

DAFTAR PUSTAKA

- Ahmed, N., Ahammed, R., Islam, M. M., Uddin, M. A., Akhter, A., Talukder, M. A. A., & Paul, B. K. (2021). Machine learning based diabetes prediction and development of smart web application. *International Journal of Cognitive Computing in Engineering*, 2, 229–241. <https://doi.org/10.1016/j.ijcce.2021.12.001>
- Anastasya, G., Aguslina Siregar, F., & Ashar, T. (2024). *The Effect of Obesity, Carbohydrate, and Fat Intake on the Incidence of Type 2 Diabetes Mellitus at the Age of 30-50 Years in Medan*.
- Balaram, A., & Vasundra, S. (2022). Prediction of software fault-prone classes using ensemble random forest with adaptive synthetic sampling algorithm. *Automated Software Engineering*, 29(1). <https://doi.org/10.1007/s10515-021-00311-z>
- Chandra Sekhar Reddy, L., Gottipalli, M., Sravanthi, P., Rajanikanth, J., Yalamarathi, G., & Gurrupu, N. (2024). Bridging Horizons in Diabetes Prediction: A Comparative Exploration of Machine Learning and Deep Learning Approaches in Pima Indian Women. *Proceedings - 2nd International Conference on Advancement in Computation and Computer Technologies, InCACCT 2024*, 386–391. <https://doi.org/10.1109/InCACCT61598.2024.10550977>
- Chang, V., Ganatra, M. A., Hall, K., Golightly, L., & Xu, Q. A. (2022). An assessment of machine learning models and algorithms for early prediction and diagnosis of diabetes using health indicators. *Healthcare Analytics*, 2. <https://doi.org/10.1016/j.health.2022.100118>
- Chen, H., Tan, C., Lin, Z., Chen, M., & Cheng, B. (2024). Applying virtual sample generation and ensemble modeling for improving the spectral diagnosis of cancer. *Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy*, 318. <https://doi.org/10.1016/j.saa.2024.124518>
- Chen, Y., Zhang, X. P., Yuan, J., Cai, B., Wang, X. L., Wu, X. L., Zhang, Y. H., Zhang, X. Y., Yin, T., Zhu, X. H., Gu, Y. J., Cui, S. W., Lu, Z. Q., & Li, X. Y. (2018). Association of body mass index and age with incident diabetes in Chinese adults: A population-based cohort study. *BMJ Open*, 8(9). <https://doi.org/10.1136/bmjopen-2018-021768>
- Chen, Z., Yu, W., & Zhou, L. (2021). ADASYN-Random Forest Based Intrusion Detection Model. *International Conference on Signal Processing and Machine Learning*.

- Cuesta, H. A., Coffman, D. L., Branas, C., & Murphy, H. M. (2019). Using decision trees to understand the influence of individual- and neighborhood-level factors on urban diabetes and asthma. *Health and Place*, 58. <https://doi.org/10.1016/j.healthplace.2019.04.009>
- Escriva, E., Lefrere, T., Martin, M., Aligon, J., Chanson, A., Excoffier, J. B., Labroche, N., Soulé-Dupuy, C., & Monsarrat, P. (2025). Effective data exploration through clustering of local attributive explanations. *Information Systems*, 127. <https://doi.org/10.1016/j.is.2024.102464>
- Farsana, K. S., & Poulouse, A. (2024). Hybrid Convolutional Neural Networks for PIMA Indians Diabetes Prediction. *International Conference on Ubiquitous and Future Networks, ICUFN*, 268–273. <https://doi.org/10.1109/ICUFN61752.2024.10624950>
- Febrian, M. E., Ferdinan, F. X., Sendani, G. P., Suryanigrum, K. M., & Yunanda, R. (2022). Diabetes prediction using supervised machine learning. *Procedia Computer Science*, 216, 21–30. <https://doi.org/10.1016/j.procs.2022.12.107>
- Feng, W., Gou, J., Fan, Z., & Chen, X. (2023). An ensemble machine learning approach for classification tasks using feature generation. *Connection Science*, 35(1). <https://doi.org/10.1080/09540091.2023.2231168>
- Fu, M., Liu, Y., Hou, Z., & Wang, Z. (2025). Interpretable prediction of acute ischemic stroke after hip fracture in patients 65 years and older based on machine learning and SHAP. *Archives of Gerontology and Geriatrics*, 129. <https://doi.org/10.1016/j.archger.2024.105641>
- Fu, Q., Wu, Y., Zhu, M., Xia, Y., Yu, Q., Liu, Z., Ma, X., & Yang, R. (2024). Identifying cardiovascular disease risk in the U.S. population using environmental volatile organic compounds exposure: A machine learning predictive model based on the SHAP methodology. *Ecotoxicology and Environmental Safety*, 286. <https://doi.org/10.1016/j.ecoenv.2024.117210>
- Gunawan, M., Sugiarto, D., & Mardianto, I. (2020). JEPIN (Jurnal Edukasi dan Penelitian Informatika) Peningkatan Kinerja Akurasi Prediksi Penyakit Diabetes Mellitus Menggunakan Metode Grid Search pada Algoritma Logistic Regression. *Jurnal Edukasi Dan Penelitian Informatika*, 6(3).
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining. Concepts and Techniques, 3rd Edition (The Morgan Kaufmann Series in Data Management Systems)*.
- International Diabetes Federation (IDF). (2025). *IDF Diabetes Atlas*.

- Jiang, L., Xia, Z., Zhu, R., Gong, H., Wang, J., Li, J., & Wang, L. (2023). Diabetes risk prediction model based on community follow-up data using machine learning. *Preventive Medicine Reports*, 35. <https://doi.org/10.1016/j.pmedr.2023.102358>
- Khokhar, P. B., Gravino, C., & Palomba, F. (2025). Advances in artificial intelligence for diabetes prediction: insights from a systematic literature review. In *Artificial Intelligence in Medicine* (Vol. 164). Elsevier B.V. <https://doi.org/10.1016/j.artmed.2025.103132>
- Kim, J. H., Shin, J. K., Lee, H., Lee, D. H., Kang, J. H., Cho, K. H., Lee, Y. G., Chon, K., Baek, S. S., & Park, Y. (2021). Improving the performance of machine learning models for early warning of harmful algal blooms using an adaptive synthetic sampling method. *Water Research*, 207. <https://doi.org/10.1016/j.watres.2021.117821>
- Kolb, H., Kempf, K., Röhling, M., & Martin, S. (2020). Insulin: Too much of a good thing is bad. In *BMC Medicine* (Vol. 18, Issue 1). BioMed Central. <https://doi.org/10.1186/s12916-020-01688-6>
- Kumari, S., Kumar, D., & Mittal, M. (2021). An ensemble approach for classification and prediction of diabetes mellitus using soft voting classifier. *International Journal of Cognitive Computing in Engineering*, 2, 40–46. <https://doi.org/10.1016/j.ijcce.2021.01.001>
- Leinonen, T., Wong, D., Vasankari, A., Wahab, A., Nadarajah, R., Kaisti, M., & Airola, A. (2024). Empirical investigation of multi-source cross-validation in clinical ECG classification. *Computers in Biology and Medicine*, 183, 109271. <https://doi.org/10.1016/j.compbimed.2024.109271>
- Mitra, R., Bajpai, A., & Biswas, K. (2023). ADASYN-assisted machine learning for phase prediction of high entropy carbides. *Computational Materials Science*, 223. <https://doi.org/10.1016/j.commatsci.2023.112142>
- Mufila Gaffar, A. W., Isthi'anah, A. L., Purnawansyah, P., Halis, A. M., Mujaddid, S., & Tenripada Syahar, A. U. (2024). Random Search Optimization of Hyperparameter in Random Forest Algorithm for Stunting Prediction. *7th International Seminar on Research of Information Technology and Intelligent Systems: Advanced Intelligent Systems in Contemporary Society, ISRITI 2024 - Proceedings*, 971–976. <https://doi.org/10.1109/ISRITI64779.2024.10963388>
- Muflikhah, L., Bachtiar, F. A., Ratnawati, D. E., & Darmawan, R. (2024a). Improving Performance for Diabetic Nephropathy Detection Using

- Adaptive Synthetic Sampling Data in Ensemble Method of Machine Learning Algorithms. *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika*, 10(1), 123. <https://doi.org/10.26555/jiteki.v10i1.28107>
- Muflikhah, L., Bachtiar, F. A., Ratnawati, D. E., & Darmawan, R. (2024b). Improving Performance for Diabetic Nephropathy Detection Using Adaptive Synthetic Sampling Data in Ensemble Method of Machine Learning Algorithms. *Jurnal Ilmiah Teknik Elektro Komputer Dan Informatika*, 10(1), 123. <https://doi.org/10.26555/jiteki.v10i1.28107>
- Mustofa, F., Safriandono, A. N., Muslikh, A. R., & Setiadi, D. R. I. M. (2023). Dataset and Feature Analysis for Diabetes Mellitus Classification using Random Forest. *Journal of Computing Theories and Applications*, 1(1), 41–48. <https://doi.org/10.33633/jcta.v1i1.9190>
- Ngo, G., Beard, R., & Chandra, R. (2022). Evolutionary bagging for ensemble learning. *Neurocomputing*, 510, 1–14. <https://doi.org/10.1016/j.neucom.2022.08.055>
- Nguyen, X. D. J., & Liu, Y. A. (2025). Methodology for hyperparameter tuning of deep neural networks for efficient and accurate molecular property prediction. *Computers and Chemical Engineering*, 193. <https://doi.org/10.1016/j.compchemeng.2024.108928>
- Nourani, V., Dehghan, M., Baghanam, A. H., & Kantoush, S. A. (2025). Dual purpose of Shapley Additive Explanation (SHAP) in model explanation and feature selection for artificial intelligence-based digital twin of wastewater treatment plant. *Journal of Water Process Engineering*, 75. <https://doi.org/10.1016/j.jwpe.2025.107947>
- Ohanyan, H., Portengen, L., Kaplani, O., Huss, A., Hoek, G., Beulens, J. W. J., Lakerveld, J., & Vermeulen, R. (2022). Associations between the urban exposome and type 2 diabetes: Results from penalised regression by least absolute shrinkage and selection operator and random forest models. *Environment International*, 170. <https://doi.org/10.1016/j.envint.2022.107592>
- Pang, K. (2025). A comparative study of explainable machine learning models with Shapley values for diabetes prediction. *Healthcare Analytics*, 7. <https://doi.org/10.1016/j.health.2025.100390>
- Preechasuk, L., Khaedon, N., Lapinee, V., Tangjittipokin, W., Srivanichakorn, W., Sriwijitkamol, A., Plengvidhya, N., Likitmaskul, S., & Thongtang, N. (2022). Cluster analysis of Thai patients with newly diagnosed type 2 diabetes mellitus to predict disease progression and treatment outcomes:

- A prospective cohort study. *BMJ Open Diabetes Research and Care*, 10(6). <https://doi.org/10.1136/bmjdr-2022-003145>
- Ramadhan, N. G. (2021). Comparative Analysis of ADASYN-SVM and SMOTE-SVM Methods on the Detection of Type 2 Diabetes Mellitus. *Scientific Journal of Informatics*, 8(2), 276–282. <https://doi.org/10.15294/sji.v8i2.32484>
- S, G., Venkata Siva Reddy, R., & Ahmed, M. R. (2024). Exploring the effectiveness of machine learning algorithms for early detection of Type-2 Diabetes Mellitus. *Measurement: Sensors*, 31. <https://doi.org/10.1016/j.measen.2023.100983>
- Saxena, R., Sharma, S. K., Gupta, M., & Sampada, G. C. (2022). A Novel Approach for Feature Selection and Classification of Diabetes Mellitus: Machine Learning Methods. *Computational Intelligence and Neuroscience*, 2022. <https://doi.org/10.1155/2022/3820360>
- Siregar, F. A., Candra, A., & Sutarman. (2024). Determining The Parameter K of K-Nearest Neighbors (KNN) using Random Grid Search. *ICSINTESA 2024 - 2024 4th International Conference of Science and Information Technology in Smart Administration: The Collaboration of Smart Technology and Good Governance for Sustainable Development Goals*, 662–665. <https://doi.org/10.1109/ICSINTESA62455.2024.10748027>
- Sivaranjani, S., Ananya, S., Aravinth, J., & Karthika, R. (2021). Diabetes Prediction using Machine Learning Algorithms with Feature Selection and Dimensionality Reduction. *2021 7th International Conference on Advanced Computing and Communication Systems, ICACCS 2021*, 141–146. <https://doi.org/10.1109/ICACCS51430.2021.9441935>
- Smith, J. W., Everhart, J. E., Dickson, W. C., Knowler, W. C., & Johannes, R. S. (1988). *Using the ADAP Learning Algorithm to Forecast the Onset of Diabetes Mellitus*.
- Thotad, P. N., Bharamagoudar, G. R., & Anami, B. S. (2023). Diabetes disease detection and classification on Indian demographic and health survey data using machine learning methods. *Diabetes and Metabolic Syndrome: Clinical Research and Reviews*, 17(1). <https://doi.org/10.1016/j.dsx.2022.102690>
- Tigga, N. P., & Garg, S. (2020). Prediction of Type 2 Diabetes using Machine Learning Classification Methods. *Procedia Computer Science*, 167, 706–716. <https://doi.org/10.1016/j.procs.2020.03.336>

- Tuppad, A., & Devi Patil, S. (2024). An efficient classification framework for Type 2 Diabetes incorporating feature interactions. *Expert Systems with Applications*, 239. <https://doi.org/10.1016/j.eswa.2023.122138>
- Verma, P., & Khatoon, A. (2024). Data Mining Applications in Healthcare: A Comparative Analysis of Classification Techniques for Diabetes Diagnosis Using the PIMA Indian Diabetes Dataset. *4th International Conference on Innovative Practices in Technology and Management 2024, ICIPTM* 2024. <https://doi.org/10.1109/ICIPTM59628.2024.10563296>
- Wardhani, K. D. K., & Akbar, M. (2022). Diabetes Risk Prediction Using Extreme Gradient Boosting (XGBoost). *Jurnal Online Informatika*, 7(2), 244–250. <https://doi.org/10.15575/join.v7i2.970>
- Xu, J., Chen, T., Fang, X., Xia, L., & Pan, X. (2024). Prediction model of pressure injury occurrence in diabetic patients during ICU hospitalization—XGBoost machine learning model can be interpreted based on SHAP. *Intensive and Critical Care Nursing*, 83. <https://doi.org/10.1016/j.iccn.2024.103715>
- Xu, T., Coco, G., & Neale, M. (2020). A predictive model of recreational water quality based on adaptive synthetic sampling algorithms and machine learning. *Water Research*, 177. <https://doi.org/10.1016/j.watres.2020.115788>
- Yadav, D. C., & Pal, S. (2021). An Experimental Study of Diversity of Diabetes Disease Features by Bagging and Boosting Ensemble Method with Rule Based Machine Learning Classifier Algorithms. *SN Computer Science*, 2(1). <https://doi.org/10.1007/s42979-020-00446-y>

SEKOLAH PASCASARJANA