

DAFTAR PUSTAKA

- Adzmi, M. S., Purnomo, A. M., & Agustini, A. (2026). Pengaruh tipe berlebihan dalam judul berita clickbait di media *online* Tribunnews.com terhadap motivasi membaca mahasiswa di Kota Bogor. *Journal of Innovative and Creativity* (Joecy), 6(1), 6720–6726. <https://doi.org/10.31004/joecy.v6i1.7493>
- Ambarwati, R., Sari, N. A. S., & Rosyid, H. A. (2025). Implementasi TF-IDF dan *Cosine Similarity* untuk deteksi relevansi berita korupsi. *Jurnal Teknologi Informasi dan Ilmu Komputer*, 8(2), 145–162.
- APJII. (2025). Profil Internet Indonesia 2025 & Segmentasi pasar ISP tahun 2025. Asosiasi Penyelenggara Jasa Internet Indonesia.
- Baeza-Yates, R., & Ribeiro-Neto, B. (2011). *Modern information retrieval: The concepts and technology behind search* (2nd ed.). Addison-Wesley Professional.
- Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python: Analyzing text with the Natural Language Toolkit*. O'Reilly Media.
- Carpineto, C., & Romano, G. (2012). A survey of automatic query expansion in information retrieval. *ACM Computing Surveys*, 44(1), 1–50. <https://doi.org/10.1145/2071389.2071390>
- Choudhary, N., Singh, P. K., & Tayal, D. K. (2020). BERT vs TF-IDF: A comparative study for *document retrieval*. *Proceedings of the International Conference on Innovative Computing & Communications (ICICC 2020)*, 2, 523–533. https://doi.org/10.1007/978-981-16-1645-2_45
- DataReportal & Kemp, S. (2025). Digital 2025: Indonesia. Retrieved from <https://datareportal.com/reports/digital-2025-indonesia>
- DataReportal & Kemp, S. (2025). Digital 2025: *Global overview report*. We Are Social & Meltwater. <https://datareportal.com/reports/digital-2025-global-overview-report>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: *Pre-training of deep bidirectional transformers for language understanding*. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 4171–4186. <https://doi.org/10.18653/v1/N19-1423>
- Fathuddin, M. A. H., Mandyartha, E. P., & Nurlaili, A. L. (2025). Penerapan *Sentence-BERT* dan *Cosine Similarity* untuk pencarian semantik dokumen skripsi dalam format PDF. *Ranah Research: Journal of Multidisciplinary Research and Development*, 8(1), 322–337. <https://doi.org/10.38035/rj.v8i1.1865>
- Guo, J., Fan, Y., Pang, L., Yang, L., Ai, Q., Zamani, H., Wu, C., Croft, W. B., & Cheng, X. (2020). A deep look into *neural ranking models for information retrieval*. *Information Processing & Management*, 57(6), 102067. <https://doi.org/10.1016/j.ipm.2019.102067>
- Goyal, P., & Sharma, A. (2022). A comparative study of TF-IDF and BM25 for *document retrieval*. *International Journal of Computer Applications*, 184(15), 1–4. <https://doi.org/10.5120/ijca2022922116>

- Hersianty, V. M., Amalia, E. L., Puspitasari, D., & Wibowo, D. W. (2025). Penerapan algoritma TF-IDF dan *Cosine Similarity* dalam sistem rekomendasi lowongan pekerjaan. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 9(1), 1619–1625. <https://doi.org/10.36040/jati.v9i1.12406>
- Jain, A., Jain, A., Chauhan, N., Singh, V., & Thakur, N. (2017). *Information retrieval using cosine and Jaccard similarity measures in Vector space Model. International Journal of Computer Applications*, 164(6), 28–30. <https://doi.org/10.5120/ijca2017913699>
- Jurafsky, D., & Martin, J. H. (2024). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (3rd ed.). Stanford University. Retrieved from <https://web.stanford.edu/~jurafsky/slp3/>
- Lin, J., Nogueira, R., & Yates, A. (2021). Pretrained transformers for text ranking: BERT and beyond. *Synthesis Lectures on Human Language Technologies*, 14(4), 1–325. Morgan & Claypool Publishers. <https://doi.org/10.2200/S01123ED1V01Y202108HLT05>
- Miller, G. A., & Fellbaum, C. (2024). *WordNet: An on-line lexical database*. Princeton University. Retrieved from <https://WordNet.princeton.edu/>
- Moffat, A., & Zobel, J. (2008). Rank-biased *precision* for measurement of *retrieval effectiveness*. *ACM Transactions on Information Systems*, 27(1), Article 2. <https://doi.org/10.1145/1416950.1416952>
- Newman, N., Fletcher, R., Robertson, C. T., Arguedas, A. R., & Nielsen, R. K. (2024). *Reuters Institute Digital News Report 2024*. Reuters Institute for the Study of Journalism.
- Newman, N., Fletcher, R., Robertson, C. T., Arguedas, A. R., & Nielsen, R. K. (2025). *Reuters Institute Digital News Report 2025*. Reuters Institute for the Study of Journalism. <https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2025>
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511809071>
- Pandit, O., Denis, P., & Ralaivola, L. (2020). Integrating knowledge graph embeddings to improve mention representation for bridging anaphora resolution. *Proceedings of the Third Workshop on Computational Models of Reference, Anaphora and Coreference (CRAC 2020)*, 55–67. <https://aclanthology.org/2020.crac-1.7/>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). *Scikit-learn: Machine learning in Python*. *Journal of Machine Learning Research*, 12, 2825–2830.
- Al Rasyid, R., & Ningsih, D. H. U. (2024). Penerapan algoritma TF-IDF dan Cosine Similarity untuk query pencarian pada dataset destinasi wisata. *Jurnal Teknologi Informasi dan Komunikasi (JTik)*, 8(1), 170–178. <https://doi.org/10.35870/jtik.v8i1.1416>
- Reimers, N., & Gurevych, I. (2019). *Sentence-BERT: Sentence embeddings using siamese BERT-networks*. *Proceedings of the 2019 Conference on Empirical Methods in Natural language processing and the 9th International Joint*

- Conference on *Natural language processing* (EMNLP-IJCNLP), 3982–3992. <https://doi.org/10.18653/v1/D19-1410>
- Ren, R., Wang, Y., Qu, Y., Zhao, W. X., Liu, J., Tian, H., Wu, H., Wen, J.-R., & Wang, H. (2023). Investigating the factual *knowledge* boundary of large *language* models with *retrieval* augmentation. arXiv preprint arXiv:2307.11019. <https://doi.org/10.48550/arXiv.2307.11019>
- Salton, G., & Buckley, C. (1988). *Term-weighting* approaches in automatic text *retrieval*. *Information Processing & Management*, 24(5), 513–523. [https://doi.org/10.1016/0306-4573\(88\)90021-0](https://doi.org/10.1016/0306-4573(88)90021-0)
- Shalihah, R., Rifai, S. N., Wiradarma, A., & Nurfajrin, S. A. A. (2025). Study literatur: *Query expansion* pada temu kembali informasi (*information retrieval*). *Jurnal Sutasoma*, 4(1), 44–50. <https://doi.org/10.58878/sutasoma.v4i1.411>
- Shi, L., Bhattacharya, N., Das, A., & Gwizdka, J. (2023). True or false? Cognitive load when reading COVID-19 *News* headlines: An eye-tracking study. *Proceedings of the 2023 Conference on Human Information Interaction and Retrieval* (CHIIR 2023), 107–116. <https://doi.org/10.1145/3576840.3578290>
- Tandon, A., Dhir, A., Islam, A. K. M. N., & Mäntymäki, M. (2022). Impact of *News* overload on social media *News* curation: Mediating role of *News* avoidance. *Frontiers in Psychology*, 13, 865246. <https://doi.org/10.3389/fpsyg.2022.865246>
- Thakur, N., Reimers, N., Rüklé, A., Srivastava, A., & Gurevych, I. (2021). BEIR: A heterogeneous *benchmark* for zero-shot evaluation of *information retrieval* models. In *Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks* (NeurIPS Datasets and Benchmarks 2021, Round 2). <https://openreview.net/forum?id=wCu6T5xFjeJ>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems* 30 (NIPS 2017), 5998–6008. Retrieved from <https://arxiv.org/abs/1706.03762>
- Wang, W., Wei, F., Dong, L., Bao, H., Yang, N., & Zhou, M. (2020). MiniLM: Deep *self-attention distillation* for task-agnostic compression of pre-trained *transformers*. *Advances in Neural Information Processing Systems*, 33, 5776–5788. Retrieved from <https://arxiv.org/abs/2002.10957>
- Widianto, A., Pebriyanto, E., Fitriyanti, F., & Marna, M. (2024). *Document similarity* using *term frequency-inverse document frequency* representation and *Cosine Similarity*. *Journal of Dinda: Data Science, Information Technology, and Data Analytics*, 4(2), 149–153. <https://doi.org/10.20895/DINDA.V4I2.1589>
- Widianto, S., Pratama, A. R., & Nugraha, F. (2024). *Document similarity* measurement using TF-IDF and *Cosine Similarity*. *KOMPUTA: Jurnal Ilmiah Komputer dan Informatika*, 13(1), 25–32. <https://doi.org/10.34010/komputa.v13i1.12401>