

DAFTAR PUSTAKA

- [1] S. Yang and G. Berdine, "Confusion matrix," *The Southwest Respiratory and Critical Care Chronicles*, vol. 12, pp. 75–79, 10 2024.
- [2] B. Jijo and A. Abdulazeez, "Classification based on decision tree algorithm for machine learning," *Journal of Applied Science and Technology Trends*, vol. 2, pp. 20–28, 01 2021.
- [3] M. Hambali, T. Oladele, A. S., A. Kumar, and W. Gao, "Feature selection and computational optimization in high-dimensional microarray cancer datasets via infogain-modified bat algorithm," *Multimedia Tools and Applications*, vol. 81, 08 2022.
- [4] H. A. Salman, A. Kalakech, and A. Steiti, "Random forest algorithm overview," *Babylonian Journal of Machine Learning*, vol. 2024, pp. 69–79, 2024.
- [5] Z. Sakhiyya and N. Martin-Anatias, "Reviving the language at risk: a social semiotic analysis of the linguistic landscape of three cities in indonesia," *International Journal of Multilingualism*, vol. 20, no. 2, pp. 290–307, 2023. [Online]. Available: <https://doi.org/10.1080/14790718.2020.1850737>
- [6] A. R. Lubis, Y. Y. Lase, D. A. Rahman, and D. Witorsyah, "Improving spell checker performance for bahasa indonesia using text preprocessing techniques with deep learning models," *Ingénierie des Systèmes d'Information*, vol. 28, no. 5, pp. 1335–1342, 2023.
- [7] D. Hladek, J. Staš, and M. Pleva, "Survey of automatic spelling correction," *Electronics*, vol. 9, p. 1670, 10 2020.
- [8] I. Suzanti, Y. Negara, A. Radiah, A. Cahyani, and H. Husni, "Koreksi kesalahan ejaan pada pencarian artikel pariwisata berbahasa indonesia menggunakan levenshtein distance," *Jurnal Simantec*, vol. 13, pp. 81–90, 12 2024.
- [9] U. Albalawi and M. Mustafa, "Current artificial intelligence (ai) techniques, challenges, and approaches in controlling and fighting covid-19: A review," *International Journal of Environmental Research and Public Health*, vol. 19, no. 10, p. 5901, May 2022.
- [10] R. Muliawanti, "Dampak teknologi terhadap jurnalisme pada era digital," *Kumparan*, 2024.
- [11] R. Silalahi and E. Susanto, "Penerapan kode etik jurnalistik media online tribunews dan efek pemberitaan pada pembacanya," *Koneksi*, vol. 4, p. 293, 10 2020.

- [12] N. Mediyawati, J. C. Young, and S. B. Nusantara, "U-tapis: Automatic spelling filter as an effort to improve indonesian language competencies of journalistic students," *Jurnal Cakrawala Pendidikan*, vol. 40, no. 2, p. 402–412, Jun. 2021. [Online]. Available: <https://jurnal.uny.ac.id/index.php/cp/article/view/34546>
- [13] G. I. Wijaya and M. V. Overbeek, "Leveraging lstm neural networks for advanced harassment detection: Insights into insults and defamation in the u-tapis module," *Engineering Proceedings*, vol. 107, no. 1, 2025. [Online]. Available: <https://www.mdpi.com/2673-4591/107/1/11>
- [14] J. Kosasih, M. V. Overbeek, N. Mediyawati, R. Sutomo, and S. B. Nusantara, "Detection and enhancement of bahasa standardized word in online news draft using levenshtein algorithm :u-tapis," in *2024 7th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, 2024, pp. 444–449.
- [15] P. Tendy and M. V. Overbeek, "U-tapis: A hybrid approach to melting word error detection and correction with damerau-levenshtein distance and roberta," *Engineering Proceedings*, vol. 107, no. 1, 2025. [Online]. Available: <https://www.mdpi.com/2673-4591/107/1/19>
- [16] K. Gleneagles, M. V. Overbeek, N. Mediyawati, R. Sutomo, and S. B. Nusantara, "U-tapis melting words : an artificial intelligence application for detecting melt word errors in indonesia online news," in *2024 7th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, 2024, pp. 438–443.
- [17] M. Overbeek, B. Glennardy, N. Mediyawati, S. Nusantara, and R. Sutomo, "U-tapis sal-tik : Typing error detection using random forest algorithm," *Ultimatics : Jurnal Teknik Informatika*, vol. 16, pp. 38–45, 07 2024.
- [18] N. R. Dwitya and M. V. Overbeek, "Development of detection and correction of errors in spelling and compound words using long short-term memory," *AIP Conference Proceedings*, vol. 3220, no. 1, p. 040005, 10 2024. [Online]. Available: <https://doi.org/10.1063/5.0235850>
- [19] A. Saputra and M. Overbeek, "Harnessing long short-term memory algorithm for enhanced di-di word error detection and correction," 01 2024, p. 040002.
- [20] F. Annisa and A. Junaidi, "Implementasi bahasa jurnalistik pada media siber (analisis wacana pada berita okezone.com kanal perjalanan)," *Koneksi*, vol. 6, p. 101, 03 2022.
- [21] Y. Yanfi, F. Gaol, B. Soewito, and H. L. H. Spits Warnars, "Spell checker for the indonesian language: Extensive review," *International Journal of Emerging Technology and Advanced Engineering*, vol. 12, pp. 1–7, 05 2022.

- [22] J. Zhu, X. Shi, and S. Zhang, "Machine learning-based grammar error detection method in english composition," *Scientific Programming*, vol. 2021, pp. 1–10, 12 2021.
- [23] J. Ali, R. Khan, N. Ahmad, and I. Maqsood, "Random forests and decision trees," *International Journal of Computer Science Issues(IJCSI)*, vol. 9, 09 2012.
- [24] E. Fitri, "Analisis sentimen terhadap aplikasi ruangguru menggunakan algoritma naive bayes, random forest dan support vector machine," *Jurnal Transformatika*, vol. 18, p. 71, 07 2020.
- [25] N. G. Ramadhan, F. Adhinata, A. Segara, and D. Putra Rakhmadani, "Deteksi berita palsu menggunakan metode random forest dan logistic regression," *JURIKOM (Jurnal Riset Komputer)*, vol. 9, p. 251, 04 2022.
- [26] A. P. Aditya, "Kesalahan berbahasa pada berita media massa online kompas.com," *Jurnal Pendidikan Bahasa dan Sastra Indonesia*, vol. 8, no. 1, pp. 1–15, 2025.
- [27] R. Sriharyati, "Analisis kesalahan berbahasa pada berita daring liputan6.com," *Jurnal Pendidikan Bahasa dan Sastra Indonesia*, vol. 7, no. 3, pp. 45–60, 2021.
- [28] EYD, "Penulisan kata ganti," <https://eyd.netlify.app/penulisan-kata/kata-ganti>, diakses: 26 September 2025.
- [29] Pujiati, "Penulisan kata ganti (ku, kau, mu, nya) yang benar," <https://penerbitdeepublish.com/penulisan-kata-ganti-yang-benar/>, 2024, diakses: 26 September 2025.
- [30] C. Supriadi, H. D. Purnomo, and I. Sembiring, "Sensitivitas sistem pencarian artikel bahasa indonesia menggunakan metode n-gram dan tanimoto cosine," *Jurnal Transformasi*, vol. 18, no. 1, p. 63, 2020.
- [31] H. Sujaini, H. Muhardi, and J. H. Simanjuntak, "Aplikasi pengoreksi ejaan (spelling correction) pada naskah jurnal bidang informatika dengan n-gram dan jaro-winkler distance," *Jurnal Edukasi dan Penelitian Informatika (JEPIN)*, vol. 8, no. 2, 2022, pontianak, Kalimantan Barat.
- [32] W. Olabiyi, D. Akinyele, and E. Joel, "The evolution of ai: From rule-based systems to data-driven intelligence," 01 2025.
- [33] L. Bergamin, M. Polato, and F. Aiolli, "Improving rule-based classifiers by bayes point aggregation," *Neurocomputing*, vol. 613, p. 128699, 2025. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S092523122401470X>

- [34] J. Ilemobayo, O. Durodola, O. Alade, O. Awotunde, T. Adewumi, O. Falana, A. Ogungbire, A. Osinuga, D. Ogunbiyi, I. Odezuligbo, O. Edu, and A. Ifeanyi, "Hyperparameter tuning in machine learning: A comprehensive review," *Journal of Engineering Research and Reports*, vol. 26, pp. 388–395, 06 2024.
- [35] P. Shetty, R. Udhayakumar, A. Patil, M. Manwal, P. Vadar, and P. Nooji, "Application of natural language processing (nlp) in machine learning," 11 2023, pp. 949–957.
- [36] H. Wei, "A random forest-based prediction for liver cirrhosis," *Applied and Computational Engineering*, vol. 78, pp. 98–106, 07 2024.
- [37] A. Masbakhah, U. Sa'adah, and M. Muslikh, "Heart disease classification using random forest and fox algorithm as hyperparameter tuning," *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, vol. 7, no. 4, pp. 964–976, Aug. 2025. [Online]. Available: <https://jeeemi.org/index.php/jeeemi/article/view/932>
- [38] H. Patel and P. Prajapati, "Study and analysis of decision tree based classification algorithms," *International Journal of Computer Sciences and Engineering*, vol. 6, pp. 74–78, 10 2018.
- [39] M. Loey, G. Manogaran, M. H. N. Taha, and N. E. M. Khalifa, "A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the covid-19 pandemic," *Measurement*, vol. 167, p. 108288, 2021. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0263224120308289>
- [40] A. Cutler, D. Cutler, and J. Stevens, *Random Forests*, 01 2011, vol. 45, pp. 157–176.
- [41] C. Sanghun, S.-M. Lee, K. E. Lee, M.-G. Kim, S.-Y. Lee, and K.-H. Shin, "Identification of biomarkers in gynecologic cancers: a machine learning approach for metabolomics," *Translational and Clinical Pharmacology*, vol. 33, p. 168, 09 2025.
- [42] W. Parker, *Model evaluation*, 08 2024, pp. 208–219.
- [43] D. K. S. Harpreet Kaur, "Evaluating the effectiveness of the proposed system using f1 score, recall, accuracy, precision and loss metrics compared to prior techniques," *International Journal of Communication Networks and Information Security (IJCNIS)*, vol. 15, no. 4, p. 368–383, Sep. 2024. [Online]. Available: <https://ijcnis.org/index.php/ijcnis/article/view/7168>