General framework for binary nonlinear classification on top samples

Václav Mácha[1], Lukáš Adam[2,1], Václav Šmídl[1]

[1] Institute of Information Theory and Automation, Czech Academy of Sciences, Prague, Czech Republic
[2] Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen, China
macha@utia.cas.cz, adam@utia.cas.cz, smidl@utia.cas.cz,

Keywords. Binary classification; duality; kernels; accuracy at the top; ranking; hypothesis testing.

In our previous work [1], we have proposed a general framework to handle binary linear classification for top samples. Our framework includes ranking problems, accuracy at the top or hypothesis testing. We have summarized known methods, such as [2, 3, 4], belonging to this framework and proposed new ones. Note that these methods were either derived in their primal form, or they did not use kernels. This forced a restriction on only linear classifiers.

In this work, we employ the convexity results derived in [1]. For all methods from our framework, we derive their equivalent dual formulation. We utilize their SVM-like structure and incorporate kernels into the dual formulation. This allows us to pass from linear to nonlinear classifiers. We show how to recover the primal solution and classification value for new samples. We propose an effective computation method and perform a numerical analysis showing the efficiency of our framework.

Acknowledgements. This work was supported by the Grant Agency of the Czech Republic (Grant No. 18-21409S), by the Program for Guangdong Introducing Innovative and Enterpreneurial Teams (Grant No. 2017ZT07X386) and by the Shenzhen Peacock Plan (Grant No. KQTD2016112514355531).

References

[1] L. Adam, V. Mácha, V. Šmídl, and T. Pevný. General framework for binary classification on top samples. *Submitted*, 2019.
[2] S. Boyd, C. Cortes, M. Mohri, and A. Radovanovic. Accuracy at the top. In *Advances in Neural Information Processing Systems*, pages 953–961, 2012.
[3] M. Grill and T. Pevný. Learning combination of anomaly detectors for security domain. *Computer Networks*, 107:55–63, 2016.
[4] N. Li, R. Jin, and Z.-H. Zhou. Top rank optimization in linear time. In *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 1*, NIPS'14, pages 1502–1510, Cambridge, MA, USA, 2014. MIT Press.