

REFERENCES

1. Mosavi A., Salimi, M., Faizollahzadeh Ardabili, S., Rabczuk, T., Shamshirband, S., and Varkonyi-Koczy, A. (2019). State of the Art of Machine Learning Models in Energy Systems, a Systematic Review. *Energies*, 12(7), p.1301. Available at: <https://doi.org/10.3390/en12071301> (Accessed: 9 September 2019)
2. Aken, D. V., Pavlo, A., Gordon, G. J., and Zhang, B.(2017). *Automatic Database Management System Tuning Through Large-scale Machine Learning*. Proceedings of the 2017 ACM International Conference on Management of Data - SIGMOD '17. Available at: <https://doi.org/10.1145/3035918.3064029> (Accessed: 9 September 2019)
3. Lyu, J. and Zhang, J. (2018). *BP neural network prediction model for suicide attempt among Chinese rural residents*. *Journal of Affective Disorders*, 246, pp.465–473. Available at: <https://doi.org/10.1016/j.jad.2018.12.111> (Accessed: 9 September 2019)
4. Mammadli, S. (2017). *Financial time series prediction using artificial neural network based on Levenberg-Marquardt algorithm*. *Procedia Computer Science*, 120, pp.602–607. Available at: <https://doi.org/10.1016/j.procs.2017.11.285> (Accessed: 9 September 2019)
5. Ahmad, S., Lavin, A., Purdy, S. and Agha, Z. (2017). *Unsupervised real-time anomaly detection for streaming data*. *Neurocomputing*, 262, pp.134–147. Available at: <https://doi.org/10.1016/j.neucom.2017.04.070> (Accessed: 9 September 2019)
6. Samuel, A. L. (1959). *Some Studies in Machine Learning Using the Game of Checkers*. *IBM Journal of Research and Development*, 3(3), pp.210–229. Available at: <https://doi.org/10.1147/rd.33.0210> (Accessed: 15 November 2019)

7. Mitchel, T. M. (1997). *Machine Learning*. New York: Mcgraw Hill.
Available at:
<http://profsite.um.ac.ir/~monsefi/machine-learning/pdf/Machine-Learning-Tom-Mitchell.pdf> (Accessed: 15 November 2019)
8. Russell, S. and Norvig, P. (2010). *Artificial intelligence: a modern approach*. New Jersey: Pearson.
9. Das, S., Dey, A., Pal, A. and Roy, N. (2015). *Applications of Artificial Intelligence in Machine Learning: Review and Prospect*. International Journal of Computer Applications, [online] 115(9), pp.31–41. Available at:
<https://pdfs.semanticscholar.org/6850/8ffc9f75462fd31de620d03093b214734011.pdf> (Accessed: 9 September 2019)
10. Xu, B., Wang, N., Chen, T. and Li, M. (2015). *Empirical Evaluation of Rectified Activations in Convolutional Network*. [online] arXiv.org. Available at: <https://arxiv.org/abs/1505.00853> (Accessed: 17 November 2019)
11. Glorot, X., Bordes, A. and Yoshua Bengio (2011). *Deep Sparse Rectifier Neural Networks*. [online] AISTATS. Available at:
<https://www.semanticscholar.org/paper/Deep-Sparse-Rectifier-Neural-Networks-Glorot-Bordes/83174a52f38c80427e237446ccda79e2a9170742>
(Accessed: 17 November 2019)
12. Nair, V. and Hinton, G. (2010). *Rectified Linear Units Improve Restricted Boltzmann Machines*. [online] Proceedings of the 27 th International Conference on Machine Learning. Available at:
<https://www.cs.toronto.edu/~hinton/absps/reluICML.pdf> (Accessed: 13 January 2020)
13. Maas, A. L., Hannun, A.Y., & Ng, A.Y. (2013). *Rectifier Nonlinearities Improve Neural Network Acoustic Models*. Proceedings of the 30 th International Conference on Machine Learning. Available at:

https://ai.stanford.edu/~amaas/papers/relu_hybrid_icml2013_final.pdf

(Accessed: 13 January 2020)

14. Nwankpa, C., Ijomah, W., Gachagan, A. and Marshall, S. (2018). *Activation Functions: Comparison of trends in Practice and Research for Deep Learning*. [online] arXiv.org. Available at: <https://arxiv.org/abs/1811.03378> (Accessed: 13 January 2020)
15. Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986). Learning representations by back-propagating errors. *Nature*, [online] 323(6088), pp.533–536. Available at: https://www.nature.com/articles/323533a?error=cookies_not_supported&code=2926f83e-9c3a-46a7-9b28-b3d19d46768a (Accessed: 13 January 2020)
16. Choi, D., Shallue, C.J., Nado, Z., Lee, J., Maddison, C.J. and Dahl, G.E. (2019). *On Empirical Comparisons of Optimizers for Deep Learning*. [online] arXiv.org. Available at: <https://arxiv.org/abs/1910.05446> (Accessed: 13 January 2020).
17. Kingma, D.P. and Ba, J. (2014). *Adam: A Method for Stochastic Optimization*. [online] arXiv.org. Available at: <https://arxiv.org/abs/1412.6980> (Accessed: 18 November 2019)
18. Ruder, S. (2016). *An overview of gradient descent optimization algorithms*. [online] arXiv.org. Available at: <https://arxiv.org/abs/1609.04747> (Accessed: 18 November 2019)
19. Godoy, D. (2018). *Hyper-parameters in action!* [online] Medium. Available at: <https://towardsdatascience.com/hyper-parameters-in-action-a524bf5bf1c> (Accessed: 20 November 2019)