

## DAFTAR PUSTAKA

- Afrianto, R.B., Kurniawati, L.Y., 2013. Kategorisasi Dokumen Teks Secara Multi Label Menggunakan Fuzzy C-Means dan K-Nearest Neighbors Pada Artikel Berbahasa Indonesia. *JUTI Jurnal Ilmiah Teknologi Informasi* .  
<https://doi.org/10.12962/j24068535.v11i1.a17>
- Arief, R., Iriawan, N.A., Lawi, D.A., Copresponent, E., 2021. Klasifikasi Audio Ucapan Emosional Menggunakan Model LSTM.
- Asha, Kumar, S., 2021. Credit card fraud detection using artificial neural network. *Global Transitions Proceedings* 2, 35–41.  
<https://doi.org/10.1016/j.gltp.2021.01.006>
- Batista, G., Prati, R., Monard, M.-C., 2004. A Study of the Behavior of Several Methods for Balancing machine Learning Training Data. *SIGKDD Explorations* 6, 20–29. <https://doi.org/10.1145/1007730.1007735>
- Chen, L., 2009. Curse of Dimensionality, in: LIU LING and ÖZSU, M.T. (Ed.), *Encyclopedia of Database Systems*. Springer US, Boston, MA, pp. 545–546.  
[https://doi.org/10.1007/978-0-387-39940-9\\_133](https://doi.org/10.1007/978-0-387-39940-9_133)
- Cho, E., Chang, T.-W., Hwang, G., 2022. Data Preprocessing Combination to Improve the Performance of Quality Classification in the Manufacturing Process. *Electronics (Basel)* 11, 477.  
<https://doi.org/10.3390/electronics11030477>
- Choirunnisa, S., Lianto, J., 2018. *Hybrid Method of Undersampling and Oversampling for Handling Imbalanced Data; Hybrid Method of Undersampling and Oversampling for Handling Imbalanced Data*.
- Domingos, P., Pazzani, M., 1997. *On the Optimality of the Simple Bayesian Classifier under Zero-One Loss*. Kluwer Academic Publishers.
- Er, M.J., Venkatesan, R., Ning, W., 2017. An Online Universal Classifier for Binary, Multi-class, and Multi-label Classification, in: 2016 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2016 - Conference Proceedings. Institute of Electrical and Electronics Engineers Inc., pp. 3701–3706. <https://doi.org/10.1109/SMC.2016.7844809>

- Fahrudin, T., Buliali, J.L., Fatichah, C., 2019. Enhancing the performance of smote algorithm by using attribute weighting scheme and new selective sampling method for imbalanced data set. *International Journal of Innovative Computing, Information and Control* 15, 423–444. <https://doi.org/10.24507/ijicic.15.02.423>
- Gameng, H.A., Gerardo, B.B., Media, R.P., 2019a. Modified Adaptive Synthetic SMOTE to Improve Classification Performance in Imbalanced Datasets. *International Conference on Engineering Technologies and Applied Sciences (ICETAS)*.
- Gameng, H.A., Gerardo, B.B., Medina, R.P., 2019b. Modified Adaptive Synthetic SMOTE to Improve Classification Performance in Imbalanced Datasets.
- Gao, Q., Jin, X., Xia, E., Wu, X., Gu, L., Yan, H., Xia, Y., Li, S., 2020. Identification of Orphan Genes in Unbalanced Datasets Based on Ensemble Learning. *Front Genet* 11. <https://doi.org/10.3389/fgene.2020.00820>
- Gao, X., Li, G., 2020. A KNN Model Based on Manhattan Distance to Identify the SNARE Proteins. *IEEE Access* 8, 112922–112931. <https://doi.org/10.1109/ACCESS.2020.3003086>
- Grandini, M., Bagli, E., Visani, G., 2020. Metrics for Multi-Class Classification: an Overview.
- Guan, H., Zhang, Y., Xian, M., Cheng, H.D., Tang, X., 2021. SMOTE-WENN: Solving class imbalance and small sample problems by oversampling and distance scaling. *Applied Intelligence* 51, 1394–1409. <https://doi.org/10.1007/s10489-020-01852-8>
- Hairani, H., Saputro, K.E., Fadli, S., 2020. K-means-SMOTE for handling class imbalance in the classification of diabetes with C4.5, SVM, and Naive Bayes. *Jurnal Teknologi dan Sistem Komputer* 8, 89–93. <https://doi.org/10.14710/jtsiskom.8.2.2020.89-93>
- Heryadi, Y., Wahyono, T., 2021. *Dasar-Dasar Deep Learning dan Implementasinya*. Gava Media, Yogyakarta.
- Heryadi, Y., Wahyono, T., 2020. *Machine Learning: Konsep dan Implementasi*. Gava Media, Yogyakarta.

- Hughes, G., 1968. On the mean accuracy of statistical pattern recognizers. *IEEE Trans Inf Theory* 14, 55–63. <https://doi.org/10.1109/TIT.1968.1054102>
- Iswanto, I., Tulus, T., Sihombing, P., 2021. Comparison of Distance Models on K-Nearest Neighbor Algorithm in Stroke Disease Detection. *Applied Technology and Computing Science Journal* 4, 63–68. <https://doi.org/10.33086/atcsj.v4i1.2097>
- Jia, W., Sun, M., Lian, J., Hou, S., 2022. Feature dimensionality reduction: a review. *Complex & Intelligent Systems* 8, 2663–2693. <https://doi.org/10.1007/s40747-021-00637-x>
- Johnson, J.M., Khoshgoftaar, T.M., 2019. Survey on Deep Learning with Class Imbalance. *J Big Data* 6. <https://doi.org/10.1186/s40537-019-0192-5>
- K. Wisaeng, 2013. A Comparison of Different Classification Techniques for Bank Direct Marketing.
- Kumar, V., Lalotra, G.S., Sasikala, P., Rajput, D.S., Kaluri, R., Lakshmana, K., Shorfuzzaman, M., Alsufyani, A., Uddin, M., 2022. Addressing Binary Classification over Class Imbalanced Clinical Datasets Using Computationally Intelligent Techniques. *Healthcare* 10, 1293. <https://doi.org/10.3390/healthcare10071293>
- Larose, D.T., 2005. *Discovering Knowledge in Data: An Introduction to Data Mining*.
- Lones, M.A., 2021. How to avoid machine learning pitfalls: a guide for academic researchers.
- Lu, T., Huang, Y., Zhao, W., Zhang, J., 2019. The Metering Automation System based Intrusion Detection Using Random Forest Classifier with SMOTE+ENN.
- M. Gorade, S., Deo, A., Purohit, P., 2017. A Study of Some Data Mining Classification Techniques. *International Research Journal of Engineering and Technology*.
- Maldonado, S., López, J., Vairetti, C., 2019a. An alternative SMOTE oversampling strategy for high-dimensional datasets. *Applied Soft Computing Journal* 76, 380–389. <https://doi.org/10.1016/j.asoc.2018.12.024>

- Maldonado, S., López, J., Vairetti, C., 2019b. An alternative SMOTE oversampling strategy for high-dimensional datasets. *Applied Soft Computing Journal* 76, 380–389. <https://doi.org/10.1016/j.asoc.2018.12.024>
- Milosevic, M.S., Ciric, V.M., 2022. Extreme minority class detection in imbalanced data for network intrusion. *Comput Secur* 123. <https://doi.org/10.1016/j.cose.2022.102940>
- Pal, A.K., Mondal, P.K., Ghosh, A.K., 2016. High dimensional nearest neighbor classification based on mean absolute differences of inter-point distances. *Pattern Recognit Lett* 74, 1–8. <https://doi.org/10.1016/j.patrec.2016.01.018>
- Pradipta, G.A., Wardoyo, R., Musdholifah, A., Sanjaya, I.N.H., 2021. Radius-SMOTE: A New Oversampling Technique of Minority Samples Based on Radius Distance for Learning from Imbalanced Data. *IEEE Access* 9, 74763–74777. <https://doi.org/10.1109/ACCESS.2021.3080316>
- Shi, H., Xu, M., 2018. A Data Classification Method Using Genetic Algorithm and K-Means Algorithm with Optimizing Initial Cluster Center. *IEEE*.
- Sutoyo, E., Almaarif, A., 2017a. Educational Data Mining untuk Prediksi Kelulusan Mahasiswa Menggunakan Algoritme Naïve Bayes Classifier. *Rekayasa Sistem dan Teknologi Informasi* 1, 95–101.
- Sutoyo, E., Almaarif, A., 2017b. Educational Data Mining untuk Prediksi Kelulusan Mahasiswa Menggunakan Algoritme Naïve Bayes Classifier. *masa berlaku* 1, 95–101.
- Tallo, T.E., Musdholifah, A., 2018. The Implementation of Genetic Algorithm in Smote (Synthetic Minority Oversampling Technique) for Handling Imbalanced Dataset Problem.
- Tanouz, D., Subramanian, R.R., Eswar, D., Reddy, G.V.P., Kumar, A.R., Praneeth, C.H.V.N.M., 2021. Credit Card Fraud Detection Using Machine Learning, in: *Proceedings - 5th International Conference on Intelligent Computing and Control Systems, ICICCS 2021*. Institute of Electrical and Electronics Engineers Inc., pp. 967–972. <https://doi.org/10.1109/ICICCS51141.2021.9432308>

- Tschannen, M., Kramer, T., Marti, G., Heinzmann, M., Wiatowski, T., 2016. Heart sound classification using deep structured features; Heart sound classification using deep structured features, 2016 Computing in Cardiology Conference (CinC). <https://doi.org/10.23919/CIC.2016.7868805>
- Turlapati, V.P.K., Prusty, M.R., 2020. Outlier-SMOTE: A refined oversampling technique for improved detection of COVID-19. *Intell Based Med* 3–4. <https://doi.org/10.1016/j.ibmed.2020.100023>
- Xia, S., Xiong, Z., Luo, Y., Weixu, Zhang, G., 2015. Effectiveness of the Euclidean distance in high dimensional spaces. *Optik (Stuttg)* 126, 5614–5619. <https://doi.org/10.1016/j.ijleo.2015.09.093>
- Xu, Z., Shen, D., Nie, T., Kou, Y., 2020. A *Hybrid* Sampling Algorithm Combining M-SMOTE and ENN Based on Random Forest for Medical Imbalanced Data. *J Biomed Inform* 107. <https://doi.org/10.1016/j.jbi.2020.103465>