

DAFTAR PUSTAKA

- Abdulkabir, m., Edem, u. A., tunde, r. S., & Kemi, b. L. (2015). *Applied & computational mathematics an empirical study of generalized linear model for count data*. 4(5), 4–6. <https://doi.org/10.4172/2168-9679.1000253>
- Admassu, t., Olalekan, a., & Sampath, k. (2024). Evaluation of adaptive synthetic resampling technique for imbalanced breast cancer identification. *Procedia computer science*, 235(2023), 1000–1007. <https://doi.org/10.1016/j.procs.2024.04.095>
- Agresti, a. (2007). *An introduction to categorical data analysis second edition* (2 ed.). Simultaneously.
- Agresti, a. (2018). *Statistical methods for the social sciences*.
- Agustina, e. B., Lestiyanti, y., Tachriri, y. I., Ulfiya, n., & fisika, p. S. (2025). *Analisis pengaruh limbah cair terhadap kualitas air sungai di kota pekalongan*. 4(2), 230–239. <https://doi.org/10.55123/insologi.v4i2.5117>
- Ahmad, a., Kallista, m., & Wibawa, p. (2024). *Klasifikasi kualitas air sungai citarum dengan artificial neural network , gaussian naive bayes , dan weighted k-nearest neighbors*. 11(4), 3156–3161.
- Ahmed, s. (2024). *A comprehensive introduction to generalized linear models*. Medium. <https://medium.com/@sahin.samia/a-comprehensive-introduction-to-generalized-linear-models-fd773d460c1d>
- Alfatihah, a., Latuconsina, h., & Prasetyo, h. D. (2022). Analisis kualitas air berdasarkan parameter fisika dan kimia di perairan sungai patrean kabupaten sumenep water quality analysis based on physical and chemical parameters in patrean river waters , sumenep regency. 1(2), 76–84. <https://doi.org/10.32734/jafs.v1i2.9174>
- Alfinsyah, d. R., & Hartato, b. P. (2025). *Evaluating the impact of random over sampling on indobert performance for indonesian sentiment analysis*. 9(6), 3270–3282.
- Andalas, u. (2021). *Pemodelan faktor-faktor yang mempengaruhi status daerah di indonesia*. 10(1), 108–115.
- Annahdli, z. F., & Pramudya, e. R. (2025). *Evaluasi model random forest dengan teknik smote dan rus untuk klasifikasi kerusakan motor di bengkel plavix ilmu komputer , teknik informatika , universitas dian nuswantoro , semarang , indonesia performance of random forest model in motorcycle malfuncti*. 5(8), 2295–2310.
- Astuti, f. D., & Lent, f. N. (2021). *Implementasi smote untuk mengatasi imbalance class pada klasifikasi car evolution menggunakan k-nn femi*. 89–98.
- Bhaskoro, r. G. E., & Ramadhan, t. E. (2018). Evaluasi kinerja instalasi pengolahan air minum (ipam) karangpilang i pdam surya sembada kota surabaya secara kuantitatif. *Jurnal presipitasi : media komunikasi dan pengembangan teknik lingkungan*, 15(2), 62. <https://doi.org/10.14710/presipitasi.v15i2.62-68>
- Bisri, a. (2015). *Penerapan adaboost untuk penyelesaian ketidakseimbangan kelas pada penentuan kelulusan mahasiswa dengan metode decision tree*. 1(1).

- Chawla, n. V., Bowyer, k. W., hall, l. O., & Kegelmeyer, w. P. (2002a). Deep synthetic minority over-sampling technique. *Artificial intelligence research*, 16, 321–357. [Http://arxiv.org/abs/2003.09788](http://arxiv.org/abs/2003.09788)
- Chawla, n. V., Bowyer, k. W., hall, l. O., & Kegelmeyer, w. P. (2002b). Deep synthetic minority over-sampling technique. *Artificial intelligence research*, 16, 321–357.
- Diah, r., & Setyowati, n. (2015). *Status kualitas air das cisanggarung , jawa barat*. 1, 37–45.
- Dobson, a. J., & Barnett, a. G. (2008). *An introduction to generalized linear models* (4 ed.).
- Dzakiroh, h. I. (2025). *Studi kualitas air sungai pada aliran sungai cileungsi di kabupaten bogor dan potensi pemanfaatan budidaya ikan air tawar skripsi*. Universitas islam negeri syarif hidayatullah jakarta.
- Elreedy, d., atiya, a. F., & Kamalov, f. (2024). A theoretical distribution analysis of synthetic minority oversampling technique (smote) for imbalanced learning. *Machine learning*, 113(7), 4903–4923. [Https://doi.org/10.1007/s10994-022-06296-4](https://doi.org/10.1007/s10994-022-06296-4)
- Estri, n., Damayanti, t., & Sila amali, m. (2023). *Dampak pelanggaran asumsi klasik terhadap estimasi model ekonometrika*. 2(1), 150–157.
- Firmansyah, y. W., setiani, o., & Darundiati, y. H. (2021). Kondisi sungai di indonesia ditinjau dari daya tampung beban pencemaran: studi literatur. *Jurnal serambi engineering*, 6(2), 1879–1890. [Https://doi.org/10.32672/jse.v6i2.2889](https://doi.org/10.32672/jse.v6i2.2889)
- Gusti, s. K., Afrianty, i., & Syafria, f. (2025). *Penerapan metode adasyn dalam mengatasi imbalanced data untuk klasifikasi penyakit stroke menggunakan support vector machine*. 5(4), 532–541. [Https://doi.org/10.47065/bulletincsr.v5i4.612](https://doi.org/10.47065/bulletincsr.v5i4.612)
- He, h., Bai, y., Garcia, e. A., & li, s. (2008). Adasyn: adaptive synthetic sampling approach for imbalanced learning. *Proceedings of the international joint conference on neural networks*, (3), 1322–1328. [Https://doi.org/10.1109/ijcnn.2008.4633969](https://doi.org/10.1109/ijcnn.2008.4633969)
- Hiola, y. P., Maulina, m., & Alamudi, a. (2025). *Performance evaluation of multinomial logistic regression , random forest , and xgboost methods in data classification*. 8(2), 355–373.
- Hosmer, d. W., & Lemeshow, s. (2013). *Applied logistic regression* (3rd ed.).
- Inayah, u. R., suyitno, & Siringoringo, m. (2021). *Upaya pencegahan pencemaran air sungai mahakam melalui pemodelan geographically weighted logistic regression pada data bod*. 12, 17–26.
- Indransyah, r., Chrisnanto, y. H., Sabrina, p. N., Informatika, p. S., jenderal, u., Yani, a., Abdillah, n., & sentimen, k. (2022). *Klasifikasi sentimen pergelaran motogp di indonesia menggunakan algoritma correlated naïve bayes clasifier*. 60–66.
- Irawan, h. (2025). *Deteksi fraud kartu kredit dengan logistic regression, random forest, dan gradient boosting*. 11(2), 92–97.
- Jiawei, h., Micheline, k., & Jian, p. (2011). *Data mining: concepts and techniques* (3 ed.). Elsevier.

- Kannan, s., Dongare, p. A., Garg, r., & Harsoor, s. S. (2019). *Describing and displaying numerical and categorical data*. 64–70. <https://doi.org/10.4103/arwy.arwy>
- Kartikasari, d. (2020). Analisis faktor-faktor yang mempengaruhi level polusi udara dengan metode regresi logistik biner. *Jurnal ilmiah matematika*, 8(1), 56–59.
- Kismiantini. (2007). Analisis peubah respon biner. *Pythagoras: jurnal matematika dan pendidikan matematika*, 3(1), 63–77. https://doi.org/https://jurnal.uny.ac.id/index.php/pythagoras/article/view/643?utm_source=chatgpt.com
- Klhk. (2017). *Petunjuk teknis restorasi kualitas air sungai*. (021), 1–106.
- Li, l. (2015). *Experimental comparisons of multi-class classifiers*. 39, 71–85.
- Li, l., Rysavy, m. A., Bobashev, g., & Das, a. (2024). *Comparing methods for risk prediction of multicategory outcomes : dichotomized logistic regression vs . Multinomial logit regression*.
- Maharani, m., Rosdiana, r., & Ndibale, w. (2023). *Analisis kualitas air sungai lasolo di kecamatan kendari barat kota kendari*. 01(10), 27–34.
- Mccullagh, p., & Frs, j. A. N. (1989). *Generalized linear models* (2 ed.).
- Mestika, w. M. (2024). *Hubungan parameter kualitas air dengan kelimpahan zooplankton di waduk pusong kota lhokseumawe*. Universitas malikussaleh aceh utara.
- Mishra, p., Pandey, c., Singh, u., & Gupta, a. (2018). *Scales of measurement and presentation of statistical data*. 419–422. <https://doi.org/10.4103/aca.aca>
- Muller, m. (2004). *Generalized linear models* *. I(1989), 1–24.
- Muniz, d. H. F., & Oliveira-filho, e. C. (2023). *Multivariate statistical analysis for water quality assessment : a review of research published between 2001 and 2020*.
- Novianti, & Rizki, s. W. (2023). *Analisis regresi logistik multinomial pada pemilihan alat kontrasepsi wanita (studi kasus di puskesmas sungai kakap)*. 08(4), 751–758.
- Nugraha, i. N., dos, d., Pinto, s., & Hanafi, m. (2025). *Optimasi deteksi penipuan kartu kredit menggunakan regresi logistik dengan particle swarm optimization*. 4(2).
- Nurdian, r. A., Ridwan, m., & Yusuf, a. (2022). *Komparasi metode smote dan adasyn dalam meningkatkan performa klasifikasi herregistrasi mahasiswa baru*. 8(april), 24–32.
- Nurhalizah, r. S., & Ardianto, r. (2024). *Analisis supervised dan unsupervised learning pada machine learning : systematic literature review*. 4(1), 61–72.
- Nurrizqi, a. I., Fitrianto, a., & Amelia, r. (2022). *Pemodelan regresi logistik berbasis backward elimination untuk mengetahui faktor yang memengaruhi tingkat kemiskinan di indonesia tahun 2021*. 6(2), 160–170.
- Pal, a. (2021). Logistic regression: a simple primer. *Cancer research, statistics, and treatment*, 4(3), 551–554. https://doi.org/10.4103/crst.crst_164_21
- Pamuji, f. Y., Dwi, s., & Putri, a. (2021). *Komparasi metode smote dan adasyn untuk penanganan data tidak seimbang multiclass*. 331–338.

- Pohan, a. B., Kurniasih, a., Informatika, b. S., & pusat, j. (2023). *Penerapan teknik smote dalam memprediksi kegagalan mesin menggunakan support vector machine dan*. 7(2), 181–187.
- Prasetia, a., & Walukow, a. F. (2021). *Analisis mutu air danau area gelanggang expo dengan metode indeks pencemaran di kota jayapura*. 8, 42–47. <https://doi.org/10.31258/dli.8.1.p.42-47>
- Purwa, n. L. G. A. D., Susilawati, m., Gautama, i., Suciptawati, n. L. P., kencana, e. N., & jayanegara, k. (2025). *Regresi logistik multinomial terhadap faktor pemanfaatan jaminan kesehatan pasien rawat jalan provinsi bali 2023*. 6(12).
- Rahmadani, a. A., Putri, a. A., happy, d. M., alensia, m., dala, d., angka, m. T. D., & rafi, m. (2023). *Analisis regresi logistik biner untuk memprediksi faktor-faktor internal yang memengaruhi keharmonisan rumah tangga menurut provinsi di indonesia pada tahun 2021*. 116–127.
- Rasyid, r., Nasrul, & Qaiyimah, d. (2025). *Analisis indeks pencemaran dan prediksi distribusi kualitas air sungai menggunakan metode spatial inverse distance weighted (studi kasus : sungai gentung , kabupaten pangkep)*. 23(5), 1213–1226. <https://doi.org/10.14710/jil.23.5.1213-1226>
- Ratnaningsih, d., Puji, r., Nazir, e., & Fauz, r. (2018). *Pengembangan indeks kualitas air sebagai alternatif penilaian kualitas air sungai the development of water quality index as an indeks kualitas air dikembangkan dengan konsep bahwa nilai indeks yang semakin tinggi*. 53–61.
- Retno, d., Saputro, s., & Widyaningsih, p. (2016). *Algoritme pendugaan parameter model regresi logistik biner (rlb) dengan maksimum likelihood dan broyden-fletcher-goldfarb-shanno (bfgs)*. (sr 1), 97–104.
- Rilvandri, q. P., nur, i. M., Damiri, b. A., & Raditir, f. (2023). *Pemodelan mutu kualitas air sungai di daerah istimewa yogyakarta tahun 2020 dengan pendekatan regresi logistik ordinal*. 1(2), 62–70.
- Ruiz, j. S., lópez, o. A. M., Ramírez, g. H., & hiriart, j. C. (2023). *Generalized linear mixed models with applications in agriculture and biology. Dalam generalized linear mixed models with applications in agriculture and biology*. <https://doi.org/10.1007/978-3-031-32800-8>
- Santi, v. M. (2018). *Pengembangan model regresi logistik multinomial untuk klasifikasi politik pada pemilihan umum*. 1(1), 37–43.
- Santi, V. M., dkk. (2018). *Klasifikasi diagnosis penyakit kanker payudara dengan pendekatan regresi logistik biner dan metode classification and regression tree (CART)*. *JMT: Jurnal Matematika dan Terapan*.
- Santi, v. M., wiyono, a., & Sudarwanto. (2021a). *Pemodelan jumlah kasus malaria di indonesia menggunakan generalized linear model*. 5(1), 112–120.
- Santi, V. M., Indriana, D., Ladayya, F., & Tonah. (2025). *Linear mixed model-LASSO with MLE and REML estimation on poverty data in Java Island*. *Jurnal Statistika dan Aplikasinya*, 9(2), 14–24.
- Santi, v. M., Wiyono, a., & Sudarwanto. (2021b). *Pemodelan jumlah kasus malaria di indonesia menggunakan generalized linear model*. 5(1), 112–120.
- Santi, V. M., Ridana, F. F., & Sumargo, B. (2025). *Bivariate binary logistic regression analysis for modeling educational level and employment status in*

- Central Java. InPrime: Indonesian Journal of Pure and Applied Mathematics*, 7(2), 181–192. <https://doi.org/10.15408/inprime.v7i2.46464>
- Santi, V. M., (2026). Analyzing severe pneumonia factors in toddlers using LASSO penalized binary logistic regression.
- Sari, n. N. (2021). *Tesis klasifikasi multilabel imbalanced data menggunakan hybrid sampling random forest (analisis data perilaku berisiko remaja jawa timur 2019)*.
- Sipahutar, l., Tri budi prayogo, & Emma yuliani. (2025). Investigasi tingkat pencemaran daerah aliran sungai brantas akibat aktivitas agrikultur serta perubahan musim di indonesia. *Jurnal teknologi dan rekayasa sumber daya air*, 5(1), 402–412. <https://doi.org/10.21776/ub.jtresda.2025.005.01.039>
- Smith, r. L. (2019). *Handout on exponential families and glms background : two formulas for likelihood functions*. 1–7.
- Sobari, m., Putri, d. I., Prama, d. R., & Suparman, y. (2024). *Multinomial logistic regression to determine factors influencing the selection of health care facilities in*. 7(2), 187–198.
- Sperandei, s. (2014a). Understanding logistic regression analysis. *Biochemia medica*, 24(1), 12–18. <https://doi.org/10.11613/bm.2014.003>
- Sperandei, s. (2014b). Understanding logistic regression analysis. *Biochemia medica*, 24(1), 12–18. <https://doi.org/10.11613/bm.2014.003>
- Sumargo, B., Handayani, D., Lubis, A. P., Firmasyah, I., & Wulansari, I. Y. (2025). Detection of factors affecting rainfall intensity in Jakarta. *Jurnal Ilmu Lingkungan*, 23(1), 133–140. <https://doi.org/10.14710/jil.23.1.133-140>
- Syahrani, r., Suhartono, & Zaman, s. (2023). *Prediksi kategori kelulusan mahasiswa menggunakan metode regresi logistik multinomial*. 8(2), 102–111.
- Tay, j. K., way, j. S., way, j. S., & way, j. S. (2023). *Elastic net regularization paths for all generalized linear models*. 1–24. <https://doi.org/10.18637/jss.v106.i01.elastic>
- Tokandari, a. F., Lestari, t. A., azi, m., & ghani, z. (2025). *Comparison of oversampling , undersampling , and smote techniques for multiclass balance data handling in random forest and multinomial logistic regression*. 9(2), 69–81.
- Unep. (2022). *Ecosystems and water quality*. United nations environment programme. <https://www.unep.org/interactives/wwqa/technical-highlights/ecosystems-and-water-quality?>
- United nations. (2015). *Sustainable development goal 6: ensure availability and sustainable management of water and sanitation for all*. <https://sdgs.un.org/goals/goal6?>
- Varsheni, s. (2025). *Adasyn: adaptive synthetic sampling for imbalanced datasets*. Train in data. <https://www.blog.trainindata.com/adasyn-adaptive-synthetic-sampling-for-imbalanced-datasets/>
- Yogafanny, e. (2015). *Pengaruh aktifitas warga di sempadan sungai terhadap kualitas air sungai winongo*. 7, 41–50.
- Zakariah, m., Alqahtani, s. A., & Al-rakhami, m. S. (2023). Machine learning-based adaptive synthetic sampling technique for intrusion detection. *Applied sciences (switzerland)*, 13(11). <https://doi.org/10.3390/app13116504>