

the-art algorithms in terms of accuracy and scalability.

Acknowledgements

This research is partially supported by the IARPA via DoD/ARL contract # W911NF-12-C-0012 and DARPA via contract # D11AP00278. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DoD/ARL, DARPA, or the U.S. Government.

References

- Argyriou, A.; Evgeniou, T.; and Pontil, M. 2008. Convex multi-task feature learning. *Mach. Learn.* 73(3):243–272.
- Bellet, A., and Habrard, A. 2012. Robustness and Generalization for Metric Learning. Technical report, arXiv:1209.1086.
- Bellet, A.; Habrard, A.; and Sebban, M. 2013. A Survey on Metric Learning for Feature Vectors and Structured Data. Technical report, arXiv:1306.6709.
- Blitzer, J.; Dredze, M.; and Pereira, F. 2007. Biographies, Bollywood, Boom-boxes and Blenders: Domain Adaptation for Sentiment Classification. In *ACL*.
- Caruana, R. 1997. Multitask Learning. *Mach. Learn.* 28(1):41–75.
- Davis, J.; Kulis, B.; Jain, P.; Sra, S.; and Dhillon, I. 2007. Information-theoretic metric learning. In *ICML*.
- Duchi, J., and Singer, Y. 2009. Efficient Online and Batch Learning Using Forward Backward Splitting. *JMLR* 10:2899–2934.
- Frome, A.; Singer, Y.; Sha, F.; and Malik, J. 2007. Learning Globally-Consistent Local Distance Functions for Shape-Based Image Retrieval and Classification. In *ICCV*.
- Goldberger, J.; Roweis, S.; Hinton, G.; and Salakhutdinov, R. 2004. Neighbourhood Components Analysis. In *NIPS*.
- Gong, P.; Ye, J.; and Zhang, C. 2012. Robust multi-task feature learning. In *KDD*.
- Hauberg, S.; Freifeld, O.; and Black, M. 2012. A Geometric take on Metric Learning. In *NIPS*.
- Hong, Y.; Li, Q.; Jiang, J.; and Tu, Z. 2011. Learning a mixture of sparse distance metrics for classification and dimensionality reduction. In *CVPR*.
- Jain, P.; Kulis, B.; Dhillon, I.; and Grauman, K. 2008. Online Metric Learning and Fast Similarity Search. In *NIPS*.
- Kulis, B. 2012. Metric Learning: A Survey. *Foundations and Trends in Machine Learning* 5(4):287–364.
- Noh, Y.-K.; Zhang, B.-T.; and Lee, D. 2010. Generative Local Metric Learning for Nearest Neighbor Classification. In *NIPS*.
- Parameswaran, S., and Weinberger, K. 2010. Large Margin Multi-Task Metric Learning. In *NIPS*.
- Ramanan, D., and Baker, S. 2011. Local Distance Functions: A Taxonomy, New Algorithms, and an Evaluation. *TPAMI* 33(4):794–806.
- Schölkopf, B.; Smola, A.; and Müller, K.-R. 1998. Non-linear component analysis as a kernel eigenvalue problem. *Neural Comput.* 10(1):1299–1319.
- Shen, C.; Kim, J.; Wang, L.; and van den Hengel, A. 2012. Positive Semidefinite Metric Learning Using Boosting-like Algorithms. *JMLR* 13:1007–1036.
- Shi, Y.; Bellet, A.; and Sha, F. 2014. Sparse Compositional Metric Learning. Technical report, arXiv:1404.4105. <http://arxiv.org/pdf/1404.4105>.
- van der Maaten, L., and Hinton, G. 2008. Visualizing Data using t-SNE. *JMLR* 9:2579–2605.
- Wang, J.; Woznica, A.; and Kalousis, A. 2012. Parametric Local Metric Learning for Nearest Neighbor Classification. In *NIPS*.
- Weinberger, K., and Saul, L. 2009. Distance Metric Learning for Large Margin Nearest Neighbor Classification. *JMLR* 10:207–244.
- Xiao, L. 2010. Dual Averaging Methods for Regularized Stochastic Learning and Online Optimization. *JMLR* 11:2543–2596.
- Xing, E.; Ng, A.; Jordan, M.; and Russell, S. 2002. Distance Metric Learning with Application to Clustering with Side-Information. In *NIPS*.
- Xu, H., and Mannor, S. 2012. Robustness and Generalization. *Mach. Learn.* 86(3):391–423.
- Yang, X.; Kim, S.; and Xing, E. 2009. Heterogeneous multitask learning with joint sparsity constraints. In *NIPS*.
- Ying, Y., and Li, P. 2012. Distance Metric Learning with Eigenvalue Optimization. *JMLR* 13:1–26.
- Yuan, M., and Lin, Y. 2006. Model selection and estimation in regression with grouped variables. *J. Roy. Statist. Soc. Ser. B* 68:49–67.
- Zhan, D.-C.; Li, M.; Li, Y.-F.; and Zhou, Z.-H. 2009. Learning instance specific distances using metric propagation. In *ICML*.