DAFTAR PUSTAKA

Acheadeth, L., Qomariyah, N. N., & Xirinda, M. M. (2022). Utilizing topic modelling in customer product review for classifying baby product. *2022 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom)*, 52–57.

Alqaryouti, O., Siyam, N., Monem, A. A., & Shaalan, K. (2020). Aspect-based sentiment analysis using smart government review data. *Applied Computing and Informatics*, 20(1/2), 142–161.

Asas, V., Juan, S. S., Kerbun, V. C., Chua, S., Labadin, J., & Lau, E. (2025). A comparative analysis of topic modelling techniques for Malaysian business news data: LDA, NMF, Top2Vec, and BERTopic. *2025 14th International Conference on Information Technology in Asia (CITA)*, 42–47.

Asgari-Chenaghlu, M., Feizi-Derakhshi, M. R., Farzinvash, L., Balafar, M. A., & Motamed, C. (2021). Topic detection and tracking techniques on Twitter: A systematic review. *Complexity*, 1–19.

Azis, I. M. A., & Wahyudi, T. (2024). Analisa sentimen rencana pemindahan Ibu Kota Nusantara dari Jakarta ke Kalimantan Timur menggunakan algoritma Naïve Bayes. *Jurnal Indonesia: Manajemen Informatika dan Komunikasi*, 5(3), 2668–2679.

Barik, A. M., Mahendra, R., & Adriani, M. (2019). Normalization of Indonesian-English code-mixed Twitter data. *Proceedings of the 5th Workshop on Noisy User-generated Text (W-NUT 2019)*, 417–424.

Blair, S. J., Bi, Y., & Mulvenna, M. D. (2020). Aggregated topic models for increasing social media topic coherence. *Applied Intelligence*, 50(1), 138–156.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993–1022.

Chen, W., Rabhi, F., Liao, W., & Al-Qudah, I. (2023). Leveraging state-of-the-art topic modeling for news impact analysis on financial markets: A comparative study. *Electronics*, 12(12), 2605.

De Araújo, P. A., Haddadi, S. J., Silva, F. D. S., Reis, M. D. S., & Dos Reis, J. C. (2024). Topic modeling influence in sentiment analysis from user-generated product reviews. *2024 32nd International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)*, 110–115.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 1, 4171–4186.

Egger, R., & Yu, J. (2022a). A topic modeling comparison between LDA, NMF, Top2Vec, and BERTopic to demystify Twitter posts. *Frontiers in Sociology*, 7, 886498.

Egger, R., & Yu, J. (2022b). Identifying hidden semantic structures in Instagram data: A topic modelling comparison. *Tourism Review*, 77(4), 1234–1246.

Feng, Y., Chen, Z., Zhang, Y., Huang, W., Zhang, X., & He, S. (2025). BERTopic_Teen: A multi-module optimization approach for short text topic modeling in adolescent health. *Frontiers in Public Health*, 13, 1608241.

Gommerman, R. B., & Motohashi, E. (2024). A comparative analysis of user-generated content across gaming platforms using deep learning-based topic modeling. *2024 IEEE International Conference on Data Mining Workshops (ICDMW)*, 580–587.

Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint arXiv:2203.05794*.

Hosseiny Marani, A., & Baumer, E. P. (2023). A review of stability in topic modeling: Metrics for assessing and techniques for improving stability. *ACM Computing Surveys*, 56(5), 1–32.

Hu, M., & Liu, B. (2004). Mining and summarizing customer reviews. *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 168–177.

Jefri, M., Fauzi, R., & Fa'rifah, R. Y. (2025). Sentiment-aware feature recommendations for maternal mental-health apps via IndoBERT and

BERTopic on Indonesian TikTok data. *2025 4th International Conference on Electronics Representation and Algorithm (ICERA)*, 575–580.

Kardos, M., Enevoldsen, K. C., & Nielbo, K. L. (2025). topicwizard—A modern, model-agnostic framework for topic model visualization and interpretation. *arXiv preprint arXiv:2505.13034*.

Kharde, V., & Sonawane, P. (2016). Sentiment analysis of twitter data: A survey of techniques. *arXiv preprint arXiv:1601.06971*.

Koto, F., Lau, J. H., & Baldwin, T. (2021). IndoBERTweet: A pretrained language model for Indonesian Twitter with effective domain-specific vocabulary initialization. *arXiv preprint arXiv:2109.04607*.

Krishnan, A. (2023). Exploring the power of topic modeling techniques in analyzing customer reviews: A comparative analysis. *arXiv preprint arXiv:2308.11520*.

Lande, J., Pillay, A., & Chandra, R. (2023). Deep learning for COVID-19 topic modelling via Twitter: Alpha, Delta and Omicron. *PLoS ONE*, 18(8), e0288681.

Laureate, C. D. P., Buntine, W., & Linger, H. (2023). A systematic review of the use of topic models for short text social media analysis. *Artificial Intelligence Review*, 56(12), 14223–14255.

Liu, H., Chatterjee, I., Zhou, M., Lu, X. S., & Abusorrah, A. (2020). Aspect-based sentiment analysis: A survey of deep learning methods. *IEEE Transactions on Computational Social Systems*, 7(6), 1358–1375.

Ma, L., Chen, R., Ge, W., Rogers, P., Lyn-Cook, B., Hong, H., ... & Zou, W. (2025). AI-powered topic modeling: Comparing LDA and BERTopic in analyzing opioid-related cardiovascular risks in women. *Experimental Biology and Medicine*, 250, 10389.

Mahfudiyah, N., & Alamsyah, A. (2023). Understanding user perception of ride-hailing services sentiment analysis and topic modelling using IndoBERT and BERTopic. *2023 International Conference on Digital Business and Technology Management (ICONDBTM)*, 1–6.

Mazarura, J., & De Waal, A. (2016). A comparison of the performance of latent Dirichlet allocation and the Dirichlet multinomial mixture model on short

text. *2016 Pattern Recognition Association of South Africa and Robotics and Mechatronics International Conference (PRASA-RobMech)*, 1–6.

Mihajlov, T., Nešić, M. I., Stanković, R., & Kitanović, O. (2024). Topic modeling of the SrpELTeC corpus: A comparison of NMF, LDA, and BERTopic. *2024 19th Conference on Computer Science and Intelligence Systems (FedCSIS)*, 649–653.

Minaee, S., Kalchbrenner, N., Cambria, E., Nikzad, N., Chenaghlu, M., & Gao, J. (2021). Deep learning–based text classification: A comprehensive review. *ACM Computing Surveys (CSUR)*, 54(3), 1–40.

Murshed, B. A. H., Abawajy, J., Mallappa, S., Saif, M. A. N., Al-Ghuribi, S. M., & Ghanem, F. A. (2022). Enhancing big social media data quality for use in short-text topic modeling. *IEEE Access*, 10, 105328–105351.

Mutsaddi, A., Jamkhande, A., Thakre, A. S., & Haribhakta, Y. (2025). BERTopic for topic modeling of Hindi short texts: A comparative study. *Proceedings of the First Workshop on Natural Language Processing for Indo-Aryan and Dravidian Languages*, 22–32.

Naim, F. A. (2021). Bangla aspect-based sentiment analysis based on corresponding term extraction. *2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)*, 65–69.

Nanayakkara, A. C., & Thennakoon, G. A. D. M. (2024). Enhancing social media content analysis with advanced topic modeling techniques: A comparative study. *International Journal on Advances in ICT for Emerging Regions (ICTer)*, 17(1).

Nigam, K., McCallum, A. K., Thrun, S., & Mitchell, T. (2000). Text classification from labeled and unlabeled documents using EM. *Machine Learning*, 39(2), 103–134.

Nurdzati, F., Tricahyono, D., & Alamsyah, A. (2025). Enhancing restaurant success insight through NLP-based topic modeling and emotion analysis of user reviews. *2025 International Conference on Data Science and Its Applications (ICoDSA)*, 843–848.

Perwira, R. I., Permadi, V. A., Purnamasari, D. I., & Agusdin, R. P. (2025). Domain-specific fine-tuning of IndoBERT for aspect-based sentiment analysis in Indonesian travel user-generated content. *Journal of Information Systems Engineering and Business Intelligence*, 11(1), 30–40.

Pita, M., Nunes, M., & Pappa, G. L. (2022). Probabilistic topic modeling for short text based on word embedding networks. *Applied Intelligence*, 52(15), 17829–17844.

Rachman, F. F., & Pramana, S. (2020). Analisis sentimen pro dan kontra masyarakat Indonesia tentang vaksin COVID-19 pada media sosial Twitter. *Indonesian of Health Information Management Journal (INOHIM)*, 8(2), 100–109.

Rahimi, H., Mimno, D., Vigly, J. H., Naacke, H., Constantin, C., & Amann, B. (2024). Contextualized topic coherence metrics. *Findings of the Association for Computational Linguistics: EACL 2024*, 1760–1773.

Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the space of topic coherence measures. *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, 399–408.

Rosid, M. A., Siahaan, D. O., & Saikhu, A. (2024). Sarcasm detection in Indonesian-English code-mixed text using multihead attention-based convolutional and bi-directional GRU. *IEEE Access*, 12, 137063–137079.

Savanur, S. R., & Sumathi, R. (2022). ABSADM-Aspect-Based Sentiment Analysis using Distance Matrix. *2022 International Conference on Futuristic Technologies (INCOFT)*, 1–7.

Siino, M., Tinnirello, I., & La Cascia, M. (2024). Is text preprocessing still worth the time? A comparative survey on the influence of popular preprocessing methods on Transformers and traditional classifiers. *Information Systems*, 121, 102342.

Sjoraida, D. F., Guna, B. W. K., & Yudhakusuma, D. (2024). Analisis sentimen film Dirty Vote menggunakan BERT (Bidirectional Encoder Representations from Transformers). *Jurnal JTIK (Jurnal Teknologi Informasi dan Komunikasi)*, 8(2), 393–404.

Suhaeni, C., Mualifah, L. N. A., & Wijayanto, H. (2025). LDA Topic Modeling

Analysis of Public Discourse on Indonesia's Free Nutritious Meals Program (MBG). *IJID (International Journal on Informatics for Development)*, 14(1), 587–600.

Sukma, E. A., Hidayanto, A. N., Pandesenda, A. I., Yahya, A. N., Widharto, P., & Rahardja, U. (2020). Sentiment analysis of the new Indonesian government policy (Omnibus Law) on social media Twitter. *2020 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*, 153–158.

Susanto, G. M., & Mauritsius, T. (2025). Product backlog recommendations using sentiment analysis with topic modelling and text summarization in the XYZ E-wallet application. *2025 29th International Conference on Information Technology (IT)*, 1–6.

Tamzila, A. S., Sulistya, A., & Lidiawaty, B. R. (2025). A comparative evaluation of LDA and BERTopic for topic modeling of traffic complaints in short social media texts. *2025 International Electronics Symposium (IES)*, 567–572.

Udupa, A., Adarsh, K. N., Aravinda, A., Godihal, N. H., & Kayarvizhy, N. (2022). An exploratory analysis of GSDMM and BERTopic on short text topic modelling. *2022 Fourth International Conference on Cognitive Computing and Information Processing (CCIP)*, 1–9.

Uğuz, S., & Tülü, Ç. N. (2024). Topic modeling analysis in the field of Large Language Models with BERTopic (2020–2024). *2024 Innovations in Intelligent Systems and Applications Conference (ASYU)*, 1–6.

Wang, B., Liu, S., Ding, K., Liu, Z., & Xu, J. (2014). Identifying technological topics and institution-topic distribution probability for patent competitive intelligence analysis: A case study in LTE technology. *Scientometrics*, 101(1), 685–704.

Wau, K. (2025). Application of fine-tuned IndoBERT for sentiment classification local product reviews on Tokopedia marketplace with limited dataset. *Journal of Artificial Intelligence and Engineering Applications (JAIEA)*, 5(1), 1377–1381.

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.

Xu, W. W., Tshimula, J. M., Dubé, È., Graham, J. E., Greyson, D., MacDonald, N. E., & Meyer, S. B. (2022). Unmasking the Twitter discourses on masks during the COVID-19 pandemic: User cluster–based BERT topic modeling approach. *Jmir Infodemiology*, 2(2), e41198.

Yamunathangam, D., Priya, C. B., Shobana, G., & Latha, L. (2021). An overview of topic representation and topic modelling methods for short texts and long corpus. *2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)*, 1–6.

Yan, H., & Zhang, Y. (2023). An optimized BERTopic framework based on cluster silhouette for improving topic coherence. *Proceedings of the 4th International Conference on Artificial Intelligence and Computer Engineering*, 559–566.

Yin, J., & Wang, J. (2014). A Dirichlet multinomial mixture model-based approach for short text clustering. *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 233–242.

Zankadi, H., Idrissi, A., Daoudi, N., & Hilal, I. (2023). Identifying learners' topical interests from social media content to enrich their course preferences in MOOCs using topic modeling and NLP techniques. *Education and Information Technologies*, 28(5), 5567–5584.

Zhu, Y., & Liu, Y. (2025). Gibbs-BERTopic: A hybrid approach for short text topic modeling. *IEEE Access*.