

DAFTAR PUSTAKA

- Aditya, M. R., Ilham, M., Surianto, D. F., & Mappalotteng, A. M. (2025). Evaluasi Pengukuran Semantik Sinonim KBBI Menggunakan Pendekatan Word Embedding. *Jurnal Nasional Teknik Elektro Dan Teknologi Informasi*, 14(2), 112–120. <https://doi.org/10.22146/JNTETI.V14I2.17117>
- Aisy, S. R., & Prasetyo, B. (2023). Sentiment Analysist of the TPKS Law on Twitter Using InSet Lexicon with Multinomial Naïve Bayes and Support Vector Machine Based on Soft Voting. *Recursive Journal of Informatics*, 1(2), 93–101. <https://doi.org/10.15294/RJI.V1I2.68324>
- Akbar, I. F., Laksana, T. G., Arifa, A. B., & Silalahi, M. R. (2023). Pengelompokan Teks Berita Utama dengan Metode LDA (Latent Dirichlet Allocation) melalui Pemahaman Pemodelan Topik. *Proceedings of the National Conference on Electrical Engineering, Informatics, Industrial Technology, and Creative Media*, 3(1), 475–484.
- Amalia, J. (2022). Membangun Slang Dictionary Untuk Normalisasi Teks Menggunakan Pre-Trained Fasttext Model. *JSR : Jaringan Sistem Informasi Robotik*, 6(2), 250–256. <https://doi.org/10.58486/JSR.V6I2.184>
- Ariansyah, A., & Indahyanti, U. (2024). Fitur Ekstraksi pada Pemodelan Topik Menggunakan Metode Latent Dirichlet Allocation pada Peristiwa Kebocoran Data. *Indonesian Journal of Applied Technology*, 1(2), 24–24. <https://doi.org/10.47134/IJAT.V1I2.3041>
- Arifin, S., & Febryanto, B. A. (2025). Analisis Sentimen Pengguna Aplikasi Merdeka Mengajar Menggunakan Metode Vader Lexicon. *SISFOTENIKA*, 15(1), 1–11. <https://doi.org/10.30700/SISFOTENIKA.V15I1.513>
- Ashmawy, M., Fakhr, M. W., & Maghraby, F. A. (2023). Lexical Normalization Using Generative Transformer Model (LN-GTM). *International Journal of Computational Intelligence Systems*, 16(1), 1–11. <https://doi.org/10.1007/S44196-023-00366-8/TABLES/8>
- Asri, J. S., & Wahyu, S. (2021). Analisis Sentimen Menerapkan Lexicon-Learning Based Untuk Melihat Opini Masyarakat Mengenai Protokol Kesehatan Dan

- Perkembangan Vaksin Covid-19 Di Indonesia Menggunakan Dataset Twitter. *Konf. Nas. Ilmu Komput*, 530–536.
- Asyhar, E. (2024). Analisis Sentimen dan Pemodelan Topik Terhadap Ulasan Aplikasi Jenius Menggunakan Metode Support Vector Machine dan Latent Dirichlet Allocation. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 8(9).
- Azzahra, S. A., Wachid, N., & Majid, A. (2025). Klasifikasi dan Analisis Semantik Cyberbullying Sosial Media X: Integrasi Web Scraping dan Natural Language Processing (NLP). *Jurnal Educatio FKIP UNMA*, 11(2), 353–360. <https://doi.org/10.31949/EDUCATIO.V11I2.12725>
- Bhatia, S., Lau, J. H., & Baldwin, T. (2018). Topic Intrusion for Automatic Topic Model Evaluation. *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, EMNLP 2018*, 844–849. <https://doi.org/10.18653/V1/D18-1098>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3.
- Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2016). Enriching Word Vectors with Subword Information. *Transactions of the Association for Computational Linguistics*, 5, 135–146. https://doi.org/10.1162/tacl_a_00051
- Budiman, A. E., & Widjaja, A. (2020). Analisis Pengaruh Teks Preprocessing Terhadap Deteksi Plagiarisme Pada Dokumen Tugas Akhir. *Jurnal Teknik Informatika Dan Sistem Informasi*, 6(3). <https://doi.org/10.28932/JUTISI.V6I3.2892>
- Bustamin, A. (2025). Text Normalization For Indonesian Slang Words In Sentiment Analysis Development. *ICIC Express Letters, Part B: Applications*, 16(2), 121–129. <https://doi.org/10.24507/icicelb.16.02.121>
- Chang, J., Boyd-Graber, J., Gerrish, S., Wang, C., & Blei, D. M. (2009). Reading Tea Leaves: How Humans Interpret Topic Models. *Advances in Neural Information Processing Systems*, 22. <http://rex.info>
- Cisatra, A., Daris, D., Ansori, M., Fathya, D. N., Akmal, A., Ramadhani, S., & Rakhmawati, N. A. (2024). Identifikasi Topik Hangat di Media Berita

- Menggunakan Latent Dirichlet Allocation. *JIEET (Journal of Information Engineering and Educational Technology)*, 8(1), 14–17. <https://doi.org/10.26740/JIEET.V8N1.P14-17>
- Clairine, A., Lestari, E. I. D., Wiyono, E. N., & R, M. W. (2025). Ekspresi Keresahan Pemuda melalui Media Sosial: Studi Kritik terhadap Narasi Indonesia Gelap. *Moderasi: Jurnal Studi Ilmu Pengetahuan Sosial*, 6(1), 36–51. <https://doi.org/10.24239/MODERASI.VOL6.ISS1.491>
- Darmawan, I. P. D. W., Pradnyana, G. A., & Pascima, I. B. N. (2023). Optimasi Parameter Support Vector Machine Dengan Algoritma Genetika Untuk Analisis Sentimen Pada Media Sosial Instagram. *SINTECH (Science and Information Technology) Journal*, 6(1), 58–67. <https://doi.org/10.31598/SINTECHJOURNAL.V6I1.1245>
- Dewi, N. L. P. R., Wijaya, I. N. S. W., Purnamawan, I. K., & Marti, N. W. (2024). Model Classifier Judul Berita Pariwisata Indonesia Berdasarkan Sentimen. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 11(1), 117–124. <https://doi.org/10.25126/JTIK.20241117617>
- Egger, R., & Gokce, E. (2022). Natural Language Processing (NLP): An Introduction. *Tourism on the Verge, Part F1051*, 307–334. https://doi.org/10.1007/978-3-030-88389-8_15
- HaCohen-Kerner, Y., Miller, D., & Yigal, Y. (2020). The influence of preprocessing on text classification using a bag-of-words representation. *PLOS ONE*, 15(5), e0232525. <https://doi.org/10.1371/JOURNAL.PONE.0232525>
- Halimi, A., Kusri, K., & Arief, M. R. (2021). Analisis Sentimen Masyarakat Indonesia Terhadap Pembelajaran Online dari Media Sosial Twitter Menggunakan Lexicon dan K-Nearest Neighbor. *COREAI: Jurnal Kecerdasan Buatan, Komputasi Dan Teknologi Informasi*, 2(1), 18–28.
- Hapsari, Y., Mujahidin, S., & Rizqiya Fadhliana, N. (2023). Analisis Sentimen Isu Vaksinasi COVID-19 Pada Twitter Menggunakan Metode Naïve Bayes dan Pembobotan TF-IDF Tokenisasi 1-2. *SPECTA Journal of Technology*, 7(2), 573–583. <https://doi.org/10.35718/SPECTA.V7I2.812>

- Herlando, F., Dzikrillah, A. R., Nufairi, F., Sinduningrum, E., & Sholeh, M. (2024). Analisis Perbandingan Sentiment Dan Perbincangan Netizen Terhadap Twitter Pasca Pergantian Nama. *JUPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 9(1), 360–367. <https://doi.org/10.29100/JUPI.V9I1.4934>
- Hoyle, A., Goel, P., Hian-Cheong, A., Peskov, D., Boyd-Graber, J., & Resnik, P. (2021). Is Automated Topic Model Evaluation Broken? The Incoherence of Coherence. *Advances in Neural Information Processing Systems*, 34, 2018–2033.
- Sindu, P., Permana, A. A. J., & Wijaya, I. N. S. W. (2024). Identifikasi Dan Normalisasi Teks Slang Dengan Fasttext Pada Twitter Dalam Bahasa Indonesia. *Jurnal Pendidikan Teknologi Dan Kejuruan*, 21(1), 33–44. <https://doi.org/10.23887/JPTKUNDIKSHA.V21I1.66381>
- Irawan, T., Mutawalli, L., Fadli, S., & Bagye, W. (2024). Topic Modelling Pola Komunikasi Pilpres 2024 Focus Web Scraping Dan Latent Dirichlet Allocation. *Jurnal Manajemen Informatika Dan Sistem Informasi*, 7(2), 186–194. <https://doi.org/10.36595/MISI.V7I2.1183>
- Kamil, A. I., Pratiwi, O. N., & Witarsyah, D. (2025). Analisis Sentimen dan Pemodelan Topik terhadap Aplikasi Pembelajaran Online pada Platform Google Play. *JUPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, 10(2), 836–849. <https://doi.org/10.29100/JUPI.V10I2.6023>
- Khadijah, U. N., & Cahyono, N. (2024). Analisis Topic Modelling Pariwisata Yogyakarta Menggunakan Latent Dirichlet Allocation (LDA). *The Indonesian Journal of Computer Science*, 13(4). <https://doi.org/10.33022/IJCS.V13I4.3816>
- Lee, K. P., & Song, S. (2022). Developing insights from the collective voice of target users in Twitter. *Journal of Big Data*, 9(1), 1–31. <https://doi.org/10.1186/S40537-022-00611-5/FIGURES/5>
- Lumbantoruan, R., Puspita, M., Nababan, S., & Saragih, L. A. (2024). Analisis Perbandingan FastText dan Word2Vec pada Sistem Temu Balik Informasi. *PROSIDING SEMINAR NASIONAL SAINS DATA*, 4(1), 1033–1041. <https://doi.org/10.33005/SENADA.V4I1.416>

- Matira, Y., Junaidi, & Setiawan, I. (2023). Pemodelan Topik pada Judul Berita Online Detikcom Menggunakan Latent Dirichlet Allocation. *ESTIMASI: Journal of Statistics and Its Application*, 53–63. <https://doi.org/10.20956/EJSA.VI.24843>
- Meddeb, A., & Romdhane, L. Ben. (2022). Using Topic Modeling and Word Embedding for Topic Extraction in Twitter. *Procedia Computer Science*, 207, 790–799. <https://doi.org/10.1016/J.PROCS.2022.09.134>
- Mramra, W. Y. Y., & Nataliani, Y. (2025). Penentuan bidang unggulan akademik universitas melalui metode topik Linear Discriminant Analysis (LDA). *IT-Explore: Jurnal Penerapan Teknologi Informasi Dan Komunikasi*, 4(1), 93–105. <https://doi.org/10.24246/ITEXPLORE.V4I1.2025.PP93-105>
- Muna, Z. N., Setiawan, B. D., & Perdana, R. S. (2024). Penerapan Pemodelan Topik Komentar Melalui Media Sosial Twitter Menggunakan Latent Dirichlet Allocation (Studi Kasus: Pemerintah Kota Malang). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 8(7).
- Negara, E. S., & Triadi, D. (2021). Topic modeling using latent dirichlet allocation (LDA) on twitter data with Indonesia keyword. *Bulletin of Social Informatics Theory and Application*, 5(2), 124–132. <https://doi.org/10.31763/BUSINTA.V5I2.455>
- Nur, A. M., Kurniawan, E., & Supriadi, R. (2025). Unveiling Public Sentiment and Ideation Patterns in the #IndonesiaGelap Discourse through Appraisal Theory: A Corpus-Based Analysis. *IJLECR (International Journal of Language Education and Cultural Review)*, 11(1), 48–58. <https://doi.org/10.21009/IJLECR.V11I1.55014>
- Nur, M. A., & Wardhani, N. (2022). Optimasi Normalisasi Kata Pada Data Twitter Untuk Meningkatkan Akurasi Analisis Sentimen (Studi Kasus Respon Masyarakat Terhadap Layanan Teman Bus). *Jurnal Fokus Elektroda : Energi Listrik, Telekomunikasi, Komputer, Elektronika Dan Kendali*, 7(4), 237–243. <https://doi.org/10.33772/JFE.V7I4.21>
- Nur, R., Jen, S. M., Kapita, S. N., & Fhadli, M. (2025). Comparison of Normalization of Indonesian Slang Words Using the FastText & Word2vec

- Model with the Natural Language Processing Approach. *Nusantara Science and Technology Proceedings*, 40–49. <https://doi.org/10.11594/NSTP.2025.4805>
- Nuraini, N., Verdikha, N. A., & Latipah, A. J. (2024). Analysis of FastText with Support Vector Machine for Hate Speech Classification on Twitter Social Media. *Jurnal Informatika*, 11(2), 73–79. <https://doi.org/10.31294/INF.V11I2.21107>
- Nurhaliza, A. F. (2024). Penerapan Pemodelan Topik menggunakan Metode Latent Dirichlet Allocation terhadap Pembahasan Pemilu Indonesia tahun 2024 di Twitter. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 8(7).
- Nurmawati, E., & Amanda, A. (2023). Analisis Sentimen Dan Pemodelan Topik Pada Tweet Terkait Data Badan Pusat Statistik. *Jurnal Sistem Informasi Dan Informatika (Simika)*, 6(2), 165–176. <https://doi.org/10.47080/SIMIKA.V6I2.2789>
- Ozyurt, O., Özköse, H., & Ayaz, A. (2024). Evaluating the latest trends of Industry 4.0 based on LDA topic model. *Journal of Supercomputing*, 80(13), 19003–19030. <https://doi.org/10.1007/S11227-024-06247-X/METRICS>
- Pradnyana, G. A., & Darmawiguna, I. G. M. (2021). Web-Based System for Bali Tourism Sentiment Analysis during The Covid-19 Pandemic using Django Web Framework and Naive Bayes Method. *Proceedings of the 4th International Conference on Innovative Research Across Disciplines (ICIRAD 2021)*, 613, 316–320. <https://doi.org/10.2991/ASSEHR.K.211222.050>
- Pramayasa, K., Md, I., Maysanjaya, D., Ayu, G., & Diatri Indradewi, A. (2023). Analisis Sentimen Program Mbkm Pada Media Sosial Twitter Menggunakan KNN Dan SMOTE. *SINTECH (Science and Information Technology) Journal*, 6(2), 89–98. <https://doi.org/10.31598/SINTECHJOURNAL.V6I2.1372>
- Purwitasari, D., Ansyah, A. S. S., Kurniawan, A. P., & Kholifah, A. N. (2023). A Hybrid Method on Emotion Detection for Indonesian Tweets of COVID-19. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 7(2), 254–262. <https://doi.org/10.29207/RESTI.V7I2.4816>

- Puspita, E., Shiddieq, F., & Roji, F. F. (2024a). Pemodelan Topik pada Media Berita Online Menggunakan Latent Dirichlet Allocation (Studi Kasus Merek Somethinc). *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(2), 481–489. <https://doi.org/10.57152/MALCOM.V4I2.1204>
- Puspita, E., Shiddieq, F., & Roji, F. F. (2024b). Pemodelan Topik pada Media Berita Online Menggunakan Latent Dirichlet Allocation (Studi Kasus Merek Somethinc). *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, 4(2), 481–489. <https://doi.org/10.57152/MALCOM.V4I2.1204>
- Rahmaniar, M. N., Hartiono, A., & Pramana, S. (2024). Implementation of Machine Learning for Spam Detection and Topic Modeling for Emails in Bahasa Indonesia. *Korean Journal of Artificial Intelligence*, 12(4), 9–19. <https://doi.org/10.24225/KJAI.2024.12.4.9>
- Raif, M. I., Hidayati, N. N., & Matulatan, T. (2024). Otomatisasi Pendeteksi Kata Baku Dan Tidak Baku Pada Data Twitter Berbasis KBBI. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 11(2), 337–348. <https://doi.org/10.25126/JTIK.20241127404>
- Ramadhy, I. F., & Sibaroni, Y. (2022). Analisis Trending Topik Twitter dengan Fitur Ekspansi FastText Menggunakan Metode Logistic Regression. *JURIKOM (Jurnal Riset Komputer)*, 9(1), 1–7. <https://doi.org/10.30865/JURIKOM.V9I1.3791>
- Riyaddulloh, R., & Romadhony, A. (2021). Normalisasi Teks Bahasa Indonesia Berbasis Kamus Slang Studi Kasus: Tweet Produk Gadget Pada Twitter. In *Agustus* (Vol. 8, Issue 4). <https://openlibrary.telkomuniversity.ac.id/home/catalog/id/170317/slug/normalisasi-teks-bahasa-indonesia-berbasis-kamus-slang-studi-kasus-tweet-produk-gadget-pada-twitter.html>
- Salsabila, N. A., Winatmoko, Y. A., Septiandri, A. A., & Jamal, A. (2018). Colloquial Indonesian Lexicon. *Proceedings of the 2018 International*

- Conference on Asian Language Processing, IALP 2018*, 226–229.
<https://doi.org/10.1109/IALP.2018.8629151>
- Sanjaya, F. R. A. D., Pudjiantoro, T. H., Ningsih, A. K., & Renaldi, F. (2022). Sentiment Analysis Of E-Wallets on Twitter social media With Naïve Bayes and Lexicon-Based Methods. *Proceedings of the International Conference on Industrial Engineering and Operations Management, Michigan, USA: IEOM Society International*, 1002–1010.
- Sholihah, U. M., Findawati, Y., & Indahyanti, U. (2023). Topic Modeling In Covid-19 Vaccination Refusal Cases Using Latent Dirichlet Allocation And Latent Semantic Analysis. *Jurnal Teknik Informatika (Jutif)*, 4(5), 1063–1074.
<https://doi.org/10.52436/1.JUTIF.2023.4.5.951>
- Sidik, P., Sunarya, I. M. G., & Gunadi, I. G. A. (2025). Comparison of Random Forest and Support Vector Machine Methods in Sentiment Analysis of Student Satisfaction Questionnaire Comments at ITB STIKOM Bali. *Journal of Applied Informatics and Computing*, 9(3), 794–802.
- Wibowo, H. A., Nityasya, M. N., Akyürek, A. F., Suci, F., Aji, A. F., Prasajo, R. E., & Wijaya, D. T. (2021). IndoCollex: A Testbed for Morphological Transformation of Indonesian Word Colloquialism. *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, 3170–3183.
<https://doi.org/10.18653/V1/2021.FINDINGS-ACL.280>
- Wijaya, A., & Enjeli, M. (2024). Analisis Sentimen Penggunaan Aplikasi Kinemaster Menggunakan Metode Naive Bayes. *Jurnal Ilmiah Computer Science*, 2(2), 89–98. <https://doi.org/10.58602/JICS.V2I2.24>
- Wijaya, I. N. S. W., Seputra, K. A., & Dewi, N. P. N. P. (2025). Fine Tuning Model Indobert Untuk Analisis Sentimen Berita Pariwisata Indonesia. *Jurnal Pendidikan Teknologi Dan Kejuruan*, 22(2), 195–204.
<https://doi.org/10.23887/JPTK-UNDIKSHA.V22I2.104056>
- Yaman, A., Sartono, B., & Soleh, A. M. (2021). Pemodelan topik pada dokumen paten terkait pupuk di Indonesia berbasis Latent Dirichlet Allocation. *Berkala Ilmu Perpustakaan Dan Informasi*, 17(2), 168–180.
<https://doi.org/10.22146/BIP.V17I2.2147>

- Yin, B., & Yuan, C. H. (2022). Detecting latent topics and trends in blended learning using LDA topic modeling. *Education and Information Technologies*, 27(9), 12689–12712. <https://doi.org/10.1007/S10639-022-11118-0/FIGURES/1>
- Yudhistira, A., & Andika, R. (2023). Pengelompokan Data Nilai Siswa Menggunakan Metode K-Means Clustering. *Journal of Artificial Intelligence and Technology Information (JAITI)*, 1(1), 20–28. <https://doi.org/10.58602/JAITI.V1I1.22>

