

DAFTAR PUSTAKA

- Afaq, S., & Rao, S. (2020). *Significance Of Epochs On Training A Neural Network*. 9(06), 485–488.
- Akbar, K., & Hayaty, M. (2020). Data Balancing untuk Mengatasi Imbalance Dataset pada Prediksi Produksi Padi. *Jurnal Ilmiah Intech : Information Technology Journal of UMUS*, 2(02). <https://doi.org/10.46772/intech.v2i02.283>
- Aksoy, T., Gülseçen, S., & Çelik, S. (2020). Data Pre-processing in Text Mining. *Who Runs the World: Data*, December, 123–144. <https://doi.org/10.26650/b/et06.2020.011.07>
- APJII, A. P. J. I. I. (2024, February 7). *APJII Jumlah Pengguna Internet Indonesia Tembus 221 Juta Orang*. <https://apjii.or.id/berita/d/apjii-jumlah-pengguna-internet-indonesia-tembus-221-juta-orang>
- Aripin, Adi Santoso, S., & Haryanto, H. (2023). *Optimizing the Accuracy of the Semantic-Based Compound Emotion Classifications using the XLM-RoBERTa*. 12(1), 29–36.
- Babazade, Y. (2025). The Bilingual Brain: Cognitive Benefits and Challenges of Multilingualism. *Euro-Global Journal of Linguistics and Language Education*, 2(3), 55–63.
- Bagheri, A., Giachanou, A., Mosteiro, P., & Verberne, S. (2023). Natural language processing and text mining: turning unstructured data into structured. *Clinical Applications of Artificial Intelligence in Real-World Data*, 1–285. <https://doi.org/10.1007/978-3-031-36678-9>
- Burkitt, I. (2021). The Emotions in Cultural-Historical Activity Theory: Personality, Emotion and Motivation in Social Relations and Activity. In *Integrative Psychological and Behavioral Science* (Vol. 55, Issue 4, pp. 797–820). <https://doi.org/10.1007/s12124-021-09615-x>
- Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., Grave, E., Ott, M., Zettlemoyer, L., & Stoyanov, V. (2020). Unsupervised cross-lingual representation learning at scale. *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, 8440–8451.

<https://doi.org/10.18653/v1/2020.acl-main.747>

- Dai, Y. fu, Gao, X. yan, Leng, W. wu, Huang, C., Yu, W. jing, & Jiang, C. hao. (2024). Congruent or conflicting? The interaction between emoji and textual sentence is not that simple! *Heliyon*, *10*(12), e32984. <https://doi.org/10.1016/j.heliyon.2024.e32984>
- Fazry, L., & Apsari, N. C. (2021). PENGARUH MEDIA SOSIAL TERHADAP PERILAKU CYBERBULLYING DI KALANGAN REMAJA. *Jurnal Pengabdian Dan Penelitian Kepada Masyarakat (JPPM)*, *2*(1), 28–36.
- Hassani, H., Beneki, C., Unger, S., Mazinani, M. T., & Yeganegi, M. R. (2020). Text mining in big data analytics. *Big Data and Cognitive Computing*, *4*(1), 1–34. <https://doi.org/10.3390/bdcc4010001>
- Hayati, R., & Anindhita, W. K. (2024). Code Switching and Code Mixing Used by Students in Social Media. *JELLE : Journal Of English Literature, Linguistic, and Education*, *5*(01), 21–27. <https://doi.org/10.31941/jele.v5i01.4138>
- Herdian, C., Kamila, A., & Agung Musa Budidarma, I. G. (2024). Studi Kasus Feature Engineering Untuk Data Teks: Perbandingan Label Encoding dan One-Hot Encoding Pada Metode Linear Regresi. *Technologia : Jurnal Ilmiah*, *15*(1), 93. <https://doi.org/10.31602/tji.v15i1.13457>
- Holstein, J., Spitzer, P., Hoell, M., Vössing, M., & Köhl, N. (2024). UNDERSTANDING DATA UNDERSTANDING: A FRAMEWORK. *Thirty-Second European Conference on Information Systems (ECIS 2024)*.
- Hussein, B. M., & Shareef, S. M. (2024). An Empirical Study on the Correlation between Early Stopping Patience and Epochs in Deep Learning. *ITM Web of Conferences*, *64*, 01003. <https://doi.org/10.1051/itmconf/20246401003>
- Hwang, Y., Jeong, S., & Wu, Y. (2025). Artificial Intelligence in Design Process: An Analysis Using Text Mining. *Applied Artificial Intelligence*, *39*(1), 1–22. <https://doi.org/10.1080/08839514.2025.2453782>
- Joseph, V. R. (2022). Optimal ratio for data splitting. *Statistical Analysis and Data Mining*, *15*(4), 531–538. <https://doi.org/10.1002/sam.11583>
- Junho, L., & Cheongtag, K. (2023). A Structure of basic emotions : A review of basic emotion theories using an emotionally fine-tuned language model.

- Proceedings of the Annual Meeting of the Cognitive Science.*
<https://escholarship.org/uc/item/2zd4f4dk>
- Kannan, S., Suryamritha, M., Balaji, V., & Venugoapalan, M. (2024). Emotion Categorization on Multiclass Textual Data Using ML and Ensemble Techniques. *RAICS - IEEE Recent Advances in Intelligent Computational Systems, 2024*, 1–7. <https://doi.org/10.1109/RAICS61201.2024.10689961>
- Kartini, Harahap, I. A., Arwana, N. Y., Wahyu, & Rambe, S. W. T. B. (2020). Teori dalam Penelitian Media. *Jurnal Edukasi Nonformal*, 3(2), 136–140.
- Koto, F., Rahimi, A., Lau, J. H., & Baldwin, T. (2020). IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP. *COLING 2020 - 28th International Conference on Computational Linguistics, Proceedings of the Conference*, 757–770. <https://doi.org/10.18653/v1/2020.coling-main.66>
- Li, Z., & Zou, Z. (2024). Punctuation and lexicon aid representation: A hybrid model for short text sentiment analysis on social media platform. *Journal of King Saud University - Computer and Information Sciences*, 36(3), 102010. <https://doi.org/10.1016/j.jksuci.2024.102010>
- Macías Loor, M. Á., Moreira Aguayo, P. Y., Venegas Loor, L. V., Loor Cevallos, J. M., Medranda Medranda, J. A., & Barcia Moreira, M. R. (2024). English as a lingua franca: an overview. *Revista Cognosis*, 9(2), 151–163. <https://doi.org/10.33936/cognosis.v9i2.5777>
- Maghfiroh, A., Findawati, Y., & Indahyanti, U. (2023). Klasifikasi Penipuan pada Rekening Bank menggunakan Pendekatan Ensemble Learning. *Building of Informatics, Technology and Science (BITS)*, 4(4), 1883–1891. <https://doi.org/10.47065/bits.v4i4.3212>
- Markoulidakis, I., Rallis, I., Georgoulas, I., Kopsiaftis, G., Doulamis, A., & Doulamis, N. (2021). Multiclass Confusion Matrix Reduction Method and Its Application on Net Promoter Score Classification Problem. *Technologies*, 9(4). <https://doi.org/10.3390/technologies9040081>
- Marthen, E., & Novita, I. (2024). Optimasi RoBERTa dengan Hyperparameter Tuning untuk Deteksi. *Jurnal Nasional Teknologi Dan Sistem Informasi*, 03,

240–248.

- Mondal, A., & Gokhale, S. S. (2020). Mining Emotions on Plutchik ' s Wheel. *Seventh International Conference on Social Networks Analysis, Management and Security (SNAMS)*., 1–6.
- Najwa, S., Winarni, S., Pravitasari, A. A., & Hidayat, Y. (2025). Emotion Classification in Social Media Posts Related To Telecommunication Services Using Bidirectional Lstm. *Communications in Mathematical Biology and Neuroscience, 2025*, 1–15. <https://doi.org/10.28919/cmbn/9006>
- Nordin, N. R. M. (2023). Code-switching and code-mixing among users of social media. *Jurnal Javanologi*, 6(2), 1267. <https://doi.org/10.20961/javanologi.v6i2.75114>
- Prasad, J., Prasad, G., & Chellamuthu, G. (2022). GJG@TamilNLP-ACL2022: Emotion Analysis and Classification in Tamil using Transformers. *DravidianLangTech 2022 - 2nd Workshop on Speech and Language Technologies for Dravidian Languages, Proceedings of the Workshop*, 86–92. <https://doi.org/10.18653/v1/2022.dravidianlangtech-1.14>
- Putra, I. F., & Purwarianti, A. (2020). Improving Indonesian Text Classification Using Multilingual Language Model. *2020 7th International Conference on Advanced Informatics: Concepts, Theory and Applications, ICAICTA 2020, September 2020*. <https://doi.org/10.1109/ICAICTA49861.2020.9429038>
- Ranasinghe, T., & Zampieri, M. (2020). Multilingual offensive language identification with cross-lingual embeddings. *EMNLP 2020 - 2020 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference, November 2020*, 5838–5844. <https://doi.org/10.18653/v1/2020.emnlp-main.470>
- Rifaldi, D., Abdul Fadlil, & Herman. (2023). Teknik Preprocessing Pada Text Mining Menggunakan Data Tweet “Mental Health.” *Decode: Jurnal Pendidikan Teknologi Informasi*, 3(2), 161–171. <https://doi.org/10.51454/decode.v3i2.131>
- Rohayati, A. S. (2023). Penggunaan Bahasa Indonesia Di Media Sosial. *Mahasiswa Kreatif, 1(2)*, 29–33.

- Saleh Alzubaidi, M., & Bourennani, F. (2021). Sentiment Analysis of Tweets Using Emojis and Texts. *ACM International Conference Proceeding Series*, 111–114. <https://doi.org/10.1145/3459955.3460608>
- Sarasati, B., & Nurvia, O. (2021). Emosi Dalam Tulisan. *Jurnal Psibernetika*, 14(1), 40–48. <https://doi.org/10.30813/psibernetika>.
- Setiawan, E. I., Kristianto, L., Hermawan, A. T., Santoso, J., Fujisawa, K., & Purnomo, M. H. (2021). Social Media Emotion Analysis in Indonesian Using Fine-Tuning BERT Model. *3rd 2021 East Indonesia Conference on Computer and Information Technology, EIconCIT 2021*, 334–337. <https://doi.org/10.1109/EIconCIT50028.2021.9431885>
- Shalahuddin, M. A., & Fajrianti, L. (2024). Dampak Emoticon dalam Komunikasi di Media Sosial. *Propaganda*, 4(2), 72–78. <https://doi.org/10.37010/prop.v4i2.1632>
- Shankar, S., & Tewari, V. (2021). Understanding the emotional intelligence discourse on social media: Insights from the analysis of twitter. *Journal of Intelligence*, 9(4). <https://doi.org/10.3390/jintelligence9040056>
- Shen, R., Gao, L., & Ma, Y.-A. (2022). *On Optimal Early Stopping: Over-informative versus Under-informative Parametrization*. 1–30. <http://arxiv.org/abs/2202.09885>
- Shoeb, A., & de Melo, G. (2020). *Are Emojis Emotional? A Study to Understand the Association between Emojis and Emotions*. May. <https://doi.org/10.48550/arXiv.2005.00693>
- 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(March 2023). <https://doi.org/10.1016/j.is.2023.102342>
- Simon. (2024). *Digital 2024: Global Overview Report*. <https://datareportal.com/reports/digital-2024-global-overview-report>
- Sudianto, S. (2022). Analisis Kinerja Algoritma Machine Learning Untuk Klasifikasi Emosi. *Building of Informatics, Technology and Science (BITS)*, 4(2), 1027–1034. <https://doi.org/10.47065/bits.v4i2.2261>

- Taha, K., Yoo, P. D., Yeun, C., Homouz, D., & Taha, A. (2024). A comprehensive survey of text classification techniques and their research applications: Observational and experimental insights. *Computer Science Review*, 54(August), 100664. <https://doi.org/10.1016/j.cosrev.2024.100664>
- Tyas, S. M. P., Sarno, R., Haryono, A. T., & Rossa Sungkono, K. (2023). A Robustly Optimized BERT using Random Oversampling for Analyzing Imbalanced Stock News Sentiment Data. *ICCoSITE 2023 - International Conference on Computer Science, Information Technology and Engineering: Digital Transformation Strategy in Facing the VUCA and TUNA Era*, 897–902. <https://doi.org/10.1109/ICCoSITE57641.2023.10127725>
- Untoro, M. C., & Yusuf, M. A. N. M. (2023). Evaluate of Random Undersampling Method and Majority Weighted Minority Oversampling Technique in Resolve Imabalanced Dataset. *IT Journal Research and Development*, 8(1), 1–13. <https://doi.org/10.25299/itjrd.2023.12412>
- Wibawa, P. W. A., & Pramatha, C. (2023). Systematic Literature Review: Machine Learning Methods in Emotion Classification in Textual Data. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*, 12(3), 425–433. <https://doi.org/10.32736/sisfokom.v12i3.1787>
- Wiciaputra, Y. K., Young, J. C., & Rusli, A. (2021). Bilingual text classification in english and indonesian via transfer learning using *XLM-RoBERTa*. *International Journal of Advances in Soft Computing and Its Applications*, 13(3), 72–87. <https://doi.org/10.15849/ijasca.211128.06>
- Yoo, B., & Rayz, J. (2021). Understanding Emojis for Sentiment Analysis. *Proceedings of the International Florida Artificial Intelligence Research Society Conference, FLAIRS*, 34. <https://doi.org/10.32473/flairs.v34i1.128562>
- Zidan, M., Elhenawy, I., Abas, A., & Othman, M. (2022). Textual Emotion Detection Approaches: a Survey. *Future Computing and Informatics Journal*, 7(1), 32–58. <https://doi.org/10.54623/fue.fcij.7.1.3>