

## References

- [1] G. Qadir, "Yeast a magical microorganism in the wastewater treatment," ~ 1498 ~ *J. Pharmacogn. Phytochem.*, vol. 8, no. 4, pp. 1498–1500, 2019.
- [2] R. Siringoringo, "Klasifikasi data tidak Seimbang menggunakan algoritma SMOTE dan k-nearest neighbor," *J. ISD*, vol. 3, no. 1, pp. 44–49, 2018.
- [3] T. M. Alam *et al.*, "An investigation of credit card default prediction in the imbalanced datasets," *IEEE Access*, vol. 8, pp. 201173–201198, 2020, doi: 10.1109/ACCESS.2020.3033784.
- [4] H. Patel, D. S. Rajput, O. P. Stan, and L. C. Miclea, "A new fuzzy adaptive algorithm to classify imbalanced data," *Comput. Mater. Contin.*, vol. 70, no. 1, pp. 73–89, 2021, doi: 10.32604/cmc.2022.017114.
- [5] L. Cao and H. Shen, "Imbalanced data classification using improved clustering algorithm and under-sampling method," *Proc. - 2019 20th Int. Conf. Parallel Distrib. Comput. Appl. Technol. PDCAT 2019*, pp. 358–363, 2019, doi: 10.1109/PDCAT46702.2019.00071.
- [6] M. C. Untoro, M. Praseptiawan, M. Widianingsih, I. F. Ashari, A. Afriansyah, and Oktafianto, "Evaluation of Decision Tree, K-NN, Naive Bayes and SVM with MWMOTE on UCI Dataset," *J. Phys. Conf. Ser.*, vol. 1477, no. 3, 2020, doi: 10.1088/1742-6596/1477/3/032005.
- [7] B. Huang, Y. Zhu, Z. Wang, and Z. Fang, "Imbalanced Data Classification Algorithm Based on Clustering and SVM," *J. Circuits, Syst. Comput.*, vol. 30, no. 2, 2021, doi: 10.1142/S0218126621500365.
- [8] L. Yu, R. Zhou, L. Tang, and R. Chen, "A DBN-based resampling SVM ensemble learning paradigm for credit classification with imbalanced data," *Appl. Soft Comput. J.*, vol. 69, pp. 192–202, 2018, doi: 10.1016/j.asoc.2018.04.049.
- [9] A. Wasilewska and E. Menasalvas, "Data preprocessing and data mining as generalization," *Stud. Comput. Intell.*, vol. 118, no. Definition 1, pp. 469–484, 2008, doi: 10.1007/978-3-540-78488-3\_27.
- [10] R. S. Somasundaram and R. Nedunchezian, "Evaluation of Three Simple Imputation Methods for Enhancing Preprocessing of Data with Missing Values," *Int. J. Comput. Appl.*, vol. 21, no. 10, pp. 14–19, 2011, doi: 10.5120/2619-3544.
- [11] A. K. Elmagarmid, P. G. Ipeirotis, and V. S. Verykios, "Duplicate record detection: A survey," *IEEE Trans. Knowl. Data Eng.*, vol. 19, no. 1, pp. 1–16, 2007, doi: 10.1109/TKDE.2007.250581.
- [12] P. Purwono, A. Wirasto, and K. Nisa, "Komparasi Algoritma Machine Learning Untuk Klasifikasi Kelompok Obat," *Sisfotenika*, vol. 11, no. 2, p. 196, 2021.
- [13] D. Deng, "DBSCAN Clustering Algorithm Based on Density," *Proc. - 2020 7th Int. Forum Electr. Eng. Autom. IFEEA 2020*, pp. 949–953, 2020, doi: 10.1109/IFEEA51475.2020.00199.
- [14] B. Mirzaei, B. Nikpour, and H. Nezamabadi-Pour, "An under-sampling technique for imbalanced data classification based on DBSCAN algorithm," *8th Iran. Jt. Congr. Fuzzy Intell. Syst. CFIS 2020*, pp. 21–26, 2020, doi: 10.1109/CFIS49607.2020.9238718.
- [15] F. E. Botchey, Z. Qin, and K. Hughes-Lartey, "Mobile money fraud prediction-A cross-case analysis on the efficiency of support vector machines, gradient boosted decision trees, and Naïve Bayes algorithms," *Inf.*, vol. 11, no. 8, 2020, doi: 10.3390/INFO11080383.
- [16] S. Huang, C. A. I. Nianguang, P. Penzuti Pacheco, S. Narandes, Y. Wang, and X. U. Wayne, "Applications of support vector machine (SVM) learning in cancer genomics," *Cancer Genomics and Proteomics*, vol. 15, no. 1, pp. 41–51, 2018, doi: 10.21873/cgp.20063.
- [17] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: Applications, challenges and trends," *Neurocomputing*, vol. 408, no. xxxx, pp. 189–215, 2020, doi: 10.1016/j.neucom.2019.10.118.
- [18] A. Fahmi Sabani, Adiwijaya, and W. Astuti, "Analisis Sentimen Review Film pada Website Rotten Tomatoes Menggunakan Metode SVM Dengan Mengimplementasikan Fitur Extraction Word2Vec," *e-Proceeding Eng.*, vol. 9, no. 3, p. 1800, 2022.
- [19] N. Tri Romadloni, I. Santoso, S. Budilaksono, and M. Ilmu Komputer STMIK Nusa Mandiri Jakarta, "Perbandingan Metode Naive Bayes, Knn Dan Decision Tree Terhadap Analisis Sentimen Transportasi Krl Commuter Line," *J. IKRA-ITH Inform.*, vol. 3, no. 2, pp. 1–9, 2019.
- [20] F. Ratnawati, "Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter," *INOVTEK Polbeng - Seri Inform.*, vol. 3, no. 1, p. 50, 2018, doi: 10.35314/isi.v3i1.335.
- [21] N. Hendrastuty *et al.*, "Analisis Sentimen Masyarakat Terhadap Program Kartu Prakerja Pada Twitter Dengan Metode Support Vector Machine," *J. Inform. J. Pengemb. IT*, vol. 6, no. 3, pp.

- 150–155, 2021.
- [22] R. Mehmood and A. Selwal, *Fingerprint biometric template security schemes: Attacks and countermeasures*, vol. 597. 2020.
- [23] V. S. Spelmen and R. Porkodi, “A Review on Handling Imbalanced Data,” *Proc. 2018 Int. Conf. Curr. Trends Towar. Converging Technol. ICCTCT 2018*, pp. 1–11, 2018, doi: 10.1109/ICCTCT.2018.8551020.