

DAFTAR PUSTAKA

- Amiroch, S., Laili, F. N., Rohmah, A. M., & Kardono, D. Y. (2026). Coronary Heart Disease Risk Prediction under Class Imbalance Using XGBoost with SHAP-Based Interpretation. *CAUCHY: Jurnal Matematika Murni Dan Aplikasi*, *11*(1), 937–955. <https://doi.org/10.18860/cauchy.v11i1.37938>
- Anggo, B., Aji, S., Setiawan, Y., Anggraini, S. D., Teknologi, J., Kampus, I. (, Surabaya, K., & Surabaya, K. (2025). *Analisis Perbandingan Algoritma Decision Tree, Random Forest, dan XGBoost untuk Klasifikasi Penyakit Infeksi Gigi dan Mulut*.
- Asemi, A., Mustafa, M. B., & Huang, H. (2023). Sentiment Analysis in E-Commerce Platforms: A Review of Current Techniques and Future Directions. *IEEE Access*, *11*, 90367–90382. <https://doi.org/10.1109/ACCESS.2023.3307308>
- Azra, S., & Shita, R. T. (2023). Analisis Sentimen Menggunakan Metode Naïve Bayes Terhadap Produk Pt. Imin Technology Berdasarkan Ulasan dari Tokopedia. *Prosiding Seminar Nasional Mahasiswa Fakultas Teknologi Informasi (SENAFTI)*, *2*(2), 445–452. <https://senafti.budiluhur.ac.id/index.php/senafti/article/view/753>
- Baloch, L., Aslam, S., Aftab, F., Marjan, S., Bazai, S., Amphawan, A., & Neo, T. (2023). A Comprehensive Survey on Sentiment Analysis Techniques. *International Journal of Technology*. <https://doi.org/10.14716/ijtech.v14i6.6632>
- Bayu Nur Setiyono, Nadia Annisa Maori, & Teguh Tamrin. (2025). Analisis Sentimen Ulasan Pengguna Aplikasi Threads di Google Play Menggunakan Algoritma XGBoost Dengan Penguatan SMOTE. *Jurnal Informatika Dan Tekonologi Komputer (JITEK)*, *5*(2), 119–130. <https://doi.org/10.55606/jitek.v5i2.6832>
- Bordoloi, M., & Biswas, S. (2023). Sentiment analysis: A survey on design framework, applications and future scopes. *Artificial Intelligence Review*, 1–56. <https://doi.org/10.1007/s10462-023-10442-2>
- BPS. (2024). *statistik-telekomunikasi-indonesia-2023*.
- Brown, M. A., Gruen, A., Maldoff, G., Messing, S., Sanderson, Z., & Zimmer, M. (2025). *Web Scraping for research: Legal, ethical, institutional, and scientific considerations*.

Big Data and Society, 12(4). <https://doi.org/10.1177/20539517251381686>

Budaya, I. G. B. A., & Suniantara, I. K. P. (2024). Comparison of Sentiment Analysis Algorithms with SMOTE Oversampling and TF-IDF Implementation on Google Reviews for Public Health Centers. *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(3), 1077–1086. <https://doi.org/10.57152/malcom.v4i3.1459>

Cai, L. J., Lv, S., & Shi, K. B. (2021). Application of an Improved CHI Feature Selection Algorithm. *Discrete Dynamics in Nature and Society*, 2021. <https://doi.org/10.1155/2021/9963382>

Chai, C. P. (2023). Comparison of text preprocessing methods. *Natural Language Engineering*, 29(3), 509–553. [https://doi.org/DOI: 10.1017/S1351324922000213](https://doi.org/DOI:10.1017/S1351324922000213)

Chandrasekaran, G., Dhanasekaran, S., Moorthy, C., & Arul Oli, A. (2025). Multimodal sentiment analysis leveraging the strength of deep neural networks enhanced by the XGBoost classifier. *Computer Methods in Biomechanics and Biomedical Engineering*, 28(6), 777–799. <https://doi.org/10.1080/10255842.2024.2313066>

Chandrasekaran, G., S., D., C., M., & and Arul Oli, A. (2025). Multimodal sentiment analysis leveraging the strength of deep neural networks enhanced by the XGBoost classifier. *Computer Methods in Biomechanics and Biomedical Engineering*, 28(6), 777–799. <https://doi.org/10.1080/10255842.2024.2313066>

Chen Sami, D., Sugiharto, A., & Jie, F. (2024). CHI SQUARE FEATURE SELECTION FOR IMPROVING SENTIMENT ANALYSIS OF NEWS DATA PRIVACY TREATS. *Journal of Theoretical and Applied Information Technology*, 102(18). www.jatit.org

communications today. (2025, March). *Indonesia's PC market sees 4.1M shipments in 2024*. <https://www.communicationstoday.co.in/indonesias-pc-market-sees-4-1m-shipments-in-2024/>

DHIKA MALITA PUSPITA ARUM, KARTIKA IMAM SANTOSO, ANDRI TRIYONO, EKO SUPRIYADI, AGUS SUSILO NUGROHO, & Widodo, E. (2025). ALGORITMA RANDOM FOREST, DECISION TREE, DAN XGBOOST UNTUK

- KLASIFIKASI STUNTING PADA BALITA. *Jurnal Transformatika*, 23(1), 67–76.
<https://doi.org/10.26623/transformatika.v23i1.12202>
- 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>
- Ernawati, A., Sari, A. O., Sofyan, S. N., Iqbal, M., & Wijaya, R. F. W. (2023). Implementasi Algoritma Naïve Bayes dalam Menganalisis Sentimen Review Pengguna Tokopedia pada Produk Kesehatan. *Bulletin of Information Technology (BIT)*, 4(4), 533–543. <https://journal.fkpt.org/index.php/BIT/article/view/1090>
- Farias, W. A. S., Melo, D. M. A., Santos, L. M. dos, de Oliveira, Â. A. S., Medeiros, R. L. B. A., & Silva, Y. K. R. O. (2024). *Web Scraping as a scientific tool for theoretical reference*. <https://doi.org/10.21203/rs.3.rs-3854342/v1>
- Ghatora, P. S., Hosseini, S. E., Pervez, S., Iqbal, M. J., & Shaukat, N. (2024). Sentiment Analysis of Product Reviews Using Machine Learning and Pre-Trained LLM. *Big Data and Cognitive Computing*, 8(12). <https://doi.org/10.3390/bdcc8120199>
- Hardika, B., Kurniawan, M. D., Adzka, M., Prastowiyono, D., Banyubasa, A., Wicaksono, A., & Nasir, M. (2024). *Pengujian Blackbox Testing Website Garuda Farm Menggunakan Teknik Equivalence Partitioning*.
- Hasan Fahreza, M., Sipayung, N. Z., Santoso, I. W., Susilowati, S., & Kunci, K. (2026). *KOMPARASI SVM-LR BERBASIS INDOBERT DAN SMOTE-TOMEK PADA DETEKSI BUZZER*.
- Hemmatian, J., Hajizadeh, R., & Nazari, F. (2025). Addressing imbalanced data classification with Cluster-Based Reduced Noise SMOTE. *PLoS ONE*, 20(2 February). <https://doi.org/10.1371/journal.pone.0317396>
- Hendrawan, I. R., Utami, E., & Hartanto, A. D. (2022). Comparison of naïve bayes algorithm and XGBoost on local product review text classification. *Edumatic: Jurnal Pendidikan Informatika*, 6(1), 143–149.
- Hossain, R., & Timmer, D. (2021). *Machine Learning Model Optimization with Hyper Parameter Tuning Approach*.

- Hssane, A., Kasri, M., & Birjali, M. (2021). A comprehensive survey on sentiment analysis: Approaches, challenges and trends. *Knowl. Based Syst.*, 226, 107134. <https://doi.org/10.1016/j.knosys.2021.107134>
- Huang, C. A., & Lam, T. (2025). *PRITES: An integrative framework for investigating and assessing web-scraped HTTP-response datasets for research applications*. <http://arxiv.org/abs/2511.13773>
- Huchaiah, M., & Belete, D. M. (2021). Grid search in hyperparameter optimization of machine learning models for prediction of HIV/AIDS test results. *International Journal of Computers and Applications*, 44, 875–886. <https://doi.org/10.1080/1206212X.2021.1974663>
- IDC. (2024, April 8). *Global PC Shipments Return to Growth and Pre-Pandemic Volumes in the First Quarter of 2024, According to IDC Tracker*. IDC.
- Ippolito, P. P. (2022). Hyperparameter Tuning. *Applied Data Science in Tourism*. https://doi.org/10.1007/978-3-030-88389-8_12
- Irawan, D., Ruliyanti, A. D., Zulfikar, M. F., Fu'adhi, M. Z., Rozi, Moch. F., Kholifah, M., & Thoib, I. (2025). Functional Testing of Nganjuk Runners Website Using Black Box Testing Method. *Jurnal Ilmiah Sistem Informasi*, 4(3), 959–972. <https://doi.org/10.51903/ht6zeq23>
- Khder, M. A. (2021). Web Scraping or web crawling: State of art, techniques, approaches and application. *International Journal of Advances in Soft Computing & Its Applications*, 13(3). <https://ijasca.zuj.edu.jo/PapersUploaded/2021.3.11.pdf>
- Komang Dharmendra, I., Made, I., Putra, A. W., & Atmojo, Y. P. (2024). Evaluasi Efektivitas SMOTE dan Random Under Sampling pada Klasifikasi Emosi Tweet. *Informatics for Educators And Professionals : Journal of Informatics*, 9(2), 192–193.
- Kumar, C. N., Hussain, S. S., Shoaib, S., Tumuluru, P., Madhu, D., & Kankanala, L. (2023). Advancing Twitter Sentiment Analysis: An Ensemble Approach with Transformer-XL, RoBERTa, and XGBoost. *2023 International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS)*, 944–950. <https://doi.org/10.1109/ICSSAS57918.2023.10331828>

- Lan, F. (2022). Research on Text Similarity Measurement Hybrid Algorithm with Term Semantic Information and TF-IDF Method. *Advances in Multimedia*, 2022(1), 7923262.
- Lee, C., Lim, K., & Tan, K. L. (2023). A Survey of Sentiment Analysis: Approaches, Datasets, and Future Research. *Applied Sciences*. <https://doi.org/10.3390/app13074550>
- Liao, L., Li, H., & Shang, W. (2022). An Empirical Study of the Impact of Hyperparameter Tuning and Model Optimization on the Performance Properties of Deep Neural Networks. *ACM Transactions on Software Engineering and Methodology (TOSEM)*, 31, 1–40. <https://doi.org/10.1145/3506695>
- Mayang Pratiwi, D., & Mufidah, L. (2024). *Perbandingan Metode Decision Tree Classifier dan XGBoost Classifier Dalam Memprediksi Penyakit Jantung* (Vol. 4, Number 1).
- Meidea, Z. A. I., Dewi, R. S., & Prabawani, B. (2024). Pengaruh Kualitas Produk Dan Persepsi Harga Terhadap Brand Switching Melalui Ketidakpuasan Konsumen Sebagai Variabel Intervening (Studi Pada Brand Switching Laptop Merek Asus Ke Merek Lain Di Kota Semarang). *Jurnal Ilmu Administrasi Bisnis*, 13(1), 186–198. <https://ejournal3.undip.ac.id/index.php/jiab/article/view/40567>
- Mujahid, M., Kina, E. R. O. L., Rustam, F., Villar, M. G., Alvarado, E. S., De La Torre Diez, I., & Ashraf, I. (2024). Data oversampling and imbalanced datasets: an investigation of performance for machine learning and feature engineering. *Journal of Big Data*, 11(1). <https://doi.org/10.1186/s40537-024-00943-4>
- Naz, S., Pamučar, D., Tasneem, R., Shafiq, A., Gonzalez, Z. C., & Butt, S. (2025). Decision-making model for selecting products through online product reviews utilizing natural language processing techniques. *Neurocomputing*, 611, 128593. <https://doi.org/10.1016/j.neucom.2024.128593>
- Nguyen, N., & Ngo, D. (2025). Comparative analysis of boosting algorithms for predicting personal default. *Cogent Economics & Finance*, 13(1), 2465971. <https://doi.org/10.1080/23322039.2025.2465971>
- Nisha Arya Ahmed. (2024, November 10). *What is A Confusion Matrix in Machine Learning? The Model Evaluation Tool Explained*. Datacamp.

- <https://www.datacamp.com/tutorial/what-is-a-confusion-matrix-in-machine-learning>
- Normawati, D., & Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. In *Jurnal Sains Komputer & Informatika (J-SAKTI)* (Vol. 5, Number 2).
- Nur Ichsanudin, M., Yusuf, M., Jurusan Rekayasa Sistem Komputer, S., Teknik Industri, J., AKPRIND Yogyakarta, I., & Artikel, R. (2022). *PENGUJIAN FUNGSIONAL PERANGKAT LUNAK SISTEM INFORMASI PERPUSTAKAAN DENGAN METODE BLACK BOX TESTING BAGI PEMULA INFO ARTIKEL ABSTRAK*. 1(2), 1–8. <https://doi.org/10.55123>
- Ogunsanya, M., Desai, S., & Isichei, J. (2023). Grid search hyperparameter tuning in additive manufacturing processes. *Manufacturing Letters*. <https://doi.org/10.1016/j.mfglet.2023.08.056>
- Rajput, D., Wang, W. J., & Chen, C. C. (2023). Evaluation of a decided sample size in machine learning applications. *BMC Bioinformatics*, 24(1). <https://doi.org/10.1186/s12859-023-05156-9>
- Ranjan, M., Sattar, A. Md., Tatkar, N., & Tiwari, S. (2024). A New Approach for Carrying Out Sentiment Analysis of Social Media Comments Using Natural Language Processing. *RAiSE-2023*. <https://doi.org/10.3390/engproc2023059181>
- Ridwan Yulianto, P., & Parti Astuti, Y. (2026). Perbandingan Kinerja Algoritma CatBoost, XGBoost, LightGBM dan Random Forest Dalam Memprediksi Risiko Infeksi Aids Dalam Dataset Kesehatan. *Technology and Science (BITS)*, 7(4). <https://doi.org/10.47065/bits.v7i4.8975>
- Rodríguez-Ibáñez, M., Casañez-Ventura, A., Castejón-Mateos, F., & Cuenca-Jiménez, P. (2023). A review on sentiment analysis from social media platforms. *Expert Syst. Appl.*, 223, 119862. <https://doi.org/10.1016/j.eswa.2023.119862>
- Sahyudi, S., & Voutama, A. (2025). Pengujian Fungsional Black Box SISKAS UNSIKA dengan Equivalence Partitioning untuk Validasi Input dan Output Sistem. *Volume*, 8, 148–155. <https://ojs.trigunadharma.ac.id/index.php/jsk/index>
- Salsabila, N. A., Sa'adah, U., & Fauzi, F. (2024). Analisis Sentimen Pada Ulasan Aplikasi

- Tokopedia Menggunakan Klasifikasi Naïve Bayes. *PRISMA, Prosiding Seminar Nasional Matematika*, 44–51. <https://proceeding.unnes.ac.id/prisma/article/view/2931>
- Samih, A., Ghadi, A., & Fennan, A. (2023). Enhanced sentiment analysis based on improved word embeddings and XGboost. *International Journal of Electrical and Computer Engineering*, 13(2), 1827–1836. <https://doi.org/10.11591/ijece.v13i2.pp1827-1836>
- Sarathan, I., Yulita, I., & Afifah, K. (2021). Sentiment Analysis on Telemedicine App Reviews using XGBoost Classifier. *2021 International Conference on Artificial Intelligence and Big Data Analytics*, 22–27. <https://doi.org/10.1109/ICAIBDA53487.2021.9689762>
- Sasongko, B. B., Malik, F., Ardiansyah, F., Rahmawati, A. F., Dharma Adhinata, F., & Rakhmadani, D. P. (2021). Pengujian Blackbox Menggunakan Teknik Equivalence Partitions pada Aplikasi Petgram Mobile. In *Jurnal ICTEE* (Vol. 2, Number 1).
- Satria, B. C., & Suryandari, R. T. (2025). Cognitive Dissonance in Online Shopping Experiences and Impulsive Buying Among E-Commerce Users in Indonesia. *International Journal of Current Science Research and Review*, 08(10). <https://doi.org/10.47191/ijcsrr/V8-i10-08>
- Sediatmoko, N. S., Nataliani, Y., & Suryady, I. (2024). Suryady (Sentiment Analysis of Customer Review Using Classification Algorithms and SMOTE for Handling Imbalanced Class). In *Indonesian Journal of Information Systems (IJIS)* (Vol. 7, Number 1).
- Semary, N. A., Ahmed, W., Amin, K., Pławiak, P., & Hammad, M. (2023). Improving sentiment classification using a RoBERTa-based hybrid model. *Frontiers in Human Neuroscience*, 17. <https://doi.org/10.3389/fnhum.2023.1292010>
- Sharma, S. (2023). *Managing Product Reviews: A Comprehensive Guide for Brands and Businesses*. <https://doi.org/10.46679/9788195732265>
- Siagian, N. A., Sipayung, S. P., Rikki, A., & Marbun, N. (2025). Integrating SMOTE with XGBoost for Robust Classification on Imbalanced Datasets: A Dual-Domain Evaluation. *Sinkron*, 9(3), 1094–1107. <https://doi.org/10.33395/sinkron.v9i3.15029>
- Sikri, A., Singh, N. P., & Dalal, S. (2023). International Journal of INTELLIGENT

SYSTEMS AND APPLICATIONS IN ENGINEERING Chi-Square Method of Feature Selection: Impact of Pre-Processing of Data. In *International Journal of Intelligent Systems and Applications in Engineering IJISAE* (Vol. 2023, Number 3s). www.ijisae.org

Sondang, S. (2024). Penerapan metode RAD dalam pengembangan sistem informasi pemesanan jasa percetakan berbasis web pada percetakan Karya Sehati Jaya. *Remik: Riset Dan E-Jurnal Manajemen Informatika Komputer*, 8(3), 871–881.

STAMFORD, Conn. (2025, January 16). *Gartner Says Worldwide PC Shipments Increased 1.4% in Fourth Quarter of 2024 and 1.3% for the Year*. Gartner.

Sun, Y., Ding, S., Zhang, Z., & Jia, W. (2021). An improved grid search algorithm to optimize SVR for prediction. *Soft Computing*, 25, 5633–5644. <https://doi.org/10.1007/s00500-020-05560-w>

Tran, N. T., Tran, T. T. G., Nguyen, T. A., & Lam, M. B. (2023). A new grid search algorithm based on XGBoost model for load forecasting. *Bulletin of Electrical Engineering and Informatics*, 12(4), 1857–1866. <https://doi.org/10.11591/eei.v12i4.5016>

Wang, Z. (2025). Sentiment Analysis of Mobile Phone Reviews Using XGBoost and Word Vectors. *ITM Web of Conferences*. <https://doi.org/10.1051/itmconf/20257003018>

Wicaksono, A. R. (2025). Analisis Kinerja Pengujian Black Box Terhadap Website Nestify Setelah Implementasi Equivalence Partitioning dan Boundary Value Analysis. *Jurnal Informatika Dan Teknologi Pendidikan*, 5(1), 77–98. <https://doi.org/10.59395/jitp.v5i1.131>

Wongvorachan, T., He, S., & Bulut, O. (2023). A Comparison of Undersampling, Oversampling, and SMOTE Methods for Dealing with Imbalanced Classification in Educational Data Mining. *Information (Switzerland)*, 14(1). <https://doi.org/10.3390/info14010054>

xgboost developers. (2025). *Introduction to Boosted Trees*. XGBoost Docs. <https://xgboost.readthedocs.io/en/stable/tutorials/model.html>

Zantvoort, K., Nacke, B., Görlich, D., Hornstein, S., Jacobi, C., & Funk, B. (2024).

Estimation of minimal data sets sizes for machine learning predictions in digital mental health interventions. *Npj Digital Medicine*, 7(1). <https://doi.org/10.1038/s41746-024-01360-w>

Zen, M., Irwan, Hafni, & Ananda, M. D. P. (2024). Implementasi dan Pengujian Menggunakan Metode BlackBox Testing Pada Sistem Informasi Tracer Study. *Bulletin of Computer Science Research*, 4(4), 327–340. <https://doi.org/10.47065/bulletincsr.v4i4.359>

Zhou, J., Ye, Z., Zhang, S., Geng, Z., Han, N., & Yang, T. (2024). Investigating response behavior through TF-IDF and Word2vec text analysis: A case study of PISA 2012 problem-solving process data. *Heliyon*, 10(16). <https://doi.org/10.1016/j.heliyon.2024.e35945>