

DAFTAR PUSTAKA

- [1] N. L. Marliana and E. Puryanto, “Problematika penggunaan ragam bahasa jurnalistik pada media massa dan implikasinya terhadap pembinaan bahasa indonesia di masyarakat,” *Diksi*, vol. 17, 2015.
- [2] M. Simatupang, R. Peter, E. Murniarti, and H. Male, “The tendency of using nonstandard indonesian in learning activity,” *Proceedings from the 1st International Conference on Law and Human Rights, ICLHR 2021, 14-15 April 2021, Jakarta, Indonesia*, 2021.
- [3] S. Suprihatma, “Analisis penggunaan bahasa indonesia dalam jurnalistik pada media massa online,” *Journal on Education*, vol. 6, pp. 11 011–11 018, 01 2024.
- [4] S. Talitha and W. Ferdianto, “Analisis kesalahan berbahasa pada teks berita daring dan implikasinya terhadap pembelajaran bahasa indonesia smp,” *Jurnal Pendidikan Bahasa dan Sastra Indonesia*, vol. I, 2022.
- [5] 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,” *Cakrawala Pendidikan*, vol. 40, 2021.
- [6] J. K. S. D. No, N. Mediyawati, R. Sutomo, S. B. Nusantara, and M. V. Overbeek, “Utapis indonesian word error detection application: Design and development,” *Indonesian Journal of Computer Science Attribution*, vol. 13, pp. 2024–626, 2024.
- [7] N. Mediyawati and S. Bintang, “Platform kecerdasan buatan sebagai media inovatif untuk meningkatkan keterampilan berkomunikasi: U-tapis,” 2021.
- [8] V. Gavriel, A. Siswanto, and M. V. Overbeek, “Development of ”kata terikat” detection and writing errors correction using rabin-karp and random forest algorithm,” 2024.
- [9] A. Kevin, B. Saputra, and M. V. Overbeek, “Harnessing long short-term memory algorithm for enhanced di-di word error detection and correction,” 2024.
- [10] N. Indurkhya and F. J. Damerau, *Handbook of natural language processing, second edition*, 2010.
- [11] N. R. Dwitya and M. V. Overbeek, “Development of detection and correction of errors in spelling and compound words using long short-term memory,” 2024.

- [12] A. M. Azmi, M. N. Almutery, and H. Aboalsamh, "Real-word errors in arabic texts: A better algorithm for detection and correction," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, pp. 1308–1320, 2019.
- [13] J. Li and X. Wang, "Combining trigram and automatic weight distribution in chinese spelling error correction," *Journal of Computer Science and Technology*, vol. 17, pp. 915–923, 2002.
- [14] L. Zhang, M. Zhou, C. Huang, and H. Pan, "Automatic detecting/correcting errors in chinese text by an approximate word-matching algorithm," pp. 248–254, 2000.
- [15] A. Islam and D. Inkpen, "Real-word spelling correction using google web 1t 3-grams," pp. 1241–1249, 2009.
- [16] N. Zukarnain, B. S. Abbas, S. Wayan, A. Trisetyarso, and C. Kang, "Spelling checker algorithm methods for many languages," *2019 International Conference on Information Management and Technology (ICIMTech)*, vol. 1, pp. 198–201, 2019.
- [17] A. Ene and A. Ene, "An application of levenshtein algorithm in vocabulary learning," *2017 9th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, pp. 1–4, 2017.
- [18] B. Berger, M. Waterman, and Y. Yu, "Levenshtein distance, sequence comparison and biological database search," *IEEE transactions on information theory*, vol. 67, pp. 3287 – 3294, 2020.
- [19] C. Gooskens and W. Heeringa, "Perceptive evaluation of levenshtein dialect distance measurements using norwegian dialect data," *Language Variation and Change*, vol. 16, pp. 189 – 207, 2004.
- [20] O. Susilawati, "Perbandingan akurasi algoritma levenshtein distance dengan algoritma cosine similarity untuk koreksi penulisan dalam kalimat," 2019. [Online]. Available: <https://etheses.uinsgd.ac.id/20367/>
- [21] T. Okuda, E. Tanaka, and T. Kasai, "A method for the correction of garbled words based on the levenshtein metric," *IEEE Transactions on Computers*, vol. C-25, pp. 172–178, 1976.
- [22] R. Devianty, "Penggunaan kata baku dan tidak baku dalam bahasa indonesia," *EUNOIA (Jurnal Pendidikan Bahasa Indonesia)*, vol. 1, 2021.
- [23] N. F. T. Utami, A. P. Y. Utomo, S. A. Buono, and N. I. Sabrina, "Analisis kesalahan sintaksis pada cerpen berjudul "warisan untuk doni" karya putu ayub," *JURNAL RISET RUMPUN ILMU BAHASA*, vol. 1, 2022.

- [24] R. V and S. Naik, "Natural language processing," *International Journal of Computing Algorithm*, 2023.
- [25] X. Sun, X. Liu, J. Hu, and J. Zhu, "Empirical studies on the nlp techniques for source code data preprocessing," pp. 32–39, 2014.
- [26] R. Dale, H. Moisl, and H. Somers, "Handbook of natural language processing," *Computational Linguistics*, vol. 27, pp. 602–603, 2001.
- [27] C. P. Chai, "Comparison of text preprocessing methods," *Natural Language Engineering*, vol. 29, pp. 509 – 553, 2022.
- [28] L. Hickman, S. Thapa, L. Tay, M. Cao, and P. Srinivasan, "Text preprocessing for text mining in organizational research: Review and recommendations," *Organizational Research Methods*, vol. 25, pp. 114 – 146, 2020.
- [29] J. Sheng-jun, "The similitude calculate of sentence based on levenshtein distance algorithm," *Computer Knowledge and Technology*, 2009.
- [30] R. W. Soukoreff and I. S. MacKenzie, "Measuring errors in text entry tasks: an application of the levenshtein string distance statistic," *CHI '01 Extended Abstracts on Human Factors in Computing Systems*, 2001.
- [31] I. S. Ardan, J. Sulastri, and A. Rakhmawati, "Analisis performansi entity matching dengan fuzzy wuzzy pada artikel fairness ai," 2023. [Online]. Available: <https://ejurnal.teknokrat.ac.id/index.php/teknoinfo/index>
- [32] Z. Karimi, "Confusion matrix some of the authors of this publication are also working on these related projects: data cleaning process view project," *researchgate*, vol. 1, 2021.
- [33] J. Coburn, "Getting started with flask," *Build Your Own Car Dashboard with a Raspberry Pi*, pp. 143 – 170, 2020.
- [34] O. Bunke and B. Droge, "Estimators of the mean squared error of prediction in linear regression," *Technometrics*, vol. 26, pp. 145–155, 1984.

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A