DAFTAR PUSTAKA

Abdullah-All-Tanvir, Khandokar, I. A., & Shatabda, S. (2023). New boosting approaches for improving cluster-based undersampling in problems with imbalanced data. *Decision Analytics Journal*, *8*. https://doi.org/10.1016/j.dajour.2023.100316

Afifah, N. (2025, January 22). *HUBUNGAN FAKTOR RISIKO DENGAN KEJADIAN DIABETES MELLITUS GESTASIONAL PADA IBU HAMIL (Di Wilayah Puskesmas Jelakombo Kabupaten Jombang)*. Repositori Institut Teknologi Sains Dan Kesehatan. https://repository.itskesicme.ac.id/

Agung Khoeruddin, Fahri Andriansyah Sudrajat, Galuh Purnama, Iman Kuwangid, Kurnia Kurnia, & Ricky Firmansyah. (2023). Optimasi Fitur Seleksi Random Forest Menggunakan GA Dalam Klasifikasi Data Penyakit Gagal Jantung. *Jurnal Penelitian Teknologi Informasi Dan Sains*, *1*(2), 01–09. https://doi.org/10.54066/jptis.v1i2.323

Alpaydin, E. (2020). *Introduction to Machine Learning* (4e ed.). https://books.google.co.id/books?id=tZnSDwAAQBAJ&printsec=frontcover&redir_esc=y#v=onepage&q&f=false

AMALIANA, L., ASTUTI, A. B., GADIS, R. S., RABBANI, N. A., & SOVIA, N. A. (2024). HARD-VOTING DAN SOFT-VOTING CLASSIFIER: MODEL KLASIFIKASI RISIKO KEMATIAN PADA PASIEN GAGAL GINJAL KRONIK. *E-Jurnal Matematika*, *13*(4), 210. https://doi.org/10.24843/mtk.2024.v13.i04.p464

Badr, W. (2019, February 22). *Having an Imbalanced Dataset? Here Is How You Can Fix It*. Towards Data Science. https://towardsdatascience.com/having-an-imbalanced-dataset-here-is-how-you-can-solve-it-1640568947eb/

Bawden, A. (2024, November 13). *More than 800 million people around the world have diabetes, study finds*. The Guardian. https://www.theguardian.com/society/2024/nov/13/diabetes-rates-increase-world-study

Britanthia, L., Tanujaya, C., Susanto, B., & Saragih, A. (2020). Perbandingan Metode Regresi Logistik dan Random Forest untuk Klasifikasi Fitur Mode Audio Spotify. *Indonesian Journal of Data and Science (IJODAS)*, *1*(3), 68–78.

Buulolo, E. (2020, January). *DATA MINING UNTUK PERGURUAN TINGGI*. https://repository.deepublish.com/media/publications/592918-uk-perguruan-tinggi-e3135731.pdf

ıni, P. (2024). Deteksi Dini Penyakit Diabetes dengan Algoritma Random Forest. *Jurnal Sains Dan Manajemen*,

earchgate.net/publication/381458063_Deteksi_Dini_Penyakit_

Diabetes_dengan_Menggunakan_Algoritma_Random_Forest/link/666e2a79de777205a32feab0/download?_tp=eyJjb250ZXh0Ijp7ImZpcnN0UGFnZSI6InB1YmxpY2F0aW9uIiwicGFnZSI6InB1YmxpY2F0aW9uIiwicHJldmlvdXNQYWdlIjoiX2RpcmVjdCJ9fQ

Cholis, D. (2023, January 13). *PENERAPAN ENSEMBLE VOTING PADA DECISION TREE C4.5, K-NEAREST NEIGHBOR DAN BACKPROPAGATION UNTUK KLASIFIKASI MAHASISWA DROP OUT*. Repository UIN Sunan Kalijaga. http://digilib.uinsa.ac.id/60774/4/Daffa%20Nur%20Cholis_H92219044%20OK.pdf

de Oliveira, G. P., Fonsêca, A., & Rodrigues, P. C. (2022). Diabetes diagnosis based on hard and soft voting classifiers combining statistical learning models. *Brazilian Journal of Biometrics*, *40*(4), 415–427. https://doi.org/10.28951/bjb.v40i4.605

Diabetes - Gejala, Penyebab, dan Pengobatan. (2024, November 4). Alodokter. https://www.alodokter.com/diabetes

Elgeldawi, E., Sayed, A., Galal, A. R., & Zaki, A. M. (2021). Hyperparameter tuning for machine learning algorithms used for arabic sentiment analysis. *Informatics*, *8*(4). https://doi.org/10.3390/informatics8040079

Fitriyani, F. (2021). Prediksi Diabetes Menggunakan Algoritma Naive Bayes dan Greedy Forward Selection. *Jurnal Nasional Teknologi Dan Sistem Informasi*, *7*(2), 61–69. https://doi.org/10.25077/teknosi.v7i2.2021.61-69

Guo, K., Wan, X., Liu, L., Gao, Z., & Yang, M. (2021). Fault diagnosis of intelligent production line based on digital twin and improved random forest. *Applied Sciences (Switzerland)*, *11*(16). https://doi.org/10.3390/app11167733

Handoko, K., & Sabda Lesmana, L. (2018). PENGELOMPOKKAN DATA MINING PADA JUMLAH PENUMPANG DI BANDARA HANG NADIM INFORMASI ARTIKEL KATA KUNCI. *CBIS JOURNAL*. http://ejournal.upbatam.ac.id/index.php/cbis

Harif, A., & Kassimi, M. (2024). Predictive Modeling of Student Performance Using RFECV-RF for Feature Selection and Machine Learning Techniques. *International Journal of Advanced Computer Science and Applications (IJACSA)*, *15*(7). https://thesai.org/Downloads/Volume15No7/Paper_23-Predictive_Modeling_of_Student_Performance.pdf

, T. Samavi, R. (2022). MLCM: Multi-Label Confusion Matrix.

ng Wahyu Utomo. (2025). Prediksi Diabetes menggunakan ɔle Learning dengan Teknik Soft Voting. *Infotekmesin*, *16*(1), /doi.org/10.35970/infotekmesin.v16i1.2534

IDF Diabetes Atlas 11th Edition. (2025, April 7). International Diabetes Federation. https://idf.org/news/idf-diabetes-atlas-11th-edition/

Kalimah, S. (2022). *KLASIFIKASI PENYAKIT DIABETES MENGGUNAKAN METODE DECISION TREE DAN RANDOM FOREST*. https://repository.unsri.ac.id/81689/1/RAMA_44201_08011181722063.pdf

Khuat, T., & Le, M. (2020, March 30). *Evaluation of Sampling-Based Ensembles of Classifiers on Imbalanced Data for Software Defect Prediction Problem*. SN Computer Science. https://www.researchgate.net/profile/Tung-Khuat/publication/340288155_Evaluation_of_Sampling-Based_Ensembles_of_Classifiers_on_Imbalanced_Data_for_Software_Defect_Prediction_Problems/links/5e85c23092851c2f5274b3e8/Evaluation-of-Sampling-Based-Ensembles-of-Classifiers-on-Imbalanced-Data-for-Software-Defect-Prediction-Problems.pdf

Kumar, A. (2020, September 6). *Hard vs Soft Voting Classifier Python Example*. Data Analytics. https://vitalflux.com/hard-vs-soft-voting-classifier-python-example/

Kumar, A., 'Gangwar, R., Zargar, A. A., Kumar, R., & Sharma Amit. (2023, May 19). *Prevalence of Diabetes in India: A Review of IDF Diabetes Atlas 10th Edition*. Current Diabetes Reviews. https://www.eurekaselect.com/article/130914

Kumar, B. (2024, August 4). *Comparison GridSearchCV and RandomSearchCV*. 360DigiTMG. https://360digitmg.com/blog/comparison-gridsearchcv-and-randomsearchcv

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

Manconi, A., Armano, G., Gnocchi, M., & Milanesi, L. (2022). A Soft-Voting Ensemble Classifier for Detecting Patients Affected by COVID-19. *Applied Sciences (Switzerland)*, *12*(15). https://doi.org/10.3390/app12157554

Marsya Finda, S., & Wahyu Utomo, D. (2024). Klasifikasi Stunting Balita menggunakan Metode Ensemble Learning dan Random Forest. *Infoteknikmesin*, *15*(02). https://doi.org/10.35970/infotekmesin.v15i2.2326

Maulid, R. (2024, January 5). *Konsep Hyperparameter Tuning pada Machine Learning*. DQLab. https://dqlab.id/konsep-hyperparameter-tuning-earning

lay 17). *Prediksi deteksi penyakit kanker payudara dengan algoritma decision tree*. Program Studi Teknik Informatika, Dan Teknologi, Universitas Islam Negeri (UIN) Maulana Malik . http://etheses.uin-malang.ac.id/53151/1/17650087.pdf

Mienye, I. D., & Sun, Y. (2022). A Survey of Ensemble Learning: Concepts, Algorithms, Applications, and Prospects. In *IEEE Access* (Vol. 10, pp. 99129–99149). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2022.3207287

Morgan-Benita, J. A., Galván-Tejada, C. E., Cruz, M., Galván-Tejada, J. I., Gamboa-Rosales, H., Arceo-Olague, J. G., Luna-García, H., & Celaya-Padilla, J. M. (2022). Hard Voting Ensemble Approach for the Detection of Type 2 Diabetes in Mexican Population with Non-Glucose Related Features. *Healthcare (Switzerland)*, *10*(8). https://doi.org/10.3390/healthcare10081362

Nugroho, K., 'Noersasongko, E., Purwanto, Muljono, Fanani, A., Affandy, & Basuki, R. (2019, December 23). *Improving Random Forest Method to Detect Hatespeech and Offensive Word*. IEEE. https://ieeexplore.ieee.org/document/8938451/figures

Nurhalisa. (2024). SISTEM KLASIFIKASI KELAYAKAN AIR MINUM DENGAN SELEKSI FITUR MENGGUNAKAN METODE RECURSIVE FEATURE ELIMINATION WITH CROSS-VALIDATION (RFECV) DAN SELECTKBEST. *Universitas Hasanuddin*.

Peppes, N., Daskalakis, E., Alexakis, T., Adamopoulou, E., & Demestichas, K. (2021). Performance of machine learning-based multi-model voting ensemble methods for network threat detection in agriculture 4.0. *Sensors*, *21*(22). https://doi.org/10.3390/s21227475

Proses Data Mining KDD. (2021, September 30). BINUS University. https://sis.binus.ac.id/2021/09/30/proses-data-mining-kdd/

Rahmadini, Lubis, E., Priansyah, A., R.W.N, Y., & Meutia, T. (2023). *PENERAPAN DATA MINING UNTUK MEMPREDIKSI HARGA BAHAN PANGAN DI INDONESIA MENGGUNAKAN ALGORITMA K-NEAREST NEIGHBOR* (Vol. 4, Issue 4). https://www.researchgate.net/publication/373982713_PENERAPAN_DATA_MINING_UNTUK_MEMPREDIKSI_HARGA_BAHAN_PANGAN_DI_INDONESIA_MENGGUNAKAN_ALGORITMA_K-NEAREST_NEIGHBOR

Riski Indra Pratama, A., Amalia Latipah, S., Nurina Sari, B., Ilmu Komputer, F., Singaperbangsa Karawang Jl HSRonggo Waluyo, U., Telukjambe Tim, K., Karawang, K., & Barat, J. (n.d.). *OPTIMASI KLASIFIKASI CURAH HUJAN MENGGUNAKAN SUPPORT VECTOR MACHINE (SVM) DAN RECURSIVE FEATURE ELIMINATION (RFE)*.

), C., & Utami, E. (n.d.). *Optimalisasi Seleksi Fitur dengan Credit Scoring menggunakan ANN: Vols. x, No.x*. .uho.ac.id/index.php/journal

July 5). *Prevalensi Diabetes Indonesia Naik Jadi 11,7% pada* Katadata. https://databoks.katadata.co.id/layanan-konsumen-

kesehatan/statistik/8a95a31a9cb29b4/prevalensi-diabetes-indonesia-naik-jadi-117-pada-2023

Saputra, A. (2023). *SKRIPSI IMPLEMENTASI SOFT VOTING CLASSIFIER UNTUK DIAGNOSIS PNEUMONIA*. *Universitas Hasanuddin*

Setyawan, N. H., & Wakhidah, N. (2025). ANALISIS PERBANDINGAN METODE LOGISTIC REGRESSION, RANDOM FOREST, GRADIENT BOOSTING UNTUK PREDIKSI DIABETES. *JIPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, *10*(1), 150–162. https://doi.org/10.29100/jipi.v10i1.5743

Shah, R. (2025, March 13). *Tune Hyperparameters with GridSearchCV*. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2025/03/tune-hyperparameters-with-gridsearchcv/

Sherazi, S. W. A., Bae, J. W., & Lee, J. Y. (2021). A soft voting ensemble classifier for early prediction and diagnosis of occurrences of major adverse cardiovascular events for STEMI and NSTEMI during 2-year follow-up in patients with acute coronary syndrome. *PLoS ONE*, *16*(6 June 2021). https://doi.org/10.1371/journal.pone.0249338

Soepangkat, B. O. P., Norcahyo, R., Effendi, M. K., & Pramujati, B. (2020). Multi-response optimization of carbon fiber reinforced polymer (CFRP) drilling using back propagation neural network-particle swarm optimization (BPNN-PSO). *Engineering Science and Technology, an International Journal*, *23*(3), 700–713. https://doi.org/10.1016/j.jestch.2019.10.002

Sunny, S., Pinky, S., Jalal, S., Kayser, M., Wadud, M., & Mansoor, N. (2024). Soft Voting Ensemble- Based Approach for Diagnosing Diabetes Mellitus. *2024 International Conference on Advances in Computing, Communication, Electrical, and Smart Systems: Innovation for Sustainability, ICACCESS 2024*. https://doi.org/10.1109/iCACCESS61735.2024.10499577

Tan, P.-N., Steinbach, M., Karpatne, A., & Kumar, V. (2019). *Introduction to Data Mining*.

Thorn, J. (2020, March 9). *Decision Trees Explained*. Towards Data Science. https://towardsdatascience.com/decision-trees-explained-3ec41632ceb6/

Topre, G. (2023, August 17). *Exploring Edited Nearest Neighbour Undersampling Technique for Imbalanced Classification*. Python in Plain English (Medium). https://python.plainenglish.io/exploring-edited-nearest-neighbour-technique-for-imbalanced-classification-9e1dbe3d17f7

awan, Odi. (2023, February 1). *IMPLEMENTASI ALGORITMA MENENTUKAN PENERIMA BANTUAN SOSIAL (STUDI IKASI KEMISKINAN DESA LURAH)*. JATI (Jurnal Mahasiswa ka).

Willy, W., Rini, D., & Samsuryadi, S. (2021). Perbandingan Algoritma Random Forest Classifier, Support Vector Machine dan Logistic Regression Clasifier Pada Masalah High Dimension (Studi Kasus: Klasifikasi Fake News). *JURNAL MEDIA INFORMATIKA BUDIDARMA*, *5*(4). https://doi.org/10.30865/mib.v5i4.3177

Yakip. (2020). RANCANG BANGUN SISTEM KLASIFIKASI MINERAL DAN BATUAN MENGGUNAKAN TENSORFLOW.JS. *Universitas Hasanuddin*.

Zhang, Y., Nie, B., Du, J., Chen, J., Du, Y., Jin, H., Zheng, X., Chen, X., & Miao, Z. (2023). Feature selection based on neighborhood rough sets and Gini index. *PeerJ Computer Science*, *9*. https://doi.org/10.7717/PEERJ-CS.1711

Zhou, X., Lu, P., Zheng, Z., Tolliver, D., & Keramati, A. (2020). Accident Prediction Accuracy Assessment for Highway-Rail Grade Crossings Using Random Forest Algorithm Compared with Decision Tree. *Reliability Engineering & System Safety*, *200*. https://www.sciencedirect.com/science/article/abs/pii/S0951832019308580?via%3Dihub