Bayesian Optimization - A Review with References

A Brief List of References for INTSYS Assignment 2

Below, you can find a list of resources where some of the IPython notebooks were extracted from, and a list of references used in the slides, the walkthroughs and further reading.

REFERENCES

[1] Kai Arulkumaran, Marc Peter Deisenroth, Miles Brundage, and Anil Anthony Bharath. 2017. A brief survey of deep reinforcement learning. *arXiv preprint arXiv:1708.05866* (2017).

[2] James Bergstra and Yoshua Bengio. 2012. Random Search for Hyper-parameter Optimization. *J. Mach. Learn. Res.* 13, 1 (Feb. 2012), 281–305. http://dl.acm.org/citation.cfm?id=2503308.2188395

[3] James Bergstra, Brent Komer, Chris Eliasmith, and David Warde-Farley. 2014. Preliminary evaluation of hyperopt algorithms on HPOLib. In *ICML workshop on AutoML*. http://ai2-s2-pdfs.s3.amazonaws.com/0da5/366fb41f66827c89144e5a949cc0e7e27d47.pdf

[4] James Bergstra, Brent Komer, Chris Eliasmith, Dan Yamins, and David D. Cox. 2015. Hyperopt: a Python library for model selection and hyperparameter optimization. *Comput. Sci. Disc.* 8, 1 (2015), 014008. https://doi.org/10.1088/1749-4699/8/1/014008

[5] J. Bergstra, D. Yamins, and D. D. Cox. 2013. Making a Science of Model Search: Hyperparameter Optimization in Hundreds of Dimensions for Vision Architectures. In *Proceedings of the 30th International Conference on International Conference on Machine Learning - Volume 28 (ICML'13)*. JMLR.org, Atlanta, GA, USA, I–115–I–123. http://dl.acm.org/citation.cfm?id=3042817.3042832

[6] James S. Bergstra, RÃľmi Bardenet, Yoshua Bengio, and BalÃązs KÃľgl. 2011. Algorithms for Hyper-Parameter Optimization. In *Advances in Neural Information Processing Systems 24*, J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, and K. Q. Weinberger (Eds.). Curran Associates, Inc., 2546–2554. http://papers.nips.cc/paper/4443-algorithms-for-hyper-parameter-optimization.pdf

[7] Bernd Bischl, Pascal Kerschke, Lars Kotthoff, Marius Lindauer, Yuri Malitsky, Alexandre FrÃľchette, Holger Hoos, Frank Hutter, Kevin Leyton-Brown, Kevin Tierney, and Joaquin Vanschoren. 2016. ASlib: A benchmark library for algorithm selection. *Artificial Intelligence* 237 (Aug. 2016), 41–58. https://doi.org/10.1016/j.artint.2016.04.003

[8] Eric Brochu, Vlad M. Cora, and Nando de Freitas. 2010. A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning. *arXiv:1012.2599 [cs]* (Dec. 2010). http://arxiv.org/abs/1012.2599 arXiv: 1012.2599.

[9] Eric Brochu, Vlad M Cora, and Nando De Freitas. 2010. A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. *arXiv preprint arXiv:1012.2599* (2010).

[10] Yutian Chen, Matthew W. Hoffman, Sergio Gomez Colmenarejo, Misha Denil, Timothy P. Lillicrap, and Nando de Freitas. 2016. Learning to Learn for Global Optimization of Black Box Functions. *arXiv:1611.03824 [cs, stat]* (Nov. 2016). http://arxiv.org/abs/1611.03824 arXiv: 1611.03824.

[11] Ian Dewancker, Michael McCourt, Scott Clark, Patrick Hayes, Alexandra Johnson, and George Ke. 2016. A Stratified Analysis of Bayesian Optimization Methods. *arXiv:1603.09441 [cs, stat]* (March 2016). http://arxiv.org/abs/1603.09441 arXiv: 1603.09441.

[12] Matthias Feurer, Aaron Klein, Katharina Eggensperger, Jost Tobias Springenberg, Manuel Blum, and Frank Hutter. 2015. Efficient and Robust Automated Machine Learning. In *Proceedings of the 28th International Conference on Neural Information Processing Systems (NIPS'15)*. MIT Press, Cambridge, MA, USA, 2755–2763. http://dl.acm.org/citation.cfm?id=2969442.2969547

[13] Matthias Feurer, Aaron Klein, Katharina Eggensperger, Jost T Springenberg, Manuel Blum, and Frank Hutter. 2015. Methods for improving bayesian optimization for AutoML. In *AutoML Workshop, International Conference on Machine Learning*, Vol. 2015.

[14] Luca Franceschi, Michele Donini, Paolo Frasconi, and Massimiliano Pontil. 2017. Forward and Reverse Gradient-Based Hyperparameter Optimization. *arXiv:1703.01785 [stat]* (March 2017). http://arxiv.org/abs/1703.01785 arXiv: 1703.01785.

[15] Matthew Hoffman, Bobak Shahriari, and Nando Freitas. 2014. On correlation and budget constraints in model-based bandit optimization with application to automatic machine learning. In *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics (Proceedings of Machine Learning Research)*, Samuel Kaski and Jukka Corander (Eds.), Vol. 33. PMLR, Reykjavik, Iceland, 365–374. http://proceedings.mlr.press/v33/hoffman14.html

[16] Aaron Klein, Stefan Falkner, Simon Bartels, Philipp Hennig, and Frank Hutter. 2016. Fast Bayesian Optimization of Machine Learning Hyperparameters on Large Datasets. *arXiv:1605.07079 [cs, stat]* (May 2016). http://arxiv.org/abs/1605.07079 arXiv: 1605.07079.

[17] Lisha Li, Kevin Jamieson, Giulia DeSalvo, Afshin Rostamizadeh, and Ameet Talwalkar. 2016. Hyperband: A Novel Bandit-Based Approach to Hyperparameter Optimization. *arXiv:1603.06560 [cs, stat]* (March 2016). http://arxiv.org/abs/1603.06560 arXiv: 1603.06560.

[18] Yuxi Li. 2017. Deep reinforcement learning: An overview. *arXiv preprint arXiv:1701.07274* (2017).

[19] Romy Lorenz, Adam Hampshire, and Robert Leech. 2017. Neuroadaptive Bayesian Optimization and Hypothesis Testing. *Trends in Cognitive Sciences* 21, 3 (March 2017), 155–167. https://doi.org/10.1016/j.tics.2017.01.006

[20] Ilya Loshchilov and Frank Hutter. 2016. CMA-ES for Hyperparameter Optimization of Deep Neural Networks. *arXiv:1604.07269 [cs]* (April 2016). http://arxiv.org/abs/1604.07269 arXiv: 1604.07269.

[21] Gang Luo. 2016. A review of automatic selection methods for machine learning algorithms and hyper-parameter values. *Netw Model Anal Health Inform Bioinforma* 5, 1 (Dec. 2016), 18. https://doi.org/10.1007/s13721-016-0125-6

[22] Gustavo Malkomes, Charles Schaff, and Roman Garnett. 2016. Bayesian optimization for automated model selection. In *Advances in Neural Information Processing Systems*. 2900–2908.

[23] Hector Mendoza, Aaron Klein, Matthias Feurer, Jost Tobias Springenberg, and Frank Hutter. 2016. Towards Automatically-Tuned Neural Networks. In *Proceedings of the Workshop on Automatic Machine Learning (Proceedings of Machine Learning Research)*, Frank Hutter, Lars Kotthoff, and Joaquin Vanschoren (Eds.), Vol. 64. PMLR, New York, New York, USA, 58–65. http://proceedings.mlr.press/v64/mendoza_towards_2016.html

[24] B. Shahriari, K. Swersky, Z. Wang, R. P. Adams, and N. de Freitas. 2016. Taking the Human Out of the Loop: A Review of Bayesian Optimization. *Proc. IEEE* 104, 1 (Jan. 2016), 148–175. https://doi.org/10.1109/JPROC.2015.2494218

[25] Jasper Snoek, Hugo Larochelle, and Ryan P Adams. 2012. Practical Bayesian Optimization of Machine Learning Algorithms. In *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger (Eds.). Curran Associates, Inc., 2951–2959. http://papers.nips.cc/paper/4522-practical-bayesian-optimization-of-machine-learning-algorithms.pdf

[26] Jasper Snoek, Oren Rippel, Kevin Swersky, Ryan Kiros, Nadathur Satish, Narayanan Sundaram, Md Mostofa Ali Patwary, Prabhat, and Ryan P. Adams. 2015. Scalable Bayesian Optimization Using Deep Neural Networks. *arXiv:1502.05700 [stat]* (Feb. 2015). http://arxiv.org/abs/1502.05700 arXiv: 1502.05700.

[27] Jaemin Son, Samarth Gupta, and Gary Tan. 2016. Bayesian Optimization in High Dimensional Input Space. In *Proceedings of the 9th EAI International Conference on Simulation Tools and Techniques (SIMUTOOLS'16)*. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), ICST, Brussels, Belgium, Belgium, 18–27. http://dl.acm.org/citation.cfm?id=3021426.3021430

[28] Jost Tobias Springenberg, Aaron Klein, Stefan Falkner, and Frank Hutter. 2016. Bayesian Optimization with Robust Bayesian Neural Networks. In *Advances in Neural Information Processing Systems 29*, D. D. Lee, M. Sugiyama, U. V. Luxburg, I. Guyon, and R. Garnett (Eds.). Curran Associates, Inc., 4134–4142. http://papers.nips.cc/paper/6117-bayesian-optimization-with-robust-bayesian-neural-networks.pdf

[29] Chris Thornton, Frank Hutter, Holger H. Hoos, and Kevin Leyton-Brown. 2013. Auto-WEKA: Combined Selection and Hyperparameter Optimization of Classification Algorithms. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '13)*. ACM, New York, NY, USA, 847–855. https://doi.org/10.1145/2487575.2487629

[30] Ziyu Wang, Frank Hutter, Masrour Zoghi, David Matheson, and Nando de Freitas. 2016. Bayesian Optimization in a Billion Dimensions via Random Embeddings. *Journal of Artificial Intelligence Research* (2016). https://doi.org/10.1613/jair.4806

[31] Chengwei Yao, Deng Cai, Jiajun Bu, and Gencai Chen. 2017. Pre-training the deep generative models with adaptive hyperparameter optimization. *Neurocomputing* 247 (July 2017), 144–155. https://doi.org/10.1016/j.neucom.2017.03.058

[32] Y. Zhang, K. Sohn, R. Villegas, G. Pan, and H. Lee. 2015. Improving object detection with deep convolutional networks via Bayesian optimization and structured prediction. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 249–258. https://doi.org/10.1109/CVPR.2015.7298621