

partially through grant NCRR 1S10RR022984-01A1.

References

- Bottou, L.; Peters, J.; Quiñero-Candela, J.; Charles, D. X.; Chickering, D. M.; Portugaly, E.; Ray, D.; Simard, P.; and Snelson, E. 2013. Counterfactual reasoning and learning systems: the example of computational advertising. *The Journal of Machine Learning Research* 14(1):3207–3260.
- Busa-Fekete, R.; Benbouzid, D.; Kégl, B.; et al. 2012. Fast classification using sparse decision dags. In *International Conference on Machine Learning*.
- Chen, M.; Weinberger, K. Q.; Chapelle, O.; Kedem, D.; and Xu, Z. 2012. Classifier cascade for minimizing feature evaluation cost. In *International Conference on Artificial Intelligence and Statistics*, 218–226.
- Dalal, N., and Triggs, B. 2005. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition*, volume 1, 886–893.
- Das, A., and Kempe, D. 2011. Submodular meets spectral: Greedy algorithms for subset selection, sparse approximation and dictionary selection. *International Conference on Machine Learning*.
- Das, A.; Dasgupta, A.; and Kumar, R. 2012. Selecting diverse features via spectral regularization. In *Advances in Neural Information Processing Systems*, 1592–1600.
- Diekhoff, G. 1992. *Statistics for the social and behavioral sciences: univariate, bivariate, multivariate*. Wm. C. Brown Publishers Dubuque, IA.
- Dredze, M.; Gevayahu, R.; and Elias-Bachrach, A. 2007. Learning fast classifiers for image spam. In *Conference on Email and Anti-Spam*.
- Efron, B.; Hastie, T.; Johnstone, I.; and Tibshirani, R. 2004. Least angle regression. *The Annals of Statistics* 32(2):407–499.
- Golovin, D., and Krause, A. 2011. Adaptive submodularity: Theory and applications in active learning and stochastic optimization. *Journal of Artificial Intelligence Research* 42(1):427–486.
- Grubb, A., and Bagnell, J. A. 2012. Speedboost: Anytime prediction with uniform near-optimality. In *International Conference on Artificial Intelligence and Statistics*.
- Hastie, T.; Tibshirani, R.; and Friedman, J. 2001. The elements of statistical learning: data mining, inference and prediction. *New York: Springer-Verlag* 1(8):371–406.
- Johnson, R. A., and Wichern, D. W. 2002. *Applied multivariate statistical analysis*, volume 5. Prentice hall Upper Saddle River, NJ.
- Kowalski, M. 2009. Sparse regression using mixed norms. *Applied and Computational Harmonic Analysis* 27(3):303–324.
- Krause, A., and Cevher, V. 2010. Submodular dictionary selection for sparse representation. In *International Conference on Machine Learning*, 567–574.
- Krause, A., and Guestrin, C. E. 2012. Near-optimal nonmyopic value of information in graphical models. *arXiv preprint arXiv:1207.1394*.
- Nemhauser, G. L.; Wolsey, L. A.; and Fisher, M. L. 1978. An analysis of approximations for maximizing submodular set functions i. *Mathematical Programming* 14(1):265–294.
- Streeter, M., and Golovin, D. 2007. An online algorithm for maximizing submodular functions. Technical report, DTIC Document.
- Wang, J., and Saligrama, V. 2012. Local supervised learning through space partitioning. In *Advances in Neural Information Processing Systems*, 91–99.
- Weinberger, K.; Dasgupta, A.; Langford, J.; Smola, A.; and Attenberg, J. 2009. Feature hashing for large scale multitask learning. In *International Conference on Machine Learning*, 1113–1120.
- Xu, Z.; Kusner, M. J.; Chen, M.; and Weinberger, K. Q. 2013a. Cost-sensitive tree of classifiers. In *International Conference on Machine Learning*.
- Xu, Z.; Kusner, M.; Huang, G.; and Weinberger, K. Q. 2013b. Anytime representation learning. In *International Conference on Machine Learning*.
- Xu, Z.; Kusner, M. J.; Weinberger, K. Q.; Chen, M.; and Chapelle, O. 2014. Budgeted learning with trees and cascades. *Journal of Machine Learning Research*.
- Zhang, D.; He, J.; Si, L.; and Lawrence, R. 2013. Mileage: Multiple instance learning with global embedding. *International Conference on Machine Learning*.
- Zheng, Z.; Zha, H.; Zhang, T.; Chapelle, O.; Chen, K.; and Sun, G. 2008. A general boosting method and its application to learning ranking functions for web search. In *Advances in Neural Information Processing Systems*. 1697–1704.
- Zolghadr, N.; Bartók, G.; Greiner, R.; György, A.; and Szepesvári, C. 2013. Online learning with costly features and labels. In *Advances in Neural Information Processing Systems*, 1241–1249.