

References

- Bengio, Y. 2012. Deep learning of representations for unsupervised and transfer learning. In *JMLR W&CP: Proc. Unsupervised and Transfer Learning challenge and workshop*, volume 27, 17–36.
- Cho, K.; Raiko, T.; and Ilin, A. 2010. Parallel tempering is efficient for learning restricted Boltzmann machines. In *Proceedings of the International Joint Conference on Neural Networks (IJCNN 2010)*.
- Coates, A., and Ng, A. Y. 2011. The importance of encoding versus training with sparse coding and vector quantization. In *ICML'2011*.
- Denil, M., and de Freitas, N. 2011. Toward the implementation of a quantum rbm. *NIPS*2011 Workshop on Deep Learning and Unsupervised Feature Learning*.
- Desjardins, G.; Courville, A.; and Bengio, Y. 2010. Tempered Markov chain Monte Carlo for training of restricted Boltzmann machine. In *JMLR W&CP: Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics (AISTATS 2010)*, volume 9, 145–152.
- Dupret, A.; Belhaire, E.; Rodier, J.-C.; Lalanne, P.; Prevost, D.; Gardia, P.; and Chavel, P. 1996. An optoelectronic CMOS circuit implementing a simulated annealing algorithm. *IEEE Journal of Solid-State Circuits* 31(7):1046–1050.
- Goodfellow, I. J.; Mirza, M.; Courville, A.; and Bengio, Y. 2013. Multi-prediction deep Boltzmann machines. In *Advances in Neural Information Processing Systems 26 (NIPS'13)*. Nips Foundation (<http://books.nips.cc>).
- Hinton, G. E.; Srivastava, N.; Krizhevsky, A.; Sutskever, I.; and Salakhutdinov, R. 2012. Improving neural networks by preventing co-adaptation of feature detectors. Technical report, arXiv:1207.0580.
- Hinton, G. E.; Osindero, S.; and Teh, Y. 2006. A fast learning algorithm for deep belief nets. *Neural Computation* 18:1527–1554.
- Ising, E. 1925. Beitrag zur Theorie des Ferromagnetismus. *Zeitschrift fur Physik* 31:253–258.
- Larochelle, H.; Bengio, Y.; and Turian, J. 2010. Tractable multivariate binary density estimation and the restricted Boltzmann forest. *Neural Computation* 22(9):2285–2307.
- LeCun, Y.; Bottou, L.; Bengio, Y.; and Haffner, P. 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE* 86(11):2278–2324.
- Long, P. M., and Servedio, R. A. 2010. Restricted Boltzmann machines are hard to approximately evaluate or simulate. In *Proceedings of the 27th International Conference on Machine Learning (ICML'10)*.
- Salakhutdinov, R., and Murray, I. 2008. On the quantitative analysis of deep belief networks. In Cohen, W. W.; McCallum, A.; and Roweis, S. T., eds., *ICML 2008*, volume 25, 872–879. ACM.
- Salakhutdinov, R. 2010a. Learning deep Boltzmann machines using adaptive MCMC. In Bottou, L., and Littman, M., eds., *Proceedings of the Twenty-seventh International Conference on Machine Learning (ICML-10)*, volume 1, 943–950. ACM.
- Salakhutdinov, R. 2010b. Learning in Markov random fields using tempered transitions. In *NIPS'2010*.
- Smolensky, P. 1986. Information processing in dynamical systems: Foundations of harmony theory. In Rumelhart, D. E., and McClelland, J. L., eds., *Parallel Distributed Processing*, volume 1. Cambridge: MIT Press. chapter 6, 194–281.
- Tieleman, T., and Hinton, G. 2009. Using fast weights to improve persistent contrastive divergence. In *ICML'2009*.
- Tieleman, T. 2008. Training restricted Boltzmann machines using approximations to the likelihood gradient. In Cohen, W. W.; McCallum, A.; and Roweis, S. T., eds., *ICML 2008*, 1064–1071. ACM.
- Younes, L. 1998. On the convergence of Markovian stochastic algorithms with rapidly decreasing ergodicity rates. In *Stochastics and Stochastics Models*, 177–228.