DAFTAR PUSTAKA

Ahuja, R., Chug, A., Kohli, S., Gupta, S., Ahuja, P., 2019., The impact of features extraction on the sentiment analysis, *Procedia Computer Science*, Vol.152, 341-348. https://doi.org/10.1016/j.procs.2019.05.008.

Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M., 2019., Optuna: A next-generation hyperparameter optimization framework, *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Anchorage AK, USA, 4-8 August 2019, 2623-2631. https://doi.org/10.1145/3292500.3330701.

Alawi, A. B., Bozkurt, F., 2024., A hybrid machine learning model for sentiment analysis and satisfaction assessment with Turkish universities using Twitter data, *Decision Analytics Journal*, Vol.11, 100473. https://doi.org/10.1016/j.dajour.2024.100473.

Almuayqil, S. N., Humayun, M., Jhanjhi, N. Z., Almufareh, M. F., Javed, D., 2022., Framework for improved sentiment analysis via random minority oversampling for user tweet review classification, *Electronics*, Vol.11 No.19, 3058. https://doi.org/10.3390/electronics11193058.

An, C., Park, Y. W., Ahn, S. S., Han, K., Kim, H., Lee, S. K., 2021., Radiomics machine learning study with a small sample size: Single random training-test set split may lead to unreliable results, *PLoS One*, Vol.16 No.8, e0256152. https://doi.org/10.1371/journal.pone.0256152.

Anisha, P. R., Nguyen, N. G., Sreelatha, G., 2021., A text mining using Web scraping for meaningful insights, *Journal of Physics: Conference Series*, Vol.2089 No.1, 012048. IOP Publishing. https://doi.org/10.1088/1742-6596/2089/1/012048.

Awajan, I., Mohamad, M., Al-Quran, A., 2021., Sentiment analysis technique and neutrosophic set theory for mining and ranking big data from online reviews, *IEEE Access*, Vol.9, 47338-47353. https://doi.org/10.1109/ACCESS.2021.3067844.

Azizah, A. N., Asy'ari, M. F., Prastya, I. W. D., Purwitasari, D., 2023., Easy Data Augmentation untuk Data yang Imbalance pada Konsultasi Kesehatan Daring, *Jurnal Teknologi Informasi dan Ilmu Komputer*, Vol.10, 1095-1104. https://doi.org/10.25126/jtiik.20231057082.

Cheng, K., Gao, S., Dong, W., Yang, X., Wang, Q., Yu, H., 2020., Boosting label weighted extreme learning machine for classifying multi-label imbalanced data, *Neurocomputing*, Vol. 403, 360-370. https://doi.org/10.1016/j.neucom.2020.04.098.

Cheng, M. Y., Kusoemo, D., Gosno, R. A., 2020., Text mining-based construction site accident classification using hybrid supervised machine

learning, *Automation in Construction*, Vol.118, 103265. https://doi.org/10.1016/j.autcon.2020.103265.

Chicco, D., Jurman, G., 2023., The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification, *BioData Mining*, Vol.16 No.4. https://doi.org/10.1186/s13040-023-00322-4.

Choudhary, R., Shukla, S., 2022., Reduced-Kernel weighted extreme learning machine using universum data in feature space (RKWELM-UFS) to handle binary class imbalanced dataset classification, *Symmetry*, Vol.14 No.2, 379. https://doi.org/10.3390/sym14020379.

Dutschmann, T. M., Kinzel, L., Ter Laak, A., Baumann, K., 2023., Large-scale evaluation of k-fold cross-validation ensembles for uncertainty estimation, *Journal of cheminformatics*, Vol.15 No.1, 49. https://doi.org/10.1186/s13321-023-00709-9.

Elistiana, K. M., Kusuma, B. A., Subarkah, P., Rozaq, H. A. A., 2023., Improvement of Naive Bayes Algorithm in Sentiment Analysis of Shopee Application Reviews on Google Play Store, *Jurnal Teknik Informatika* (JUTIF), Vol.4 No.6, 1431-1436. https://doi.org/10.52436/1.jutif.2023.4.6.1486.

Feng, X., 2019., Research of sentiment analysis based on AdaBoost algorithm, 2019 International Conference on Machine learning, Big Data and Business Intelligence (MLBDBI), 279-282. IEEE. https://doi.org/10.1109/MLBDBI48998.2019.00062.

Ganesan, A., 2024., Multi Feature Descriptor Based Ship Wake Detection Using AdaBoost-Weighted Extreme Learning Machine, *2024 IEEE Space, Aerospace and Defence Conference* (SPACE), Bangalore, India, 22-23 July 2024, 673-677. IEEE. https://doi.org/10.1109/SPACE63117.2024.10667783.

Gao, M., Wang, J., Liu, O., 2024., Is UGC sentiment helpful for recommendation? An application of sentiment-based recommendation model, *Industrial Management & Data Systems*, Vol.124, 1356-1384. https://doi.org/10.1108/IMDS-05-2023-0335.

Ghatora, P. S., Hosseini, S. E., Pervez, S., Iqbal, M. J., Shaukat, N., 2024., Sentiment Analysis of Product Reviews Using Machine Learning and Pre-Trained LLM, *Big Data and Cognitive Computing*, Vol.8 No.12, 199. https://doi.org/10.3390/bdcc8120199.

Gnanakumaran, R., Rohatgi, D., Sampath, A. K., Nagar, N., Amuthaguka, D., Gupta, R. K., 2023., Robust Extreme Learning Machine based Sentiment Analysis and Classification, *2023 5th International Conference on Smart Systems and Inventive Technology* (ICSSIT), Tirunelveli, India, 23-25 January 2023, 817-822. IEEE. https://doi-org/10.1109/ICSSIT55814.2023.10061017.

Gumaei, A., Hassan, M. M., Hassan, M. R., Alelaiwi, A., Fortino, G., 2019., A hybrid feature extraction method with regularized Extreme Learning Machine for brain tumor classification, *IEEE Access*, Vol.7, 36266-36273. https://doi.org/10.1109/ACCESS.2019.2904145.

Ilemobayo, J. A., Durodola, O., Alade, O., Awotunde, O. J., Olanrewaju, A. T., Falana, O., Ogungbire, A., Osinuga, A., Ogunbiyi, D., Ifeanyi, A., Odezuligbo, I. E., Edu, O. E., 2024., Hyperparameter Tuning in Machine Learning: A Comprehensive Review, *Journal of Engineering Research and Reports*, Vol.26 No.6, 388-395. https://doi.org/10.9734/jerr/2024/v26i61188.

Iranzad, R., Liu, X., 2024., A review of Random Forest-based feature selection methods for data science education and applications, *International Journal of Data Science and Analytics*, 1-15. https://doi.org/10.1007/s41060-024-00509-w.

Jiang, H., Tang, Y., Liu, J., Zhang, X., 2020., Extreme Learning Machine based ensemble methods for sentiment classification, *Information Sciences*, Vol.520, 18-34. https://doi.org/10.1016/j.ins.2020.02.015.

Joy, J., Selvan, M. P., 2022., A comprehensive study on the performance of different Multi-class Classification Algorithms and Hyperparameter Tuning Techniques using Optuna, *2022 International Conference on Computing, Communication, Security and Intelligent Systems* (IC3SIS), Kochi, India, 23-25 June 2022, 1-5. IEEE. https://doi.org/10.1109/IC3SIS54991.2022.9885695.

Karthika, P., Murugeswari, R., Manoranjithem, R., 2019., Sentiment analysis of social media network using Random Forest algorithm, *2019 IEEE international conference on intelligent techniques in control, optimization and signal processing* (INCOS), 1-5. IEEE. https://doi.org/10.1109/INCOS45849.2019.8951367.

Kumar, R., Sachdeva, S., 2020., Web scraping: Data extraction techniques for collecting structured information from the web, *Theoretical Computer Science*, Vol.801, 73-84. https://doi.org/10.1016/j.tcs.2019.11.014.

Li, K., Kong, X., Lu, Z., Wenyin, L., Yin, J., 2014., Boosting weighted ELM for imbalanced learning, *Neurocomputing*, Vol.128, 15-21. https://doi.org/10.1016/j.neucom.2013.05.051.

Lilhore, U. K., Simaiya, S., Prasad, D., Verma, D. K., 2021., Hybrid weighted random forests method for prediction & classification of online buying customers, *Journal of Information Technology Management* (JITM), Vol.13 No.2, 245-259. https://doi.org/10.22059/jitm.2021.310062.2607.

Liu, Z., Tang, D., Li, J., Wang, R., 2017., Objective cost-sensitive-boosting-WELM for handling multi class imbalance problem, *2017 International Joint Conference on Neural Networks* (IJCNN), Anchorage, AK, USA, 14-19

May 2017, 1975-1982. IEEE. https://doi.org/10.1109/IJCNN.2017. 7966093.

Mahendran, N., Vincent, P. D. R., Srinivasan, K., Sharma, V., Jayakody, D. K., 2020., Realizing a stacking generalization model to improve the prediction accuracy of major depressive disorder in adults, *IEEE Access*, Volume 8, 49509-49522. https://doi.org/10.1109/ACCESS.2020.2977887.

Mei, N. C., Tiun, S., Sastria, G., 2024., Multi-Label Aspect-Sentiment Classification on Indonesian Cosmetic Product Reviews with IndoBERT Model, *International Journal of Advanced Computer Science & Applications*, Vol.15 No.11, 712-720. https://dx.doi.org/10.14569/IJACSA. 2024.0151168.

Mishra, P., Biancolillo, A., Roger, J. M., Marini, F., Rutledge, D. N., 2020., New data pre-processing trends based on ensemble of multiple pre-processing techniques, *TrAC Trends in Analytical Chemistry*, Vol.132, 116045. https://doi.org/10.1016/j.trac.2020.116045.

Natras, R., Soja, B., Schmidt, M., 2022., Ensemble Machine Learning of Random Forest, AdaBoost and XGBoost for Vertical Total Electron Content Forecasting, *Remote Sensing*, Vol.14 No.15, 3547. https://doi.org/ 10.3390/rs14153547.

Ng, C. Y., Law, K. M., 2020., Investigating consumer preferences on product designs by analyzing opinions from social networks using evidential reasoning, *Computers & Industrial Engineering*, Vol.139, 106180. https://doi.org/10.1016/j.cie.2019.106180.

Nufus, Gina Khayatun., 2022., *Analisis Sentimen Berbasis Aspek Untuk Kepuasan Pengguna Aplikasi Video On Demand Dengan Model Long Short Term Memory*, Master's Thesis, Semarang, Indonesia.

Özmen, C. G., Gündüz, S., 2025., Comparison of Machine Learning Models for Sentiment Analysis of Big Turkish Web-Based Data, *Applied Sciences*, Vol.15 No.5, 2297. https://doi.org/10.3390/app15052297.

Padhy, M., Modibbo, U. M., Rautray, R., Tripathy, S. S., Bebortta, S., 2024., Application of Machine Learning Techniques to Classify Twitter Sentiments Using Vectorization Techniques, *Algorithms*, Vol.17 No.11, 486. https://doi.org/10.3390/a17110486.

Palomino, M. A., Aider, F., 2022., Evaluating the effectiveness of text pre-processing in sentiment analysis, *Applied Sciences*, Vol.12 No.17, 8765. https://doi.org/10.3390/app12178765.

Prabhakar, E., Santhosh, M., Krishnan, A. H., Kumar, T., Sudhakar, R., 2019., Sentiment analysis of US airline twitter data using new adaboost approach, *International Journal of Engineering Research & Technology* (IJERT), Vol.7 No.1, 1-6. DOI: 10.17577/IJERTCONV7IS01003.

Qiu, J., 2024., An analysis of model evaluation with cross-validation: techniques, applications, and recent advances, *Advances in Economics, Management and Political Sciences*, Vol.99, 69-72. https://doi.org/10.54254/2754-1169/99/2024OX0213.

Raghuwanshi, B. S., Shukla, S., 2019., Class-specific cost-sensitive boosting weighted elm for class imbalance learning, *Memetic Computing*, Vol.11, 263-283. https://doi.org/10.1007/s12293-018-0267-4.

Rahmi, N. A., Defit, S. Okfalisa, O., 2024., The Use of Hyperparameter Tuning in Model Classification: A Scientific Work Area Identification, *JOIV: International Journal on Informatics Visualization*, Vol.8 No.4, 2181-2188. https://dx.doi.org/10.62527/joiv.8.4.3092.

Rainio, O., Teuho, J., Klén, R. 2024., Evaluation metrics and statistical tests for machine learning, *Scientific Reports*, Vol.14 No.1, 6086. https://doi.org/10.1038/s41598-024-56706-x.

Rathod, N., Wankhade, S., 2021., Review of Optimization in Improving Extreme Learning Machine, *EAI Endorsed Transactions on Industrial Networks and Intelligent Systems*, Vol.8 No.28, e2, 1-13. https://doi.org/10.4108/eai.17-9-2021.170960.

Salman, A. H., Al-Jawher, W. A. M., 2024., Performance Comparison of Support Vector Machines, AdaBoost, and Random Forest for Sentiment Text Analysis and Classification, *Journal Port Science Research*, Vol.7 No.3, 300–311. https://doi.org/10.36371/port.2024.3.8.

Setiawan, I., Widodo, A. M., Rahaman, M., Tugiman, T., Hadi, M. A., Anwar, N., Ulum, M. B., Mulyani, E. Y., Erzed, N., 2022., Utilizing Random Forest Algorithm for Sentiment Prediction Based on Twitter Data, *First Mandalika International Multi-Conference on Science and Engineering 2022*, MIMSE 2022 (Informatics and Computer Science) (MIMSE-IC-2022), 446-456. Atlantis Press. https://doi.org/10.2991/978-94-6463-084-8_37.

Sharma, S. R., Singh, B., Kaur, M., 2023., A novel approach of ensemble methods using the stacked generalization for high-dimensional datasets, *IETE journal of research*, Vol.69 No.10, 6802-6817. https://doi.org/10.1080/03772063.2022.2028582

Shekhar, S., Bansode, A., Salim, A., 2021., A comparative study of hyper-parameter optimization tools, *2021 IEEE Asia-Pacific Conference on Computer Science and Data Engineering* (CSDE), Brisbane, Australia, 08-10 December 2021, 1-6. IEEE. https://doi.org/10.1109/CSDE53843.2021.9718485.

Singrodia, V., Mitra, A., Paul, S., 2019., A review on web scrapping and its applications, *2019 International Conference on Computer Communication and Informatics* (ICCCI), 1-6. IEEE. https://doi.org/10.1109/ICCCI.2019.8821809.

Srivastava, A., Srivastava, V., Kumar, K., Srivastava, S., Garg, N., 2023., Hybrid Machine Learning Method for Sentiment Analysis, *2023 3rd International Conference on Innovative Mechanisms for Industry Applications* (ICIMIA), Bengaluru, India, 2023, 646-652, doi: 10.1109/ICIMIA60377.2023. 10426420.

Stephenie, S., Warsito, B., Prahutama, A., 2020., Sentiment analysis on tokopedia product online reviews using Random Forest method, *E3S Web of Conferences*, ICENIS 2020, Semarang, Indonesia, 12-13 August 2020, Vol.202, 16006. https://doi.org/10.1051/e3sconf/202020216006.

Suryadi, M. K., Herteno, R., Saputro, S. W., Faisal, M. R., Nugroho, R. A., 2024., Comparative study of various hyperparameter tuning on random forest classification with SMOTE and feature selection using genetic algorithm in software defect prediction, *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, Vol.6 No.2, 137-147. https://doi.org/ 10.35882/jeeemi.v6i2.375.

Talekar, B., Agrawal, S., 2020., A detailed review on decision tree and random forest, *Biosci. Biotechnol. Res. Commun*, Vol.13 No.14, 245-248. http://dx.doi.org/10.21786/bbrc/13.14/57.

van de Kamp, Rogier., 2022., *A Hybrid approach of extreme gradient boosting and Random Forest to nowcast US GDP*, Master's Thesis, Rotterdam, Belanda.

Wang, T., Cao, J., Lai, X., Chen, B., 2018., Deep weighted extreme learning machine, *Cognitive Computation*, 10, 890-907. https://doi.org/10.1007/ s12559-018-9602-9.

Xiao, X., Zou, Y., Huang, J., Luo, X., Yang, L., Li, M., Yang, P., Ji, X., Li, Y., 2024., An interpretable model for landslide susceptibility assessment based on Optuna hyperparameter optimization and Random Forest, *Geomatics, Natural Hazards and Risk*, Vol.15 No.1, 2347421. https://doi.org/10.1080/ 19475705.2024.2347421.

Zheng, D., 2024., Sentiment Analysis for Film Reviews Based on Random Forest, Science and Technology of Engineering, *Chemistry and Environmental Protection*, Vol.1 No.7, 1-5. https://doi.org/10.61173/5t8epb44.

Zulfikri, M., Yudaningtyas, E., Rahmadwati, R., 2019., Sistem Penegakan Speed Bump Berdasarkan Kecepatan Kendaraan yang Diklasifikasikan Haar Cascade Classifier, *Jurnal Teknologi dan Sistem Komputer*, Vol.7, 12-18. https://doi.org/10.14710/jtsiskom.7.1.2019.12-18.