

References

- [1] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [2] Y. Bengio, P. Frasconi, and P. Simard, “Problem of learning long-term dependencies in recurrent networks,” *1993 IEEE Int. Conf. Neural Networks*, no. February, pp. 1183–1188, 1993, doi: 10.1109/icnn.1993.298725.
- [3] M. Sundermeyer, R. Schlüter, and H. Ney, “LSTM neural networks for language modeling,” *13th Annu. Conf. Int. Speech Commun. Assoc. 2012, INTERSPEECH 2012*, vol. 1, pp. 194–197, 2012.
- [4] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan, “Show and tell: A neural image caption generator,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 07-12-June, pp. 3156–3164, 2015, doi: 10.1109/CVPR.2015.7298935.
- [5] A. A. Nugraha, A. Arifianto, and Suyanto, “Generating Image Description on Indonesian Language using Convolutional Neural Network and Gated Recurrent Unit,” in *2019 7th International Conference on Information and Communication Technology (ICoICT)*, Jul. 2019, pp. 1–6, doi: 10.1109/ICoICT.2019.8835370.
- [6] Q. V. Le Ilya Sutskever Oriol Vinyals, “Sequence to Sequence Learning with Neural Networks,” *Neural Inf. Process. Syst. Proc.*, 2014.
- [7] G. Melis, C. Dyer, and P. Blunsom, “On the state of the art of evaluation in neural language models,” *6th Int. Conf. Learn. Represent. ICLR 2018 - Conf. Track Proc.*, pp. 1–10, 2018.
- [8] S. Merity, N. Shirish, and K. Richard, “An Analysis of Neural Language Modeling at Multiple Scales,” *CoRR*, 2018.
- [9] F. Hutter, *Automated Machine Learning*. 2019.
- [10] J. Bergstra and Y. Bengio, “Random search for hyper-parameter optimization,” *J. Mach. Learn. Res.*, vol. 13, pp. 281–305, 2012.
- [11] A. Thengade and R. Dondal, “Genetic Algorithm – Survey Paper,” *MPGI Natl. Multi Conf. Int. J. Comput. Appl.*, no. March, pp. 975–8887, 2012.
- [12] Suyanto, *An informed genetic algorithm for university course and student timetabling problems*, vol. 6114 LNAI, no. PART 2. 2010, doi: https://doi.org/10.1007/978-3-642-13232-2_28
- [13] R. M. Karim and S. Suyanto, “Optimizing Parameters of Automatic Speech Segmentation into Syllable Units,” *Int. J. Intell. Syst. Appl.*, vol. 11, no. 5, pp. 9–17, 2019, doi: 10.5815/ijisa.2019.05.02.
- [14] N. R. Emillia, Suyanto, and W. Maharani, “Isolated word recognition using ergodic hidden markov models and genetic algorithm,” *Telkomnika*, vol. 10, no. 1, pp. 129–136, 2012, doi: 10.12928/telkomnika.v10i1.769.
- [15] H. R. Ahmed and J. I. Glasgow, “Swarm Intelligence: Concepts, Models and Applications,” *Queen’s Univ. Sch. Comput. Tech. Reports*, vol. 585, no. February, pp. 1–50, 2012, doi: 10.13140/2.1.1320.2568.
- [16] A. P. Pertiwi and Suyanto, “Globally Evolved Dynamic Bee Colony Optimization,” in *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, 12-14 September 2011, Published in Part I, LNAI vol. 6881, pp. 52-61, Springer-Verlag Berlin Heidelberg 2011, Print ISBN: 978-3-642-23, 2011, no. 1, pp. 52–61, doi: https://doi.org/10.1007/978-3-642-23851-2_6.*
- [17] V. Clarissa and S. Suyanto, “New Reward-Based Movement to Improve Globally-Evolved BCO in Nurse Rostering Problem,” in *2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, Dec. 2019, pp. 114–117, doi: 10.1109/ISRITI48646.2019.9034669.
- [18] F. Ghaisani and S. Suyanto, “Discrete Firefly Algorithm for an Examination Timetabling,” in *2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, Dec. 2019, pp. 1–4, doi: 10.1109/ISRITI48646.2019.9034668.
- [19] U. Abdillah and S. Suyanto, “Clustering Nodes and Discretizing Movement to Increase the Effectiveness of HEFA for a CVRP,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 4, pp. 774–779, 2020, doi: <https://dx.doi.org/10.14569/IJACSA.2020.01104100>.

- [20] G. K. Jati and S. Suyanto, "Evolutionary Discrete Firefly Algorithm for Travelling Salesman Problem," in *Adaptive and Intelligent Systems: Second International Conference, ICAIS 2011, Klagenfurt, Austria, September 6-8, 2011*, 2011, no. 1, pp. 393–403, doi: 10.1007/978-3-642-23857-4_38.
- [21] G. K. Jati and S. Suyanto, "Discrete Cuckoo Search for Traveling Salesman Problem," in *2012 7th International Conference on Computing and Convergence Technology (ICCCCT)*, 2014, pp. 993–997.
- [22] G. K. Jati, R. Manurung, and Suyanto, *Discrete Firefly Algorithm for Traveling Salesman Problem: A New Movement Scheme*. Elsevier Inc., 2013.
- [23] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, 2014, doi: 10.1016/j.advengsoft.2013.12.007.
- [24] T. Mikolov, A. Deoras, S. Kombrink, L. Burget, and J. H. Černocký, "Empirical evaluation and combination of advanced language modeling techniques," *Proc. Annu. Conf. Int. Speech Commun. Assoc. INTERSPEECH*, no. August, pp. 605–608, 2011.
- [25] P. Koehn, "Neural Machine Translation," no. 15, p. 66, 2015, doi: 10.5565/rev/tradumatica.203.
- [26] S. Mirjalili, "How effective is the Grey Wolf optimizer in training multi-layer perceptrons," *Appl. Intell.*, vol. 43, no. 1, pp. 150–161, 2015, doi: 10.1007/s10489-014-0645-7.
- [27] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," pp. 1–15, 2014.
- [28] H. Faris, I. Aljarah, M. A. Al-Betar, and S. Mirjalili, "Grey wolf optimizer: a review of recent variants and applications," *Neural Comput. Appl.*, vol. 30, no. 2, pp. 413–435, 2018, doi: 10.1007/s00521-017-3272-5.
- [29] M. MARCUS, B. SANTORINI, and M. MARCINKIEWICZ, "Building a Large Annotated Corpus of English: The Penn Treebank," *Comput. Linguist.*, vol. 19, no. 2, p. 313, 1993.
- [30] H. Inan, K. Khosravi, and R. Socher, "Tying word vectors and word classifiers: A loss framework for language modeling," *5th Int. Conf. Learn. Represent. ICLR 2017 - Conf. Track Proc.*, pp. 1–13, 2019.