Daftar Pustaka

[1] A. F. Aji, R. E. P. Tirana Noor Fatyanosa, P. Arthur, S. Fitriany, S. Qonitah, N. Zulfa, T. Santoso, and M. Data. Paracotta: Synthetic multilingual paraphrase corpora from the most diverse translation sample pair. page 533–542. Proceedings of the 35th Pacific Asia Conference on Language, Information and Computation, Association for Computational Lingustics, 11 2021.

[2] A. F. Aji, G. I. Winata, F. Koto, S. Cahyawijaya, A. Romadhony, R. Mahendra, K. Kurniawan, D. Moeljadi, R. E. Prasojo, T. Baldwin, J. H. Lau, and S. Ruder. One country, 700+ languages: Nlp challenges for underrepresented languages and dialects in indonesia. arXiv:2203.13357 [cs], 03 2022.

[3] X. Chen, H. Fang, T.-Y. Lin, R. Vedantam, S. Gupta, P. Dollar, and C. L. Zitnick. Microsoft coco captions: Data collection and evaluation server. *arXiv:1504.00325 [cs]*, 04 2015.

[4] L. Deléger and P. Zweigenbaum. Paraphrase acquisition from comparable medical corpora of specialized and lay texts. volume 2008, page 146–150. AMIA Annual Symposium Proceedings, AMIA Symposium, 2008.

[5] L. Dong, J. Mallinson, S. Reddy, and M. Lapata. Learning to paraphrase for question answering. page 875–886. Proceedings of 2017 Empirical Methods in Natural Language Processing, Association for Computational Linguistics, 09 2017.

[6] A. Fader, L. Zettlemoyer, and O. Etzioni. Paraphrase-driven learning for open question answering. volume 1, page 1608–1618. Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, 08 2013.

[7] W. C. Gan and H. T. Ng. Improving the robustness of question answering systems to question paraphrasing. page 6065–6075. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, 07 2019.

[8] S. Gao, Y. Zhang, Z. Ou, and Z. Yu. Paraphrase augmented task-oriented dialog generation. page 639–649. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, 07 2020.

[9] J. E. Hu, R. Rudinger, M. Post, and B. Van Durme. Parabank: Monolingual bitext generation and sentential paraphrasing via lexically-constrained neural machine translation. *arXiv:1901.03644 [cs]*, 01 2019.

[10] J. E. Hu, A. Singh, N. Holzenberger, M. Post, and B. Van Durme. Large-scale, diverse, paraphrastic bitexts via sampling and clustering. page 44–54. Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL), Association for Computational Linguistics, 11 2019.

[11] S. Iyer, N. Dandekar, and K. Csernai. First quora dataset release: Question pairs, 2012.

[12] X. Jia, W. Zhou, X. Sun, and Y. Wu. How to ask good questions? try to leverage paraphrases. page 6130–6140. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, 07 2020.

[13] M. Klemen and M. Robnik-Šikonja. Extracting and filtering paraphrases by bridging natural language inference and paraphrasing. *arXiv:2111.07119 [cs]*, 11 2021.

[14] F. Koto, J. H. Lau, and T. Baldwin. Liputan6: A large-scale indonesian dataset for text summarization. page 598–608. 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, Association for Computational Linguistics, 12 2020.

[15] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin. IndoLEM and IndoBERT: A benchmark dataset and pre-trained language model for Indonesian NLP. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 757–770, Barcelona, Spain (Online), Dec. 2020. International Committee on Computational Linguistics.

[16] A. Kumar, S. Bhattamishra, M. Bhandari, and P. Talukdar. Submodular optimization-based diverse paraphrasing and its effectiveness in data augmentation. volume 1, page 3609–3619. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Association for Computational Linguistics, 06 2019.

[17] W. Lan, S. Qiu, H. He, and W. Xu. A continuously growing dataset of sentential paraphrases. page 1224–1234. Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, 09 2017.

[18] R. Mahendra, A. F. Aji, S. Louvan, F. Rahman, and C. Vania. Indonli: A natural language inference dataset for indonesian. page 10511–10527. Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, 11 2021.

[19] J. Mallinson, R. Sennrich, and M. Lapata. Paraphrasing revisited with neural machine translation. volume 1, page 881–893. Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics, Association for Computational Linguistics, 04 2017.

[20] R. Mehdizadeh Seraj, M. Siahbani, and A. Sarkar. Improving statistical machine translation with a multilingual paraphrase database. page 1379–1390. Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, 09 2015.

[21] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu. Bleu: a method for automatic evaluation of machine translation. Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, 2001.

[22] M. Post. A call for clarity in reporting BLEU scores. In *Proceedings of the Third Conference on Machine Translation: Research Papers*, pages 186–191, Brussels, Belgium, Oct. 2018. Association for Computational Linguistics.

[23] S. Rothe, S. Narayan, and A. Severyn. Leveraging pre-trained checkpoints for sequence generation tasks. volume 8, pages 264–280. Transactions of the Association for Computational Linguistics, 07 2020.

[24] B. Thompson and M. Post. Automatic machine translation evaluation in many languages via zero-shot paraphrasing. page 90–121. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), Association for Computational Linguistics, 11 2020.

[25] Y.-M. Tseng, Y.-T. Huang, M. C. Chen, and Y. S. Sun. Generating comprehension questions using paraphrase. volume 8916, pages 310–321. Technologies and Applications of Artificial Intelligence, Lecture Notes in Computer Science, 2014.

[26] J. Wieting and K. Gimpel. Paranmt-50m: Pushing the limits of paraphrastic sentence embeddings with millions of machine translations. volume 1, page 451–462. Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, 07 2018.

[27] B. Wilie, K. Vincentio, G. I. Winata, S. Cahyawijaya, X. Li, Z. Y. Lim, S. Soleman, R. Mahendra, P. Fung, S. Bahar, and A. Purwarianti. Indonlu: Benchmark and resources for evaluating indonesian natural language understanding. page 843–857. 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, Association for Computational Linguistics, 12 2020.

[28] A. W. Yu, D. Dohan, M.-T. Luong, R. Zhao, K. Chen, M. Norouzi, and Q. V. Le. Qanet: Combining local convolution with global self-attention for reading comprehension. Proceedings of 6th International Conference on Learning Representations, 02 2018.

[29] T. Zhang, V. Kishore, F. Wu, K. Q. Weinberger, and Y. Artzi. Bertscore: Evaluating text generation with bert. *arXiv:1904.09675 [cs]*, 02 2020.

[30] J. Zhou and S. Bhat. Paraphrase generation: A survey of the state of the art. page 5075–5086. Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, 11 2021.

[31] S. Zhu, X. Cheng, S. Su, and S. Lang. Knowledge-based question answering by jointly generating, copying and paraphrasing. page 2439–2442. Proceedings of the 2017 ACM on Conference on Information and Knowledge Management (CIKM '17), Association for Computing Machinery, 11 2017.