

DAFTAR PUSTAKA

- Akiba, T., Sano, S., Yanase, T., Ohta, T., and Koyama, M. (2019). Optuna: A Next-generation Hyperparameter Optimization Framework, *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 2623–2631.
- Bazarra, M. S., Sherali, H. D., and Shetty, C. M. (2006). *Nonlinear Programming - Theory and Algorithms*. 3rd ed. Wiley Interscience.
- Bengio, Y., Simard, P., and Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult, *IEEE Transactions on Neural Network*, vol. 5, no. 2, pp. 157-166.
- Bergstra, J., Bardenet, R., Bengio, Y., and Kégl, B. (2011). Algorithms for Hyperparameter Optimization, *NIPS'11: Proceedings of the 24th International Conference on Neural Information Processing Systems*, vol. 24, pp. 2546-2554.
- Bergstra, J. and Bengio, Y. (2012). Random search for hyper-parameter optimization, *The Journal of Machine Learning Research*, vol. 13, no. 1, pp. 281–305.
- Bergstra, J., Yamins, D., and Cox, D. (2013). Making a Science of Model Search: Hyperparameter Optimization in Hundreds of Dimensions for Vision Architectures, *Proceedings of the 30th International Conference on Machine Learning*, vol. 28, no. 1, pp. 115-123.
- Chong, E. K. P. and Zak, S. H. (2013). *An Introduction to Optimization*. 4th ed. Wiley.
- Falkner, S., Klein, A., and Hutter, F. (2018). BOHB: Robust and Efficient Hyperparameter Optimization at Scale, *Proceedings of the 35th International Conference on Machine Learning*, pp. 1437-1446.

- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. Cambridge: MIT Press.
- Graves, A., Jaitly, N., and Mohamed, A. (2013). Hybrid speech recognition with Deep Bidirectional LSTM, *2013 IEEE Workshop on Automatic Speech Recognition and Understanding*, pp. 273-278.
- Graves, A., Mohamed, A., and Hinton, G. E. (2013). Speech recognition with deep recurrent neural networks, *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 6645-6649.
- Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory, *Neural Computation*, pp. 1735-1780.
- Jamieson, K. and Talwakar, A. (2016). Non-stochastic Best Arm Identification and Hyperparameter Optimization, *Journal of Machine Learning Research*, vol. 41, pp. 240-248.
- Kennedy, M. C. and O'Hagan, A. (2001). Bayesian calibration of computer models, *Journal of the Royal Statistical Society: Statistical Methodology Series B*, vol. 63, no. 3, pp. 425-464.
- Larochelle, H., Erhan, D., Courville, A., Bergstra, J., and Bengio, Y. (2007). An empirical evaluation of deep architectures on problems with many factors of variation, *ICML '07: Proceedings of the 24th international conference on Machine learning*, pp. 473-480.
- Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A., and Talwakar, A. (2018). Hyperband: A Novel Bandit-Based Approach to Hyperparameter Optimization, *Journal of Machine Learning Research*, vol. 18, pp. 1-52.
- Li, L., Jamieson, K., Rostamizadeh, A., Gonina, E., Ben-Tzur, J., Hardt, M., Recht, B., and Talwakar, A. (2020). A System for Massively Parallel Hyperparameter Tuning, *Third Conference on Systems and Machine Learning*, vol. 2, pp. 230-246.

- Lu, Y., Wang, S., Zhao, Y., and Yan, C. (2014). Renewable energy system optimization of low/zero energy buildings using single-objective and multi-objective optimization methods, *Energy and Buildings*, vol. 89, pp. 61-75.
- Ludwig, S. A. (2011). Single-objective versus multi-objective Genetic algorithms for workflow composition based on service level agreements, *2011 IEEE International Conference on Service-Oriented Computing and Applications (SOCA)*, pp. 1-8.
- Mahrach, M., Miranda, G., León, C., and Segredo, E. (2020). Comparison between Single and Multi-Objective Evolutionary Algorithms to Solve the Knapsack Problem and the Travelling Salesman Problem, *Mathematics*, vol. 8, no. 11.
- Mockus, J., Tiesis, V., and Zilinskas, A. (1978). The application of Bayesian methods for seeking the extremum, *Towards Global Optimization*, vol. 2, pp. 117-129.
- Miettinen, K. (1998). *Nonlinear Multiobjective Optimization*, New York: Springer
- Parzen, E. (1962). On estimation of a probability density function and mode, *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1065-1076.
- Purwarianti, A. and Crisdayanti, I. A. P. A. (2019). Improving Bi-LSTM Performance for Indonesian Sentiment Analysis Using Paragraph Vector, *2019 International Conference of Advanced Informatics: Concepts, Theory and Applications (ICAICTA)*, pp. 1-5.
- Quadrana, M., Larreche-Mouly, A., and Mauch, M. (2022). Multi-Objective Hyper-Parameter Optimization of Behavioral Song Embeddings, *Proceeding of the 23rd International Society for Music Information Retrieval Conference*.
- Rasmussen, C. E and Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*, MIT Press.
- Rehman, A.U., Malik, A.K., Raza, B., and Ali, W. (2019). A Hybrid CNN-LSTM Model for Improving Accuracy of Movie Reviews Sentiment Analysis, *Multimedia Tools and Applications*, vol. 78, pp. 26597–26613.

- Robinson, T. D., Eldred, M. S., Willcox, K. E., and Haimes, R. (2008). Surrogate-Based Optimization Using Multifidelity Models with Variable Parameterization and Corrected Space Mapping, *AIAA Journal*, vol. 46, no. 11, pp. 2814-2822.
- Royden, H. L. and Fitzpatrick, P. M. (2010). *Real Analysis*, 4th ed. Pearson.
- Rumelhart, D., Hinton, G., and Williams, R. (1986). Learning representations by back-propagating errors, *Nature*, vol. 323, pp. 533–536.
- Sagheer, A. and Kotb, M. (2019). Time series forecasting of petroleum production using deep LSTM recurrent networks, *Neurocomputing*, vol. 323, pp. 203-213.
- Saputri, M. S., Mahendra, R., and Adriani, M. (2018). Emotion classification on Indonesian twitter dataset, *2018 International Conference on Asian Language Processing (IALP)*, pp. 90-95.
- Simonoff, J. S. (1996). *Smoothing Methods in Statistics*, New York: Springer.
- Shahriari, B., Swersky, K., Wang, Z., Adams, R. P., and de Freitas, N. (2016). Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proceedings of the IEEE*, vol. 104, no. 1, pp. 148-175.
- Snoek, J., Larochelle, H., and Adams, R. P. (2012). Practical Bayesian Optimization of Machine Learning Algorithms, *NIPS'12: Proceedings of the 25th International Conference on Neural Information Processing Systems*, vol. 2, pp. 2951–2959.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting, *Journal of Machine Learning Research*, vol. 15, pp. 1929-1958.
- Viadinugroho, R. A. A. and Rosadi, D. (2021). Long Short-Term Memory Neural Network Model for Time Series Forecasting: Case Study of Forecasting IHSG during Covid-19 Outbreak, *Journal of Physics: Conference Series*, vol. 1863.
- Wang, B., Wang, A., Chen, F., Wang, Y., and Jay Kuo, C. C. (2019). Evaluating word embedding models: methods and experimental results, *APSIPA Transactions on Signal and Information Processing*, vol. 8, pp. 1-14.

- Wilie, B., Vincentio, K., Winata, G. I., Cahyawijaya, S., Li, X., Lim, Z. Y., Soleman, S., Mahendra, R., Fung, P., Bahar, S., and Purwarianti, A. (2020). IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding, *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*, pp. 843-857.
- Wilson, J. T., Hutter, F., and Deisenroth, M. P. (2018). Maximizing acquisition functions for Bayesian optimization, *32rd Conference on Neural Information Processing Systems (NeurIPS 2018)*, pp. 9906-9917.
- Wu, J., Chen, X., Zhang, H., Xiong, L., Lei, H., and Deng, S. (2019). Hyperparameter Optimization for Machine Learning Models Based on Bayesian Optimization, *Journal of Electronic Science and Technology*, vol. 17, no. 1, pp. 26-40.
- Xiangxue, W., Lunhui, X., and Kaixun, X. (2019). Data-Driven Short-Term Forecasting for Urban Road Network Traffic Based on Data Processing and LSTM-RNN, *Arabian Journal for Science and Engineering*, vol. 44, pp. 3043-3060.
- Zadeh, L. (1963). Optimality and Non-Scalar-Valued Performance Criteria, *IEEE Transactions on Automatic Control*, pp. 59-60.
- Zeleny, M. (1973). "Compromise Programming", in Cochrane, J. L. and Zeleny, M. (eds) *Multiple Criteria Decision Making*. pp. 262-301. Columbia: University of South Carolina Press.