

## DAFTAR PUSTAKA

- Ahlawat, S. (2023). Recurrent Neural Networks. Dalam S. Ahlawat, *Reinforcement Learning for Finance* (hlm. 177–232). Apress. [https://doi.org/10.1007/978-1-4842-8835-1\\_4](https://doi.org/10.1007/978-1-4842-8835-1_4)
- Aji, A. F., Winata, G. I., Koto, F., Cahyawijaya, S., Romadhony, A., Mahendra, R., Kurniawan, K., Moeljadi, D., Prasajo, R. E., Baldwin, T., Lau, J. H., & Ruder, S. (2022). One Country, 700+ Languages: NLP Challenges for Underrepresented Languages and Dialects in Indonesia. *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 7226–7249. <https://doi.org/10.18653/v1/2022.acl-long.500>
- Ali, A. H., Mohanad G. Yaseen, Mohammad Aljanabi, & Saad Abbas Abed. (2023). Transfer Learning: A New Promising Techniques. *Mesopotamian Journal of Big Data*, 2023, 29–30. <https://doi.org/10.58496/MJBD/2023/004>
- Ali, A., & Hussain, A. (2018). An Incremental Technique of Improving Translation. *International Journal of Advanced Computer Science and Applications*, 9(8). <https://doi.org/10.14569/IJACSA.2018.090860>
- Al-Wesabi, F. N., Alshahrani, H. J., Osman, A. E., & Abd Elhameed, E. S. (2023). Low-Resource Language Processing Using Improved Deep Learning with Hunter–Prey Optimization Algorithm. *Mathematics*, 11(21), 4493. <https://doi.org/10.3390/math11214493>
- Amorokhman, F. I., Romadhony, A., & Ihsan, A. F. (2023). Indonesian-Kailinese Machine Translation. *2023 International Conference on Data Science and Its Applications (ICoDSA)*, 397–402. <https://doi.org/10.1109/ICoDSA58501.2023.10277136>
- Arsyad, T. D., Berutu, I. A., Azis, K. R., & Chairunisa, H. (2025). *PELESTARIAN BUDAYA*

*LOKAL MELALUI DIGITALISASI.*

- Babych, B., & Hartley, A. (2021). Automated error analysis for multiword expressions: Using BLEU-type scores for automatic discovery of potential translation errors. *Linguistica Antverpiensia, New Series – Themes in Translation Studies*, 8.  
<https://doi.org/10.52034/lanstts.v8i.246>
- Bahdanau, D., Cho, K., & Bengio, Y. (2014). *Neural Machine Translation by Jointly Learning to Align and Translate* (Versi 7). arXiv. <https://doi.org/10.48550/ARXIV.1409.0473>
- Barrett, C., De Moura, L., Ranise, S., Stump, A., & Tinelli, C. (2011). The SMT-LIB Initiative and the Rise of SMT: (HVC 2010 Award Talk). Dalam S. Barner, I. Harris, D. Kroening, & O. Raz (Ed.), *Hardware and Software: Verification and Testing* (Vol. 6504, hlm. 3–3). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-19583-9\\_2](https://doi.org/10.1007/978-3-642-19583-9_2)
- Bengi, M., Amalia, A., & Budiman, M. A. (2024). A Transformer-Based Approach with OpenNMT for Indonesian to Gayo Language Translation. *2024 2nd International Symposium on Information Technology and Digital Innovation (ISITDI)*, 149–153.  
<https://doi.org/10.1109/ISITDI62380.2024.10797000>
- Biermann, H. (2024). Transfer Learning. Dalam D. Memmert (Ed.), *Computer Science in Sport* (hlm. 193–200). Springer Berlin Heidelberg.  
[https://doi.org/10.1007/978-3-662-68313-2\\_23](https://doi.org/10.1007/978-3-662-68313-2_23)
- Bonandrini, R., & Gatti, D. (2024). fastText (sub)word Vectors. Dalam *Reference Module in Social Sciences* (hlm. B9780323955041000326). Elsevier.  
<https://doi.org/10.1016/B978-0-323-95504-1.00032-6>
- Cahyawijaya, S., Winata, G. I., Wilie, B., Vincentio, K., Li, X., Kuncoro, A., Ruder, S., Lim, Z. Y., Bahar, S., Khodra, M., Purwarianti, A., & Fung, P. (2021). IndoNLG: Benchmark and

- Resources for Evaluating Indonesian Natural Language Generation. *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, 8875–8898.  
<https://doi.org/10.18653/v1/2021.emnlp-main.699>
- Chan, S. (Ed.). (2023). *Routledge encyclopedia of translation technology* (Second edition).  
Routledge, Taylor & Francis Group.
- Dabre, R., Nakagawa, T., & Kazawa, H. (2017). *An Empirical Study of Language Relatedness for Transfer Learning in Neural Machine Translation*. <https://aclanthology.org/Y17-1038/>
- De Silva, D. I., & Gallage, I. S. (2024). The Role of Syntax and Semantics in Rule-Based Translation: A Comprehensive Review. *2024 International Conference on Information and Communication Technology for Development for Africa (ICT4DA)*, 229–234.  
<https://doi.org/10.1109/ICT4DA62874.2024.10777198>
- Djoko Sutrisno. (2024). *Bahasa, budaya dan masyarakat*.
- Dou, Q., & Gales, M. (2022). *Parallel Attention Forcing for Machine Translation* (Versi 1).  
arXiv. <https://doi.org/10.48550/ARXIV.2211.03237>
- Farhansyah, M. R., Johari, M. Z. F., Amiral, A., Purwarianti, A., Yuana, K. A., & Wijaya, D. T. (2024). *DriveThru: A Document Extraction Platform and Benchmark Datasets for Indonesian Local Language Archives* (Versi 2). arXiv.  
<https://doi.org/10.48550/ARXIV.2411.09318>
- Fiddienika, A. & Inriani. (2024). *ANCAMAN PERGESERAN BAHASA DAERAH DALAM ERA GLOBALISASI: TINJAUAN KASUS DI KABUPATEN BARRU*.
- Haque, R., Liu, C.-H., & Way, A. (2021). Recent advances of low-resource neural machine translation. *Machine Translation*, 35(4), 451–474.  
<https://doi.org/10.1007/s10590-021-09281-1>

- Harefa, K. R., & Harefa, K. H. (2024). *PERAN BAHASA DALAM PEMBENTUKAN IDENTITAS BUDAYA DI INDONESIA. 01.*
- Henuarianto, I., Yadi, I. Z., Kunang, Y. N., & Purnamasari, S. D. (2023). Komerling—Indonesian Machine Translation Using Embedding RNN. *2023 IEEE 7th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, 163–167. <https://doi.org/10.1109/ICITISEE58992.2023.10404777>
- Heryadi, Y., Wijanarko, B. D., Fitria Murad, D., Tho, C., & Hashimoto, K. (2023). Indonesian-Sundanese Language Machine Translation using Bidirectional Long Short-term Memory Model. *2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE)*, 945–950. <https://doi.org/10.1109/ICCoSITE57641.2023.10127691>
- Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Johnson, M., Schuster, M., Le, Q. V., Krikun, M., Wu, Y., Chen, Z., Thorat, N., Viégas, F., Wattenberg, M., Corrado, G., Hughes, M., & Dean, J. (2017). Google’s Multilingual Neural Machine Translation System: Enabling Zero-Shot Translation. *Transactions of the Association for Computational Linguistics*, 5, 339–351. [https://doi.org/10.1162/tacl\\_a\\_00065](https://doi.org/10.1162/tacl_a_00065)
- Khaikal, M. F., & Suryani, A. A. (2021). Statistical Machine Translation Dayak Language – Indonesia Language. *Informatika Mulawarman : Jurnal Ilmiah Ilmu Komputer*, 16(1), 49. <https://doi.org/10.30872/jim.v16i1.5315>
- Khairunnisa, S. O., Chen, Z., & Komachi, M. (2023). Dataset Enhancement and Multilingual Transfer for Named Entity Recognition in the Indonesian Language. *ACM Transactions*

- on Asian and Low-Resource Language Information Processing*, 22(6), 1–21.  
<https://doi.org/10.1145/3592854>
- Klein, G., Kim, Y., Deng, Y., Senellart, J., & Rush, A. (2017). OpenNMT: Open-Source Toolkit for Neural Machine Translation. *Proceedings of ACL 2017, System Demonstrations*, 67–72. <https://doi.org/10.18653/v1/P17-4012>
- Kondrak, G. (2005). N-Gram Similarity and Distance. Dalam M. Consens & G. Navarro (Ed.), *String Processing and Information Retrieval* (Vol. 3772, hlm. 115–126). Springer Berlin Heidelberg. [https://doi.org/10.1007/11575832\\_13](https://doi.org/10.1007/11575832_13)
- Kudo, T., & Richardson, J. (2018). SentencePiece: A simple and language independent subword tokenizer and detokenizer for Neural Text Processing. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, 66–71. <https://doi.org/10.18653/v1/D18-2012>
- Lakew, S. M., Cettolo, M., & Federico, M. (2018). *A Comparison of Transformer and Recurrent Neural Networks on Multilingual Neural Machine Translation* (Versi 2). arXiv. <https://doi.org/10.48550/ARXIV.1806.06957>
- Lakew, S. M., Federico, M., Negri, M., & Turchi, M. (2018). Multilingual Neural Machine Translation for Low-Resource Languages. *Italian Journal of Computational Linguistics*, 4(1), 11–25. <https://doi.org/10.4000/ijcol.531>
- Lankford, S., Afli, H., & Way, A. (2024). *Transformers for Low-Resource Languages: Is Féidir Linn!* <https://doi.org/10.48550/ARXIV.2403.01985>
- Levenshtein, V. I. (1966). Binary Codes Capable of Correcting Deletions, Insertions and Reversals. *Soviet Physics Doklady*, 10, 707.
- Lin, C.-Y. (2004). *ROUGE: A Package for Automatic Evaluation of Summaries*.

- Lin, W. (2024). Neural Machine Translation Based on Multi Translation Parallel Corpus. *2024 3rd International Conference on Data Analytics, Computing and Artificial Intelligence (ICDACAI)*, 69–73. <https://doi.org/10.1109/ICDACAI65086.2024.00020>
- Liu Jianwei, Liu Junwen, & Luo Xionglin. (2021). *Research Progress on Attention Mechanism in Deep Learning—CNKI*. <https://dx.doi.org/10.13374/j.issn2095-9389.2021.01.30.005>
- Liu, Y., Gu, J., Goyal, N., Li, X., Edunov, S., Ghazvininejad, M., Lewis, M., & Zettlemoyer, L. (2020). *Multilingual Denoising Pre-training for Neural Machine Translation* (No. arXiv:2001.08210). arXiv. <https://doi.org/10.48550/arXiv.2001.08210>
- Lopo, J. A., & Tanone, R. (2024). *Constructing and Expanding Low-Resource and Underrepresented Parallel Datasets for Indonesian Local Languages* (Versi 1). arXiv. <https://doi.org/10.48550/ARXIV.2404.01009>
- Mahendra, R., Aji, A. F., Louvan, S., Rahman, F., & Vania, C. (2021). IndoNLI: A Natural Language Inference Dataset for Indonesian. *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, 10511–10527. <https://doi.org/10.18653/v1/2021.emnlp-main.821>
- Mallick, R., Susan, S., Agrawal, V., Garg, R., & Rawal, P. (2021). Context- and sequence-aware convolutional recurrent encoder for neural machine translation. *Proceedings of the 36th Annual ACM Symposium on Applied Computing*, 853–856. <https://doi.org/10.1145/3412841.3442099>
- Mazi, C. C., Anya, O. I., Nwanakwaugwu, A. C., & Anichebe, G. (2024). Machine Translation and Natural Language Processing. *International Journal of Science and Management Studies (IJSMS)*, 65–69. <https://doi.org/10.51386/25815946/ijsms-v7i1p110>
- Mohapatra, S., & Mohapatra, S. (2022). *The (In)Effectiveness of Intermediate Task Training For*

- Domain Adaptation and Cross-Lingual Transfer Learning* (Versi 2). arXiv.  
<https://doi.org/10.48550/ARXIV.2210.01091>
- Naveen, P., & Trojovský, P. (2024). Overview and challenges of machine translation for contextually appropriate translations. *iScience*, 27(10), 110878.  
<https://doi.org/10.1016/j.isci.2024.110878>
- Neubig, G., & Hu, J. (2018). Rapid Adaptation of Neural Machine Translation to New Languages. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 875–880. <https://doi.org/10.18653/v1/D18-1103>
- Nguyen, T. Q., & Chiang, D. (2017). *Transfer Learning across Low-Resource, Related Languages for Neural Machine Translation* (Versi 2). arXiv.  
<https://doi.org/10.48550/ARXIV.1708.09803>
- Östling, R., & Tiedemann, J. (2016). Efficient Word Alignment with Markov Chain Monte Carlo. *The Prague Bulletin of Mathematical Linguistics*, 106(1), 125–146.  
<https://doi.org/10.1515/pralin-2016-0013>
- Ott, M., Edunov, S., Baevski, A., Fan, A., Gross, S., Ng, N., Grangier, D., & Auli, M. (2019). fairseq: A Fast, Extensible Toolkit for Sequence Modeling. *Proceedings of the 2019 Conference of the North*, 48–53. <https://doi.org/10.18653/v1/N19-4009>
- Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2001). BLEU: A method for automatic evaluation of machine translation. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics - ACL '02*, 311.  
<https://doi.org/10.3115/1073083.1073135>
- Paul, M., Finch, A., & Sumita, E. (2013). How to Choose the Best Pivot Language for Automatic Translation of Low-Resource Languages. *ACM Transactions on Asian Language*

- Information Processing*, 12(4), 1–17. <https://doi.org/10.1145/2505126>
- Popović, M. (2015). chrF: Character n-gram F-score for automatic MT evaluation. *Proceedings of the Tenth Workshop on Statistical Machine Translation*, 392–395.  
<https://doi.org/10.18653/v1/W15-3049>
- Puram, V., Bobbili, R. R., & Thomas, J. P. (2024). *Quantum Algorithm for Jaccard Similarity* (Versi 1). arXiv. <https://doi.org/10.48550/ARXIV.2408.08940>
- Rukua, A. W., Lesnussa, Y. A., Rahakbauw, D. L., & Tomasouw, B. P. (2023). Application of Neural Machine Translation with Attention Mechanism for Translation of Indonesian to Seram Language (Geser). *Pattimura International Journal of Mathematics (PIJMath)*, 2(2), 53–62. <https://doi.org/10.30598/pijmathvol2iss2pp53-62>
- Sewell, Y. (2022). Linguistic pragmatism, lingua francae, and language death in Indonesia. *Journal of Language Teaching*, 2(11), 15–19. <https://doi.org/10.54475/jlt.2022.015>
- Sihite, M. R., & Sibarani, B. (2024). Technology and Language Revitalization in Indonesia: A Literature Review of Digital Tools for Preserving Endangered Languages. *International Journal of Educational Research Excellence (IJERE)*, 3(2), 610–620.  
<https://doi.org/10.55299/ijere.v3i2.988>
- Son, T. N., Tu, N. A., & Tri, N. M. (2025). *An Efficient Approach for Machine Translation on Low-resource Languages: A Case Study in Vietnamese-Chinese* (Versi 1). arXiv.  
<https://doi.org/10.48550/ARXIV.2501.19314>
- Suhassatya, G. K. (2025). *KEPUNAHAN BAHASA DAERAH DI INDONESIA DALAM PERSPEKTIF BAHASA SEBAGAI IDENTITAS BUDAYA*. 4.
- Sulistyo, D. A. (2023). LSTM-Based Machine Translation for Madurese-Indonesian. *Journal of Applied Data Sciences*, 4(3), 189–199. <https://doi.org/10.47738/jads.v4i3.113>

- Tan, P.-N., Steinbach, M., & Kumar, V. (2014). *Introduction to data mining* (New international edition). Pearson.
- Tang, Y., Tran, C., Li, X., Chen, P.-J., Goyal, N., Chaudhary, V., Gu, J., & Fan, A. (2020). *Multilingual Translation with Extensible Multilingual Pretraining and Finetuning* (No. arXiv:2008.00401). arXiv. <https://doi.org/10.48550/arXiv.2008.00401>
- Tang, Y., Tran, C., Li, X., Chen, P.-J., Goyal, N., Chaudhary, V., Gu, J., & Fan, A. (2021). Multilingual Translation from Denoising Pre-Training. *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, 3450–3466. <https://doi.org/10.18653/v1/2021.findings-acl.304>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention Is All You Need* (Versi 7). arXiv. <https://doi.org/10.48550/ARXIV.1706.03762>
- Winata, G. I., Aji, A. F., Cahyawijaya, S., Mahendra, R., Koto, F., Romadhony, A., Kurniawan, K., Moeljadi, D., Prasajo, R. E., Fung, P., Baldwin, T., Lau, J. H., Sennrich, R., & Ruder, S. (2023). *NusaX: Multilingual Parallel Sentiment Dataset for 10 Indonesian Local Languages* (No. arXiv:2205.15960). arXiv. <https://doi.org/10.48550/arXiv.2205.15960>
- Wongso, W., Joyoadikusumo, A., Buana, B. S., & Suhartono, D. (2023). Many-to-Many Multilingual Translation Model for Languages of Indonesia. *IEEE Access*, 11, 91385–91397. <https://doi.org/10.1109/ACCESS.2023.3308818>
- Wu, T.-W., Zhao, C., Chang, E., Shi, Y., Chuang, P., Chandra, V., & Juang, B. (2023). Towards Zero-Shot Multilingual Transfer for Code-Switched Responses. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 7551–7563. <https://doi.org/10.18653/v1/2023.acl-long.417>

- Xia, F., He, S., Liu, K., Liu, S., & Zhao, J. (2021). Toward a Better Text Data Augmentation via Filtering and Transforming Augmented Instances. Dalam B. Qin, Z. Jin, H. Wang, J. Pan, Y. Liu, & B. An (Ed.), *Knowledge Graph and Semantic Computing: Knowledge Graph Empowers New Infrastructure Construction* (Vol. 1466, hlm. 198–210). Springer Singapore. [https://doi.org/10.1007/978-981-16-6471-7\\_15](https://doi.org/10.1007/978-981-16-6471-7_15)
- Yamin, M., Sarno, R., & Tambunan, T. (2023). Enhancing machine translation: Syntax and semantics-based word type and function extraction through multi-task transfer learning in Indonesian, Tolaki, and English. *Periodicals of Engineering and Natural Sciences (PEN)*, 12(1), 223–235. <https://doi.org/10.21533/pen.v12.i1.32>
- Yang, M., Zhu, J., Li, J., Wang, L., Qi, H., Li, S., & Daxin, L. (2008). Extending BLEU Evaluation Method with Linguistic Weight. *2008 The 9th International Conference for Young Computer Scientists*, 1683–1688. <https://doi.org/10.1109/ICYCS.2008.362>
- Yumnam, G., & Singh, C. I. (2024). A Bibliometric Perspective of Regional Languages on Select Scholarly Articles. *DESIDOC Journal of Library & Information Technology*, 44(1), 37–44. <https://doi.org/10.14429/djlit.44.1.18938>
- Zein, S. (2020). *Language Policy in Superdiverse Indonesia* (1 ed.). Routledge. <https://doi.org/10.4324/9780429019739>
- Zhou, S., Shan, H., Plank, B., & Litschko, R. (2024). *MaiNLP at SemEval-2024 Task 1: Analyzing Source Language Selection in Cross-Lingual Textual Relatedness* (Versi 1). arXiv. <https://doi.org/10.48550/ARXIV.2404.02570>
- Zhou, Z. (2024). *Massively Multilingual Text Translation For Low-Resource Languages* (Versi 1). arXiv. <https://doi.org/10.48550/ARXIV.2401.16582>
- Zoph, B., Yuret, D., May, J., & Knight, K. (2016). Transfer Learning for Low-Resource Neural

Machine Translation. *Proceedings of the 2016 Conference on Empirical Methods in*

*Natural Language Processing*, 1568–1575. <https://doi.org/10.18653/v1/D16-1163>

*Transfer learning*. (t.t.). Diakses pada 19 Juni 2025, dari

[https://en.wikipedia.org/wiki/Transfer\\_learning](https://en.wikipedia.org/wiki/Transfer_learning)