

DAFTAR PUSTAKA

- [1] M. K. Wulan, “Jumlah Wajib Pajak Meningkatkan Signifikan, Pemerintah Bakal Luncurkan Coretax System,” Kompas.id. Accessed: Feb. 17, 2025. [Online]. Available: <https://www.kompas.id/baca/ekonomi/2024/08/01/jumlah-wajib-pajak-meningkat-signifikan-pemerintah-bakal-luncurkan-core-tax-system>
- [2] L. O. Anto, Husin, W. Hamid, and N. L. Bulan, “Taxpayer awareness, tac knowledge, tac sanctions, public service account ability and taxpayer compliance,” *Accounting*, vol. 7, no. 1, pp. 49–58, Jan. 2021, doi: 10.5267/j.ac.2020.10.015.
- [3] N. Wilantika and S. B. Wibisono, “Evaluating the quality of public services through social media,” *Asian Journal for Public Opinion Research*, vol. 9, no. 3, pp. 240–265, 2021, doi: 10.15206/ajpor.2021.9.3.240.
- [4] T. Taufiqurokhman, E. Satispi, A. Andriansyah, M. Murod, and E. Sulastrri, “The impact of e-service quality on public trust and public satisfaction in e-government public services,” *International Journal of Data and Network Science*, vol. 8, no. 2, pp. 765–772, Mar. 2024, doi: 10.5267/j.ijdns.2024.1.002.
- [5] M. A. Khadija and W. Nurharjadmo, “Enhancing Indonesian customer complaint analysis: LDA topic modelling with BERT embeddings,” *Sinergi (Indonesia)*, vol. 28, no. 1, pp. 153–162, 2024, doi: 10.22441/sinergi.2024.1.015.
- [6] N. F. AL-Bakri, J. F. Yonan, and A. T. Sadiq, “Tourism Companies Assessment via Social Media Using Sentiment Analysis,” *Baghdad Science Journal*, vol. 19, no. 2, p. 0422, Apr. 2022, doi: 10.21123/bsj.2022.19.2.0422.
- [7] R. Alfred, J. H. Obit, Y. Lim, H. Havaluddin, and A. Azman, “Social Media Mining: A Genetic Based Multiobjective Clustering Approach to Topic Modelling.”
- [8] A. F. Ibrahim, M. Hassaballah, A. A. Ali, Y. Nam, and I. A. Ibrahim, “COVID19 Outbreak: A Hierarchical Framework for User Sentiment Analysis,” *Computers, Materials & Continua*, vol. 70, no. 2, pp. 2507–2524, 2022, doi: 10.32604/cmc.2022.018131.
- [9] W. Nur Syahirah Wan Min, N. Zareen Zulkarnain, and F. Teknologi Maklumat Dan Komunikasi, “Comparative Evaluation of Lexicons in

Performing Sentiment Analysis,” 2020. [Online]. Available:
<https://www.researchgate.net/publication/342451911>

- [10] D. I. Ulya, A. Kunaefi, D. Rolliawati, and B. A. Nugroho, “Unpacking Public Perceptions of Qris with Twitter Data: A Vader And LDA Methodology,” *Jurnal ELTIKOM*, vol. 7, no. 2, pp. 145–159, Dec. 2023, doi: 10.31961/eltikom.v7i2.742.
- [11] G. B. Bergamini Gomes and R. Attux, “Contributions to Social Media Analysis Based on Topic Modelling,” in *Anais do XI Symposium on Knowledge Discovery, Mining and Learning (KDMiLe 2023)*, Sociedade Brasileira de Computação - SBC, Sep. 2023, pp. 113–120. doi: 10.5753/kdmile.2023.231795.
- [12] A. Baird, Y. Xia, and Y. Cheng, “Consumer perceptions of telehealth for mental health or substance abuse: a Twitter-based topic modeling analysis,” *JAMIA Open*, vol. 5, no. 2, Apr. 2022, doi: 10.1093/jamiaopen/ooac028.
- [13] Imamah and F. H. Rachman, “Twitter Sentiment Analysis of Covid-19 Using Term Weighting TF-IDF And Logistic Regresion,” in *2020 6th Information Technology International Seminar (ITIS)*, IEEE, Oct. 2020, pp. 238–242. doi: 10.1109/ITIS50118.2020.9320958.
- [14] J. Setiawan, V. Gousander, and I. Prasetiawan, “Unmasking the Sentiments of Labuan Bajo: An Instagram-based Analysis for Tourism Insights through VADER Sentiment Analysis,” *G-Tech: Jurnal Teknologi Terapan*, vol. 7, no. 3, pp. 967–976, Jul. 2023, doi: 10.33379/gtech.v7i3.2615.
- [15] Statista, “X/Twitter: Global Audience 2024.”
- [16] S. H. W. Ilyas, Z. T. Soomro, A. Anwar, H. Shahzad, and U. Yaquub, “Analyzing Brexit’s impact using sentiment analysis and topic modeling on Twitter discussion,” in *The 21st Annual International Conference on Digital Government Research*, New York, NY, USA: ACM, Jun. 2020, pp. 1–6. doi: 10.1145/3396956.3396973.
- [17] C. Hutto and E. Gilbert, “VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text,” *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 8, no. 1, pp. 216–225, May 2014, doi: 10.1609/icwsm.v8i1.14550.
- [18] M. J. Lavin, “Analyzing Documents with TF-IDF,” *Programming Historian*, no. 8, May 2019, doi: 10.46430/phen0082.
- [19] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” in

Proceedings of the 2019 Conference of the North, Stroudsburg, PA, USA: Association for Computational Linguistics, 2019, pp. 4171–4186. doi: 10.18653/v1/N19-1423.

- [20] A. C. Nanayakkara and G. A. D. M. Thennakoon, “Enhancing Social Media Content Analysis with Advanced Topic Modeling Techniques: A Comparative Study,” *International Journal on Advances in ICT for Emerging Regions (ICTer)*, vol. 17, no. 1, pp. 40–47, May 2024, doi: 10.4038/icter.v17i1.7276.
- [21] V. Arya, A. K. Mishra, and A. González-Briones, “Sentiments Analysis of Covid-19 Vaccine Tweets Using Machine Learning and Vader Lexicon Method,” *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal*, vol. 11, no. 4, pp. 507–518, Jun. 2023, doi: 10.14201/adcaij.27349.
- [22] F. Fajri, B. Tutuko, and S. Sukemi, “Membandingkan Nilai Akurasi BERT dan DistilBERT pada Dataset Twitter,” *JUSIFO (Jurnal Sistem Informasi)*, vol. 8, no. 2, pp. 71–80, Dec. 2022, doi: 10.19109/jusifo.v8i2.13885.
- [23] 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.
- [24] M. Chiny, M. Chihab, O. Bencharef, and Y. Chihab, “LSTM, VADER and TF-IDF based Hybrid Sentiment Analysis Model,” *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 7, 2021, doi: 10.14569/IJACSA.2021.0120730.
- [25] Z. Jin, X. Lai, and J. Cao, “Multi-label Sentiment Analysis Base on BERT with modified TF-IDF,” in *2020 IEEE International Symposium on Product Compliance Engineering-Asia (ISPCE-CN)*, IEEE, Nov. 2020, pp. 1–6. doi: 10.1109/ISPCE-CN51288.2020.9321861.
- [26] L. Z. S. Sudar, J. L. Imbenay, I. Budi, A. Ramadiah, P. K. Putra, and A. B. Santoso, “Textual Analysis for Public Sentiment Toward National Police Using CRISP-DM Framework,” *International Information and Engineering Technology Association*, vol. 38, no. 1, pp. 63–72, Feb. 2024, doi: 10.18280/ria.380107.
- [27] S. H. Utami, A. A. Purnama, and A. N. Hidayanto, “Fintech Lending in Indonesia: A Sentiment Analysis, Topic Modelling, and Social Network Analysis using Twitter Data,” 2022.
- [28] D. Jiang, C. Zhang, and Y. Song, *Probabilistic Topic Models*. Singapore: Springer Nature Singapore, 2023. doi: 10.1007/978-981-99-2431-8.

- [29] R. Alghamdi and K. Alfalqi, "A Survey of Topic Modeling in Text Mining," *International Journal of Advanced Computer Science and Applications*, vol. 6, no. 1, 2015, doi: 10.14569/IJACSA.2015.060121.
- [30] M. Taufiq Anwar, D. Riandhita Arief Permana, P. STMI Jakarta, P. Sistem Informasi Industri Otomotif, J. Letjen Suprpto No, and J. Pusat, "Analisis Sentimen Masyarakat Indonesia Terhadap Produk Kendaraan Listrik Menggunakan VADER," vol. 10, no. 1, pp. 783–792, 2023.
- [31] N. Garg and K. Sharma, "Text pre-processing of multilingual for sentiment analysis based on social network data," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 1, p. 776, Feb. 2022, doi: 10.11591/ijece.v12i1.pp776-784.
- [32] S. Samsir, R. S. Saragih, S. Subagio, R. Aditiya, and R. Watrianthos, "BERTopic Modeling of Natural Language Processing Abstracts: Thematic Structure and Trajectory," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 7, no. 3, p. 1514, Jul. 2023, doi: 10.30865/mib.v7i3.6426.
- [33] A. Fantaye, "Text Mining Technique for Driving Potentially Valuable Information from Text," *Information and Knowledge Management*, Jan. 2020, doi: 10.7176/IKM/10-1-01.
- [34] S. Qaiser and R. Ali, "Text Mining: Use of TF-IDF to Examine the Relevance of Words to Documents," *Int J Comput Appl*, vol. 181, no. 1, pp. 25–29, Jul. 2018, doi: 10.5120/ijca2018917395.
- [35] J. Wang, W. Xu, W. Yan, and C. Li, "Text Similarity Calculation Method Based on Hybrid Model of LDA and TF-IDF," in *Proceedings of the 2019 3rd International Conference on Computer Science and Artificial Intelligence*, New York, NY, USA: ACM, Dec. 2019, pp. 1–8. doi: 10.1145/3374587.3374590.
- [36] A. Sakho, E. Malherbe, and E. Scornet, "Do we need rebalancing strategies? A theoretical and empirical study around SMOTE and its variants," Feb. 2024.
- [37] F. Kamalov, S. E. Choutri, and A. F. Atiya, "Analytical formulation of synthetic minority oversampling technique (SMOTE) for imbalanced learning," *Gulf Journal of Mathematics*, vol. 19, no. 1, pp. 400–415, Jan. 2025, doi: 10.56947/gjom.v19i1.2639.
- [38] J. Ah-Pine and E.-P. Soriano-Morales, "A Study of Synthetic Oversampling for Twitter Imbalanced Sentiment Analysis," HAL Open

sciences. Accessed: Mar. 15, 2025. [Online]. Available: <https://hal.science/hal-01504684v1>

- [39] Z. Su, “Applications of BERT in sentimental analysis,” *Applied and Computational Engineering*, vol. 92, no. 1, pp. 147–152, Oct. 2024, doi: 10.54254/2755-2721/92/20241711.
- [40] Y. Wu, Z. Jin, C. Shi, P. Liang, and T. Zhan, “Research on the Application of Deep Learning-based BERT Model in Sentiment Analysis,” Mar. 2024.
- [41] V. Sanh, L. Debut, J. Chaumond, and T. Wolf, “DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter,” Oct. 2019.
- [42] U. Narayan and D. Kumar, “SENTIMENT ANALYSIS USING TRANSFORMER BASED MODEL,” in *International Journal of Innovative Research in Computer Science and Technology (IJIRCST)*, Innovative Research Publication, Mar. 2024, pp. 264–269. doi: 10.55524/CSISTW.2024.12.1.46.
- [43] S. Arora, A. Malhotra, A. Kumar, and N. Singh, “Sentiment Analysis Unleashed: Leveraging VADER for Evaluation of Public Opinion,” in *2024 11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, IEEE, Mar. 2024, pp. 1–6. doi: 10.1109/ICRITO61523.2024.10522172.
- [44] A. Das, “Logistic Regression,” in *Encyclopedia of Quality of Life and Well-Being Research*, Cham: Springer International Publishing, 2021, pp. 1–2. doi: 10.1007/978-3-319-69909-7_1689-2.
- [45] Z. Wang, S. Ameenuddin Irfan, C. Teoh, and P. Hriday Bhoyar, “Logistic Regression,” in *Numerical Machine Learning*, BENTHAM SCIENCE PUBLISHERS, 2023, pp. 71–96. doi: 10.2174/9789815136982123010005.
- [46] S. Kumar, N. Kaur, Kavita, and A. Joshi, “Tweet sentiment analysis using logistic regression,” *IET Conference Proceedings*, vol. 2023, no. 11, pp. 332–336, Sep. 2023, doi: 10.1049/icp.2023.1801.
- [47] D. Medvecki, B. Bašaragin, A. Ljajić, and N. Milošević, “Multilingual Transformer and BERTopic for Short Text Topic Modeling: The Case of Serbian,” 2024, pp. 161–173. doi: 10.1007/978-3-031-50755-7_16.
- [48] N. Khodeir and F. Elghannam, “Efficient topic identification for urgent MOOC Forum posts using BERTopic and traditional topic modeling techniques,” *Educ Inf Technol (Dordr)*, vol. 30, no. 5, pp. 5501–5527, Apr. 2025, doi: 10.1007/s10639-024-13003-4.

- [49] Jude Chukwura Obi, “A comparative study of several classification metrics and their performances on data,” *World Journal of Advanced Engineering Technology and Sciences*, vol. 8, no. 1, pp. 308–314, Feb. 2023, doi: 10.30574/wjaets.2023.8.1.0054.
- [50] D. A. Kristiyanti, S. A. Sanjaya, V. C. Tjokro, and J. Suhali, “Dealing imbalance dataset problem in sentiment analysis of recession in Indonesia,” *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 13, no. 2, p. 2060, Jun. 2024, doi: 10.11591/ijai.v13.i2.pp2060-2072.
- [51] H. Satria, “Crawl Data Twitter Menggunakan Tweet Harvest .”
- [52] “Hunspell.” Accessed: Mar. 12, 2025. [Online]. Available: <https://hunspell.github.io/>
- [53] S. D. A. Simanjuntak, “Dua Bulan Coretax Berlaku, Penerimaan Pajak Turun 30% per Februari 2025.” Accessed: Jun. 04, 2025. [Online]. Available: <https://ekonomi.bisnis.com/read/20250313/259/1860996/dua-bulan-coretax-berlaku-penerimaan-pajak-turun-30-per-februari-2025>

