

DAFTAR PUSTAKA

- Aldabbas, H., Bajahzar, A., Alruily, M., Qureshi, A. A., Amir Latif, R. M., & Farhan, M. (2020). Google Play content scraping and knowledge engineering using natural language processing techniques with the analysis of user reviews. *Journal of Intelligent Systems*, 30(1), 192–208.
- Arjovsky, M., Chintala, S., & Bottou, L. (2017). *Wasserstein generative adversarial networks*. arXiv. <https://arxiv.org/abs/1701.07875>
- Bahdanau, D., Cho, K., & Bengio, Y. (2014). *Neural machine translation by jointly learning to align and translate*. arXiv. <https://arxiv.org/abs/1409.0473>
- Bhimantoros. (2021). *Pre-trained Word2Vec Indonesia* [Dataset]. Kaggle. <https://www.kaggle.com/datasets/bhimantoros/pretrained-word2vec-indonesia>
- BPJS Kesehatan. (2025). *Mobile JKN* [Aplikasi mobile]. Google Play Store. <https://play.google.com/store/apps/details?id=app.bpjs.mobile>
- Chahal, A., & Gulia, P. (2019). Machine learning and deep learning. *International Journal of Innovative Technology and Exploring Engineering*, 8(12), 4910–4914.
- Cooper, N., Chen, L., Dwivedy, S., & Gurari, D. (2025). Logit-Based Losses Limit the Effectiveness of Feature Knowledge Distillation. arXiv preprint arXiv:2511.14981.
- Dani, A. H., Puspaningrum, E. Y., & Mumpuni, R. (2024). Studi performa TF-IDF dan Word2Vec pada analisis sentimen cyberbullying. *Router: Jurnal Teknik Informatika dan Terapan*, 2(2), 94–106.
- Fernandez-Lozano, C., Gestal, M., Munteanu, C. R., Dorado, J., & Pazos, A. (2016). A methodology for the design of experiments in computational intelligence with multiple regression models. *PeerJ*, 4, e2721.
- Foerderer, J. (2023). *Should we trust web-scraped data?* arXiv. <https://arxiv.org/abs/2308.02231>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. Cambridge, MA; London, England: MIT Press.



st-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., & Bengio, Y. (2014). *Generative adversarial networks*. arXiv. <https://arxiv.org/abs/1406.2661>

schmidhuber, J. (1997). Long short-term memory. *Neural* 9(8), 1735–1780.

- Husain, G., Nasef, D., Jose, R., Mayer, J., Bekbolatova, M., Devine, T., & Toma, M. (2025). SMOTE vs. SMOTEENN: A study on the performance of resampling algorithms for addressing class imbalance in regression models. *Algorithms*, 18(1), 37.
- Keskar, N. S., Mudigere, D., Nocedal, J., Smelyanskiy, M., & Tang, P. T. P. (2017). *On large-batch training for deep learning: Generalization gap and sharp minima*. arXiv. <https://arxiv.org/abs/1609.04836>
- Kim, Y. (2014). *Convolutional neural networks for sentence classification*. arXiv. <https://arxiv.org/abs/1408.5882>
- Kingma, D. P., & Ba, J. (2015). *Adam: A method for stochastic optimization*. arXiv. <https://arxiv.org/abs/1412.6980>
- Li, D. C., Shi, Q. S., Lin, Y. S., & Lin, L. S. (2022). A boundary-information-based oversampling approach to improve learning performance for imbalanced datasets. *Entropy*, 24(3), 322.
- Lipton, Z. C., Berkowitz, J., & Elkan, C. (2015). *A critical review of recurrent neural networks for sequence learning*. arXiv. <https://arxiv.org/abs/1506.00019>
- Markoulidakis, I., & Markoulidakis, G. (2024). Probabilistic confusion matrix: A novel method for machine learning algorithm generalized performance analysis. *Technologies*, 12(7), 113.
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient estimation of word representations in vector space*. arXiv. <https://arxiv.org/abs/1301.3781>
- Mukhamediev, R. I., Popova, Y., Kuchin, Y., Zaitseva, E., Kalimoldayev, A., Symagulov, A., ... & Yelis, M. (2022). Review of artificial intelligence and machine learning technologies: Classification, restrictions, opportunities and challenges. *Mathematics*, 10(15), 2552.
- Radford, A., Metz, L., & Chintala, S. (2015). *Unsupervised representation learning with deep convolutional generative adversarial networks*. arXiv. <https://arxiv.org/abs/1511.06434>
- Russell, S. J., & Norvig, P. (2022). *Artificial intelligence: A modern approach* (4th ed.). Harlow, United Kingdom: Pearson.



- P. K., & Chandra, R. (2022). SMOTified-GAN for class
 attern classification problems. *IEEE Access*, 10, 30655–30665.
- 370). Natural language question-answering systems: 1969.
 ons of the ACM, 13(1), 15–30.

- Srivastava, A., Valkov, L., Russell, C., Gutmann, M. U., & Sutton, C. (2017). Veegan: Reducing mode collapse in GANs using implicit variational learning. *Advances in Neural Information Processing Systems*, 30.
- TensorFlow. (2023). *ExponentialDecay learning rate schedule*. https://www.tensorflow.org/api_docs/python/tf/keras/optimizers/schedules/ExponentialDecay
- Weizenbaum, J. (1966). ELIZA—A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, 9(1), 36–45.
- Xue, X., Yu, X. N., Zhou, D. Y., Wang, X., Zhou, Z. B., & Wang, F. Y. (2022). *Computational experiments: Past, present and future*. arXiv. <https://arxiv.org/abs/2202.13690>
- Xu, L., Skoularidou, M., Cuesta-Infante, A., & Veeramachaneni, K. (2019). Modeling tabular data using conditional GAN. *Advances in Neural Information Processing Systems*, 32.
- Yani, D. D. A., Pratiwi, H. S., & Muhandi, H. (2019). Implementasi web scraping untuk pengambilan data pada situs marketplace. *JUSTIN (Jurnal Sistem dan Teknologi Informasi)*, 7(4), 257–262.

