

Machine Learning from Weak Supervision

Masashi Sugiyama, Marthinus Christoffel du Plessis, Gang Niu,
Yao Ma, and Tomoya Sakai

The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

E-mail: sugi@k.u-tokyo.ac.jp

Keyword: Machine learning, Supervised learning, Unsupervised learning, Semi-supervised learning

Machine learning from big training data is making a great success in various real-world applications such as speech, image, and natural language processing. However, there are various application domains that prohibit the use of massive labeled data. In this talk, I will introduce our recent advances in machine learning from weak supervision, including unsupervised classification [1], classification only from positive and unlabeled data [2, 3, 4, 5, 6], and semi-supervised learning [7].

References

- [1] du Plessis, M. C., Niu, G., & Sugiyama, M. Clustering unclustered data: Unsupervised binary labeling of two datasets having different class balances. In *Proceedings of Conference on Technologies and Applications of Artificial Intelligence (TAAI2013)*, pp.1–6, Taipei, Taiwan, Dec. 6–8, 2013.
- [2] du Plessis, M. C., Niu, G., & Sugiyama, M. Analysis of learning from positive and unlabeled data. In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems 27*, pp.703–711, 2014. (Presented at *Neural Information Processing Systems (NIPS2014)*, Montreal, Quebec, Canada, Dec. 8–11, 2014)
- [3] du Plessis, M. C., Niu, G., & Sugiyama, M. Convex formulation for learning from positive and unlabeled data. In F. Bach and D. Blei (Eds.), *Proceedings of 32nd International Conference on Machine Learning (ICML2015)*, *JMLR Workshop and Conference Proceedings*, vol.37, pp.1386–1394, Lille, France, Jul. 6–11, 2015.
- [4] Niu, G., du Plessis, M. C., Sakai, T., Ma, Y., & Sugiyama, M. Theoretical comparisons of positive-unlabeled learning against positive-negative learning. *arXiv:1603.03130*, 2016.
- [5] du Plessis, M. C. & Sugiyama, M. Class prior estimation from positive and unlabeled data. *IEICE Transactions on Information and Systems*, vol.E97-D, no.5, pp.1358–1362, 2014.
- [6] du Plessis, M. C., Niu, G., & Sugiyama, M. Class-prior estimation for learning from positive and unlabeled data. *Proceedings of the 7th Asian Conference on Machine Learning (ACML2015)*, *JMLR Workshop and Conference Proceedings*, vol.45, pp.221–236, Hong Kong, China, Nov. 20–22, 2015.
- [7] Sakai, T., du Plessis, M. C., Niu, G., & Sugiyama, M. Beyond the low-density separation principle: A novel approach to semi-supervised learning. *arXiv:1605.06955*, 2016.