DAFTAR PUSTAKA

[1] D. Ramesh and S. K. Sanampudi, "An automated essay scoring systems: a systematic literature review," *Artif Intell Rev*, vol. 55, no. 3, pp. 2495–2527, Mar. 2022, doi: 10.1007/s10462-021-10068-2.

[2] H. Abdeljaber, "Automatic Arabic Short Answers Scoring Using Longest Common Subsequence and Arabic WordNet," *IEEE Access*, vol. PP, pp. 1–1, Jan. 2021, doi: 10.1109/ACCESS.2021.3082408.

[3] J. Wang and Y. Dong, "Measurement of Text Similarity: A Survey," *Information*, vol. 11, no. 9, Art. no. 9, Sep. 2020, doi: 10.3390/info11090421.

[4] M. M. Deza and E. Deza, *Encyclopedia of Distances*. Springer Science & Business Media, 2009.

[5] D. Prakoso, A. Abdi, and C. Amrit, "Short text similarity measurement methods: a review," *Soft Computing*, vol. 25, pp. 1–25, Mar. 2021, doi: 10.1007/s00500-020-05479-2.

[6] A.-R. Muhammad, A. E. Permanasari, and I. Hidayah, "Enhancing GAN-LCS Performance Using an Abbreviations Checker in Automatic Short Answer Scoring," *Computers*, vol. 11, no. 7, Art. no. 7, Jul. 2022, doi: 10.3390/computers11070108.

[7] J. Zhou and S. Bhat, "Paraphrase Generation: A Survey of the State of the Art," in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, Online and Punta Cana, Dominican Republic: Association for Computational Linguistics, Nov. 2021, pp. 5075–5086. doi: 10.18653/v1/2021.emnlp-main.414.

[8] F. S. Pribadi, A. E. Permanasari, and T. B. Adji, "Short answer scoring system using automatic reference answer generation and geometric average normalized-longest common subsequence (GAN-LCS)," *Educ Inf Technol*, vol. 23, no. 6, pp. 2855–2866, Nov. 2018, doi: 10.1007/s10639-018-9745-z.

[9] U. Hasanah, A. E. Permanasari, S. S. Kusumawardani, and F. S. Pribadi, "A scoring rubric for automatic short answer grading system," *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 17, no. 2, pp. 763–770, 2019.

[10] F. F. Lubis *et al.*, "Automated Short-Answer Grading using Semantic Similarity based on Word Embedding," *IJTech*, vol. 12, no. 3, p. 571, Jul. 2021, doi: 10.14716/ijtech.v12i3.4651.

[11] R. A. Rajagede, "Improving Automatic Essay Scoring for Indonesian Language using Simpler Model and Richer Feature," *Kinetik Game Technology Information System Computer Network Computing Electronics and Control*, vol. 6, pp. 11–18, Feb. 2021, doi: 10.22219/kinetik.v6i1.1196.

[12] K. Yang *et al.*, "GCPG: A General Framework for Controllable Paraphrase Generation," in *Findings of the Association for Computational Linguistics: ACL 2022*, Dublin, Ireland: Association for Computational

Linguistics, May 2022, pp. 4035–4047. doi: 10.18653/v1/2022.findings-acl.318.

[13] T.-C. Bui, V.-D. Le, H.-T. To, and S. K. Cha, "Generative Pre-training for Paraphrase Generation by Representing and Predicting Spans in Exemplars," in *2021 IEEE International Conference on Big Data and Smart Computing (BigComp)*, Jan. 2021, pp. 83–90. doi: 10.1109/BigComp51126.2021.00025.

[14] A. Kumar, K. Ahuja, R. Vadapalli, and P. Talukdar, "Syntax-Guided Controlled Generation of Paraphrases," *Transactions of the Association for Computational Linguistics*, vol. 8, pp. 329–345, 2020, doi: 10.1162/tacl_a_00318.

[15] J. Sun, X. Ma, and N. Peng, "AESOP: Paraphrase Generation with Adaptive Syntactic Control," in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, Online and Punta Cana, Dominican Republic: Association for Computational Linguistics, Nov. 2021, pp. 5176–5189. doi: 10.18653/v1/2021.emnlp-main.420.

[16] L. Weng, "From GAN to WGAN." arXiv, Apr. 18, 2019. doi: 10.48550/arXiv.1904.08994.

[17] X. Chen, S. Jia, and Y. Xiang, "A review: Knowledge reasoning over knowledge graph," *Expert Systems with Applications*, vol. 141, p. 112948, Sep. 2019, doi: 10.1016/j.eswa.2019.112948.

[18] R. T. Wahyuni, D. Prastiyanto, and E. Supraptono, "Penerapan Algoritma Cosine Similarity dan Pembobotan TF-IDF pada Sistem Klasifikasi Dokumen Skripsi," *Jurnal Teknik Elektro*, vol. 9, no. 1, Art. no. 1, 2017, doi: 10.15294/jte.v9i1.10955.

[19] A. K. Patidar, J. Agrawal, and N. Mishra, "Analysis of different similarity measure functions and their impacts on shared nearest neighbor clustering approach," *International Journal of Computer Applications*, vol. 40, no. 16, pp. 1–5, 2012.

[20] A. R. Lahitani, A. E. Permanasari, and N. A. Setiawan, "Cosine similarity to determine similarity measure: Study case in online essay assessment," in *2016 4th International Conference on Cyber and IT Service Management*, IEEE, 2016, pp. 1–6. Accessed: Dec. 06, 2023. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7577578/

[21] S. Wahyu SJ and M. Faisal, "Pengenalan karakter pada proses digitalisasi dokumen menggunakan cosine similarity," *Pengenalan Karakter Pada Proses Digitalisasi Dokumen Menggunakan Cosine Similarity*, pp. 51–56, 2013.

[22] J. Carbonell and J. Goldstein, "The use of MMR, diversity-based reranking for reordering documents and producing summaries," in *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*, in SIGIR '98. New York, NY, USA: Association for Computing Machinery, Aug. 1998, pp. 335–336. doi: 10.1145/290941.291025.

[23] 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, 2002, pp. 311–318. Accessed: Dec. 06, 2023. [Online]. Available: https://aclanthology.org/P02-1040.pdf

[24] A. Lavie and A. Agarwal, "METEOR: An Automatic Metric for MT Evaluation with High Levels of Correlation with Human Judgments," in *Proceedings of the Second Workshop on Statistical Machine Translation*, C. Callison-Burch, P. Koehn, C. S. Fordyce, and C. Monz, Eds., Prague, Czech Republic: Association for Computational Linguistics, Jun. 2007, pp. 228–231. Accessed: Dec. 06, 2023. [Online]. Available: https://aclanthology.org/W07-0734

[25] C.-Y. Lin, "Rouge: A package for automatic evaluation of summaries," in *Text summarization branches out*, 2004, pp. 74–81. Accessed: Dec. 06, 2023. [Online]. Available: https://aclanthology.org/W04-1013.pdf

[26] T. Zhang, V. Kishore, F. Wu, K. Q. Weinberger, and Y. Artzi, "BERTScore: Evaluating Text Generation with BERT." arXiv, Feb. 24, 2020. doi: 10.48550/arXiv.1904.09675.

[27] R. Bhagat and E. Hovy, "What Is a Paraphrase?," *Computational Linguistics*, vol. 39, no. 3, pp. 463–472, Sep. 2013, doi: 10.1162/COLI_a_00166.

[28] A. Vaswani *et al.*, "Attention Is All You Need." arXiv, Dec. 05, 2017. doi: 10.48550/arXiv.1706.03762.

[29] D. Bahdanau, K. Cho, and Y. Bengio, "Neural Machine Translation by Jointly Learning to Align and Translate." arXiv, May 19, 2016. doi: 10.48550/arXiv.1409.0473.

[30] S. Hochreiter and J. Schmidhuber, "Long Short-term Memory," *Neural computation*, vol. 9, pp. 1735–80, Dec. 1997, doi: 10.1162/neco.1997.9.8.1735.

[31] K. Cho, B. van Merrienboer, D. Bahdanau, and Y. Bengio, "On the Properties of Neural Machine Translation: Encoder-Decoder Approaches." arXiv, Oct. 07, 2014. doi: 10.48550/arXiv.1409.1259.

[32] M. Lewis *et al.*, "BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension." arXiv, Oct. 29, 2019. doi: 10.48550/arXiv.1910.13461.

[33] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, "Improving language understanding by generative pre-training," 2018, Accessed: Oct. 15, 2023. [Online]. Available: https://www.mikecaptain.com/resources/pdf/GPT-1.pdf

[34] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." arXiv, May 24, 2019. doi: 10.48550/arXiv.1810.04805.

[35] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, "Language models are unsupervised multitask learners," *OpenAI blog*, vol. 1, no. 8, p. 9, 2019.

[36] J. Howard and S. Ruder, "Universal Language Model Fine-tuning for Text Classification." arXiv, May 23, 2018. doi: 10.48550/arXiv.1801.06146.

[37] J. Bergstra and Y. Bengio, "Random search for hyper-parameter

optimization.," *Journal of machine learning research*, vol. 13, no. 2, 2012, Accessed: Oct. 30, 2023. [Online]. Available: https://www.jmlr.org/papers/volume13/bergstra12a/bergstra12a.pdf?ref=blog.floydhub.com

[38] M. Chen, Q. Tang, S. Wiseman, and K. Gimpel, "Controllable Paraphrase Generation with a Syntactic Exemplar," in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, Florence, Italy: Association for Computational Linguistics, Jul. 2019, pp. 5972–5984. doi: 10.18653/v1/P19-1599.

[39] J. Wieting and K. Gimpel, "ParaNMT-50M: Pushing the Limits of Paraphrastic Sentence Embeddings with Millions of Machine Translations." arXiv, Apr. 20, 2018. Accessed: Jun. 04, 2023. [Online]. Available: http://arxiv.org/abs/1711.05732

[40] M. Mohler, R. Bunescu, and R. Mihalcea, "Learning to Grade Short Answer Questions using Semantic Similarity Measures and Dependency Graph Alignments," in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, Portland, Oregon, USA: Association for Computational Linguistics, Jun. 2011, pp. 752–762. Accessed: Jun. 22, 2023. [Online]. Available: https://aclanthology.org/P11-1076