

DAFTAR PUSTAKA

- Anindyatri, A. O. and Mufidah, I. (2020). *Gambaran kondisi vitalitas bahasa daerah di Indonesia: berdasarkan data Tahun 2018 – 2019*. Pusat Data Informasi dan Teknologi, Kementerian Pendidikan dan Kebudayaan, Jakarta.
- Aranta, A., Djunaidy, A., and Suciati, N. (2024). Preserving sasak dialectal features in english to sasak machine translation through locked tokenization with transformer models. In *2024 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, pages 19–24.
- Ba, J. L., Kiros, J. R., and Hinton, G. E. (2016). Layer normalization.
- Bahdanau, D., Cho, K., and Bengio, Y. (2016). Neural Machine Translation by Jointly Learning to Align and Translate. arXiv:1409.0473.
- Bird, S. and Chiang, D. (2012). Machine translation for language preservation. In Kay, M. and Boitet, C., editors, *Proceedings of COLING 2012: Posters*, pages 125–134, Mumbai, India. The COLING 2012 Organizing Committee.
- BPS (2023). Hasil Long Form Sensus Penduduk 2020.
- Brown, R. (2002). Example-based machine translation. In Richardson, S. D., editor, *Proceedings of the 5th Conference of the Association for Machine Translation in the Americas: Tutorial Descriptions*, Tiburon, USA. Springer.
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., and Amodei, D. (2020). Language models are few-shot learners. In *Proceedings of the 34th International Conference on Neural Information Processing Systems, NIPS '20*, Red Hook, NY, USA. Curran Associates Inc.
- Cahyawijaya, S., Winata, G. I., Wilie, B., Vincentio, K., Li, X., Kuncoro, A., Ruder, S., Lim, Z. Y., Bahar, S., Khodra, M., Purwarianti, A., and Fung, P. (2021). In-doNLG: Benchmark and Resources for Evaluating Indonesian Natural Language

- Generation. In Moens, M.-F., Huang, X., Specia, L., and Yih, S. W.-t., editors, *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 8875–8898, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.
- Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Burstein, J., Doran, C., and Solorio, T., editors, *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)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.
- Do, P. H., Quan, P. V., Duc Le, T., Duy Dinh, T., Nang Hung Van, N., and Tuan Pham, M. (2023). Unveiling the Power of Pretrained Models for Neural Machine Translation in Vietnamese Language: A Comparative Analysis. In *2023 RIVF International Conference on Computing and Communication Technologies (RIVF)*, pages 452–457. ISSN: 2473-0130.
- Fauziyah, Y., Ilyas, R., and Kasyidi, F. (2022). Mesin Penterjemah Bahasa Indonesia-Bahasa Sunda Menggunakan Recurrent Neural Networks. *Jurnal Teknoinfo*, 16(2):313–322. Number: 2.
- Gheini, M., Ren, X., and May, J. (2021). Cross-Attention is All You Need: Adapting Pretrained Transformers for Machine Translation. In Moens, M.-F., Huang, X., Specia, L., and Yih, S. W.-t., editors, *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 1754–1765, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.
- Ian Goodfellow, Y. B. and Courville, A. (2016). *Deep Learning*. The MIT Press, Massachutes.
- Johan, P. (2016). MESIN PENERJEMAH BAHASA INDONESIA-JAWA MENGGUNAKAN METODE STATISTIK BERBASIS FRASA. *Skripsi, Fakultas Ilmu Komputer*.
- Joulin, A., Grave, E., Bojanowski, P., and Mikolov, T. (2017). Bag of tricks for efficient text classification. In Lapata, M., Blunsom, P., and Koller, A., editors, *Proceedings of the 15th Conference of the European Chapter of the Association*

- for Computational Linguistics: Volume 2, Short Papers*, pages 427–431, Valencia, Spain. Association for Computational Linguistics.
- Kalchbrenner, N. and Blunsom, P. (2013). Recurrent Continuous Translation Models. In Yarowsky, D., Baldwin, T., Korhonen, A., Livescu, K., and Bethard, S., editors, *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 1700–1709, Seattle, Washington, USA. Association for Computational Linguistics.
- Kingma, D. P. and Ba, J. (2017). Adam: A method for stochastic optimization.
- Koehn, P. (2009). *Statistical Machine Translation*. Cambridge University Press, Cambridge.
- Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Commun. ACM*, 60(6):84–90.
- Kudo, T. (2018). Subword regularization: Improving neural network translation models with multiple subword candidates. In Gurevych, I. and Miyao, Y., editors, *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 66–75, Melbourne, Australia. Association for Computational Linguistics.
- Kudo, T. and Richardson, J. (2018). SentencePiece: A simple and language independent subword tokenizer and detokenizer for neural text processing. In Blanco, E. and Lu, W., editors, *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 66–71, Brussels, Belgium. Association for Computational Linguistics.
- Lesatari, A. E. P., Ardiyanti, A., and Asror, I. (2021). Phrase Based Statistical Machine Translation Javanese-Indonesian. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(2):378–386. Number: 2.
- Mikolov, T., Chen, K., Corrado, G. S., and Dean, J. (2013). Efficient estimation of word representations in vector space. In *International Conference on Learning Representations*.
- Miranda, B. O. S., Yuliansyah, H., and Biddinika, M. K. (2024). Machine Translation Indonesian Bengkulu Malay Using Neural Machine Translation-LSTM. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 18(3). Number: 3.

- Miyagawa, S. (2023). Machine Translation for Highly Low-Resource Language: A Case Study of Ainu, a Critically Endangered Indigenous Language in Northern Japan. In Hämäläinen, M., Öhman, E., Pirinen, F., Alnajjar, K., Miyagawa, S., Bizzoni, Y., Partanen, N., and Rueter, J., editors, *Proceedings of the Joint 3rd International Conference on Natural Language Processing for Digital Humanities and 8th International Workshop on Computational Linguistics for Uralic Languages*, pages 120–124, Tokyo, Japan. Association for Computational Linguistics.
- Nirenburg, S. (1989). Knowledge-based machine translation. *Machine Translation*, 4(1):5–24.
- Ogunfunmi, T., Ramachandran, R. P., Togneri, R., Zhao, Y., and Xia, X. (2019). A Primer on Deep Learning Architectures and Applications in Speech Processing. *Circuits, Systems, and Signal Processing*, 38(8):3406–3432.
- Papineni, K., Roukos, S., Ward, T., and Zhu, W.-J. (2002). Bleu: a method for automatic evaluation of machine translation. In Isabelle, P., Charniak, E., and Lin, D., editors, *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 311–318, Philadelphia, Pennsylvania, USA. Association for Computational Linguistics.
- Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., and Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, 21(140):1–67.
- Sholikhatin, S. A., Magnolia, C., and Putra, R. D. M. (2021). LOCAL LANGUAGE KEYPAD: KEYBOARD BAHASA DAERAH BERBASIS SMARTPHONE SEBAGAI MEDIA PELESTARIAN BAHASA DAN SASTRA DAERAH. *Jurnal RESISTOR (Rekayasa Sistem Komputer)*, 4(2):186–191.
- Suharyo, S. (2018). Nasib bahasa jawa dan bahasa indonesia dalam pandangan dan sikap bahasa generasi muda jawa. *Nusa: Jurnal Ilmu Bahasa dan Sastra*, 13(2):244–255.
- Sutomo, S. D. (2022). Evaluasi Pengaruh Mekanisme Attention Terhadap Model Mesin Penerjemah Bahasa Jawa-Indonesia Berbasis Neural Machine Translation. Skripsi, Universitas Gadjah Mada.

- Sutskever, I., Vinyals, O., and Le, Q. V. (2014). Sequence to sequence learning with neural networks. In *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2, NIPS'14*, pages 3104–3112, Cambridge, MA, USA. MIT Press.
- Tu, Z., Liu, Y., Lu, Z., Liu, X., and Li, H. (2017). Context gates for neural machine translation. *Transactions of the Association for Computational Linguistics*, 5:87–99.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2017). Attention is all you need. In *Proceedings of the 31st International Conference on Neural Information Processing Systems, NIPS'17*, page 6000–6010, Red Hook, NY, USA. Curran Associates Inc.
- Wang, H., Wu, H., He, Z., Huang, L., and Church, K. W. (2022). Progress in machine translation. *Engineering*, 18:143–153.
- Wibawa, A. P., Nafalski, A., Kadarisman, A. E., and Mahmudy, W. F. (2013). Indonesian-to-Javanese Machine Translation. *International journal of innovation, management and technology*.
- Wilie, B., Vincentio, K., Winata, G. I., Cahyawijaya, S., Li, X., Lim, Z. Y., Soleman, S., Mahendra, R., Fung, P., Bahar, S., and Purwarianti, A. (2020). IndoNLU: Benchmark and resources for evaluating Indonesian natural language understanding. In Wong, K.-F., Knight, K., and Wu, H., editors, *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*, pages 843–857, Suzhou, China. Association for Computational Linguistics.
- Wongso, W., Joyoadikusumo, A., Buana, B. S., and Suhartono, D. (2023). Many-to-Many Multilingual Translation Model for Languages of Indonesia. *IEEE Access*, 11:91385–91397. Conference Name: IEEE Access.
- Wongso, W., Setiawan, D. S., and Suhartono, D. (2021). Causal and Masked Language Modeling of Javanese Language using Transformer-based Architectures. In *2021 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, pages 1–7. ISSN: 2330-4588.



Yati, D. (2015). MENYELAMATKAN BAHASA DAERAH MELALUI PEMBE-
LAJARAN BAHASA YANG KOMUNIKATIF. In *Seminar Nasional Bulan Baha-
sa 2015*, pages 157–170. Unit Penerbitan FKIP Universitas Bengkulu.