetes Prediction,” *ACM Trans. Multimed. Comput. Commun. Appl.*, vol. 17, no. 2s, 2021, doi: 10.1145/3415155.
- [63] R. Imron and Z. Arief, *Panduan Praktis Menulis Pengembangan Keprofesian Berkelanjutan (PKB)*. PENERBIT PT KANISIUS, 2020.
- [64] A. R. Hevner and S. Chatterjee, *Design research in information systems : theory and practice*. Springer, 2010.
- [65] S. Tripathi, D. Muhr, M. Brunner, H. Jodlbauer, M. Dehmer, and F. Emmert-Streib, “Ensuring the Robustness and Reliability of Data-Driven Knowledge Discovery Models in Production and Manufacturing,” *Front. Artif. Intell.*, vol. 4, no. June, pp. 1–20, 2021, doi: 10.3389/frai.2021.576892.
- [66] R. Wirth and J. Hipp, “CRISP-DM: towards a standard process model for data mining. Proceedings of the Fourth International Conference on the Practical Application of Knowledge Discovery and Data Mining, 29-39,” *Proc. Fourth Int. Conf. Pract. Appl. Knowl. Discov. Data Min.*, no. 24959, pp. 29–39, 2000, [Online]. Available: https://www.researchgate.net/publication/239585378_CRISP-DM_Towards_a_standard_process_model_for_data_mining.
- [67] F. Martinez-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, 2021, doi: 10.1109/TKDE.2019.2962680.
- [68] C. Schröer, F. Kruse, and J. M. Gómez, “A systematic literature review on applying CRISP-DM process model,” *Procedia Comput. Sci.*, vol. 181, no. 2019, pp. 526–534, 2021, doi: 10.1016/j.procs.2021.01.199.
- [69] A. Engelbrecht, J. Sen, and S. Mehtab, *Machine Learning Algorithms, Models and Applications*. 2021.
- [70] H. Allende and C. Valle, “Ensemble Methods for Time Series Forecasting,” *Claudio Moraga A Passion for Multi-Valued Log. Soft Comput.*, vol. 349, pp. 201–215, 2017, doi: 10.1007/978-3-319-48317-7.
- [71] Y. Zhang, J. Liu, and W. Shen, “A Review of Ensemble Learning Algorithms Used in Remote Sensing Applications,” *Appl. Sci.*, vol. 12, no. 17, 2022, doi:

10.3390/app12178654.

- [72] V. C. Osamor and A. F. Okezie, “Enhancing the weighted voting ensemble algorithm for tuberculosis predictive diagnosis,” *Sci. Rep.*, vol. 11, no. 1, pp. 1–11, 2021, doi: 10.1038/s41598-021-94347-6.
- [73] R. A. Atmoko, *Dasar Implementasi Protokol MQTT Menggunakan Python dan NodeMCU*. Mokosoft Media, 2019.
- [74] Jubilee Enterprise, *Trik Cepat Menguasai Pemrograman Python*. Elex Media Computindo, 2016.
- [75] D. Paper, *Hands-on Scikit-Learn for machine learning applications : data science fundamentals with Python*. 2019.
- [76] E. Bressert, *SciPy and NumPy: An Overview for Developers*. O’Reilly Media, Inc, 2012.
- [77] Python Software Foundation, “The Python Package Index (PyPI),” 2019. <https://pypi.org/project/pandas/>.
- [78] B. V Root, *Interactive applications using Matplotlib: don’t just see your data, experience it!* 2015.
- [79] M. R. Faisal and E. Kurniawan, *Data Science Pengenalan Azure Machine Learning Studio*. 2019, 2019.
- [80] H. Rajaguru and S. K. Prabhakar, *KNN Classifier and K-Means Clustering for Robust Classification of Epilepsy from EEG Signals. A Detailed Analysis*. Anchor Academic Publishing, 2017.
- [81] R. Ye, Z. Le, and P. N. Suganthan, “K-nearest neighbor based bagging SVM pruning,” in *Proceedings of the 2013 IEEE Symposium on Computational Intelligence and Ensemble Learning, CIEL 2013 - 2013 IEEE Symposium Series on Computational Intelligence, SSCI 2013*, 2013, no. July 2015, pp. 25–30, doi: 10.1109/CIEL.2013.6613136.
- [82] O. Maimon and R. Lior, *Data Mining and Knowledge Discovery Handbook*. Springer.
- [83] M. Sadikin, *Teknik Machine Learning Untuk Menangani Permasalahan Data Bias Pada Transaksi Aplikasi Pos*. Zahira Media Publisher, 2022.
- [84] A. Yanto, *Machine Learning*. Global Eksekutif Teknologi, 2023.
- [85] S. A. Huda, R. M. Awangga, and N. S. Fathonah, *Prediksi Penerimaan*

Pegawai Baru Dengan Metode Naive Bayes. KREATIF, 2020.

- [86] S. Tewari and U. D. Dwivedi, "A comparative study of heterogeneous ensemble methods for the identification of geological lithofacies," *J. Pet. Explor. Prod. Technol.*, vol. 10, no. 5, pp. 1849–1868, 2020, doi: 10.1007/s13202-020-00839-y.
- [87] A. A. Aburomman and M. Bin Ibne Reaz, "A novel SVM-kNN-PSO ensemble method for intrusion detection system," *Appl. Soft Comput. J.*, vol. 38, pp. 360–372, 2016, doi: 10.1016/j.asoc.2015.10.011.
- [88] Y. Ning, G. Wang, J. Yu, and H. Jiang, "A Feature Selection Algorithm Based on Variable Correlation and Time Correlation for Predicting Remaining Useful Life of Equipment Using RNN," *2018 Cond. Monit. Diagnosis, C. 2018 - Proc.*, pp. 1–6, 2018, doi: 10.1109/CMD.2018.8535843.
- [89] R. Das, G. Kasieczka, and D. Shih, "Feature Selection with Distance Correlation," pp. 1–14, 2022, [Online]. Available: <https://arxiv.org/abs/2212.00046v1>.
- [90] N. Pudjihartono, T. Fadason, A. W. Kempa-Liehr, and J. M. O'Sullivan, "A Review of Feature Selection Methods for Machine Learning-Based Disease Risk Prediction," *Front. Bioinforma.*, vol. 2, no. June, pp. 1–17, 2022, doi: 10.3389/fbinf.2022.927312.
- [91] Y. Akhiat, M. Chahhou, and A. Zinedine, "Feature Selection Based on Graph Representation," in *Colloquium in Information Science and Technology, CIST*, 2018, vol. 2018-October, pp. 232–237, doi: 10.1109/CIST.2018.8596467.
- [92] L. Al-Shalabi, "New Feature Selection Algorithm Based on Feature Stability and Correlation," *IEEE Access*, vol. 10, pp. 4699–4713, 2022, doi: 10.1109/ACCESS.2022.3140209.
- [93] I. S. Leon Andretti Abdillah, Sufyati HS, Puji Muniarty, Indra Nanda, Septina Dwi Retnandari, Wulandari Wulandari, Adirasa Hadi Prasetyo, Sarton Sinambela, Sukarman Purba, Triana Zuhrotun Aulia, Amir Hamzah, Hamdan Firmansyah, Soetji Andari, Bambang Rismadi, Ga, *Metode Penelitian dan Analisis Data Comprehensive*. Penerbit Insania, 2021.

- [94] P. V. Anusha, C. Anuradha, P. S. R. Chandra Murty, and C. S. Kiran, "Detecting outliers in high dimensional data sets using Z-score methodology," *Int. J. Innov. Technol. Explor. Eng.*, vol. 9, no. 1, pp. 48–53, 2019, doi: 10.35940/ijitee.A3910.119119.
- [95] F. I. Mowbray, S. M. Fox-Wasylyshyn, and M. M. El-Masri, "Univariate Outliers: A Conceptual Overview for the Nurse Researcher," *Can. J. Nurs. Res.*, vol. 51, no. 1, pp. 31–37, 2019, doi: 10.1177/0844562118786647.