

BIBLIOGRAPHY

- [1] Pittara, "Heart Disease," <https://www.alodokter.com/penyakit-jantung>. Accessed: Jun. 25, 2025.
- [2] H. Susanto, D. Yanto, N. Rusdiana, and DT Rahayu, "Prediction of Heart and Blood Vessel Disease Risk Using the K-Nearest Neighbor (KNN) Algorithm," 2024.
- [3] JC Bastiaans, J. Hartojo, RA Pramunendar, and N. Andono, "Evaluating the Impact of Particle Swarm Optimization Based Feature Selection on Support Vector Machine Performance in Coral Reef Health Classification," *International Journal of New Media Technology*, vol. 11, no. 2, p. 90, 2024.
- [4] G. Manikandan, B. Pragadeesh, V. Manojkumar, AL Karthikeyan, R. Manikandan, and AH Gandomi, "Classification models combined with Boruta feature selection for heart disease prediction," *Informatics in Medicine Unlocked*, vol. 44, Jan. 2024, doi: 10.1016/j.imu.2023.101442.
- [5] A. Tri Arsanto, A. Faizin, M. Lutfi, Z. Nikmatus Saadah, T. Informatika, and U. Yudharta Pasuruan, "Optimization of the Naive Bayes Algorithm with SMOTETomek Combination for Imbalance Class Fraud Detection," *Sistemasi: Jurnal Sistem Informasi*, 2024, [Online]. Available: <http://sistemasi.ftik.unisi.ac.id>
- [6] KA Barry, Y. Manzali, R. Flouchi, and M. Elfar, "Heart disease approach using modified random forest and particle swarm optimization," *IAES International Journal of Artificial Intelligence*, vol. 14, no. 2, pp. 1242–1251, Apr. 2025, doi: 10.11591/ijai.v14.i2.pp1242-1251.
- [7] H. Zhou, Y. Xin, and S. Li, "A diabetes prediction model based on Boruta feature selection and ensemble learning," *BMC Bioinformatics*, vol. 24, no. 1, Dec. 2023, doi: 10.1186/s12859-023-05300-5.
- [8] MC Ahire et al., "Analysis of Data using PSO model and Random forest," 2019. [Online]. Available: www.ijrti.org
- [9] B. Or, "Improving Requirements Classification with SMOTE-Tomek Preprocessing," Jan. 2025, [Online]. Available: <http://arxiv.org/abs/2501.06491>
- [10] H. Hairani, A. Anggrawan, and D. Priyanto, "Improvement Performance of the Random Forest Method on Unbalanced Diabetes Data Classification Using Smote-Tomek Link," 2023. [Online]. Available: www.joiv.org/index.php/joiv
- [11] A. Indrawati, "APPLICATION OF COMBINATION OF OVERSAMPLING AND UNDERSAMPLING TECHNIQUES TO OVERCOME IMBALANCED DATASET PROBLEMS," *Journal of Informatics and Computers) Accreditation of the Ministry of Research, Technology and Higher Education*, vol. 4, no. 1, 2021, doi: 10.33387/jiko.
- [12] R. M. Prabha and S. Sasikala, "Data Analytics for Imbalanced Datasets," *Journal of Computer Science*, vol. 20, no. 2, pp. 207–217, 2024, doi: 10.3844/jcssp.2024.207.217.
- [13] H. Ahmadpour, O. Bazrafshan, E. Rafiei-Sardooi, H. Zamani, and T. Panagopoulos, "Gully erosion susceptibility assessment in the kondoran watershed using machine learning algorithms and the boruta feature selection," *Sustainability (Switzerland)*, vol. 13, no. 18, Sept. 2021, doi: 10.3390/su131810110.
- [14] SMFD Syed Mustapha, "Predictive Analysis of Students' Learning Performance Using Data Mining Techniques: A Comparative Study of Feature Selection Methods," *Applied System Innovation*, vol. 6, no. 5, Oct. 2023, doi: 10.3390/asi6050086.

- [15] Erdania, Faizal M., and Anggraini RB, "FACTORS RELATED TO THE INCIDENCE OF CORONARY HEART DISEASE (CHD) AT DR. (HC) IR. SOEKARNO HOSPITAL, BANGKA BELITUNG PROVINCE," 2023.
- [16] D.A. Ryfai, N. Hidayat, and E. Santoso, "Classification of Heart Attack Risk Levels Using the K-Nearest Neighbor Method," 2022. [Online]. Available: <http://j-ptiik.ub.ac.id>
- [17] R. Sumara, SH Saputra, N. Mukarommah, NA Wibowo, FL Yumni, and S. Supriyanto, "IDENTIFICATION OF RISK FACTORS FOR CORONARY HEART DISEASE INCIDENTS AT THE HAJI GENERAL HOSPITAL IN SURABAYA," 2021.
- [18] NE Almansouri et al., "Early Diagnosis of Cardiovascular Diseases in the Era of Artificial Intelligence: An In-Depth Review," *Cureus*, Mar. 2024, doi: 10.7759/cureus.55869.
- [19] A. Groenewegen et al., "Improving early diagnosis of cardiovascular disease in patients with type 2 diabetes and COPD: Protocol of the RED-CVD cluster randomized diagnostic trial," *BMJ Open*, vol. 11, no. 10, Oct. 2021, doi: 10.1136/bmjopen-2020-046330.
- [20] A. Paul, D. P. Mukherjee, P. Das, A. Gangopadhyay, A. R. Chintla, and S. Kundu, "Improved Random Forest for Classification," *IEEE Transactions on Image Processing*, vol. 27, no. 8, pp. 4012–4024, Aug. 2018, doi: 10.1109/TIP.2018.2834830.
- [21] L. Breiman, "Random Forests," 2001.
- [22] J. Kennedy and R. Eberhart, "Particle Swarm Optimization," 1995.
- [23] MB Kursu and WR Rudnicki, "Feature Selection with the Boruta Package," 2010. [Online]. Available: <http://www.jstatsoft.org/>
- [24] GEAPA Batista, ALC Bazzan, and MC Monard, "Balancing Training Data for Automated Annotation of Keywords: a Case Study," 2003. Accessed: June. 30, 2025. [Online]. Available: <http://www.inf.ufrgs.br/maslab/pergamus/pubs/balancing-training-data-for.pdf>
- [25] Rohit Kundu, "Confusion Matrix: How To Use It & Interpret Results [Examples]," v7labs. Accessed: Jul. 09, 2025. [Online]. Available: <https://www.v7labs.com/blog/confusion-matrix-guide>