

DAFTAR PUSTAKA

- Ahmed, A., Joorabchi, A., & Hayes, M. J. (2022). On Deep Learning Approaches to Automated Assessment: Strategies for Short Answer Grading. *International Conference on Computer Supported Education, CSEDU - Proceedings, 2*, 85–94. <https://doi.org/10.5220/0011082100003182>
- Alpaydin, E. (2020). *Introduction to machine learning* (4th ed.). MIT Press.
- Amalia, A., Lydia, M. S., Muchtar, M. A., Manik, F. Y., Sinu, & Gunawan, D. (2025). Mitigating Bias and Assessment Inconsistencies with BERT-Based Automated Short Answer Grading for the Indonesian Language. *IAENG International Journal of Computer Science*, 52(3), 533–545. https://www.iaeng.org/IJCS/issues_v52/issue_3/IJCS_52_3_01.pdf
- Bhandari, P. (2020). *How to Find Interquartile Range (IQR)*. Scribbr. Diambil 11 Juni 2025, dari <https://www.scribbr.com/statistics/interquartile-range/>
- Burrows, S., Gurevych, I., & Stein, B. (2015a). The eras and trends of automatic short answer grading. In *International Journal of Artificial Intelligence in Education* (Vol. 25, Issue 1, pp. 60–117). Springer New York LLC. <https://doi.org/10.1007/s40593-014-0026-8>
- Camus, L., & Filighera, A. (2020a). Investigating Transformers for Automatic Short Answer Grading. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12164 LNAI, 43–48. https://doi.org/10.1007/978-3-030-52240-7_8
- Chaudhary, S. (2024). *Why 1.5 Is Used in the IQR Rule for Outlier Detection*. Built In. Diambil 11 Juni 2025, dari <https://builtin.com/articles/1-5-iqr-rule>
- Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 7, 1–24. <https://doi.org/10.7717/PEERJ-CS.623>
- Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., ... & Stoyanov, V. (2020). Unsupervised Cross-lingual Representation Learning at Scale. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 8440–8451. <https://doi.org/10.18653/v1/2020.acl-main.747>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of NAACL-HLT 2019*, 4171–4186. <https://doi.org/10.18653/v1/N19-1423>

- Finn, C., Abbeel, P., & Levine, S. (2017). Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks. *Proceedings of the 34th International Conference on Machine Learning*, 70, 1126–1135. <https://doi.org/10.5555/3305381.3305498>
- Florida Institute of Technology Libraries. (2024). *Introduction to Text Mining*. Diambil 11 Juni 2025, dari <https://libguides.lib.fit.edu/c.php?g=968160>
- Gardner, M. W., & Dorling, S. R. (1998). Artificial neural networks (the multilayer perceptron)-A review of applications in the atmospheric sciences. *Atmospheric Environment*, 32(14), 2627–2636. [https://doi.org/10.1016/S1352-2310\(97\)00447-0](https://doi.org/10.1016/S1352-2310(97)00447-0)
- Hock-Chuan, C. (2018). *Regular Expression (Regex) Tutorial*. Diambil 11 Juni 2025, dari <https://www3.ntu.edu.sg/home/ehchua/programming/howto/Regexe.html>
- Learnsity. (2025). *US Teachers Overwhelmed by Grading, New Survey Finds*. Diambil 11 Juni 2025, dari <https://learnsity.com/edtech-blog/a-third-of-us-teachers-considered-leaving-education-in-last-12-months-due-to-grading-workload/>
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *An introduction to information retrieval*. Cambridge University Press.
- Nichol, A., Achiam, J., & Schulman, J. (2018). *On First-Order Meta-Learning Algorithms*. arXiv preprint. <http://arxiv.org/abs/1803.02999>
- OpenAI. (2018). *Reptile: A scalable meta-learning algorithm*. Diambil 11 Juni 2025, dari <https://openai.com/index/reptile/>
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., ... Chintala, S. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, 8026–8037.
- Rahadika Diana, K. D., & Khodra, M. L. (2023). IndoSBERT: Enhancing Indonesian Sentence Embeddings with Siamese Networks Fine-tuning. *2023 10th International Conference on Advanced Informatics: Concept, Theory and Application, ICAICTA 2023*. <https://doi.org/10.1109/ICAICTA59291.2023.10390469>
- Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*, 3982–3992. <https://doi.org/10.18653/v1/D19-1410>
- Salim, H. R., De, C., Pratamaputra, N. D., & Suhartono, D. (2022). Indonesian automatic short answer grading system. *Bulletin of Electrical Engineering and Informatics*, 11(3), 1586–1603. <https://doi.org/10.11591/eei.v11i3.3531>

- Sun, Q. (2024). LSTM-based Trend Prediction of Public Opinion on Urban Fires. *2024 4th International Conference on Neural Networks, Information and Communication Engineering, NNICE 2024*, 929–933. <https://doi.org/10.1109/NNICE61279.2024.10498760>
- Wijaya, M. C. (2021). Automatic Short Answer Grading System in Indonesian Language Using BERT Machine Learning. *Revue d'Intelligence Artificielle*, 35(6), 503–509. <https://doi.org/10.18280/ria.350609>
- Wilie, B., Vincentio, K., Indra Winata, G., Cahyawijaya, S., Li, X., Lim, Z. Y., Soleman, S., Mahendra, R., Fung, P., Bahar, S., Purwarianti, A., & Bandung, I. T. (2020). IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding. *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*, 843–857. <https://doi.org/10.18653/v1/2020.aacl-main.85>
- Zhao, C., Silva, M., & Poulsen, S. (2025). *Language models are few-shot graders*. arXiv preprint. <http://arxiv.org/abs/2502.13337>