DAFTAR PUSTAKA

[1] A. N. Ihsan and S. Tresnawati, "Analisis Sentimen Pada Platform X Terhadap Layanan Provider Tri Menggunakan Naïve Bayes Dan Support Vector Machine," Jurnal Informatika dan Teknik Elektro Terapan, vol. 12, no. 3S1, Oct. 2024, doi: 10.23960/jitet.v12i3s1.5264.

[2] We Are Social and Meltwater, "Digital 2025: Indonesia," Feb. 2025. Available: https://datareportal.com/reports/digital-2025-indonesia

[3] I. Villanueva-Miranda, Y. Xie, and G. Xiao, "Sentiment analysis in public health: a systematic review of the current state, challenges, and future directions," Front. Public Health, vol. 13, p. 1609749, Jun. 2025, doi: 10.3389/fpubh.2025.1609749.

[4] M. Kholilullah, M. Martanto, and U. Hayati, "Analisis Sentimen Pengguna Twitter(X) Tentang Piala Dunia Usia 17 Menggunakan Metode Naive Bayes," JATI (Jurnal Mahasiswa Teknik Informatika), vol. 8, no. 1, pp. 392– 398, Feb. 2024, doi: 10.36040/jati.v8i1.8378.

[5] M. Amien, "Sejarah dan Perkembangan Teknik Natural Language Processing (NLP) Bahasa Indonesia: Tinjauan tentang sejarah, perkembangan teknologi, dan aplikasi NLP dalam bahasa Indonesia," arXiv (Cornell University), Mar. 2023, doi: 10.48550/arxiv.2304.02746. Available: http://arxiv.org/abs/2304.02746

[6] A. Kiftiyah, F. A. Palestina, F. U. Abshar, and K. Rofiah, "Program Makan Bergizi Gratis (MBG) dalam Perspektif Keadilan Sosial dan Dinamika Sosial – Politik," Pancasila: Jurnal Keindonesiaan, vol. 5, no. 1, pp. 101–112, Apr. 2025, doi: 10.52738/pjk.v5i1.726.

[7] E. B. I. Panjiwinata, "Kebijakan MBG: Menabur janji manis, menuai ekonomi kritis," Panjiwinata | Konferensi Nasional Ilmu Administrasi, Oct. 18, 2025. Available: http://knia.stialanbandung.ac.id/index.php/knia/article/view/1197

[8] Badan Gizi Nasional, "BGN: Dari Rp268T, 93% Anggaran Digunakan Langsung untuk Bantuan MBG." bgn.go.id, Mar.2026. Available: https://www.bgn.go.id/news/siaran-pers/bgn-dari-rp268t-93-anggaran-digunakan-langsung-untuk-bantuan-mbg

[9] W. Ningsih, B. Alfianda, R. Rahmaddeni, and D. Wulandari, "Perbandingan Algoritma SVM dan Naïve Bayes dalam Analisis Sentimen Twitter pada 50 Penggunaan Mobil Listrik di Indonesia," MALCOM: Indonesian Journal of

Machine Learning and Computer Science, vol. 4, no. 2, pp. 556–562, Feb. 2024, doi: 10.57152/malcom.v4i2.1253.

[10] Hunga Baba, U. R. (2024). Analisa Sentimen Menjelang Pemilihan Umum Presideen 2024 di Indonesia Menggunakan Perbandingan Performa Support Vector Machine (SVM) dan Naïve Bayes. Innovative: Journal Of Social Science Research, 4(3), 11972–11990. Available: https://jinnovative.org/index.php/Innovative/article/view/10761

[11] L. Luthfanida, "Analisis Sentimen Data Twitter Menggunakan Metode Naive Bayes Dan Support Vector Machine (Svm) Tentang Presiden Jokowi 3 Periode," Djtechno Jurnal Teknologi Informasi, vol. 3, no. 1, pp. 5–11, Jul. 2022, doi: 10.46576/djtechno.v3i1.2143.

[12] Ramadhani, A., Permana, I., Afdal, M., & Fronita, M. (2024). Analisis Sentimen Tanggapan Publik di Twitter Terkait Program Kerja Makan Siang Gratis Prabowo–Gibran Menggunakan Algoritma Naïve Bayes Classifier dan Support Vector Machine. Building of Informatics, Technology and Science (BITS), 6(3), 1509−1516. https://doi.org/10.47065/bits.v6i3.6188

[13] B. Liu, "Sentiment Analysis and Opinion Mining," Morgan & Claypool Publishers, 2012.

[14] M. A. Subarkah, W. M. Rahmawati, S. R. Wardhana, and R. K. Hapsari, "Analisis sentimen terhadap video ulasan produk menggunakan metode Support Vector Machine dengan sequential Minimal optimization," KERNEL Jurnal Riset Inovasi Bidang Informatika Dan Pendidikan Informatika, vol. 3, no. 2, pp. 17–24, Dec. 2022, doi: 10.31284/j.kernel.2022.v3i2.4039.

[15] T. Juniardi and C. A. Sugianto, "Analisis Sentimen Tim Nasional Sepak Bola Indonesia Di Turnamen Piala Dunia U-17 Indonesia Pada Twitter (X) Menggunakan Algortima Naïve Bayes," Jurnal Informatika dan Teknik Elektro Terapan, vol. 12, no. 3S1, Oct. 2024, doi: 10.23960/jitet.v12i3s1.5188.

[16] E. Saputri, "Teknik dan aplikasi data mining di Indonesia: tinjauan literatur satu dekade (2015-2024)," IT-Explore: Jurnal Penerapan Teknologi Informasi dan Komunikasi, vol. 4, no. 2, pp. 138–149, Jun. 2025, doi: 10.24246/itexplore.v4i2.2025.pp138-149.

[17] P. W. Rahayu et al., Buku ajar data mining. PT. Sonpedia Publishing Indonesia, 2024. [Online]. Available: (Pdf) Buku Ajar Data Mining

[18] H. Mondal *et al.*, "Science communication on X (formerly Twitter): A H. Mondal et al., "Science communication on X (formerly Twitter): A picture is

worth a thousand characters?," Exploration of Digital Health Technologies, pp. 28–34, Nov. 2023, doi: 10.37349/edht.2023.00005.

[19] R. Mustaqlillah, O. Widyaningtyas, and T. Wantoro, "Efektivitas penggunaan Twitter Sebagai Sarana Peningkatan Berpikir Kritis Mahasiswa Ilmu Komunikasi," MUKASI Jurnal Ilmu Komunikasi, vol. 2, no. 1, pp. 18–28, Feb. 2023, doi: 10.54259/mukasi.v2i1.1346.

[20] D. S. Nugroho, I. F. Hanif, M. A. Hasbi, F. Fredianto, A. M. Saputra, and R. Zildjian, "Analisis Sentimen Dugaan Pelanggaran Pemilu 2024 Berdasarkan Tweet Menggunakan Algoritma Naïve Bayes Classifier," MALCOM: Indonesian Journal of Machine Learning and Computer Science, vol. 4, no. 3, pp. 1169–1176, Jul. 2024, doi: 10.57152/malcom.v4i3.1496.

[21] Aprillia, N. P. ., & Azzahra, M. U. D. . (2025). Program Makan Bergizi Gratis (MBG): Antara Legacy Project dan Solusi Stunting. Journal of Information Systems and Management (JISMA), 4(5), 19–22. Available: https://jisma.org/index.php/jisma/article/view/1229

[22] H. A. Rahmah, A. Anggraini, Y. P. Nilasari, and E. P. Salsabilla, "Analisis Efektivitas Program Makan Bergizi Gratis Di Sekolah Dasar Indonesia Tahun 2025," Integrative Perspectives of Social and Science Journal, vol. 2, no. 2, p. 2855, 2025. Available: https://ipssj.com/index.php/ojs/article/view/380

[23] W. Widyananda, Maskur, and A. Fauzi, "Machine Learning and Transformer-based Model for Sentiment Analysis of Indonesian ECommerce Reviews," The Indonesian Journal of Computer Science, vol. 14, no. 4, Aug. 2025, doi: 10.33022/ijcs.v14i4.4980.

[24] C. Adiwinata and A. Afiyati, "Sentiment analysis of X users toward electric motorcycles using SVM and BERT algorithms," JUITA Jurnal Informatika, pp. 119–126, Aug. 2025, doi: 10.30595/juita.v13i2.26152.

[25] N. A. Salsabila, Y. A. Winatmoko, A. A. Septiandri, and A. Jamal, "Colloquial Indonesian Lexicon," in 2018 International Conference on Asian Language Processing (IALP), Bandung, Indonesia, Nov. 2018, pp. 226–229, doi: 10.1109/IALP.2018.8629151. Available: https://github.com/nasalsabila/kamus-alay.

[26] Wibowo, H. A., Nityasya, M. N., Akyürek, A. F., Fitriany, S., Aji, A. F., Prasojo, R. E., & Wijaya, D. T. (2021). *IndoCollex: A Testbed for Morphological Transformation of Indonesian Colloquial Words*. Findings of ACL-IJCNLP 2021, 3170–3183.

[27] Kralj Novak, P., Smailović, J., Sluban, B., & Mozetič, I. (2015). *Sentiment of Emojis*. PLOS ONE, 10(12), e0144296.

[28] M. Adriani, J. Asian, B. Nazief, S. M. M. Tahaghoghi, and H. E. Williams, "Stemming Indonesian: A confix-stripping approach," ACM Transactions on Asian Language Information Processing, vol. 6, no. 4, pp. 1–33, Dec. 2007, doi: 10.1145/1316457.1316459.

[29] A. S. Rizkia, W. Wufron, and F. F. Roji, "Analisis Sentimen Coretax: Perbandingan Pelabelan Data Manual, Transformers-Based, dan LexiconBased pada Performa IndoBERT," MALCOM: Indonesian Journal of Machine Learning and Computer Science, vol. 5, no. 3, Jul. 2025, doi: 10.57152/malcom.v5i3.2151.

[30] A. R. Ismail dan R. B. F. Hakim, "Implementasi Lexicon Based Untuk Analisis Sentimen Dalam Mengetahui Trend Wisata Pantai Di DI Yogyakarta Berdasarkan Data Twitter" Emerging Statistics and Data Science Journal, vol. 1, no. 1, pp. 37–46, Jan. 2023, doi: 10.20885/esds.vol1.iss.1.art5.

[31] M. F. N. Fathoni, E. Y. Puspaningrum, dan A. N. Sihananto, "Perbandingan Performa Labeling Lexicon InSet dan VADER pada Analisa Sentimen Rohingya di Aplikasi X dengan SVM," Modem: Jurnal Informatika dan Sains Teknologi, vol. 2, no. 3, pp. 62–76, Jul. 2024, doi: 10.62951/modem.v2i3.112.

[32] Y. Mao, Q. Liu, and Y. Zhang, "Sentiment analysis methods, applications, and challenges: A systematic literature review," Journal of King Saud University - Computer and Information Sciences, vol. 36, no. 4, p. 102048, Apr. 2024, doi: 10.1016/j.jksuci.2024.102048.

[33] T. Wilson, J. Wiebe, and P. Hoffmann, "Recognizing contextual polarity in phrase-level sentiment analysis," in *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*, 2005, pp. 347–354. [Online]. Available: ACL Anthology.

[34] J. Wiebe, R. F. Bruce, and T. P. O'Hara, "Development and use of a gold-standard data set for subjectivity classifications," in *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics*, 1999, pp. 246–253. [Online]. Available: ACL Anthology.

[35] R. Artstein and M. Poesio, "Inter-coder agreement for computational linguistics," *Computational Linguistics*, vol. 34, no. 4, pp. 555–596, 2008, doi: 10.1162/coli.07-034-R2.

[36] J. Cohen, "A coefficient of agreement for nominal scales," *Educational and Psychological Measurement*, vol. 20, no. 1, pp. 37–46, 1960, doi: 10.1177/001316446002000104.

[37] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, no. 1, pp. 159–174, 1977, doi: 10.2307/2529310.

[38] E. Riloff, A. Qadir, P. Surve, L. De Silva, N. Gilbert, and R. Huang, "Sarcasm as contrast between a positive sentiment and negative situation," in *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, 2013, pp. 704–714. [Online]. Available: ACL Anthology.

[39] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002, doi: 10.1613/jair.953.

[40] G. E. A. P. A. Batista, R. C. Prati, and M. C. Monard, "A study of the behavior of several methods for balancing machine learning training data," *ACM SIGKDD Explorations Newsletter*, vol. 6, no. 1, pp. 20–29, 2004, doi: 10.1145/1007730.1007735.

[41] J. E. Bartlett II, J. W. Kotrlik, and C. C. Higgins, "Organizational Research: Determining Appropriate Sample Size in Survey Research," *Information Technology, Learning, and Performance Journal*, vol. 19, no. 1, pp. 43–50, 2001. [Online]. Available: https://www.opalco.com/wp-content/uploads/2014/10/Reading-Sample-Size1.pdf

[42] Y. A. Singgalen, "Analisis Sentimen Wisatawan Melalui Data Ulasan Candi Borobudur di Tripadvisor Menggunakan Algoritma Naïve Bayes Classifier," Building of Informatics, Technology and Science (BITS), vol. 4, no. 3, Dec. 2022, doi: 10.47065/bits.v4i3.2486.

[43] A. Pamuji, "Prediksi Otorisasi Pengguna Sistem Berkas pada Algoritma Klasifikasi dengan Teknik Naïve Bayes," Infomatek, vol. 24, no. 1, pp. 35–44, Jun. 2022, doi: 10.23969/infomatek.v24i1.4604.

[44] L. Nursinggah, R. Ruuhwan, and T. Mufizar, "Analisis Sentimen Pengguna Aplikasi X Terhadap Program Makan Siang Gratis Dengan Metode Naïve Bayes Classifier," Jurnal Informatika dan Teknik Elektro Terapan, vol. 12, no. 3, Aug. 2024, doi: 10.23960/jitet.v12i3.4336.

[45] W. M. P. Dhuhita and F. Zone, "Perbandingan Kinerja Algoritma Multinomial dan Bernoulli Naïve Bayes dalam Mengklasifikasikan

Komentar Cyberbullying," *Komputika: Jurnal Sistem Komputer*, vol. 12, no. 2, pp. 109–117, Oct. 2023, doi: 10.34010/komputika.v12i2.9767.

[46] A. M. Yolanda, R. T. Mulya, and R. T. Mulya, "Implementasi Metode Support Vector Machine untuk Analisis Sentimen pada Ulasan Aplikasi Sayurbox di Google Play Store," VARIANSI: Journal of Statistics and Its Application on Teaching and Research, vol. 6, no. 2, pp. 76–83, 2024, doi: 10.35580/variansiunm258.

[47] P. Chowdhuri, P. Pal, and T. Si, "A novel steganographic technique for medical image using SVM and IWT," Multimedia Tools and Applications, vol. 82, no. 13, pp. 20497–20516, Jan. 2023, doi: 10.1007/s11042-022-14301-0.

[48] S. D. Wahyuni and R. H. Kusumodestoni, "Optimalisasi Algoritma Support Vector Machine (SVM) Dalam Klasifikasi Kejadian Data Stunting," *Bulletin of Information Technology (BIT)*, vol. 5, no. 2, pp. 56–64, 2024, doi: 10.47065/bit.v5i2.1247.

[49] M. Q. Alqowiy and E. Utami, "Parameter Optimisation of Support Vector Machine using Genetic Algorithm for Cyberbullying Detection," SISTEMASI, vol. 14, no. 1, p. 332, Jan. 2025, doi: 10.32520/stmsi.v14i1.4870.

[50] A. A. Baskara, N. M. Piranti, and M. F. Romdendine, "Framework Data Mining: Sebuah Survei," JATI (Jurnal Mahasiswa Teknik Informatika), vol. 9, no. 3, pp. 4886–4895, May 2025, doi: 10.36040/jati.v9i3.13803.

[51] S. Pambudi, P. Setiaji, and W. A. Triyanto, "Sentiment Analysis of Fizzo Novel Application Using Support Vector Machine and Naïve Bayes Algorithm with SEMMA Framework," Jurnal Teknik Informatika (Jutif), vol. 6, no. 4, pp. 1861–1880, Aug. 2025, doi: 10.52436/1.jutif.2025.6.4.4875.

[52] L. Hakim, A. Sobri, L. Sunardi, and D. Nurdiansyah, "Prediksi penyakit jantung berbasis mesin learning dengan menggunakan metode k-nn," Jurnal Digital Teknologi Informasi, vol. 7, no. 2, p. 14, Feb. 2025, doi: 10.32502/digital.v7i2.9429.

[53] M. N. Hidayat and R. Pramudita, "Analisis Sentimen Terhadap Pembelajaran Secara Daring Pasca Pandemi Covid-19 Menggunakan Metode IndoBERT," Information Management for Educators and Professionals, Journal of Information Management,, vol. 8, no. 2, pp. 161–170, 2023, doi: 10.51211/imbi.v8i2.2719.

[54] M. Sokolova and G. Lapalme, “A systematic analysis of performance measures for classification tasks,” *Information Processing & Management*, vol. 45, no. 4, pp. 427–437, May 2009, doi: 10.1016/j.ipm.2009.03.002.

[55] A. U. E. Majid and R. Nuari, “Performance Comparison Of Bert Metrics And Classical Machine Learning Models (SVM, Naive Bayes) For Sentiment Analysis,” INOVTEK Polbeng - Seri Informatika, vol. 10, no. 2, pp. 741–752, May 2025, doi: 10.35314/wmh3rg23.

[56] A. Sultoni, D. A. Putra, H. N. Wahidah, M. M. Arief, P. J. P. RM, and S. B. Utami, “Public Sentiment analysis and distribution optimization of Indonesia’s Makan Bergizi Gratis (MBG) program,” Sultoni | Jurnal Matematika Thales, Oct. 2025, doi: 10.22146/jmt.108835.

[57] D. Septiani and I. Isabela, “Analisis Term Frequency Inverse Document Frequency (TF-IDF) Dalam Temu Kembali Informasi Pada Dokumen Teks,” 2022. Available: https://journal.unj.ac.id/unj/index.php/SINTESIA/article/view/39364

[58] W. Wijiyanto, A. I. Pradana, S. Sopingi, and V. Atina, “Teknik K-Fold Cross Validation untuk Mengevaluasi Kinerja Mahasiswa,” Jurnal Algoritma, vol. 21, no. 1, May 2024, doi: 10.33364/algoritma/v.21-1.1618.

[59] R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," in Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI), Montreal, Canada, 1995, pp. 1137 - 1145.

[60] J. A. Wibowo, V. C. Mawardi, and T. Sutrisno, “Visualisasi Word Cloud Hasil Analisis Sentimen Berbasis Fitur Layanan Aplikasi Gojek Dengan Support Vector Machine,” Jurnal Serina Sains, Teknik dan Kedokteran, vol. 2, no. 1, pp. 61–70, Mar. 2024, doi: 10.24912/jsstk.v2i1.32058.

[61] M. G. Pradana, “Penggunaan Fitur Wordcloud Dan Document Term Matrix Dalam Text Mining,” Jurnal Ilmiah Informatika, vol. 8, no. 01, pp. 38–43, Mar. 2020, doi: 10.33884/jif.v8i01.1838.

[62] M. Sivakumar, S. Parthasarathy, and T. Padmapriya, “A simplified approach for efficiency analysis of machine learning algorithms,” PeerJ Computer Science, vol. 10, p. e2418, Nov. 2024, doi: 10.7717/peerj-cs.2418.

[63] N. A. Semary, W. Ahmed, K. Amin, P. Pławiak, and M. Hammad, “Enhancing machine learning-based sentiment analysis through feature extraction techniques,” PLoS One, vol. 19, no. 2, p. e0294968, Feb. 2024, doi: 10.1371/journal.pone.0294968.

[64] L. B. Ilmawan and M. A. Mude, “Perbandingan Metode Klasifikasi Support Vector Machine dan Naïve Bayes untuk Analisis Sentimen pada Ulasan Tekstual di Google Play Store,” ILKOM Jurnal Ilmiah, vol. 12, no. 2, pp. 154–161, Aug. 2020, doi: 10.33096/ilkom.v12i2.597.154-161.

[65] A. R. D. P. A. Rama and A. Afiyati, “Comparative Sentiment Analysis of Indonesian Leadership Transitions on Platform X Using LSTM and Naïve Bayes,” Jurnal Teknologi Dan Open Source, vol. 8, no. 2, pp. 639–649, Dec. 2025, doi: 10.36378/jtos.v8i2.4596.

[66] D. A. P. Hapsari, W. K. Nofa, dan S. Santoso, "Analisis Kompleksitas Algoritma Filter IIR Shen-Castan untuk Deteksi Tepi pada Citra Digital," CCIT Journal, vol. 12, no. 2, pp. 218 - 228, 2019.

[67] B.-M. Hsu, “Comparison of supervised classification models on textual data,” *Mathematics*, vol. 8, no. 5, p. 851, May 2020, doi: 10.3390/math8050851.

[68] V. Cutting and N. Stephen, “A Review on using Python as a Preferred Programming Language for Beginners,” International Research Journal of Engineering and Technology, 2021, [Online]. Available: A-Review-onusing-Python-as-a-Preferred-Programming-Language-for-Beginners.pdf

[69] T. Hidayansyah and S. Budi, “Analisis eksploratif terhadap faktor kunci pada dataset pemesanan hotel,” May 23, 2025. Available: https://strategi.it.maranatha.edu/index.php/strategi/article/view/582.

[70] F. Pedregosa et al., “SciKit-Learn: Machine Learning in Python,” arXiv (Cornell University), Jan. 2012, doi: 10.48550/arxiv.1201.0490.

[71] Google Research, “Colaboratory Frequently Asked Questions,” 2024. Accessed: May 17, 2026. [Online]. Available: https://research.google.com/colaboratory/faq.html

[72] M. R. Saputra and P. Parjito, “Analisis Sentimen Twitter Terhadap Konflik Di Papua Menggunakan Perbandingan Naive Bayes Dan Svm,” JIPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika), vol. 10, no. 2, pp. 1197–1208, Mar. 2025, doi: 10.29100/jipi.v10i2.6180.

[73] J. Vallen and A. H. Azizah, “Analisis Sentimen Media Sosial X Program Makanan Sehat Gratis dengan Support Vector Machine dan Naive Bayes,” Jurnal Ilmiah Sistem Informasi Dan Ilmu Komputer, vol. 5, no. 3, pp. 308–318, Oct. 2025, doi: 10.55606/juisik.v5i3.1665.

[74] R. W. Harwenda, M. D. Angelo, I. Budi, A. B. Santoso, and P. K. Putra, “Sentiment Analysis on Government Public Policies: A Systematic Literature Review,” *Dinasti International Journal of Education*

Management and Social Science, vol. 6, no. 5, pp. 4192–4211, Jul. 2025, doi: 10.38035/dijemss.v6i5.4699.

[75] H. Satria, "Tweet-Harvest," GitHub repository, 2024. Available: https://github.com/helmisatria/tweet-harvest

[76] Puspitasari, L., Sumbodo, Y. P., & Rofi'ah, R. A content analysis of public discourse on Indonesia's free nutritious meal program (Makan Bergizi Gratis) on platform X. SOSIOHUMANIORA: Jurnal Ilmiah Ilmu Sosial Dan Humaniora., 12(1). 54-70. https://doi.org/10.30738/sosio.v12i1.20432.

[77] I. Mozetič, L. Torgo, V. Cerqueira, dan J. Smailović, "How to evaluate sentiment classifiers for Twitter time-ordered data?" *PLoS ONE*, vol. 13, no. 3, art. e0194317, 2018, doi: 10.1371/journal.pone.0194317.

[78] A. Alves, C. de S. Baptista, D. O. Serrano de Andrade, M. G. de Oliveira, dan A. B. de Oliveira, "A spatiotemporal approach for social media sentiment analysis," *First Monday*, vol. 26, no. 8, 2021, doi: 10.5210/fm.v26i8.10757.

[79] M. N. Y. Utomo, "python-sentianalysis-id," GitHub repository, 2017. Available: https://github.com/yasirutomo/python-sentianalysis-id.

[80] F. Koto dan G. Y. Rahmaningtyas, "InSet Lexicon: Evaluation of a Word List for Indonesian Sentiment Analysis in Microblogs," 2017 International Conference on Asian Language Processing (IALP), Singapura, 2017, pp. 391 sampai 394, doi: 10.1109/IALP.2017.8300625.

[81] O. Kolchyna, T. T. P. Souza, P. C. Treleaven, dan T. Aste, "Twitter Sentiment Analysis: Lexicon Method, Machine Learning Method and Their Combination," arXiv preprint arXiv:1507.00955, 2015. https://doi.org/10.48550/arXiv.1507.00955.

[82] D. Suhartono, W. Wongso, and A. T. Handoyo, "IdSarcasm: Benchmarking and Evaluating Language Models for Indonesian Sarcasm Detection," *IEEE Access*, vol. 12, pp. 87323–87332, 2024. https://doi.org/10.1109/ACCESS.2024.3416955.

[83] S. Rony dan J. Patman, "Standard Occupation Classifier - A Natural Language Processing Approach," arXiv preprint arXiv:2511.23057, 2025. https://arxiv.org/pdf/2511.23057.

[84] H. He and E. A. Garcia, "Learning from Imbalanced Data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 9, pp. 1263–1284, 2009, doi: 10.1109/TKDE.2008.239.

[85] R. K. Shyrokykh, M. Girnyk, and L. Dellmuth, "Short text classification with machine learning in the social sciences: The case of climate change on

Twitter," *arXiv (Cornell University)*, Oct. 2023, doi: 10.48550/arxiv.2310.04452.

[86] Python Software Foundation, "time—Time access and conversions," *Python Documentation*, bagian "time.perf_counter." [Online]. Available: https://docs.python.org/3/library/time.html#time.perf_counter.

[87] A. Mueller, "wordcloud.WordCloud," *WordCloud Documentation*. [Online]. Available: https://amueller.github.io/word_cloud/generated/wordcloud.WordCloud.html

[88] Scikit-learn developers, "Common pitfalls and recommended practices," *Scikit-learn documentation*. Accessed: Sep. 14, 2026. [Online]. Available: Dokumentasi pencegahan data leakage https://scikit-learn.org/stable/common_pitfalls.html#how-to-avoid-data-leakag