Robust AUC Maximization for Classification with Pairwise Confidence Comparisons

Haochen SHI, Mingkun XIE, Shengjun HUANG

Frontiers of Computer Science, DOI: 10.1007/s11704-023-2709-5

Problems & Ideas

- Problems of traditional supervised learning:
  - Traditional supervised learning requires a large number of labeled examples to train the classification model.
  - It is difficult and costly to annotate a huge number of unlabeled examples in many real-world scenarios.
- Ideas: To mitigate this issue, we focus on a recently proposed weakly-supervised learning framework called pairwise comparison (Pcomp) classification, and propose a novel learning framework called PC-AUC, which performs AUC maximization to avoid the class-imbalance issue.

Theorem 4.1. *The classification risk Eq.*(4) *can be equivalently re-written as*

$$R(f) = \frac{1}{\pi_{10}}\tilde{R}(f) - \frac{\pi_{11}}{2\pi_{10}} - \frac{\pi_{00}}{2\pi_{10}},$$

Theorem 5.1. *Learning from Pcomp data with the pairwise surrogate loss Φ is consistent w.r.t. AUC, if ϕ is a convex, differential and non-increasing function with $\phi'(0) < 0$.*

Main results proposed in the paper showing the unbiasedness and consistency of PC-AUC.

Main Contributions

- Contributions:
  - A novel Pcomp-AUC (PC-AUC for short) framework for robust AUC maximization by minimizing pairwise surrogate loss with only Pcomp data;
  - Disclosing the robustness of PC-AUC maximization, i.e., maximizing PC-AUC is equivalent to maximizing AUC;
  - Showing the consistency of learning from Pcomp data, i.e., the learner under Pcomp data approximating the optimal Bayes learner.

Class Prior	Datasets	PC-AUC	Pcomp-Unbiased	Pcomp-ReLU	Noisy-Unbiased	Binary-Biased
$\pi_1 = 0.1$	MNIST	**96.69 ± 0.28**	93.28 ± 0.29	93.25 ± 0.43	90.32 ± 0.32	57.46 ± 3.86
	Kuzushiji	**82.70 ± 1.00**	78.61 ± 0.99	78.50 ± 1.11	74.52 ± 0.68	50.95 ± 0.29
	Fashion	**98.70 ± 0.13**	97.69 ± 0.11	98.23 ± 0.03	98.28 ± 0.12	68.77 ± 4.36
	CIFAR10	**79.49 ± 1.02**	62.77 ± 1.10	78.44 ± 0.49	62.13 ± 0.86	51.92 ± 0.72
$\pi_1 = 0.2$	MNIST	**97.75 ± 0.09**	93.50 ± 0.23	93.44 ± 0.18	94.37 ± 0.23	54.98 ± 0.89
	Kuzushiji	**85.94 ± 0.72**	79.36 ± 0.83	79.35 ± 0.60	80.28 ± 0.55	52.04 ± 1.94
	Fashion	**99.05 ± 0.04**	97.46 ± 0.08	98.20 ± 0.03	98.67 ± 0.08	64.90 ± 2.07
	CIFAR10	**81.39 ± 0.62**	68.53 ± 1.16	80.66 ± 0.59	65.65 ± 1.75	51.52 ± 0.53
$\pi_1 = 0.3$	MNIST	**97.89 ± 0.13**	93.20 ± 0.16	93.06 ± 0.22	95.31 ± 0.20	57.61 ± 1.74
	Kuzushiji	**86.65 ± 0.38**	79.28 ± 0.38	78.27 ± 0.45	82.40 ± 0.22	51.27 ± 0.86
	Fashion	**99.07 ± 0.03**	97.21 ± 0.08	98.06 ± 0.04	98.68 ± 0.09	65.07 ± 2.17
	CIFAR10	**82.41 ± 0.37**	71.21 ± 1.09	80.62 ± 0.25	69.40 ± 1.00	52.54 ± 0.78
$\pi_1 = 0.4$	MNIST	**97.90 ± 0.13**	93.31 ± 0.30	92.97 ± 0.28	95.74 ± 0.14	54.94 ± 1.49
	Kuzushiji	**87.10 ± 0.49**	77.90 ± 0.67	76.52 ± 0.51	82.71 ± 0.48	50.36 ± 0.22
	Fashion	**99.07 ± 0.05**	97.32 ± 0.13	98.01 ± 0.06	98.77 ± 0.09	63.66 ± 2.12
	CIFAR10	**82.69 ± 0.23**	72.11 ± 1.71	80.37 ± 0.49	74.21 ± 1.17	52.37 ± 1.03
$\pi_1 = 0.5$	MNIST	**98.03 ± 0.13**	93.69 ± 0.17	93.34 ± 0.22	96.13 ± 0.13	56.32 ± 1.82
	Kuzushiji	**86.64 ± 0.40**	76.16 ± 0.62	74.64 ± 0.60	81.92 ± 0.51	50.86 ± 1.05
	Fashion	**99.01 ± 0.03**	97.30 ± 0.07	97.89 ± 0.04	98.71 ± 0.04	64.61 ± 2.15
	CIFAR10	**83.66 ± 0.60**	72.44 ± 1.28	80.38 ± 0.75	72.63 ± 1.75	52.36 ± 0.94

AUC of each comparing method on four benchmark image classification datasets with different class prior probabilities.