

DAFTAR PUSTAKA

- [1]. Adennia, R. (2021) *IMPLEMENTASI CONVOLUTIONAL NEURAL NETWORK UNTUK EKSTRAKSI FITUR CITRA DAUN DALAM KASUS DETEKSI PENYAKIT PADA TANAMAN MANGGA MENGGUNAKAN RANDOM FOREST* - UMM Institutional Repository. Available at: <https://eprints.umm.ac.id/79344/> (Accessed: 3 April 2022).
- [2]. Albahri, A.S. *et al.* (2020) 'Role of biological Data Mining and Machine Learning Techniques in Detecting and Diagnosing the Novel Coronavirus (COVID-19): A Systematic Review', *Journal of Medical Systems*, 44(7). doi:10.1007/s10916-020-01582-x.
- [3]. AlShboul, R. *et al.* (2018) 'A visualization cybersecurity method based on features' dissimilarity', *Computers and Security*, 77. doi:10.1016/j.cose.2018.04.007.
- [4]. Anwar, D.S. *et al.* (2012) 'No Title', (x).
- [5]. Appiahene, P., Missah, Y.M. and Najim, U. (2020) 'Predicting Bank Operational Efficiency Using Machine Learning Algorithm: Comparative Study of Decision Tree, Random Forest, and Neural Networks', *Advances in Fuzzy Systems*, 2020. doi:10.1155/2020/8581202.
- [6]. Apriliah, W. *et al.* (2021) 'Prediksi Kemungkinan Diabetes pada Tahap Awal Menggunakan Algoritma Klasifikasi Random Forest', *Sistemasi*, 10(1), p. 163. doi:10.32520/stmsi.v10i1.1129.
- [7]. Azrar, A. *et al.* (2018) 'Data mining models comparison for diabetes prediction', *International Journal of Advanced Computer Science and Applications*, 9(8), pp. 320–323. doi:10.14569/ijacsa.2018.090841.
- [8]. Balasaraswathi, V.R., Sugumaran, M. and Hamid, Y. (2017) 'Feature selection techniques for intrusion detection using non-bio-inspired and bio-inspired optimization algorithms', *Journal of Communications and Information Networks*, 2(4). doi:10.1007/s41650-017-0033-7.
- [9]. Benali, L. *et al.* (2019) 'Solar radiation forecasting using artificial neural

- network and random forest methods: Application to normal beam, horizontal diffuse and global components’, *Renewable Energy*, 132, pp. 871–884. doi:10.1016/j.renene.2018.08.044.
- [10]. Buulolo, E. (2020) *Buku Data Mining Untuk Perguruan Tinggi*. I. Edited by E. Bulolo.
- [11]. C, M.A. (2020) ‘Reduksi Atribut Pada Dataset Penyakit Jantung dan Klasifikasi’, 4, pp. 994–1006. doi:10.30865/mib.v4i4.2355.
- [12]. California, U., Irvine, I. and Serikat, A. (2021) ‘Machine Learning dalam Memprediksi Diabetes Sejak Dini’, pp. 167–172.
- [13]. Chicco, D. and Jurman, G. (2020) ‘The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation’, *BMC Genomics*, 21(1). doi:10.1186/s12864-019-6413-7.
- [14]. Dipto, I.C. *et al.* (2020) ‘Comparison of Different Machine Learning Algorithms for the Prediction of Coronary Artery Disease’, *Journal of Data Analysis and Information Processing*, 08(02). doi:10.4236/jdaip.2020.82003.
- [15]. Gárate-Escamila, A.K., Hajjam El Hassani, A. and Andrés, E. (2020) ‘Classification models for heart disease prediction using feature selection and PCA’, *Informatics in Medicine Unlocked*, 19. doi:10.1016/j.imu.2020.100330.
- [16]. Geeksforgeeks (2019a) *KDD Process in Data Mining*. Available at: <https://www.geeksforgeeks.org/kdd-process-in-data-mining/?ref=lbp>.
- [17]. Geeksforgeeks (2019b) *Relationship between Data Mining and Machine Learning*, *geeksforgeeks*. Available at: <https://www.geeksforgeeks.org/relationship-between-data-mining-and-machine-learning/>.
- [18]. Hendro, G., Adji, T.B. and Setiawan, N.A. (2012) ‘Penggunaan Metodologi Analisa Komponen Utama (PCA) untuk Mereduksi Faktor-Faktor yang Mempengaruhi Penyakit Jantung Koroner’, *Seminar Nasional ScrETec*, pp. 1–5.

- [19]. Islam Ayon, S. and Milon Islam, M. (2019) 'Diabetes Prediction: A Deep Learning Approach', *International Journal of Information Engineering and Electronic Business*, 11(2), pp. 21–27. doi:10.5815/ijieeb.2019.02.03.
- [20]. Johnson, J.M. and Khoshgoftaar, T.M. (2019) 'Survey on deep learning with class imbalance', *Journal of Big Data*, 6(1). doi:10.1186/s40537-019-0192-5.
- [21]. Lei, C. *et al.* (2019) 'A comparison of random forest and support vector machine approaches to predict coal spontaneous combustion in gob', *Fuel*, 239(September 2018), pp. 297–311. doi:10.1016/j.fuel.2018.11.006.
- [22]. Lv, Z. and Qiao, L. (2020) 'Analysis of healthcare big data', *Future Generation Computer Systems*, 109, pp. 103–110. doi:10.1016/j.future.2020.03.039.
- [23]. Ma, J. (2020) 'Machine Learning in Predicting Diabetes in the Early Stage', *Proceedings - 2020 2nd International Conference on Machine Learning, Big Data and Business Intelligence, MLBDBI 2020*, pp. 167–172. doi:10.1109/MLBDBI51377.2020.00037.
- [24]. Putra, F.M. and Sitanggang, I.S. (2020) 'Classification model of air quality in Jakarta using decision tree algorithm based on air pollutant standard index', in *IOP Conference Series: Earth and Environmental Science*. doi:10.1088/1755-1315/528/1/012053.
- [25]. Sneha, N. and Gangil, T. (2019) 'Analisis diabetes mellitus untuk prediksi awal menggunakan pemilihan fitur yang optimal', 0.
- [26]. Suyanto (2017) *Data Mining untuk Klasifikasi dan Klasterisasi Data*. Bandung: Informatika Bandung.
- [27]. Thakkar, A. and Lohiya, R. (2021) 'Attack classification using feature selection techniques: a comparative study', *Journal of Ambient Intelligence and Humanized Computing*, 12(1). doi:10.1007/s12652-020-02167-9.
- [28]. Thaseen, I.S., Kumar, C.A. and Ahmad, A. (2019) 'Integrated Intrusion Detection Model Using Chi-Square Feature Selection and Ensemble of Classifiers', *Arabian Journal for Science and Engineering*, 44(4).

doi:10.1007/s13369-018-3507-5.

- [29]. Utomo, D.P. (2020) ‘Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung’, 4(April), pp. 437–444. doi:10.30865/mib.v4i2.2080.
- [30]. Xaverius, F., Jayawardana, H. and Hidayat, S.H. (2020) ‘Comparing euclidean distance and nearest neighbor algorithm in an expert system for diagnosis of diabetes mellitus’, *Enfermeria Clinica*, 30, pp. 374–377. doi:10.1016/j.enfcli.2019.07.121.
- [31]. Yan, H. *et al.* (2020) ‘Data mining in the construction industry: Present status, opportunities, and future trends’, *Automation in Construction*. Elsevier B.V., p. 103331. doi:10.1016/j.autcon.2020.103331.
- [32]. Zhang, Y. *et al.* (2019) ‘A novel ensemble method for k-nearest neighbor’, *Pattern Recognition*, 85, pp. 13–25. doi:10.1016/j.patcog.2018.08.003.
- [33]. Zhao, Y. *et al.* (2020) ‘A review of data mining technologies in building energy systems: Load prediction, pattern identification, fault detection and diagnosis’, *Energy and Built Environment*, 1(2), pp. 149–164. doi:10.1016/j.enbenv.2019.11.003.
- [34]. Zhu, Y. *et al.* (2020) ‘Inspector: a lysine succinylation predictor based on edited nearest-neighbor undersampling and adaptive synthetic oversampling’, *Analytical Biochemistry*, 593(January), p. 113592. doi:10.1016/j.ab.2020.113592.

UNIVERSITAS
MIKROSKIL