

DAFTAR PUSTAKA

- Allen, J. F. (2016). Natural Language Understanding. *Artificial Intelligence*.
<https://api.semanticscholar.org/CorpusID:46208633>
- Church, K. W., Sun, J., Yue, R., Vickers, P., Saba, W., & Chandrasekar, R. (2024). Emerging trends: a gentle introduction to RAG. *Natural Language Engineering*, 30(4), 870–881.
<https://doi.org/10.1017/S1351324924000044>
- Dadure, P., Pakray, P., & Bandyopadhyay, S. (2022). A formula embedding approach for semantic similarity and relatedness between formulas. *Concurrency and Computation: Practice and Experience*, 34(22). <https://doi.org/10.1002/cpe.7146>
- Deutsch, D., & Roth, D. (2021). Understanding the Extent to which Content Quality Metrics Measure the Information Quality of Summaries. *CoNLL 2021 - 25th Conference on Computational Natural Language Learning, Proceedings*, 300–309.
<https://doi.org/10.18653/V1/2021.CONLL-1.24>
- Ehsan, T., Khalid, J., Ambreen, S., Mustafa, A., & Hussain, S. (2022). Improving Phrase Chunking by using Contextualized Word Embeddings for a Morphologically Rich Language. *Arabian Journal for Science and Engineering*, 47(8), 9781–9799.
<https://doi.org/10.1007/s13369-021-06343-7>
- Elsayed, A. A., Hafez, A. I., Ceprián, R., Martínez, G., Granados, A., Soriano, B., Llorens, C., & Sempere, J. M. (2024). Practical Design and Implementation of Virtual Chatbot Assistants for Bioinformatics Based on a NLU Open Framework. *Big Data and Cognitive Computing*, 8(11). <https://doi.org/10.3390/bdcc8110163>
- Fergencs, T., & Meier, F. (2021). Engagement and Usability of Conversational Search – A Study of a Medical Resource Center Chatbot. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12645 LNCS, 328–345. https://doi.org/10.1007/978-3-030-71292-1_26
- Giabelli, A. (2022). A Unified Framework for Intrinsic Evaluation of Word-Embedding Algorithms. *IJCAI International Joint Conference on Artificial Intelligence*, 5853–5854. <https://doi.org/10.24963/ijcai.2022/829>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- Hao, X., Zhang, G., & Ma, S. (2016). Deep Learning. *International Journal of Semantic Computing*, 10(3), 417–439. <https://doi.org/10.1142/S1793351X16500045>
- Jain, A. M., & Jain, A. (2025). Scaling LLM Inference Architectures: A Performance Analysis for Chatbot Applications. *2025 6th International Conference on Artificial*

- Intelligence, Robotics, and Control, AIRC 2025*, 8–16.
<https://doi.org/10.1109/AIRC64931.2025.11077484>
- Jeong, T. (2024). 4bit-Quantization in Vector-Embedding for RAG. *Proceedings - 2024 International Conference on Machine Learning and Applications, ICMLA 2024*, 1037–1042. <https://doi.org/10.1109/ICMLA61862.2024.00156>
- Jiang, S., Cormier, S., Angarita, R., & Rousseaux, F. (2023). Improving text mining in plant health domain with GAN and/or pre-trained language model. *Frontiers in Artificial Intelligence*, 6. <https://doi.org/10.3389/frai.2023.1072329>
- Joshi, D., Tekade, S., Zanzane, S., Sakharwade, D., Tripathi, A., & Tiwadi, S. (2024). Generative AI-Driven Chatbot for Personalized University Information and Assistance. *2024 International Conference on Artificial Intelligence and Quantum Computation-Based Sensor Applications, ICAIQSA 2024 - Proceedings*. <https://doi.org/10.1109/ICAIQSA64000.2024.10882421>
- Jurafsky, D., & Martin, J. H. (2026). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models* (3rd ed.). <https://web.stanford.edu/~jurafsky/slp3/>
- Kang, G., Zhang, X., Ge, Z., Wang, L., Hu, J., & Zhan, J. (2025). BigVectorBench: Heterogeneous Data Embedding and Compound Queries are Essential in Evaluating Vector Databases. *Proceedings of the VLDB Endowment*, 18(5), 1536–1550. <https://doi.org/10.14778/3718057.3718078>
- Kaster, M., Zhao, W., & Eger, S. (2021). Global Explainability of BERT-Based Evaluation Metrics by Disentangling along Linguistic Factors. *EMNLP 2021 - 2021 Conference on Empirical Methods in Natural Language Processing, Proceedings*, 8912–8925. <https://doi.org/10.18653/V1/2021.EMNLP-MAIN.701>
- Kim, S., Moon, J., Oh, J., Choi, I., & Yang, J. S. (2025). Survey and Evaluation of Converging Architecture in LLMs Based on Footsteps of Operations. *IEEE Open Journal of the Computer Society*. <https://doi.org/10.1109/OJCS.2025.3587005>
- Kong, X., & Wang, G. (2021). *Conversational AI with Rasa : build, automate, and deploy AI-powered text and voice-based assistants and chatbots*. Packt Publishing Ltd.
- Lee, J., Kim, S., Jo, W., & Yoo, H. J. (2023). An Energy-Efficient Heterogeneous Fourier Transform-Based Transformer Accelerator with Frequency-Wise Dynamic Bit-Precision. *2023 IEEE Asian Solid-State Circuits Conference, A-SSCC 2023*. <https://doi.org/10.1109/A-SSCC58667.2023.10347933>
- Liu, L., & Xu, X. (2023). Self-attention Mechanism at the Token Level: Gradient Analysis and Algorithm Optimization. *Knowledge-Based Systems*, 277, 110784. <https://doi.org/10.1016/J.KNOSYS.2023.110784>

- Luise, R. S. A. De, Arevalillo-Herraez, M., & Arnau, D. (2023). On Using Conversational Frameworks to Support Natural Language Interaction in Intelligent Tutoring Systems. *IEEE Transactions on Learning Technologies*, 16(5), 722–735. <https://doi.org/10.1109/TLT.2023.3245121>
- Maheswari, B., & Nagarajan, R. G. (2023). Educational Chatbot for Adopting Effective Teaching Learning Process using Long Short-Term Memory with Feature Extraction. *Proceedings of the 2023 2nd International Conference on Augmented Intelligence and Sustainable Systems, ICAISS 2023*, 1749–1757. <https://doi.org/10.1109/ICAISS58487.2023.10250704>
- Mungai, P. K., & Huang, R. (2017). Chunking mechanisms for a self improving associative memory model. *2017 IEEE Symposium Series on Computational Intelligence, SSCI 2017 - Proceedings, 2018-January*, 1–6. <https://doi.org/10.1109/SSCI.2017.8285215>
- NGO, T. T. A., BUI, C. T., CHAU, H. K. L., TRAN, N. P. N., NGUYEN, P. P. T., & TRAN, N. K. T. (2025). Influence of university service quality on student experiences, academic performance and institutional loyalty: A case study in Vietnam. *Acta Psychologica*, 260, 105599. <https://doi.org/10.1016/J.ACTPSY.2025.105599>
- Oliveira, P. F., & Matos, P. (2023). Introducing a Chatbot to the Web Portal of a Higher Education Institution to Enhance Student Interaction †. *Engineering Proceedings*, 56(1). <https://doi.org/10.3390/ASEC2023-16621>
- Osman, Z., Yusoof, F., ... A. A.-Int. J. Acad. Res. P., & 2024, undefined. (2024). Enhancing student retention through academic support and course design: Satisfaction as mediator. *kwpublications.com Z Osman, F Yusoof, A Ali, RC Aziz Int. J. Acad. Res. Progress. Educ. Dev, 2024•kwpublications.com, 13(4)*. <https://doi.org/10.6007/IJARPED/v13-i4/23679>
- Peña-Lang, M. B., Barrutia, J. M., & Echebarria, C. (2023). Service quality and students' academic achievement. *Quality Assurance in Education*, 31(2), 247–262. <https://doi.org/10.1108/QAE-02-2022-0039>
- Plazas, R. A., & Avella Rodríguez, E. J. (2024). Chatbots: The Future of Education? *International Journal of Engineering Pedagogy*, 14(8), 107–119. <https://doi.org/10.3991/ijep.v14i8.49715>
- Purba, C. N. Y., Parulian, P. V. A., Iswara, R. W., Suhandana, A. A., & Sari, R. (2024). The Implementation of Natural Language Processing on Chatbot as Academic Information System: A Case Study of IC2IE 2024. *Proceedings 7th IC2IE 2024 - 2024 International Conference of Computer and Informatics Engineering: Generative AI in Democratizing Access to Knowledge and Skills*. <https://doi.org/10.1109/IC2IE63342.2024.10748156>
- Rogers, S., Girolami, M., Rish, I., Grabarnik, G., Liu, X., Datta, A., Lim, E.-P., Suykens, J. A. K., Signoretto, M., Argyriou, A., Marsland, S., Programming, B., Bessiere, P.,

- Mazer, E., Ahuactzin, J. M., Mekhnacha, K., Lu, H., Plataniotis, K. N., Venetsanopoulos, A., ... Lappin, S. (2022). Transformers for Machine Learning: A Deep Dive. *Transformers for Machine Learning*. <https://doi.org/10.1201/9781003170082>
- Rothman, Denis. (2024). *RAG-Driven Generative AI: Build Custom Retrieval Augmented Generation Pipelines with LlamaIndex, Deep Lake, and Pinecone*. Packt Publishing Ltd.
- S., K., Kailath, B. J., & B.S., R. (2024). CiMComp: An Energy Efficient Compute-in-Memory Based Comparator for Convolutional Neural Networks. *2024 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, 1–2. <https://doi.org/https://doi.org/10.23919/date58400.2024.10546864>
- Salloum, S. A., Shalan, K., Basiouni, A., Salloum, A., & Alfaisal, R. (2024). Building and Evaluating a Chatbot Using a University FAQs Dataset. *Communications in Computer and Information Science, 2162 CCIS*, 207–218. https://doi.org/10.1007/978-3-031-65996-6_18
- Sellam, T., Das, D., & Parikh, A. P. (2020). BLEURT: Learning robust metrics for text generation. *Proceedings of the Annual Meeting of the Association for Computational Linguistics*. <https://doi.org/https://doi.org/10.18653/v1/2020.acl-main.704>
- Shenoy, A., Y., S. R., Sharma, A., Rajan, A., & Gv, A. (2019). NLP models behind RASA stack. *International Journal of Engineering and Advanced Technology*, 8(5s), 62–66. <https://doi.org/https://doi.org/10.35940/ijeat.e1013.0585s19>
- Sonkar, A., Singh, S. P., Sahu, K., Sahu, A., & Mishra, S. (2025). Dynamic Query Handling with RAG Fusion for PDF-Based Knowledge Retrieval Systems. *2025 4th OPJU International Technology Conference on Smart Computing for Innovation and Advancement in Industry 5.0, OTCON 2025*. <https://doi.org/10.1109/OTCON65728.2025.11070378>
- Szmurlo, H., & Akhtar, Z. (2024). Digital Sentinels and Antagonists: The Dual Nature of Chatbots in Cybersecurity. *Information (Switzerland)*, 15(8). <https://doi.org/10.3390/info15080443>
- 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, 2017-December*, 5999–6009.
- Yang, J., Cai, Q., & Tian, X. (2020). How do we segment text? Two-stage chunking operation in reading. *eNeuro*, 7(3), 1–14. <https://doi.org/10.1523/ENEURO.0425-19.2020>

- Yin, B., Li, D., Xu, Z., & Hu, N. (2025). A Semantic-Structural Synergistic Chunking Method for Enhancing Domain-Specific LLM Fine-Tuning. *Lecture Notes in Computer Science*, 507–518. https://doi.org/10.1007/978-981-95-0020-8_43
- Zhang, H., & Shafiq, M. O. (2024). Survey of transformers and towards ensemble learning using transformers for natural language processing. *Journal of Big Data*, 11(1). <https://doi.org/10.1186/s40537-023-00842-0>
- Zhang, T., Kishore, V., Wu, F., Weinberger, K. Q., & Artzi, Y. (2020). Bertscore: Evaluating Text Generation With Bert. *8th International Conference on Learning Representations, ICLR 2020*.
- Zhang, T., Patterson, L., Beiting-Parrish, M., Webb, B., Abeysinghe, B., Bailey, P., & Sikali, E. (2024). Ask NAEP: A Generative AI Assistant for Querying Assessment Information. *Journal of Measurement and Evaluation in Education and Psychology*, 15, 378–394. <https://doi.org/10.21031/epod.1548128>
- Zhou, Q., Liu, C., Duan, Y., Sun, K., Li, Y., Kan, H., Gu, Z., Shu, J., & Hu, J. (2024). GastroBot: a Chinese gastrointestinal disease chatbot based on the retrieval-augmented generation. *frontiersin.org*, 11, 1392555. <https://doi.org/10.3389/FMED.2024.1392555/FULL>
- Zhu, J., Yan, L., Shi, H., Yin, D., & Sha, L. (2024). ATM: Adversarial Tuning Multi-agent System Makes a Robust Retrieval-Augmented GENERATOR. *EMNLP 2024 - 2024 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference*, 10902–10919. <https://doi.org/10.18653/v1/2024.emnlp-main.610>
- Zhu, Y., Yang, X., Wu, Y., Zhu, M., & Zhang, W. (2023). Differentiable N-gram objective on abstractive summarization. *Expert Systems with Applications*, 215, 119367. <https://doi.org/10.1016/J.ESWA.2022.119367>