

DAFTAR PUSTAKA

- Abro, A. A., Talpur, M. S. H., & Jumani, A. K. (2023). Natural Language Processing Challenges and Issues: A Literature Review. *Gazi University Journal of Science*, 36(4), 1522–1536. <https://doi.org/10.35378/gujs.1032517>
- Aggarwal, Charu C.; Hinneburg, Alexander; Keim, D. A. (2001). On the Surprising Behavior of Distance Metrics in High Dimensional Space. *Database Theory – ICDT 2001*, 420–434. https://doi.org/https://doi.org/10.1007/3-540-44503-X_27
- Allaoui, M., Kherfi, M. L., & Cheriet, A. (2020). Considerably Improving Clustering Algorithms Using UMAP Dimensionality Reduction Technique: A Comparative Study. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*: Vol. 12119 LNCS (Issue July). Springer International Publishing. https://doi.org/10.1007/978-3-030-51935-3_34
- Asadi, L., Amiri, F., & Safinejad, H. (2020). Investigating The Effect of Health Literacy Level on Improving The Quality of Care During Pregnancy in Pregnant Women Covered by Health Centers. *Journal of Education and Health Promotion*, 9(1). https://doi.org/10.4103/jehp.jehp_204_20
- Aung, K. Z., & Myo, N. N. (2017). Sentiment Analysis of Students' Comment Using Lexicon Based Approach. *Proceedings - 16th IEEE/ACIS International Conference on Computer and Information Science, ICIS 2017*, 149–154. <https://doi.org/10.1109/ICIS.2017.7959985>
- Banik, D., Chalil Madathil, S., Lopes, A. J., Luna Fong, S. A., & Mukka, S. K. (2024). An Evaluation of the Maternal Patient Experience through Natural Language Processing Techniques: The Case of Twitter Data in the United States during COVID-19. *Applied Sciences (Switzerland)*, 14(19). <https://doi.org/10.3390/app14198762>
- Blei, D. M. (2012). Probabilistic Topic Models. *Communications of the ACM*, 55(4), 77–84. <https://doi.org/10.1145/2133806.2133826>
- Beyer, K., Goldstein, J., Ramakrishnan, R., & Shaft, U. (1997). *When Is "Nearest Neighbor" Meaningful?*.
- Catelli, R., Pelosi, S., & Esposito, M. (2022). Lexicon-Based vs. Bert-Based Sentiment Analysis: A Comparative Study in Italian. *Electronics (Switzerland)*, 11(3). <https://doi.org/10.3390/electronics11030374>
- Chai, C. P. (2023). Comparison of Text Preprocessing Methods. *Natural Language Engineering*, 29(3), 509–553. <https://doi.org/10.1017/S1351324922000213>
- 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 - Proceedings of the 2009 Conference*, 288–296.
- Dikiyanti, T. D., Rukmi, A. M., & Irawan, M. I. (2021). Sentiment Analysis and Topic Modeling of BPJS Kesehatan Based on Twitter Crawling Data Using Indonesian Sentiment Lexicon and Latent Dirichlet Allocation Algorithm. *Journal of Physics: Conference Series*, 1821(1). <https://doi.org/10.1088/1742-6596/1821/1/012054>

- Dinas Kesehatan Kota Semarang (2015). Profil Kesehatan Kota Semarang. In dinkes.semarang.go.id.
http://www.depkes.go.id/resources/download/profil/PROFIL_KAB_KOTA_2015/3374_Jateng_Kota_Semarang_2015.pdf
- Fatani, I. (2024). Twitter, Instagram, Youtube Speak: Understanding Sentiments on LRT Jabodebek Services via Inset Lexicon, IndoBERT and BERTopic Approaches. *Journal of Electrical Systems*, 20(4s), 1028–1035. <https://doi.org/10.52783/jes.2147>
- Fayyad, Usama; Piatetsy-Shapiro, Gregory; Smyth, P. (1996). From Data Mining to Knowledge Discovery in Databases. In *AI Magazine* (Vol. 17, Issue 3). American Association for Artificial Intelligence. https://doi.org/10.1007/978-3-319-18032-8_50
- Firdaus, R., Asror, I., & Herdiani, A. (2021). Lexicon-Based Sentiment Analysis of Indonesian Language Student Feedback Evaluation. *Indonesia Journal of Computing*, 6(1). <https://doi.org/10.34818/indojc.2021.6.1.408>
- Grootendorst, M. (2022). BERTopic: Neural Topic Modeling with a Class-based TF-IDF Procedure. <http://arxiv.org/abs/2203.05794>
- Han, Jiawei; Kamber, Micheline; Pei, J. (2012). *Data Mining: Concepts and Techniques* (Third).
- Han, Y., & Moghaddam, M. (2021). Analysis of Sentiment Expressions for User-centered Design. *Expert Systems with Applications*, 171. <https://doi.org/10.1016/j.eswa.2021.114604>
- Handayani, S., & Mubarakah, K. (2015). Kondisi Demografi Ibu dan Suami pada Kasus Kematian Ibu. *Who* 2015, 3(5), 99–108. <http://journal.unnes.ac.id/sju/index.php/higeia>
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. In *Design Science in IS Research MIS Quarterly* (Vol. 28, Issue 1).
- Hou, K., & Hou, T. (2022). Investigating Pregnant Women's Health Information Needs During Pregnancy on Internet Platforms. *Frontiers in Physiology*, 13(November), 1–13. <https://doi.org/10.3389/fphys.2022.1038048>
- Kemendes RI. (2023). *Profil Kesehatan Indonesia 2022*. Kementerian Kesehatan Republik Indonesia, Jakarta.
- Khadidos, A. O., Saleem, F., Selvarajan, S., Ullah, Z., & Khadidos, A. O. (2024). Ensemble Machine Learning Framework for Predicting Maternal Health Risk During Pregnancy. *Scientific Reports*, 14(1). <https://doi.org/10.1038/s41598-024-71934-x>
- Krishnamurti, T., Allen, K., Hayani, L., Rodriguez, S., & Davis, A. L. (2022). Identification of Maternal Depression Risk From Natural Language Collected in a Mobile Health App. *Procedia Computer Science*, 206, 132–140. <https://doi.org/10.1016/j.procs.2022.09.092>
- Margaretha, D. C., & Wati, L. (2024). Harga Diri dan Problematic Social Media Use Pengguna Tiktok. *RISOMA : Jurnal Riset Sosial Humaniora Dan Pendidikan*, 2(5), 170–177. <https://doi.org/10.62383/risoma.v2i5.326>
- Marini, I., & Koniasari. (2023). Factors That Influenced A Generation Z In Search Of Health Information On The Internet under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC4.0).

- Jurnal Eduhealth, 14(02), 2023.
<http://ejournal.seaninstitute.or.id/index.php/healt>
- Maryani, H., Kristiana, L., & Notobroto, H. B. (2012). Indeks Kesehatan Ibu Tingkat Provinsi di Indonesia Berdasarkan Riskesdas 2010. 2010.
- McCarthy, J., & Maine, D. (1992). A Framework for Analyzing the Determinants of Maternal Mortality. *Studies in Family Planning*, 23(1), 23. <https://doi.org/10.2307/1966825>
- McInnes, L., & Healy, J. (2017). Accelerated Hierarchical Density Based Clustering. *IEEE International Conference on Data Mining Workshops, ICDMW, 2017-Novem*, 33–42. <https://doi.org/10.1109/ICDMW.2017.12>
- McInnes, L., Healy, J., Saul, N., & Großberger, L. (2018). UMAP: Uniform Manifold Approximation and Projection. *Journal of Open Source Software*, 3(29), 861. <https://doi.org/10.21105/joss.00861>
- Mehanna, Y. S., & Mahmuddin, M. bin. (2021). A Semantic Conceptualization Using Tagged Bag-of-Concepts for Sentiment Analysis. *IEEE Access*, 9, 118736–118756. <https://doi.org/10.1109/ACCESS.2021.3107237>
- Meldgaard, M., Gamborg, M., & Terkildsen Maindal, H. (2022). Health Literacy Levels Among Women in The Prenatal Period: A Systematic Review. In *Sexual and Reproductive Healthcare (Vol. 34)*. Elsevier B.V. <https://doi.org/10.1016/j.srhc.2022.100796>
- Mukhtar, N., Khan, M. A., & Chiragh, N. (2018). Lexicon-based Approach Outperforms Supervised Machine Learning Approach for Urdu Sentiment Analysis in Multiple Domains. *Telematics and Informatics*, 35(8), 2173–2183.
- Murthy, D., Keshari, S., Arora, S., Yang, Q., Loukas, A., Schwartz, S. J., Harrell, M. B., Hébert, E. T., & Wilkinson, A. v. (2024). Categorizing E-cigarette-related Tweets Using BERT Topic Modeling. *Emerging Trends in Drugs, Addictions, and Health*, 4. <https://doi.org/10.1016/j.etdah.2024.100160>
- Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural Language Processing: An Introduction. In *Journal of the American Medical Informatics Association (Vol. 18, Issue 5, pp. 544–551)*. <https://doi.org/10.1136/amiajnl-2011-000464>
- Nawabi, F., Krebs, F., Vennedey, V., Shukri, A., Lorenz, L., & Stock, S. (2021). Health Literacy in Pregnant Women: A Systematic Review. In *International Journal of Environmental Research and Public Health (Vol. 18, Issue 7)*. MDPI AG. <https://doi.org/10.3390/ijerph18073847>
- Newman, D., Lau, J. H., Grieser, K., & Baldwin, T. (2010). Automatic Evaluation of Topic Coherence. *NAACL HLT 2010 - Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Proceedings of the Main Conference, June*, 100–108.
- Pandove, D., Goel, S., & Rani, R. (2018). Systematic Review of Clustering High-Dimensional and Large Datasets. *ACM Transactions on Knowledge Discovery from Data*, 12(2), 1–68. <https://doi.org/10.1145/3132088>
- Petersen, C. L., Halter, R., Kotz, D., Loeb, L., Cook, S., Pidgeon, D., Christensen, B. C., & Batsis, J. A. (2020). Using Natural Language Processing and Sentiment Analysis to Augment Traditional User-centered Design:

- Development and Usability Study. *JMIR MHealth and UHealth*, 8(8). <https://doi.org/10.2196/16862>
- Plotnikova, V., Dumas, M., & Milani, F. (2020). Adaptations of Data Mining Methodologies: A Systematic Literature Review. *PeerJ Computer Science*, 6, 1–43. <https://doi.org/10.7717/PEERJ-CS.267>
- Qi, Y., & Shabrina, Z. (2023). Sentiment Analysis Using Twitter Data: A Comparative Application of Lexicon-Based and Machine-Learning-Based Approach. *Social Network Analysis and Mining*, 13(1). <https://doi.org/10.1007/s13278-023-01030-x>
- Qiao, F., & Williams, J. (2021). Topic Modelling and Sentiment Analysis of Global Warming Tweets. *Journal of Organizational and End User Computing*, 34(3), 1–18. <https://doi.org/10.4018/joeuc.294901>
- Qin, Z., & Ronchieri, E. (2022). Exploring Pandemics Events on Twitter by Using Sentiment Analysis and Topic Modelling. *Applied Sciences (Switzerland)*, 12(23). <https://doi.org/10.3390/app122311924>
- Roder, Michael; Both, Andreas; Hinneburg, A. (2015). Exploring The Space of Topic Coherence Measures. *WSDM '15: Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, 399–408.
- Roselina, E., Asmiyanto, T., & Andriany, M. (2021). Health Information-Seeking Behavior on the COVID-19 Pandemic: Social Media Usage by Gen Z in Jakarta, Indonesia. *Library Philosophy and Practice*, 1–7.
- Sadia, A., Khan, F., & Bashir, F. (2018). An Overview of Lexicon-Based Approach for Sentiment Analysis. 2018 3rd International Electrical Engineering Conference (IEEC 2018).
- Sajjadian, F., Amiri-Farahani, L., Haghani, S., & Pezaro, S. (2024). Investigating The Relationship Between Health Literacy and Preconceptual Care Components During The First 14 Weeks of Pregnancy: a cross-sectional study. *BMC Primary Care*, 25(1). <https://doi.org/10.1186/s12875-024-02467-5>
- Shafique, U., & Qaiser, H. (2014). A Comparative Study of Data Mining Process Models (KDD, CRISP-DM and SEMMA). *International Journal of Innovation and Scientific Research*, November, 217–222. <https://ijisr.issr-journals.org/abstract.php?article=IJISR-14-281-04>
- Steinbach, M., & Kumar, V. (2004). The Challenges of Clustering High Dimensional Data (Issue January 2003). <https://doi.org/10.1007/978-3-662-08968-2>
- Syairaji, M., Nurdianti, D. S., Wiratama, B. S., Prüst, Z. D., Bloemenkamp, K. W. M., & Verschueren, K. J. C. (2024). Trends and Causes of Maternal Mortality in Indonesia: a systematic review. *BMC Pregnancy and Childbirth*, 24(1). <https://doi.org/10.1186/s12884-024-06687-6>
- Thakur, N., Reimers, N., Daxenberger, J., & Gurevych, I. (2021). Augmented SBERT: Data Augmentation Method for Improving Bi-Encoders for Pairwise Sentence Scoring Tasks. *NAACL-HLT 2021 - 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Conference*, 2, 296–310. <https://doi.org/10.18653/v1/2021.naacl-main.28>

WHO. (2021). Trends in Maternal Mortality 2000 to 2023. In WHO, Geneva.
<https://www.who.int/reproductivehealth/publications/maternal-mortality-2000-2017/en/>