

DAFTAR PUSTAKA

- Anjani, M., & Irmanda, H. N. (2024). Comparison Performance Of Word2Vec, GLOVE, FastText Using Support Vector Machine Method For Sentiment Analysis. *Jurnal Teknik Informatika (Jutif)*, 5(3), 669–674. <https://doi.org/10.52436/1.jutif.2024.5.3.1366>
- Aydoğan Kılıç, D., Kılıç, D. K., & Nielsen, I. E. (2025). Comparative study of feature extraction methods for automated ICD code classification using MIMIC-III medical notes and deep learning models. *Mathematical Modelling and Numerical Simulation with Applications*, 5(2), 421–450. <https://doi.org/10.53391/mmnsa.1666223>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding*.
- Ding, J., Qiao, Y., & Zhang, L. (2023). Plant Disease Prescription Recommendation Based on Electronic Medical Records and Sentence Embedding Retrieval. *Plant Methods*, 19(1), 91. <https://doi.org/10.1186/s13007-023-01070-6>
- Douze, M., Guzhva, A., Deng, C., Johnson, J., Szilvasy, G., Mazaré, P.-E., Lomeli, M., Hosseini, L., & Jégou, H. (2025). *The Faiss Library*. <http://arxiv.org/abs/2401.08281>
- Gao, J., Galley, M., & Li, L. (2019). Neural Approaches to Conversational AI. *Foundations and Trends® in Information Retrieval*, 13(2–3), 127–298. <https://doi.org/10.1561/15000000074>
- Garg, S. (2021). Drug Recommendation System based on Sentiment Analysis of Drug Reviews using Machine Learning. *2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, 175–181. <https://doi.org/10.1109/Confluence51648.2021.9377188>
- Heryawan, L., Novitaningrum, D., Nastiti, K. R., & Mahmudah, S. N. (2024). Medical Record Document Search with TF-IDF and Vector Space Model (VSM). *International Journal on Advanced Science, Engineering and Information Technology*, 14(3), 847–852. <https://doi.org/10.18517/ijaseit.14.3.19606>
- Indriyani, T., Pradana, A. I., & Hartanti, D. (2025). Penerapan Metode Content-Based Filtering Pada Sistem Rekomendasi Pemilihan Produk Obat Studi Kasus : Apotek Hero Farma. *Infotek: Jurnal Informatika Dan Teknologi*, 8(2), 532–542. <https://doi.org/10.29408/jit.v8i2.30498>
- Jadon, A., & Patil, A. (2024). *A Comprehensive Survey of Evaluation Techniques for Recommendation Systems*. <http://arxiv.org/abs/2312.16015>

- Jin, Q., Kim, W., Chen, Q., Comeau, D. C., Yeganova, L., Wilbur, W. J., & Lu, Z. (2023). MedCPT: Contrastive Pre-Trained Transformers With Large-Scale Pubmed Search Logs For Zero-Shot Biomedical Information Retrieval. *Bioinformatics*, 39(11). <https://doi.org/10.1093/bioinformatics/btad651>
- Kartika, B. V., Alfredo, M. J., & Kusuma, G. P. (2023). Fine-Tuned IndoBERT Based Model and Data Augmentation for Indonesian Language Paraphrase Identification. *Revue d'Intelligence Artificielle*, 37(3), 733–743. <https://doi.org/10.18280/ria.370322>
- Khurana, D., Koli, A., Khatter, K., & Singh, S. (2023). Natural Language Processing: State of The Art, Current Trends and Challenges. *Multimedia Tools and Applications*, 82(3), 3713–3744. <https://doi.org/10.1007/s11042-022-13428-4>
- Koto, F., Rahimi, A., Lau, J. H., & Baldwin, T. (2020). IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP. *Proceedings of the 28th International Conference on Computational Linguistics*, 757–770. <https://doi.org/10.18653/v1/2020.coling-main.66>
- Limbong, M., Hidayah, N., & Atjo, N. M. (2023). *Farmakologi Sosial dan Pengelolaan Obat*.
- Minaee, S., Kalchbrenner, N., Cambria, E., Nikzad, N., Chenaghlu, M., & Gao, J. (2022). Deep Learning - Based Text Classification. *ACM Computing Surveys*, 54(3), 1–40. <https://doi.org/10.1145/3439726>
- Naseralallah, L., Stewart, D., Azfar Ali, R., & Paudyal, V. (2022). An Umbrella Review of Systematic Reviews on Contributory Factors to Medication Errors in Health-Care Settings. *Expert Opinion on Drug Safety*, 21(11), 1379–1399. <https://doi.org/10.1080/14740338.2022.2147921>
- Palomino, M. A., & Aider, F. (2022). Evaluating the Effectiveness of Text Pre-Processing in Sentiment Analysis. *Applied Sciences*, 12(17), 1–21. <https://doi.org/10.3390/app12178765>
- Pandey, A., Gupta, A., & Pudi, V. (2022). *CitRet: A Hybrid Model for Cited Text Span Retrieval*. <http://www.nist.gov/tac/2014/BiomedSumm/>
- Pirdaus, D. I., & Hidayana, R. A. (2024). Analysis Testing Black Box and White Box on Application To-Do List Based Web. *International Journal of Mathematics, Statistics, and Computing*, 2(2), 68–75.
- Pradana, A. W., & Hayaty, M. (2019). The Effect of Stemming and Removal of Stopwords on the Accuracy of Sentiment Analysis on Indonesian-language Texts. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, 375–380. <https://doi.org/10.22219/kinetik.v4i4.912>

- Putri, S. K., Amalia, A., Nababan, E. B., & Sitompul, O. S. (2021). Bahasa Indonesia pre-trained word vector generation using word2vec for computer and information technology field. *Journal of Physics: Conference Series*, 1898(1), 1–10. <https://doi.org/10.1088/1742-6596/1898/1/012007>
- Qiu, X., Sun, T., Xu, Y., Shao, Y., Dai, N., & Huang, X. (2020). Pre-trained models for natural language processing: A survey. *Science China Technological Sciences*, 63(10), 1872–1897. <https://doi.org/10.1007/s11431-020-1647-3>
- R, N., Khan, S. B., kumar, A. V., T R, M., Alojail, M., Sangwan, S. R., & Saraee, M. (2025). Enhancing Drug Discovery and Patient Care Through Advanced Analytics with The Power of NLP and Machine Learning in Pharmaceutical Data Interpretation. *SLAS Technology*, 31, 1–11. <https://doi.org/10.1016/j.slast.2024.100238>
- Reimers, N., & Gurevych, I. (2019). *Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks*.
- Schiavo, G., Forgerini, M., Lucchetta, R. C., Silva, G. O., & Mastroianni, P. d. C. (2022). Cost of Adverse Drug Events Related to Potentially Inappropriate Medication Use: a Systematic Review. *Journal of the American Pharmacists Association*, 62(5), 1463–1476. <https://doi.org/10.1016/j.japh.2022.04.008>
- Selvaraj, A., Ponkalpana, P., Nazeeha Sherin, S., Sri Harini, G., Professor, A., & Year Students, F. (2024). Medicine Recommendation System Using Cosine Similarity. *International Journal of Scientific Development and Research*, 9(4), 1366–1370. www.ijedr.org
- Septiani, D., & Isabela, I. (2022). Analisis Term Frequency Inverse Document Frequency (TF-IDF) dalam Temu Kembali Informasi pada Dokumen Teks. *SINTESIA: Jurnal Sistem Dan Teknologi Informasi Indonesia*, 01(2), 81–88.
- Shukla, D., & Dwivedi, S. K. (2024). The Study of The Effect of Preprocessing Techniques for Emotion Detection on Amazon Product Review Dataset. *Social Network Analysis and Mining*, 14(1), 191. <https://doi.org/10.1007/s13278-024-01352-4>
- Siino, M., Tinnirello, I., & La Cascia, M. (2024). Is text preprocessing still worth the time? A comparative survey on the influence of popular preprocessing methods on Transformers and traditional classifiers. *Information Systems*, 121, 1–19. <https://doi.org/10.1016/j.is.2023.102342>
- Siti Khomsah, Rima Dias Ramadhani, & Sena Wijaya. (2022). The Accuracy Comparison Between Word2Vec and FastText On Sentiment Analysis of Hotel Reviews. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 6(3), 352–358. <https://doi.org/10.29207/resti.v6i3.3711>
- Sutriawan, Rustad, S., Shidik, G. F., & Pujiono. (2025). Performance Evaluation of Text Embedding Models for Ambiguity Classification in Indonesian News

- Corpus: A Comparative Study of TF-IDF, Word2Vec, FastText BERT, and GPT. *Ingenierie Des Systemes d'Information*, 30(6), 1469–1482. <https://doi.org/10.18280/isi.300606>
- Tang, G., Izumi, K., Izumisawa, M., & Koyama, S. (2023). Analysis of Japanese Consumers' Attitudes Toward The Digital Transformation of OTC Medicine Purchase Behavior and Ehealth Literacy: an Online Survey for Digital Experience Design. *Frontiers in Digital Health*, 5. <https://doi.org/10.3389/fdgth.2023.1173229>
- Turchin, A., Masharsky, S., & Zitnik, M. (2023). Comparison of BERT Implementations for Natural Language Processing of Narrative Medical Documents. *Informatics in Medicine Unlocked*, 36, 1. <https://doi.org/10.1016/j.imu.2022.101139>
- Weber, F., & Toddenroth, D. (2024). *Comparative Evaluation of Pre-Trained Language Models for Biomedical Information Retrieval*. <https://doi.org/10.3233/SHTI240539>
- Wibawa, A. D., Ramadhani, P. A., Buntoro, G. A., Hariadi, R. R., Siswanto, P. A., & Sabilla, S. I. (2023). Leveraging Text-Mining Techniques On Electronic Medical Records to Analyze National Drug-insured Medication Use. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, 8(2), 581–590. <https://doi.org/10.22219/kinetik.v8i2.1695>
- Young, J. C., & Rusli, A. (2019). Review and Visualization of Facebook's FastText Pretrained Word Vector Model. *2019 International Conference on Engineering, Science, and Industrial Applications (ICESI)*, 1–6. <https://doi.org/10.1109/ICESI.2019.8863015>
- Zomorodi, M., Ghodsollahee, I., Martin, J. H., Talley, N. J., Salari, V., Plawiak, P., Rahimi, K., & Rajendra Acharya, U. (2022). *RECOMED: A Comprehensive Pharmaceutical Recommendation System*.