

DAFTAR PUSTAKA

- Aggarwal, C. C. (2015). *Data mining: The textbook*. Springer.
<https://doi.org/10.1007/978-3-319-14142-8>
- Bahri, D., Mobahi, H., & Tay, Y. (2022). Sharpness-Aware Minimization Improves Language Model Generalization. In S. Muresan, P. Nakov, & A. Villavicencio (Eds.), *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 7360–7371). Dublin, Ireland: Association for Computational Linguistics.
<https://doi.org/10.18653/v1/2022.acl-long.508>
- Bishop, C. (2006). *Pattern Recognition and Machine Learning*. Springer
- Bommasani, R., Hudson, D., Adeli, E., Altman, R., Arora, S., Arx, S., ... Liang, P. (2021). *On the Opportunities and Risks of Foundation Models*.
<https://doi.org/10.48550/arXiv.2108.07258>
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... Amodei, D. (2020). Language Models are Few-Shot Learners. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, & H. Lin (Eds.), *Advances in Neural Information Processing Systems* (Vol. 33, pp. 1877–1901). Curran Associates, Inc. Retrieved from https://proceedings.neurips.cc/paper_files/paper/2020/file/1457c0d6bfcb4967418bfb8ac142f64a-Paper.pdf
- Buda, M., Maki, A., & Mazurowski, M. A. (2018). A systematic study of the class imbalance problem in convolutional neural networks. *Neural Networks*, 106, 249–259. <https://doi.org/10.1016/j.neunet.2018.07.011>
- Chowdhary, K. R. (2020). Natural Language Processing. In *Fundamentals of Artificial Intelligence* (pp. 603–649). New Delhi: Springer India.
https://doi.org/10.1007/978-81-322-3972-7_19
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 4171–4186). Association for Computational Linguistics. <https://doi.org/10.18653/v1/N19-1423>
- Elmagarmid, A. K., Ipeirotis, P. G., & Verykios, V. S. (2007). Duplicate record detection: A survey. *IEEE Transactions on Knowledge and Data Engineering*, 19(1), 1–16. <https://doi.org/10.1109/TKDE.2007.250581>
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861–874. <https://doi.org/10.1016/j.patrec.2005.10.010>
- Feldman, R., & Sanger, J. (2006). *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge University Press.

- Foret, P., Kleiner, A., Mobahi, H., & Neyshabur, B. (2021). Sharpness-Aware Minimization for Efficiently Improving Generalization. *ICLR 2021 - 9th International Conference on Learning Representations*.
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137–144. <https://doi.org/10.1016/j.ijinfomgt.2014.10.007>
- Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow* (2nd ed.). O'Reilly Media.
- Goldberg, Y. (2017). *Neural Network Methods for Natural Language Processing*. Morgan & Claypool Publishers. <https://doi.org/10.1007/978-3-031-02165-7>
- Goodfellow, I., Courville, A., & Bengio, Y. (2016). *Deep Learning*. Request Exam Copy View Preview *Adaptive Computation and Machine Learning Series*, 1–23.
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques* (3rd ed.). Boston: Morgan Kaufmann.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer. <https://doi.org/10.1007/978-0-387-84858-7>
- He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, 21(9), 1263–1284. <https://doi.org/10.1109/TKDE.2008.239>
- Hirschberg, J., & Manning, C. D. (2015). Advances in natural language processing. *Science (New York, N.Y.)*, 349(6245), 261–266. <https://doi.org/10.1126/science.aaa8685>
- Johnson, J. M., & Khoshgoftaar, T. M. (2019). Survey on deep learning with class imbalance. *Journal of Big Data*, 6(1), 27. <https://doi.org/10.1186/s40537-019-0192-5>
- Jurafsky, D., & Martin, J. H. (2023). *Speech and Language Processing* (3rd ed.). Pearson.
- Klein, D., & Manning, C. D. (2003). Accurate unlexicalized parsing. In *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics (ACL)* (pp. 423–430). Association for Computational Linguistics. <https://doi.org/10.3115/1075096.1075150>
- Keskar, N., Mudigere, D., Nocedal, J., Smelyanskiy, M., & Tang, P. (2016). On Large-Batch Training for Deep Learning: Generalization Gap and Sharp Minima. <https://doi.org/10.48550/arXiv.1609.04836>
- Kingma, D. P., & Ba, J. (2015). *Adam: A method for stochastic optimization*. In *Proceedings of the International Conference on Learning Representations (ICLR)*.

- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Lee, K., Feldman, J., Kumar, S., et al. (2022). Deduplicating training data makes language models better. In *Proceedings of the International Conference on Learning Representations (ICLR)*.
- Li, H., Xu, Z., Taylor, G., Studer, C., & Goldstein, T. (2018). Visualizing the loss landscape of neural nets. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems* (pp. 6391–6401). Red Hook, NY, USA: Curran Associates Inc.
- Liu, B. (2012). *Sentiment analysis and opinion mining*. Morgan & Claypool Publishers. <https://doi.org/10.2200/S00416ED1V01Y201204HLT016>
- Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). *RoBERTa: A robustly optimized BERT pretraining approach*. arXiv preprint arXiv:1907.11692. <https://doi.org/10.48550/arXiv.1907.11692>
- Lorena, A. C., Garcia, L. P. F., Lehmann, J., Souto, M. C. P., & Ho, T. K. (2019). How complex is your classification problem? A survey on measuring classification complexity. *ACM Computing Surveys*, 52(5), Article 107. <https://doi.org/10.1145/3347711>
- Loshchilov, I., & Hutter, F. (2019). Decoupled weight decay regularization. In *Proceedings of the International Conference on Learning Representations (ICLR)*.
- Manning, C., & Schütze, H. (1999). *Foundations of Statistical Natural Language Processing*. Cambridge, MA: MIT Press.
- Munika, M., Shaky, S., & Shrestha, A. (2019). Fine-grained sentiment classification using BERT. In *Proceedings of the 2019 Artificial Intelligence for Transforming Business and Society Conference (AITB)*. <https://doi.org/10.1109/AITB48515.2019.8947435>
- Murphy, K. P. (2012). *Machine Learning A Probabilistic Perspective*. Cambridge, MA: MIT Press.
- Na, C., Mehta, S. V., & Strubell, E. (2022). Train Flat, Then Compress: Sharpness-Aware Minimization Learns More Compressible Models. In *Findings of the Association for Computational Linguistics: EMNLP 2022* (pp. 4938–4965). Association for Computational Linguistics (ACL). <https://doi.org/10.18653/v1/2022.findings-emnlp.200>
- Otter, D. W., Medina, J. R., & Kalita, J. K. (2021). A Survey of the Usages of Deep Learning for Natural Language Processing. *IEEE Transactions on Neural Networks and Learning Systems*, 32(2), 604–624. <https://doi.org/10.1109/TNNLS.2020.2979670>
- Oxford University Press. (n.d.). *Effectiveness*. In *Oxford Learner's Dictionary*. <https://www.oxfordlearnersdictionaries.com/>

- Pang, B., & Lee, L. (2005). Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL)* (pp. 115–124). Association for Computational Linguistics. <https://doi.org/10.3115/1219840.1219855>
- Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1–2), 1–135. <https://doi.org/10.1561/15000000011>
- Powers, D. M. W. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, 2(1), 37–63.
- Pusat Bahasa. (n.d.). *Efektivitas*. Dalam *Kamus Besar Bahasa Indonesia (KBBI)*. <https://kbbi.kemdikbud.go.id/>
- Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). *Improving language understanding by generative pre-training*. OpenAI. <https://doi.org/10.48550/arXiv.1801.06146>
- Rogers, A., Kovaleva, O., & Rumshisky, A. (2020). A Primer in BERTology: What We Know About How BERT Works. *Transactions of the Association for Computational Linguistics*, 8, 842–866. https://doi.org/10.1162/tacl_a_00349
- Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C. D., Ng, A., & Potts, C. (2013). Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank. In D. Yarowsky, T. Baldwin, A. Korhonen, K. Livescu, & S. Bethard (Eds.), *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing* (pp. 1631–1642). Seattle, Washington, USA: Association for Computational Linguistics.
- Strang, G. (2016). *Introduction to Linear Algebra* (5th ed.). Wellesley, MA: Wellesley-Cambridge Press.
- Sun, C., Qiu, X., Xu, Y., & Huang, X. (2019). How to fine-tune BERT for text classification? In *Chinese Computational Linguistics* (pp. 194–206). Springer. https://doi.org/10.1007/978-3-030-32381-3_16
- Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational Linguistics*, 37(2), 267–307. https://doi.org/10.1162/COLI_a_00049
- Tan, P.-N., Steinbach, M., Karpatne, A., & Kumar, V. (2019). *Introduction to Data Mining. Global Edition* (2nd ed.). Pearson Deutschland.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... Polosukhin, I. (2017). Attention is All you Need. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems* (Vol. 30). Curran Associates, Inc. Retrieved from

https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547de91fbd053c1c4a845aa-Paper.pdf

- Wolf, T., Debut, L., Sanh, V., et al. (2020). Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations* (pp. 38–45). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.emnlp-demos.6>
- Wu, Y., Schuster, M., Chen, Z., et al. (2016). Google’s neural machine translation system: Bridging the gap between human and machine translation. *Transactions of the Association for Computational Linguistics*, 4, 339–351. https://doi.org/10.1162/tacl_a_00080
- Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent Trends in Deep Learning Based Natural Language Processing [Review Article]. *IEEE Computational Intelligence Magazine*, 13(3), 55–75. <https://doi.org/10.1109/MCI.2018.2840738>