IEEE TRANSACTIONS ON
PATTERN ANALYSIS AND
MACHINE INTELLIGENCE

# Weakly Supervised AUC Optimization: A Unified Partial AUC Approach

**Zheng Xie, Yu Liu, Hao-Yuan He,**
**Ming Li,** and **Zhi-Hua Zhou**

National Key Laboratory for Novel Software Technology
Nanjing University
{xiez, liuy, hehy, lim, zhouzh}@lamda.nju.edu.cn

## 1. A Unified Weakly Supervised AUC Optimization Framework

In this paper, we propose WSAUC, a unified weakly supervised AUC optimization framework. It solves the AUC optimization problems in the following scenarios:

1. Label Incomplete (e.g., semi-supervised, positive-unlabeled, …),
2. Label Inaccurate (e.g., noisy labeled, crowd sourced label, …),
3. Label Inexact (e.g., multi-instance, …),

and even combination of these scenarios.

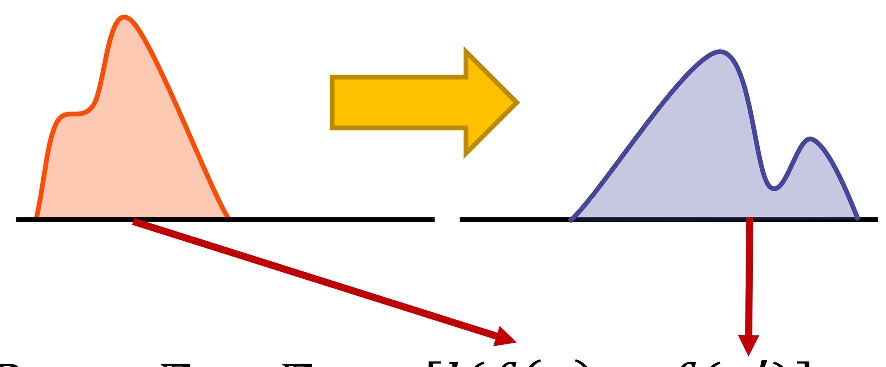
WSAUC framework consists of two key steps:

1. convert the optimization problem into a unified formulation,
2. achieve robust AUC optimization by minimizing rpAUC risk.

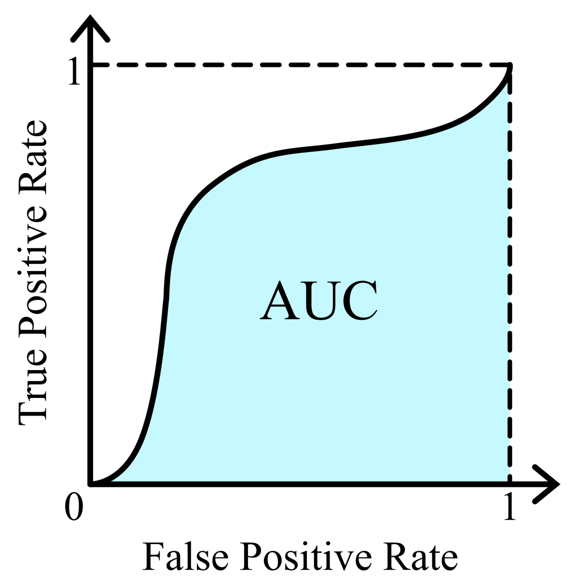
Thanks to these two steps, WSAUC can solve multiple weakly supervised AUC optimization problem in a unified way, and achieve superior learning performance on these tasks.

### Preliminary: AUC Optimization

Optimizing AUC is equivalent to maximizing the probability of the model ranking instances from the positive distribution P before those from the negative distribution N.



$$R_{PN} = \mathbb{E}_{x \sim P}\mathbb{E}_{x' \sim N}[l(f(x) - f(x'))]$$




## 2. The Unified Formulation

We show that **different weakly supervised scenarios can be converted into a unified formulation**:

$$R_{AB} = aR_{PN} + b$$

$a$: coefficient $R_{PN}$: true risk $b$: bias term

$A$: $\theta_A$ positive + $(1-\theta_A)$ negative

$B$: $\theta_B$ positive + $(1-\theta_B)$ negative

Different from the original AUC optimization formulation, this formulation can be regarded as maximizing AUC from two contaminated instance sets A and B, both of which contains positive and negative instances of different proportions.

### Weakly Supervised AUC Optimization Cases

**Noisy label AUC optimization**. By regarding the noisy positive and negative sets as A and B, we have:

$$R_{\tilde{P}\tilde{N}} = (1 - \eta_P - \eta_N)R_{PN} + \frac{\eta_P + \eta_N}{2}$$

$\eta_P$: positive noisy rate $\eta_N$: negative noisy rate

**Positive-unlabeled AUC optimization**. By regarding the positive and unlabeled sets as A and B, we have:

$$R_{PU} = \pi_N R_{PN} + \frac{\pi_P}{2}$$

$\pi_P$: positive prior $\pi_N$: negative prior

**Multi-instance AUC optimization**. By regarding the positive bag instance as A and negative bag instance as B, we have:

$$R_{\tilde{P}N} = (1 - \eta_P)R_{PN} + \frac{\eta_P}{2}$$

**Semi-supervised AUC optimization**. We can construct the following risk formulation:

$$R_{PU} + R_{UN} - \frac{1}{2} = R_{PN}$$

## 3. rpAUC as Robust AUC Optimization Objective

We propose **rpAUC** (reversed partial AUC) to **mitigate the negative impact of the contamination in the unified formulation**. The core idea is that optimizing rpAUC is equivalent as removing noisy pairs during training, and thus yield similar effect as the well-known "small-loss trick".

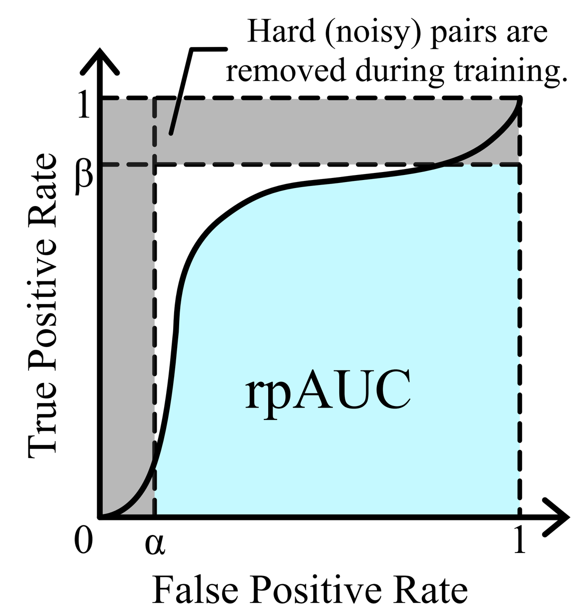
**Definition 11.** *Two-way Reverse Partial AUC with FPR threshold $\alpha$ and TPR threshold $\beta$ of model $f$ can be defined as:*

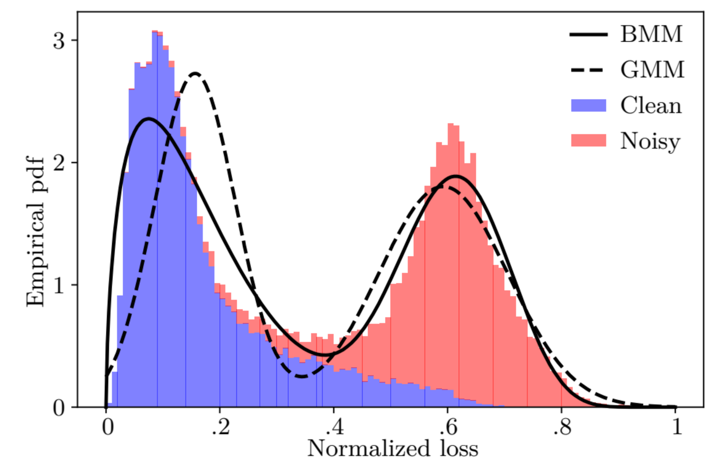
$$\mathrm{rpAUC}(f;\alpha,\beta) = 1 - \mathop{\mathbb{E}}_{\boldsymbol{x}\sim p_A^+(\boldsymbol{x})}\left[\mathop{\mathbb{E}}_{\boldsymbol{x}'\sim p_B^-(\boldsymbol{x})}[\ell_{01}(f(\boldsymbol{x},\boldsymbol{x}'))]\right],$$

$$\text{where } p_A^+(\boldsymbol{x}) = p_A(\boldsymbol{x}|f(\boldsymbol{x}) \in [\mathrm{TPR}_f^{-1}(\beta),\infty)),$$

$$p_B^-(\boldsymbol{x}) = p_B(\boldsymbol{x}|f(\boldsymbol{x}) \in (-\infty,\mathrm{FPR}_f^{-1}(\alpha)]).$$

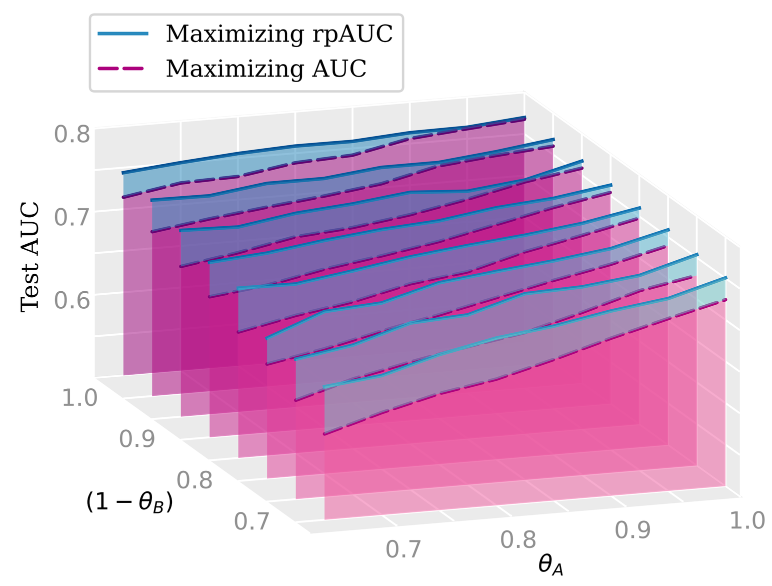
### rpAUC Illustration




[Arazo et al., ICML'19]

**Left**: rpAUC cuts out the left and upper margin of AUC, so that those pairs with largest losses are removed from training. **Right**: Noisy labels yield large loss during training.




**rpAUC vs AUC as training objective:** The blue shades indicate the relative improvement achieved by using rpAUC. The more contaminated the dataset is, the greater the improvement achieved.

## 4. Conclusions

In this paper, we show that:

1. different weakly supervised AUC optimization problem can be converted into a unified formulation,
2. rpAUC can be used as a robust training objective for (contaminated) AUC optimization.

Based on these findings, we propose WSAUC framework that can solve weakly supervised AUC optimization under different scenarios.

There are comprehensive theoretical and empirical analysis in the paper. For more details, please scan the OR code to download the full paper PDF.