

DAFTAR PUSTAKA

- [1] Y. Pusparisa, “Indonesia Peringkat ke-3 Global Memanfaatkan Aplikasi Kesehatan,” Katadata.co.id. Accessed: Oct. 27, 2025. [Online]. Available: https://databoks.katadata.co.id/produk-konsumen/statistik/e7e5f757216d534/indonesia-peringkat-ke-3-global-memanfaatkan-aplikasi-kesehatan#:~:text=%20Minyak%20Mentah.%20*%20Nilai%20Tukar.%20*%20Harga%20Emas.
- [2] Badan Pusat Statistik, “Cerita Data Statistik untuk Indonesia: Transformasi Digital – Menelusuri Asosiasi Digitalisasi dengan Pendidikan dan Kesehatan,” Apr. 2025. Accessed: Oct. 27, 2025. [Online]. Available: <https://www.bps.go.id/id/publication/2025/04/30/675c728bd4ed47751a4f3c36/cerita-data-statistik-untuk-indonesia---transformasi-digital--menelusuri-asosiasi-digitalisasi-dengan-pendidikan-dan-kesehatan.html>
- [3] A. Muhawarman, “Digitalisasi Permudah Akses Layanan Kesehatan Primer,” *Kementerian Kesehatan RI*, p. 1, Dec. 10, 2024. Accessed: Oct. 27, 2025. [Online]. Available: <https://kemkes.go.id/id/digitalisasi-permudah-akses-layanan-kesehatan-primer?>
- [4] Tim Penulis, “Kemenkes: 8.362 Faskes di Indonesia Sudah Terkoneksi ke Aplikasi SatuSehat,” *Koran Jakarta*, p. 1, Feb. 17, 2024. Accessed: Oct. 27, 2025. [Online]. Available: <https://koran-jakarta.com/2024-02-17/kemenkes-8-362-faskes-di-indonesia-sudah-terkoneksi-ke-aplikasi-satusehat?>
- [5] Raidi, “Survey Shows 69% of Young Indonesians Rely on Telehealth Apps for Healthcare Needs,” *Indonesia Sentinel*. Accessed: Oct. 27, 2025. [Online]. Available: <https://indonesiasentinel.com/survey-shows-69-of-young-indonesians-rely-on-telehealth-apps-for-healthcare-needs/?>
- [6] PT Siloam International Hospitals Tbk, “Laporan Tahunan Annual Report,” 2024. [Online]. Available: www.siloamhospitals.com
- [7] A. Lia, A. Rahim, and T. A. Yoga Siswa, “Analisis Sentimen Aplikasi Mysiloam Menggunakan Metode Naïve Bayes,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 13, no. 1, Jan. 2025, doi: 10.23960/jitet.v13i1.5997.

- [8] Google Playstore, “Aplikasi MySiloam,” Google Playstore. Accessed: Nov. 02, 2025. [Online]. Available: <https://play.google.com/store/apps/details?id=com.siloam.android>
- [9] Y. C. Hua, P. Denny, J. Wicker, and K. Taskova, “A systematic review of aspect-based sentiment analysis: domains, methods, and trends,” *Artif. Intell. Rev.*, vol. 57, no. 11, p. 296, Sep. 2024, doi: 10.1007/s10462-024-10906-z.
- [10] A. Sari Pranasti, “Penerapan Analisis Sentimen Berbasis Aspek Menggunakan Mesin Vektor Pendukung pada Aplikasi Layanan Medis Digital,” *Journal of Comprehensive Science (JCS)*, vol. 4, no. 2, pp. 780–797, Feb. 2025, doi: 10.59188/jcs.v4i2.3047.
- [11] H. Jelodar *et al.*, “Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey,” *Multimed. Tools Appl.*, vol. 78, no. 11, pp. 15169–15211, Jun. 2019, doi: 10.1007/s11042-018-6894-4.
- [12] R. Egger and J. Yu, “A Topic Modeling Comparison Between LDA, NMF, Top2Vec, and BERTopic to Demystify Twitter Posts,” *Frontiers in Sociology*, vol. 7, May 2022, doi: 10.3389/fsoc.2022.886498.
- [13] M. Hankar, M. Kasri, and A. Beni-Hssane, “A comprehensive overview of topic modeling: Techniques, applications and challenges,” *Neurocomputing*, vol. 628, p. 129638, May 2025, doi: 10.1016/j.neucom.2025.129638.
- [14] M. Grootendorst, “BERTopic: Neural topic modeling with a class-based TF-IDF procedure,” Mar. 2022, [Online]. Available: <http://arxiv.org/abs/2203.05794>
- [15] L. Ma *et al.*, “AI-powered topic modeling: comparing LDA and BERTopic in analyzing opioid-related cardiovascular risks in women,” *Exp. Biol. Med.*, vol. 250, Feb. 2025, doi: 10.3389/ebm.2025.10389.
- [16] B. Ogunleye, T. Maswera, L. Hirsch, J. Gaudoin, and T. Brunson, “Comparison of Topic Modelling Approaches in the Banking Context,” *Applied Sciences*, vol. 13, no. 2, p. 797, Jan. 2023, doi: 10.3390/app13020797.
- [17] M. Muhajir, D. Rosadi, and Danardono, “Integrating Decision Tree and BIRCH Clustering Algorithms of BERTopic for Analyzing Public Sentiment on Dirtyvote Movie,” *Mathematical Modelling of Engineering Problems*, vol. 11, no. 12, pp. 3391–3401, Dec. 2024, doi: 10.18280/mmep.111217.
- [18] A. Wibowo, I. Indarti, and D. Laraswati, “Komparasi Algoritma Decision Tree, Random Forest dan SVM untuk Prognosis COVID-19,” *IMTechno: Journal of*

- Industrial Management and Technology*, vol. 5, no. 2, pp. 10–15, Jul. 2024, doi: 10.31294/imtechno.v5i2.2868.
- [19] A. Nurian, M. S. Ma'arif, I. N. Amalia, and C. Rozikin, "Analisis Sentimen Pengguna Aplikasi Shopee Pada Situs Google Play Menggunakan Naive Bayes Classifier," *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 1, Jan. 2024, doi: 10.23960/jitet.v12i1.3631.
- [20] S. A. R. Rizaldi, S. Alam, and I. Kurniawan, "Analisis Sentimen Pengguna Aplikasi JMO (Jamsostek Mobile) Pada Google Play Store Menggunakan Metode Naive Bayes," *STORAGE: Jurnal Ilmiah Teknik dan Ilmu Komputer*, vol. 2, no. 3, pp. 109–117, Aug. 2023, doi: 10.55123/storage.v2i3.2334.
- [21] Mafda Khoirotul Fatha, Seftin Fitri Ana Wati, Bhagas Satrya Dewa, and Krisna Eko Prasetyo, "Peran Big Data Pada Intelijen Bisnis Sebagai Sistem Pendukung Keputusan (Systematic Literature Review)," *Prosiding Seminar Nasional Teknologi dan Sistem Informasi*, vol. 3, no. 1, pp. 318–326, Nov. 2023, doi: 10.33005/sitasi.v3i1.612.
- [22] W. Zhu, S. Proksch, D. M. German, M. W. Godfrey, L. Li, and S. McIntosh, "What is an app store? The software engineering perspective," *Empir. Softw. Eng.*, vol. 29, no. 1, p. 35, Jan. 2024, doi: 10.1007/s10664-023-10362-3.
- [23] S. Jansen and E. Bloemendal, "Defining App Stores: The Role of Curated Marketplaces in Software Ecosystems," 2013, pp. 195–206. doi: 10.1007/978-3-642-39336-5_19.
- [24] D. Martens and W. Maalej, "Towards understanding and detecting fake reviews in app stores," *Empir. Softw. Eng.*, vol. 24, no. 6, pp. 3316–3355, Dec. 2019, doi: 10.1007/s10664-019-09706-9.
- [25] Siloam Hospital, "Press Release For Immediate Release Siloam Hospitals Reported Sustained Growth through Execution of Its 5-Year Strategy," 2024. [Online]. Available: www.siloamhospitals.com.
- [26] M. Amien, *Sejarah dan Perkembangan Teknik Natural Language Processing (NLP) Bahasa Indonesia: Tinjauan tentang sejarah, perkembangan teknologi, dan aplikasi NLP dalam bahasa Indonesia*. 2023. doi: 10.48550/arXiv.2304.02746.
- [27] Badry Ali Mustofa and Wawan Laksito Yuly Saptomo, "Use of Natural Language Processing in Social Media Text Analysis," *Journal of Artificial*

- Intelligence and Engineering Applications (JAIEA)*, vol. 4, no. 2, pp. 1235–1238, Feb. 2025, doi: 10.59934/jaiea.v4i2.875.
- [28] Z. Kastrati, F. Dalipi, A. S. Imran, K. Pireva Nuci, and M. A. Wani, “Sentiment Analysis of Students’ Feedback with NLP and Deep Learning: A Systematic Mapping Study,” *Applied Sciences*, vol. 11, no. 9, p. 3986, Apr. 2021, doi: 10.3390/app11093986.
- [29] A. Hotho, A. Nürnberger, and G. Paaß, “A Brief Survey of Text Mining,” *Journal for Language Technology and Computational Linguistics*, vol. 20, no. 1, pp. 19–62, Jul. 2005, doi: 10.21248/jlcl.20.2005.68.
- [30] C. Uma, S. Krithika, and C. Kalaivani, “A Survey Paper on Text Mining Techniques,” *International Journal of Engineering Trends and Technology*, vol. 40, no. 4, pp. 225–229, Oct. 2016, doi: 10.14445/22315381/IJETT-V40P237.
- [31] V. B. Kobayashi, S. T. Mol, H. A. Berkers, G. Kismihók, and D. N. Den Hartog, “Text Mining in Organizational Research,” *Organ. Res. Methods*, vol. 21, no. 3, pp. 733–765, Jul. 2018, doi: 10.1177/1094428117722619.
- [32] N. Nofiyani and W. Wulandari, “Implementasi Electronic Data Processing Untuk meningkatkan Efektifitas dan Efisiensi Pada Text Mining,” *Jurnal Media Informatika Budidarma*, vol. 6, no. 3, p. 1621, Jul. 2022, doi: 10.30865/mib.v6i3.4332.
- [33] S. Shevira, I. M. A. D. Suarjaya, and P. W. Buana, “Pengaruh Kombinasi dan Urutan Pre-Processing pada Tweets Bahasa Indonesia,” *JITTER : Jurnal Ilmiah Teknologi dan Komputer*, vol. 3, no. 2, p. 1074, Jul. 2022, doi: 10.24843/JTRTI.2022.v03.i02.p06.
- [34] I. S. Lundam and J. B. B. Darmawan, “Analisis Sentimen Ulasan Aplikasi Gojek Di Google Play Store Menggunakan Multinomial Naïve Bayes,” *ReTII*, pp. 149–156, Nov. 2024, [Online]. Available: <https://journal.itny.ac.id/index.php/ReTII/article/view/5432>
- [35] M. F. Juna and M. Hayaty, “The observed preprocessing strategies for doing automatic text summarizing,” *Computer Science and Information Technologies*, vol. 4, no. 2, pp. 119–126, Jul. 2023, doi: 10.11591/csit.v4i2.pp119-126.
- [36] Pande sindu, Agus Aan Jiwa Permana, and I Nyoman Saputra Wahyu Wijaya, “Identifikasi Dan Normalisasi Teks Slang Dengan Fasttext Pada Twitter Dalam Bahasa Indonesia,” *Jurnal Pendidikan Teknologi dan Kejuruan*, vol. 21, no. 1, pp. 33–44, Jan. 2024, doi: 10.23887/jptkundiksha.v21i1.66381.

- [37] N. P. A. Widiari, I. M. A. D. Suarjaya, and D. P. Githa, “Teknik Data Cleaning Menggunakan Snowflake untuk Studi Kasus Objek Pariwisata di Bali,” *Jurnal Ilmiah Merpati (Menara Penelitian Akademika Teknologi Informasi)*, p. 137, Jul. 2020, doi: 10.24843/JIM.2020.v08.i02.p07.
- [38] L. A. Mullen, K. Benoit, O. Keyes, D. Selivanov, and J. Arnold, “Fast, Consistent Tokenization of Natural Language Text,” *J. Open Source Softw.*, vol. 3, no. 23, p. 655, Mar. 2018, doi: 10.21105/joss.00655.
- [39] J. J. Webster and C. Kit, “Tokenization as the initial phase in NLP,” in *Proceedings of the 14th conference on Computational linguistics -*, Morristown, NJ, USA: Association for Computational Linguistics, 1992, p. 1106. doi: 10.3115/992424.992434.
- [40] D. J. Ladani and N. P. Desai, “Stopword Identification and Removal Techniques on TC and IR applications: A Survey,” in *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)*, IEEE, Mar. 2020, pp. 466–472. doi: 10.1109/ICACCS48705.2020.9074166.
- [41] A. F. Hidayatullah, “Pengaruh Stopword Terhadap Performa Klasifikasi Tweet Berbahasa Indonesia,” *JISKA (Jurnal Informatika Sunan Kalijaga)*, vol. 1, no. 1, pp. 1–4, May 2016, doi: 10.14421/jiska.2016.11-01.
- [42] A. A. Magriyanti, “Analisis Pengembangan Algoritma Porter Stemming Dalam Bahasa Indonesia,” Sep. 29, 2018. doi: 10.31227/osf.io/7ge4v.
- [43] I. P. S. A. Taufik, “Analisis Sentimen Terhadap Tokoh Publik Menggunakan Algoritma Support Vector Machine (SVM),” 2018. Accessed: Oct. 27, 2025. [Online]. Available: https://repository.uinjkt.ac.id/dspace/bitstream/123456789/59802/1/ILHAM%20TAUFIK%20-%20FST.pdf?utm_source=chatgpt.com
- [44] T. Walasary, “Survey Paper tentang Analisis Sentimen,” *KONSTELASI: Konvergensi Teknologi dan Sistem Informasi*, vol. 2, no. 1, Apr. 2022, doi: 10.24002/konstelasi.v2i1.5378.
- [45] C. Matulesy and A. Purnama, “Analisis Sentimen Terhadap Review Pengguna Indrive Di Google Playstore Menggunakan Algoritma Support Vector Machine,” *Jurnal Darma Agung*, vol. 31, no. 1, p. 1015, Apr. 2023, doi: 10.46930/ojsuda.v31i1.3078.
- [46] N. Laskari and S. Sanampudi, “Aspect Based Sentiment Analysis Survey,” *IOSR Journals*, vol. 18, pp. 2278–2661, Mar. 2016, doi: 10.9790/0661-18212428.

- [47] E. I. Setiawan, “Aspect-Based Sentiment Analysis of Healthcare Reviews from Indonesian Hospitals based on Weighted Average Ensemble,” *Journal of Applied Data Sciences*, vol. 5, no. 4, pp. 1579–1596, Dec. 2024, doi: 10.47738/jads.v5i4.328.
- [48] Gading Sectio Aryoseto, Is Mardianto, and Anung B. Ariwibowo, “Recommendation System for Mental Health Article on Circle Application,” *Intelmatiks*, vol. 3, no. 1, pp. 1–6, Jan. 2023, doi: 10.25105/itm.v3i1.16343.
- [49] M. Allaoui, M. L. Kherfi, and A. Cheriet, “Considerably Improving Clustering Algorithms Using UMAP Dimensionality Reduction Technique: A Comparative Study,” 2020, pp. 317–325. doi: 10.1007/978-3-030-51935-3_34.
- [50] H. Al Rasyid, B. F. K. Soebari, and D. S. Y. Kartika, “Implementasi Algoritma K-Means Clustering Untuk Pengelompokan Penjualan Produk Pada Online Shop Toko Gizi,” *Prosiding Seminar Nasional Teknologi dan Sistem Informasi*, vol. 2, no. 1, pp. 242–248, Sep. 2022, doi: 10.33005/sitasi.v2i1.304.
- [51] P. R. Sampaio and H. Maxcici, “Unsupervised Document and Template Clustering using Multimodal Embeddings,” 2025. [Online]. Available: <https://arxiv.org/abs/2506.12116>
- [52] A. Eklund, M. Forsman, and F. Drewes, “Empirical Configuration Study of a Common Document Clustering Pipeline,” *Northern European Journal of Language Technology*, vol. 9, no. 1, Sep. 2023, doi: 10.3384/nejlt.2000-1533.2023.4396.
- [53] P. J. Rousseeuw, “Silhouettes: A graphical aid to the interpretation and validation of cluster analysis,” *J. Comput. Appl. Math.*, vol. 20, pp. 53–65, Nov. 1987, doi: 10.1016/0377-0427(87)90125-7.
- [54] D. L. Davies and D. W. Bouldin, “A Cluster Separation Measure,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-1, no. 2, pp. 224–227, Apr. 1979, doi: 10.1109/TPAMI.1979.4766909.
- [55] I. F. Ashari, E. Dwi Nugroho, R. Baraku, I. Novri Yanda, and R. Liwardana, “Analysis of Elbow, Silhouette, Davies-Bouldin, Calinski-Harabasz, and Rand-Index Evaluation on K-Means Algorithm for Classifying Flood-Affected Areas in Jakarta,” *Journal of Applied Informatics and Computing*, vol. 7, no. 1, pp. 89–97, Jul. 2023, doi: 10.30871/jaic.v7i1.4947.
- [56] J. P. Lim and H. Lauw, “Large-Scale Correlation Analysis of Automated Metrics for Topic Models,” in *Proceedings of the 61st Annual Meeting of the Association*

for Computational Linguistics (Volume 1: Long Papers), Stroudsburg, PA, USA: Association for Computational Linguistics, 2023, pp. 13874–13898. doi: 10.18653/v1/2023.acl-long.776.

- [57] F. Rosner, A. Hinneburg, M. Röder, M. Nettling, and A. Both, “Evaluating topic coherence measures,” 2014. [Online]. Available: <https://arxiv.org/abs/1403.6397>
- [58] D. Mimno, H. Wallach, E. Talley, M. Leenders, and A. McCallum, “Optimizing Semantic Coherence in Topic Models,” in *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, R. Barzilay and M. Johnson, Eds., Edinburgh, Scotland, UK.: Association for Computational Linguistics, Jul. 2011, pp. 262–272. [Online]. Available: <https://aclanthology.org/D11-1024/>
- [59] S. Syed and M. Spruit, “Full-Text or Abstract? Examining Topic Coherence Scores Using Latent Dirichlet Allocation,” in *2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, IEEE, Oct. 2017, pp. 165–174. doi: 10.1109/DSAA.2017.61.
- [60] G. Bouma, “Normalized (Pointwise) Mutual Information in Collocation Extraction,” *Proceedings of the Biennial GSCL Conference 2009*, Jan. 2009.
- [61] J. H. Lau, D. Newman, and T. Baldwin, “Machine Reading Tea Leaves: Automatically Evaluating Topic Coherence and Topic Model Quality,” in *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2014, pp. 530–539. doi: 10.3115/v1/E14-1056.
- [62] A. B. Dieng, F. J. R. Ruiz, and D. M. Blei, “Topic Modeling in Embedding Spaces,” *Trans. Assoc. Comput. Linguist.*, vol. 8, pp. 439–453, Dec. 2020, doi: 10.1162/tacl_a_00325.
- [63] Klaus Krippendorff, “Computing Krippendorff’s Alpha-Reliability,” 2011. Accessed: May 18, 2026. [Online]. Available: <https://repository.upenn.edu/entities/publication/034a6030-c584-4d14-9d3d-7b7e8d16df20>
- [64] M. H. Setiono, “A Komparasi Algoritma Decision Tree, Random Forest, Svm Dan K-Nn Dalam Klasifikasi Kepuasan Penumpang Maskapai Penerbangan,” *INTI Nusa Mandiri*, vol. 17, no. 1, pp. 32–39, Oct. 2022, doi: 10.33480/inti.v17i1.3420.

- [65] M. Fahmy Amin, “Confusion Matrix in Binary Classification Problems: A Step-by-Step Tutorial,” *Journal of Engineering Research*, vol. 6, no. 5, pp. 0–0, Dec. 2022, doi: 10.21608/erjeng.2022.274526.
- [66] Google Developers, “Classification: Accuracy, Precision, Recall, and Related Metrics.” Accessed: Oct. 28, 2025. [Online]. Available: https://developers.google.com/machine-learning/crash-course/classification/accuracy-precision-recall?utm_source=chatgpt.com&hl=id
- [67] Miguel Grinberg, *Flask Web Development: Developing Web Applications with Python*, 2nd ed. Sebastopol: O’Reilly Media, 2018.
- [68] Jon Duckett, *HTML and CSS: Design and Build Websites*. Indianapolis: John Wiley & Sons, 2011.
- [69] M. Muhajir, G. Gunardi, D. Danardono, and D. Rosadi, “Integrating IndoBERT and balanced iterative reducing and clustering using hierarchies of BERTopic in Indonesian short text,” *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 14, no. 5, p. 4192, Oct. 2025, doi: 10.11591/ijai.v14.i5.pp4192-4201.
- [70] N. N. Hidayati and S. Shaleha, “BERTopic Analysis of Indonesian Biodiversity Policy on Social Media,” *ECTI Transactions on Computer and Information Technology (ECTI-CIT)*, vol. 18, no. 3, pp. 260–271, May 2024, doi: 10.37936/ecti-cit.2024173.255058.
- [71] A. Eklund, M. Forsman, and F. Drewes, “Empirical Configuration Study of a Common Document Clustering Pipeline,” *Northern European Journal of Language Technology*, vol. 9, no. 1, Sep. 2023, doi: 10.3384/nejlt.2000-1533.2023.4396.
- [72] Baromim Triwijaya, Setyoningsih Wibowo, and Nur Latifah Dwi Mutiara Sari, “Performance Comparison of K-Means Algorithm and BIRCH Algorithm in Clustering Earthquake Data in Indonesia with Web-Based Map Visualization,” *Jurnal Teknologi Dan Open Source*, vol. 8, no. 1, pp. 278–287, Jun. 2025, doi: 10.36378/jtos.v8i1.4400.
- [73] M. Röder, A. Both, and A. Hinneburg, “Exploring the Space of Topic Coherence Measures,” in *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, New York, NY, USA: ACM, Feb. 2015, pp. 399–408. doi: 10.1145/2684822.2685324.

- [74] A. B. Dieng, F. J. R. Ruiz, and D. M. Blei, “Topic Modeling in Embedding Spaces,” *Trans. Assoc. Comput. Linguist.*, vol. 8, pp. 439–453, Dec. 2020, doi: 10.1162/tacl_a_00325.
- [75] G. Marzi, M. Balzano, and D. Marchiori, “K-Alpha Calculator–Krippendorff’s Alpha Calculator: A user-friendly tool for computing Krippendorff’s Alpha inter-rater reliability coefficient,” *MethodsX*, vol. 12, p. 102545, Jun. 2024, doi: 10.1016/j.mex.2023.102545.
- [76] N. P. Rahmawati, Irwan Budiman, Muhammad Itqan Mazdadi, Andi Farmadi, and Friska Abadi, “Comparison Between K-Fold Cross Validation And Percentage Split In Decision Tree Algorithms For Anemia Classification,” *Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics*, vol. 8, no. 1, pp. 105–114, Feb. 2026, doi: 10.35882/ijeeemi.v8i1.315.
- [77] M. N. R. N. Suryono, “Aspect-Based Sentiment Analysis Pada Ulasan Aplikasi Access By Kai Dengan Menggunakan Metode Tf-Idf Dan Algoritma Support Vector Machine,” UPN “Veteran” Jawa Timur, 2025. Accessed: Nov. 14, 2025. [Online]. Available: <https://repository.upnjatim.ac.id/38522/>
- [78] Y. C. Hua, P. Denny, J. Wicker, and K. Taskova, “A systematic review of aspect-based sentiment analysis: domains, methods, and trends,” *Artif. Intell. Rev.*, vol. 57, no. 11, p. 296, Sep. 2024, doi: 10.1007/s10462-024-10906-z.
- [79] A. Nugroho, Wiyanto, and D. Maulana, “Comparative Analysis Of Classification Algorithms In Handling Imbalanced Data With Smote Oversampling Approach,” *JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer)*, vol. 11, no. 2, pp. 487–495, Nov. 2025, doi: 10.33480/jitk.v11i2.6956.
- [80] Hyuto, “Analisis Sentimen Corona DKI Jakarta,” 2020, *Github*.
- [81] Mariya Mansurova, “Topics per Class Using BERTopic: How to understand the differences in texts by categories,” Medium (TDS Archive). Accessed: May 10, 2026. [Online]. Available: <https://medium.com/data-science/topics-per-class-using-bertopic-252314f2640>
- [82] C. Yuan and H. Yang, “Research on K-Value Selection Method of K-Means Clustering Algorithm,” *J (Basel)*, vol. 2, no. 2, pp. 226–235, Jun. 2019, doi: 10.3390/j2020016.

- [83] Muhammad Raqib Syahkur, D. Hartama, and S. Solikhun, “Evaluasi Jumlah Cluster pada Algoritma K-Means++ Menggunakan Silhouette dan Elbow dengan Validasi Nilai DBI dalam Mengelompokkan Gizi Balita,” *JST (Jurnal Sains dan Teknologi)*, vol. 13, no. 3, pp. 487–496, Oct. 2024, doi: 10.23887/jstundiksha.v13i3.86419.
- [84] S. Terragni, E. Fersini, B. G. Galuzzi, P. Tropeano, and A. Candelieri, “OCTIS: Comparing and Optimizing Topic models is Simple!,” in *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*, Stroudsburg, PA, USA: Association for Computational Linguistics, 2021, pp. 263–270. doi: 10.18653/v1/2021.eacl-demos.31.
- [85] Radim Řehůřek and Petr Sojka, “Software Framework for Topic Modelling with Large Corpora,” in *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, Valletta, Malta: ELRA, 2010. Accessed: May 08, 2026. [Online]. Available: <https://is.muni.cz/publication/884893/en>