

DAFTAR PUSTAKA

- Abdelmotaleb, H., Wojtyś, M., & McNeile, C. (2023). GSDMM Model Evaluation Techniques with Application to British Telecom Data. *Proceedings of the 5th International Conference on Statistics: Theory and Applications*. <https://doi.org/10.11159/icsta23.116>
- Anuratha, K., Joshi, S., Sharmila, P., Nandhini, J. M. N., & Paravthy, M. (2021). Topical Sentiment Classification to Unmask the Concerns of General Public during COVID-19 Pandemic using Indian Tweets. *Proceedings of the 2021 4th International Conference on Computing and Communications Technologies, ICCCT 2021*, 508–514. <https://doi.org/10.1109/ICCCT53315.2021.9711863>
- Apriantoni, Thooriqoh, H. A., Faticah, C., & Purwitasari, D. (2021). Topic Detection in Sentiment Analysis of Twitter Texts for Understanding the COVID-19 Effect in Local Economic Activities. *Proceedings of 2021 13th International Conference on Information and Communication Technology and System, ICTS 2021*, 354–359. <https://doi.org/10.1109/ICTS52701.2021.9608519>
- Asmussen, C. B., & Møller, C. (2019). Smart Literature Review: A Practical Topic Modelling Approach to Exploratory Literature Review. *Journal of Big Data*, 6(1), 1–18. <https://doi.org/10.1186/s40537-019-0255-7>
- Aysha, A. (2021, June 17). *Short-Text Topic Modelling: LDA vs GSDMM*. <https://towardsdatascience.com/short-texttopic-modelling-lda-vs-gsdmm-20f1db742e14>, diakses pada 25 April 2024 pukul 17.00.
- Blei, D. M. (2012). Probabilistic Topic Models. *Communications of the ACM*, 55(4), 77–84. <https://doi.org/10.1145/2133806.2133826>
- Chawla, N. V, Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. In *Journal of Artificial Intelligence Research* (Vol. 16).
- Czakon, J. (2023). *24 Evaluation Metrics for Binary Classification (And When to Use Them)*. <https://neptune.ai/blog/evaluation-metrics-binary-classification>, diakses pada 26 Desember 2023 pukul 20.15.

- Data Miner. (2021). *Data Miner is the Most Powerful Web Scraping Tool for Professional Data Miners*. <https://dataminer.io/>, diakses pada 23 Oktober 2023 pukul 20.22.
- Devndra, G. (2020). *Comparative Study On Python Web Frameworks: Flask and Django*. Metropolia University of Applied Sciences.
- Eckerson, W. W. (2011). *Performance Dashboards: Measuring, Monitoring, and Managing Your Business* (2nd ed.). Wiley.
- Facebook. (2024). *FastText*. Facebook Open Source. <https://fasttext.cc/>, diakses pada 28 Februari 2024 pukul 16.10.
- Feldman, R., & Sanger, J. (2007). *The Text Mining Handbook: Advanced Approaches In Analyzing Unstructured Data*. Cambridge University Press.
- George, S., & Sumathi, B. (2020). Grid Search Tuning of Hyperparameters in Random Forest Classifier for Customer Feedback Sentiment Prediction. *IJACSA) International Journal of Advanced Computer Science and Applications*, 11(9). <http://dx.doi.org/10.14569/IJACSA.2020.0110920>.
- Hakim, N. S. (2021). *Analisis Sentimen Persepsi Pengguna MyIndihome*. Universitas Islam Indonesia.
- Han, J., Kamber, M., & Pei, J. (2011). *Data Mining Concepts and Techniques* (3rd ed.). Elsevier.
- Husna, N., & Asri, Z. (2022). Perkembangan Trans Padang Sebagai Transportasi Umum di Kota Padang (2014-2020). *Jurnal Kronologi*, 4(3), 136–146.
- Indrawati, A., Subagyo, H., Sihombing, A., Wagiyah, & Afandi, S. (2020). Analyzing The Impact Of Resampling Method For Imbalanced Data Text In Indonesian Scientific Articles Categorization. *Jurnal Dokumentasi Dan Informasi*, 42(2), 133–141. <https://doi.org/10.14203/j.baca.v41i2.563>
- Işik, M., & Dağ, H. (2020). The Impact of Text Preprocessing on The Prediction of Review Ratings. *Turkish Journal of Electrical Engineering and Computer Sciences*, 28(3), 1405–1421. <https://doi.org/10.3906/elk-1907-46>
- Jefrimon. (2023, December 12). *Resmi! Bus Trans Padang Kini Miliki 6 Jalur, Terbaru Koridor 2 dan 3*. <https://www.harianhaluan.com/news/1011171567/resmi-bus-trans-padang-kini-miliki-6-jalur-terbaru-koridor-2-dan-3>, diakses pada 14 Mei 2024 pukul 19.33.

- Kalra, V., & Rashmi, D. A. (2018). Importance of Text Data Preprocessing & Implementation in RapidMiner. *Proceedings of the First International Conference on Information Technology and Knowledge Management*, 71–75.
- Kherwa, P., & Bansal, P. (2020). Topic Modeling: A Comprehensive Review. *EAI Endorsed Transactions on Scalable Information Systems*, 7(24), 1–16. <https://doi.org/10.4108/eai.13-7-2018.159623>
- Kirk, A. (2019). *Data Visualisation: A Handbook for Data Driven Design*. SAGE Publications.
- Kulkarni, A., & Shivananda, A. (2019). Natural Language Processing Recipes: Unlocking Text Data With Machine Learning and Deep Learning Using Python. In *Natural Language Processing Recipes: Unlocking Text Data with Machine Learning and Deep Learning using Python*. Apress Media LLC. <https://doi.org/10.1007/978-1-4842-4267-4>
- Lee, X. J., Yap, T. T. V., Ng, H., & Goh, V. T. (2022). Comparison of Word Embeddings for Sentiment Classification with Preconceived Subjectivity. In *Proceedings of the International Conference on Computer, Information Technology and Intelligent Computing (CITIC 2022)* (pp. 488–502). Atlantis Press International BV. https://doi.org/10.2991/978-94-6463-094-7_39
- Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers.
- Mahing, N. F., Gunawan, A. L., Zen, A. F. A., Bachtiar, F. A., & Wicaksono, S. A. (2023). Klasifikasi Tingkat Stres Dari Data Berbentuk Teks Dengan Menggunakan Algoritma Support Vector Machine (SVM) dan Random Forest. *Jurnal Teknologi Informasi Dan Ilmu Komputer (JTIK)*, 10(7). <https://doi.org/10.25126/jtiik2023108010>
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. *International Conference on Learning Representations*, 1–12.
- Parmar, A., Katariya, R., & Patel, V. (2019). A Review on Random Forest: An Ensemble Classifier. In *Lecture Notes on Data Engineering and Communications Technologies* (Vol. 26, pp. 758–763). Springer Science and

- Business Media Deutschland GmbH. https://doi.org/10.1007/978-3-030-03146-6_86
- Pintas, J. T., Fernandes, L. A. F., & Garcia, A. C. B. (2021). Feature Selection Methods for Text Classification: A Systematic Literature Review. *Artificial Intelligence Review*, 54(8), 6149–6200. <https://doi.org/10.1007/s10462-021-09970-6>
- Purbo, O. W. (2019). *Text Mining, Analisis Medsos, Kekuatan Brand dan Intelijen di Internet*. Andi Publisher.
- Putra, R. A. (2023). *Studi Penerapan Standar Pelayanan Minimum Pada Bus Trans Padang Pada Koridor VI (Pusat Kota-Universitas Andalas)*. Universitas Andalas.
- Rahimi, H., Hoover, J. L., Mimno, D., Naacke, H., Constantin, C., & Amann, B. (2023). *Contextualized Topic Coherence Metrics*. <https://doi.org/https://doi.org/10.48550/arXiv.2305.14587>
- Rehurek, R. (2022). *Gensim*. <https://radimrehurek.com/gensim/intro.html>, diakses pada 12 Desember 2023 pukul 21.04.
- Ribeiro, M. B. (2020). *Sentiment Analysis and Topic Modelling on Brand-Related Social Media Data*. Universitas Porto.
- Rizaty, M. A. (2023, August 7). *Pengguna Instagram di Indonesia Capai 109,3 Juta per April 2023*. 2023. <https://dataindonesia.id/digital/detail/pengguna-instagram-di-indonesia-capai-1093-juta-per-april-2023>, diakses pada 7 Agustus 2023 pukul 14.44.
- Robbani, H. A. (2018). *PySastrawi*. <https://github.com/har07/PySastrawi>, diakses pada 10 November 2023 pukul 21.16.
- Ross, K. A. (2018). Cache-Conscious Query Processing. In *Encyclopedia of Database Systems* (2nd ed.). Springer.
- Sakti, A. T. (2024). *Studi Penilaian Penerapan Standar Pelayanan Minimum Pada Bus Trans Padang Koridor V (Pusat Kota-Indarung)*. Universitas Andalas.
- Silvana, M., Akbar, R., & Tifani, R. (2017). Penerapan Dashboard System di Perpustakaan Universitas Andalas Menggunakan Tableau Public. *Seminar Nasional Sains Dan Teknologi Fakultas Teknik Universitas Muhammadiyah Jakarta*.

- Sommerville, I. (2011). *Software Engineering* (9th ed.). Pearson.
- Suraya, & Sholeh, M. (2022). Designing and Implementing a Database for Thesis Data Management by Using the Python Flask Framework. *International Journal of Engineering, Science & Information Technology (IJESTY)*, 2(1), 9–14. <https://doi.org/10.52088/ijesty.v1i1.197>
- Suwarno, & Abdillah, A. (2016). Penerapan Algoritma Bayesian Regularization Backpropagation Untuk Memprediksi Penyakit Diabetes. *Jurnal MIPA*, 39(2), 150–158. <https://doi.org/10.15294/ijmns.v39i2.9293>
- Syed, S., & Spruit, M. (2017). Full-Text or Abstract? Examining Topic Coherence Scores Using Latent Dirichlet Allocation. *Proceedings - 2017 International Conference on Data Science and Advanced Analytics, DSAA 2017, 2018-January*, 165–174. <https://doi.org/10.1109/DSAA.2017.61>
- Tougui, I., Jilbab, A., & Mhamdi, J. El. (2021). Impact of The Choice Of Cross-Validation Techniques On The Results of Machine Learning-Based Diagnostic Applications. *Healthcare Informatics Research*, 27(3), 189–199. <https://doi.org/10.4258/HIR.2021.27.3.189>
- Tresnasari, N. A., Adji, T. B., & Permanasari, A. E. (2020). Social-Child-Case Document Clustering Based On Topic Modeling using Latent Dirichlet Allocation. *Indonesian Journal of Computing and Cybernetics Systems (IJCCS)*, 14(2), 179–188. <https://doi.org/10.22146/ijccs.54507>
- Walker, R. (2021). *GSDMM*. Github. <https://github.com/rwalk/gsdmm>, diakses pada 4 Maret 2024 pukul 14.08.
- Weiss, S. M., Indurkha, N., & Zhang, T. (2010). *Fundamentals of Predictive Text Mining*. Springer.
- Weisser, C., Gerloff, C., Thielmann, A., Python, A., Reuter, A., Kneib, T., & Säfken, B. (2023). Pseudo-Document Simulation for Comparing LDA, GSDMM and GPM Topic Models On Short and Sparse Text Using Twitter Data. *Computational Statistics*, 38(2), 647–674. <https://doi.org/10.1007/s00180-022-01246-z>
- Yin, J., & Wang, J. (2014). A Dirichlet Multinomial Mixture Model-Based Approach for Short Text Clustering. *Proceedings of the ACM SIGKDD*

International Conference on Knowledge Discovery and Data Mining, 233–242. <https://doi.org/10.1145/2623330.2623715>

Zhang, H., & Wang, Z. (2011). A Normal Distribution-Based Over-Sampling Approach to Imbalanced Data Classification. *International Conference on Advanced Data Mining and Applications*, 83–96.

Žižka, J., Dařena, F., & Svoboda, A. (2020). *Text Mining with Machine Learning : Principles and Techniques*.

