

References

- [1] M. Iqbal, M. Zahid, D. Habib, and L. John, “Efficient prediction of network traffic for real-time applications,” *Journal of Computer Networks and Communications*, pp. 1–11, 2019. doi: 10.1155/2019/4067135
- [2] A. Azari, P. Papapetrou, S. Denic, and G. Peters, *Cellular Traffic Prediction and Classification: A Comparative Evaluation of LSTM and ARIMA*, 2019, pp. 129–144. ISBN 978-3-030-33777-3
- [3] T. Aldhyani, M. Alrasheed, A. Alqarni, Y. Alzahrani, and A. Bamhdi, “Intelligent hybrid model to enhance time series models for predicting network traffic,” *IEEE Access*, 2020. doi: 10.1109/ACCESS.2020.3009169
- [4] D. Harkut and Z. Shaikh, “An overview of network traffic classification methods,” *Int. J. Recent Innovation Trends Comput. Commun*, vol. 3, pp. 482–488, 2015.
- [5] H. D. Trinh, L. Giupponi, and P. Dini, “Mobile traffic prediction from raw data using lstm networks,” in *IEEE 29th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2018. doi: 10.1109/PIMRC.2018.8581000 pp. 1827–1832.
- [6] M. Aliyari Shoorehdeli, M. Teshnehlab, and A. Sedigh, “Training anfis as an identifier with intelligent hybrid stable learning algorithm based on particle swarm optimization and extended kalman filter,” *Fuzzy Sets and Systems*, vol. 160, pp. 922–948, 2009. doi: 10.1016/j.fss.2008.09.011
- [7] M. sup Kim, Y. J. Won, H. jo Lee, J. W. Hong, and R. Boutaba, “Flow-based characteristic analysis of internet application traffic,” in *in Proc. E2EMON*, 2004, pp. 62–67.

- [8] Juan Sebastián Rojas. (2019) Labeled network traffic flows - 141 applications. Universidad Del Cauca. [Online]. Available: <https://www.kaggle.com/jsrojas/labeled-network-traffic-flows-114-applications>
- [9] United Nations (U.N). (2015) Transforming our world: the 2030 agenda for sustainable development. [Online]. Available: <https://EconPapers.repec.org/RePEc:ess:wpaper:id:7559>
- [10] J. Clement. (2019) Mobile internet usage worldwide - Statistics & Facts. Statista Research Department. [Online]. Available: <https://www.statista.com/topics/779/mobile-internet/>
- [11] Internet Engineering Task Force (IETF). (2021) Request for Comments (RFC) Index. [Online]. Available: <https://tools.ietf.org/rfc/index>
- [12] J. F. Kurose and K. W. Ross, *Computer Networking: A Top-Down Approach*, 7th ed. Pearson, 2016. ISBN 978-0-13-359414-0
- [13] Internet Engineering Task Force (IETF). (2020) IANA IPv4 Special-Purpose Address Registry. [Online]. Available: <https://www.iana.org/assignments/iana-ipv4-special-registry/iana-ipv4-special-registry.xhtml>
- [14] J. Rajahalme, A. Conta, B. Carpenter, and S. Deering, “IPv6 Flow Label Specification,” IETF, RFC 3697, 2004. [Online]. Available: <http://tools.ietf.org/rfc/rfc3697.txt>
- [15] M. Kubat, *An Introduction to Machine Learning*, 1st ed. Springer Publishing Company, Incorporated, 2015. ISBN 3319200097
- [16] A. Singh, N. Thakur, and A. Sharma, “A review of supervised machine learning algorithms,” in *3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, 2016, pp. 1310–1315.
- [17] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to Linear Regression Analysis (4th ed.)*. Wiley & Sons, 2006. ISBN 0471754951
- [18] S. Weisberg, *Applied Linear Regression*, 3rd ed. Hoboken NJ: Wiley, 2005. ISBN 9780471704096. [Online]. Available: <http://www.stat.umn.edu/alr>
- [19] L. Rokach and O. Maimon, *Data Mining With Decision Trees: Theory and Applications*, 2nd ed. World Scientific Publishing Co., Inc., 2014. ISBN 9789814590075

- [20] M. Hjorth-Jensen, “Data analysis and machine learning: From decision trees to forests and all that.” Department of Physics and Astronomy and National Superconducting Cyclotron Laboratory, Michigan State University, 2019.
- [21] I. N. Da Silva, D. H. Spatti, R. A. Flauzino, L. H. B. Liboni, and S. F. dos Reis Alves, “Artificial neural network architectures and training processes,” in *Artificial neural networks*. Springer, 2017, pp. 21–28.
- [22] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. The MIT Press, 2016. ISBN 0262035618
- [23] K. Hinkelmann. (2021) Neural networks. University of Applied Sciences Northwestern Switzerland. [Online]. Available: http://didattica.cs.unicam.it/lib/exe/fetch.php?media\=didattica:magistrale:kebi:ay_1718:ke-11_neural_networks.pdf
- [24] B. Mehlig, “Machine learning with neural networks,” vol. ArXiv abs/1901.05639. Department of Physics, University of Gothenburg, 2019.
- [25] K. He, X. Zhang, S. Ren, and J. Sun, “Delving deep into rectifiers: Surpassing human-level performance on imagenet classification,” in *IEEE International Conference on Computer Vision (ICCV)*, 2015. doi: 10.1109/ICCV.2015.123 pp. 1026–1034.
- [26] S. Ioffe and C. Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” in *Proceedings of the 32nd International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, vol. 37. PMLR, 2015, pp. 448–456.
- [27] A. Sherstinsky, “Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network,” *ArXiv*, vol. abs/1808.03314, 2018.
- [28] R. M. Schmidt, “Recurrent neural networks (rnns): A gentle introduction and overview,” *ArXiv*, vol. abs/1912.05911, 2019.
- [29] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Comput.*, vol. 9, p. 1735–1780, 1997. doi: 10.1162/neco.1997.9.8.1735. [Online]. Available: <https://doi.org/10.1162/neco.1997.9.8.1735>
- [30] A. Graves, N. Jaitly, and A.-r. Mohamed, “Hybrid speech recognition with deep bidirectional lstm,” in *IEEE workshop on automatic speech recognition and understanding*, 2013, pp. 273–278.

- [31] N. Kalchbrenner, L. Espeholt, K. Simonyan, A. van den Oord, A. Graves, and K. Kavukcuoglu, “Neural machine translation in linear time,” vol. ArXiv abs/1610.10099, 2017. [Online]. Available: <https://arxiv.org/abs/1610.10099>
- [32] W. Noble, “What is a support vector machine?” *Nature biotechnology*, vol. 24, pp. 1565–1567, 2007. doi: 10.1038/nbt1206-1565
- [33] X.-h. Tan, W.-h. Bi, X.-l. Hou, and W. Wang, “Reliability analysis using radial basis function networks and support vector machines,” *Computers and Geotechnics*, vol. 38, no. 2, pp. 178–186, 2011.
- [34] H. Drucker, C. J. Burges, L. Kaufman, A. Smola, V. Vapnik *et al.*, “Support vector regression machines,” *Advances in neural information processing systems*, vol. 9, pp. 155–161, 1997.
- [35] C. Qiu, Y. Zhang, Z. Feng, P. Zhang, and S. Cui, “Spatio-temporal wireless traffic prediction with recurrent neural network,” *IEEE Wireless Communications Letters*, vol. 7, no. 4, pp. 554–557, 2018. doi: 10.1109/LWC.2018.2795605
- [36] W. Yuankai, H. Tan, B. Ran, and Z. Jiang, “A hybrid deep learning based traffic flow prediction method and its understanding,” *Transportation Research Part C: Emerging Technologies*, vol. 90, 2018. doi: 10.1016/j.trc.2018.03.001
- [37] J. Wang, J. Tang, Z. Xu, Y. Wang, G. Xue, X. Zhang, and D. Yang, “Spatiotemporal modeling and prediction in cellular networks: A big data enabled deep learning approach,” in *IEEE Conference on Computer Communications*, 2017. doi: 10.1109/INFOCOM.2017.8057090 pp. 1–9.
- [38] M. Ring, A. Dallmann, D. Landes, and A. Hotho, “Ip2vec: Learning similarities between ip addresses,” in *IEEE International Conference on Data Mining Workshops (ICDMW)*, 2017. doi: 10.1109/ICDMW.2017.93 pp. 657–666.
- [39] T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient estimation of word representations in vector space,” in *1st International Conference on Learning Representations*, Y. Bengio and Y. LeCun, Eds., 2013. [Online]. Available: <http://arxiv.org/abs/1301.3781>
- [40] M. Li, C. Lumezanu, B. Zong, and H. Chen, “Deep learning ip network representations,” in *Proceedings of the 2018 Workshop on Big Data Analytics and Machine Learning for Data Communication Networks*, ser.

- Big-DAMA '18. Association for Computing Machinery, 2018. ISBN 9781450359047 p. 33–39.
- [41] R. Beverly, K. Sollins, and A. Berger, “Svm learning of ip address structure for latency prediction,” in *Proceedings of the 2006 SIGCOMM Workshop on Mining Network Data*. Association for Computing Machinery, 2006. ISBN 159593569X p. 299–304.
- [42] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016. doi: 10.1109/CVPR.2016.90 pp. 770–778.
- [43] D. Chen, F. Hu, G. Nian, and T. Yang, “Deep residual learning for nonlinear regression,” *Entropy*, vol. 22, p. 193, 2020. doi: 10.3390/e22020193
- [44] X. Qiu, L. Zhang, Y. Ren, P. Suganthan, and G. Amaratunga, “Ensemble deep learning for regression and time series forecasting,” 2014. doi: 10.1109/CIEL.2014.7015739
- [45] K. Pace and R. Barry, “Sparse spatial autoregressions,” *Statistics Probability Letters*, vol. 33, no. 3, pp. 291–297, 1997. [Online]. Available: <https://EconPapers.repec.org/RePEc:eee:stapro:v:33:y:1997:i:3:p:291-297>
- [46] B. Das, D. Science, T. O. (Australia), Electronics, and S. R. L. (Australia)., *Representing uncertainties using Bayesian networks / Balaram Das*. DSTO Electronics and Surveillance Research Laboratory Salisbury, S. Aust, 1999. [Online]. Available: <http://www.dsto.defence.gov.au/corporate/reports/DSTO-TR-0918.pdf>
- [47] Y. Gal and Z. Ghahramani, “Dropout as a bayesian approximation: Representing model uncertainty in deep learning,” *ArXiv*, vol. abs/1506.02142, 2016.
- [48] D. Nix and A. Weigend, “Estimating the mean and variance of the target probability distribution,” *Proceedings of 1994 IEEE International Conference on Neural Networks (ICNN'94)*, vol. 1, pp. 55–60 vol.1, 1994.
- [49] T. Heskes, “Practical confidence and prediction intervals,” in *Advances in Neural Information Processing Systems 9*. MIT press, 1997, pp. 176–182.

- [50] E. Haglund, “Estimating prediction intervals with machine learning and monte carlo methods in online advertising,” Master’s thesis, KTH, School of Electrical Engineering and Computer Science (EECS), 2020.
- [51] D. Pevec and I. Kononenko, “Prediction intervals in supervised learning for model evaluation and discrimination,” *Applied Intelligence*, vol. 42, 2015. doi: 10.1007/s10489-014-0632-z
- [52] A. Habibi Lashkari. (2018) Cicflowmeter: Ethernet traffic bi-flow generator and analyzer for anomaly detection. University of New Brunswick. [Online]. Available: <https://github.com/ISCX/CICFlowMeter>
- [53] N. Moustafa and J. Slay, “Unsw-nb15: a comprehensive data set for network intrusion detection systems (unsw-nb15 network data set),” pp. 1–6, 2015. doi: 10.1109/MilCIS.2015.7348942
- [54] E. D. Source, “Intrusion detection evaluation dataset (iscxids2012) (06/11/2010 to 06/17/2010),” 2018. [Online]. Available: https://www.impatcybertrust.org/dataset_view?idDataset=916
- [55] F. Liu, X. Wu, W. Li, and X. Liu, “The packet size distribution patterns of the typical internet applications,” pp. 325–332, 2012. doi: 10.1109/ICNIDC.2012.6418769
- [56] Cooperative Association for Internet Data Analysis (CAIDA), “Packet length distributions,” 2020. [Online]. Available: https://www.caida.org/research/traffic-analysis/AIX/plen_hist/
- [57] K. Lan and J. Heidemann, “A measurement study of correlations of internet flow characteristics,” *Comput. Netw.*, vol. 50, no. 1, p. 46–62, 2006.
- [58] M.-S. Kim, Y. J. Won, and J. W. Hong, “Characteristic analysis of internet traffic from the perspective of flows,” *Comput. Commun.*, vol. 29, no. 10, p. 1639–1652, 2006. doi: 10.1016/j.comcom.2005.07.015
- [59] Y. Zhang, L. Breslau, V. Paxson, and S. Shenker, “On the characteristics and origins of internet flow rates,” *SIGCOMM Comput. Commun. Rev.*, vol. 32, no. 4, p. 309–322, 2002. doi: 10.1145/964725.633055. [Online]. Available: <https://doi.org/10.1145/964725.633055>

- [60] Internet Engineering Task Force (IETF). (2021) IANA Assigned Internet Protocol Numbers. [Online]. Available: <https://www.iana.org/assignments/protocol-numbers/protocol-numbers.xhtml>
- [61] M. S. Lewis-Beck and A. Skalaban, “The r-squared: Some straight talk,” *Political Analysis*, vol. 2, pp. 153–171, 1990.
- [62] C. Tofallis, “A better measure of relative prediction accuracy for model selection and model estimation,” *Journal of the Operational Research Society*, vol. 66, pp. 1352–1362, 2015. doi: 10.1057/jors.2014.103
- [63] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” in *3rd International Conference on Learning Representations Conference Track Proceedings*, 2015. [Online]. Available: <http://arxiv.org/abs/1412.6980>
- [64] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, “Gans trained by a two time-scale update rule converge to a local nash equilibrium,” in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, 2017, p. 6629–6640.
- [65] S. Wenke and J. Fleming, “Contextual recurrent neural networks,” *ArXiv*, vol. abs/1902.03455, 2019.

