DAFTAR PUSTAKA

Adamopoulou, E., & Moussiades, L. (2020). An Overview of Chatbot Technology. Dalam I. Maglogiannis, L. Iliadis, & E. Pimenidis (Ed.), *Artificial Intelligence Applications and Innovations* (Vol. 584, hlm. 373–383). Springer International Publishing.

Ajiz, M. F., Ramadan, M. F. S., Mutia, H. D., & Yanuari, P. D. (2023). Pengembangan Aplikasi Chatbot Informasi Akademik Berbasis Web Menggunakan Metode Artificial Intelligence Markup Language (AIML). *Media Jurnal Informatika*, *15*(2), 143.

Al-Amin, M., Ali, M. S., Salam, A., Khan, A., Ali, A., Ullah, A., ... & Chowdhury, S. K. (2024). History of generative Artificial Intelligence (AI) chatbots: past, present, and future development. arXiv preprint arXiv:2402.05122.

Caragea, C., & Honavar, V. (t.t.). *Machine Learning in Computational Biology*.

Colace, F., Santo, M. D., Lombardi, M., Pascale, F., Pietrosanto, A., & Lemma, S. (2018). Chatbot for E-Learning: A Case of Study. *International Journal of Mechanical Engineering and Robotics Research*, 528–533. https://doi.org/10.18178/ijmerr.7.5.528-533

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding* (No. arXiv:1810.04805). arXiv. https://doi.org/10.48550/arXiv.1810.04805

Du, K.-L., Leung, C.-S., Mow, W. H., & Swamy, M. N. S. (2022). Perceptron:

Learning, Generalization, Model Selection, Fault Tolerance, and Role in the Deep Learning Era. *Mathematics*, *10*(24), 4730. https://doi.org/10.3390/math10244730

Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., Wang, M., & Wang, H. (2024). *Retrieval-Augmented Generation for Large Language Models: A Survey* (No. arXiv:2312.10997). arXiv. https://doi.org/10.48550/arXiv.2312.10997

Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). Deep learning (Vol. 1, No. 2). Cambridge: MIT press.

Henderson, M., Al-Rfou, R., Strope, B., Sung, Y., Lukacs, L., Guo, R., Kumar, S., Miklos, B., & Kurzweil, R. (2017). *Efficient Natural Language Response Suggestion for Smart Reply* (No. arXiv:1705.00652). arXiv. https://doi.org/10.48550/arXiv.1705.00652

Khan, A. A., Hasan, M. T., Kemell, K. K., Rasku, J., & Abrahamsson, P. (2024). *Developing Retrieval Augmented Generation (RAG) based LLM Systems from PDFs: An Experience Report* (No. arXiv:2410.15944). arXiv. https://doi.org/10.48550/arXiv.2410.15944

Kim, S., Hooper, C., Wattanawong, T., Kang, M., Yan, R., Genc, H., Dinh, G., Huang, Q., Keutzer, K., Mahoney, M. W., Shao, Y. S., & Gholami, A. (2023). *Full Stack Optimization of Transformer Inference: A Survey* (No. arXiv:2302.14017). arXiv. https://doi.org/10.48550/arXiv.2302.14017

Lantana, D. A., Ningsih, S., & Waluyo, T. (t.t.). *RANCANG BANGUN CHATBOT BERBASIS RULE-BASED SEBAGAI PUSAT INFORMASI CALON MAHASISWA BARU DI UNIVERSITAS NASIONAL*.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Computation*, *1*(4), 541–551. https://doi.org/10.1162/neco.1989.1.4.541

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2021). *Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks* (No. arXiv:2005.11401). arXiv. https://doi.org/10.48550/arXiv.2005.11401

Malkov, Y. A., & Yashunin, D. A. (2020). Efficient and Robust Approximate Nearest Neighbor Search Using Hierarchical Navigable Small World Graphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *42*(4), 824–836. https://doi.org/10.1109/TPAMI.2018.2889473

Nasution, A. R. S. (2021). Identifikasi permasalahan penelitian. ALACRITY: Journal of Education, 13-19.

Neupane, S., Hossain, E., Keith, J., Tripathi, H., Ghiasi, F., Golilarz, N. A., Amirlatifi, A., Mittal, S., & Rahimi, S. (2024). *From Questions to Insightful*

Answers: Building an Informed Chatbot for University Resources (No. arXiv:2405.08120). arXiv. https://doi.org/10.48550/arXiv.2405.08120

Pan, J. J., Wang, J., & Li, G. (2023). *Survey of Vector Database Management Systems* (No. arXiv:2310.14021). arXiv. https://doi.org/10.48550/arXiv.2310.14021

Reimers, N., & Gurevych, I. (2019). *Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks* (No. arXiv:1908.10084). arXiv. https://doi.org/10.48550/arXiv.1908.10084

Samudra, G., & Zy, A. T. (2025). *Implementasi Retrieval Augmented Generation (RAG) Dalam Perancangan Chatbot Kesehatan Pencernaan*. *8*(1).

Shalhan, A. (2024). Implementasi teknik retrieval augmented generation untuk menghilangkan halusinasi pada large language model berbasis vector database [Skripsi, Universitas Multimedia Nusantara]. Repositori Universitas Multimedia Nusantara. https://kc.umn.ac.id/id/eprint/34785/

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., & Lample, G. (2023). *LLaMA: Open and Efficient Foundation Language Models* (No. arXiv:2302.13971). arXiv. https://doi.org/10.48550/arXiv.2302.13971

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2023). *Attention Is All You Need* (No.

arXiv:1706.03762). arXiv. https://doi.org/10.48550/arXiv.1706.03762

Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., Min, Y., Zhang, B., Zhang, J., Dong, Z., Du, Y., Yang, C., Chen, Y., Chen, Z., Jiang, J., Ren, R., Li, Y., Tang, X., Liu, Z., … Wen, J.-R. (2024). *A Survey of Large Language Models* (No. arXiv:2303.18223). arXiv. https://doi.org/10.48550/arXiv.2303.18223

www.itk.ac.id

www.itk.ac.id

(Halaman ini sengaja dikosongkan)

www.itk.ac.id