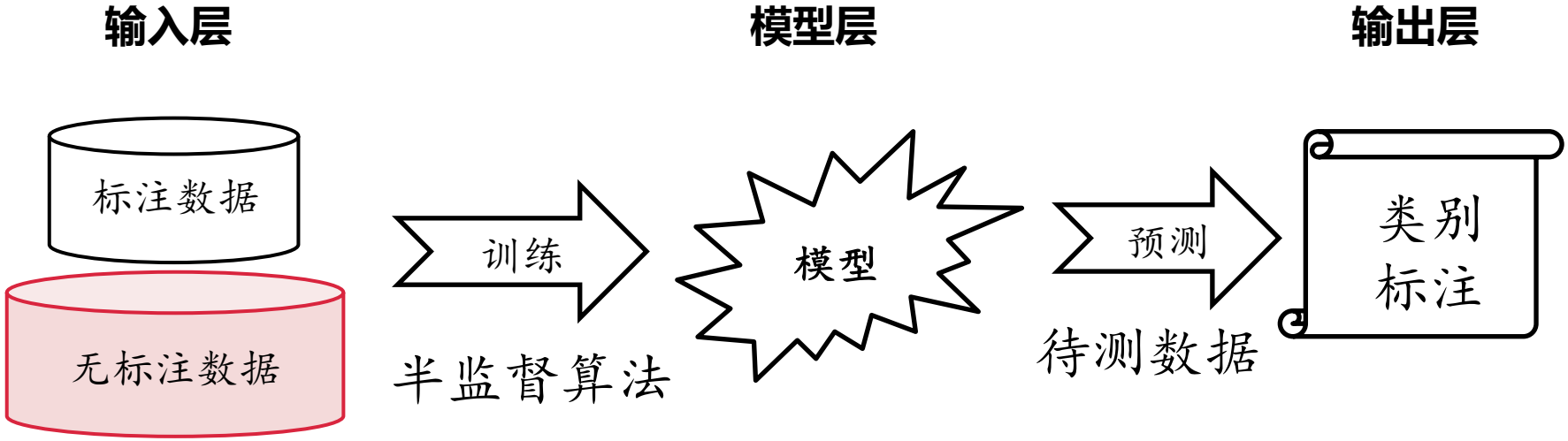
高级机器学习

安全半监督学习

半监督学习

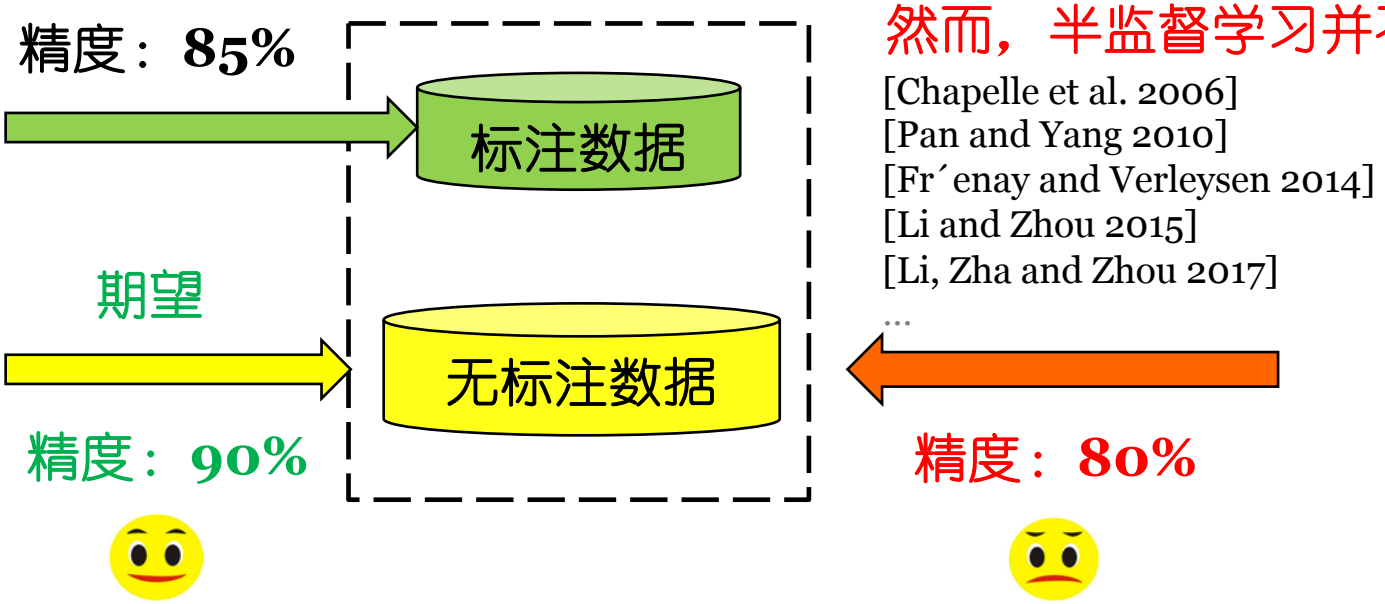
当标注数据不足时，利用大量无标注数据辅助提升性能



半监督学习

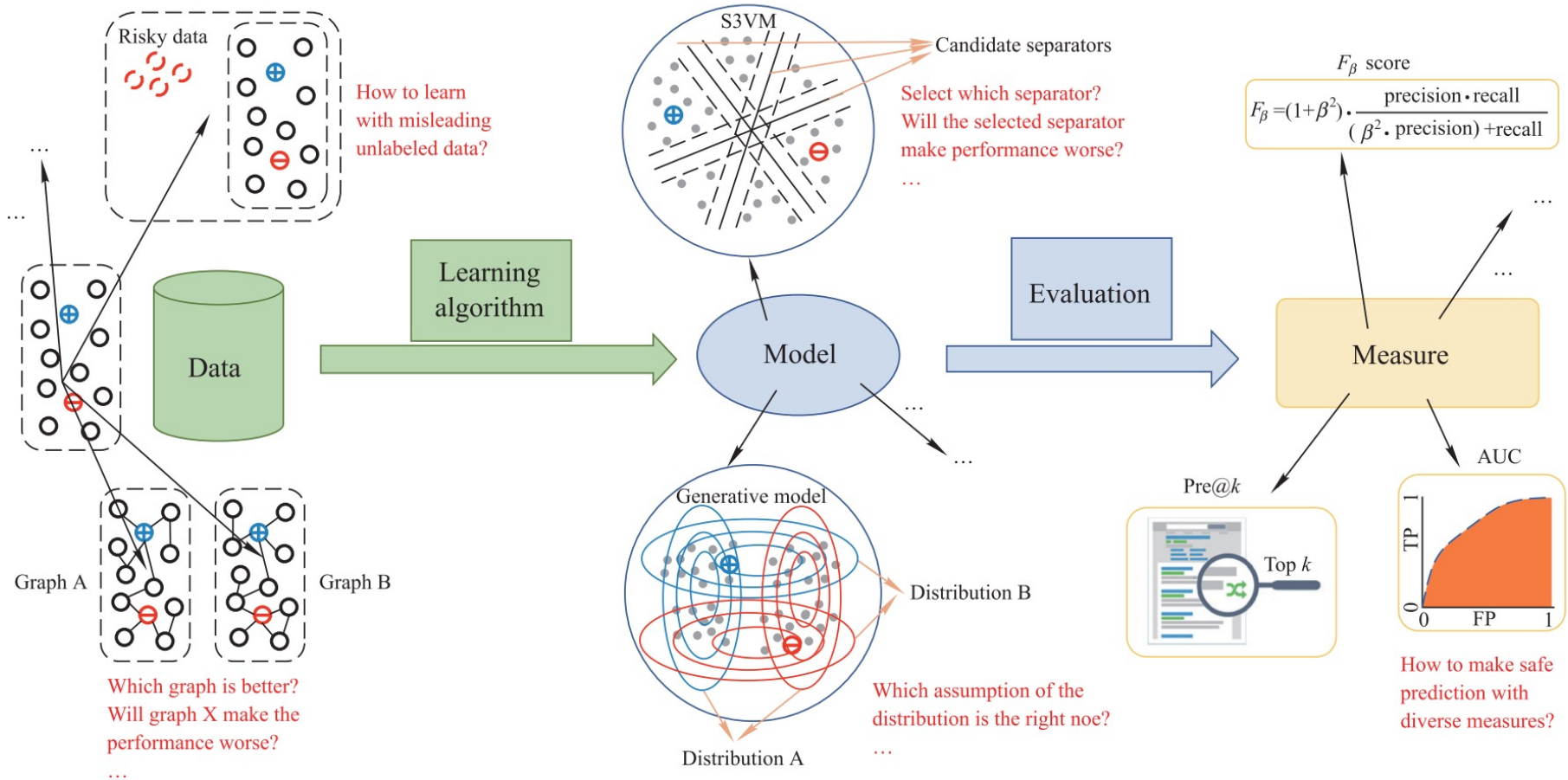
- 统计学习时期：自训练，生成式半监督、半监督SVM、图半监督、基于分歧的半监督
- 深度学习时期：熵最小化方法、一致性正则方法 混合式方法

安全半监督学习



安全半监督学习：利用了无标注数据，模型一直能够优于简单监督学习，是半监督学习领域的一个圣杯问题

半监督学习不安全的原因



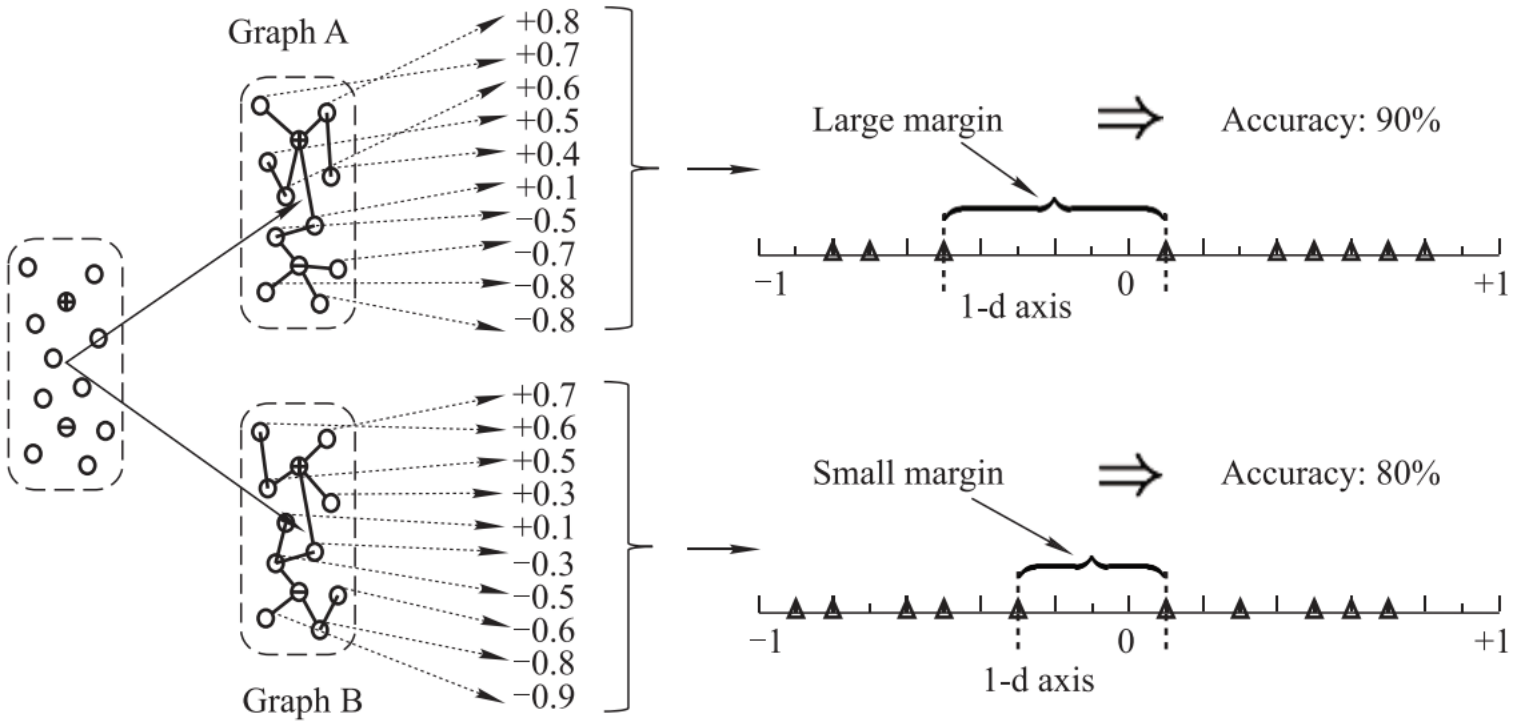
图半监督学习

- 图中的节点
  - 有标注样本+无标注样本
- 图中的边：基于样本特征计算相似度
  - K近邻图：与K近邻样本连边
  - ϵ近邻图：距离$\leq \epsilon$的样本连边
  - 全连接图：
  - …

图的质量影响半监督学习的安全性

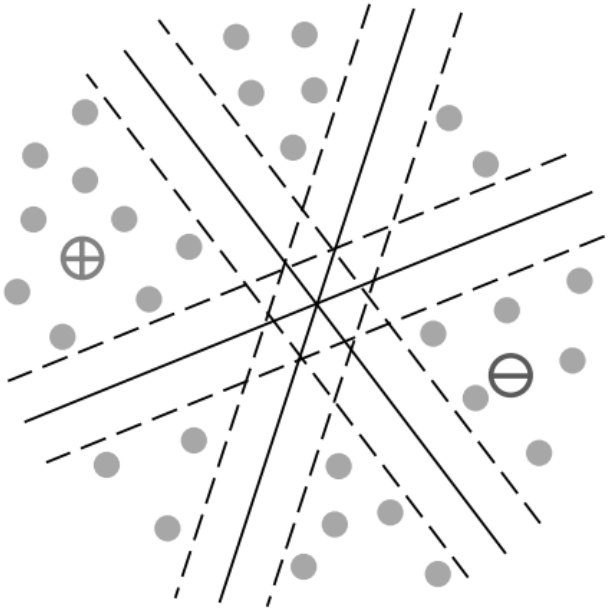
图半监督学习

基于大间隔准则判断图的质量



Li et al., Graph Quality Judgement: A Large Margin Expedition. IJCAI, 2016, pp.1725-1731.

模型不确定性（Model Uncertainty）



如果存在多个大间隔分类超平面，如何选择？

最优化最坏情况下的性能

综合考虑多个模型预测结果
以提升半监督学习安全性

- 多个半监督学习模型预测结果： $\{\mathrm{f}_1, \cdots, \mathrm{f}_b\}$
- 基线监督学习模型预测结果： f_0

假定真实标注f^*
已知

最大化性能增益

$$\max_{\mathbf{f} \in \mathbb{H}^u} \ \ell(\mathbf{f}_0, \mathbf{f}^*) - \ell(\mathbf{f}, \mathbf{f}^*)$$

$\ell(\cdot)$ 表示损失函数, f表示最终预测结果

最优化最坏情况下的性能

显然真实标注f^*未知

假设 f^*可由基学习器预测结果组合构造得到

$$\mathbf{f}^* = \sum_{i=1}^{b} \alpha_i \mathbf{f}_i$$

优化目标

$$\max_{\mathbf{f} \in \mathbb{H}^u} \ \ell(\mathbf{f}_0, \mathbf{f}^*) - \ell(\mathbf{f}, \mathbf{f}^*) \quad \Longleftrightarrow \quad \max_{\mathbf{f} \in \mathbb{H}^u} \ \ell(\mathbf{f}_0, \sum_{i=1}^{b} \alpha_i \mathbf{f}_i) - \ell(\mathbf{f}, \sum_{i=1}^{b} \alpha_i \mathbf{f}_i)$$

不失一般性，假设权重α来自凸集$\mathcal{M}$，$\mathcal{M}$可以灵活嵌入基学习器先验知识

最优化最坏情况下的性能

先验知识不足时，优化最坏情况下的性能增益

安全
半监督学习

$$\max_{\mathbf{f}\in\mathbb{H}^u}\ \min_{\boldsymbol{\alpha}\in\mathcal{M}}\ \ell(\mathbf{f}_0,\sum_{i=1}^{b}\alpha_i\mathbf{f}_i)-\ell(\mathbf{f},\sum_{i=1}^{b}\alpha_i\mathbf{f}_i)$$

最大最小优化[Li and Zhou, ICML2011/TPAMI2015; Balsubramani and Freund, COLT2015]

针对Hinge Loss, Cross-Entropy Loss, Mean Square Loss证明了该目标可以转化为凸优化问题从而高效求解

理论分析

- 如果无标注数据的真实标注可以由基学习器预测结果组合构造得到，则可以保障半监督学习的安全性

Theorem 1. *Suppose the ground-truth* $\mathbf{f}^*$ *can be constructed by base learners, i.e.,* $\mathbf{f}^* \in \{\mathbf{f} | \sum_{i=1}^{b} \alpha_i \mathbf{f}_i, \boldsymbol{\alpha} \in \mathcal{M}\}$. *Let* $\hat{\mathbf{f}}$ *and* $\hat{\boldsymbol{\alpha}}$ *be the optimal solution to Eq. (1). We have* $\ell(\hat{\mathbf{f}}, \mathbf{f}^*) \leq \ell(\mathbf{f}_0, \mathbf{f}^*)$ *and* $\hat{\mathbf{f}}$ *has already achieved the maximal performance gain against* $\mathbf{f}_0$.

实验验证

- 既有方法均出现性能不如监督学习的现象
- **SafeW方法在所有场景下均优于监督学习**

5 个标注样本							
数据集	1NN	Self-kNN	Self-LS	COREG	Voting	OpW	SafeW
abalone	.017 ± .007	**.014 ± .003**	**.013 ± .004**	**.013 ± .003**	**.012 ± .003**	**.005 ± .001**	**.013 ± .003**
bodyfat	.024 ± .008	.025 ± .009	.054 ± .016	.026 ± .008	.031 ± .011	**.018 ± .003**	.025 ± .009
cadata	.090 ± .031	**.073 ± .023**	**.067 ± .022**	**.069 ± .028**	**.069 ± .022**	**.039 ± .014**	**.070 ± .023**
cpusmall	.027 ± .012	.031 ± .008	.050 ± .021	.031 ± .009	.024 ± .006	**.014 ± .003**	.028 ± .009
eunite2001	.052 ± .017	**.037 ± .015**	**.024 ± .012**	**.037 ± .011**	**.031 ± .013**	**.018 ± .005**	**.032 ± .010**
housing	.042 ± .007	.043 ± .009	.048 ± .012	.041 ± .008	.042 ± .009	**.024 ± .002**	.041 ± .009
mg	.071 ± .035	**.057 ± .015**	**.053 ± .011**	**.054 ± .019**	**.054 ± .013**	**.028 ± .009**	**.053 ± .013**
mpg	.029 ± .012	.030 ± .012	.040 ± .014	.031 ± .012	.031 ± .012	**.016 ± .002**	.030 ± .012
pyrim	.032 ± .009	**.027 ± .005**	.063 ± .012	.029 ± .011	**.025 ± .007**	**.013 ± .002**	**.025 ± .005**
space_ga	.005 ± .002	.005 ± .003	.030 ± .005	**.004 ± .002**	.008 ± .002	**.001 ± .000**	**.004 ± .002**
平均性能	.039	.034	.044	.033	.033	.020	.032
胜出/打平/打败		5/4/1	4/0/6	5/4/1	5/3/2	**9/0/0**	**6/4/0**

Measure Diversity

在分类和回归任务中，通常采用正确率和均方误差，然而在现实机器学习应用中，评价指标可能是多样的

- 文本分类任务中，经常采用F1-Score和Precision-Recall
- 信息检索任务中，通常采用Precision和Recall
- 排序任务中，通常采用AUCROC、Top-K

针对特定的任务，半监督学习模型需要在特定指标下实现安全性

Measure Diversity

- 多个半监督学习模型预测结果： $\{f_1, \cdots, f_b\}$
- 基线监督学习模型预测结果：f_0

$$\max_{\hat{y}_u \in \mathcal{Y}} \sum_{i=1}^{b} \alpha_i \left(perf(\hat{\mathbf{y}}_u, \mathbf{y}_u^i) - perf(\hat{\mathbf{y}}_u^0, \mathbf{y}_u^i) \right)$$

优化**最坏情况下的性能增益**

$$\max_{\hat{y}_u \in \mathcal{Y}} \min_{\alpha \in \mathcal{M}} \sum_{i=1}^{b} \alpha_i \left(perf(\hat{\mathbf{y}}_u, \mathbf{y}_u^i) - perf(\hat{\mathbf{y}}_u^0, \mathbf{y}_u^i) \right)$$

既有半监督学习

主要针对封闭环境（标注与无标注数据大多是“一致”的）

Labeled Data

Unlabeled Data

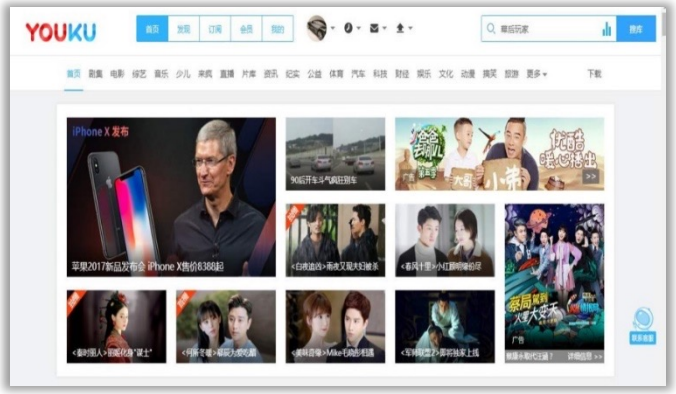
- 数据分布一致
- 样本类别一致
- 样本属性一致
- 类别比例一致
- …

现实应用常面临开放环境，无标注数据可能与标注数据不一致

数据分布不一致

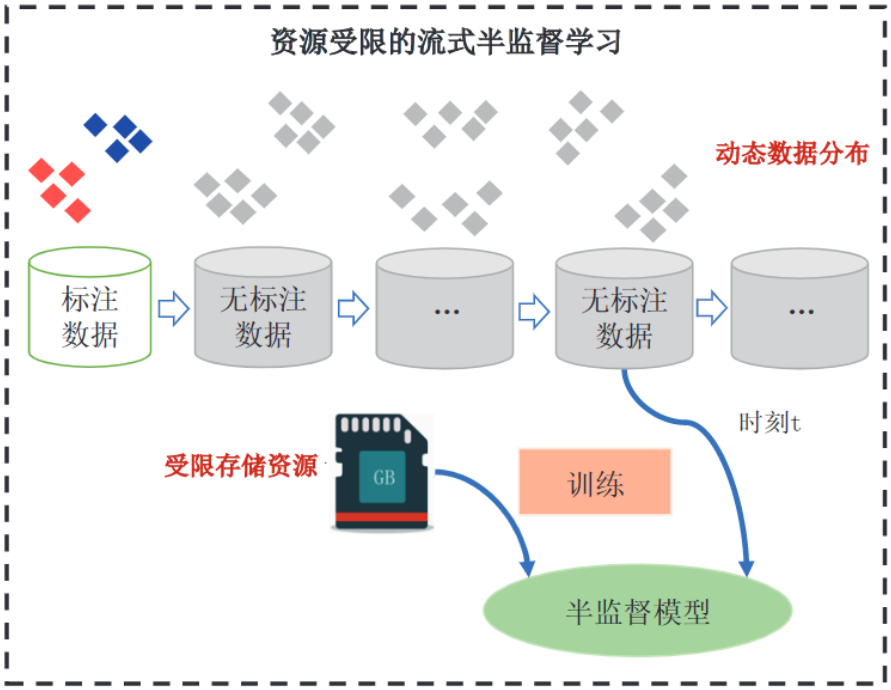
封闭环境假设数据分布一致，开放环境中数据分布动态变化

现实场景：以在线视频分类为例



- 视频内容每日更新（流式数据）
- 视频标注代价昂贵（标注不全）
- 数据海量难以存储（资源受限）

问题设置：
“资源受限的流式半监督学习”

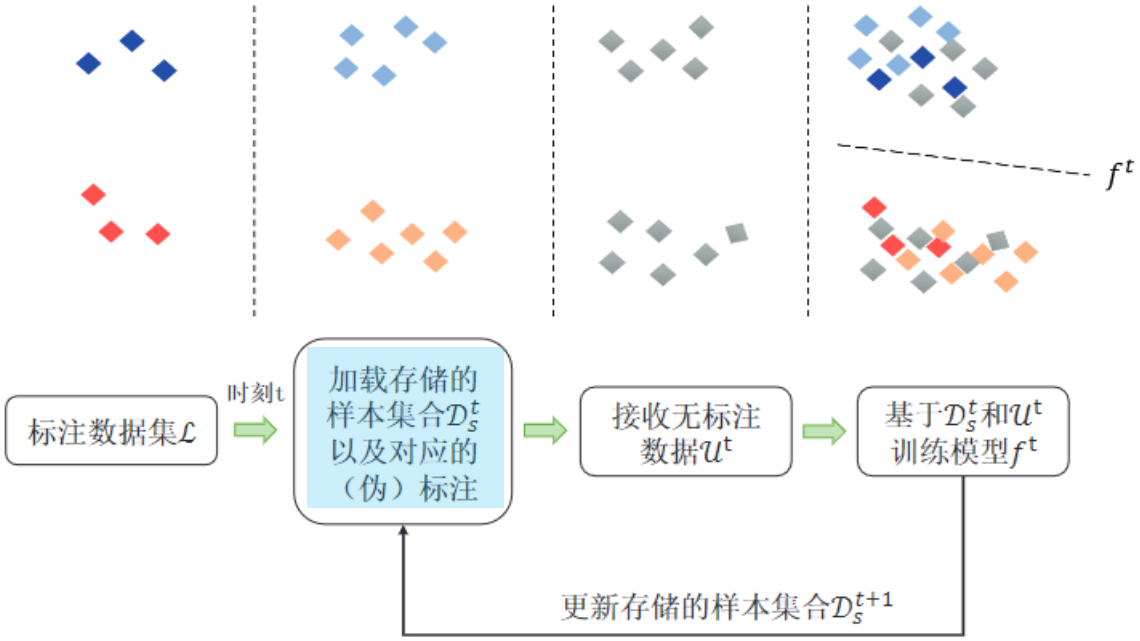


Record学习框架

主要挑战

- 数据分布不断变化，如何使模型适应分布变化，避免性能下降？
- 数据无法完整存储，如何高效利用无标注样，最大化性能增益？

Record框架



维护存储样本集合$\mathcal{D}_s^t$
满足资源约束的条件下，
追踪数据分布变化

$$\max_{\mathcal{D}_s^t} \quad \text{performance}(f)$$

$$\text{s.t.} \quad f = \mathcal{A}(\mathcal{D}_s^t, \mathcal{U}^t)$$

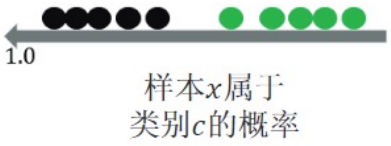
$$\mathcal{D}_s^t \subseteq \mathcal{D}_s^{t-1} \cup \mathcal{U}^{t-1}$$

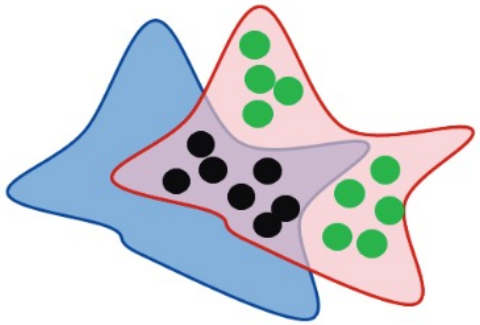
$$size(\mathcal{D}_s^t) \leq B$$

Record学习框架

如何选择样本进行存储？

基于置信度的分布偏移样本检测





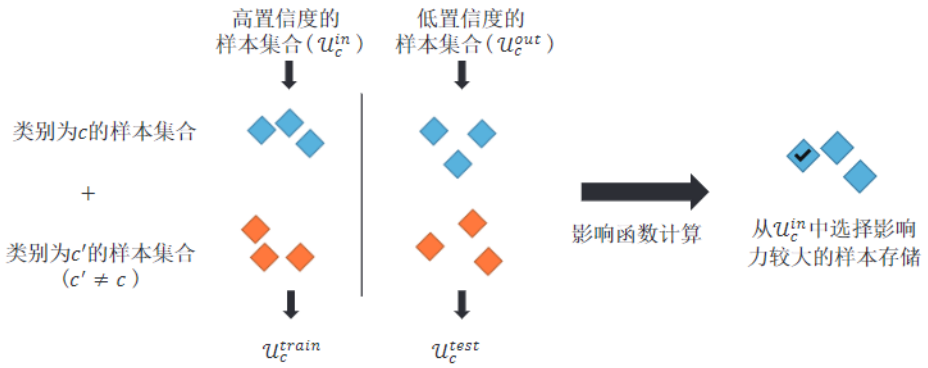
根据模型预测概率，将样本分为集合$\mathcal{U}_c^{in}$ 和集合 $\mathcal{U}_c^{out}$

- ✓ $\mathcal{U}_c^{in}$中样本置信度更高，更容易落在当前分布
- ✓ $\mathcal{U}_c^{out}$中样本置信度更低，更容易落在偏移后的分布

Record学习框架

如何选择样本进行存储？

基于影响力机制的样本选择



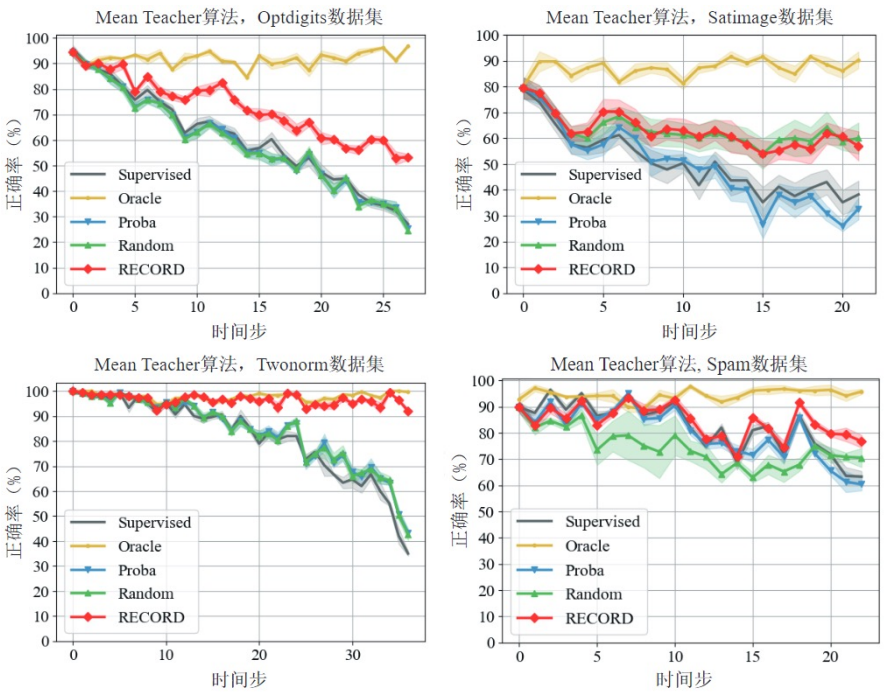
计算集合$\mathcal{U}_c^{in}$ 中的样本在集合$\mathcal{U}_c^{out}$上的影响函数

$$\mathrm{IF}(\mathbf{x}) = \frac{1}{|\mathcal{U}_c^{test}|} \sum_{\mathbf{z}_{test} \in \mathcal{U}_c^{test}} \mathcal{I}(\mathbf{x}, \mathbf{z}_{test})$$

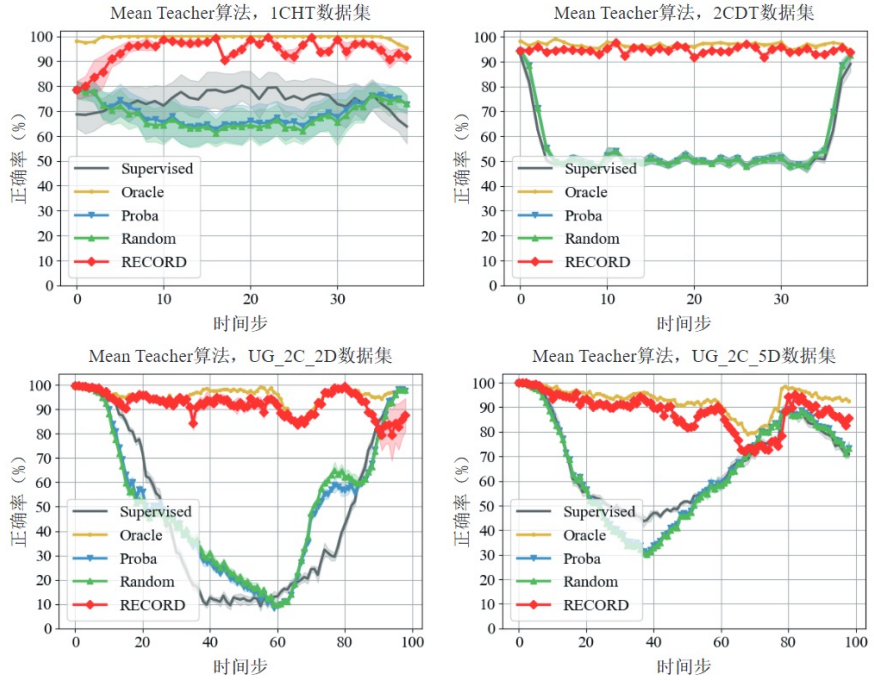
选择在偏移后的数据分布上影响力更大的样本进行存储

实验结果

- 红色：本文方法
- 黄色：利用所有数据标注的Oracle方法
- 其它：对比方法



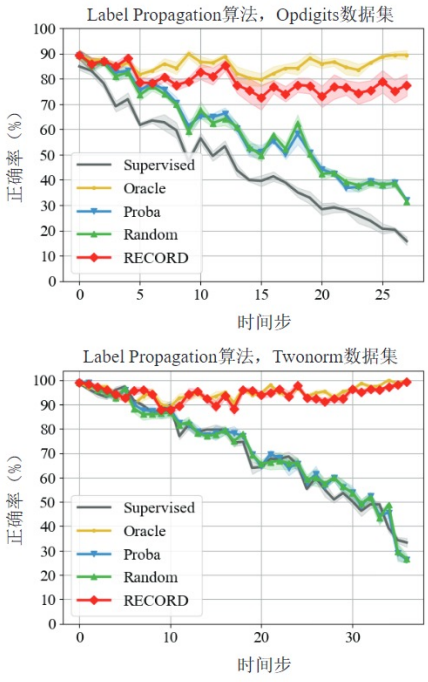
4种常用分类数据集

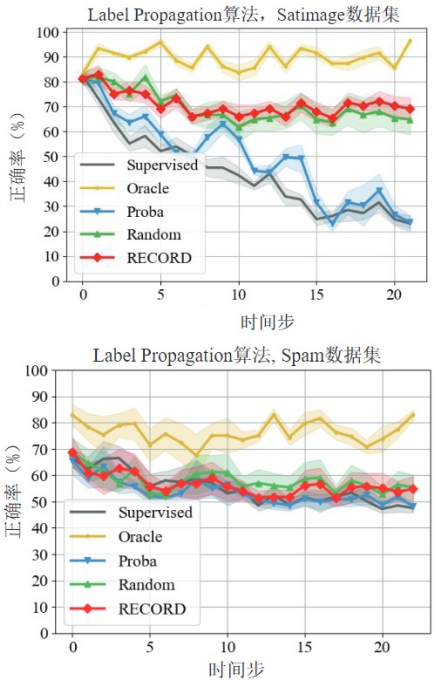


4种流数据基准数据集

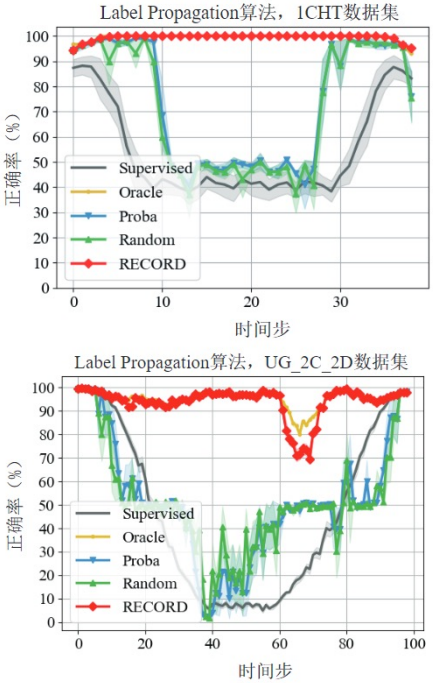
实验结果

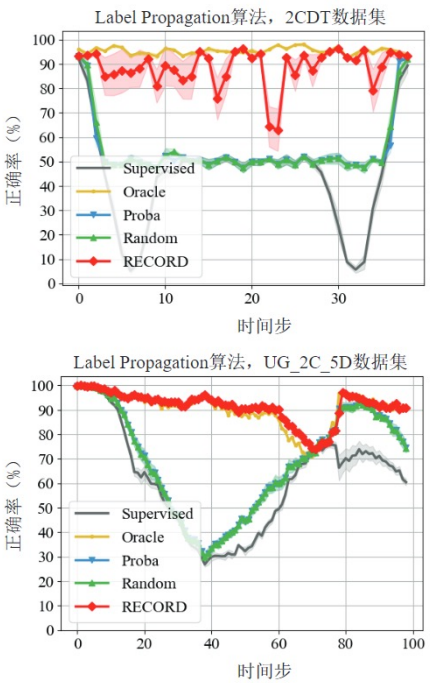
- 红色：本文方法
- 黄色：利用所有数据标注的Oracle方法
- 其它：对比方法





4种常用分类数据集

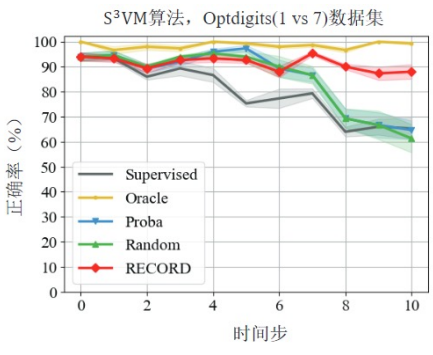


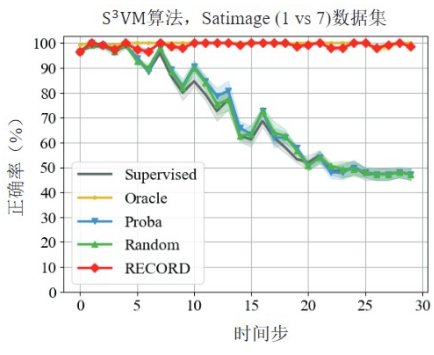


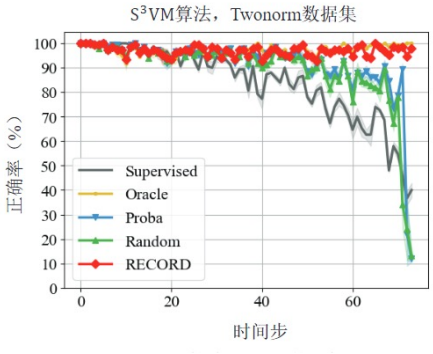
4种流数据基准数据集

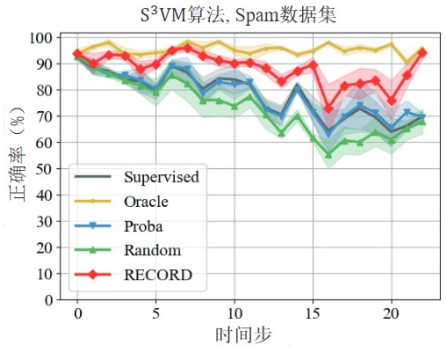
实验结果

- 红色：本文方法
- 黄色：利用所有数据标注的Oracle方法
- 其它：对比方法

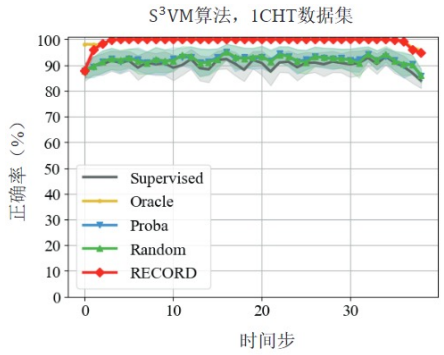


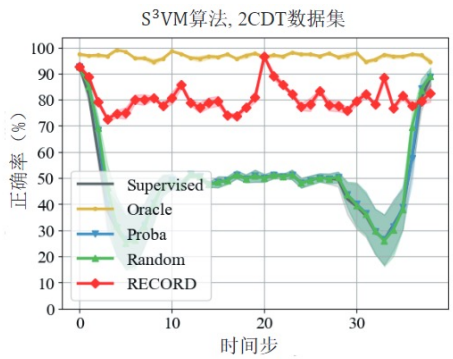


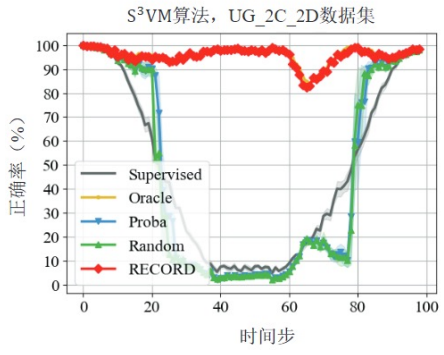


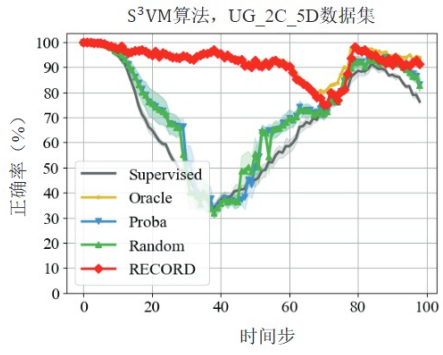


4种常用分类数据集









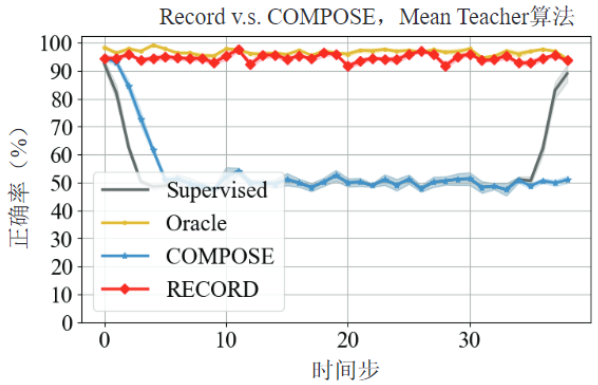
4种流数据基准数据集

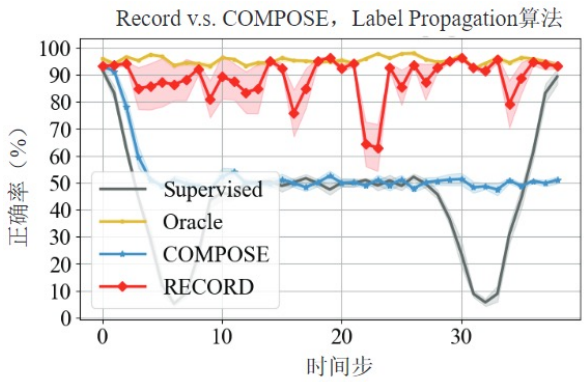
实验结果

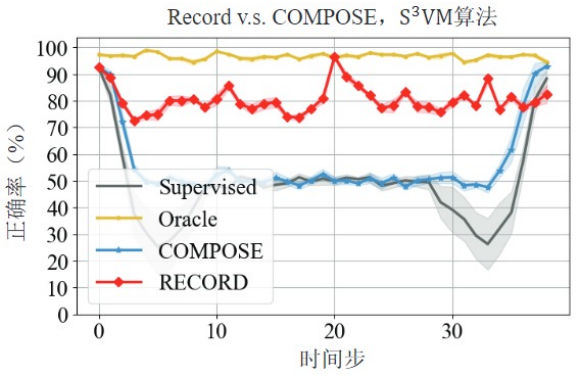
数据集	Supervised	TLP	Record	Oracle
Optdigits	46.64 ± 1.04	53.56 ± 1.42	**79.20 ± 2.01**	**85.61 ± 0.25**
Twonorm	70.84 ± 0.36	57.24 ± 0.86	**94.31 ± 0.38**	**94.99 ± 0.13**
Satimage	43.47 ± 2.61	58.43 ± 2.54	**71.02 ± 2.93**	**89.59 ± 0.70**
Spam	55.12 ± 3.48	54.18 ± 1.39	**56.55 ± 4.72**	**76.35 ± 2.41**
1CHT	54.79 ± 4.93	95.78 ± 1.27	**99.44 ± 0.05**	**99.44 ± 0.01**
2CDT	44.47 ± 0.46	54.28 ± 0.48	**88.68 ± 1.57**	**95.36 ± 0.13**
UG_2C_2D	46.40 ± 0.23	47.56 ± 1.12	**94.11 ± 0.12**	**95.35 ± 0.04**
UG_2C_5D	61.00 ± 1.01	54.40 ± 0.17	**91.39 ± 0.08**	**90.41 ± 0.05**
平均性能	52.84 ± 1.77	59.43 ± 1.16	84.34 ± 1.48	90.89 ± 0.47

实验结果

半监督算法	Supervised	COMPOSE	Record	Oracle
Mean Teacher	54.40 ± 0.36	54.09 ± 0.24	**94.60 ± 0.13**	**96.72 ± 0.18**
Label Propagation	44.47 ± 0.46	53.19± 0.20	**88.70 ± 1.57**	**95.36 ± 0.13**
S^3VM	48.40 ± 2.29	56.05 ± 0.70	**80.41 ± 0.19**	**96.71 ± 0.16**
平均性能	49.04 ± 1.04	54.44 ± 0.38	87.90 ± 0.63	96.26 ± 0.47







样本类别不一致

现象

无标注数据的收集缺少人工监督，容易包含未见类样本

标注数据

无标注数据

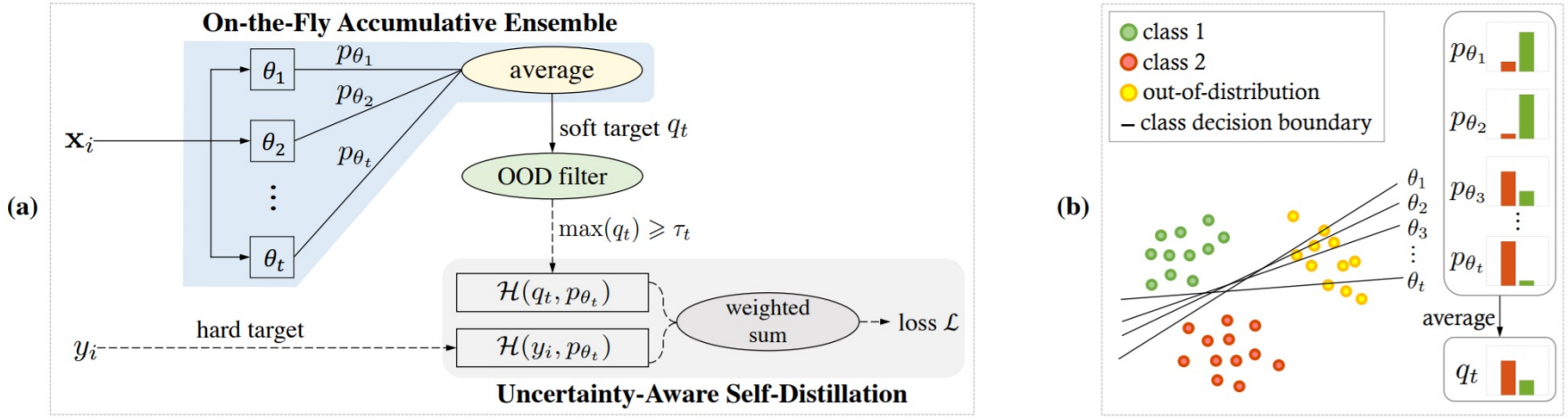
科学问题

如何降低未见类样本对学习性能的影响，

进一步，能够对未见类样本进行分类

UASD

观察 模型训练过程中，对OOD样本的预测值是不一致的

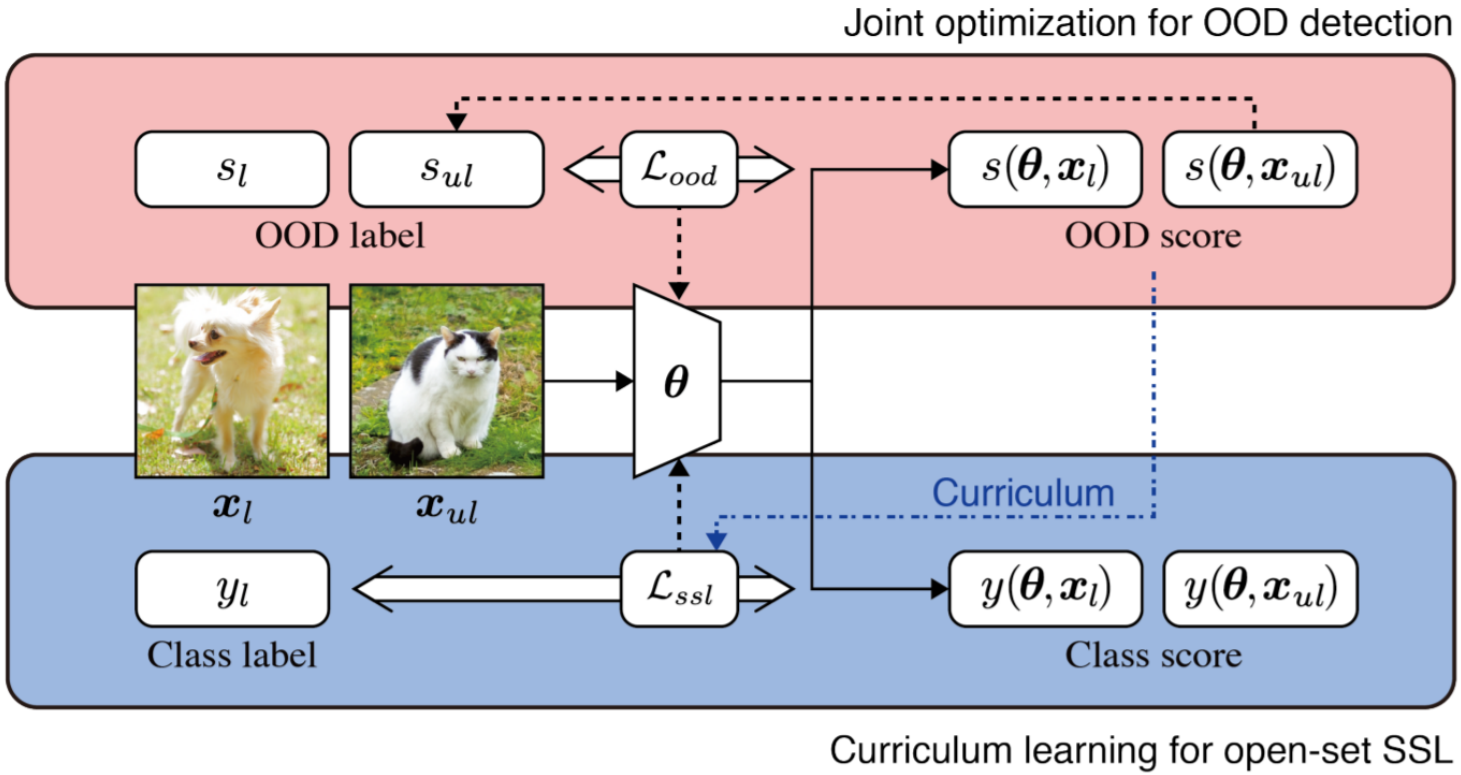


主要想法 利用集成模型预测结果的置信度筛选样本

Chen et al., Semi-supervised learning under class distribution mismatch. AAAI 2020

MTCF

主要想法　将OOD检测和半监督模型训练建模成多任务学习



Yu et al., Multi-task curriculum framework for open-set semi-supervised learning. ECCV 2020

DS3L方法

既有深度半监督学习方法

优化目标：

$$\min_{\theta\in\Theta}\sum_{i=1}^{n}\mathcal{L}_s(f(\mathbf{x}_i;\theta),\mathbf{y}_i)+\sum_{i=n+1}^{n+m}\mathcal{L}_u(\mathbf{x}_i;\theta)$$

监督损失　　无监督损失

同等利用所有无标注样本

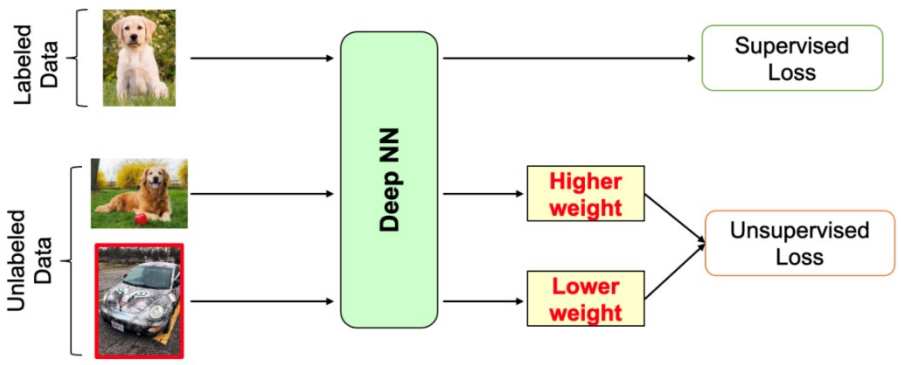
DS3L方法

优化目标：

$$\min_{\theta\in\Theta}\sum_{i=1}^{n}\mathcal{L}_s(f(\mathbf{x}_i;\theta),\mathbf{y}_i)+\sum_{i=n+1}^{n+m}w(\mathbf{x}_i;\alpha)\mathcal{L}_u(\mathbf{x}_i;\theta)$$

监督损失　　加权无监督损失

选择性利用无标注样本



基本思路：通过赋予分布外样本更低的权重，降低其对模型性能的负面影响

如何获得样本权重？

DS3L方法

双层优化

内层优化得到加权经验风险最小的模型
外层优化得到使模型经验风险最小的权重

标注数据
经验风险
最小化

$$\min_{\alpha \in \mathbb{B}^d} \sum_{i=1}^{n} \ell(h(\mathbf{x}_i; \hat{\theta}), \mathbf{y}_i)$$

s.t.

所有数据
加权经验风险
最小化

$$\hat{\theta} = \operatorname*{argmin}_{\theta \in \Theta} \sum_{i=1}^{n} \ell(h(\mathbf{x}_i; \theta), \mathbf{y}_i) + \sum_{i=n+1}^{n+m} w(\mathbf{x}_i; \alpha) \Omega(\mathbf{x}_i; \theta)$$

监督损失

加权无监督损

优化方法

双层优化

$$\min_{\alpha \in \mathbb{B}^d} \sum_{i=1}^{n} \ell(h(\mathbf{x}_i; \hat{\theta}), \mathbf{y}_i)$$

s.t.

$$\hat{\theta} = \operatorname*{argmin}_{\theta \in \Theta} \sum_{i=1}^{n} \ell(h(\mathbf{x}_i; \theta), \mathbf{y}_i) + \sum_{i=n+1}^{n+m} w(\mathbf{x}_i; \alpha) \Omega(\mathbf{x}_i; \theta)$$

双层优化求解困难，传统方法复杂度高

近似迭代优化方法：

交替优化模型参数和权重参数

$$\theta_{t+1} = \theta_t - \eta_\theta \nabla_\theta \mathcal{L}^{inner}(\theta_t, \alpha_t)$$

$$\alpha_{t+1} = \alpha_t - \eta_\alpha \nabla_\alpha \mathcal{L}^{outer}(\theta_{t+1})$$

主要结果

- 迭代优化方法可以收敛
- 收敛率为$O(\frac{1}{\sqrt{T}})$

定理 2-1 （**收敛性**）假设损失函数是 L 利普希茨连续的，令模型参数 θ 的优化步长满足 $\eta_\theta \leq \frac{2G}{L}$，$G$ 为大于 0 的常数，那么，基于本文提出的优化算法，模型在标注数据上的监督损失将会随着训练轮数的增加单调递减，即，

$$\mathcal{L}^{outer}(\theta_{t+1}) \leq \mathcal{L}^{outer}(\theta_t) \tag{2-15}$$

进一步，上式等号成立，当且仅当外层优化目标对权重参数 α 的梯度为 0，即，

$$\mathcal{L}^{outer}(\theta_{t+1}) = \mathcal{L}^{outer}(\theta_t) \tag{2-16}$$

当且仅当

$$\nabla_\alpha \mathcal{L}^{outer}(\theta_t) = 0 \tag{2-17}$$

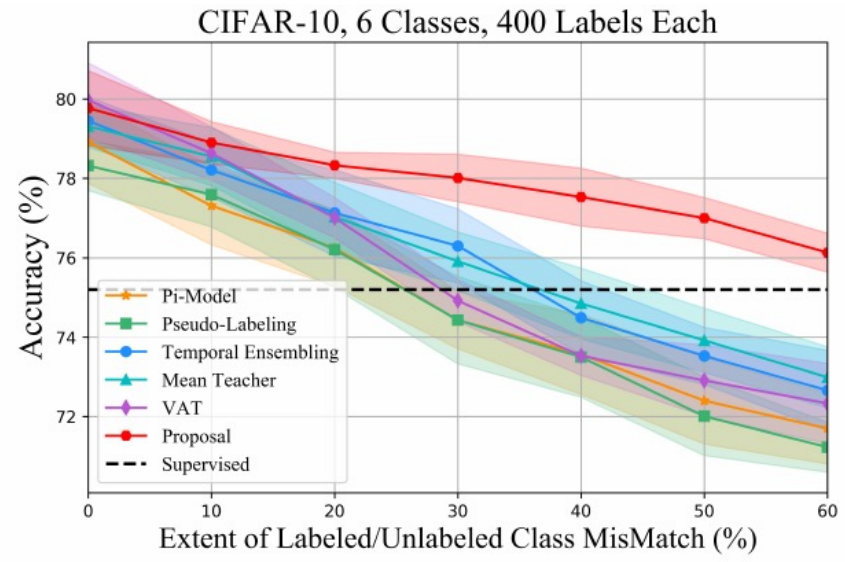
定理 2-2 （**收敛率**）假设损失函数是 L 利普希茨连续的，并且梯度是 ρ 限制的。令模型参数 θ 的优化步长 η_θ 满足 $\eta_\theta = \min\{1, \frac{k}{T}\}$，$k > 0$ 并且 $\frac{k}{T} < 1$；权重参数 α 的优化步长满足 $\eta_\alpha = \min\{\frac{1}{L}, \frac{c}{\sqrt{T}}\}$，$c > 0$ 并且 $\frac{\sqrt{T}}{c} \geq L$。那么，本文的优化算法能够以 $O(1/\epsilon^2)$ 的阶数实现 $\mathbb{E}[\|\nabla_\alpha \mathcal{L}^{outer}(\theta_t)\|_2^2] \leq \epsilon$，即：

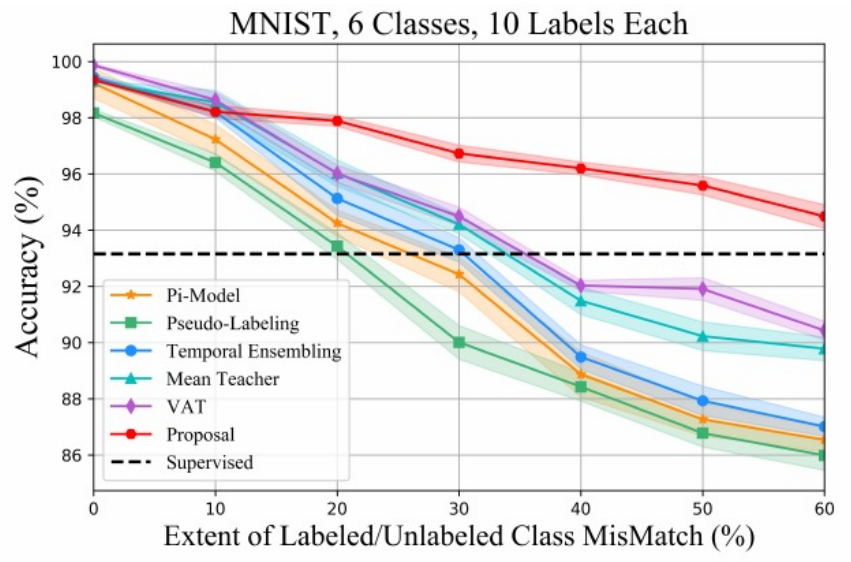
$$\min_{0 \leq t \leq T} \mathbb{E}[\|\nabla_\alpha \mathcal{L}^{outer}(\theta_t)\|_2^2] \leq O(\frac{1}{\sqrt{T}}) \tag{2-18}$$

实验验证

主要结果

- 以往方法在分布失配程度40%时表现不如简单监督学习
- DS3L方法在分布失配程度60%时仍然可以取得性能提升

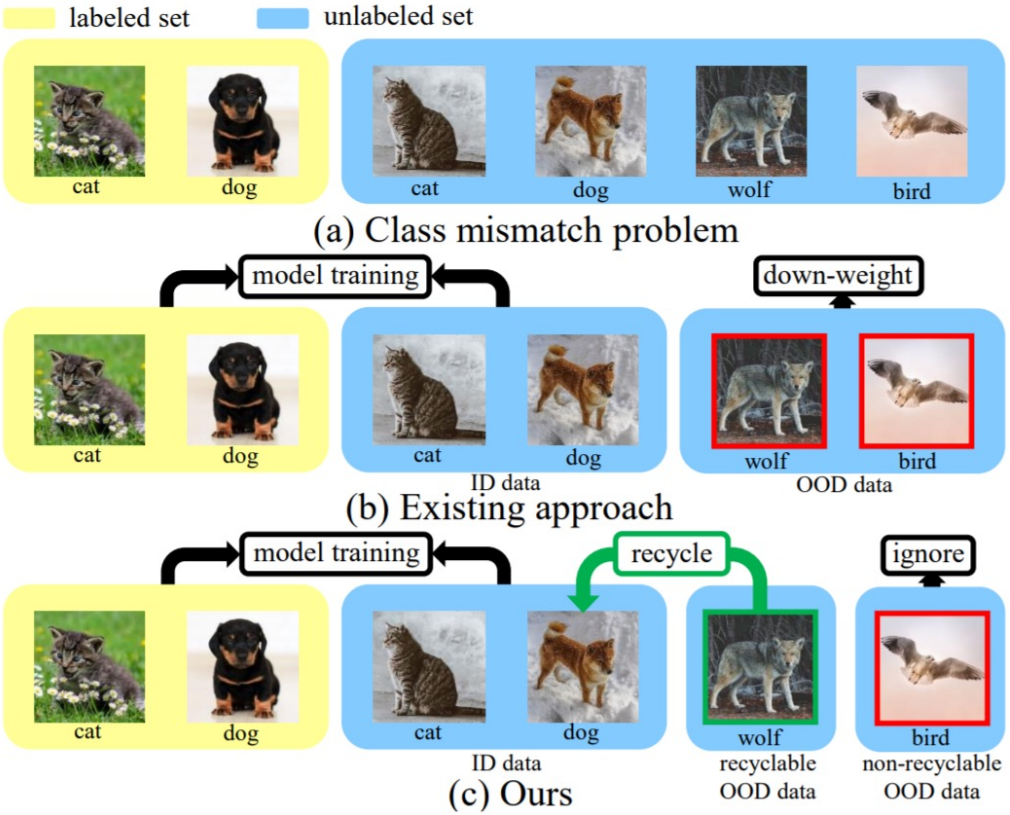




TOOR

观察

存在一些OOD的无标注数据是对学习任务有帮助的



主要想法

Adversarial Domain Adaptation

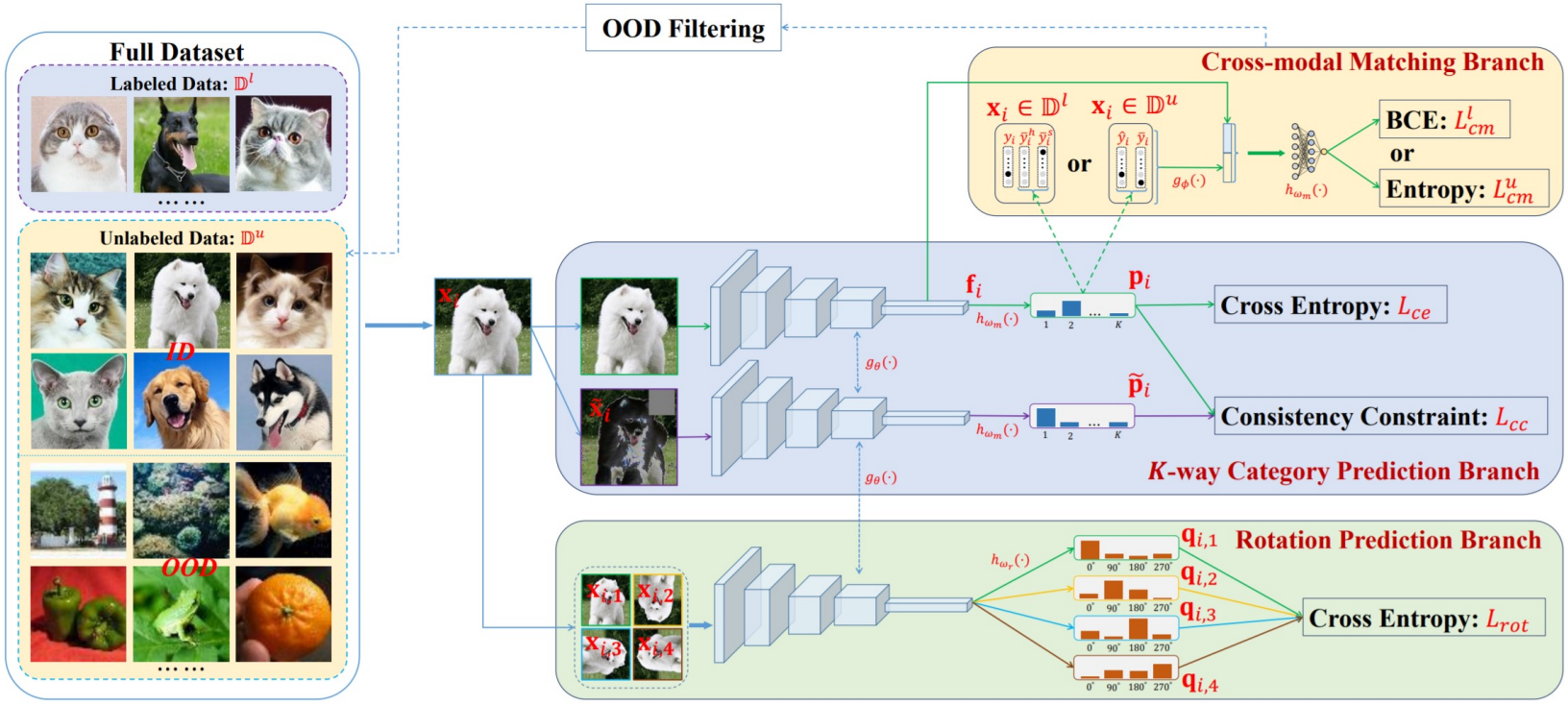
- 将OOD样本作为Source Domain
- ID样本+标记数据作为Target Domain

Huang et al., They are not completely useless: Towards recycling transferable unlabeled data for class-mismatched semi-supervised learning

OOD辅助表示学习

观察

OOD的无标注数据对表示学习有帮助

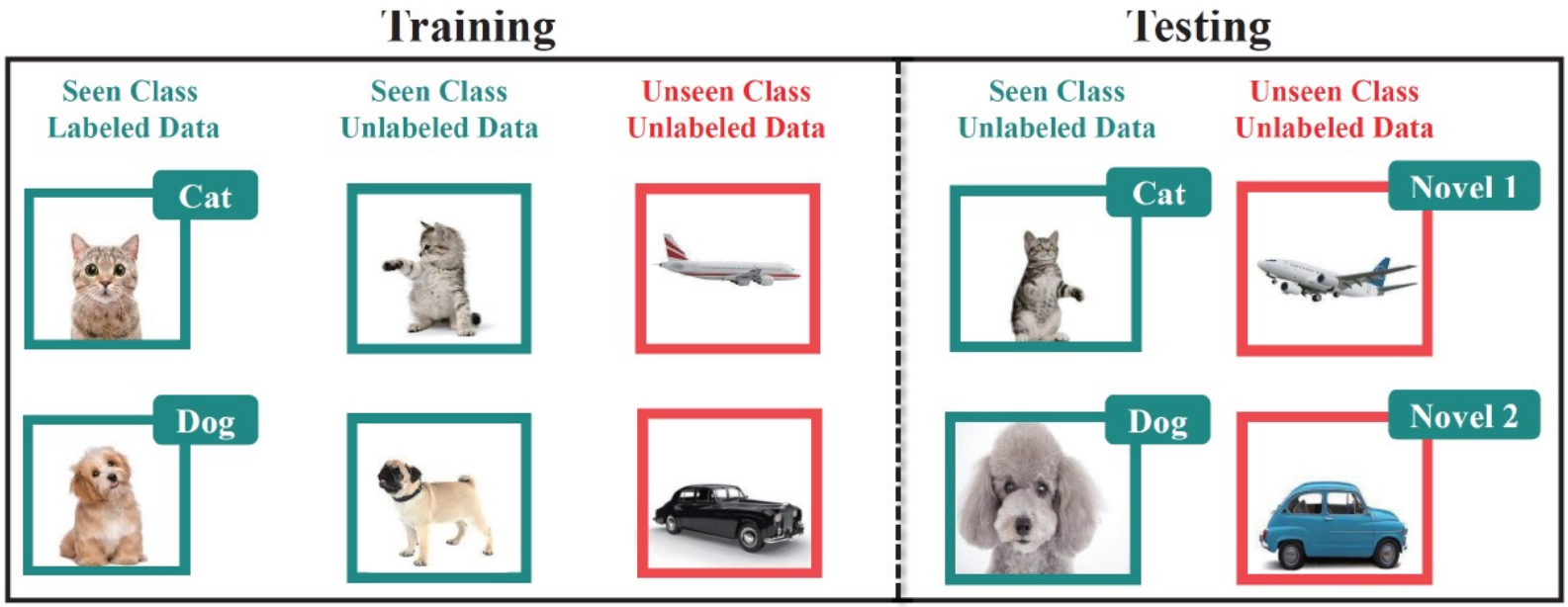


Huang et al., Trash to treasure: Harvesting ood data with cross-modal matching for open-set semi-supervised learning. ICCV 2021

如何对未见类别进行分类

现象

无标注数据中包含的未见类别是任务相关的



科学问题

如何为已知类产生正确的标注

同时将未见类样本划分为合理的簇

NACH

难点1：如何对未见类无标注样本进行划分？

可行方案：

利用样本成对相似度进行聚类

Pairwise Objective

难点2：如何平衡已知类和未见类的学习速率？

可行方案：

利用动态阈值同步学习速率

学习过程中，降低未见类的伪标注选择阈值，增加选择的样本数量

实验验证

主要结果

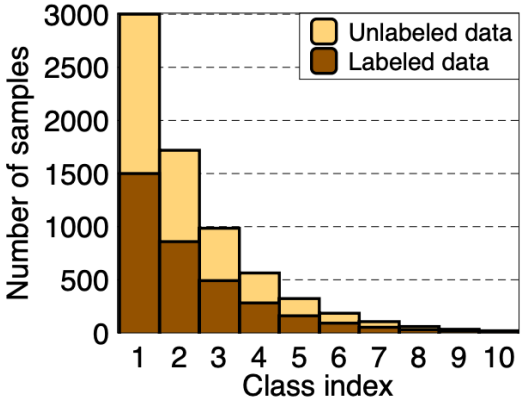
- 能够准确区分未见类，同时已知类性能不退化

Classes	Dataset	SSL	Open-Set SSL		NCD			
		Fixmatch	DS3L	CGDL	DTC	RankStats	ORCA	OURS
Seen	CIFAR-10	71.5	77.6	72.3	53.9	86.6	88.2	**89.5**
	CIFAR-100	39.6	55.1	49.3	31.3	36.4	66.9	**68.7**
	ImageNet-100	65.8	71.2	67.3	25.6	47.3	89.1	**91.0**
	Average	59.0	68.0	63.0	36.9	56.8	81.4	**83.1**
Unseen	CIFAR-10	50.4	45.3	44.6	39.5	81.0	90.4	**92.2**
	CIFAR-100	23.5	23.7	22.5	22.9	28.4	43.0	**47.0**
	ImageNet-100	36.7	32.5	33.8	20.8	28.7	72.1	**75.5**
	Average	36.9	33.9	33.6	27.7	46.0	68.5	**71.6**
All	CIFAR-10	49.5	40.2	39.7	38.3	82.9	89.7	**91.3**
	CIFAR-100	20.3	24.0	23.5	18.3	23.1	48.1	**52.1**
	ImageNet-100	34.9	30.8	31.9	21.3	40.3	77.8	**79.6**
	Average	34.9	31.7	31.7	26.0	48.8	71.9	**74.3**

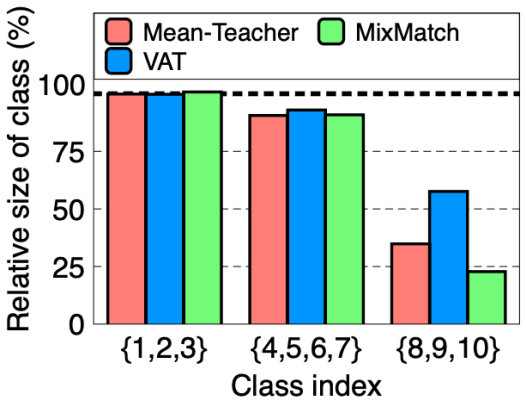
L.-Z. Guo, Y.-G. Zhang, Z.-F. Wu, J.-J. Shao, Y.-F. Li. Robust Semi-Supervised Learning when Not All Classes have Labels. In: Advances in Neural Information Processing Systems (NeurIPS), 2022.

类别不平衡

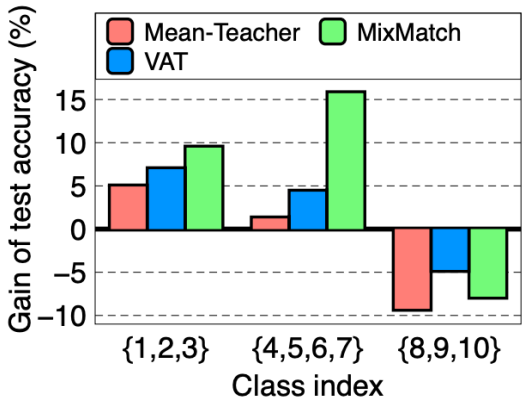
既有半监督学习假设类别比例平衡，然而，现实应用中：



(a) Imbalanced class distribution



(b) Bias of pseudo-labels



(c) Accuracy gain from SSL

半监督学习模型在少数类上性能下降严重

DARP

假设真实各类别样本数量为$\mathbf{M_k}$

$$\text{minimize} \quad \sum_{m=1}^{M} w_m D_{KL}(\hat{y}_m \,||\, \hat{y}_m^{\texttt{unlabeled}})$$

$$\text{subject to} \quad \sum_{m=1}^{M} \hat{y}_m(k) = M_k,\ \forall k,\ \sum_{k=1}^{K} \hat{y}_m(k) = 1,\ \forall m,\ \hat{y}_m(k) \in [0,1],\ \forall m,k$$

Kim et al., Distribution Aligning Refinery of Pseudo-label for Imbalanced Semi-supervised Learning. NeurIPS 2020.

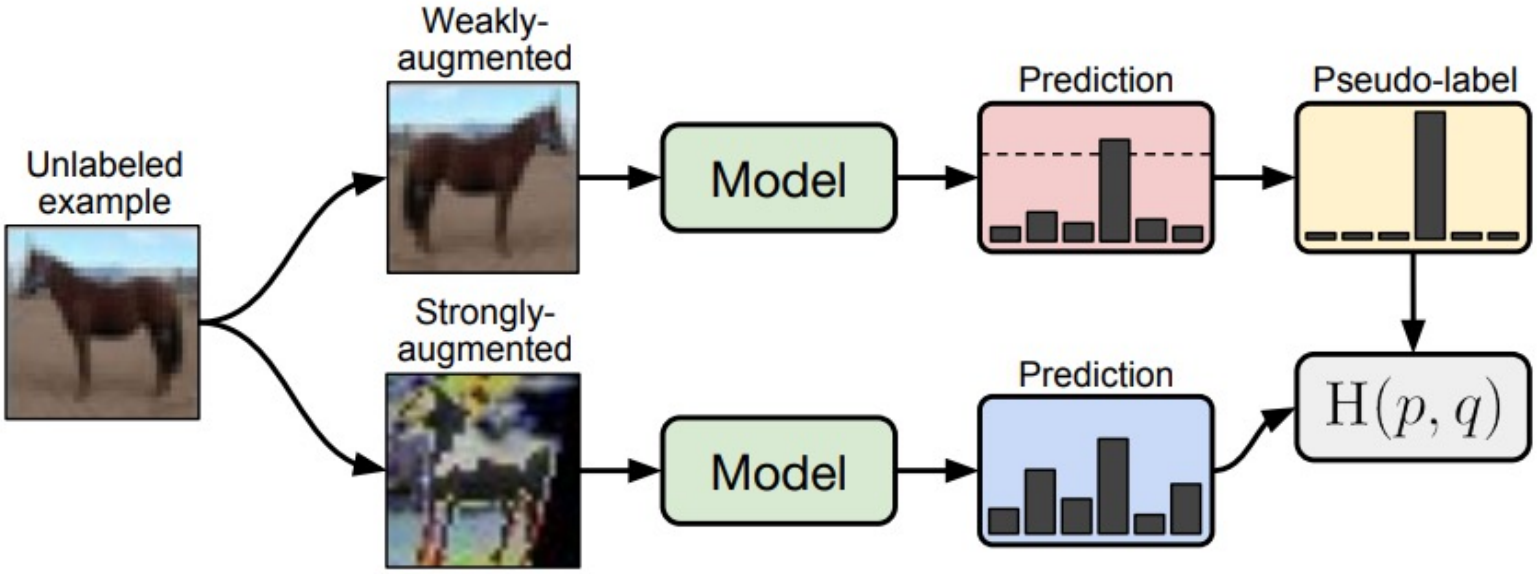
类别比例一致

Algorithm	SSL	RB	CIFAR-10 ($\gamma = \gamma_l = \gamma_u$) $\gamma = 50$	$\gamma = 100$	$\gamma = 150$
Vanilla	-	-	65.2±0.05 / 61.1±0.09	58.8±0.13 / 51.0±0.11	55.6±0.43 / 44.0±0.98
Re-sampling [22]	-	✓	64.3±0.48 / 60.6±0.67	55.8±0.47 / 45.1±0.30	52.2±0.05 / 38.2±1.49
LDAM-DRW [8]	-	✓	68.9±0.07 / 67.0±0.08	62.8±0.17 / 58.9±0.60	57.9±0.20 / 50.4±0.30
cRT [23]	-	✓	67.8±0.13 / 66.3±0.15	63.2±0.45 / 59.9±0.40	59.3±0.10 / 54.6±0.72
VAT [31]	✓	-	70.6±0.29 / 67.8±0.19	62.6±0.40 / 55.1±0.56	57.9±0.42 / 46.3±0.47
Mean-Teacher [40]	✓	-	68.8±1.05 / 64.9±1.53	60.9±0.33 / 52.8±0.81	54.5±0.22 / 39.8±0.73
MixMatch [5]	✓	-	73.2±0.56 / 68.9±1.15	64.8±0.28 / 49.0±2.05	62.5±0.31 / 42.5±1.68
MixMatch + DARP	✓	-	75.2±0.47 / 72.8±0.63	67.9±0.14 / 61.2±0.15	65.8±0.52 / 56.5±2.08
			(-7.41% / -12.6%)	(-8.77% / -23.8%)	(-8.69% / -24.4%)
ReMixMatch [4]	✓	-	81.5±0.26 / 80.2±0.32	73.8±0.38 / 69.5±0.84	69.9±0.47 / 62.5±0.35
ReMixMatch + DARP	✓	-	**82.1**±0.14 / 80.8±0.09	**75.8**±0.09 / 72.6±0.24	**71.0**±0.27 / 64.5±0.68
			(-3.45% / -3.52%)	(-7.84% / -10.2%)	(-3.60% / -5.19%)
FixMatch [39]	✓	-	79.2±0.33 / 77.8±0.36	71.5±0.72 / 66.8±1.51	68.4±0.15 / 59.9±0.43
FixMatch + DARP	✓	-	81.8±0.24 / **80.9**±0.28	75.5±0.05 / **73.0**±0.09	70.4±0.25 / **64.9**±0.17
			(-12.9% / -14.1%)	(-14.0% / -18.8%)	(-22.4% / -20.3%)

类别比例不一致

Algorithm	SSL	RB	CIFAR-10 ($\gamma_l = 100$)			
			$\gamma_u = 1$	$\gamma_u = 50$	$\gamma_u = 150$	$\gamma_u = 100$ (reversed)
Vanilla	-	-	58.8$\pm$0.13 / 51.0$\pm$0.11	58.8$\pm$0.13 / 51.0$\pm$0.11	58.8$\pm$0.13 / 51.0$\pm$0.11	58.8$\pm$0.13 / 51.0$\pm$0.11
Re-sampling [22]	-	✓	55.8$\pm$0.47 / 45.1$\pm$0.30	55.8$\pm$0.47 / 45.1$\pm$0.30	55.8$\pm$0.47 / 45.1$\pm$0.30	55.8$\pm$0.47 / 45.1$\pm$0.30
LDAM-DRW [8]	-	✓	62.8$\pm$0.17 / 58.9$\pm$0.60	62.8$\pm$0.17 / 58.9$\pm$0.60	62.8$\pm$0.17 / 58.9$\pm$0.60	62.8$\pm$0.17 / 58.9$\pm$0.60
cRT [23]	-	✓	63.2$\pm$0.45 / 59.9$\pm$0.40	63.2$\pm$0.45 / 59.9$\pm$0.40	63.2$\pm$0.45 / 59.9$\pm$0.40	63.2$\pm$0.45 / 59.9$\pm$0.40
VAT [31]	✓	-	65.2$\pm$0.12 / 59.5$\pm$0.26	64.0$\pm$0.31 / 57.3$\pm$0.66	62.8$\pm$0.19 / 55.1$\pm$0.70	59.4$\pm$0.36 / 50.6$\pm$0.61
Mean-Teacher [40]	✓	-	73.9$\pm$1.19 / 71.7$\pm$1.42	61.2$\pm$0.51 / 53.5$\pm$0.84	59.7$\pm$0.50 / 50.0$\pm$1.61	61.0$\pm$0.82 / 56.4$\pm$1.64
MixMatch [5]	✓	-	41.5$\pm$0.76 / 12.0$\pm$1.34	64.1$\pm$0.58 / 48.3$\pm$0.70	65.5$\pm$0.64 / 51.1$\pm$2.41	47.9$\pm$0.09 / 20.5$\pm$0.85
MixMatch + DARP	✓	-	86.7$\pm$0.80 / 86.2$\pm$0.82	68.3$\pm$0.47 / 62.2$\pm$1.21	66.7$\pm$0.25 / 58.8$\pm$0.42	72.9$\pm$0.24 / 71.0$\pm$0.32
			(-77.2% / -84.4%)	(-11.8% / -27.0%)	(-3.62% / -15.7%)	(-48.0% / -63.6%)
ReMixMatch [4]	✓	-	48.3$\pm$0.14 / 19.5$\pm$0.85	75.1$\pm$0.43 / 71.9$\pm$0.77	72.5$\pm$0.10 / 68.2$\pm$0.32	49.0$\pm$0.55 / 17.1$\pm$1.48
ReMixMatch*	✓	-	85.0$\pm$1.35 / 84.3$\pm$1.55	77.0$\pm$0.12 / 74.7$\pm$0.04	72.8$\pm$0.10 / 68.8$\pm$0.21	75.3$\pm$0.03 / 72.3$\pm$0.04
ReMixMatch* + DARP	✓	-	**89.7**$\pm$0.15 / **89.4**$\pm$0.17	**77.4**$\pm$0.22 / 75.0$\pm$0.25	**73.2**$\pm$0.11 / 69.2$\pm$0.31	**80.1**$\pm$0.11 / **78.5**$\pm$0.17
			(-31.4% / -32.5%)	(-1.72% / -1.49%)	(-1.53% / -2.64%)	(-19.5% / -22.5%)
FixMatch [39]	✓	-	68.9$\pm$1.95 / 42.8$\pm$8.11	73.9$\pm$0.25 / 70.5$\pm$0.52	69.6$\pm$0.60 / 62.6$\pm$1.11	65.5$\pm$0.05 / 26.0$\pm$0.44
FixMatch + DARP	✓	-	85.4$\pm$0.55 / 85.0$\pm$0.65	77.3$\pm$0.17 / **75.5**$\pm$0.21	72.9$\pm$0.24 / **69.5**$\pm$0.18	74.9$\pm$0.51 / 72.3$\pm$1.13
			(-53.1% / -73.8%)	(-13.3% / -17.0%)	(-10.9% / -18.4%)	(-31.3% / -60.3%)

FixMatch



$$\mathcal{L}_u = \frac{1}{\mu B} \sum_{b=1}^{\mu B} \mathbb{I}(\max(\mathbf{q}_b) \geq \tau) H(\hat{\mathbf{y}}_b^u, f(\mathbf{y}|\mathcal{A}(\mathbf{x}_b^u); \theta))$$

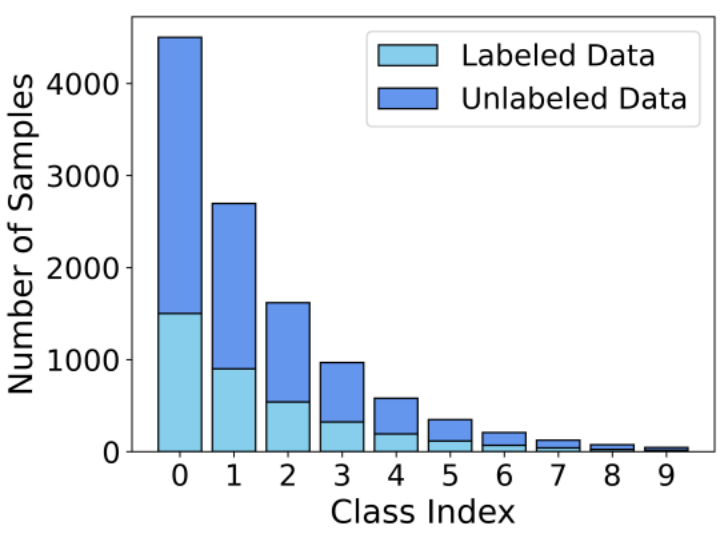
类相关阈值

$$\mathcal{L}_u = \frac{1}{\mu B} \sum_{b=1}^{\mu B} \mathbb{I}(\max(\mathbf{q}_b) \geq \tau) H(\hat{\mathbf{y}}_b^u, f(\mathbf{y}|\mathcal{A}(\mathbf{x}_b^u); \theta))$$

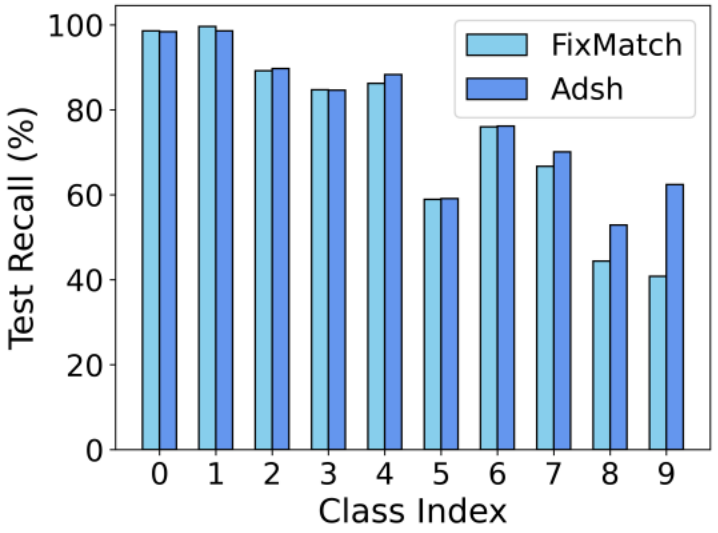
如何为不同的类别选择不同的τ？

- 伪标注数据类别比例符合真实分布
- 有标注数据和伪标注数据类别比例一致
- ...

类相关阈值:Adsh

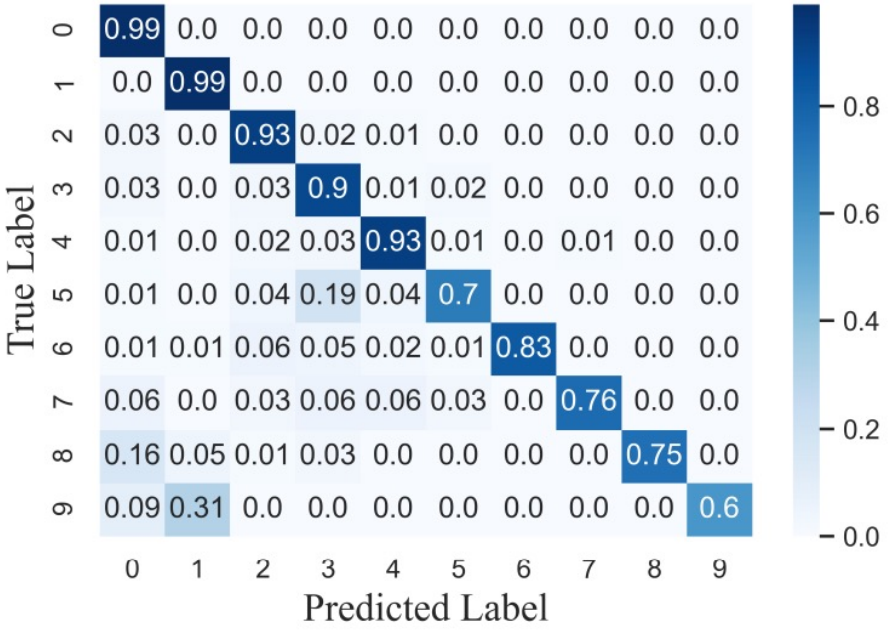


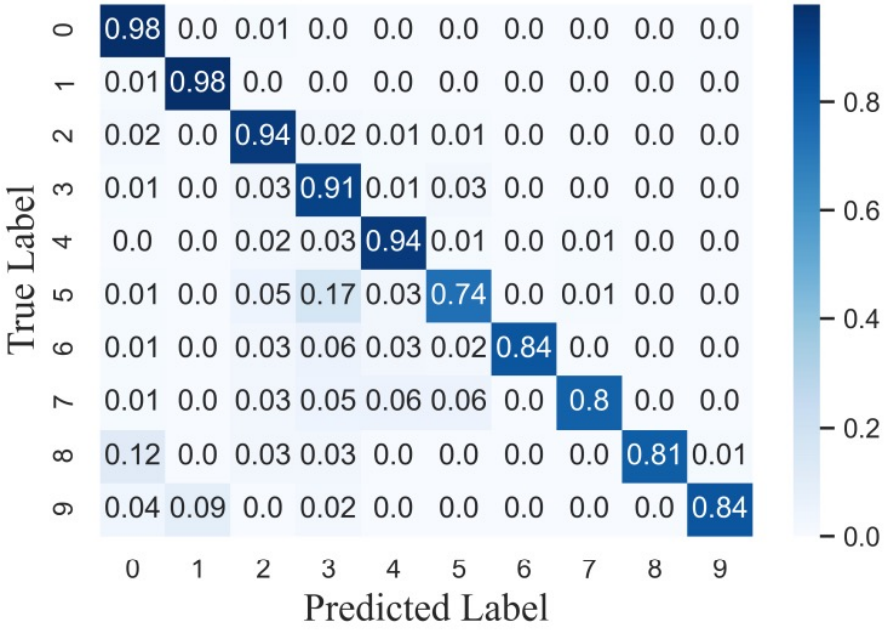
(a) Imbalanced Dataset



(b) Test Recall (%)

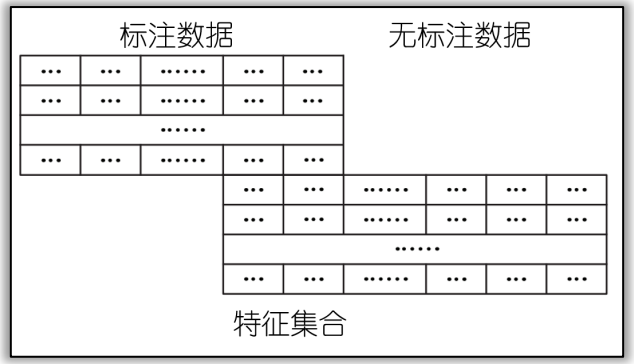
类相关阈值:Adsh


True Label
Predicted Label
0 1 2 3 4 5 6 7 8 9
0.99 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
0.0 0.99 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
0.03 0.0 0.93 0.02 0.01 0.0 0.0 0.0 0.0 0.0
0.03 0.0 0.03 0.9 0.01 0.02 0.0 0.0 0.0 0.0
0.01 0.0 0.02 0.03 0.93 0.01 0.0 0.01 0.0 0.0
0.01 0.0 0.04 0.19 0.04 0.7 0.0 0.0 0.0 0.0
0.01 0.01 0.06 0.05 0.02 0.01 0.83 0.0 0.0 0.0
0.06 0.0 0.03 0.06 0.06 0.03 0.0 0.76 0.0 0.0
0.16 0.05 0.01 0.03 0.0 0.0 0.0 0.0 0.75 0.0
0.09 0.31 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.6
0.8
0.6
0.4
0.2
0.0


True Label
Predicted Label
0 1 2 3 4 5 6 7 8 9
0.98 0.0 0.01 0.0 0.0 0.0 0.0 0.0 0.0 0.0
0.01 0.98 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0
0.02 0.0 0.94 0.02 0.01 0.01 0.0 0.0 0.0 0.0
0.01 0.0 0.03 0.91 0.01 0.03 0.0 0.0 0.0 0.0
0.0 0.0 0.02 0.03 0.94 0.01 0.0 0.01 0.0 0.0
0.01 0.0 0.05 0.17 0.03 0.74 0.0 0.01 0.0 0.0
0.01 0.0 0.03 0.06 0.03 0.02 0.84 0.0 0.0 0.0
0.01 0.0 0.03 0.05 0.06 0.06 0.0 0.8 0.0 0.0
0.12 0.0 0.03 0.03 0.0 0.0 0.0 0.0 0.81 0.01
0.04 0.09 0.0 0.02 0.0 0.0 0.0 0.0 0.0 0.84
0.8
0.6
0.4
0.2
0.0

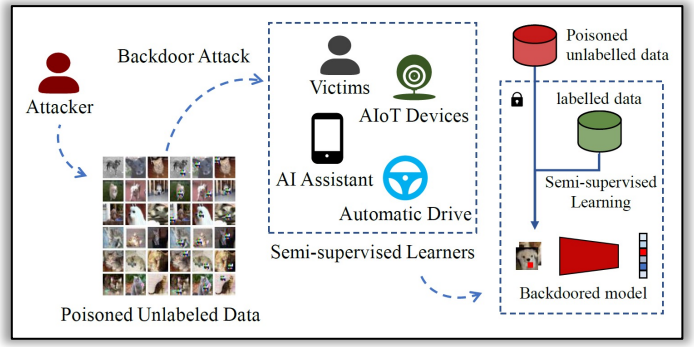
开放问题

样本属性不一致



✓ 如何有效利用与标注数据特征集合有差异的无标注样本？

对抗样本



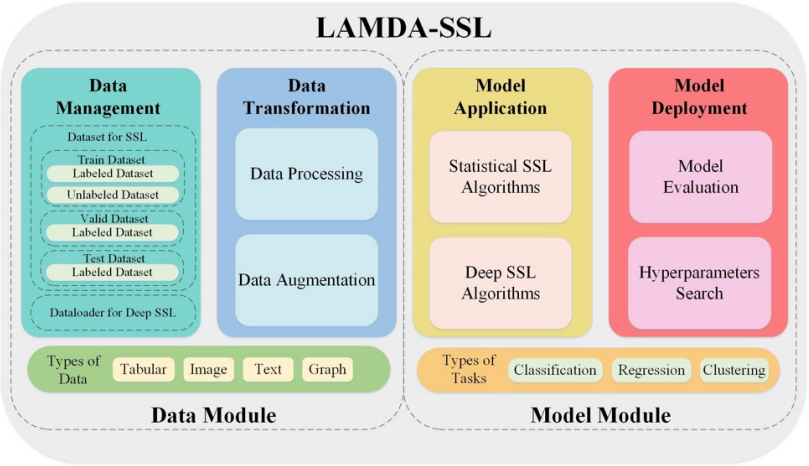
✓ 如何生成针对半监督学习的对抗攻击样本

✓ 如何提升半监督学习对对抗攻击的鲁棒性

SSL工具包

半监督学习工具包：LAMDA-SSL

- ✓ 支持30余种半监督学习算法，4种数据类型，16种评价指标
- ✓ Github: https://github.com/YGZWQZD/LAMDA-SSL




```python
from LAMDA_SSL.Dataset.Vision.CIFAR10 import CIFAR10
from LAMDA_SSL.Algorithm.Classification.FixMatch import FixMatch
from LAMDA_SSL.Evaluation.Classifier.Accuracy import Accuracy
# Initialize CIFAR10 dataset
dataset=CIFAR10(root='..\Download\cifar-10-python',labeled_size=4000)
labeled_X, labeled_y=dataset.labeled_X,dataset.labeled_Y
unlabeled_X=dataset.unlabeled_X
test_X, test_y=dataset.test_X, dataset.test_y
# Initialize FixMatch algorithm
model=FixMatch(threshold=0.95,lambda_u=1.0,T=0.5,mu=7,
               epoch=1,num_it_epoch=2**20,device='cuda:0')
# Call the fit() method to Train the model
model.fit(X=labeled_X,y=labeled_y,unlabeled_X=unlabeled_X)
# Call the predict() method to predict the labels of new samples
y_pred=model.predict(test_X)
# Evaluate the model' s performance.
performance=Accuracy().scoring(test_y,y_pred)
```