

DAFTAR PUSTAKA

- Adoe, V. S., Yusefiana, M., Diana, A., Lubis, R., & Harahap, M. (2022). *Buku Ajar E-Commerce*. CV. Feniks Muda Sejahtera.
- Ahdiat, A. (2024). *Rata-rata Jumlah Kunjungan ke 5 Situs E-Commerce Terbesar di Indonesia (Kuartal I-Kuartal IV 2023)*. databoks. <https://databoks.katadata.co.id/teknologi-telekomunikasi/statistik/f056a701023f0fe/tren-e-commerce-2023-kunjungan-ke-shopee-dan-blibli-meningkat>
- Ahmadi, S. (2020). *A Tokenization System for the Kurdish Language*.
- Alshehri, A. H. (2024). An Online Fake Review Detection Approach Using Famous Machine Learning Algorithms. *Computers, Materials & Continua*, 78(2), 2767–2786. <https://doi.org/10.32604/cmc.2023.046838>
- Asaad, W. H., Allami, R., & Ali, Y. H. (2023). Fake Review Detection Using Machine Learning. *Revue d'Intelligence Artificielle*, 37(5). <https://doi.org/10.18280/ria.370507>
- Aydin, A. (2023). *Text Preprocessing Techniques for NLP*.
- Bengio, Y., Lecun, Y., & Hinton, G. (2021). Deep learning for AI. *Communications of the ACM*, 64(7), 58–65. <https://doi.org/10.1145/3448250>
- Beyari, H. (2021). *Recent E-Commerce Trends And Learnings For E-Commerce System Development From A Quality Perspective*. *International Journal for Quality Research*, 15(3), 797–810. <https://doi.org/10.24874/IJQR15.03-07>
- Chefasa, A. (2024, Oktober 7). Pengaruh E-commerce Terhadap Perubahan Pola Konsumsi Masyarakat di Indonesia. *Qubisa*. <https://www.qubisa.com/article/pengaruh-e-commerce-terhadap-perubahan-pola-konsumsi-masyarakat-di-indonesia>
- Cholissodin, I. (2020). *Buku Ajar AI, Machine Learning & Deep Learning. Machine Learning*.

- Çorbacioğlu, Ş. K., & Aksel, G. (2023). Receiver operating characteristic curve analysis in diagnostic accuracy studies: A guide to interpreting the area under the curve value. *Turkish Journal of Emergency Medicine*, 23(4), 195–198. https://doi.org/10.4103/tjem.tjem_182_23
- Côté, P.-O., Nikanjam, A., Ahmed, N., Humeniuk, D., & Khomh, F. (2024). *Data Cleaning and Machine Learning: A Systematic Literature Review* (No. arXiv:2310.01765). arXiv. <https://doi.org/10.48550/arXiv.2310.01765>
- Crall, J. (2023). *The MCC approaches the geometric mean of precision and recall as true negatives approach infinity* (No. arXiv:2305.00594). arXiv. <https://doi.org/10.48550/arXiv.2305.00594>
- Dhewayani, F. N., Amelia, D., Alifah, D. N., Sari, B. N., & Jajuli, M. (2022). Implementasi K-Means Clustering untuk Pengelompokkan Daerah Rawan Bencana Kebakaran Menggunakan Model CRISP-DM. *Jurnal Teknologi dan Informasi*, 12(1), 64–77. <https://doi.org/10.34010/jati.v12i1.6674>
- Eom, G., & Byeon, H. (2023). Searching for Optimal Oversampling to Process Imbalanced Data: Generative Adversarial Networks and Synthetic Minority Over-Sampling Technique. *Mathematics*, 11(16), 3605. <https://doi.org/10.3390/math11163605>
- Fayyad, U. (1996). *From Data Mining to Knowledge Discovery in Databases*.
- Işık, M., & Dağ, H. (2020). The impact of text preprocessing on the prediction of review ratings. *TURKISH JOURNAL OF ELECTRICAL ENGINEERING & COMPUTER SCIENCES*, 28(3), 1405–1421. <https://doi.org/10.3906/elk-1907-46>
- Jana, V. (2015). Online consumer reviews. *European Parliamentary Research Service*, 2–5.
- Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, 31(3), 685–695. <https://doi.org/10.1007/s12525-021-00475-2>
- Jiang, Z., Pan, T., Zhang, C., & Yang, J. (2021). A New Oversampling Method Based on the Classification Contribution Degree. *Symmetry*, 13(2), 194. <https://doi.org/10.3390/sym13020194>

- Jurafsky, D., & H.Marten, J. (2024). NLP Application. Dalam *Speech and Language Processing*.
- Kang, Y., Cai, Z., Tan, C.-W., Huang, Q., & Liu, H. (2020). Natural language processing (NLP) in management research: A literature review. *Journal of Management Analytics*, 7(2), 139–172. <https://doi.org/10.1080/23270012.2020.1756939>
- Karimi, Z. (2021). Confusion Matrix. *researchgate*.
- Kowsari, K., Jafari Meimandi, K., Heidarysafa, M., Mendu, S., Barnes, L., & Brown, D. (2019). Text Classification Algorithms: A Survey. *Information*, 10(4), 150. <https://doi.org/10.3390/info10040150>
- Liu, K., Hu, X., Zhou, H., Tong, L., Widanage, W. D., & Marco, J. (2021). *Feature Analyses and Modelling of Lithium-ion Batteries Manufacturing based on Random Forest Classification* (No. arXiv:2102.06029). arXiv. <https://doi.org/10.48550/arXiv.2102.06029>
- Mahesh, B. (2020). Machine Learning Algorithms—A Review. *International Journal of Science and Research (IJSR)*, 9(1), 381–386. <https://doi.org/10.21275/ART20203995>
- Mohammed, R., Rawashdeh, J., & Abdullah, M. (2020). Machine Learning with Oversampling and Undersampling Techniques: Overview Study and Experimental Results. *2020 11th International Conference on Information and Communication Systems (ICICS)*, 243–248. <https://doi.org/10.1109/icics49469.2020.239556>
- Navisa, S., Luqman Hakim, & Aulia Nabilah. (2021). Komparasi Algoritma Klasifikasi Genre Musik pada Spotify Menggunakan CRISP-DM: Indonesia. *Jurnal Sistem Cerdas*, 4(2), 114–125. <https://doi.org/10.37396/jsc.v4i2.162>
- Nurhakiki, J., & Yahfizham, Y. (2024). *Studi Kepustakaan: Pengenalan 4 Algoritma Pada Pembelajaran Deep Learning Beserta Implikasinya*.
- Parodi, S., Verda, D., Bagnasco, F., & Muselli, M. (2022). The clinical meaning of the area under a receiver operating characteristic curve for the evaluation of the

- performance of disease markers. *Epidemiology and Health*, 44, e2022088. <https://doi.org/10.4178/epih.e2022088>
- Plotnikova, V., Dumas, M., & Milani, F. (2020). Adaptations of data mining methodologies: A systematic literature review. *PeerJ Computer Science*, 6, e267. <https://doi.org/10.7717/peerj-cs.267>
- Proserpio, D., Hollenbeck, B., & He, S. (2020, November 24). How Fake Customer Reviews Do and Don't Work. *Harvard Business Review*. <https://hbr.org/2020/11/how-fake-customer-reviews-do-and-dont-work>
- Radiansyah, E. (2022). Pengaruh Ulasan Palsu Pada Konsumen: Tinjauan Literatur dan Arah Penelitian MASA DEPAN. *KALIANDA HALOK GAGAS*, 5(1), 42–49. <https://doi.org/10.52655/khg.v5i1.47>
- Rifaldi, D., Abdul Fadlil, & Herman. (2023). Teknik Preprocessing Pada Text Mining Menggunakan Data Tweet “Mental Health.” *Decode: Jurnal Pendidikan Teknologi Informasi*, 3(2), 161–171. <https://doi.org/10.51454/decode.v3i2.131>
- Rotondo, A., & Quilligan, F. (2020). Evolution Paths for Knowledge Discovery and Data Mining Process Models. *SN Computer Science*, 1(2). <https://doi.org/10.1007/s42979-020-0117-6>
- Sajid, M., Kumar Sagar, A., Singh, J., Khalaf, O. I., & Prasad, M. (2023). *Intelligent Techniques for Cyber-Physical Systems* (1 ed.). CRC Press. <https://doi.org/10.1201/9781003438588>
- Saputro, T. H., & Hermawan, A. (2021). *The Accuracy Improvement of Text Mining Classification on Hospital Review through The Alteration in The Preprocessing Stage*. 10(4).
- Shoeb, A. A. M., & De Melo, G. (2021). Assessing Emoji Use in Modern Text Processing Tools. *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 1379–1388. <https://doi.org/10.18653/v1/2021.acl-long.110>
- Statista. (2025). *Number of users of e-commerce in Indonesia 2020-2029*. Statista. <https://www.statista.com/forecasts/251635/e-commerce-users-in-indonesia>

Suyaman, D. J. (2022). *The Influence Of Brand Image, Perceived Quality & Price On Purchasing Decisions For Dunkin Donuts Products In Bekasi*.

Tabany, M., & Gueffal, M. (2024). Sentiment Analysis and Fake Amazon Reviews Classification Using SVM Supervised Machine Learning Model. *Journal of Advances in Information Technology*, 15(1), 49–58. <https://doi.org/10.12720/jait.15.1.49-58>

Vajjala, S., Majumder, B., Gupta, A., & Surana, H. (2020). *Practical Natural Language Processing*. O'Reilly Media.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2023). *Attention Is All You Need* (No. arXiv:1706.03762). arXiv. <https://doi.org/10.48550/arXiv.1706.03762>

Wang, G., Shang, G., Pu, P., Li, X., & Peng, H. (2022). Fake Review Identification Methods Based on Multidimensional Feature Engineering. *Mobile Information Systems*, 2022, 1–12. <https://doi.org/10.1155/2022/5229277>

Wijoyo, A., Saputra, A. Y., Ristanti, S., Sya'Ban, S. R., Amalia, M., & Febriansyah, R. (2024). *Pembelajaran Machine Learning*. 3(2).

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., Von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Le Scao, T., Gugger, S., ... Rush, A. (2020). Transformers: State-of-the-Art Natural Language Processing. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, 38–45. <https://doi.org/10.18653/v1/2020.emnlp-demos.6>

Yusuf, M., & Ifada, N. (2021). *E-commerce Konsep & Teknologi*. Media Nusa Creative.

https://books.google.co.id/books?id=bE6eEAAAQBAJ&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false