

## DAFTAR PUSTAKA

- Agusta, Z. P., & Adiwijaya. (2019). Modified balanced random forest for improving imbalanced data prediction. *International Journal of Advances in Intelligent Informatics*, 5(1), 58–65. <https://doi.org/10.26555/ijain.v5i1.255>
- Akinbo, R. S., & Daramola, O. A. (2021) *Ensemble Machine Learning Algorithms for Prediction and Classification of Medical Images*. IntechOpen
- Alassaf, M., & Qamar, A. M. (2022). Improving Sentiment Analysis of Arabic Tweets by One-way ANOVA. *Journal of King Saud University - Computer and Information Sciences*, 34(6), 2849–2859. <https://doi.org/10.1016/j.jksuci.2020.10.023>
- Al Mansour, M. A. (2020). The Prevalence and Risk Factors of Type 2 Diabetes Mellitus (DMT2) in a Semi-Urban Saudi Population. *International Journal of Enviromental Research and Public Health*, 17(7). doi:10.3390/ijerph17010007
- Alshdaifat, E., Alshdaifat, D., Alsarhan, A., Hussein, F., & El-Salhi, S. M. F. S. (2021). The Effect of Preprocessing Techniques, Applied to Numeric Features, on Classification Algorithms' Performance. *Data*, 6(2), 11. doi:10.3390/data6020011
- Arowolo, M., Abdulsalam, S., Yakub, S., & M.D, S. (2016). A Feature Selection Based on One Way Anova For Microarray Data Classification A Feature Selection Based on One-Way-Anova for Microarray Data Classification. *Al-Hikmah Journal of Pure and Applied Science*, 3(December), 30–35.
- Asril, N. M., Tabuchi, K., Tsunematsu, M., Kobayashi, T., & Kakehashi, M. (2019). Qualitative Rural Indonesian Study of Diabetes Knowledge, Health Beliefs, and Behaviors in Type 2 Diabetes Patients. *Health*, 11(02), 263–275. <https://doi.org/10.4236/health.2019.112023>
- Bansal, N., (2015), Prediabetes diagnosis and treatment: A review, *World Journal of Diabetes*, 6, 2, 296–303.
- Bauder, R. A., & Khoshgoftaar, T. M. (2018). The effects of varying class distribution on learner behavior for medicare fraud detection with imbalanced big data. *Health Information Science and Systems*, 6(9), 1–14. <https://doi.org/10.1007/s13755-018-0051-3>
- Bauder, R. A., Khoshgoftaar, T. M., & Hasanin, T. (2018). An Empirical Study on Class Rarity in Big Data. *Proceedings - 17th IEEE International Conference on Machine Learning and Applications, ICMLA 2018*, 785–790. <https://doi.org/10.1109/ICMLA.2018.00125>

- Beckman, J. (2013). IDF Diabetes Atlas. In *International Diabetes Federation*.
- Berrar, D. (2018a). Bayes' theorem and naive bayes classifier. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics*, 1–3(September), 403–412. [https://doi.org/10.1016/B978-0:12-809633-8.20473-1](https://doi.org/10.1016/B978-0-12-809633-8.20473-1)
- Berrar, D. (2018b). Cross-Validation. *Encyclopedia of Bioinformatics and Computational Biology: ABC of Bioinformatics*, 1–3(April), 542–545. <https://doi.org/10.1016/B978-0:12-809633-8.20349-X>
- Booth, F.W., Roberts, C.K. & Laye, M.J., (2012), Lack of exercise is a major cause of chronic diseases, *Comprehensive Physiology*, 2, 2, 1143–1211.
- Budiastutik, I., Kartasurya, M. I., Subagio, H. W., & Widjanarko, B. (2022). High Prevalence of Prediabetes and Associated Risk Factors in Urban Areas of Pontianak, Indonesia: A Cross-Sectional Study. *Journal of Obesity*, 2022. <https://doi.org/10.1155/2022/4851044>
- Care, D. (2021). Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2021. *Diabetes Care*, 44(January), S15–S33. <https://doi.org/10.2337/dc21-S002>
- Canadian Diabetes Association. (2013). Canadian Diabetes Association Clinical Practice Guidelines. In *Elsevier* (Vol. 37). <https://doi.org/10.1503/cmaj.080554>
- Campagna, D., Alamo, A., Di Pino, A., Russo, C., Calogero, A.E., Purrello, F. & Polosa, R. (2019). Smoking and diabetes: Dangerous liaisons and confusing relationships, *Diabetology and Metabolic Syndrome*, 11, 1, 1–12. <https://doi.org/10.1186/s13098-019-0482-2>.
- Caruana, R. & Niculescu-Mizil, A. (2004). Data Mining in Metric Space: An Empirical Analysis of Supervised Learning Performance Criteria. *International Conference on Knowledge Discovery and Data Mining*, 69–78. <https://doi.org/10.1145/1014052.1014063>
- Chaudhary, N., & Tyagi, N. (2018). Diabetes Mellitus: An Overview. *International Journal of Research and Development in Pharmacy and Life Sciences*, 7(4), 3030–3033. [https://doi.org/10.21276/IJRDPL.2278-0238.2018.7\(4\).3030:3033](https://doi.org/10.21276/IJRDPL.2278-0238.2018.7(4).3030:3033)
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 321–357. <https://doi.org/10.1002/eap.2043>



- Deberneh, H. M., & Kim, I. (2021). Prediction of Type 2 Diabetes Based on Machine Learning Algorithm. *International Journal of Environmental Research and Public Health*, 18(6), 9–11. <https://doi.org/10.3390/ijerph18063317>
- Dong, Y., Zou, Z., Yang, Z., Wang, Z., Yang, Y., Ma, J., Dong, B., Ma, Y., & Arnold, L. (2018). Prevalence of excess body weight and underweight among 26 Chinese ethnic minority children and adolescents in 2014: A cross-sectional observational study. *BMC Public Health*, 18(562), 1–9. <https://doi.org/10.1186/s12889-018-5352-6>
- Ekinci, E.I., Clarke, S., Thomas, M.C., Moran, J.L., Cheong, K., Macisaac, R.J. & Jerums, G. (2011), Dietary Salt Intake and Mortality in Patients With Type 2 Diabetes, *Diabetes Care*, 34, 3, 703–709.
- Elssied, N. O. F., Ibrahim, O., & Osman, A. H. (2014). A Novel Feature Selection Based on One-Way ANOVA F-Test for E-Mail Spam Classification. *Research Journal of Applied Sciences, Engineering and Technology*, 7(3), 625–638. <https://doi.org/10.19026/rjaset.7.299>
- Elseddawy, A. I., Karim, F. K., Hussein, A. M., & Khafaga, D. S. (2022). Predictive Analysis of Diabetes-Risk with Class Imbalance. *Computational Intelligence and Neuroscience*, 2022, 1–16. <https://doi.org/10.1155/2022/3078025>
- Ewa, O.-B., & Piłśniak, A. (2020). Can Type 2 Diabetes Be Cured? *International Journal of Science and Research Methodology*, 12(2), 76–77. <https://doi.org/10.17925/USE.2016.12.02.76>
- Fan, J., Upadhye, S., & Worster, A. (2006). Understanding receiver operating characteristic (ROC) curves. *CJEM*, 8(01), 19–20. <https://doi.org/10.1017/S1481803500013336>
- Fountaine, T., Lembong, J., Nair, R. & Dyckerhoff, C.S., (2016), *Tackling Indonesia's diabetes challenge: Eight approaches from around the world*, [https://www.mckinsey.com/~media/McKinsey/Industries/Healthcare/Systems\\_and\\_Services/Our\\_Insights/Tackling\\_Indonesias\\_diabetes\\_challenge\\_Eight\\_approaches\\_from\\_around\\_the\\_world/Tackling-Indonesias-diabetes-challenge-Eight-approaches-from-around-the-world.a.](https://www.mckinsey.com/~media/McKinsey/Industries/Healthcare/Systems_and_Services/Our_Insights/Tackling_Indonesias_diabetes_challenge_Eight_approaches_from_around_the_world/Tackling-Indonesias-diabetes-challenge-Eight-approaches-from-around-the-world.a.)
- García, R., Luengo, J., & Herrera, F. (2015). Data Preprocessing in Data Mining. Springer, 72.
- Grandini, M., Bagli, E., & Visani, G. (2020). Metrics for Multi-Class Classification: an Overview. *ArXiv*, 1–17. <http://arxiv.org/abs/2008.05756>

- Guan, H., & Zhang, C. (2022). Predicting Diabetes in Imbalanced Datasets using Neural Networks. *Proceedings of the 13th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics, BCB 2022*. <https://doi.org/10.1145/3535508.3545540>
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining Concepts and Techniques* (3rd Editio). Morgan Kaufmann.
- Hasanin, T., Khoshgoftaar, T. M., Leevy, J. L., & Bauder, R. A. (2019). Severely imbalanced Big Data challenges: investigating data sampling approaches. *Journal of Big Data*, 6(107), 1–25. <https://doi.org/10.1186/s40537-019-0274-4>
- Hassine, K., Erbad, A., & Hamila, R. (2019). Important Complexity Reduction of Random Forest in Multi-Classification Problem. *2019 15th International Wireless Communications and Mobile Computing Conference, IWCMC 2019*, 226–231. <https://doi.org/10.1109/IWCMC.2019.8766544>.
- Hastie, T., Tibshirani, R., & Friedman, J. (2008). *The Elements of Statistical Learning*, 2<sup>nd</sup> edition, Springer, New York/Berlin/Heidelberg.
- Hidayat, B., Ramadani, R.V., Rudijanto, A., Soewondo, P., Suastika, K. & Siu Ng, J.Y. (2022). Direct Medical Cost of Type 2 Diabetes Mellitus and Its Associated Complications in Indonesia, *Elsevier*, 28, 82–89. <https://doi.org/10.1016/j.vhri.2021.04.006>,.
- Hossin, M. & Sulaiman, M.N. (2015). A Review On Evaluation Metrics For Data Classification Evaluations. *International Journal of Data Mining & Knowledge Management Process (IJDKP)*, 5(2). DOI : 10.5121/ijdkp.2015.5201
- Hurwitz, J., & Kirsch, D. (2018). Machine Learning for dummies. In *Journal of the American Society for Information Science*. John Wiley & Sons, Inc. <https://doi.org/10.1002/asi.4630350509>
- Joseph, S. R., Hlomani. & Letsholo, K. (2016). Data Mining Algorithms: An Overview. *International Journal of Computers and Technology*, 15(6), 6806–6813. <http://dx.doi.org/10.24297/ijct.v15i6.1615>
- Kececi, A., Yildirak, A., Ozyazici, K., Ayluctarhan, G., Agbulut, O., & Zincir, I. (2020). Implementation of machine learning algorithms for gait recognition. *Engineering Science and Technology, an International Journal*, 23(4), 931–937. <https://doi.org/10.1016/j.jestch.2020.01.005>
- Kotsiantis, S., Kanellopoulos, D., & Pintelas, P. (2006). Handling imbalanced datasets : A review. *International Transactions on Computer Science and Engineering*, 30(1), 25–36.

- Kurniawaty, E. (2014). Diabetes Mellitus. *Jurnal Kedokteran Universitas Lampung (JK Unila)*, 4(7), 114-119.
- Kumar, A., Arora, A., Sharma, P., Anikhindi, S. A., Bansai, N., Singla, V., Khare, S., & Srivastava, A. (2020). Is diabetes mellitus associated with mortality and severity of COVID-19? A meta-analysis. *Elsevier Public Health Emergency Collection*, 14(4), 535-545. doi: <https://dx.doi.org/10.1016%2Fj.dsx.2020.04.044>
- Larose, D. T., & Larose, C. D. (2014). Discovering Knowledge in Data. In *Discovering Knowledge in Data: An Introduction to Data Mining: Second Edition* (Vol. 9780470908). Wiley. <https://doi.org/10.1002/9781118874059>
- Latiifah, I. R. N & Rosyid, F. N. (2020). Hubungan Antara Karakteristik Responden Dengan Kadar Glukosa Darah Puasa Pada Penderita Diabetes Mellitus Tipe 2 Di Puskesmas Purwosari Surakarta. Skripsi thesis, Universitas Muhammadiyah Surakarta.
- Ley, S. H., Schulze, M. B., Hivert, M.-F., Meigs, J. B., & Hu, F. B. (2002). Risk factors for type 2 diabetes mellitus. In *Diabetes in America* (3rd Editio). National Institute of Diabetes and Digestive and Kidney Diseases
- Mahapoonyanont, N., Mahapoonyanont, T., Pengkaew, N., & Kamhangkit, R. (2010). Power of the test of one-way ANOVA after transforming with large sample size data. *Procedia - Social and Behavioral Sciences*, 9, 933–937. <https://doi.org/10.1016/j.sbspro.2010.12.262>
- Mehra, N., & Gupta, S. (2013). Survey on Multiclass Classification Methods. *International Journal of Computer Science and Information Technologies*, 4(November), 572–576. <https://www.cs.utah.edu/~piyush/teaching/aly05multiclass.pdf>
- Mitchell, T. M. (1997). *Machine Learning*. McGraw Hill.
- Moral, P. Del, Nowaczyk, S., & Pashami, S. (2022). Why Is Multiclass Classification Hard? *IEEE Access*, 10, 80448–80462. <https://doi.org/10.1109/ACCESS.2022.3192514>
- Morteza, A., Jenab, Y., Nargesi, A. A., & Ghazizadeh, Z. (2012). Urea and Oxidative Stress in Type 2 Diabetes. *Journal of Metabolic Syndrome*, 01(02). <https://doi.org/10.4172/2167-0943.1000105>
- Muntean, M., V˘alean, H., Joldes, R., & Ceuca, E. (2011). Feature Selection for Classifier Accuracy Improvement. *Acta Universitatis Apulensis*, 26, 203–216.
- Nathan, D. M., Balkau, B., Bonora, E., Borch-Johnsen, K., Buse, J. B., Colagiuri,

- S., Davidson, M. B., DeFronzo, R., Genuth, S., Holman, R. R., Ji, L., Kirkman, S., Knowler, W. C., Schatz, D., Shaw, J., Sobngwi, E., Steffes, M., Vaccaro, O., Wareham, N., ... Kahn, R. (2009). International Expert Committee Report on The Role of the A1C Assay in the Diagnosis of Diabetes. *Diabetes Care*, 32(7), 1327–1334. <https://doi.org/10.2337/dc09-9033>
- Nexøe, J., Halvorsen, P. A., & Kristiansen, I. S. (2007). Review Article: Critiques of the risk concept — valid or not? *Scandinavian Journal of Public Health*, 35, 648–654. <https://doi.org/10.1080/14034940701418897>
- Nilsson, N. J. (2005). INTRODUCTION TO MACHINE LEARNING. In *Stanford University*. <http://www.ncbi.nlm.nih.gov/pubmed/21172442>
- Ortiz-Martínez, M., González-González, M., Martagón, A. J., Hlavinka, V., Willson, R. C., & Rito-Palomares, M. (2022). Recent Developments in Biomarkers for Diagnosis and Screening of Type 2 Diabetes Mellitus. *Current Diabetes Reports*, 22, 95–115. <https://doi.org/10.1007/s11892-022-01453-4>
- Ost, K. J., Anderson, D. W., & Cadotte, D. W. (2021). *Delivering Precision Medicine to Patients with Spinal Cord Disorders; Insights into Applications of Bioinformatics and Machine Learning from Studies of Degenerative Cervical Myelopathy*. IntechOpen
- Pangestu, A. D. (2020). HUBUNGAN ASUPAN NATRIUM DAN KALIUM, STATUS GIZI, AKTIVITAS FISIK DENGAN TEKanan DARAH PADA PASIEN DIABETES MELITUS DI PUSKESMAS KECAMATAN PESANGGRAHAN Association. *ARGIPA (Arsip Gizi Dan Pangan)*, 4(2), 54–64. <https://doi.org/10.22236/argipa.v4i2.3875>
- Pavya, K., & Srinivasan, D. B. (2017). Feature Selection Techniques in Data Mining: A Study. *International Journal of Scientific Development and Research*, 2(6), 594–598.
- Practice, D. R. and C. (2014). Diagnostic criteria and classification of hyperglycaemia first detected in pregnancy: a World Health Organization Guideline. *Diabetes Research and Clinical Practice*, 103(3), 341–363. <https://doi.org/10.1016/j.diabres.2013.10.012>
- Priya, M., & Muthulakshmi, M. (2017). Optimization Techniques Used in Various Field – A Study. *International Journal on Applications in Basic and Applied Sciences*, 3(1), 7–9.
- Pujari, P., & Gupta, J. B. (2012). Improving Classification Accuracy by Using Feature Selection and Ensemble Model. *International Journal of Soft Computing and Engineering (IJSCE)*, 2(2), 380–386.



- Purbolaksono, M. D., Irvan Tantowi, M., Imam Hidayat, A., & Adiwijaya, A. (2021). Perbandingan Support Vector Machine dan Modified Balanced Random Forest dalam Deteksi Pasien Penyakit Diabetes. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(2), 393–399. <https://doi.org/10.29207/resti.v5i2.3008>
- Rajput, M. R., & Khedgikar, S. S. (2022). Diabetes prediction and analysis using medical attributes: A Machine learning approach. *Journal of Xi'an University of Architecture & Technology*, XIV(1), 98–103. <https://doi.org/10.37896/JXAT14.01/314405>
- Rani, K. J. (2020). Diabetes Prediction Using Machine Learning Techniques. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 6(4), 294–305. [https://doi.org/10.1007/978-981-19-2980-9\\_26](https://doi.org/10.1007/978-981-19-2980-9_26)
- Rashid, A. (2020). “Diabetes Dataset”, Mendeley Data, V1, doi: 10.17632/wj9rwkp9c2.1
- Reshmi, S., Biswas, S. K., Boruah, A. N., Thounaojam, D. M., & Purkayastha, B. (2022). Diabetes Prediction Using Machine Learning Analytics. 2022 *International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON)*, May, 108–112. <https://doi.org/10.1109/COM-IT-CON54601.2022.9850922>
- Ridwan, A. (2020). Penerapan Algoritma *Naïve Bayes* Untuk Klasifikasi Penyakit Diabetes Mellitus. *Jurnal Sistem Komputer dan Kecerdasan Buatan*, 4(1), 15–21.
- Rifiana, A. J., & Suharyanto, T. (2020). Hubungan Diabetes Mellitus dan Hipertensi Dengan Kejadian Corona Virus Deases-19 (Covid-19) di Wisma Atlet Jakarta Tahun 2020. Universitas Nasional, 19, 1–15
- Riset Kesehatan Dasar (Riskesdas). (2013). Badan Penelitian dan Pengembangan Kesehatan Kementerian RI tahun 2013.
- Riset Kesehatan Dasar (Riskesdas). (2018). Badan Penelitian dan Pengembangan Kesehatan Kementerian RI tahun 2018.
- Riset Kesehatan Dasar (Riskesdas). (2020). Badan Penelitian dan Pengembangan Kesehatan Kementerian RI tahun 2020.
- Ross, B. C. (2014). Mutual Information between Discrete and Continuous Data Sets. *PLoS ONE*, 9(2). <https://doi.org/10.1371/journal.pone.0087357>
- Rout, N., Kuhoo, & Mishra, D. (2018). Analysing the Multi-class Imbalanced

Datasets using Boosting Methods and Relevant Information. *International Journal of Pure and Applied Mathematics*, 120(6), 6171–6191.

Sacks, D. B. (2011). A1C versus Glucose Testing: A Comparison. *Diabetes Care*, 34, 518–523. <https://doi.org/10.2337/dc10-1546>

Sadeghi, E., Hosseini, S. M., Vossoughi, M., Aminorroaya, A., & Amini, M. (2020). Association of Lipid Profile with Type 2 diabetes in First-Degree Relatives: A 14-Year Follow-Up Study in Iran. *Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy*, 13, 2743–2750. <https://doi.org/10.2147/DMSO.S259697>

Sari, Y., Yusuf, S., Haryanto, Kusumawardani, L.H., Sumeru, A., Sutrisna, E. & Saryono. (2022) The cultural beliefs and practices of diabetes self-management in Javanese diabetic patients: An ethnographic study, *Heliyon*, 8, e08873. <https://doi.org/10.1016/j.heliyon.2022.e08873>,.

Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. In *Advances in Computational Approaches for Artificial Intelligence, Image Processing, IoT and Cloud Applications*, edited by Bhanu Prakash K N and M. Shivakumar, 159-160. Springer Nature, 2021. <https://doi.org/10.1007/s42979-021-00592-x>

Sarwar, N., Gao, P., Kondapally Seshasai, S. R., Gobin, R., Kaptoge, S., Di Angelantonio, E., Ingelsson, E., Lawlor, D. A., Selvin, E., Stampfer, M., Stehouwer, C. D. A., Lewington, S., Pennells, L., Thompson, A., Sattar, N., White, I. R., Ray, K. K., Danesh, J., Tipping, R. W., ... Wormser, D. (2010). Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: A collaborative meta-analysis of 102 prospective studies. *The Lancet*, 375(9733), 2215–2222. [https://doi.org/10.1016/S0140-6736\(10\)60484-9](https://doi.org/10.1016/S0140-6736(10)60484-9)

Seiffert, C., Khoshgoftaar, T. M., Van Hulse, J., & Napolitano, A. (2007). Mining Data with Rare Events: A Case study. *Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI*, 2, 132–139. <https://doi.org/10.1109/ICTAI.2007.71>

Series, I. (2021). *Machine Learning Algorithms, Models and Applications* (Jaydip Sen (ed.); Vol. 7). IntechOpen.

Simbolon, D., Siregar, A., & Talib, R. A. (2020). Physiological factors and physical activity contribute to the incidence of type 2 diabetes mellitus in Indonesia. *Jurnal Kesehatan Masyarakat Nasional*, 15(3), 120–127. <https://doi.org/10.21109/KESMAS.V15I3.3354>

Sliwinska-Mosson, M. & Milnerowicz, H. (2017) The impact of smoking on the development of diabetes and its complications, *Diabetes and Vascular Disease Research*, 14, 4, 265–276.



- Soewondo, P., & Pramono, L. A. (2011). Prevalence, characteristics, and predictors of pre-diabetes in Indonesia. *Medical Journal of Indonesia*, 20(4), 283–294. <https://doi.org/10.13181/mji.v20i4.465>
- Somasundaram, A., & Reddy, U. S. (2016). Data Imbalance: Effects and Solutions for Classification of Large and Highly Imbalanced Data. *Proceedings of 1st International Conference on Research in Engineering, Computers and Technology (ICRECT 2016)*, November, 28–34.
- Song, D. K., Hong, Y. S., Sung, Y. A., & Lee, H. (2022). Association of serum creatinine levels and risk of type 2 diabetes mellitus in Korea: a case control study. *BMC Endocrine Disorders*, 22(1), 1–7. <https://doi.org/10.1186/s12902-021-00915-2>
- Sugiarta, I. G. R. M., & Darmita, I. G. K. (2020). Profil penderita Diabetes Mellitus Tipe-2 (DM-2) dengan komplikasi yang menjalani rawat inap di Rumah Sakit Umum Daerah (RSUD) Klungkung, Bali tahun 2018. *Intisari Sains Medis*, 11(1), 7–12. <https://doi.org/10.15562/ism.v11i1.515>
- Sun, S., Cao, Z., Zhu, H., & Zhao, J. (2020). A Survey of Optimization Methods from a Machine Learning Perspective. *IEEE Transactions on Cybernetics*, 50(8), 3668–3681. <https://doi.org/10.1109/TCYB.2019.2950779>
- Suryanegara, G. A. B., Adiwijaya., & Purbolaksono, M. D. (2021). Peningkatan Hasil Klasifikasi pada Algoritma Random Forest untuk Deteksi Pasien Penderita Diabetes Menggunakan Metode Normalisasi. *Jurnal Rekayasa Sistem dan Teknologi Informasi*, 5(1), 114-122.
- Swamynathan, M. (2019). Mastering Machine Learning with Python in Six Steps. In *Apress*. <https://doi.org/10.1007/978-1-4842-4947-5>
- Tabák, A. G., Herder, C., Rathmann, W., Brunner, E. J., & Kivimäki, M. (2012). Prediabetes: A high-risk state for developing diabetes. *The Lancet*, 379(9833), 2279–2290. [https://doi.org/10.1016/S0140-6736\(12\)60283-9](https://doi.org/10.1016/S0140-6736(12)60283-9)
- Vergara, J. R., & Estévez, P. A. (2014). A Review of Feature Selection Methods Based on Mutual Information. *Neural Computing and Applications*, 24, 175–186. <https://doi.org/10.1007/s00521-013-1368-0>
- Wasserstein, R. L., & Lazar, N. A. (2016). The ASA’s Statement on p-Values: Context, Process, and Purpose. *American Statistician*, 70(2), 129–133. <https://doi.org/10.1080/00031305.2016.1154108>
- World Health Organization (WHO). (2015). World Health Statistics 2015.



- World Health Organization (WHO). (2018). World Health Statistics 2018.
- World Health Organization (WHO). (2019). World Health Statistics 2019.
- World Health Organization (WHO). (2020). World Health Statistics 2020.
- World Health Organization (WHO). (2021). World Health Statistics 2021.
- Widayanti, A.W., Heydon, S., Norris, P. & Green, J.A. (2019). Lay perceptions and illness experiences of people with type 2 diabetes in Indonesia: a qualitative study, *Health Psychology and Behavioral Medicine*, 8(1), 1–15.
- Yu, L., & Zhou, N. (2021). *Survey of Imbalanced Data Methodologies*. <http://arxiv.org/abs/2104.02240>
- Zamzami, F. N., Adiwijaya, A., & P, M. D. (2021). Analisis Sentimen Terhadap Review Film Menggunakan Metode Modified Balanced Random Forest dan Mutual Information. *Jurnal Media Informatika Budidarma*, 5(2), 415–421. <https://doi.org/10.30865/mib.v5i2.2835>