

## DAFTAR PUSTAKA

- Aitchison, J. and Silvey, S. (1958). Maximum-likelihood Estimation of Parameters Subject to Restraints, *Annals of Mathematical Statistics*, vol. 29, pp. 813-828.
- Aitchison, J. and Silvey, S. (1960). Maximum-likelihood Estimation Procedures and Associated Tests of Significance, *Journal of the Royal Statistical Society*, vol. 22, pp. 154-171.
- Aizenberg, I., Sheremetov, L., Villa-Vargas, L., and Martinez-Muñoz, J. (2016). Multilayer Neural Network with Multi-Valued Neurons in time series forecasting of oil production, *Neurocomputing*, vol. 175, pp. 980–989.
- Aggarwal, C. C. (2018). *Neural Networks and Deep Learning*. Springer.
- Anton, H. and Rorres, C. (2014). *Elementary Linear Algebra*. 11th ed. Wiley.
- Bain, L. J. and Engelhardt, M. (1992). *Introduction to Probability and Mathematical Statistics*. 2nd ed. Duxbury Press.
- 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.
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
- Brownlee, J. (2017). *Long Short-Term Memory Networks With Python*. Jason Brownlee: Melbourne.
- Cleveland, R. B., Cleveland, W. S., McRae, J. E., and Terpenning, I. (1990). STL: A Seasonal-Trend Decomposition Procedure Based on Loess, *Journal of Official Statistics*, vol. 6, no. 1, pp. 3-73.

- Davey, N., Hunt, S., and Frank, R. (1999). Time Series Prediction and Neural Networks, *Procs 5th Int Conf on Engineering Applications of Neural Networks (EANN'99)*, pp. 93-98.
- Duchi, J., Hazan, E., and Singer, Y. (2011). Adaptive subgradient methods for online learning and stochastic optimization, *The Journal of Machine Learning Research*, vol. 12, pp. 2121-2159.
- Finance.yahoo.com. (2020). *Yahoo! Finance*. [online] Available at: <https://finance.yahoo.com/quote/%5EJKSE?p=%5EJKSE> [Accessed 5 Feb. 2020].
- Finance.yahoo.com. (2020). *Yahoo! Finance*. [online] Available at: <https://finance.yahoo.com/quote/USDIDR%3DX/history?p=USDIDR%3DX> [Accessed 3 Mar. 2020].
- Firmino, P. R. A., de Mattos Neto, P. S. G., and Ferreira, T. A. E. (2014). Correcting and combining time series forecasters, *Neural Networks*, vol. 50, pp. 1-11.
- 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.
- Han, J., Kamber, M., and Pei, J. (2012). *Data Mining Concepts and Techniques*. 3rd ed. Boston: Elsevier.
- Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory, *Neural Computation*, pp. 1735-1780.
- Hogg, R. V., McKean, J., and Craig, A. T. (2014). *Introduction to Mathematical Statistics*. 7th ed. New York: Pearson Education.

Hosseini, MP., Lu, S., Kamaraj, K., Slowikowski, A., and Venkatesh, H. C. (2020). "Deep Learning Architectures" in Pedrycz, W. and Chen, SM. (eds) *Deep Learning: Concepts and Architectures. Studies in Computational Intelligence*. vol. 866, pp. 1-23. Springer.

Kelleher, J. D. (2019). *Deep Learning*. Cambridge: MIT Press.

Kingma, D. P. and Ba, J. (2015). A method for stochastic optimization, *Proceedings of 3rd International Conference on Learning Representations*.

McCulloch, W. S. and Pitts, W. H. (1943). A logical calculus of the ideas immanent in nervous activity, *Bulletin of Mathematical Biophysics*, vol. 5, pp. 115-133.

Picard, R. R. and Berk, K. N. (1990). Data Splitting, *The American Statistician*, vol. 44, no. 2, pp. 140-147.

Rosadi, D. (2011). *Analisis Ekonometrika dan Runtun Waktu Terapan dengan R*. Yogyakarta: Andi Offset.

Rosadi, D. (2014). *Analisis Runtun Waktu dan Aplikasinya dengan R*. Yogyakarta: UGM Press.

Rumelhart, D., Hinton, G., and Willams, 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.

Schuster, M. and Paliwal, K. K. (1992). Bidirectional Recurrent Neural Networks, *IEEE Transactions On Signal Processing*, vol. 45, no. 11, pp. 2673-2681.

Schmidhuber, J. (1992). Learning complex, extended sequences using the principle of history compression, *Neural Computation*, vol. 4, no. 2, pp. 234-242.

Sedki, A., Ouazar, D., and El Mazoudi, E. (2009). Evolving neural network using real coded genetic algorithm for daily rainfall-runoff forecasting, *Expert Systems with Applications*, vol. 36, pp. 4523-4527.

- 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.
- Subanar (2013). *Statistika Matematika: Probabilitas, Distribusi, dan Asimtotis dalam Statistika*. Yogyakarta: Graha Ilmu.
- Theodosiou, M. (2011). Disaggregation & aggregation of time series components: A hybrid forecasting approach using generalized regression neural networks and the theta method, *Neurocomputing*, vol. 74, pp. 896–905.
- Wei, W. W. S. (2006). *Time Series Analysis Univariate and Multivariate Methods*. Boston: Pearson Education.
- 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.
- Zhu, W., Lan, C., Xing, J., Zeng, W., Li, Y., Shen, L., and Xie, X. (2016). Co-Occurrence Feature Learning for Skeleton Based Action Recognition Using Regularized Deep LSTM Networks, *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)*, pp. 3697-3703.