ICML | 2022
Thirty-ninth International
Conference on Machine Learning

GenLabel: Mixup Relabeling using Generative Models

Jy-yong Sohn, Liang Shang, Hongxu Chen,
Jaekyun Moon (KAIST), Dimitris Papailiopoulos, Kangwook Lee

WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

KAIST

Preview

- **Goal:** Data augmentation (DA) for robust ML
- **Motivation:** Classifiers are brittle to adversarial attacks

- **Key results:**

Mixup + **GenLabel** ➡ Margin, Accuracy, Robustness ↑

(A famous DA)

Preliminary: Mixup

cat dog

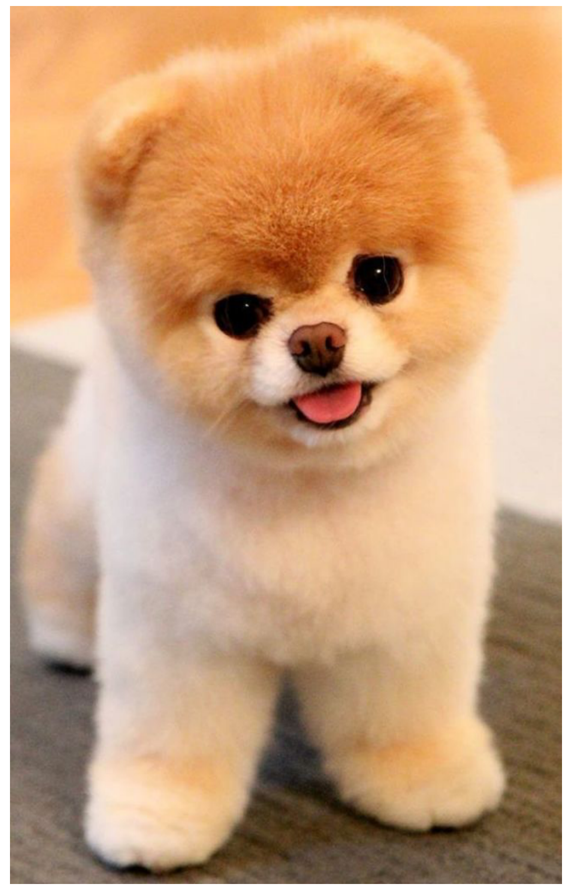
[1 0]

Feature $\boldsymbol{x}$

Label $\boldsymbol{y}$

cat dog

[0 1]



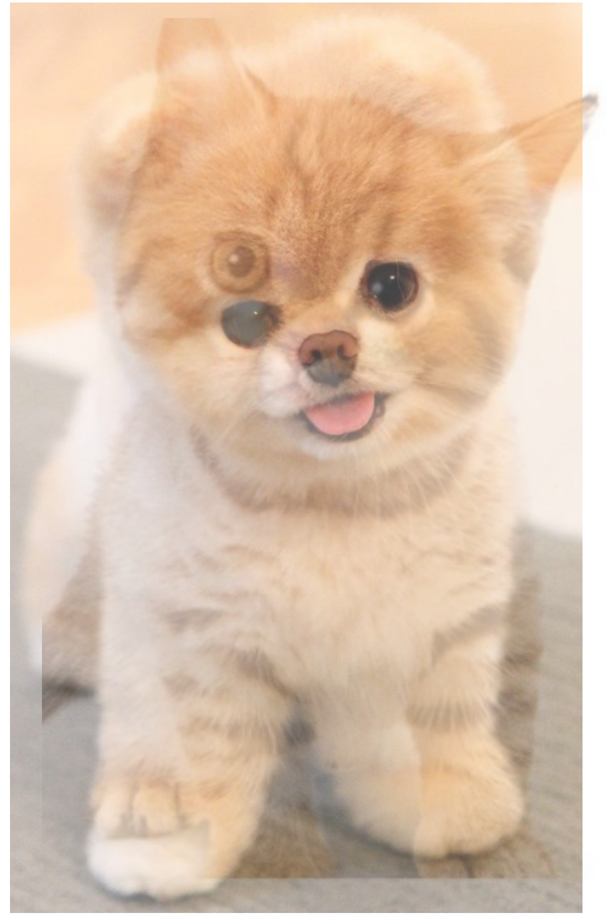
Feature $\boldsymbol{x}'$

Label $\boldsymbol{y}'$

Preliminary: Mixup

cat dog
[1 0]

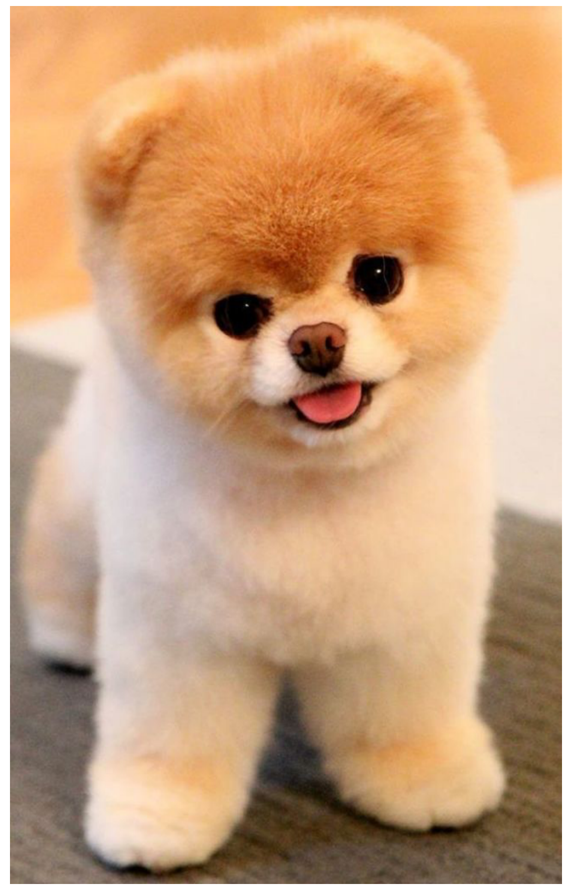
Feature $\boldsymbol{x}$
Label $\boldsymbol{y}$



cat dog
[0.5 0.5]

Feature $\boldsymbol{x}^{\text{mix}} = \lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{x}'$
Label $\boldsymbol{y}^{\text{mix}} = \lambda \boldsymbol{y} + (1-\lambda)\boldsymbol{y}'$

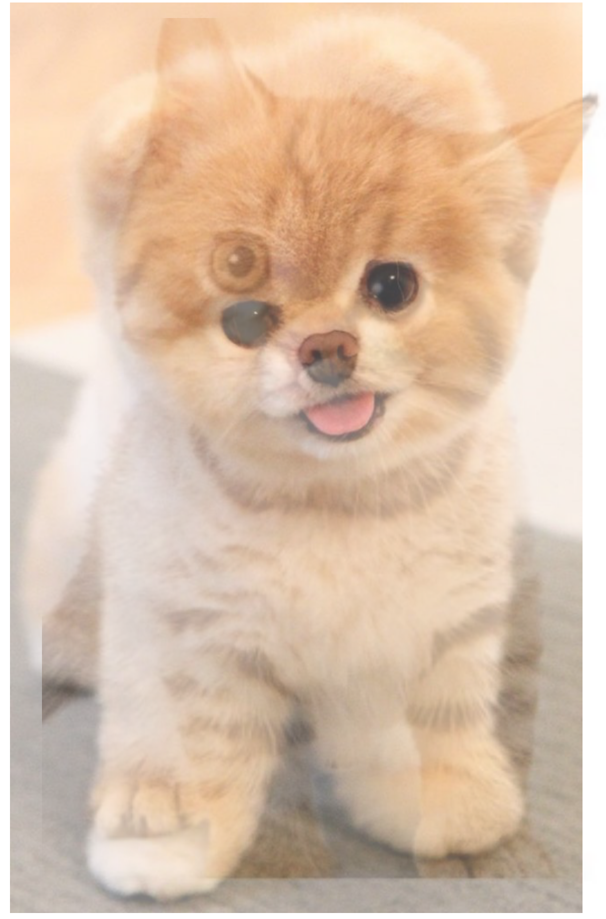
cat dog
[0 1]



Feature $\boldsymbol{x}'$
Label $\boldsymbol{y}'$

Mixup: Convex combination in feature & label domain

Preliminary: Mixup



cat dog

[0.5 0.5]

cat dog

[1 0]

Feature $\boldsymbol{x}$

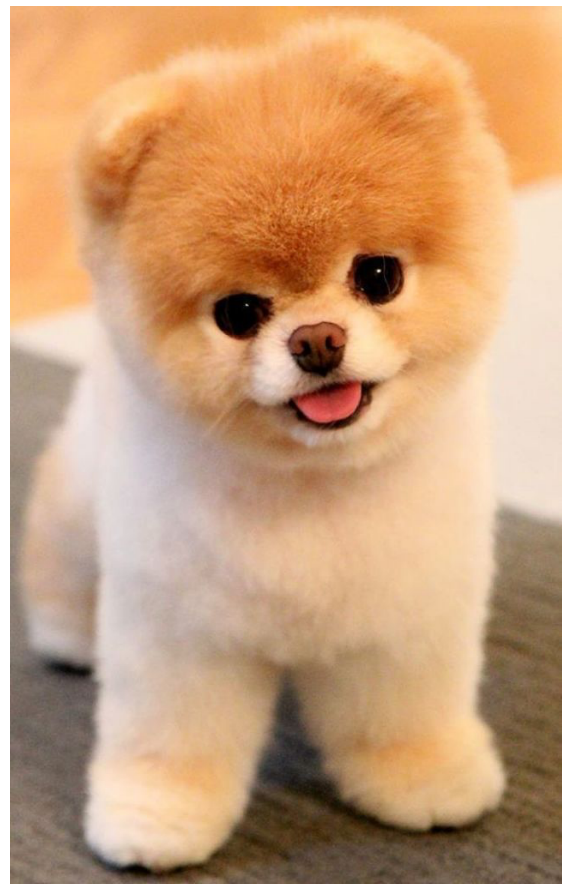
Label $\boldsymbol{y}$

$$\text{Feature } \boldsymbol{x}^{\text{mix}} = \lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{x}'$$

$$\text{Label } \boldsymbol{y}^{\text{mix}} = \lambda \boldsymbol{y} + (1-\lambda)\boldsymbol{y}'$$

cat dog

[0 1]

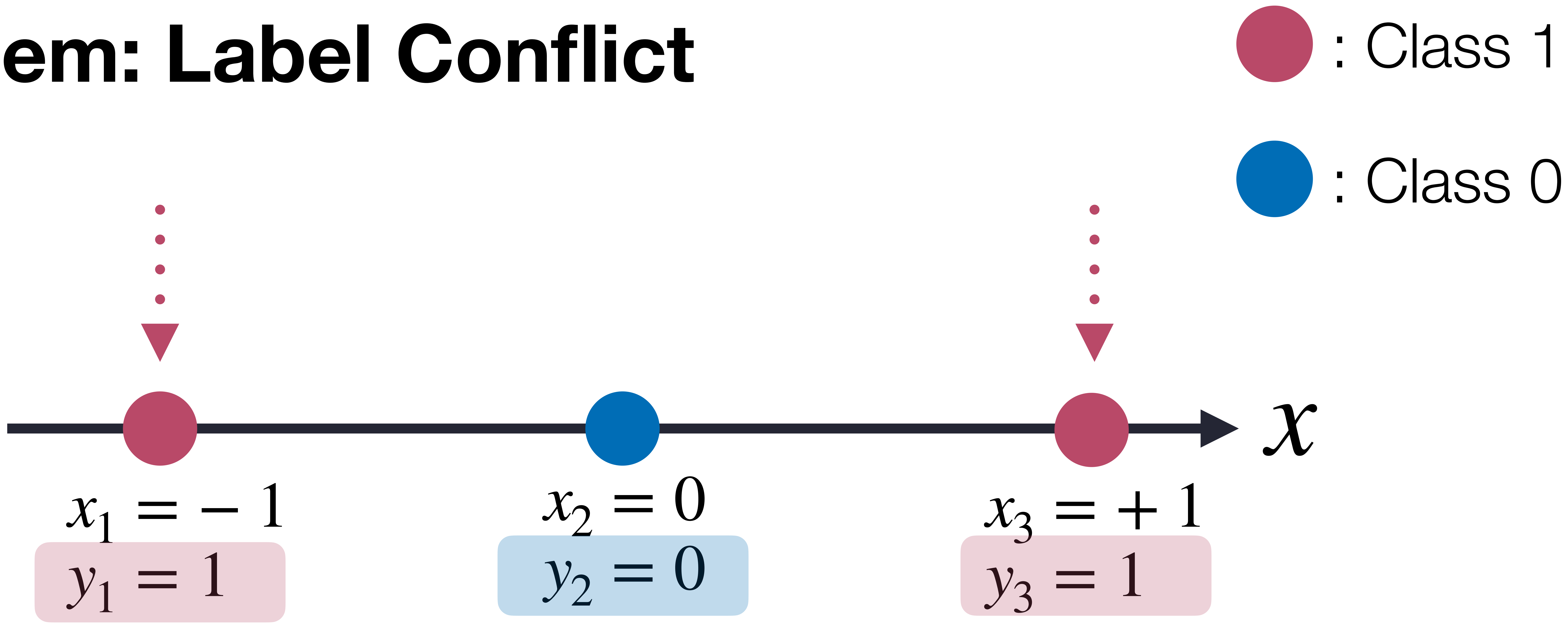


Feature $\boldsymbol{x}'$

Label $\boldsymbol{y}'$

Train with mixup sample improves accuracy/robustness

Problem: Label Conflict

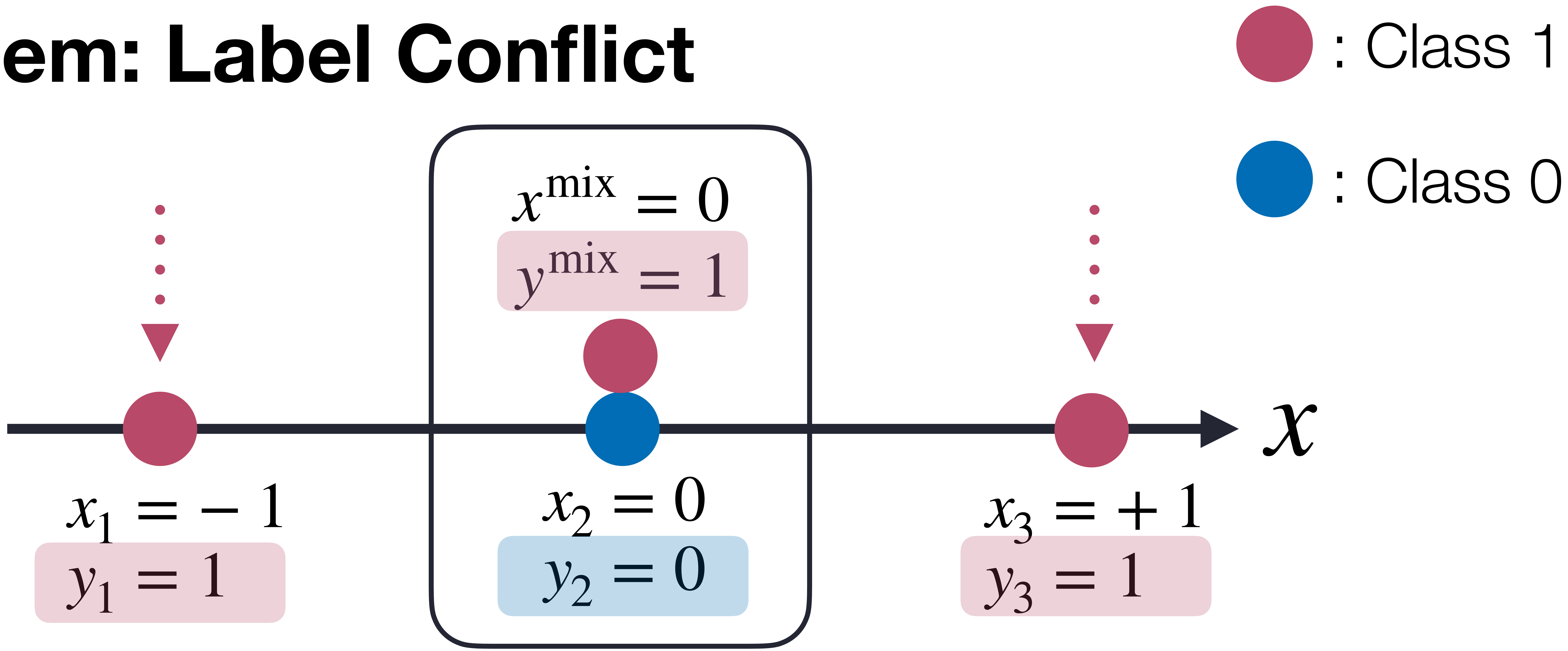


Mixing (x_1, y_1) and (x_3, y_3) generates

$$x^{\text{mix}} = 0.5x_1 + 0.5x_3 = 0$$

$$y^{\text{mix}} = 0.5y_1 + 0.5y_3 = 1$$

Problem: Label Conflict



Mixing (x_1, y_1) and (x_3, y_3) generates

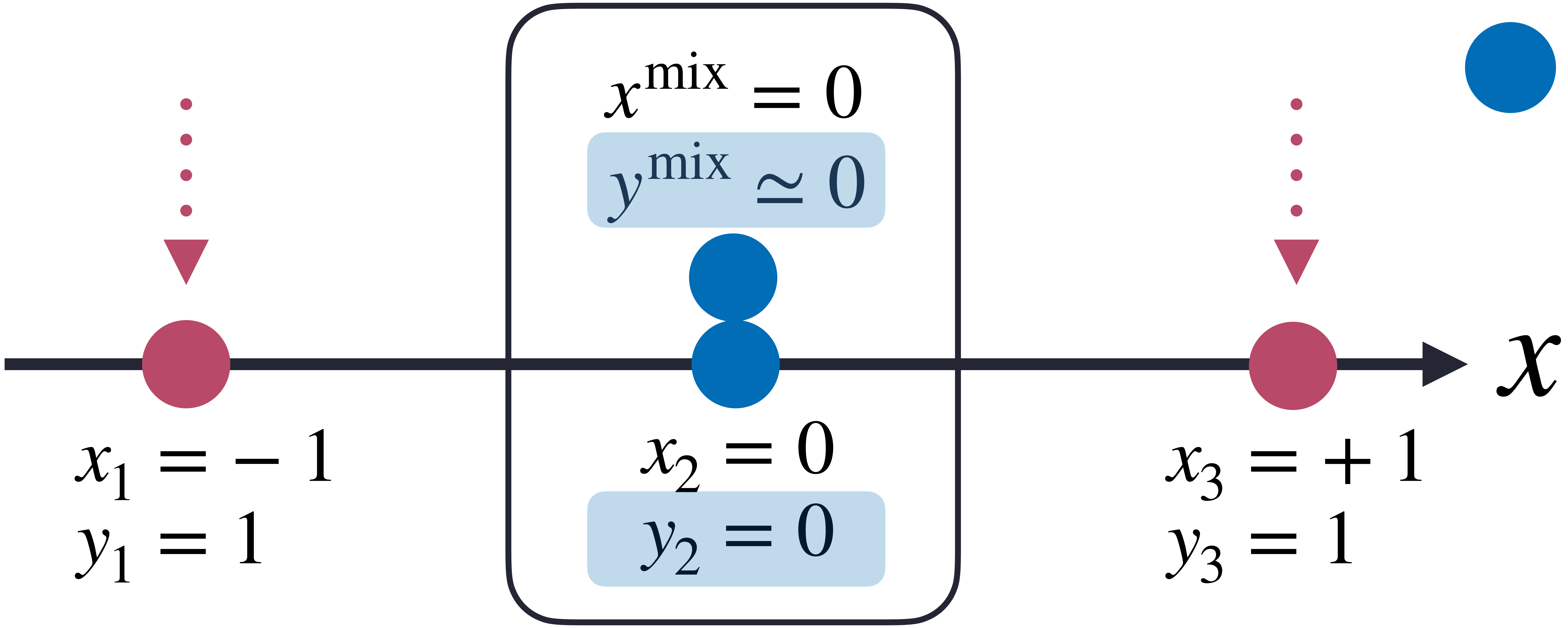
$$x^{\text{mix}} = 0.5x_1 + 0.5x_3 = 0$$

$$y^{\text{mix}} = 0.5y_1 + 0.5y_3 = 1$$

Our Solution: Re-Label

: Class 1

: Class 0

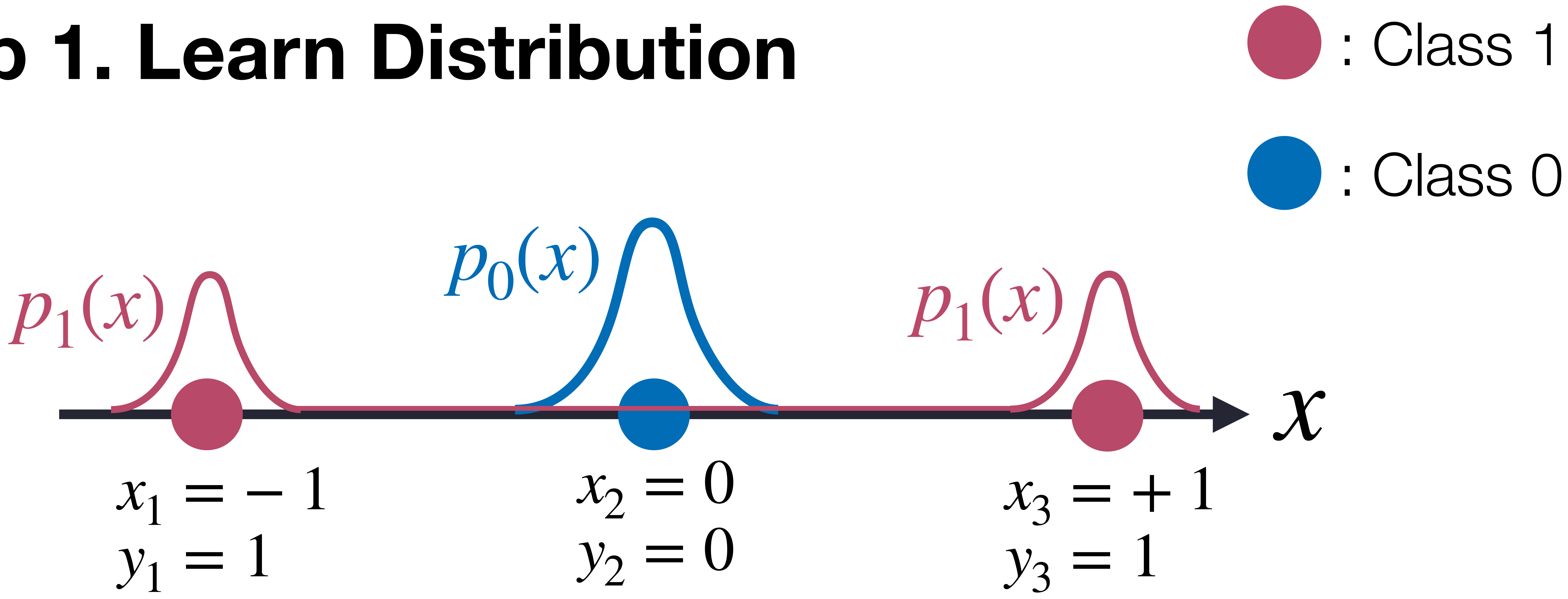


Mixing (x_1, y_1) and (x_3, y_3) generates

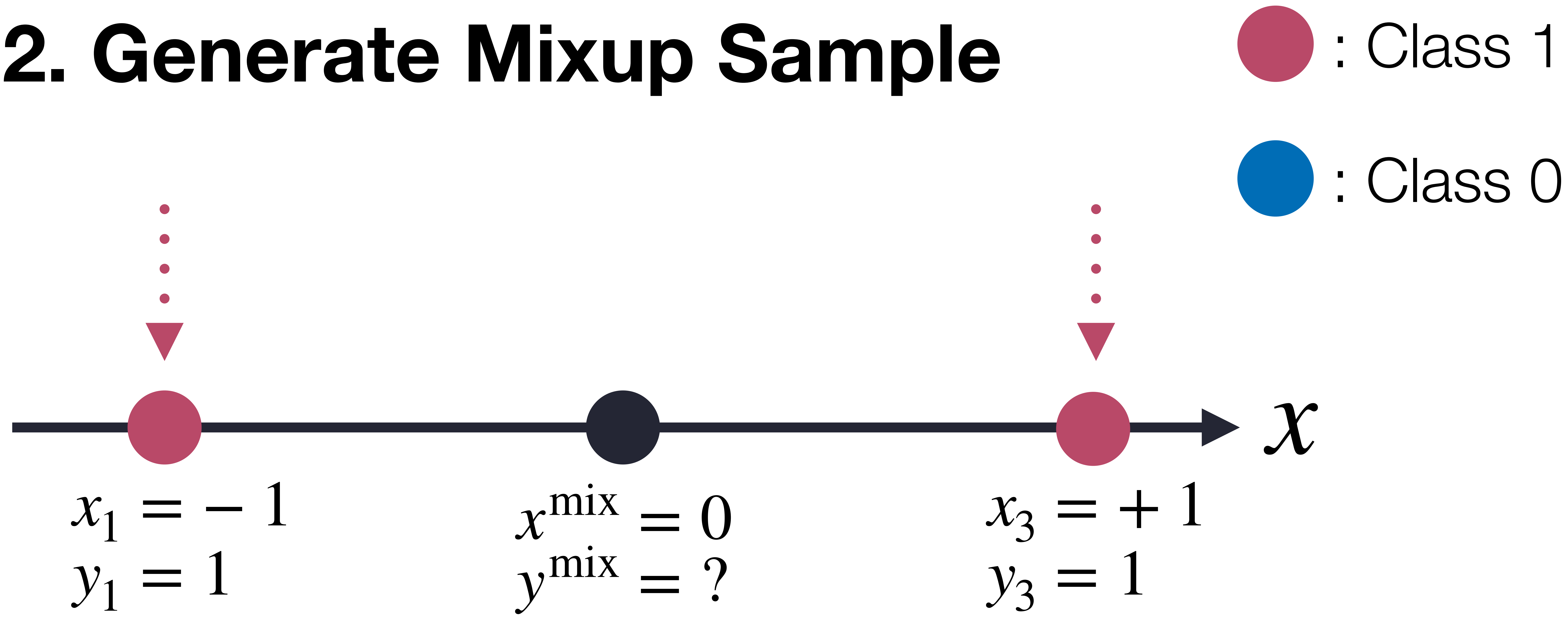
$$x^{\text{mix}} = 0.5x_1 + 0.5x_3 = 0$$

$$y^{\text{mix}} = \cancel{0.5y_1 + 0.5y_3 = 1}$$

Step 1. Learn Distribution



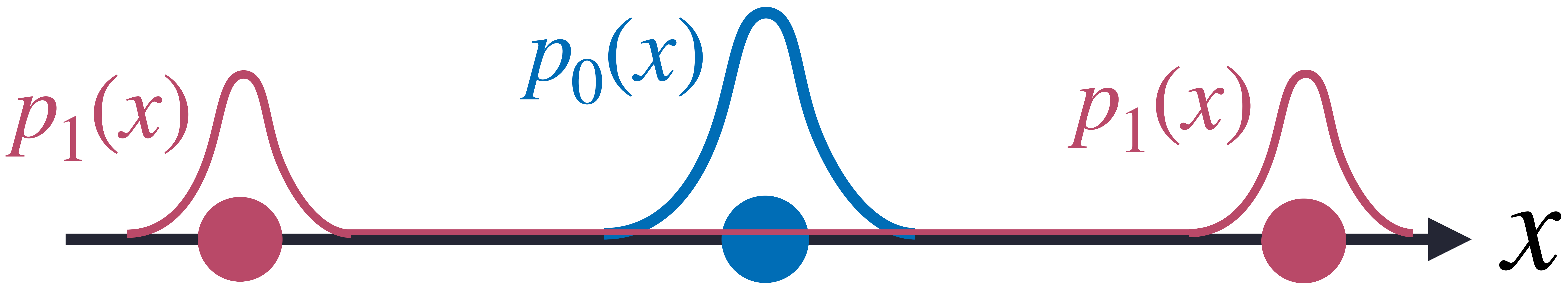
Step 2. Generate Mixup Sample



Step 3. Label Mixup Sample

: Class 1

: Class 0



$x_1 = -1$
$y_1 = 1$

$x^{\text{mix}} = 0$

$x_3 = +1$
$y_3 = 1$

$p_0(x^{\text{mix}}) \gg p_1(x^{\text{mix}})$

$$y^{\text{mix}} = \frac{p_1(x^{\text{mix}})}{p_0(x^{\text{mix}}) + p_1(x^{\text{mix}})} \cdot 1 \simeq 0$$

Step 3. Label Mixup Sample

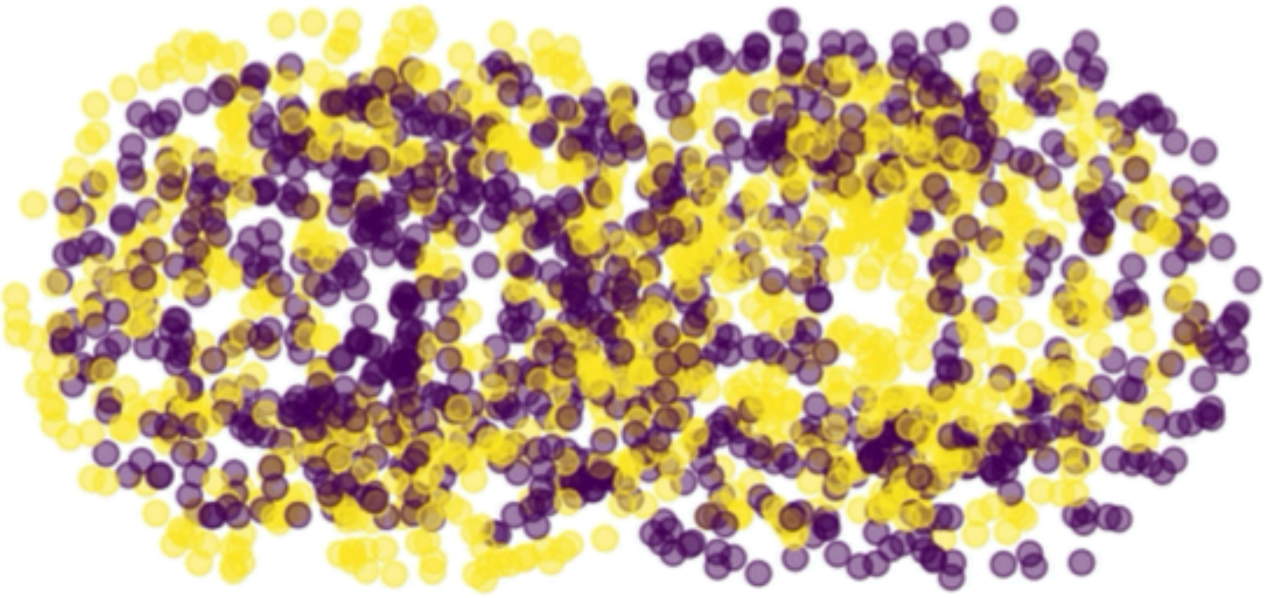
GenLabel
(Generative Model-based Labeling)

$p_0(x^{\text{mix}}) \gg p_1(x^{\text{mix}})$

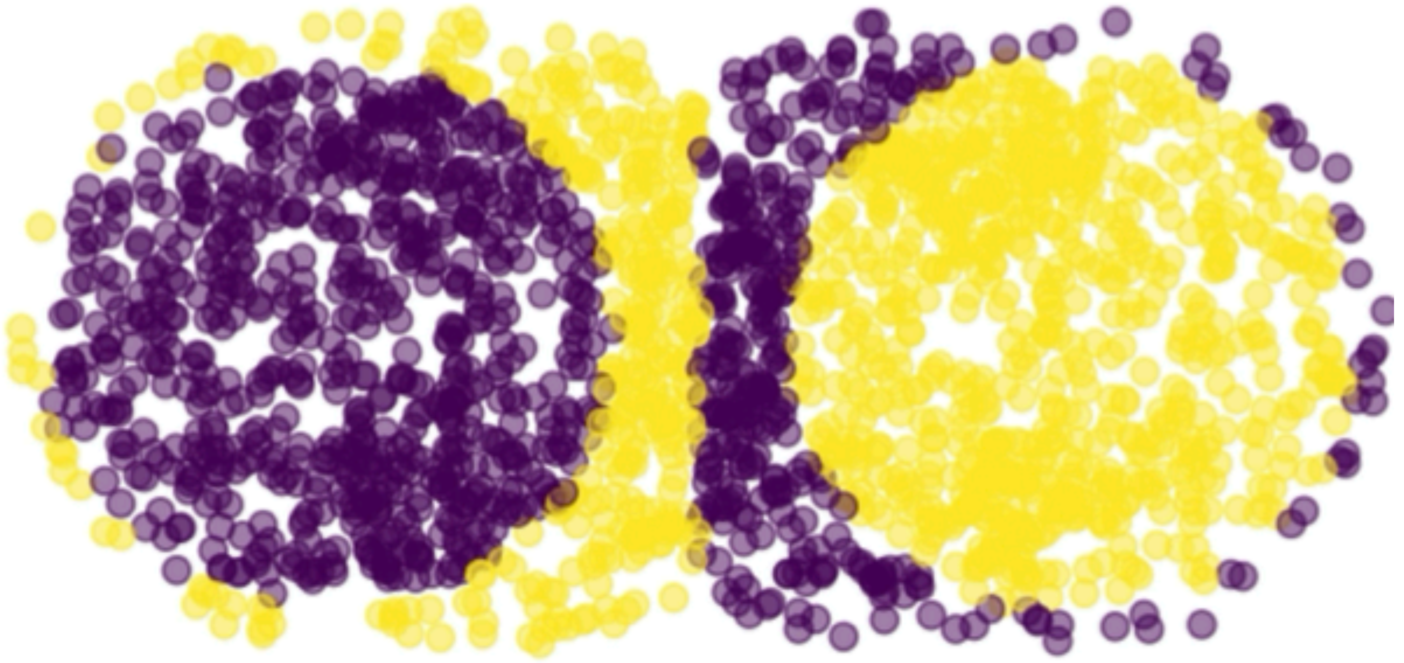
$$y^{\text{mix}} = \frac{p_1(x^{\text{mix}})}{p_0(x^{\text{mix}}) + p_1(x^{\text{mix}})} \cdot 1 \simeq 0$$

Key Results: Margin

Dataset



Top-1 label of **mixup**



Top-1 label of **mixup+GenLabel**

Decision boundary of **mixup**

Decision boundary of **mixup+GenLabel**

Key Results: Accuracy

Methods \ OpenML Dataset ID	721	777	792	830	855	913	1413	1498
Vanilla	79.67	58.67	73.20	77.60	63.33	70.80	95.56	66.91
AdaMixup	80.33	**64.00**	73.87	78.40	66.67	70.53	92.44	66.76
Mixup	79.33	62.67	73.47	76.27	66.00	69.87	88.00	66.76
Mixup + Excluding MI	79.67	62.67	74.53	78.13	66.40	71.47	93.33	66.33
Mixup + GenLabel (GM)	**81.00**	58.67	75.47	**86.13**	66.40	71.47	96.00	67.63
Mixup + GenLabel (KDE)	79.67	58.67	**75.87**	77.33	**67.60**	72.67	96.00	66.33
Mixup + GenLabel (CV)	80.33	**64.00**	75.60	84.53	**67.33**	**73.20**	**96.44**	**67.77**

GenLabel improves accuracy of mixup up to 8 – 10%

Key Results: Robustness

Methods \ OpenML ID	446	468	683	755	763	1413
Vanilla	29.67	34.55	51.11	41.05	64.27	68.00
AdaMixup	30.33	37.27	51.11	37.89	63.20	67.11
Mixup	30.67	37.27	50.00	36.84	65.07	67.56
Mixup + Excluding MI	31.67	31.82	**52.22**	38.95	63.20	70.67
Mixup + GenLabel (GM)	37.00	**42.73**	**52.22**	**43.16**	61.87	71.11
Mixup + GenLabel (NN)	**38.00**	32.73	46.67	**43.16**	**66.93**	**77.33**

GenLabel improves robustness of mixup up to 7 – 10%

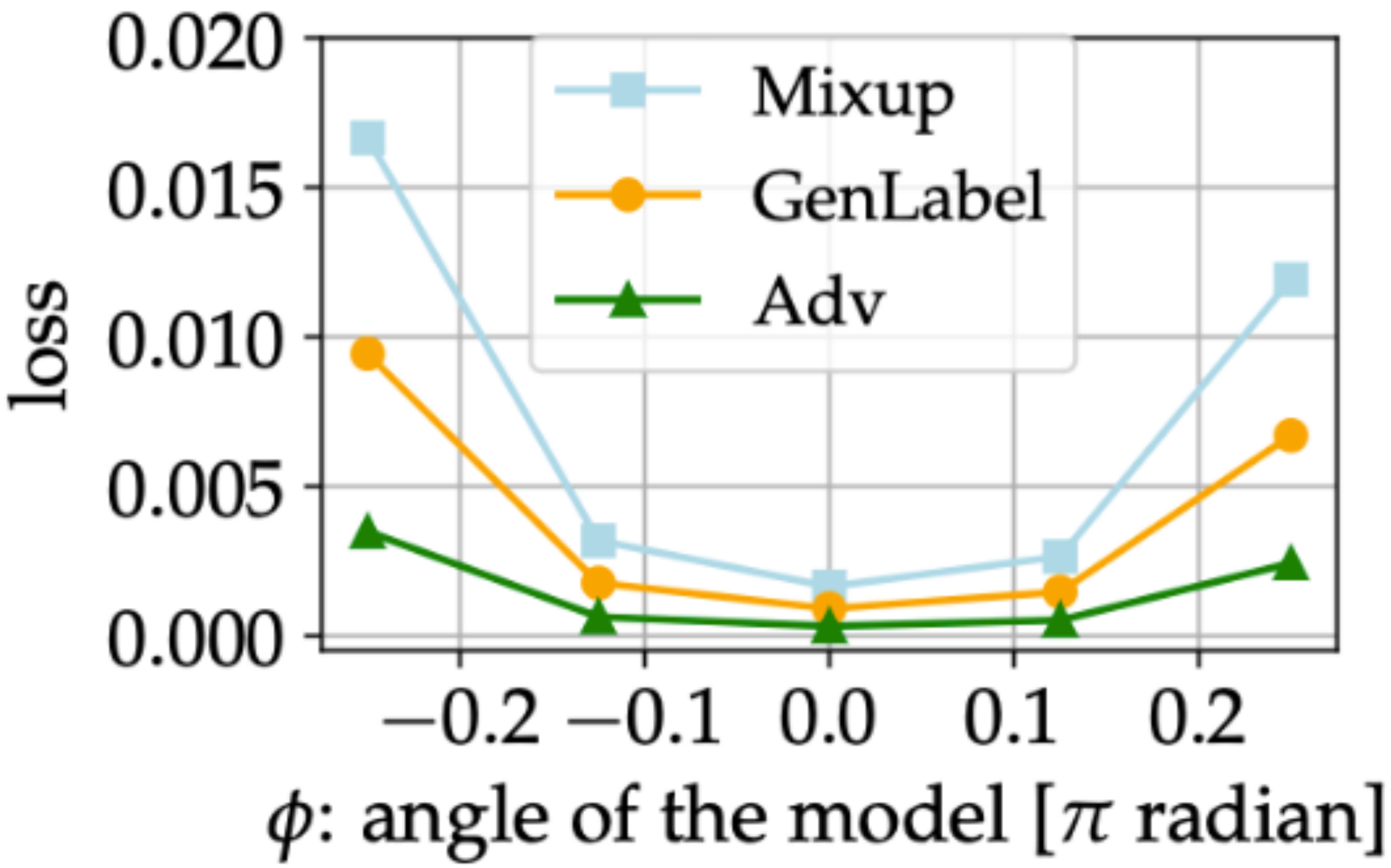
* black-box attack, $\varepsilon = 0.1$

Key Results: Robustness

[Thm] For logistic regression model & FC ReLU networks,

Mixup loss $\geq$ Mixup+**GenLabel** loss $\geq$ Adversarial loss

(Tighter Upper Bound)



Full version available at
https://arxiv.org/pdf/2201.02354.pdf

Hall E, Poster Session 2, #525