

## DAFTAR PUSTAKA

- Abraham, M. P., Udaya Kumar Reddy, K. R., & Abraham, S. (2023). Aspect Based Sentiment Analysis System using Machine Learning Techniques. *2023 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics, DISCOVER* 2023. <https://doi.org/10.1109/DISCOVER58830.2023.10316716>
- Al-Thubaity, A., Alhoshan, M., & Hazzaa, I. (2015). Using word N-grams as features in Arabic text classification. *Studies in Computational Intelligence*, 569. [https://doi.org/10.1007/978-3-319-10389-1\\_3](https://doi.org/10.1007/978-3-319-10389-1_3)
- Altrad, A., Pathmanathan, P. R., Al Moaiad, Y., Endara, Y. M., Aseh, K., El-Ebiary, Y. A. B., Farea, M. M., Latiff, N. A. A., & Saany, S. I. A. (2021). Amazon in business to customers and overcoming obstacles. 2021 2nd International Conference on Smart Computing and Electronic Enterprise: Ubiquitous, Adaptive, and Sustainable Computing Solutions for New Normal, ICSCEE 2021. <https://doi.org/10.1109/ICSCEE50312.2021.9498129>
- Amirabadi, M. A., Kahaei, M. H., & Nezamalhosseini, S. A. (2020). Novel suboptimal approaches for hyperparameter tuning of deep neural network [under the shelf of optical communication]. *Physical Communication*, 41. <https://doi.org/10.1016/j.phycom.2020.101057>
- Asosiasi Penyelenggara Jasa Internet Indonesia (APJII). (2025). Survei Penetrasi Internet dan Perilaku Penggunaan Internet 2025. <https://survei.apjii.or.id/survei/group/11>
- Aydemir, E., Barua, P. D., Chakraborty, S., Hafeez-Baig, A., Dogan, S., Tuncer, T., Tan, R.-S., & Acharya, U. R. (2025). Iterated and Linear SVMs in Text Classification: An Ensemble Feature Engineering Model. *Arabian Journal for Science and Engineering*. <https://doi.org/10.1007/s13369-025-10666-0>
- Bensoltane, R., & Zaki, T. (2021). Comparing word embedding models for Arabic aspect category detection using a deep learning-based approach. *E3S Web of Conferences*, 297. <https://doi.org/10.1051/e3sconf/202129701072>
- Bhansali, A., Chandravadiya, A., Panchal, B. Y., Bohara, M. H., & Ganatra, A. (2022). Language Identification Using Combination of Machine Learning Algorithms and Vectorization Techniques. *2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering Icacite 2022*, 1329–1334. <https://doi.org/10.1109/ICACITE53722.2022.9823628>
- Bounabi, M., Moutaouakil, K. E., & Satori, K. (2018). A probabilistic vector representation and neural network for text classification. Dalam

*Communications in Computer and Information Science* (Vol. 872).  
[https://doi.org/10.1007/978-3-319-96292-4\\_27](https://doi.org/10.1007/978-3-319-96292-4_27)

- Camilleri, M. A. (2022). E-commerce websites, consumer order fulfillment and after-sales service satisfaction: the customer is always right, even after the shopping cart check-out. *Journal of Strategy and Management*, 15(3).  
<https://doi.org/10.2139/ssrn.3853156>
- Chebolu, S. U. S., Rosso, P., Kar, S., & Solorio, T. (2023). Survey on Aspect Category Detection. *ACM Computing Surveys*, 55(7). <https://doi.org/10.1145/3544557>
- Chong, K., & Shah, N. (2022). Comparison of Naive Bayes and SVM Classification in Grid-Search Hyperparameter Tuned and Non-Hyperparameter Tuned Healthcare Stock Market Sentiment Analysis. *International Journal of Advanced Computer Science and Applications*, 13(12), 90–94.  
<https://doi.org/10.14569/IJACSA.2022.0131213>
- Chuang, H.-Y., & Chou, W. (2025). SilhouetteScoreinR: Beyond traditional network layouts by leveraging local cohesion and nearest neighbor separation. *Methodsx*, 15. <https://doi.org/10.1016/j.mex.2025.103622>
- Danish, S. M. H., Hasnain, S. M. E., Ashraf, H., & Rukaiya, R. (2024). Comparative Analysis of BERT and TF-IDF for Textual Semantic Similarity Assessment. *Proceedings of the IEEE International Multi Topic Conference Inmic*, (2024).  
<https://doi.org/10.1109/INMIC64792.2024.11004377>
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *NAACL HLT 2019 - 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies - Proceedings of the Conference, 1*.
- Dong, S. (2025). Research on the Enhancement and Improvement of Logistic Regression Models. *2025 6th International Conference on Electrical Electronic Information and Communication Engineering Eeice 2025*, 492–496.  
<https://doi.org/10.1109/EEICE65049.2025.11033904>
- Efimov, V. (2023, September 12). *Large Language Models: SBERT – Sentence-BERT*.  
<https://towardsdatascience.com/sbert-deb3d4aef8a4/>.
- Flores, V., Luengo, R., Ulloa, C. H., & Peralta, B. (2024). Toward Automatic Vehicle Classification for Toll Booths on Chilean Interurban Highways. *Proceedings International Conference of the Chilean Computer Science Society Sccc*.  
<https://doi.org/10.1109/SCCC63879.2024.10767607>
- Gautam, A. K., & Bansal, A. (2022). Effect of Features Extraction Techniques on Cyberstalking Detection Using Machine Learning Framework. *Journal of*

- Advances in Information Technology*, 13(5), 486–502.  
<https://doi.org/10.12720/jait.13.5.486-502>
- Geisler, T., Ray, H., & Xie, Y. (2023). Finding the Proverbial Needle: Improving Minority Class Identification Under Extreme Class Imbalance. *Journal of Classification*, 40(1). <https://doi.org/10.1007/s00357-023-09431-5>
- Ghag, K. V., & Shah, K. (2018). Conceptual sentiment analysis model. *International Journal of Electrical and Computer Engineering*, 8(4), 2358–2366. <https://doi.org/10.11591/ijece.v8i4.pp2358-2366>
- Glahn, B. (2014). A nonsymmetric logit model and grouped predictand category development. *Monthly Weather Review*, 142(8), 2991–3002. <https://doi.org/10.1175/MWR-D-13-00300.1>
- Gomes, J. R. S. A., Garcia, E. A. S., Junior, A. F. B., Rodrigues, R. C., Silva, D. F. C., Maia, D. F., da Silva, N. F. F., Filho, A. R. G., & da Silva Soares, A. (2022). Deep Learning Brasil at ABSAPT 2022: Portuguese Transformer Ensemble Approaches. *CEUR Workshop Proceedings*, 3202.
- Graff, M., Moctezuma, D., & Téllez, E. S. (2025). Bag-of-Word approach is not dead: A performance analysis on a myriad of text classification challenges. *Natural Language Processing Journal*, 11, 100154. <https://doi.org/10.1016/j.nlp.2025.100154>
- Guzmán-Monteza, Y. (2023). Assessment of an annotation method for the detection of Spanish argumentative, non-argumentative, and their components. *Telematics and Informatics Reports*, 11. <https://doi.org/10.1016/j.teler.2023.100068>
- Ha, T.-T., Nguyen, V.-N., Nguyen, K.-H., Nguyen, K.-A., & Than, Q.-K. (2021). Utilizing SBERT For Finding Similar Questions in Community Question Answering. *Proceedings International Conference on Knowledge and Systems Engineering Kse*, 2021-Novem. <https://doi.org/10.1109/KSE53942.2021.9648830>
- Hermawan, Indra. (2020). *Product Cognition, Platform Emotion, Behavior Intention, and Actual Behavior Stage in Cross Border E-commerce (Case Study: Shopee as The Biggest Cross Border E-Commerce in Indonesia)*. 212–217. <https://doi.org/10.1109/IC2IE50715.2020.9274563>
- Idris, I., Mustofa, Y., & Salihi, I. (2023). Analisis Sentimen Terhadap Penggunaan Aplikasi Shopee Menggunakan Algoritma Support Vector Machine (SVM) (Universitas Negeri Gorontalo, Penerj.). *Jambura Journal of Electrical and Electronics Engineering*, 5(61), 32–35. <https://doi.org/10.37905/JJEEE.V5I1.16830>

- Jiang, W. (2022). Applications of singular value decomposition in data reduction. *Proceedings of SPIE the International Society for Optical Engineering*, 12345. <https://doi.org/10.1117/12.2648696>
- Kadhim, A. I. (2019). Term Weighting for Feature Extraction on Twitter: A Comparison between BM25 and TF-IDF. *2019 International Conference on Advanced Science and Engineering Icoase 2019*, 124–128. <https://doi.org/10.1109/ICOASE.2019.8723825>
- Kasthuriarachchy, B. H., de Zoysa, K., & Premarathne, H. L. (2015). *Context-Aware Sentiment Classification*. <https://ieeexplore.ieee.org/document/7377709>
- Keivandarian, N., & Carvalho, M. (2025). A Comprehensive Categorization and Comparative Analysis of Methodologies in Aspect-Based Sentiment Analysis. *Proceedings 2025 IEEE Conference on Artificial Intelligence Cai 2025*, 132–138. <https://doi.org/10.1109/CAI64502.2025.00028>
- Khader, M., & Zarzour, H. (2025). Enhancing Sentiment Analysis via Advanced Deep Learning Techniques. *Pais 2025 Proceeding 7th International Conference on Pattern Analysis and Intelligent Systems*. <https://doi.org/10.1109/PAIS66004.2025.11126468>
- Khalastchi, E., Kalech, M., & Rokach, L. (2017). A hybrid approach for improving unsupervised fault detection for robotic systems. *Expert Systems with Applications*, 81, 372–383. <https://doi.org/10.1016/j.eswa.2017.03.058>
- Kocon, J., Radom, J., Kaczmarz-Wawryk, E., Wabnic, K., Zajackowska, A., & Zasko-Zielinska, M. (2021). AspectEmo: Multi-Domain Corpus of Consumer Reviews for Aspect-Based Sentiment Analysis. *IEEE International Conference on Data Mining Workshops Icdmw, 2021-Decem*, 166–173. <https://doi.org/10.1109/ICDMW53433.2021.00027>
- Kumar, A. K. M., Anusha, S. G., Sharath, H. N., Milind, M., & Charan, M. (2024). Impact of Different Embeddings on Text Classification. *IEEE International Conference on Recent Advances in Science and Engineering Technology Icraset 2024*. <https://doi.org/10.1109/ICRASET63057.2024.10895636>
- Kushadiani, S. K., Nugroho, B., Mardian, S., Muhammad-Bello, B., & Hermawan, A. (2021). Evaluating Digital Economy in the Covid-19 Pandemic Era: A Review. *ACM International Conference Proceeding Series*. <https://doi.org/10.1145/3479645.3479653>
- Kuznetsov, D., Gavrilov, I., Benkovich, N., Charnetskaya, T., & Yavorskiy, R. (2019). Media placement: Using sentiment analysis in brand reputation maintaining. *CEUR Workshop Proceedings*, 2478.

- Leonid, Thomas., Kumar, Praveen., Subramanian, Bala., & Pandi, Senthil. (2025). Research and Analysis of Market Trends with Sentimental Analysis Using SBERT. *International Conference on Emerging Trends in Engineering and Technology Icetet*. <https://doi.org/10.1109/ICETETSIP64213.2025.11156722>
- Li, Y., Adams, N., & Bellotti, T. (2022). A Relabeling Approach to Handling the Class Imbalance Problem for Logistic Regression. *Journal of Computational and Graphical Statistics*, *31*(1), 241–253. <https://doi.org/10.1080/10618600.2021.1978470>
- Llaca-Sánchez, B. A., García-Noguez, L. R., Aceves-Fernández, M. A., Takacs, A., & Tovar-Arriaga, S. (2025). Exploring LLM Embedding Potential for Dementia Detection Using Audio Transcripts. *Eng*, *6*(7). <https://doi.org/10.3390/eng6070163>
- Lutfiana, D., Yanuar, A., & Sulistiyaningsih, F. (2023). Analisis Kualitas Pelayanan Shopee dan Tokopedia Menggunakan Metode Gap Analysis dan Importance Performance Analysis (IPA). <https://ojs.cahayamandalika.com/>
- Mandala, V., Rajavarman, R., Jamunadevi, C., Janani, R., & Avudaiappan, T. (2023). Recognition of E-Commerce through Big Data Classification and Data Mining Techniques Involving Artificial Intelligence. *Proceedings of the 8th International Conference on Communication and Electronics Systems, ICCES 2023*. <https://doi.org/10.1109/ICCES57224.2023.10192673>
- Marchevsky, A. M., Walts, A. E., Lissenberg-Witte, B. I., & Thunnissen, E. (2020). Pathologists should probably forget about kappa. Percent agreement, diagnostic specificity and related metrics provide more clinically applicable measures of interobserver variability. *Annals of Diagnostic Pathology*, *47*. <https://doi.org/10.1016/j.anndiagpath.2020.151561>
- Marutho, D., Muljono, Rustad, S., & Purwanto. (2024). Optimizing aspect-based sentiment analysis using sentence embedding transformer, bayesian search clustering, and sparse attention mechanism. *Journal of Open Innovation: Technology, Market, and Complexity*, *10*(1). <https://doi.org/10.1016/j.joitmc.2024.100211>
- McKenzie, K. A., & Mahnken, J. D. (2023). Simulating and estimating agreement in the presence of multiple raters and covariates. *Statistics in Medicine*, *42*(11), 1687–1698. <https://doi.org/10.1002/sim.9694>
- Minn, T. Y., Khant Myat, A., Nyo Theint, K. N., Thinn Wai, T., Tar, H. H., Win Thant, W., & Zin, T. T. (2025). Systematic Comparison of Vectorization Methods in Topic Classification for Myanmar Language. *Proceedings of the 22nd IEEE International Conference on Computer Applications Icca 2025*. <https://doi.org/10.1109/ICCA65395.2025.11011194>

- Paramita, A. S., & Jusak, J. (2025). Fine-Grained Sentiment Analysis Approach on Customer Reviews Based on Aspect-Level Emotion Detection. *Journal of Applied Data Sciences*, 6(3). <https://doi.org/10.47738/jads.v6i3.964>
- Pitoyo, D., Seraya, N. N., Shichiyakh, R. A., Laxmi Lydia, E., & Shankar, K. (2020). Strengthening environment resistance for improving quality growth. *Journal of Environmental Treatment Techniques*, 8(3).
- Poplewell, M., Reizes, J., & Zaslowski, C. (2019). Appropriate Statistics for Determining Chance-Removed Interpractitioner Agreement. *Journal of Alternative and Complementary Medicine*, 25(11), 1115–1120. <https://doi.org/10.1089/acm.2017.0297>
- Porcello, J. C. (2019). Designing and Implementing SVMs for High-Dimensional Knowledge Discovery Using FPGAs. *IEEE Aerospace Conference Proceedings, 2019-March*. <https://doi.org/10.1109/AERO.2019.8741916>
- Punithamalar, M. S., & Sasikala, V. K. (2024). An Empirical Analysis and Comprehensive Investigation on Customer Quality Satisfaction in Online Shopping. *MSW Management*, 34(1). <https://doi.org/10.7492/f754er14>
- Rafiani, K. M., Yunanda, R. A., & Rusmanto, T. (2024). Determinants of QRIS Usage as a Digital Payment Tool for MSMEs. *Journal of Theoretical and Applied Information Technology*, 102(3).
- Ran, X., Xi, Y., Lu, Y., Wang, X., & Lu, Z. (2023). Comprehensive survey on hierarchical clustering algorithms and the recent developments. *Artificial Intelligence Review*, 56(8), 8219–8264. <https://doi.org/10.1007/s10462-022-10366-3>
- Remesh, R., & Pattabiraman, V. (2017). A survey on the cures for the curse of dimensionality in big data. *Asian Journal of Pharmaceutical and Clinical Research*, 10, 355–360. <https://doi.org/10.22159/ajpcr.2017.v10s1.19755>
- Rizkya, D. D., Retno, S., & Yunizar, Z. (2025). Evaluating the Quality of Agglomerative Hierarchical Clustering on Crime Data in Indonesia. *International Journal of Engineering Science and Information Technology*, 5(2), 455–461. <https://doi.org/10.52088/ijesty.v5i2.863>
- Sheridan, P., & Onsjö, M. (2024). The hypergeometric test performs comparably to TF-IDF on standard text analysis tasks. *Multimedia Tools and Applications*, 83(10), 28875–28890. <https://doi.org/10.1007/s11042-023-16615-z>
- Shrivash, B. K., Verma, D. K., & Pandey, P. (2025). A Novel Framework for Text Preprocessing using NLP Approaches and Classification using Random Forest Grid Search Technique for Sentiment Analysis. *Economic Computation and*

*Economic Cybernetics Studies and Research*, 59(2).  
<https://doi.org/10.24818/18423264/59.2.25.06>

- Silva, B., Leite, P., & Ribeiro, O. R. (2025). Evaluating Spam Detection Techniques: A Comparison of TF-IDF and Sentence Embeddings with Machine Learning Models. *ISDFS 2025 - 13th International Symposium on Digital Forensics and Security*. <https://doi.org/10.1109/ISDFS65363.2025.11012070>
- Singh, H., Thakral, P., Rathor, K., & Verma, D. C. (2025). A Machine Learning Based Diabetes Mellitus Classification on Weighted Class using Inverse Class Frequency Method. *Proceedings of IEEE International Conference on Signal Processing Computing and Control*, 788–793. <https://doi.org/10.1109/ISPC66872.2025.11039549>
- Soni, M., Kumari, N., Singh, P., & Kashtariya, V. (2024). Analysis on Movie Reviews: A Comparative Study. *International Conference on Parallel Distributed and Grid Computing Pdgc*, (2024), 252–257. <https://doi.org/10.1109/PDGC64653.2024.10984113>
- Souza, R., Lucena, O., Bento, M., Garrafa, J., Appenzeller, S., Rittner, L., Lotufo, R., & Frayne, R. (2018). Reliability of using single specialist annotation for designing and evaluating automatic segmentation methods: A skull stripping case study. *Proceedings International Symposium on Biomedical Imaging, 2018-April*, 1344–1347. <https://doi.org/10.1109/ISBI.2018.8363820>
- Taneja, K., Vashishtha, J., & Ratnoo, S. (2023). Efficient Deep Pre-trained Sentence Embedding Model for Similarity Search. *International Journal of Computer Information Systems and Industrial Management Applications*, 15(2023), 605–615.
- Tian, Y., Chen, W., Hu, B., Song, Y., & Xia, F. (2023). End-to-end Aspect-based Sentiment Analysis with Combinatory Categorical Grammar. *Proceedings of the Annual Meeting of the Association for Computational Linguistics*. <https://doi.org/10.18653/v1/2023.findings-acl.859>
- Umar, N., & Nur, M. A. (2022). Application of Naïve Bayes Algorithm Variations On Indonesian General Analysis Dataset for Sentiment Analysis. *Jurnal Resti*, 6(4), 585–590. <https://doi.org/10.29207/resti.v6i4.4179>
- Wang, X., & Xu, Y. (2019). An improved index for clustering validation based on Silhouette index and Calinski-Harabasz index. *Iop Conference Series Materials Science and Engineering*, 569(5). <https://doi.org/10.1088/1757-899X/569/5/052024>

- Westin, F. (2024). Time Period Categorization in Fiction: A Comparative Analysis of Machine Learning Techniques. *Cataloging and Classification Quarterly*, 62(2), 124–153. <https://doi.org/10.1080/01639374.2024.2315548>
- Widhianto, R. A., & Romadhony, A. (2021). Aspect Term Extraction Using Deep Learning-Based Approach on Indonesian Restaurant Reviews. *2021 9th International Conference on Information and Communication Technology, ICoICT 2021*. <https://doi.org/10.1109/ICoICT52021.2021.9527472>
- Wu, T., Wang, Y., & Quanch, N. (2025). Advancements in Natural Language Processing: Exploring Transformer-Based Architectures for Text Understanding. *2025 5th International Conference on Artificial Intelligence and Industrial Technology Applications Aiita 2025*, 1384–1388. <https://doi.org/10.1109/AIITA65135.2025.11048063>
- Wu, Y. (2018). Power calculation of adjusted McNemar's test based on clustered data of varying cluster size. *Biometrical Journal*, 60(6), 1190–1200. <https://doi.org/10.1002/bimj.201800034>
- Yu, H., & Hou, X. (2022). Hierarchical clustering in astronomy. *Scientia Sinica Physica Mechanica Et Astronomica*, 52(8). <https://doi.org/10.1360/SSPMA-2022-0053>
- Zhang, W., Jiang, L., Zhang, H., & Chen, L. (2021). A Two-Layer Bayes Model: Random Forest Naive Bayes. *Jisuanji Yanjiu Yu Fazhan Computer Research and Development*, 58(9), 2040–2051. <https://doi.org/10.7544/issn1000-1239.2021.20200521>