

DAFTAR PUSTAKA

- Aggarwal, C. C., 2015. *Data Mining: The Textbook*. Springer, New York.
- Amoretti, M., Amsler, C., Bonomi, G., Bouchta, A., Bowe, P., Carraro, C., Cesar, C. L., Chaliton, M., Collier, M. J. T., Doser, M., Filippini, V., Fine, F. S., Fontana, A., Fujiwara, M. C., Funakoshi, R., Genova, P., Hangst, J. S., Hayano, R. S., Holzscheiter, M. H., ... Van Der Werf, D. P. 2002. Production and detection of cold antihydrogen atoms. *Nature*, 419(6906), 456–459.
- Bain, L. J. dan Engelhardt, M., 1992. *Introduction to Probability and Mathematical Statistics*. Duxbury Press, California.
- Barus, F. 2020. Anteraja dorong digitalisasi layanan logistik di Indonesia. *Industry.co.id*. <https://www.industry.co.id/read/75149/anteraja-dorong-digitalisasi-layanan-logistik-di-indonesia>, 25 September 2020, diakses 16 September 2024.
- Blei, D. M. 2012. Probabilistic topic models. *Communications of the ACM*, 55(4), 77–84.
- Blei, D. M., & Lafferty, J. D. 2009. *Topic models*. In A. Srivastava & M. Sahami, eds. *Text mining: Classification, clustering, and applications*. Chapman and Hall/CRC.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. 2003. Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- Casella, G., & Berger, R. L. 2002. *Statistical inference* (edisi kedua). Duxbury, Pacific Grove.
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. 2018. *BERT: Pre-training of deep bidirectional transformers for language understanding*. arXiv:1810.04805.
- Ediyanto, Mara, M.N., & Satyahadewi, N. 2013. *Pengklasifikasian Karakteristik Dengan Metode K-Means Cluster Analysis*. *Bimaster: Buletin Ilmiah Matematika, Statistika dan Terapannya*. 2(2). 133-136.
- Feldman, R., & Sanger, J. 2007. *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge University Press, Cambridge.
- George, L., & Sumathy, P. 2023. An integrated clustering and BERT framework for improved topic modeling. *International Journal of Information Technology*, 15(4), 2187–2195.
- Goldberg, D. E., & Holland, J. H. 1988. Genetic algorithms and machine learning. *Machine Learning*, 3(2), 95–99.
- Grootendorst, M. 2022. *BERTopic: Neural topic modeling with a class-based TF-IDF procedure*. arXiv.
- Hirschberg, J., & Manning, C. D. 2015. Advances in natural language processing. *Science*, 349(6245), 261–266.

- Kaelbling, L.P., Littman, M.L. dan Moore, A.W., 1996, Reinforcement Learning: A Survey, *Journal of Artificial Intelligence Research (JAIR)*, 4, 237–285.
- Khadija, M. A., & Nurharjadm, W. 2024. Enhancing Indonesian customer complaint analysis: LDA topic modelling with BERT embeddings. *Sinergi*, 28(1), 153–162.
- Khudzaeva, E., Mintarsih, F., Muharam, A. T., & Wirawan, C. 2018. Application of clustering method in data mining for determining SNMPTN quota invitation UIN Syarif Hidayatullah Jakarta. *Proceedings of the 6th International Conference on Cyber and IT Service Management (CITSM)*, 1–6.
- Khurana, D., Koli, A., Khatter, K., & Singh, S. 2017. Natural language processing: State of the art, current trends and challenges. *arXiv preprint arXiv:1708.05148*.
- Kulkarni, A., & Shivananda, A. 2019. *Natural language processing recipes: Unlocking text data with machine learning and deep learning using Python*. Apress Media LLC
- Matira, Y., & Setiawan, I. 2023. Pemodelan Topik pada Judul Berita Online Detikcom Menggunakan Latent Dirichlet Allocation. *Estimasi: Journal of Statistics and Its Application*, 4(1), 2721–379.
- McInnes, L., Healy, J. dan Melville, J., 2018. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction, arxiv.org/abs/1802.03426.
- Mitchell, R., 2018. *Web Scraping with Python: Data Extraction from the Modern Web*, edisi 2, O'Reilly Media, Sebastopol.
- Paul, P. C., Rahman, M., Begum, A., Ahmed, M. T., Chakraborty, D., & Rahman, M. S. 2025, Combining BERT with LDA: Improved topic modeling in Bengali language. *IAENG International Journal of Computer Science*, 52(2), 383–393.
- Piyush, & Rishabh, 2020, 3 Types of Machine Learning, 22 Juni 2020, *New Tech Dojo*, <https://www.newtechdojo.com/3-types-of-machine-learning/>, diakses 29 April 2025.
- Qaiser, S. dan Ali, R., 2018, Text Mining: Use of TF-IDF to Examine the Relevance of Words to Documents, *International Journal of Computer Applications*, 181(1), 25–29
- Rajendra, R., 2023, Bisnis AnterAja Lesu, Laba Adi Sarana (ASSA) Turun 39 Persen Semester I/2023, *Bisnis.com*, https://market.bisnis.com/read/20230802/192/1680744/bisnis-anteraja-lesu-laba-adi-sarana-assa-turun-39-persen-semester-i2023?utm_source=desktop&utm_medium=search, diakses 26 Agustus 2024.
- Reimers, N. dan Gurevych, I., 2019, Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks, *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, 3980–3990.

- Röder, M., Both, A. dan Hinneburg, A., 2015, Exploring the Space of Topic Coherence Measures, *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining (WSDM 2015)*, 399–408.
- Rosner, F., Hinneburg, A., Röder, M., Nettling, M., dan Both, A., 2014, Evaluating Topic Coherence Measures, *arXiv preprint arXiv:1403.6397*
- Russell, S. dan Norvig, P., 2010, *Artificial Intelligence: A Modern Approach* (edisi 3), Pearson, New Jersey.
- Sahoo, P., 2008, *Probability and Mathematical Statistics*, Department of Mathematics, University of Louisville, Louisville.
- Sari, I. N., 2021, Anteraja, Bisnis Logistik Triputra Group yang Makin Ekspansif, *Katadata.co.id*, 17 September 2021, <https://katadata.co.id/finansial/korporasi/61432d81d51f7/anteraja-bisnis-logistik-triputra-group-yang-makin-ekspansif>, diakses 5 September 2024.
- Soemanagara, D. F., 2022, Analis Nilai Kinerja Anteraja Dongkrak Laba Bersih ASSA, *Bisnis.com*, 20 April 2022, <https://market.bisnis.com/read/20220420/192/1524848/analisis-nilai-kinerja-anteraja-dongkrak-laba-bersih-assa>, diakses 5 September 2024.
- Stewart, G., & Al-Khassaweneh, M. 2022, An Implementation of the HDBSCAN* Clustering Algorithm. *Applied Sciences (Switzerland)*, 12(5).
- Tan, P.-N., Steinbach, M. dan Kumar, V., 2005, *Introduction to Data Mining*, Addison Wesley, Boston
- Vayansky, I., & Kumar, S. A. P. 2020, A review of topic modeling methods. *Information Systems*, 94.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., dan Polosukhin, I., 2017, Attention Is All You Need, *arXiv:1706.03762*
- Wang, W., Wei, F., Dong, L., Bao, H., Yang, N., dan Zhou, M., 2020, MiniLM: Deep Self-Attention Distillation for Task-Agnostic Compression of Pre-Trained Transformers, *Proceedings of the 34th Conference on Neural Information Processing Systems (NeurIPS 2020)*, Vancouver, Kanada, 1–11.
- Wilie, B., Vincentio, K., Winata, G.I., Cahyawijaya, S., Li, X., Lim, Z.Y., Soleman, S., Mahendra, R., Fung, P., Bahar, S., dan Purwarianti, A., 2020, IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding, *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*, Suzhou, China, 843–857.
- Wu, Y., Schuster, M., Chen, Z., Le, Q.V., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K., Klingner, J., Shah, A., Johnson, M., Liu, X., Kaiser, Ł., Gouws, S., Kato, Y., Kudo, T., Kazawa, H., Stevens, K., Kurian, G., Patil, N., Wang, W., Young, C., Smith, J., Riesa, J., Rudnick, A., Vinyals, O., Corrado, G.,

Hughes, M., & Dean, J., 2016, *Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation*, *Proceedings of the 2016 Conference on Neural Information Processing Systems (NeurIPS 2016)*, Barcelona, Spanyol, 1–10.

Yang, H., Luo, L., Chueng, L.P., Ling, D., dan Chin, F., 2017, Deep Learning and Its Applications to Natural Language Processing, *Journal of Process Control*, 49, 36–44.