



DAFTAR PUSTAKA

- [1] I. Sutskever, O. Vinyals, and Q. V. Le, “Sequence to sequence learning with neural networks,” in *Advances in Neural Information Processing Systems 27 (NIPS 2014)*, 2014, pp. 3104–3112.
- [2] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, “Attention is all you need,” in *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, 2017, pp. 5998–6008.
- [3] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, and W. Chen, “LoRA: Low-rank adaptation of large language models,” in *International Conference on Learning Representations (ICLR)*, 2022.
- [4] T. Zhang, V. Kishore, F. Wu, K. Q. Weinberger, and Y. Artzi, “Bertscore: Evaluating text generation with bert,” in *International Conference on Learning Representations (ICLR)*, 2020.
- [5] D. Bahdanau, K. Cho, and Y. Bengio, “Neural machine translation by jointly learning to align and translate,” in *International Conference on Learning Representations (ICLR)*, 2015.
- [6] H. A. Wibowo, A. F. Aji, T. A. Prawiro, R. E. Prasajo, M. Ihsan, and R. Mahendra, “Semi-supervised low-resource style transfer of indonesian informal to formal language with iterative forward-translation,” *arXiv preprint arXiv:2011.03286*, 2020.
- [7] A. Aranta, A. Djunaidy, and N. Suciati, “Developing an english-to-indonesian speech-to-text as a foundation for sasak language translation using the mbart algorithm,” in *International Conference on Green Energy, Computing and Intelligent Technology 2024 (GEN-CITY 2024)*, vol. 2024, 2024, pp. 568–573.
- [8] B. R. Irnawan and R. Adi, “Improving indonesian informal to formal style transfer via pre-training unlabelled augmented data,” in *2023 6th International Conference of Computer and Informatics Engineering (IC2IE)*, 2023, pp. 25–29.
- [9] A. M. Barik, R. Mahendra, and M. Adriani, “Normalization of Indonesian-English code-mixed Twitter data,” in *Proceedings of the 5th Workshop on Noisy User-generated Text (W-NUT 2019)*, W. Xu, A. Ritter, T. Baldwin, and A. Rahimi, Eds. Hong Kong, China: Association for Computational Linguistics, Nov. 2019, pp. 417–424. [Online]. Available: <https://aclanthology.org/D19-5554/>
- [10] A. S. Ariesandy, M. Amien, A. F. Aji, and R. E. Prasajo, “Synthetic source language augmentation for colloquial neural machine translation,” *arXiv preprint arXiv:2012.15178*, 2020.
- [11] Llama Team, AI @ Meta, “The Llama 3 herd of models,” Meta, Tech. Rep. arXiv:2407.21783, July 2024.
- [12] H. Touvron, T. Lavril, G. Izacard, X. Martinet, M.-A. Lachaux, T. Lacroix, B. Rozière, N. Goyal, E. Hambro, F. Azhar, A. Rodriguez, A. Joulin, E. Grave, and G. Lample, “Llama: Open and efficient foundation language models,” 2023. [Online]. Available: <https://arxiv.org/abs/2302.13971>



- [13] J. Hoffmann, S. Borgeaud, A. Mensch, E. Buchatskaya, T. Cai, E. Rutherford, D. de Las Casas, L. A. Hendricks, J. Welbl, A. Clark, T. Hennigan, E. Noland, K. Millican, G. van den Driessche, B. Damoc, A. Guy, S. Osindero, K. Simonyan, E. Elsen, J. W. Rae, O. Vinyals, and L. Sifre, “Training compute-optimal large language models,” 2022. [Online]. Available: <https://arxiv.org/abs/2203.15556>
- [14] B. Zhang and R. Sennrich, “Root mean square layer normalization,” 2019. [Online]. Available: <https://arxiv.org/abs/1910.07467>
- [15] N. Shazeer, “Glu variants improve transformer,” 2020. [Online]. Available: <https://arxiv.org/abs/2002.05202>
- [16] J. Su, Y. Lu, S. Pan, A. Murtadha, B. Wen, and Y. Liu, “Roformer: Enhanced transformer with rotary position embedding,” 2023. [Online]. Available: <https://arxiv.org/abs/2104.09864>
- [17] Y. Liu, J. Gu, N. Goyal, X. Li, S. Edunov, M. Ghazvininejad, M. Lewis, and L. Zettlemoyer, “Multilingual denoising pre-training for neural machine translation,” in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL)*. Association for Computational Linguistics, 2020, pp. 725–740.
- [18] Y. Tang, C. Tran, X. Li, P.-J. Chen, N. Goyal, V. Chaudhary, J. Gu, and A. Fan, “Multilingual translation with extensible multilingual pretraining and finetuning,” in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Association for Computational Linguistics, 2020, pp. 5244–5255.
- [19] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, “BLEU: a method for automatic evaluation of machine translation,” in *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)*, 2002, pp. 311–318.
- [20] M. Post, “A call for clarity in reporting BLEU scores,” in *Proceedings of the Third Conference on Machine Translation: Research Papers*. Association for Computational Linguistics, 2018, pp. 186–191.
- [21] H. A. Wibowo, M. N. Nityasya, S. Fitriany, A. F. Aji, R. E. Prasojo, A. F. Akyürek, and D. T. Wijaya, “Indocollex: A testbed for morphological transformation of Indonesian colloquial words,” in *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, August 2021, pp. 3170–3183.
- [22] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, “Language models are unsupervised multitask learners,” OpenAI, Tech. Rep., 2019, openAI Blog.