

DAFTAR PUSTAKA

- Baptista, J. P., & Gradim, A. (2020). Understanding Fake News Consumption: A Review. *Social Sciences*, 9(10), 185.
- Bentéjac, C., Csörgő, A., & Martínez-Muñoz, G. (2021). A comparative analysis of gradient boosting algorithms. *Artificial Intelligence Review*, 54(3), 1937–1967.
- Bergstra, J., Bardenet, R., Bengio, Y., & Kégl, B. (2011). Algorithms for hyper-parameter optimization. *Advances in Neural Information Processing Systems*, 24.
- Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13(2), 281-305.
- Bischl, B., Binder, M., Lang, M., Pielok, T., Richter, J., Coors, S., ... & Lindauer, M. (2023). *Hyperparameter* optimization: Foundations, algorithms, best practices, and open challenges. *WIREs Data Mining and Knowledge Discovery*, 13(2), e1484.
- Bondielli, A., & Marcelloni, F. (2019). A survey on fake news and rumour detection techniques. *Information Sciences*, 497, 38–55.
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (2000). CRISP-DM 1.0: Step-by-step Data Mining Guide. CRISP-DM Consortium.
- Chen, T., & Guestrin, C. (2016). *XGBoost*: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794).
- Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and *Accuracy* in binary classification *Evaluation*. *BMC Genomics*, 21(1), 1-13.
- Das, M., Kamalanathan, S., & Alphonse, P. J. A. (2023). A Comparative Study on TF-IDF Feature Weighting Method and its Analysis using Unstructured Dataset.

- Elgeldawi, E., Sayed, A., Galal, A. R., & Zaki, A. M. (2021). *Hyperparameter tuning for Machine Learning algorithms used for Arabic sentiment analysis*. Informatics, 8(4), 79.
- Essa, E., Omar, K., & Alqahtani, A. (2023). Fake news detection based on a hybrid BERT and *LightGBM* models. *Complex & Intelligent Systems*.
- Faucheux, R., Marchand, J. L., & Uchida, S. (2021). Light Gradient Boosting Machine as a Regression Method for Quantitative Structure-Activity Relationships. *ArXiv preprint*.
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5), 1189-1232.
- Gao, L., & Ding, Y. (2020). Disease prediction via Bayesian *Hyperparameter* optimization and ensemble learning. *BMC Research Notes*, 13(1), 1-7.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., ... & Liu, T. Y. (2017). *LightGBM*: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems (NeurIPS)*, 30.
- Koto, F., Rahimi, A., Lau, J. H., & Baldwin, T. (2020). IndoLEM and IndoBERT: A benchmark dataset and pre-trained language model for Indonesian NLP. *Proceedings of the 28th International Conference on Computational Linguistics*, 757–770.
- Luz de Araujo, P. H., & Rodrigues Chavez, F. J. (2023). Comparison of *Text Preprocessing* methods. *Natural Language Engineering*, 29(3), 509-553.
- Minaee, S., Kalchbrenner, N., Cambria, E., Nikzad, N., Chenaghlu, M., & Gao, J. (2022). Deep Learning Based Text Classification: A Comprehensive Review. *ACM Computing Surveys*, 54(3).
- Naeem, S., Ali, A., Anam, S., & Ahmed, M. M. (2023). An Unsupervised *Machine Learning* Algorithms: Comprehensive Review. *International Journal of Computing and Digital Systems*, 13(1), 911–921.
- Oladunni, T. I., *et al.* (2024). *Confusion Matrix-Based Performance Evaluation Metrics*. *African Journal of Biomedical Research*.

- Olan, F., Jayawickrama, U., Arakpogun, E. O., Suklan, J., & Liu, S. (2022). Fake news on SocialMedia: the Impact on Society. *Information Systems Frontiers*.
- Pennycook, G., & Rand, D. G. (2021). The psychology of fake news. *Trends in Cognitive Sciences*, 25(5), 388–402.
- Pinto, G., *et al.* (2024). Improving classification *Accuracy* of fine-Tuned CNN models: Impact of *Hyperparameter* optimization. *PMC*.
- Plotnikova, V., Dumas, M., & Milani, F. P. (2022). Designing a *Data Mining* process for the financial services domain. *Journal of Decision Systems (Applying the CRISP-DM Data Mining process)*. *Data & Knowledge Engineering*, 139, 102013.
- Pratama, B. Y., & Widodo, A. (2021). Analisis penyebaran berita hoaks di Indonesia dan dampaknya terhadap kohesi sosial. *Jurnal Ilmu Komunikasi*, 18(2), 145–160.
- Probst, P., Wright, M. N., & Boulesteix, A.-L. (2019). *Hyperparameters* and *Tuning* strategies for random forest. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 9(3), e1301.
- Putatunda, S., & Rama, K. (2020). A Modified Bayesian Optimization based Hyper-Parameter *Tuning* Approach for Extreme Gradient Boosting. *ArXiv*.
- Rais, *et al.* (2025). Implementation of *Machine Learning* Algorithm with Extreme Gradient Boosting (*XGBoost*). *Journal of Applied Science, Engineering, Technology, and Education*, 7(1), 126-136.
- Ramos, J. (2003). Using TF-IDF to determine word relevance in document queries. *Proceedings of the First Instructional Conference on Machine Learning*, 242, 29–48.
- Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2020). Fake news detection on social media: A *Data Mining* perspective. *ACM SIGKDD Explorations Newsletter*, 19(1), 22–36.
- Shyrokykh, K., Girnyk, M., & Dellmuth, L. (2023). Short text classification with *Machine Learning* in the social sciences: The case of climate change on Twitter. *PLOS ONE*, 18(9), e0290762.

- Siino, M., Tinnirello, I., & La Cascia, M. (2024). Is *Text Preprocessing* still worth the time? A comparative survey on the influence of popular *Preprocessing* methods on Transformers and traditional classifiers. *Information Systems*, 121, 102342.
- Sitarz, M. (2022). Extending F1 metric, probabilistic approach. ArXiv preprint.
- Tho, Q. C., Nguyen, L. A. T., & Nguyen, T. T. (2021). Fake news detection using TF-IDF and *Machine Learning* techniques. *International Journal of Advanced Computer Science and Applications*, 12(9), 514–521.
- Trisna, S., *et al.* (2025). Classifying Indonesian *Hoax* News Titles with SVM, *XGBoost*, and BiLSTM. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*.
- Wendland, A., Zenere, M., & Niemann, J. (2021). Introduction to Text Classification: Impact of *Stemming* and Comparing TF-IDF and Count Vectorization as Feature Extraction Technique. In *EuroSPI 2021*. Springer, Cham.
- Wiens, M. T., *et al.* (2025). A Tutorial and Use Case Example of the eXtreme Gradient Boosting (*XGBoost*) Artificial Intelligence Algorithm for Drug Development Applications. *Clinical and Translational Science*.
- Zhou, H. (2022). Research of Text Classification Based on TF-IDF and CNN-LSTM. *Journal of Physics: Conference Series*, 2171, 012021.
- Zubiaga, A., Aker, A., Bontcheva, K., Liakata, M., & Procter, R. (2020). Detection and resolution of rumours in social media: A survey. *ACM Computing Surveys*, 51(2), 1–36.