Daftar Pustaka

Ainapure, B. S., Pise, R. N., Reddy, P., Appasani, B., Khan, M. S., Bizon, N., & Srinivasulu, A. (2023). Sentiment *Analysis* of COVID-19 Tweets Using Deep *Learning* and Lexicon-*Based* Approaches. *Sustainability (Switzerland)*, *15*(3). https://doi.org/10.3390/su15032573

Alabduallah, B., Maray, M., Alruwais, N., Alabdan, R., Darem, A. A., Alallah, F. S., Alsini, R., & Yafoz, A. (2024). *Class ImBalanced* data *Handling* with cyberattack *Class*ification using *Hybrid* Salp Swarm Algorithm with deep *learning* approach. *Alexandria Engineering Journal*, *106*(July), 654–663. https://doi.org/10.1016/j.aej.2024.08.061

Ashraf Al-Quran, and M. U. R. Q. H. H. Z. R. N. H. K. F. A.-S. (2024). Triangular Fuzzy Merec (TFMEREC) And Its *Application*s In Multi Criteria *Decision Making*. *Journal of Fuzzy Extension and Applications*, *5*(Mcdm), 1–27.

Azmi, P. A. R., Yusoff, M., & Sallehud-din, M. T. M. (2025). Improving transformer failure *Class*ification *on ImBalanced* DGA data using data-level *Techniques* and *machine learning*. *Energy Reports*, *13*(November 2024), 264–277. https://doi.org/10.1016/j.egyr.2024.12.006

Badan Pusat Statistik Indonesia. (2024). Statistik Wisatawan Nasional Outbound Tourism Statistics 2023. In *Badan Pusat Statistik/BPS-Statistics Indonesia* (Vol. 3).

Beygelzimer, A., Street, E., & Langford, J. (2006). Cover Trees for Nearest Neighbor. *Proceedings of the 23rd International Con- Ference on Machine Learning*, 97–104.

Bhowmik, N. R., Arifuzzaman, M., & Mondal, M. R. H. (2022). Sentiment *analysis on* Bangla text using extended lexicon dictionary and deep *learning* algorithms. *Array*, *13*, 100123. https://doi.org/10.1016/j.array.2021.100123

SEKOLAH PASCASARJANA

Cahyawijaya, S., Winata, G. I., Wilie, B., Vincentio, K., Li, X., Kuncoro, A., Ruder, S., Lim, Z. Y., Bahar, S., Khodra, M. L., Purwarianti, A., & Fung, P. (2021). IndoNLG : Benchmark and Resources for Evaluating Indonesian Natural Language *Generation*. *Proceedings Ofthe 2021 Conference on Empirical Methods in Natural Language Processing*, 8875–8898.

Çalişir, D., & Doğantekin, E. (2011). An automatic diabetes diagnosis *system based on* LDA-Wavelet *Support* Vector *Machine Classifier*. *Expert Systems with Applications*, *38*(7), 8311–8315. https://doi.org/10.1016/j.eswa.2011.01.017

Cao, K., Wei, C., Gaidon, A., Arechiga, N., & Ma, T. (2019). *Learning ImBalanced Datasets* with Label-*Distribution*-Aware Margin *Loss*. *Neural Information Processing Systems, NeurIPS*, 1–18.

Chatterjee, S., & Byun, Y. C. (2025). Leveraging generative adversarial *networks* for data augmentation to improve fault detection in wind turbines with *ImBalanced* data. *Results in Engineering*, *25*(January). https://doi.org/10.1016/j.rineng.2025.103991

Chatterjee, S., Chakrabarti, K., Garain, A., & Schwenker, F. (2021). A Sentiment *Analysis Dataset* for Movie *Recommend*ation. *Applied Sciences (Switzerland)*, *11*, 1–24.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: *Synthetic Minority Over-sampling Technique*. *Journal of Artificial Intelligence Research*, *16*(1), 321–357. https://doi.org/10.46880/jmika.vol4no1.pp67-72

Dash, P. K., Rekha Pattnaik, S., N.V.D.V. Prasad, E., & Bisoi, R. (2023). Detection and *Class*ification of DC and feeder faults in DC microgrid using new morphological operators with multi *Class* AdaBoost algorithm. *Applied Energy*, *340*(December 2022), 121013. https://doi.org/10.1016/j.apenergy.2023.121013

Devlin, J., Chang, M.-W., Lee, K., Google, K. T., & Language, A. I. (2018). BERT: Pre-*training* of Deep Bidirectional Transformers for Language Understanding. *Naacl-Hlt 2019*, *Mlm*, 4171–4186. https://aclanthology.org/N19-1423.pdf

DinarStandard. (2023). State of the Global Islamic Economy Report. *DinarStandard*, 65–66. https://haladinar.io/hdn/doc/report2018.pdf

Farou, Z., Wang, Y., & Horváth, T. (2024). Cluster-*based oversampling* with area extraction from representative points for *Class* imbalance *learning*. *Intelligent Systems with Applications*, *22*(August 2023), 200357. https://doi.org/10.1016/j.iswa.2024.200357

Fernández, A., García, S., Galar, M., Prati, R. C., Krawczyk, B., & Herrera, F. (2018). *Learning* from *ImBalanced* Data Sets. In *Learning from ImBalanced Data Sets*. https://doi.org/10.1007/978-3-319-98074-4_5

Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., & Sun, J. (2024). *Retrieval-Augmented Generation* for *Large* Language Models : A Survey. *Computation and Language*, 1–21.

*Good*fel*low*, I. J., Pouget-abadie, J., Mirza, M., Xu, B., Warde-farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative Adversarial Nets. *Advances in Neural Information Processing Systems 27 (NIPS 2014)*, 1–9.

Góra, G., & Skowron, A. (2025). RIONIDA: A novel algorithm for *ImBalanced* data combining instance-*based learning* and rule induction. *Information Sciences*, *708*(February). https://doi.org/10.1016/j.ins.2025.122015

Guu, K., Lee, K., Tung, Z., Pasupat, P., & Chang, M. (2020). REALM : Retrieval-Augmented Language Model Pre-*Training*. *Computation and Language*.

Han, H., Wang, W. Y., & Mao, B. H. (2005). Borderline-SMOTE: A new *over-sampling* method in *ImBalanced* data sets *learning*. *Lecture Notes in Computer Science*, *3644*(PART I), 878–887. https://doi.org/10.1007/11538059_91

He, H., Bai, Y., Garcia, E., & Li, S. (2008). ADASYN: *Adaptive Synthetic Sampling* approach for *ImBalanced learning*. In IEEE International Joint Conference *on* Neural *Networks*, 2008. *IJCNN 2008.(IEEE World Congress on Computational Intelligence) (Pp. 1322– 1328)*, *3*, 1322– 1328.

He, P., Liu, X., Gao, J., & Chen, W. (2021). DEBERTA: DECODING-ENHANCED BERT WITH DIS- ENTANGLED *ATTENTION*. *ICLR 2021*, 1–23.

He, W., Farrahi, K., Chen, B., Peng, B., & Villavicencio, A. (2024a). Representation *Transfer* and data cleaning in multi-views for text simplification. *Pattern Recognition Letters*, *177*(November 2023), 40–46. https://doi.org/10.1016/j.patrec.2023.11.011

He, W., Farrahi, K., Chen, B., Peng, B., & Villavicencio, A. (2024b). Representation *Transfer* and data cleaning in multi-views for text simplification. *Pattern Recognition Letters*, *177*(July 2023), 40–46. https://doi.org/10.1016/j.patrec.2023.11.011

He, Y. X., Liu, D. X., Lyu, S. H., Qian, C., & Zhou, Z. H. (2024). Multi-*Class* imbalance problem: A multi-*Objective* solution. *Information Sciences*, *680*(July), 121156. https://doi.org/10.1016/j.ins.2024.121156

Hou, D., Chen, J. H., Cheng, R., Hu, X., & Shi, P. (2024). A bearing remaining life prediction method under variable ope*Rating* conditions *based on* cross-transformer fusioning segmented data cleaning. *Reliability Engineering and System Safety*, *245*(January), 110021. https://doi.org/10.1016/j.ress.2024.110021

Japkowicz, N., & Stephen, S. (2002). The *Class* imbalance problem : A *system*atic study. *Intelligent Data Analysis*, *6*, 429–449.

Johnson, J., Douze, M., & J´egou, H. (2019). Billion-scale *Similarity* search with GPUs. *IEEE Transactions on Big Data*, *7*(3).

Johnson, J. M., & Khoshgoftaar, T. M. (2019). Deep *Learning* and Data *Sampling* with *ImBalanced* Big Data. *2019 IEEE 20th International Conference on Information Reuse and Integration for Data Science (IRI)*, *July*. https://doi.org/10.1109/IRI.2019.00038

Kamari, M. S. B. M., Rodzi, Z. B. M., Al-Obaidi, R. H., Al-Sharqi, F., Al-Quran, A., & Shlaka, R. A. (2025). Deciphering the Geometric Bonferroni Mean Operator in Pythagorean Neutrosophic Sets *Framework*. *Neutrosophic Sets and Systems*, *75*, 139–161. https://doi.org/10.5281/zenodo.13932052

Kementerian Koperasi dan UKM. (2022). Paten Program Adaptasi Dan Transformasi Ekonomi Nasional. In *Tabloid Kementerian Koperasi dan UKM*.

SEKOLAH PASCASARJANA

https://satudata.kemenkopukm.go.id/file/arsip/752ded2a-b156-4032-b155-21c874f06931.pdf?type=download

Koto, F., Lau, J. H., & Baldwin, T. (2021). INDOBERTWEET: A Pretrained Language Model for Indonesian Twitter with Effective Domain-Specific Vocabulary Initialization. *Proceedings Ofthe 2021 Conference on Empirical Methods in Natural Language Processing*, 10660–10668.

Koto, F., Rahimi, A., Lau, J. H., & Baldwin, T. (2020). IndoLEM and IndoBERT: A Benchmark *Dataset* and Pre-trained Language Model for Indonesian NLP. *Computation and Language*.

Krawczyk, B. (2018). *Learning* from *ImBalanced* data : Open challenges and future directions *Learning* from *ImBalanced* data : open challenges and future directions. *Progress in Artificial Intelligence, April 2016*. https://doi.org/10.1007/s13748-016-0094-0

Kubat, M., & Matwin, S. (1997). *Addressing* the Curse of *ImBalanced training* Sets: One-Sided Selection. *International Conference on Machine Learning*, *4*.

Kurniawati, D. A., & Rochman, M. A. (2023). Two stages of halal food *distribution* model for perishable food products. *International Journal of Production Management and Engineering*, *11*(2), 147–166. https://doi.org/10.4995/ijpme.2023.18233

Lan, Z., Chen, M., *Good*man, S., Gimpel, K., Sharma, P., & Soricut, R. (2020). ALBERT: A Lite Bert For Self-Supervised *Learning* Of Language Representations. *ICLR 2020*, 1–17.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2021). *Retrieval-Augmented Generation* for *Knowledge*-Intensive NLP Tasks. *Computation and Language*.

Li, Y., Cao, J., Xu, Y., Zhu, L., & Dong, Z. Y. (2024). Deep *learning based on* Transformer architecture for power *system* short-term voltage stability *Assessment* with *Class* imbalance. *Renewable and Sustainable Energy Reviews*, *189*(PB), 113913. https://doi.org/10.1016/j.rser.2023.113913

SEKOLAH PASCASARJANA

Liu, Y., Li, B., Yang, S., & Li, Z. (2024). *Handling* missing values and *ImBalanced Class*es in *machine learning* to predict consumer preference: Demonstrations and *Comparison*s to prominent methods. *Expert Systems with Applications*, *237*(PC), 121694. https://doi.org/10.1016/j.eswa.2023.121694

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). RoBERTa: A *Robust*ly Optimized BERT Pre*training* Approach. *Computation and Language*, *1*.

LPPOM. (2024). Membangun Ekosistem Halal, Menuju Indonesia Sebagai Pusat Halal Dunia. In *Jurnal Halal*.

Mahardika, Ri., Yuliendi, R. R., & Yandri, Y. (2022). Sistem Rekomendasi Pemilihan Lokasi Kuliner Halal Area Kota Pekanbaru Menggunakan Metode *Weight*ed Product. *JEKIN - Jurnal Teknik Informatika*, *2*(3), 175–185. https://doi.org/10.58794/jekin.v3i1.419

Mustapha, M., Zineddine, M., Kaufman, E., Friedman, L., Gmira, M., Majikumna, K. U., & Alaoui, A. E. H. (2025). A *hybrid machine learning* approach for *ImBalanced* irrigation water *quality Class*ification. *Desalination and Water Treatment*, *321*(November 2024), 100910. https://doi.org/10.1016/j.dwt.2024.100910

Penny, R. N. (1995). Model Assisted Survey *Sampling* . *Journal of the Royal Statistical Society Series D (The Statistician)*, *September*. https://doi.org/10.2307/2348627

Rahman, T., Agustin, F. E. M., & Rozy, N. F. (2019). Normalization of Unstructured Indonesian Tweet Text for Presidential Candidates Sentiment *Analysis*. *7th International Conference on Cyber and IT Service Management, CITSM*, 9–14. https://doi.org/10.1109/CITSM47753.2019.8965324

Ramos, M. M., Bijani, R., Santos, F. V., Lupinacci, W. M., & Freire, A. F. M. (2023). *Analysis* of alternative strategies *Applied* to Naïve-Bayes *Classifier* into the recognition of electrofacies: *Application* in well-log data at Recôncavo Basin, North-East Brazil. *Geoenergy Science and Engineering*, *227*(May 2022), 211889. https://doi.org/10.1016/j.geoen.2023.211889

SEKOLAH PASCASARJANA

Rosline Mary, A., & Kavitha, P. (2023). Automated Diabetic Retinopathy detection and *Class*ification using stochastic coordinate descent deep *learning* architectures. *Materials Today: Proceedings*, *80*, 3333–3345. https://doi.org/10.1016/j.matpr.2021.07.250

Sakib, N., Paul, T., Anwari, N., & Hadiuzzaman, M. (2024). *Ensemble-based* model to investigate *factor*s influencing road crash fatality for *ImBalanced* data. *Transportation Engineering*, *18*(October), 100284. https://doi.org/10.1016/j.treng.2024.100284

Silverman, B. W. (1986). Density Estimation for Statistics and Data *Analysis*. *Monographs on Statistics and Applied Probability*.

Sokolova, M., & Lapalme, G. (2009). A *system*atic *analysis* of performance *measure*s for *Class*ification tasks. *Information Processing and Management*, *45*(4), 427–437. https://doi.org/10.1016/j.ipm.2009.03.002

Tomek, I. (1976). Tomek Link: Two Modifications of CNN. *IEEE Trans. Systems, Man and Cybernetics*, 769–772. https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4309452

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention* Is All You Need. *31st Conference on Neural Information Processing Systems (NIPS 2017), Nips*.

Weiss, G. M., & Provost, F. (2003). *Learning* when *training* data are costly: The effect of *Class distribution on* tree induction. *Journal of Artificial Intelligence Research*, *19*, 315–354. https://doi.org/10.1613/jair.1199

Wilie, B., Vincentio, K., Winata, G. I., Cahyawijaya, S., Li, X., Lim, Z. Y., Soleman, S., Mahendra, R., Fung, P., Bahar, S., & Purwarianti, A. (2020). IndoNLU : Benchmark and Resources for Evaluating Indonesian Natural Language Understanding. *Proceedings Ofthe 1st Conference Ofthe Asia-Pacific Chapter Ofthe Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*, 843–857.

Wojarnik, G. (2021). Sentiment *analysis* as a *factor* included in the forecasts of price changes in the stock exchange. *Procedia Computer Science*, *192*, 3176–

SEKOLAH PASCASARJANA

3183. https://doi.org/10.1016/j.procs.2021.09.090

Yang, F., Li, X., Wang, B., Zhang, T., Yu, X., Yi, X., & Zhu, R. (2025). MMSeg: A novel multi-task *learning framework* for *Class* imbalance and label scarcity in medical image segmentation. *Knowledge-Based Systems*, *309*(November 2024), 112835. https://doi.org/10.1016/j.knosys.2024.112835

Yang, Y., Zha, K., Hao, Y. C., & Dina, W. (2021). Delving into Deep *ImBalanced* Regression. *Machine Learning*.

Yu, Y., Liu, L., Chang, Z., Li, Y., & Shi, K. (2024). Detecting Forest Fires in Southwest China From Remote Sensing Nighttime Lights Using the Random Forest *Class*ification Model. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *17*, 10759–10769. https://doi.org/10.1109/JSTARS.2024.3410172

Zhang, J., & Mani, I. (2003). kNN Approach to Un*Balanced* Data *Distribution*s: A Case Study involvoing Information Extraction. *Learning from ImBalanced Datasets II, ICML*.

Zheng, M., Wang, F., Hu, X., Hu, L., Yu, Q., & Zheng, X. (2025). UFIDSF: An *undersampling* approach *based on Feature Importance* and double side filter for *ImBalanced* data *Class*ification. *Future Generation Computer Systems*, *167*(February). https://doi.org/10.1016/j.future.2025.107750

Zhong, K., Tan, X., Liu, S., Lu, Z., Hou, X., & Wang, Q. (2025). Prediction of slope failure *Probability based on machine learning* with genetic-ADASYN algorithm. *Engineering Geology*, *346*(December 2024). https://doi.org/10.1016/j.enggeo.2024.107885

SEKOLAH PASCASARJANA