

DAFTAR PUSTAKA

- [1] W. Gustiawan, B. K. Dwi Putri, T. Azra, and M. Permata Sari, "Internship programme and work readiness among vocational students," *Issues and Perspectives in Business and Social Sciences*, vol. 5, no. 1, pp. 15–26, Jan. 2025, doi: 10.33093/ipbss.2025.5.1.2.
- [2] F. Furtasan Ali Yusuf and Basrowi, "The Effectiveness of the Internship Program for Vocational High School Students Using the CIPP Method," *JTP - Jurnal Teknologi Pendidikan*, vol. 25, no. 1, pp. 15–28, Apr. 2023, doi: 10.21009/jtp.v25i1.33553.
- [3] R. Pal, S. Shaikh, S. Satpute, and S. Bhagwat, "Resume Classification using various Machine Learning Algorithms," *ITM Web of Conferences*, vol. 44, p. 03011, 2022, doi: 10.1051/itmconf/20224403011.
- [4] Bhushan Kinge, Shrinivas Mandhare, Pranali Chavan, and S. M. Chaware, "Resume Screening using Machine Learning and NLP: A proposed system," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, pp. 253–258, Apr. 2022, doi: 10.32628/cseit228240.
- [5] R. Jagtap and R. Bharti, "Data Operation with Regular Expression," 2019.
- [6] K. Singh, "Agile Methodology for Product Development A Conceptual Study," *International Journal of Recent Technology and Engineering (IJRTE)*, vol. 10, no. 1, pp. 209–215, May 2021, doi: 10.35940/ijrte.A5899.0510121.
- [7] Gaurav Dutta, "Resume Dataset," Kaggle. Accessed: Jan. 25, 2025. [Online]. Available: <https://www.kaggle.com/datasets/gauravduttakiit/resume-dataset/data>
- [8] C. Chapman and K. T. Stolee, "Exploring regular expression usage and context in python," in *ISSTA 2016 - Proceedings of the 25th International Symposium on Software Testing and Analysis*, Association for Computing Machinery, Inc, Jul. 2016, pp. 282–293. doi: 10.1145/2931037.2931073.
- [9] J.-C. Zhao, "Enhancement Encoding: A Novel Imbalanced Classification Approach via Encoding the Training Labels," Aug. 2022, [Online]. Available: <http://arxiv.org/abs/2208.11056>
- [10] N. A. Saputra, K. Aeni, and N. M. Saraswati, "Indonesian Hate Speech Text Classification Using Improved K-Nearest Neighbor with TF-IDF-ICSpF," *Scientific Journal of Informatics*, vol. 11, no. 1, pp. 21–30, Feb. 2024, doi: 10.15294/sji.v11i1.48085.

- [11] D. E. Wijayanti, S. Afriyani, S. Surono, and D. Ad, "Optimization of feature selection on semi-supervised data," *Bulletin of Applied Mathematics and Mathematics Education*, vol. 4, no. 2, pp. 61–74, Oct. 2024, doi: 10.12928/bamme.v4i2.11104.
- [12] J. Cai, J. Luo, S. Wang, and S. Yang, "Feature selection in machine learning: A new perspective," *Neurocomputing*, vol. 300, pp. 70–79, Jul. 2018, doi: 10.1016/j.neucom.2017.11.077.
- [13] Y. Li, T. Li, and H. Liu, "Recent advances in feature selection and its applications," *Knowl Inf Syst*, vol. 53, no. 3, pp. 551–577, Dec. 2017, doi: 10.1007/s10115-017-1059-8.
- [14] S. Bahassine, A. Madani, M. Al-Sarem, and M. Kissi, "Feature selection using an improved Chi-square for Arabic text classification," *Journal of King Saud University - Computer and Information Sciences*, vol. 32, no. 2, pp. 225–231, Feb. 2020, doi: 10.1016/j.jksuci.2018.05.010.
- [15] A. M. Mesleh, "Feature sub-set selection metrics for Arabic text classification," *Pattern Recognit Lett*, vol. 32, no. 14, pp. 1922–1929, Oct. 2011, doi: 10.1016/j.patrec.2011.07.010.
- [16] F. Sebastiani, "Machine learning in automated text categorization," *ACM Comput Surv*, vol. 34, no. 1, pp. 1–47, Mar. 2002, doi: 10.1145/505282.505283.
- [17] M. Kuhn and K. Johnson, *Applied Predictive Modeling*. New York, NY: Springer New York, 2013. doi: 10.1007/978-1-4614-6849-3.
- [18] Y. Yin *et al.*, "IGRF-RFE: a hybrid feature selection method for MLP-based network intrusion detection on UNSW-NB15 dataset," *J Big Data*, vol. 10, no. 1, Dec. 2023, doi: 10.1186/s40537-023-00694-8.
- [19] A. Bolívar, V. García, R. Alejo, R. Florencia-Juárez, and J. S. Sánchez, "Data-Centric Solutions for Addressing Big Data Veracity with Class Imbalance, High Dimensionality, and Class Overlapping," *Applied Sciences (Switzerland)*, vol. 14, no. 13, Jul. 2024, doi: 10.3390/app14135845.
- [20] S. A. Alex, "Imbalanced data learning using SMOTE and deep learning architecture with optimized features," *Neural Comput Appl*, 2024, doi: 10.1007/s00521-024-10481-y.
- [21] M. Schubach, M. Re, P. N. Robinson, and G. Valentini, "Imbalance-Aware Machine Learning for Predicting Rare and Common Disease-Associated Non-Coding Variants," *Sci Rep*, vol. 7, no. 1, p. 2959, Jun. 2017, doi: 10.1038/s41598-017-03011-5.

- [22] D. Elreedy and A. F. Atiya, "A Comprehensive Analysis of Synthetic Minority Oversampling Technique (SMOTE) for handling class imbalance," *Inf Sci (N Y)*, vol. 505, pp. 32–64, Dec. 2019, doi: 10.1016/j.ins.2019.07.070.
- [23] H. Hairani, T. Widiyaningtyas, and D. Dwi Prasetya, "Addressing Class Imbalance of Health Data: A Systematic Literature Review on Modified Synthetic Minority Oversampling Technique (SMOTE) Strategies," *JOIV : International Journal on Informatics Visualization*, vol. 8, no. 3, p. 1310, Sep. 2024, doi: 10.62527/joiv.8.3.2283.
- [24] A. Hoarau, A. Martin, J. C. Dubois, and Y. Le Gall, "Evidential Random Forests," *Expert Syst Appl*, vol. 230, Nov. 2023, doi: 10.1016/j.eswa.2023.120652.
- [25] L. Breiman, "Bagging predictors," *Mach Learn*, vol. 24, no. 2, pp. 123–140, Aug. 1996, doi: 10.1007/BF00058655.
- [26] S. González, S. García, J. Del Ser, L. Rokach, and F. Herrera, "A practical tutorial on bagging and boosting based ensembles for machine learning: Algorithms, software tools, performance study, practical perspectives and opportunities," *Information Fusion*, vol. 64, pp. 205–237, Dec. 2020, doi: 10.1016/j.inffus.2020.07.007.
- [27] L. Yang and A. Shami, "On hyperparameter optimization of machine learning algorithms: Theory and practice," *Neurocomputing*, vol. 415, pp. 295–316, Nov. 2020, doi: 10.1016/j.neucom.2020.07.061.
- [28] D. R. Wijaya, R. Handayani, T. Fahrudin, G. P. Kusuma, and F. Afianti, "Electronic Nose and Optimized Machine Learning Algorithms for Noninfused Aroma-Based Quality Identification of Gambung Green Tea," *IEEE Sens J*, vol. 24, no. 2, pp. 1880–1893, Jan. 2024, doi: 10.1109/JSEN.2023.3337264.
- [29] C. Goutte and E. Gaussier, "A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation," 2005, pp. 345–359. doi: 10.1007/978-3-540-31865-1_25.
- [30] "Classification: Accuracy, recall, precision, and related metrics," Google Developer. Accessed: Jan. 25, 2025. [Online]. Available: <https://developers.google.com/machine-learning/crash-course/classification/accuracy-precision-recall>
- [31] M. A. H. Frasanta, D. R. Wijaya, H. Nugroho, and T. Fahrudin, "Heart Diagnose Application Using Bagging Algorithm," in *2022 1st International Conference on Information System & Information Technology (ICISIT)*, IEEE, Jul. 2022, pp. 342–346. doi: 10.1109/ICISIT54091.2022.9872984.

- [32] D. R. Wijaya, R. Sarno, and E. Zulaika, "DWTLSTM for electronic nose signal processing in beef quality monitoring," *Sens Actuators B Chem*, vol. 326, Jan. 2021, doi: 10.1016/j.snb.2020.128931.