

DAFTAR PUSTAKA

- [1] A. Feger, "Worldwide ecommerce sales to break \$6 trillion, make up a fifth of total retail sales," *emarketer*. Accessed: Nov. 27, 2025. [Online]. Available: <https://www.emarketer.com/content/worldwide-ecommerce-sales-break-6-trillion>
- [2] S. Kemp, "Digital 2025: Indonesia," *datareportal*. Accessed: Nov. 27, 2025. [Online]. Available: <https://datareportal.com/reports/digital-2025-indonesia>
- [3] G. Indonesia, "e-Conomy SEA 2024: Perekonomian digital Indonesia akan mencapai GMV \$90 miliar pada tahun 2024," *google*. Accessed: Nov. 27, 2025. [Online]. Available: <https://blog.google/intl/id-id/e-conomy-sea-2024-perekonomian-digital-indonesia-akan-mencapai-gmv-90-miliar-pada-tahun-2024/>
- [4] P. and C. M. Intelligence, "A Comprehensive Look at Indonesia's E-Commerce Market," *Payments and Commerce Market Intelligence*. Accessed: Oct. 30, 2025. [Online]. Available: <https://paymentscmi.com/insights/indonesia-ecommerce-market-data/>
- [5] T. Group, "11 Marketplaces in Indonesia for Businesses Selling Online," *tmogroup asia*. [Online]. Available: <https://www.tmogroup.asia/insights/top-online-marketplaces-indonesia/>
- [6] S. Kemp, "Digital 2025: online shopping still popping," *datareportal*. [Online]. Available: <https://datareportal.com/reports/digital-2025-sub-section-online-shopping>
- [7] T. Chen, P. Samaranayake, X. Cen, M. Qi, and Y.-C. Lan, "The Impact of Online Reviews on Consumers' Purchasing Decisions: Evidence From an Eye-Tracking Study," *Front. Psychol.*, vol. 13, Jun. 2022, doi: 10.3389/fpsyg.2022.865702.
- [8] D. A. Kristiyanti and Sri Hardani, "Sentiment Analysis of Public Acceptance of Covid-19 Vaccines Types in Indonesia using Naïve Bayes, Support Vector Machine, and Long Short-Term Memory (LSTM)," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 7, no. 3, pp. 722–732, Jun. 2023, doi: 10.29207/resti.v7i3.4737.
- [9] S. S. Felina Khasanah, Muhammad Hatta, "ANALISIS SENTIMEN PADA ULASAN PEMBELIAN PRODUK DENGAN METODE SUPPORT VECTOR MACHINE," *J. Jar. Sist. Inf. Robot.*, vol. Vol 8, No, 2024, doi: 10.58486/jsr.v8i2.426.
- [10] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," May 2019, [Online]. Available: <http://arxiv.org/abs/1810.04805>
- [11] D. K. R. M.P. Geetha, "Improving the performance of aspect based sentiment analysis using fine-tuned Bert Base Uncased model," *Int. J. Intell. Networks*,

vol. Volume 2, p. Pages 64-69, 2021, doi: 10.1016/j.ijin.2021.06.005.

- [12] D. Cortiz, “Exploring Transformers in Emotion Recognition: a comparison of BERT, DistillBERT, RoBERTa, XLNet and ELECTRA,” Apr. 2021, [Online]. Available: <http://arxiv.org/abs/2104.02041>
- [13] F. Koto, A. Rahimi, J. H. Lau, and T. Baldwin, “IndoLEM and IndoBERT: A Benchmark Dataset and Pre-trained Language Model for Indonesian NLP,” in *COLING 2020 - 28th International Conference on Computational Linguistics, Proceedings of the Conference*, 2020. doi: 10.18653/v1/2020.coling-main.66.
- [14] S. Aras, M. Yusuf, R. Y. Ruimassa, E. A. B. Wambrauw, and E. B. Pala’lengan, “Sentiment Analysis on Shopee Product Reviews Using IndoBERT,” *J. Inf. Syst. Informatics*, vol. 6, no. 3, pp. 1616–1627, Sep. 2024, doi: 10.51519/journalisi.v6i3.814.
- [15] R. A. Rudiyanto and E. B. Setiawan, “Sentiment Analysis Using Convolutional Neural Network (CNN) and Particle Swarm Optimization on Twitter,” *JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer)*, vol. 9, no. 2, pp. 188–195, Feb. 2024, doi: 10.33480/jitk.v9i2.5201.
- [16] G. K. Pandey and S. Srivastava, “ResNet-18 comparative analysis of various activation functions for image classification,” in *2023 International Conference on Inventive Computation Technologies (ICICT)*, IEEE, Apr. 2023, pp. 595–601. doi: 10.1109/ICICT57646.2023.10134464.
- [17] J. Lin, “Application of ResNet-50 Based User Sentiment Analysis in Digital Media Advertising Design,” 2024. doi: 10.3233/FAIA240037.
- [18] R. R. Balqis and D. A. Kristiyanti, “BERT-DeiT Model: Multimodal Sentiment Analysis for E-Commerce Product Review,” in *2025 4th International Conference on Electronics Representation and Algorithm (ICERA)*, IEEE, Jun. 2025, pp. 221–226. doi: 10.1109/ICERA66156.2025.11087379.
- [19] J. Ren, “Multimodal Sentiment Analysis Based on BERT and ResNet,” Dec. 2024, [Online]. Available: <http://arxiv.org/abs/2412.03625>
- [20] C. N. Karaeng and D. A. Kristiyanti, “Multimodal Sentiment Analysis Approach Combining Transfer Learning and Generative AI of E-Commerce Product Reviews,” in *2025 4th International Conference on Electronics Representation and Algorithm (ICERA)*, IEEE, Jun. 2025, pp. 01–06. doi: 10.1109/ICERA66156.2025.11087376.
- [21] A. Z. Apurba, M. F. S. Titu, M. A. Pavel, I. T. Naheen, and R. Khan, “Fusion of Image Filtering and Knowledge-Distilled YOLO Models for Root Canal Failure Diagnosis,” *IEEE Access*, vol. 13, pp. 66557–66573, 2025, doi: 10.1109/ACCESS.2025.3560998.
- [22] J. Li, Y. Huang, Y. Lu, L. Wang, Y. Ren, and R. Chen, “Sentiment Analysis

- Using E-Commerce Review Keyword-Generated Image with a Hybrid Machine Learning-Based Model,” *Comput. Mater. Contin.*, vol. 80, no. 1, pp. 1581–1599, 2024, doi: 10.32604/cmc.2024.052666.
- [23] G. Rizzi and P. Bertola, “Exploring the generative AI potential in the fashion design process: an experimental experience on the collaboration between fashion design practitioners and generative AI tools,” *Eur. J. Cult. Manag. Policy*, vol. 15, May 2025, doi: 10.3389/ejcmp.2025.13875.
- [24] K. Y. Ram Potham, Jinkai Qiu, Andrew Zhang, “Enhancing Fashion E-Commerce with AI: A Multi-LoRA Approach to Digital Clothing Representation,” 2024, doi: 10.13140/RG.2.2.13704.64001.
- [25] D. S. Bisht, A. Pant, and J. Bhatt, “A Sentiment Analysis Based on the Comparison of BERT and Naive Bayes,” in *2024 First International Conference on Technological Innovations and Advance Computing (TIACOMP)*, 2024, pp. 419–423. doi: 10.1109/TIACOMP64125.2024.00076.
- [26] S. N. Aymen Omar Alharm, “Enhancing Natural Disaster Response: A Deep Learning Approach to Disaster Sentiment Analysis Using Bert and Lstm,” *Soc. Sci. Res. Netw.*, 2024, [Online]. Available: <http://dx.doi.org/10.2139/ssrn.4755638>
- [27] R. S. S. C. R. Gupta, “Sentiment Analysis of Customer Reviews from E-Commerce Website,” *2025 Int. Conf. Adv. Power, Signal, Inf. Technol.*, 2025, doi: 10.1109/APSIT63993.2025.11086266.
- [28] M. I. M. Irwan Budiman, Mohammad Reza Faisal, Astina Faridhah, Andi Farmadi and F. A. Triando Hamonangan Saragih, “Classification Performance Comparison of BERT and IndoBERT on Self-Report of COVID-19 Status on Social Media,” *J. Comput. Sci. Inst.*, vol. 30, pp. 61–67, 2024, doi: 10.35784/jcsi.5564.
- [29] T. B. S. Asyer Aprinando Pratama Simarmata, “Sentiment Analysis on BRImo Application Reviews Using IndoBERT,” *J. Appl. Informatics Comput.*, vol. Vol. 9 No., 2025, [Online]. Available: <https://doi.org/10.30871/jaic.v9i3.8162>
- [30] S. A. Shiping Ye, V. Sorokina, “Clothing Image Retrieval Algorithm Using YOLOv8 and Attention Model,” *Proc. 9TH Int. Work. “IMAGE Min. THEORY Appl.*, vol. Volume 34, pp. 1105–1110, 2025, doi: 10.1134/S1054661824701153.
- [31] G. Z. Zheyu Zhao, Zhongqing Wang, Shichen Li, Hongling Wang, “Bridging Modality Gap for Effective Multimodal Sentiment Analysis in Fashion-related Social Media,” *Proc. 31st Int. Conf. Comput. Linguist.*, pp. 1813–1823, 2025, [Online]. Available: <https://aclanthology.org/2025.coling-main.123/>

- [32] D. J. Simjanović, N. Zdravković, and N. O. Vesić, “On the Factors of Successful e-Commerce Platform Design during and after COVID-19 Pandemic Using Extended Fuzzy AHP Method,” *Axioms*, vol. 11, no. 3, p. 105, Feb. 2022, doi: 10.3390/axioms11030105.
- [33] A. L. Kilay, B. H. Simamora, and D. P. Putra, “The Influence of E-Payment and E-Commerce Services on Supply Chain Performance: Implications of Open Innovation and Solutions for the Digitalization of Micro, Small, and Medium Enterprises (MSMEs) in Indonesia,” *J. Open Innov. Technol. Mark. Complex.*, vol. 8, no. 3, p. 119, Sep. 2022, doi: 10.3390/joitmc8030119.
- [34] A. V. Mali and R. Rachmawati, “The Utilization of E-Commerce in the SMEs Development to Support Smart Village in Turi District, Sleman Regency,” in *IOP Conference Series: Earth and Environmental Science*, 2022. doi: 10.1088/1755-1315/1039/1/012068.
- [35] I. Firmansyah, “Model Bisnis E Commerce: Panduan & Contoh Praktis,” *softwareseni*. Accessed: Aug. 09, 2025. [Online]. Available: <https://www.softwareseni.co.id/blog/5-model-bisnis-ecommerce-b2b-b2c-c2c-c2b-b2g>
- [36] O. Wijaya, “E-Commerce: Perkembangan, Tren, dan Peraturan Perundang-Undangan,” *E-Bisnis J. Ilm. Ekon. dan Bisnis*, vol. 16, no. 1, pp. 41–47, May 2023, doi: 10.51903/e-bisnis.v16i1.1083.
- [37] E. Y. Nasution, P. Hariani, L. S. Hasibuan, and W. Pradita, “Perkembangan Transaksi Bisnis E-Commerce terhadap Pertumbuhan Ekonomi di Indonesia,” *Jesya*, 2020, doi: 10.36778/jesya.v3i2.227.
- [38] D. Dwidienawati, D. Tjahjana, S. B. Abdinagoro, D. Gandasari, and Munawaroh, “Customer review or influencer endorsement: which one influences purchase intention more?,” *Heliyon*, vol. 6, no. 11, p. e05543, Nov. 2020, doi: 10.1016/j.heliyon.2020.e05543.
- [39] T. T. A. Ngo, C. T. Bui, H. K. L. Chau, and N. P. N. Tran, “Electronic word-of-mouth (eWOM) on social networking sites (SNS): Roles of information credibility in shaping online purchase intention,” *Heliyon*, vol. 10, no. 11, p. e32168, Jun. 2024, doi: 10.1016/j.heliyon.2024.e32168.
- [40] N. Peña-García, M. Losada-Otálora, D. P. Auza, and M. P. Cruz, “Reviews, trust, and customer experience in online marketplaces: the case of Mercado Libre Colombia,” *Front. Commun.*, vol. 9, Sep. 2024, doi: 10.3389/fcomm.2024.1460321.
- [41] R. Rachmiani, N. Kintan Oktadinna, and T. Rachmat Fauzan, “The Impact of Online Reviews and Ratings on Consumer Purchasing Decisions on E-commerce Platforms,” *Int. J. Manag. Sci. Inf. Technol.*, vol. 4, no. 2, pp. 504–515, Dec. 2024, doi: 10.35870/ijmsit.v4i2.3373.
- [42] M. Khder, “Web Scraping or Web Crawling: State of Art, Techniques,

- Approaches and Application,” *Int. J. Adv. Soft Comput. its Appl.*, vol. 13, no. 3, pp. 145–168, Dec. 2021, doi: 10.15849/IJASCA.211128.11.
- [43] N. Adila, “Implementation of Web Scraping for Journal Data Collection on the SINTA Website,” *Sinkron*, 2022, doi: 10.33395/sinkron.v7i4.11576.
 - [44] C. Lotfi, S. Srinivasan, M. Ertz, and I. Latrous, “Web Scraping Techniques and Applications: A Literature Review,” in *SCRS CONFERENCE PROCEEDINGS ON INTELLIGENT SYSTEMS*, Soft Computing Research Society, 2021, pp. 381–394. doi: 10.52458/978-93-91842-08-6-38.
 - [45] J. R. Jim, M. A. R. Talukder, P. Malakar, M. M. Kabir, K. Nur, and M. F. Mridha, “Recent advancements and challenges of NLP-based sentiment analysis: A state-of-the-art review,” *Nat. Lang. Process. J.*, vol. 6, p. 100059, Mar. 2024, doi: 10.1016/j.nlp.2024.100059.
 - [46] P. Nandwani and R. Verma, “A review on sentiment analysis and emotion detection from text,” *Soc. Netw. Anal. Min.*, vol. 11, no. 1, p. 81, Dec. 2021, doi: 10.1007/s13278-021-00776-6.
 - [47] D. A. Kristiyanti, A. Al Kaafi, E. Purwaningsih, E. Nurelasari, and B. Nisa, “Deep learning for Twitter sentiment analysis about the pros and cons of Covid-19 vaccines in Indonesia,” 2023, p. 030025. doi: 10.1063/5.0128686.
 - [48] D. Ajeng Kristiyanti and A. Hairul Umam, “S-Shaped Versus V-Shaped Transfer Functions for Salp Swarm Algorithm in Sentiment Analysis Feature Selection Problem,” in *2025 2nd International Conference on Advanced Innovations in Smart Cities (ICAISC)*, IEEE, Feb. 2025, pp. 01–08. doi: 10.1109/ICAISC64594.2025.10959425.
 - [49] H. Yu and Y. Guo, “Generative artificial intelligence empowers educational reform: current status, issues, and prospects,” *Front. Educ.*, vol. 8, Jun. 2023, doi: 10.3389/educ.2023.1183162.
 - [50] S. S. Sengar, A. Bin Hasan, S. Kumar, and F. Carroll, “Generative Artificial Intelligence: A Systematic Review and Applications,” May 2024, [Online]. Available: <http://arxiv.org/abs/2405.11029>
 - [51] Gabriel Torres, “Generative AI: What is it and How Does it Work?,” voiceflow. Accessed: Aug. 09, 2025. [Online]. Available: <https://www.voiceflow.com/blog/generative-ai>
 - [52] L. Martini, S. Iacono, D. Zolezzi, and G. V. Vercelli, “Advancing Persistent Character Generation: Comparative Analysis of Fine-Tuning Techniques for Diffusion Models,” *AI*, vol. 5, no. 4, pp. 1779–1792, Sep. 2024, doi: 10.3390/ai5040088.
 - [53] M. Lee, “Recent Advances in Deep Learning for Protein-Protein Interaction Analysis: A Comprehensive Review,” 2023. doi: 10.3390/molecules28135169.

- [54] A. Dosovitskiy *et al.*, “AN IMAGE IS WORTH 16X16 WORDS: TRANSFORMERS FOR IMAGE RECOGNITION AT SCALE,” in *ICLR 2021 - 9th International Conference on Learning Representations*, 2021.
- [55] M. T. Yılmaz *et al.*, “Synthetic Image Generation with Stable Diffusion and Trained LoRA Model for Industrial Areas,” in *2024 Innovations in Intelligent Systems and Applications Conference (ASYU)*, IEEE, Oct. 2024, pp. 1–4. doi: 10.1109/ASYU62119.2024.10757052.
- [56] S. Gong, “Application of Stable Diffusion and LoRA Models in AI Drawing,” *Appl. Comput. Eng.*, vol. 97, no. 1, pp. 12–17, Nov. 2024, doi: 10.54254/2755-2721/97/20241294.
- [57] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, “High-Resolution Image Synthesis with Latent Diffusion Models,” Apr. 2022, [Online]. Available: <http://arxiv.org/abs/2112.10752>
- [58] I. Ulku and E. Akagunduz, “A Survey on Deep Learning-based Architectures for Semantic Segmentation on 2D images,” Mar. 2022, doi: 10.1080/08839514.2022.2032924.
- [59] M. J. Willemink *et al.*, “Preparing Medical Imaging Data for Machine Learning,” *Radiology*, vol. 295, no. 1, pp. 4–15, Apr. 2020, doi: 10.1148/radiol.2020192224.
- [60] A. Peláez-Vegas, P. Mesejo, and J. Luengo, “A Survey on Semi-Supervised Semantic Segmentation,” Feb. 2023, [Online]. Available: <http://arxiv.org/abs/2302.09899>
- [61] A. S. Gillis, “Definition data splitting,” TechTarget. Accessed: Dec. 01, 2024. [Online]. Available: <https://www.techtarget.com/searchenterpriseai/definition/data-splitting>
- [62] T. T. Wong and P. Y. Yeh, “Reliable Accuracy Estimates from k-Fold Cross Validation,” *IEEE Trans. Knowl. Data Eng.*, 2020, doi: 10.1109/TKDE.2019.2912815.
- [63] M. Sivakumar, S. Parthasarathy, and T. Padmapriya, “Trade-off between training and testing ratio in machine learning for medical image processing,” *PeerJ Comput. Sci.*, vol. 10, p. e2245, Sep. 2024, doi: 10.7717/peerj-cs.2245.
- [64] O. Rainio, J. Teuho, and R. Klén, “Evaluation metrics and statistical tests for machine learning,” *Sci. Rep.*, 2024, doi: 10.1038/s41598-024-56706-x.
- [65] J. Terven, D. M. Cordova-Esparza, A. Ramirez-Pedraza, E. A. Chavez-Urbiola, and J. A. Romero-Gonzalez, “Loss Functions and Metrics in Deep Learning,” Jul. 2023, doi: <https://doi.org/10.48550/arXiv.2307.02694>.
- [66] R. Yacouby and D. Axman, “Probabilistic Extension of Precision, Recall, and F1 Score for More Thorough Evaluation of Classification Models,” 2020. doi: 10.18653/v1/2020.eval4nlp-1.9.

- [67] N. W. S. Wardhani, M. Y. Rochayani, A. Iriany, A. D. Sulistyono, and P. Lestantyo, "Cross-validation Metrics for Evaluating Classification Performance on Imbalanced Data," in *2019 International Conference on Computer, Control, Informatics and its Applications: Emerging Trends in Big Data and Artificial Intelligence, IC3INA 2019*, 2019. doi: 10.1109/IC3INA48034.2019.8949568.
- [68] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "Knowledge Discovery and Data Mining: Towards a Unifying Framework," in *Proceedings - 2nd International Conference on Knowledge Discovery and Data Mining, KDD 1996*, 1996.
- [69] L. M. M. Trechera and F. M. Varela, "Independent Component Analysis in Knowledge Discovery in Databases Process: A Fuzzy and Genetic Approach," in *MELECON 2006 - 2006 IEEE Mediterranean Electrotechnical Conference*, IEEE, pp. 816–819. doi: 10.1109/MELCON.2006.1653223.
- [70] F. Gullo, "From Patterns in Data to Knowledge Discovery: What Data Mining Can Do," *Phys. Procedia*, vol. 62, pp. 18–22, 2015, doi: 10.1016/j.phpro.2015.02.005.
- [71] M. Alnoukari and A. El Sheikh, "Knowledge Discovery Process Models," 2012, pp. 72–100. doi: 10.4018/978-1-61350-050-7.ch004.
- [72] B. Khalili and A. W. Smyth, "SOD-YOLOv8—Enhancing YOLOv8 for Small Object Detection in Aerial Imagery and Traffic Scenes," *Sensors*, vol. 24, no. 19, p. 6209, Sep. 2024, doi: 10.3390/s24196209.
- [73] M. Gu, X. Dong, and M. Ding, "Occlusive and multi-scale clothing detection algorithm based on improved YOLOv8," *Signal, Image Video Process.*, vol. 19, no. 10, p. 859, Oct. 2025, doi: 10.1007/s11760-025-04462-4.
- [74] Z. Xu, Y. Bao, and B. Tian, "Improved YOLOv8 garment sewing defect detection method based on attention mechanism," *J. Meas. Eng.*, vol. 12, no. 4, pp. 706–721, Dec. 2024, doi: 10.21595/jme.2024.24283.
- [75] W. Zhang and Z. Zeng, "Research Progress of Convolutional Neural Network and its Application in Object Detection," Jul. 2020, [Online]. Available: <http://arxiv.org/abs/2007.13284>
- [76] X. Zhao, L. Wang, Y. Zhang, X. Han, M. Deveci, and M. Parmar, "A review of convolutional neural networks in computer vision," *Artif. Intell. Rev.*, 2024, doi: 10.1007/s10462-024-10721-6.
- [77] Nafiz Shahriar, "What is Convolutional Neural Network — CNN (Deep Learning)," Medium. Accessed: Dec. 01, 2023. [Online]. Available: <https://nafizshahriar.medium.com/what-is-convolutional-neural-network-cnn-deep-learning-b3921bdd82d5>
- [78] M. Krichen, "Convolutional Neural Networks: A Survey," *Computers*, 2023,

doi: 10.3390/computers12080151.

- [79] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” Dec. 2015, [Online]. Available: <http://arxiv.org/abs/1512.03385>
- [80] S. Zagoruyko and N. Komodakis, “Wide Residual Networks,” Jun. 2017, [Online]. Available: <http://arxiv.org/abs/1605.07146>
- [81] H. Touvron, M. Cord, M. Douze, F. Massa, A. Sablayrolles, and H. Jégou, “Training data-efficient image transformers & distillation through attention,” Jan. 2021, [Online]. Available: <http://arxiv.org/abs/2012.12877>
- [82] H. Anzum, M. N. S. Sammo, and S. Akhter, “Leveraging Data Efficient Image Transformer (DeIT) for Road Crack Detection and Classification,” in *2024 International Conference on Advances in Computing, Communication, Electrical, and Smart Systems (iCACCESS)*, IEEE, Mar. 2024, pp. 1–6. doi: 10.1109/iCACCESS61735.2024.10499539.
- [83] A. Pangestu, B. Purnama, and R. Risnandar, “Vision Transformer untuk Klasifikasi Kematangan Pisang,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 11, no. 1, pp. 75–84, Feb. 2024, doi: 10.25126/jtiik.20241117389.
- [84] Y. Zhuang and J. Kim, “A BERT-Based Multi-Criteria Recommender System for Hotel Promotion Management,” *Sustainability*, vol. 13, no. 14, p. 8039, Jul. 2021, doi: 10.3390/su13148039.
- [85] N. J. Prottasha *et al.*, “Transfer Learning for Sentiment Analysis Using BERT Based Supervised Fine-Tuning,” *Sensors*, vol. 22, no. 11, p. 4157, May 2022, doi: 10.3390/s22114157.
- [86] L. Martín-Hoz, S. Yanes-Luis, J. Huerta Cejudo, D. Gutiérrez-Reina, and E. Franco Álvarez, “A Comparative Study of BERT-Based Models for Teacher Classification in Physical Education,” *Electronics*, vol. 14, no. 19, p. 3849, Sep. 2025, doi: 10.3390/electronics14193849.
- [87] S. Benabderrahmane, P. Valtchev, J. Cheney, and T. Rahwan, “APT-LLM: Embedding-Based Anomaly Detection of Cyber Advanced Persistent Threats Using Large Language Models,” Feb. 2025, [Online]. Available: <http://arxiv.org/abs/2502.09385>
- [88] Z. Lan, M. Chen, S. Goodman, K. Gimpel, P. Sharma, and R. Soricut, “ALBERT: A Lite BERT for Self-supervised Learning of Language Representations,” Feb. 2020, [Online]. Available: <http://arxiv.org/abs/1909.11942>
- [89] L. F. Simanjuntak, R. Mahendra, and E. Yulianti, “We Know You Are Living in Bali: Location Prediction of Twitter Users Using BERT Language Model,” *Big Data Cogn. Comput.*, vol. 6, no. 3, p. 77, Jul. 2022, doi: 10.3390/bdcc6030077.

- [90] C. Bao and Y. Wang, "A Survey of Word Cloud Visualization," *J. Comput. Des. Comput. Graph.*, vol. 33, no. 4, pp. 532–544, Apr. 2021, doi: 10.3724/SP.J.1089.2021.18811.
- [91] M. Skeppstedt, M. Ahltop, K. Kucher, and M. Lindström, "From word clouds to Word Rain: Revisiting the classic word cloud to visualize climate change texts," *Inf. Vis.*, vol. 23, no. 3, pp. 217–238, Jul. 2024, doi: 10.1177/14738716241236188.
- [92] V. Ibrahim, J. Abu Bakar, N. H. Harun, and A. F. Abdulateef, "A Word Cloud Model based on Hate Speech in an Online Social Media Environment," *Baghdad Sci. J.*, vol. 18, no. 2, Jun. 2021, doi: 10.21123/bsj.2021.18.2(Suppl.).0937.
- [93] O. Ozyurt, H. Özköse, and A. Ayaz, "Evaluating the latest trends of Industry 4.0 based on LDA topic model," *J. Supercomput.*, vol. 80, no. 13, pp. 19003–19030, Sep. 2024, doi: 10.1007/s11227-024-06247-x.
- [94] S. C. Joonwoo Baik, Sehwan Chung, "Evaluation of Topic Modeling Performance for Overseas Construction Market Analysis Using LDA and BERTopic on News Articles," *KSCE J. Civ. Eng.*, vol. 43, no. 811–819, 2023, doi: <https://doi.org/10.12652/Ksce.2023.43.6.0811>.
- [95] A. J. Dhruv, R. Patel, and N. Doshi, "Python: The Most Advanced Programming Language for Computer Science Applications," 2021. doi: 10.5220/0010307902920299.
- [96] Python, "General Python FAQ," python. [Online]. Available: <https://docs.python.org/3/faq/general.html>
- [97] Python, "Install packages in a virtual environment using pip and venv," python. [Online]. Available: <https://packaging.python.org/en/latest/guides/installing-using-pip-and-virtual-environments/>
- [98] Numpy, "NumPy documentation," numpy. [Online]. Available: <https://numpy.org/doc/stable/>
- [99] A. Lawrence, "Apa Itu Visual Studio Code? Pengertian, Kelebihan, dan Cara Installnya," 2022.
- [100] A. Del Sole, *Visual Studio Code Distilled: Evolved Code Editing for Windows, macOS, and Linux, Second Edition*. 2021. doi: 10.1007/978-1-4842-6901-5.
- [101] Google, "Welcome to Colab!," Google Colab Documentation. [Online]. Available: <https://colab.research.google.com/>
- [102] J. Praveen Gujjar, H. R. Prasanna Kumar, and N. N. Chiplunkar, "Image classification and prediction using transfer learning in colab notebook," *Glob. Transitions Proc.*, vol. 2, no. 2, pp. 382–385, Nov. 2021, doi:

10.1016/j.gltp.2021.08.068.

- [103] S. Umair and Q. Haseeb, “A Comparative Study of Data Mining Process Models (KDD, CRISP-DM and SEMMA),” *Int. J. Innov. Sci. Res.*, 2014.
- [104] U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, “From data mining to knowledge discovery in databases,” *AI Mag.*, 1996.
- [105] SAS, “Semma,” *Sas*, 2006.
- [106] C. Shearer *et al.*, “The CRISP-DM model: The New Blueprint for Data Mining,” *J. Data Warehous.*, 2000.
- [107] L. Sai Lakshmi, D. L. Sumithra, K. Jhansi, and J. Vandana, “Web Scraping And Data Analysis For Online Shopping With Selenium,” *IARJSET*, vol. 11, no. 3, Mar. 2024, doi: 10.17148/IARJSET.2024.11329.
- [108] A. Abodayeh, R. Hejazi, W. Najjar, L. Shihadeh, and R. Latif, “Web Scraping for Data Analytics: A BeautifulSoup Implementation,” in *2023 Sixth International Conference of Women in Data Science at Prince Sultan University (WiDS PSU)*, IEEE, Mar. 2023, pp. 65–69. doi: 10.1109/WiDS-PSU57071.2023.00025.
- [109] M. I. Firdian, E. Darwiyanto, and M. Adrian, “Web Scraping with HTML DOM Method for Website News API creation,” *JUPI (Jurnal Ilm. Penelit. dan Pembelajaran Inform.*, vol. 7, no. 4, pp. 1211–1219, Nov. 2022, doi: 10.29100/jupi.v7i4.3235.
- [110] A. Nazarkar, H. Kuchulakanti, C. S. Paidimarry, and S. Kulkarni, “Impact of Various Data Splitting Ratios on the Performance of Machine Learning Models in the Classification of Lung Cancer,” 2023, pp. 96–104. doi: 10.2991/978-94-6463-252-1_12.
- [111] K. Okoye and S. Hosseini, “Understanding Dependent and Independent Variables in Research Experiments and Hypothesis Testing,” in *R Programming*, Singapore: Springer Nature Singapore, 2024, pp. 99–107. doi: 10.1007/978-981-97-3385-9_5.
- [112] D. Podell *et al.*, “SDXL: Improving Latent Diffusion Models for High-Resolution Image Synthesis,” Jul. 2023, [Online]. Available: <http://arxiv.org/abs/2307.01952>
- [113] F. Pedregosa *et al.*, “Scikit-learn: Machine Learning in Python,” Jun. 2018, [Online]. Available: <http://arxiv.org/abs/1201.0490>