

DAFTAR PUSTAKA

- Abdullah, A. A., Hafidz, S. A., & Khairunizam, W. (2020, May). Performance comparison of machine learning algorithms for classification of chronic kidney disease (CKD). In *Journal of Physics: Conference Series* (Vol. 1529, No. 5, p. 052077). IOP Publishing.
- Alam, T. M., Iqbal, M. A., Ali, Y., Wahab, A., Ijaz, S., Baig, T. I., ... & Abbas, Z. (2019). A model for early prediction of diabetes. *Informatics in Medicine Unlocked*, 16, 100204.
- Alshdaifat, E. A., Alshdaifat, D. A., Alsarhan, A., Hussein, F., & El-Salhi, S. M. D. F. S. (2021). The effect of preprocessing techniques, applied to numeric features, on classification algorithms' performance. *Data*, 6(2), 11.
- Ameena, R. R., & Ashadevi, B. (2020). Predictive analysis of diabetic women patients using R. In *Systems Simulation and Modeling for Cloud Computing and Big Data Applications* (pp. 99-113). Academic Press.
- Anggoro, D.A., & Novitaningrum, D. (2021). *Comparison Of Accuracy Level Of Support Vector Machine (Svm) And Artificial Neural Network (Ann) Algorithms In Predicting Diabetes Mellitus Disease. ICIC Express Letters*. DOI: 10.24507/icicel.15.01.9.
- Anggoro, D. A., & Supriyanti, W. (2019). *Improving Accuracy By Applying Z-Score Normalization In Linear Regression and Polynomial Regression Model for Real Estate Data. International Journal of Emerging Trends in Engineering Research*, 7(11), 549555. <https://doi.org/10.30534/ijeter/2019/247112019>.
- Anwar, F., Qurat-Ul-Ain, Ejaz, M. Y., & Mosavi, A. (2020). A comparative analysis on diagnosis of diabetes mellitus using different approaches – A survey. *Informatics in Medicine Unlocked*, 21, 100482.
- Asha Kiranmai, S., & Jaya Laxmi, A. (2018). Data mining for classification of power quality problems using WEKA and the effect of attributes on classification accuracy. *Protection and Control of Modern Power Systems*, 3, 1-12.

- Alic, B., Gurbeta, L., & Badnjevic, A. (2017, June). Machine learning techniques for classification of diabetes and cardiovascular diseases. In *2017 6th mediterranean conference on embedded computing (MECO)* (pp. 1-4). IEEE.
- Bahani, K., Moujabbir, M., & Ramdani, M. (2021). An accurate fuzzy rule-based classification systems for heart disease diagnosis. *Scientific African*, 14, e01019.
- Bali, J., & Kant, A. (2016). *Basics of biostatistics: A manual for medical practitioners*. JP Medical Ltd.
- Barhate, R., & Kulkarni, P. (2018). *Analysis of Classifiers for Prediction of Type II Diabetes Mellitus. 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*. doi:10.1109/iccubea.2018.8697856.
- Birjais, R., Mourya, A. K., Chauhan, R., & Kaur, H. (2019). Prediction and diagnosis of future diabetes risk: a machine learning approach. *SN Applied Sciences*, 1, 1-8.
- Breiman, L. (2001). Random forests. *Machine learning*, 45, 5-32.
- Bukhari, M. M., Alkhamees, B. F., Hussain, S., Gumaei, A., Assiri, A., & Ullah, S. (2021). An improved artificial neural network model for effective diabetes prediction. *Complexity*, 2021, 1-10.
- Chandrasekara, E. S. K., Kanchana, W. K. T., & Nandani, E. J. K. P. (2021, December). Comprehensive Study for Diabetes Identification Ability of Various Optimizers in Deep Learning Neural Network. In *2021 5th SLAAI International Conference on Artificial Intelligence (SLAAI-ICAI)* (pp. 1-6). IEEE.
- Crone, E. A., Wendelken, C., Donohue, S., van Leijenhorst, L., & Bunge, S. A. (2006). Neurocognitive development of the ability to manipulate information in working memory. *Proceedings of the National Academy of Sciences*, 103(24), 9315-9320.
- Cutler, A., Cutler, D. R., & Stevens, J. R. (2012). Random forests. *Ensemble machine learning: Methods and applications*, 157-175.

- Cynthia, E. P., Nazir, A., & Syafria, F. (2021). Random Forest Algorithm to Investigate the Case of Acute Coronary Syndrome. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 5(2), 369-378.
- Dharmasaputro, A. A., Fauzan, N. M., Kallista, M., Wibawa, I. P. D., & Kusuma, P. D. (2022, January). Handling Missing and Imbalanced Data to Improve Generalization Performance of Machine Learning Classifier. In *2021 International Seminar on Machine Learning, Optimization, and Data Science (ISMODE)* (pp. 140-145). IEEE.
- Doreswamy, G. S., Nandish, and Santosh, K. J. (2018). Prediction accuracy comparison of predictive models using machine learning for diabetes data set. *Int. J. Adv. Res. Comput. Sci*, 9, 86.
- Dzakiyullah, N. R., Burhanuddin, M. A., Ikram, R. R., Ghani, K. A., & Setyonugroho, W. (2019). Machine learning methods for diabetes prediction. *Int J Innov Technol Explor Eng [Internet]. Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM), Melaka, Malaysia*, 8, 2199-205.
- Edelmann, D., Móri, T. F., & Székely, G. J. (2021). On relationships between the Pearson and the distance correlation coefficients. *Statistics & probability letters*, 169, 108960.
- Fausett, L. *Fundamentals of Neural Networks*; Prentice-Hall: Upper Saddle River, NJ, USA, 1994; ISBN 978-0-13-334186-7
- Fatima, M., & Pasha, M. (2017). Survey of machine learning algorithms for disease diagnostic. *Journal of Intelligent Learning Systems and Applications*, 9(01), 1-16.
- Favaretto, M., De Clercq, E., Schneble, C. O., & Elger, B. S. (2020). What is your definition of Big Data? Researchers' understanding of the phenomenon of the decade. *PloS one*, 15(2), e0228987.
- Friedman, M. A Comparison of Alternative Tests of Significance for the Problem of m Rankings. *Ann. Math. Stat.* 1940, 11, 86–92. [CrossRef] 13. Nemenyi, P. Distribution-Free Multiple Comparisons; Princeton University: Princeton, NJ, USA, 1963.

- Gulati, V., & Raheja, N. (2021, October). Efficiency Enhancement of Machine Learning Approaches through the Impact of Preprocessing Techniques. In *2021 6th International Conference on Signal Processing, Computing and Control (ISPCC)* (pp. 191-196). IEEE.
- Gupta, I., Sharma, V., Kaur, S., & Singh, A. K. (2022). PCA-RF: an efficient Parkinson's disease prediction model based on random forest classification. arXiv preprint arXiv:2203.11287.
- Henderi, H., Wahyuningsih, T., & Rahwanto, E. (2021). 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. *International Journal of Informatics and Information Systems*, 4(1), 13-20.
- Herwanto, H. W., Handayani, A. N., Wibawa, A. P., Chandrika, K. L., & Arai, K. (2021, October). Comparison of Min-Max, Z-Score and Decimal Scaling Normalization for Zoning Feature Extraction on Javanese Character Recognition. In *2021 7th International Conference on Electrical, Electronics and Information Engineering (ICEEIE)* (pp. 1-3). IEEE.
- Ilyas, H., Ali, S., Ponum, M., Hasan, O., Mahmood, M. T., Iftikhar, M., & Malik, M. H. (2021). Chronic kidney disease diagnosis using decision tree algorithms. *BMC nephrology*, 22(1), 1-11.
- Islam, M. T., Raihan, M., Aktar, N., Alam, M. S., Ema, R. R., & Islam, T. (2020, July). Diabetes mellitus prediction using different ensemble machine learning approaches. In *2020 11th international conference on computing, communication and networking technologies (ICCCNT)* (pp. 1-7). IEEE.
- Karun, S., Raj, A., & Attigeri, G. (2019). *Comparative Analysis of Prediction Algorithms for Diabetes. Advances in Computer Communication and Computational Sciences*, 177–187. doi:10.1007/978-981-13-0341-8_16.
- Larose, D. T. (2004). Discovery Knowledge in Data Mining. In *Structure and Bonding* (Vol. 134).
- Loola Bokonda, P., Sidibe, M., Souissi, N., & Ouazzani-Touhami, K. (2023). Machine Learning Model for Predicting Epidemics. *Computers*, 12(3), 54.

- Malakar, S., Roy, S. D., Das, S., Sen, S., Velásquez, J. D., & Sarkar, R. (2022). Computer Based Diagnosis of Some Chronic Diseases: A Medical Journey of the Last Two Decades. *Archives of Computational Methods in Engineering*, 1-43.
- Maniruzzaman, M., Rahman, M. J., Al-MehediHasan, M., Suri, H. S., Abedin, M. M., El-Baz, A., & Suri, J. S. (2018). *Accurate Diabetes Risk Stratification Using Machine Learning: Role of Missing Value and Outliers. Journal of Medical Systems*, 42(5). doi:10.1007/s10916-018-0940-7.
- Mohamed, E. S., Naqishbandi, T. A., Bukhari, S. A. C., Rauf, I., Sawrikar, V., & Hussain, A. (2023). A hybrid mental health prediction model using Support Vector Machine, Multilayer Perceptron, and Random Forest algorithms. *Healthcare Analytics*, 3, 100185.
- Nasution, D. A., Khotimah, H. H., & Chamidah, N. (2019). Perbandingan normalisasi data untuk klasifikasi wine menggunakan algoritma K-NN. *CESS (Journal of Computer Engineering, System and Science)*, 4(1), 78-82.
- Nayeem, M. J., Rana, S., Alam, F., & Rahman, M. A. (2021, February). Prediction of hepatitis disease using K-nearest neighbors, Naive Bayes, support vector machine, multi-layer perceptron and random forest. In *2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)* (pp. 280-284). IEEE.
- Yuvaraj, N., & Sri Preethaa, K. R. (2019). Diabetes prediction in healthcare systems using machine learning algorithms on Hadoop cluster. *Cluster Computing*, 22(Suppl 1), 1-9.
- Patro, S. G. O. P. A. L., & Sahu, K. K. (2015). Normalization: A preprocessing stage. *arXiv preprint arXiv:1503.06462*.
- Prasad, K. S., Reddy, N., & Puneeth, B. N. (2020). A Framework for Diagnosing Kidney Disease in Diabetes Patients Using Classification Algorithms. *SN Computer Science*, 1(2), 1-6.
- Q. Zou, K. Qu, Y. Luo, D. Yin, Y. Ju, H. Tang, Predicting diabetes mellitus with machine learning techniques. *Front. Genet.* 9, 515 (2018).

- Rahman, A. (2019). Statistics-based data preprocessing methods and machine learning algorithms for big data analysis. *International Journal of Artificial Intelligence*, 17(2), 44-65.
- Ramdhani, Y., Susanti, S., Adiwisastro, M. F., & Topiq, S. (2018). Penerapan Algoritma Neural Network Untuk Klasifikasi Kardiotokografi. *Jurnal Informatika*, 5(1), 43-49.
- Raut, J.R., Sharma, Y., Shinde, V.D. (2020). *Performance Evaluation of Various Supervised Machine Learning Algorithms For Diabetes Prediction. European Journal of Molecular & Clinical Medicine*.
- Sajjadnia, Z., Khayami, R., & Moosavi, M. R. (2020). Preprocessing breast cancer data to improve the data quality, diagnosis procedure, and medical care services. *Cancer Informatics*, 19, 1176935120917955.
- Sarwar, M. A., Kamal, N., Hamid, W., & Shah, M. A. (2018). *Prediction of Diabetes Using Machine Learning Algorithms in Healthcare. 2018 24th International Conference on Automation and Computing (ICAC)*. doi:10.23919/iconac.2018.8748992.
- Singh, M., Kaiser, J., & Hahn, H. (2015). Thick electrodes for high energy lithium ion batteries. *Journal of The Electrochemical Society*, 162(7), A1196.
- Singh, A., Thakur, N., & Sharma, A. (2016, March). A review of supervised machine learning algorithms. In *2016 3rd international conference on computing for sustainable global development (INDIACom)* (pp. 1310-1315). IEEE.
- Singh, N., & Singh, P. (2021, September). Exploring the effect of normalization on medical data classification. In *2021 International Conference on Artificial Intelligence and Machine Vision (AIMV)* (pp. 1-5). IEEE.
- Sopian, A., Wiyatno, A., & Riyandi, A. (2019). Komparasi Algoritma Support Vector Machines dengan Algoritma Artificial Neural Network untuk Memprediksi Nilai Persetujuan Kredit Modal Kerja yang Diberikan Bank Umum. *Jurnal Teknologi Informatika dan Komputer*, 5(1), 88-95.
- Supriyanti, W., & Puspitasari, N. (2018). IMPLEMENTASI TEKNIK SELEKSI FITUR FORWARD SELECTION PADA ALGORITMA KLASIFIKASI

DATA MINING UNTUK PREDIKSI MASA STUDI MAHASISWA POLITEKNIK INDONUSA SURAKARTA. *Jurnal Informa: Jurnal Penelitian dan Pengabdian Masyarakat*, 4(2), 49-54.

Suvarnamukhi, B., & Seshashayee, M. (2019). Big data processing system for diabetes prediction using machine learning technique. *Int. J. Innov. Technol. Explor. Eng*, 8(12), 4478-4483.

Syukron, M., Santoso, R., & Widiharih, T. (2020). Perbandingan metode smote random forest dan smote xgboost untuk klasifikasi tingkat penyakit hepatitis C pada imbalance class data. *Jurnal Gaussian*, 9(3), 227-236.

Taz, N. H., Islam, A., & Mahmud, I. (2021). *A Comparative Analysis of Ensemble Based Machine Learning Techniques for Diabetes Identification. 2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST)*. doi:10.1109/icrest51555.2021.9331036.

Tripathi, G., & Kumar, R. (2020). *Early Prediction of Diabetes Mellitus Using Machine Learning. 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*. doi:10.1109/icrito48877.2020.9197832.

Wahono, R. S., Herman, N. S., & Ahmad, S. (2014). A comparison framework of classification models for software defect prediction. *Advanced Science Letters*, 20(10-11), 1945-1950.

Wang, Q., Cao, W., Guo, J., Ren, J., Cheng, Y., Davis, D.N. (2019). *DMP_MI: An Effective Diabetes Mellitus Classification Algorithm on Imbalanced Data with Missing Values. IEEE (Institute of Electrical and Electronics Engineers)*. DOI: 10.1109/ACCESS.2019.2929866.

Wang, Y., Li, Y., Song, Y., & Rong, X. (2020). The influence of the activation function in a convolution neural network model of facial expression recognition. *Applied Sciences (Switzerland)*, 10(5). <https://doi.org/10.3390/app10051897>

Wu, H., Yang, S., Huang, Z., He, J., & Wang, X. (2018). Type 2 diabetes mellitus prediction model based on data mining. *Informatics in Medicine Unlocked*, 10, 100-107.

- Xu, W., Zhang, J., Zhang, Q., & Wei, X. (2017). *Risk prediction of type II diabetes based on random forest model. 2017 Third International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB)*. doi:10.1109/aeecb.2017.7972337.
- Yokoyama, A., & Yamaguchi, N. (2020). Comparison between ANN and random forest for leakage current alarm prediction. *Energy Reports*, 6, 150-157.
- Yuniarti, W. D., Winarko, E., & Musdholifah, A. (2020, November). Data Mining for Student Assessment in e-Learning: A Survey. In *2020 Fifth International Conference on Informatics and Computing (ICIC)* (pp. 1-6). IEEE.
- Zailani, A. U., Perdananto, A., & Sholihin, N. (2020). Pengenalan sejak dini siswa smp tentang machine learning untuk klasifikasi gambar dalam menghadapi revolusi 4.0. *KOMMAS: Jurnal Pengabdian Kepada Masyarakat*, 1(1), 7-15.
- Zhang, C., & Ma, Y. (Eds.). (2012). *Ensemble machine learning: methods and applications*. Springer Science & Business Media.
- Zou, Q., Qu, K., Luo, Y., Yin, D., Ju, Y., & Tang, H. (2018). *Predicting Diabetes Mellitus With Machine Learning Techniques. Frontiers in Genetics*, 9. doi:10.3389/fgene.2018.00515.