

- Ng, A. (2015). Advice for applying machine learning. <https://see.stanford.edu/materials/aimlcs229/ML-advice.pdf>. 421
- Niesler, T. R., Whittaker, E. W. D., and Woodland, P. C. (1998). Comparison of part-of-speech and automatically derived category-based language models for speech recognition. In *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 177–180. 463
- Ning, F., Delhomme, D., LeCun, Y., Piano, F., Bottou, L., and Barbano, P. E. (2005). Toward automatic phenotyping of developing embryos from videos. *Image Processing, IEEE Transactions on*, 14(9), 1360–1371. 360
- Nocedal, J. and Wright, S. (2006). *Numerical Optimization*. Springer. 92, 96
- Norouzi, M. and Fleet, D. J. (2011). Minimal loss hashing for compact binary codes. In *ICML’2011*. 525
- Nowlan, S. J. (1990). Competing experts: An experimental investigation of associative mixture models. Technical Report CRG-TR-90-5, University of Toronto. 450
- Nowlan, S. J. and Hinton, G. E. (1992). Simplifying neural networks by soft weight-sharing. *Neural Computation*, 4(4), 473–493. 139
- Olshausen, B. and Field, D. J. (2005). How close are we to understanding V1? *Neural Computation*, 17, 1665–1699. 16
- Olshausen, B. A. and Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381, 607–609. 147, 255, 370, 496
- Olshausen, B. A., Anderson, C. H., and Van Essen, D. C. (1993). A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *J. Neurosci.*, 13(11), 4700–4719. 450
- Opper, M. and Archambeau, C. (2009). The variational Gaussian approximation revisited. *Neural computation*, 21(3), 786–792. 689
- Oquab, M., Bottou, L., Laptev, I., and Sivic, J. (2014). Learning and transferring mid-level image representations using convolutional neural networks. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pages 1717–1724. IEEE. 536
- Osindero, S. and Hinton, G. E. (2008). Modeling image patches with a directed hierarchy of Markov random fields. In J. Platt, D. Koller, Y. Singer, and S. Roweis, editors, *Advances in Neural Information Processing Systems 20 (NIPS’07)*, pages 1121–1128, Cambridge, MA. MIT Press. 632
- Ovid and Martin, C. (2004). *Metamorphoses*. W.W. Norton. 1