



面向标记分布学习的 标记增强

2018.5.15, 成都

耿 新

模式学习与挖掘实验室 (PALM)
<http://palm.seu.edu.cn>

东南大学计算机科学与工程学院



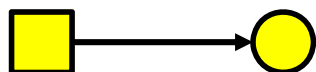
报告内容

- 研究背景
- 标记增强
 - 概念定义
 - 学习算法
 - 实验
- 结论

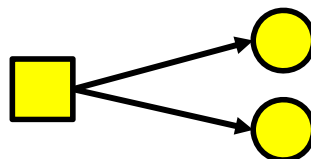


标记端多义性

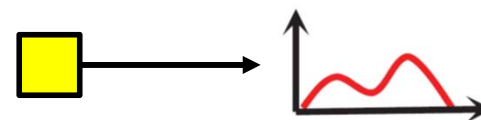
标记分布: 覆盖所有可能的标记, 并显式表示每个标记对示例的重要程度



单标记学习



多标记学习



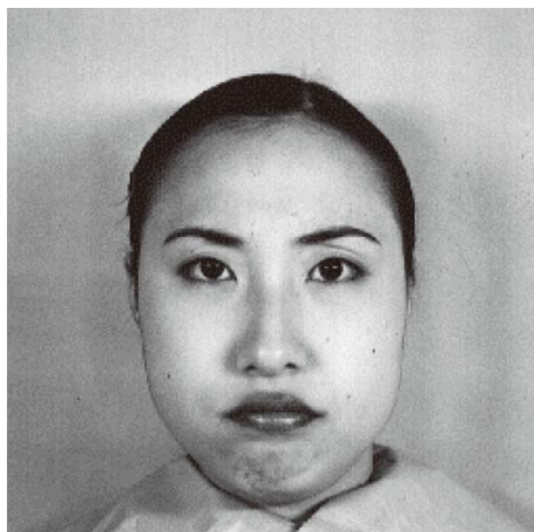
标记分布学习

明确

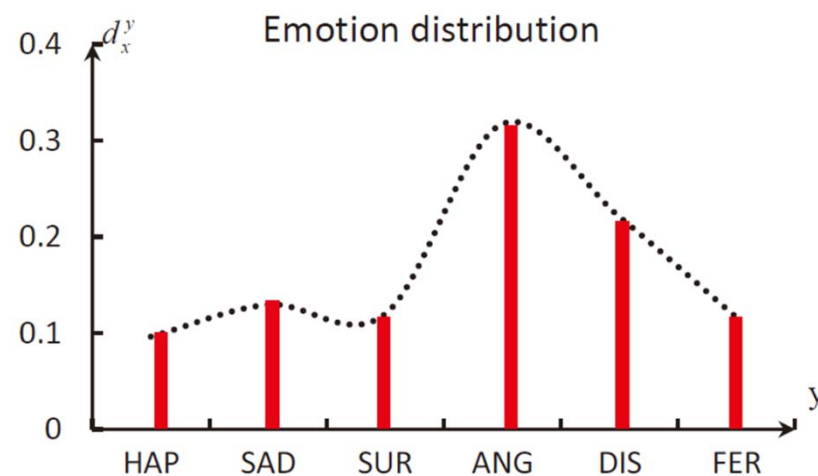
标记端多义性

多义

举例1：表情情感分布



情感	人工打分
高兴 (HAP)	1.48
悲伤 (SAD)	2.48
惊讶 (SUR)	2.19
气愤 (ANG)	4.48
恶心 (DIS)	3.42
害怕 (FER)	2.16

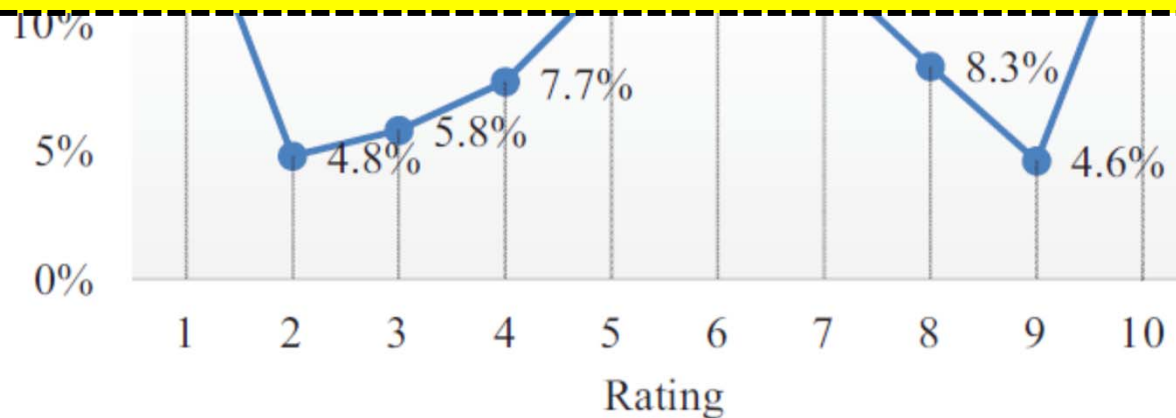


举例2：电影评分分布



Title	Twilight
Average Rating	5.2/10
Budget	\$ 37 Million
Gross	\$ 191 Million

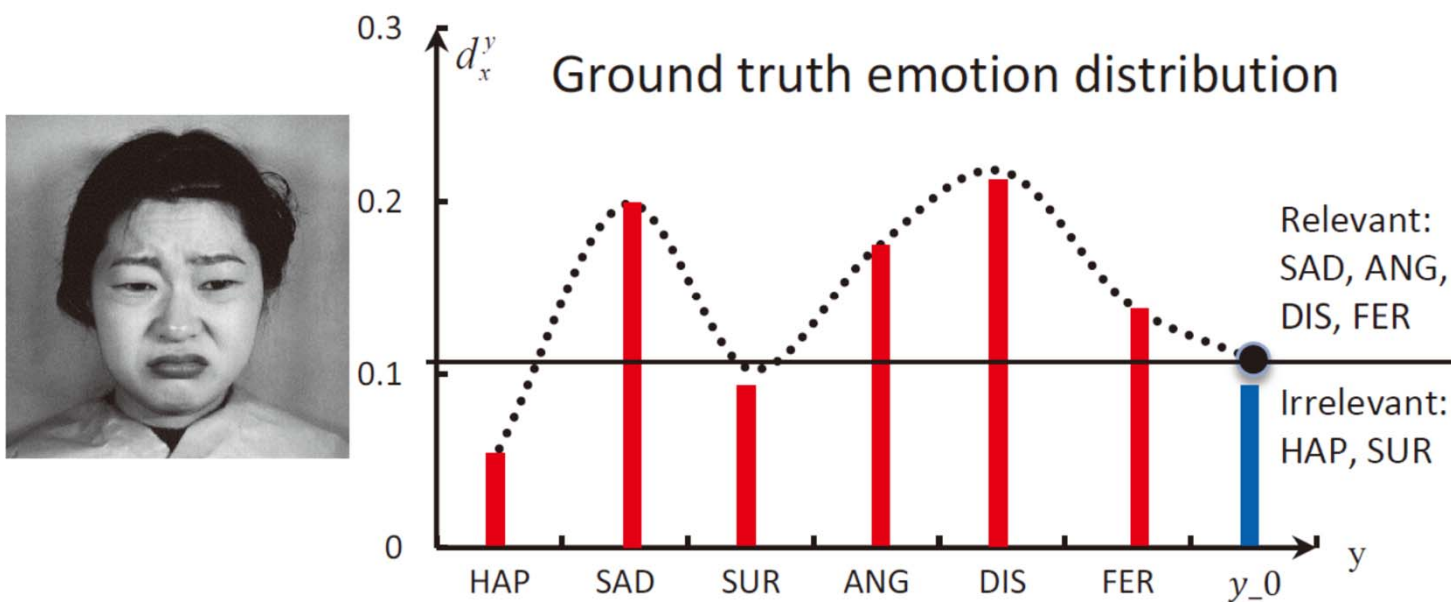
更多含有标记分布的真实问题和数据，可从如下链接下载：
<http://cse.seu.edu.cn/PersonalPage/xgeng/LDL>



本质上的普适性

- 本质上，标记与示例的相关或不相关是相对的
 - 相关与不相关的划分

将标记集合分为相关和不相关两个子集往往是对实际问题的简化，而实际上两者的“边界”可能没那么明确



本质上的普适性

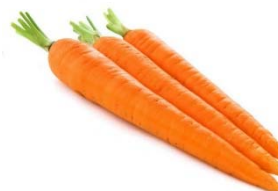
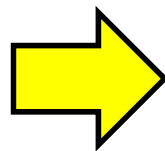
- 本质上，标记与示例的相关或不相关是相对的
 - 对于相关标记

当一个示例与多个标记相关时，这些标记对于这个示例的重要程度一般情况下不会恰好完全一样



本质上的普适性

- 本质上，标记与示例的相关或不相关是相对的
 - 对于不相关标记
 - 对一个示例来说，众多不相关标记的“不相关性”可能差异巨大



实践上的局限性

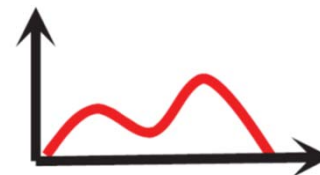
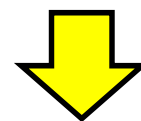
- 实践中，直接获得每个标记的重要程度在许多应用中并不现实
 - 代价高昂
 - 难以量化
- 多数现有数据的监督信息是对问题的简化，即逻辑标记
 - 1: 相关标记
 - 0: 不相关标记

需要一种方法，将简单逻辑
标记转化为标记分布

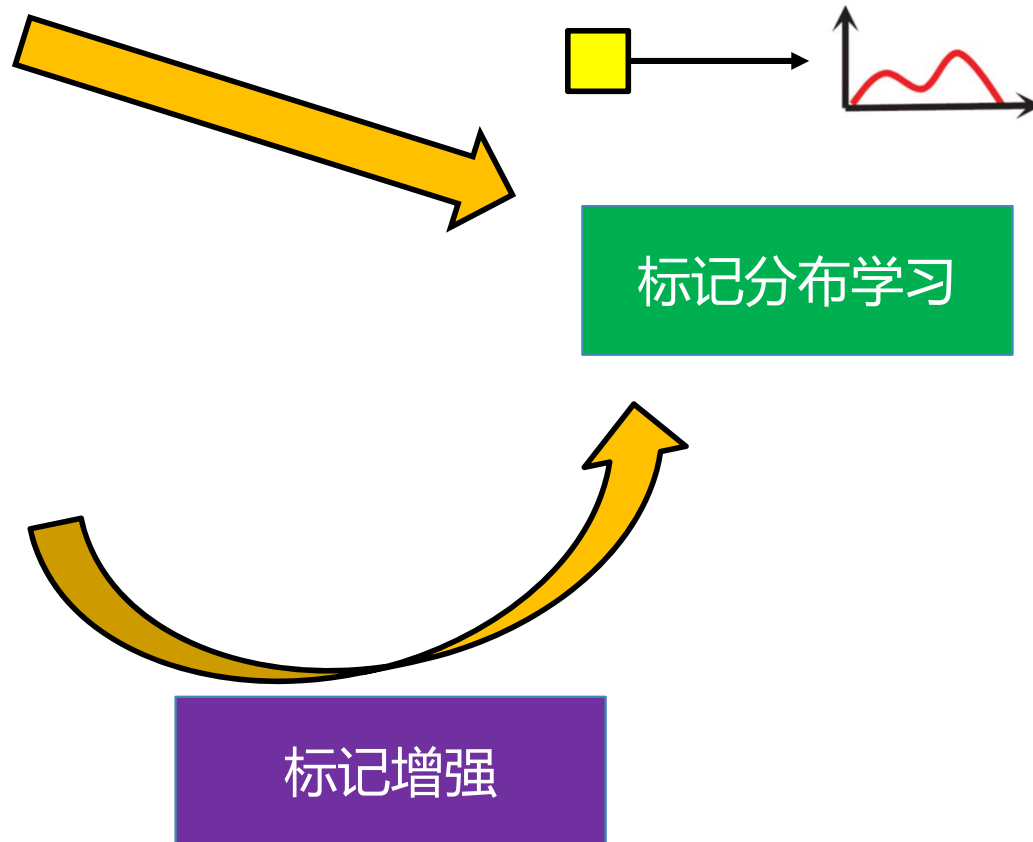
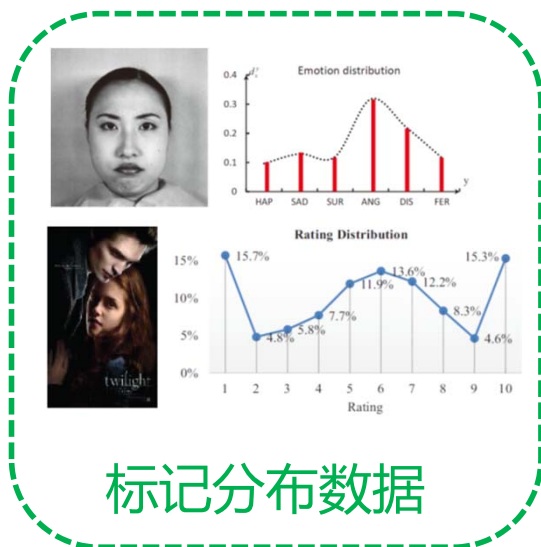


标记增强

$\{0, 1, 0, 1, 0\}$



标记分布 & 标记增强



报告内容

- 研究背景
- 标记增强
 - 概念定义
 - 学习算法
 - 实验
- 结论





标记增强形式化定义

[耿新 et al., 《计算机研发》, 2017]

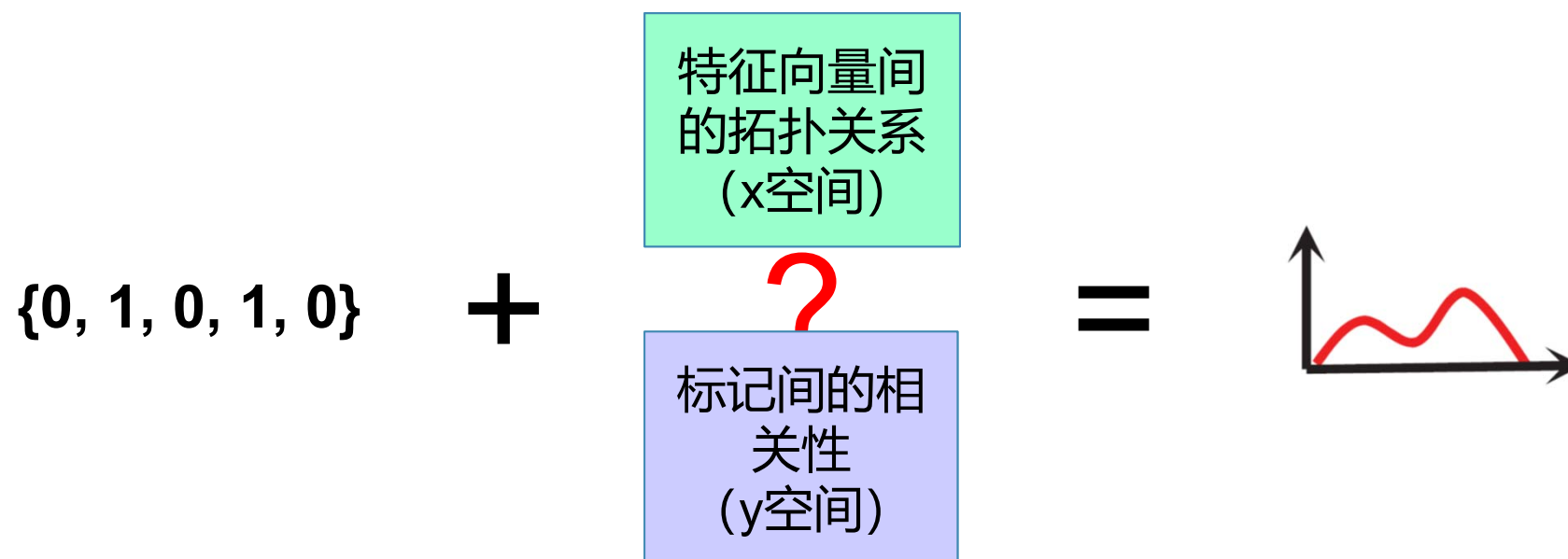
示例 x_i 的**逻辑标记** $L_i = [l_{x_i}^{y_1}, l_{x_i}^{y_2}, \dots, l_{x_i}^{y_c}]$, 其中 $l_{x_i}^{y_j} \in \{0,1\}$ 表示 y_j 是否是 x_i 的相关标记, c 是可能的标记数目, 则 $L_i \in \{0,1\}^c$

x_i 的**标记分布** $D_i = [d_{x_i}^{y_1}, d_{x_i}^{y_2}, \dots, d_{x_i}^{y_c}]$, 其中 $d_{x_i}^{y_j} \in [0,1]$ 并且 $\sum_j d_{x_i}^{y_j} = 1$ 表示 y_j 对 x_i 的描述度, 则 $D_i \in [0,1]^c$

标记增强定义如下:

给定训练集 $S = \{(x_i, L_i) | 1 \leq i \leq n\}$, 标记增强即将每个示例 x_i 的逻辑标记 L_i 转化为相应的标记分布 D_i , 从而得到标记分布训练集 $E = \{(x_i, D_i) | 1 \leq i \leq n\}$ 的过程。

增加了什么?



报告内容

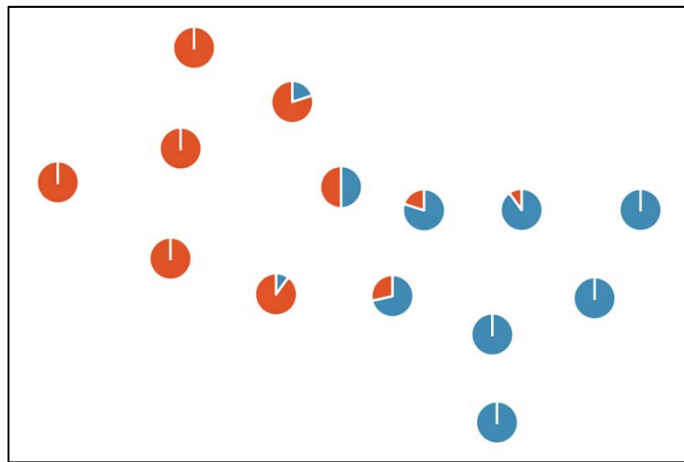
- 研究背景
- 标记增强
 - 概念定义
 - 学习算法
 - 实验
- 结论



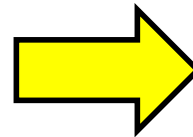
标记增强算法

- 基于模糊方法的标记增强
 - 基于模糊聚类的标记增强
[Gayar et al., ANNPR'06]
 - 基于核隶属度的标记增强
[Jiang et al., NCA, 2006]
- 基于图的标记增强
 - 基于标记传播的标记增强
[Li et al., ICDM'15]
 - 基于流形学习的标记增强
[Hou et al., AACL'16]
 - 基于图拉普拉斯矩阵的标记增强
[Xu and Geng, IJCAI'18]

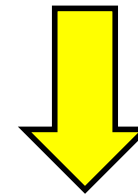
基于模糊聚类的标记增强 (FCM) [Gayar et al., ANNPR'06]



模糊聚类
(示例对聚类的隶属度)



叠加属于同一类（标记）
的示例的隶属度，形成聚
类-标记关联矩阵



利用模糊合成运算，将示
例对聚类的隶属度，通过
聚类-标记关联矩阵转化为
示例对标记的隶属度

基于模糊聚类的标记增强 (FCM) [Gayar et al., ANNPR'06]

- **第一步**: 模糊C均值聚类 (FCM)

1. 选定聚类数 p , 初始化 $n \times p$ 聚类隶属度矩阵 M (m_{ik} 表示 x_i 对聚类 k 的隶属度)
2. 计算聚类中心

$$\mu_k = \frac{\sum_{i=1}^n (m_{ik})^\beta x_i}{\sum_{i=1}^n (m_{ik})^\beta}$$

3. 更新聚类隶属度矩阵 M

$$m_{ik} = \frac{1}{\sum_{j=1}^p \left(\frac{Dist(x_i, \mu_k)}{Dist(x_i, \mu_j)} \right)^{\frac{2}{\beta-1}}}$$

4. 重复以上2、3步直至 M 收敛

M 的每一行 m_i 示例 x_i 对于每个聚类的隶属度

基于模糊聚类的标记增强 (FCM)

[Gayar et al., ANNPR'06]

- **第二步**: 计算聚类-标记关联矩阵

1. 初始化 $c \times p$ 矩阵零矩阵 A

2. 用如下公式更新 A 的第 j 行 A_j

$$A_j = A_j + m_i, \quad \text{if } l_{x_i}^{y_j} = 1$$

3. 矩阵 A 列归一化

4. 矩阵 A 行归一化

A 的元素 a_{jk} 表示标记 j 与聚类 k 的关联强度

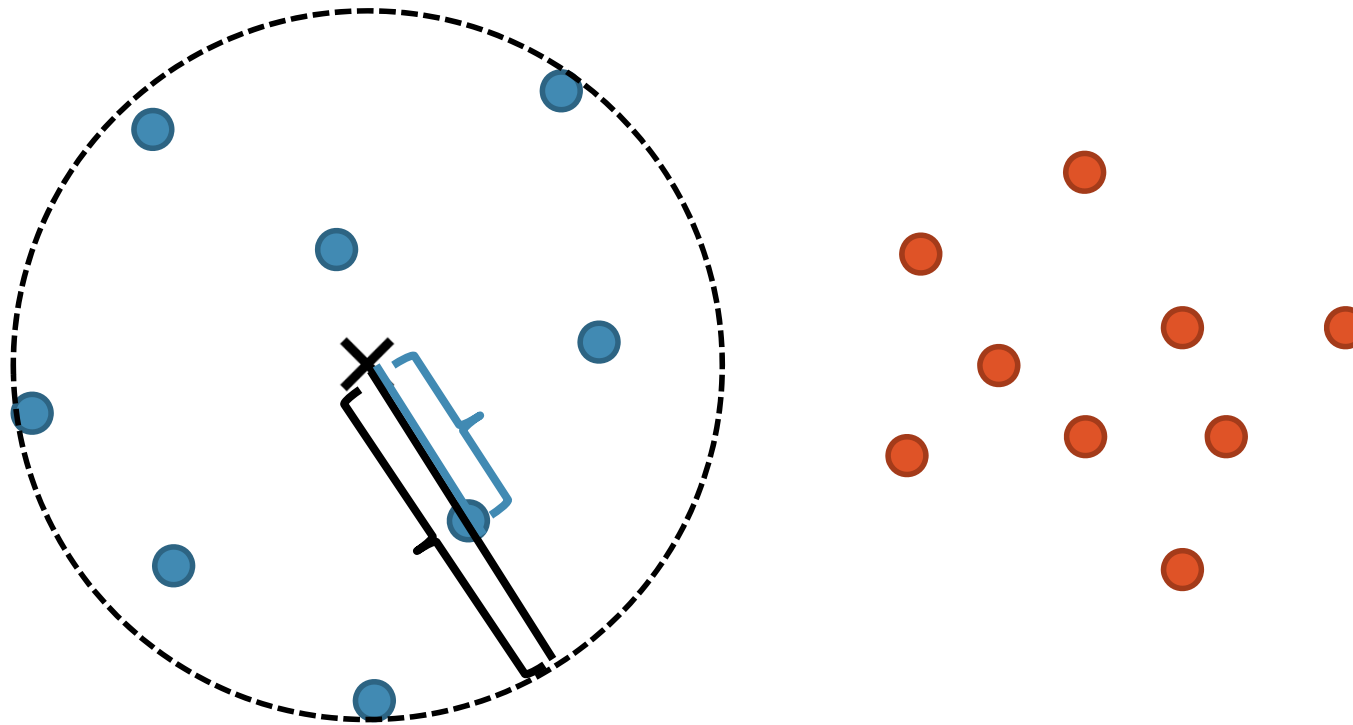
- **第三步**: 计算每个示例 x_i 的标记分布

1. $D_i = A \circ m_i$ (模糊合成运算)

$$D_i^j = \max_k (a_{jk} \times m_{ik})$$

2. D_i 归一化即得到 x_i 的标记分布

基于核隶属度的标记增强 (KM) [Jiang et al., NCA, 2006]



利用核技巧引入非线性关系



基于核隶属度的标记增强 (KM) [Jiang et al., NCA, 2006]

- **第一步**: 对标记 y_j , 假设所有用 y_j 标注的示例集合为 C^{y_j} , 个数为 n_j , 计算类中心 Ψ^{y_j} :

$$\Psi^{y_j} = \frac{1}{n_j} \sum_{x_i \in C^{y_j}} \phi(x_i)$$

其中 $\phi(x_i)$ 是一个非线性映射函数, 由核函数 $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$ 确定

- **第二步**: 计算类半径

$$r_j = \max_{x_i \in C^{y_j}} \|\Psi^{y_j} - \phi(x_i)\|,$$

r_j^2 的计算可用 $\phi(x_i)$ 的内积实现

- **第三步**: 计算示例 x_i 到类中心的平方距离

$$d_{ij}^2 = \|\phi(x_i) - \Psi^{y_j}\|^2$$

d_{ij}^2 的计算可用 $\phi(x_i)$ 的内积实现

基于核隶属度的标记增强 (KM) [Jiang et al., NCA, 2006]

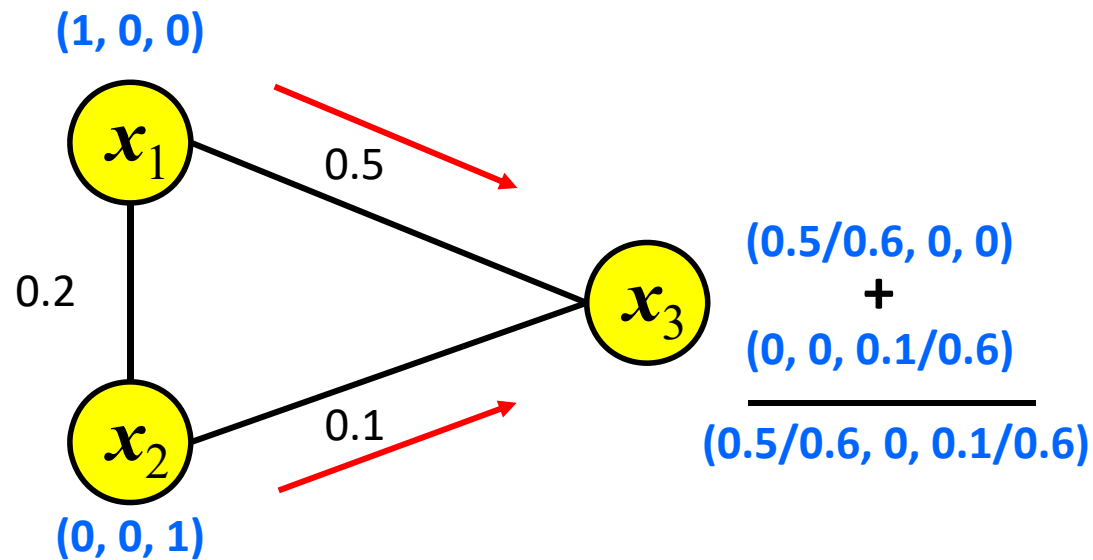
- 第四步：计算示例 x_i 对于标记 y_j 的隶属度

$$m_{x_i}^{y_j} = \begin{cases} 1 - \sqrt{\frac{\|d_{ij}^2\|}{(r_j^2 + \delta)}} & \text{if } l_{x_i}^{y_j} = 1 \\ 0 & \text{if } l_{x_i}^{y_j} = 0 \end{cases}$$

无法改变不相关标记的隶属度

- 第五步：对 $\mathbf{m}_{x_i} = [m_{x_i}^{y_1}, m_{x_i}^{y_2}, \dots, m_{x_i}^{y_c}]$ 归一化，即可得到 x_i 的标记分布 \mathbf{D}_i

基于标记传播的标记增强 (LP) [Li, Zhang and Geng, ICDM'15]



训练集中的标记传播 (Label Propagation)

基于标记传播的标记增强 (LP) [Li, Zhang and Geng, ICDM'15]

$$G = (V, E) \quad V = \{\mathbf{x}_i \mid 1 \leq i \leq m\}$$

$$\forall_{i,j=1}^m: w_{ij} = \begin{cases} \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|_2^2}{2\sigma^2}\right), & \text{if } i \neq j \\ 0, & \text{if } i = j \end{cases}$$

Propagation Matrix

$$\mathbf{P} = \mathbf{D}^{-\frac{1}{2}} \mathbf{W} \mathbf{D}^{-\frac{1}{2}} \quad \mathbf{D} = \text{diag}[d_1, d_2, \dots, d_m] \quad d_i = \sum_{j=1}^m w_{ij}$$

$$\mathbf{F}^{(0)} = \Phi \longrightarrow \forall_{i=1}^m \forall_{l=0}^q: \phi_{il} = \begin{cases} \tau, & \text{if } y_l = y_0 \\ 1, & \text{if } y_l \in Y_i \\ 0, & \text{otherwise} \end{cases}$$

Label Propagation

$$\mathbf{F}^{(t)} = \alpha \mathbf{P} \mathbf{F}^{(t-1)} + (1 - \alpha) \Phi$$

Converge to

$$\mathbf{F}^* = (1 - \alpha)(\mathbf{I} - \alpha \mathbf{P})^{-1} \Phi$$

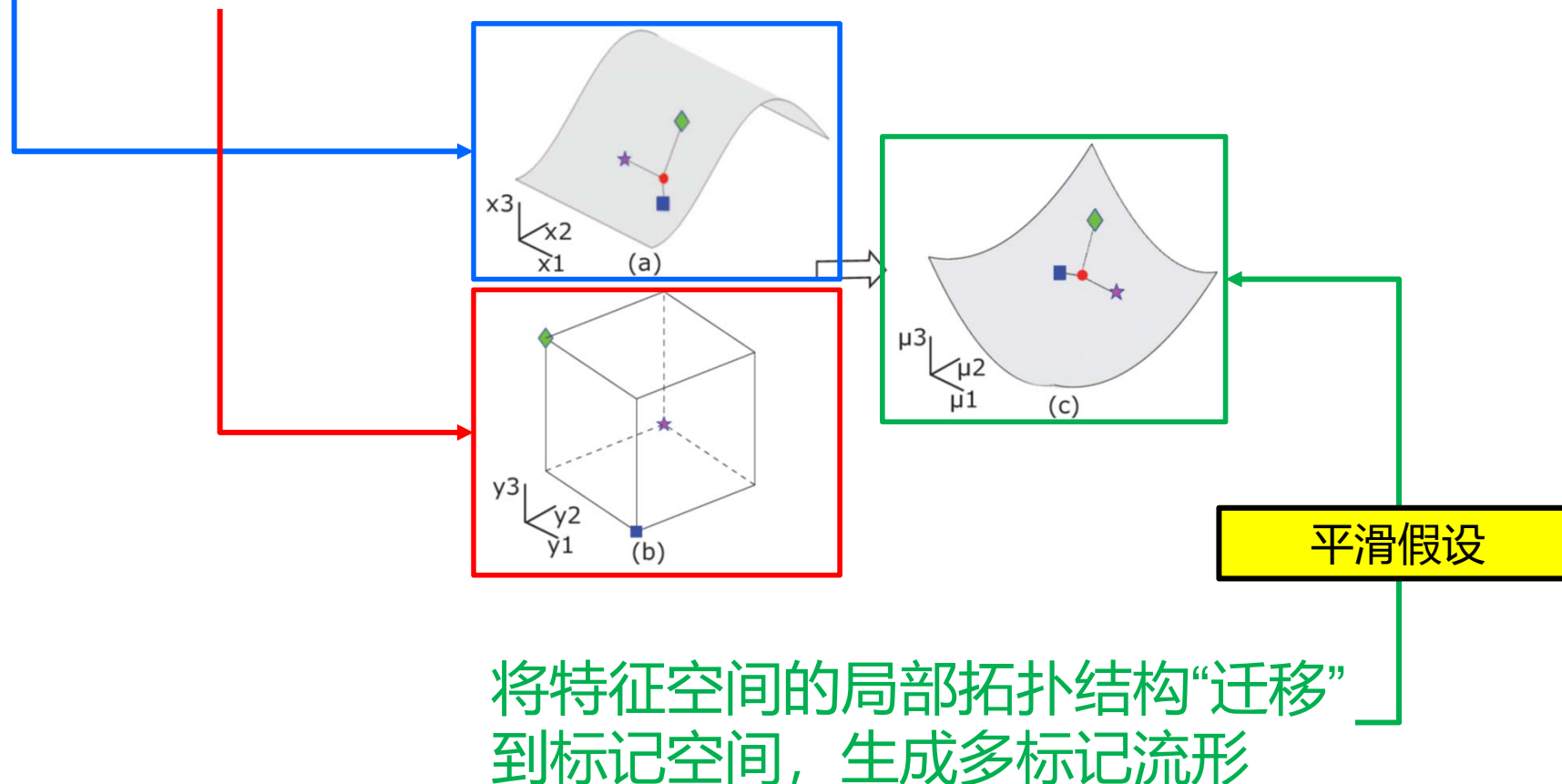
$$\mu_{\mathbf{x}_i}^{y_l} = \frac{f_{il}^*}{\sum_{k=0}^q f_{ik}^*}$$

Label Distribution

基于流形学习的标记增强 (ML)

[Hou, Geng and Zhang, AAAI'16]

- **特征空间**: 连续欧式空间, 示例分布在流形上
- **标记空间**: 离散逻辑空间, 标记分布在超立方体顶点上



基于流形学习的标记增强 (ML)

[Hou, Geng and Zhang, AAAI'16]

- 特征空间的流形学习 [Roweis & Saul, Science, 2000]

$$\arg \min_W \sum_{i=1}^n \|\mathbf{x}_i - \sum_{j \neq i} W_i^j \mathbf{x}_j\|^2$$

$$\text{s.t. } \mathbf{1}^T \mathbf{W}_i = 1$$

如果 \mathbf{x}_j 不是 \mathbf{x}_i 的 k 近邻, 则 $W_i^j = 0$

局部拓扑
关系迁移

- 标记空间的流形学习

$$\arg \min_{\mu} \sum_{i=1}^n \|\mu_i - \sum_{j \neq i} W_i^j \mu_j\|^2$$

$$\text{s.t. } \forall 1 \leq i \leq n, 1 \leq l \leq q$$

$$y_i^l = \begin{cases} +1, & \text{if the } l\text{-th label is relevant} \\ -1, & \text{if the } l\text{-th label is irrelevant} \end{cases}$$

$$y_i^l \mu_i^l \geq \lambda, \lambda > 0$$

控制符号和尺度

基于图拉普拉斯矩阵的标记增强 (GLLE)

[Xu and Geng, IJCAI'18]

- 模型

非线性变换

$$\mathbf{D}_i = \mathbf{W}^\top \varphi(\mathbf{x}_i) + \mathbf{b} = \widehat{\mathbf{W}} \boldsymbol{\phi}_i$$

Goal

找到最优的参数 $\widehat{\mathbf{W}}^*$, 使得 \mathbf{D}_i 最合理

- 优化目标

逻辑标记的信息

$$\min_{\widehat{\mathbf{W}}} L(\widehat{\mathbf{W}}) + \lambda R(\widehat{\mathbf{W}})$$

特征空间的拓扑关系约束了标记间的相关性

基于图拉普拉斯矩阵的标记增强 (GLLE)

[Xu and Geng, IJCAI'18]

- 目标函数第一部分

$$L(\widehat{\mathbf{W}}) = \sum_{i=1}^n \|\widehat{\mathbf{W}} \boldsymbol{\phi}_i - \mathbf{L}_i\|^2 \quad \text{平方损失 (LS)}$$

- 目标函数第二部分

平滑假设

$$\begin{aligned} R(\widehat{\mathbf{W}}) &= \sum_{i,j} a_{ij} \|\mathbf{D}_i - \mathbf{D}_j\|^2 \\ &= \text{tr}(\mathbf{D} \mathbf{G} \mathbf{D}^\top) \end{aligned} \quad \begin{aligned} &\xrightarrow{\text{red arrow}} a_{ij} = \begin{cases} \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right) & \text{if } \mathbf{x}_j \in N(i) \\ 0 & \text{otherwise} \end{cases} \\ &\quad \mathbf{x}_i \text{ 与 } \mathbf{x}_j \text{ 的相似度} \end{aligned}$$

Graph Laplacian

$$\mathbf{G} = \widehat{\mathbf{A}} - \mathbf{A}, \quad \hat{a}_{ij} = \sum_{j=1}^n a_{ij}$$

报告内容

- 研究背景
- 标记增强
 - 概念定义
 - 学习算法
 - 实验
- 结论



数据集

No.	Dataset	#Examples	#Features	#Labels
1	Artificial	2601	3	3
2	SJAFFE	213	243	6
3	Natural Scene	2,000	294	9
4	Yeast-spoem	2,465	24	2
5	Yeast-spo5	2,465	24	3
6	Yeast-dtt	2,465	24	4
7	Yeast-cold	2,465	24	4
8	Yeast-heat	2,465	24	6
9	Yeast-spo	2,465	24	6
10	Yeast-diau	2,465	24	7
11	Yeast-elu	2,465	24	14
12	Yeast-cdc	2,465	24	15
13	Yeast-alpha	2,465	24	18
14	SBU_3DFE	2,500	243	6
15	Movie	7,755	1,869	5

人造数据集

- 示例 $\mathbf{x} = [x_1, x_2, x_3]$ 的标记分布 $\mathbf{D} = [d_x^{y_1}, d_x^{y_2}, d_x^{y_3}]$ 由以下方法产生:

$$t_i = ax_i + bx_i^2 + cx_i^3 + d, i = 1, \dots, 3,$$

$$\psi_1 = (\boldsymbol{\omega}_1^T \mathbf{t})^2, \quad \psi_2 = (\boldsymbol{\omega}_2^T \mathbf{t} + \lambda_1 \psi_1)^2, \quad \psi_3 = (\boldsymbol{\omega}_3^T \mathbf{t} + \lambda_2 \psi_2)^2,$$

$$d_x^{y_i} = \frac{\psi_i}{\psi_1 + \psi_2 + \psi_3}, i = 1, \dots, 3,$$

其中 $a = 1, b = 0.5, c = 0.2, d = 1, \boldsymbol{\omega}_1^T = [4, 2, 1], \boldsymbol{\omega}_2^T = [1, 2, 4], \boldsymbol{\omega}_3^T = [1, 4, 2], \lambda_1 = \lambda_2 = 0.01$ 。

- 训练样本采样

\mathbf{x} 中的 x_1, x_2 分量在 $[-1, 1]^2$ 范围内按照间隔为 0.04 的网格采样 (2601 个采样点), 而 x_3 由下式计算得到:

$$x_3 = \sin((x_1 + x_2) \times \pi)$$

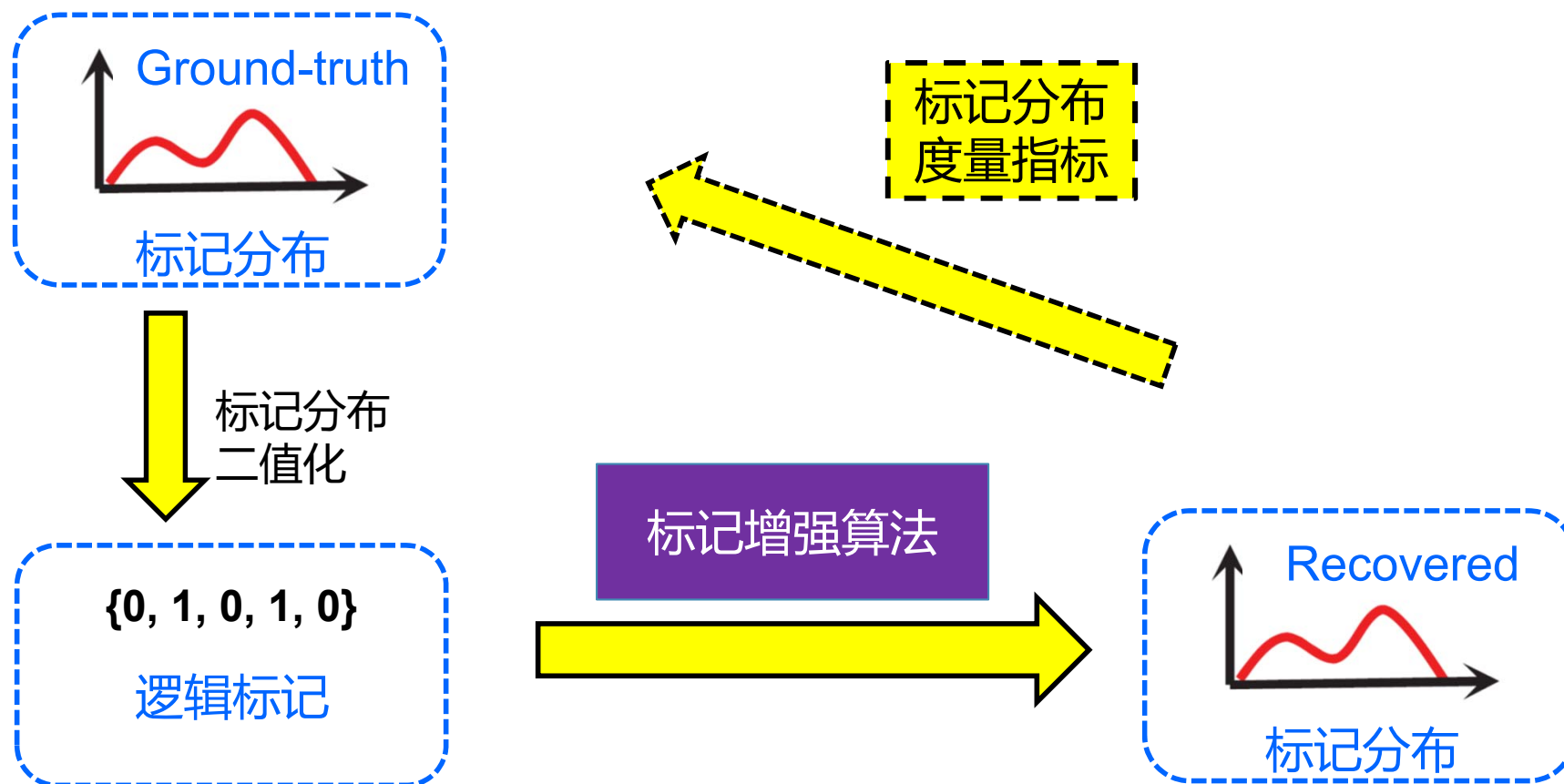
数据集

- 标记分布二值化

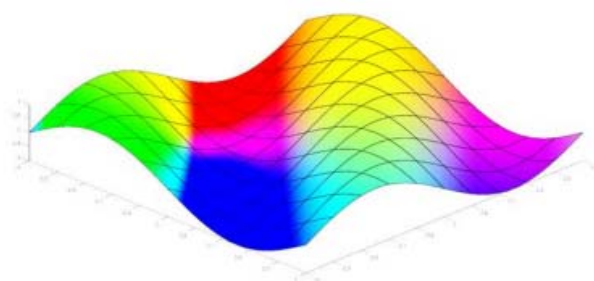
1. 相关标记集合 $y^+ = \emptyset$, 不相关标记集合 $y^- = y$, $L = 0$
2. 选取 y^- 中描述度最大的标记 y_j 作为相关标记, $y^+ = y^+ + y_j$,
 $y^- = y^- - y_j$, $l_x^{y_j} = 1$;
3. $t = \sum_{y_j \in y^+} d_x^{y_j}$, 如果 $t < T$ 则回到步骤2, 否则结束

$T = 0.5$

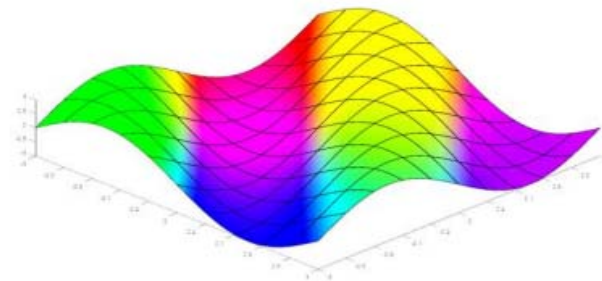
标记分布恢复实验



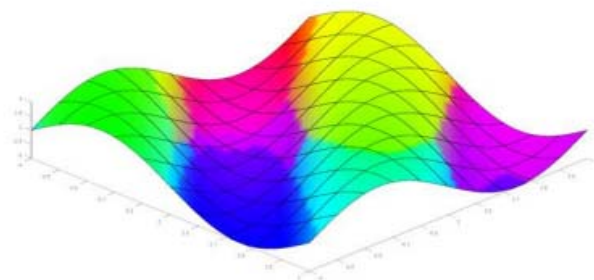
实验结果



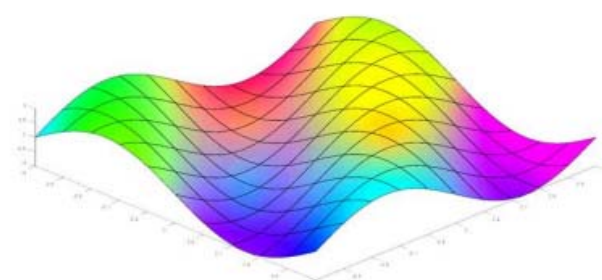
(a) Ground-Truth



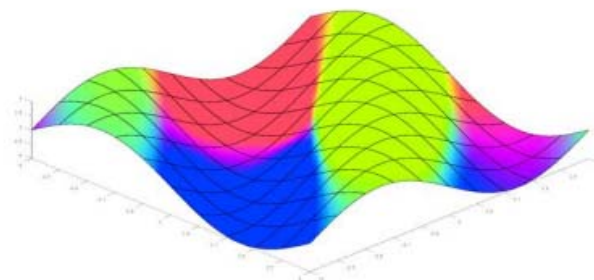
(b) GLLE



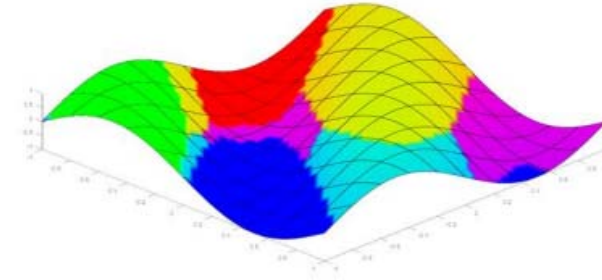
(c) LP



(d) ML



(e) FCM



(f) KM

实验结果

Recovery Results (value(rank)) Measured by Cheb ↓

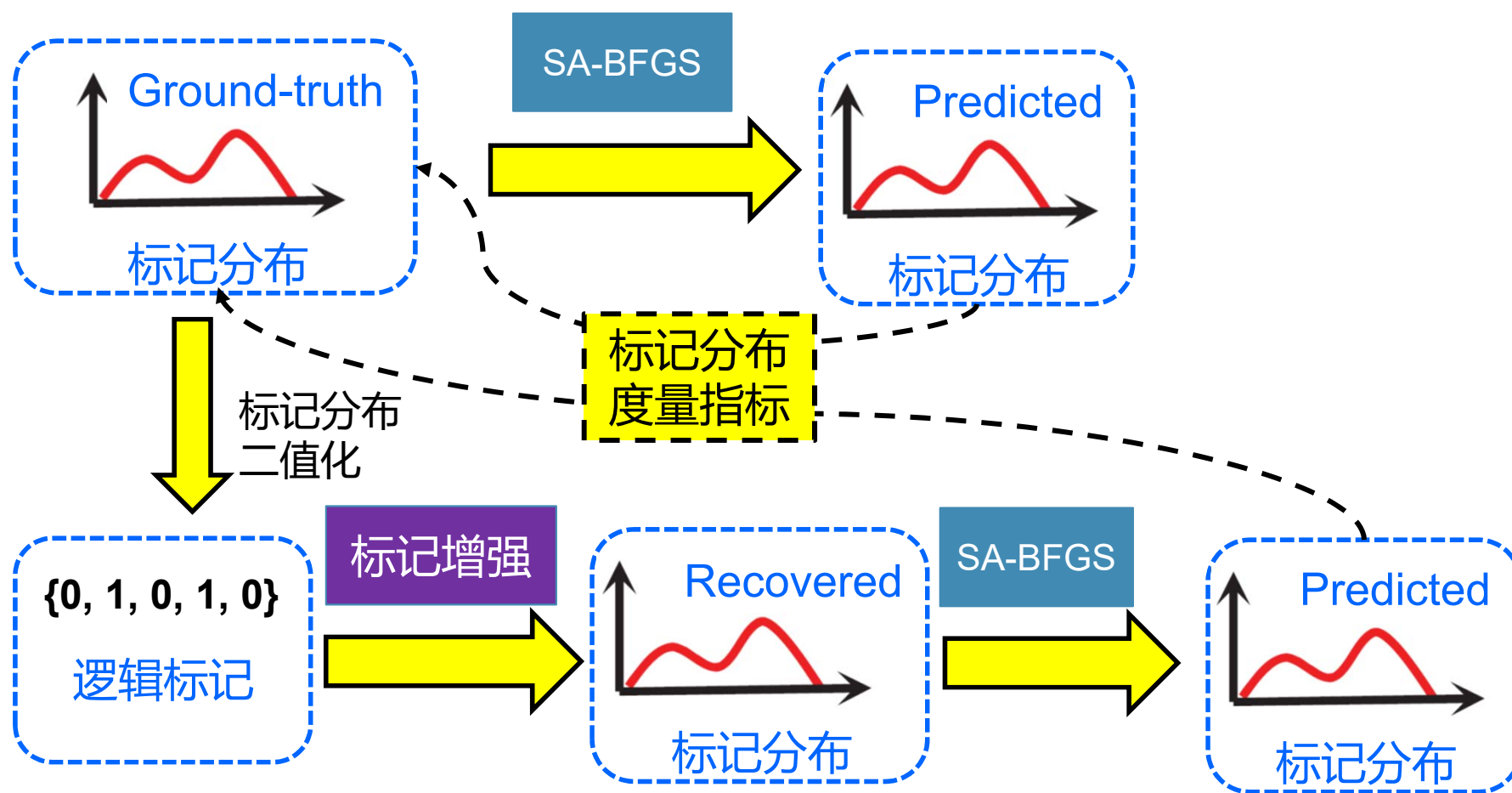
Datasets	FCM	KM	LP	ML	GLLE
Artificial	0.188(3)	0.260(5)	0.130(2)	0.227(4)	0.108(1)
SJAFFE	0.132(3)	0.214(5)	0.107(2)	0.190(4)	0.100(1)
Natural Scene	0.368(5)	0.306(4)	0.275(1)	0.295(2)	0.296(3)
Yeast-spoem	0.233(3)	0.408(5)	0.163(2)	0.400(4)	0.108(1)
Yeast-spo5	0.162(3)	0.277(5)	0.114(2)	0.273(4)	0.092(1)
Yeast-dtt	0.097(2)	0.257(5)	0.128(3)	0.244(4)	0.065(1)
Yeast-cold	0.141(3)	0.252(5)	0.137(2)	0.242(4)	0.093(1)
Yeast-heat	0.169(4)	0.175(5)	0.086(2)	0.165(3)	0.056(1)
Yeast-spo	0.130(3)	0.175(5)	0.090(2)	0.171(4)	0.067(1)
Yeast-diau	0.124(3)	0.152(5)	0.099(2)	0.148(4)	0.084(1)
Yeast-elu	0.052(3)	0.078(5)	0.044(2)	0.072(4)	0.030(1)
Yeast-cdc	0.051(3)	0.076(5)	0.042(2)	0.071(4)	0.038(1)
Yeast-alpha	0.044(3)	0.063(5)	0.040(2)	0.057(4)	0.033(1)
SBU_3DFE	0.135(2)	0.238(5)	0.123(1)	0.233(4)	0.141(3)
Movie	0.230(4)	0.234(5)	0.161(2)	0.164(3)	0.160(1)
Avg. Rank	3.13	4.93	1.93	3.73	1.27

实验结果

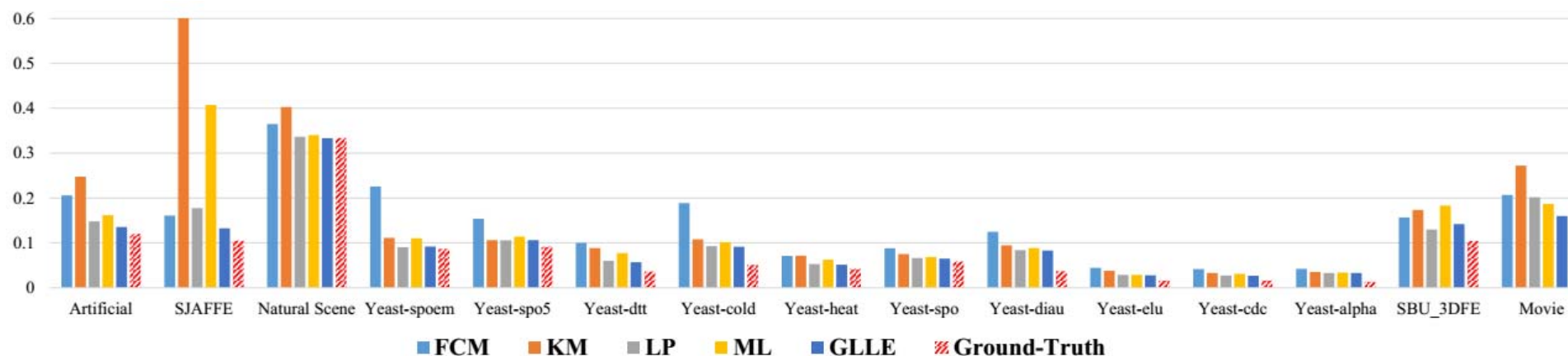
Recovery Results (value(rank)) Measured by Cosine \uparrow

Datasets	FCM	KM	LP	ML	GLLE
Artificial	0.933(3)	0.918(5)	0.974(2)	0.925(4)	0.980(1)
SJAFFE	0.906(3)	0.827(5)	0.941(2)	0.857(4)	0.946(1)
Natural Scene	0.593(5)	0.748(4)	0.860(1)	0.818(2)	0.769(3)
Yeast-spoem	0.878(3)	0.812(5)	0.950(2)	0.815(4)	0.968(1)
Yeast-spo5	0.922(3)	0.882(5)	0.969(2)	0.884(4)	0.974(1)
Yeast-dtt	0.959(2)	0.759(5)	0.921(3)	0.763(4)	0.983(1)
Yeast-cold	0.922(3)	0.779(5)	0.925(2)	0.784(4)	0.969(1)
Yeast-heat	0.883(3)	0.779(5)	0.932(2)	0.783(4)	0.980(1)
Yeast-spo	0.909(3)	0.800(5)	0.939(2)	0.803(4)	0.968(1)
Yeast-diau	0.882(3)	0.799(5)	0.915(2)	0.803(4)	0.939(1)
Yeast-elu	0.950(2)	0.758(5)	0.918(3)	0.763(4)	0.978(1)
Yeast-cdc	0.929(2)	0.754(5)	0.916(3)	0.759(4)	0.959(1)
Yeast-alpha	0.922(2)	0.751(5)	0.911(3)	0.756(4)	0.973(1)
SBU_3DFE	0.912(2)	0.812(5)	0.922(1)	0.815(4)	0.900(3)
Movie	0.773(5)	0.880(4)	0.929(1)	0.919(2)	0.900(3)
Avg. Rank	2.93	4.87	2.07	3.73	1.40

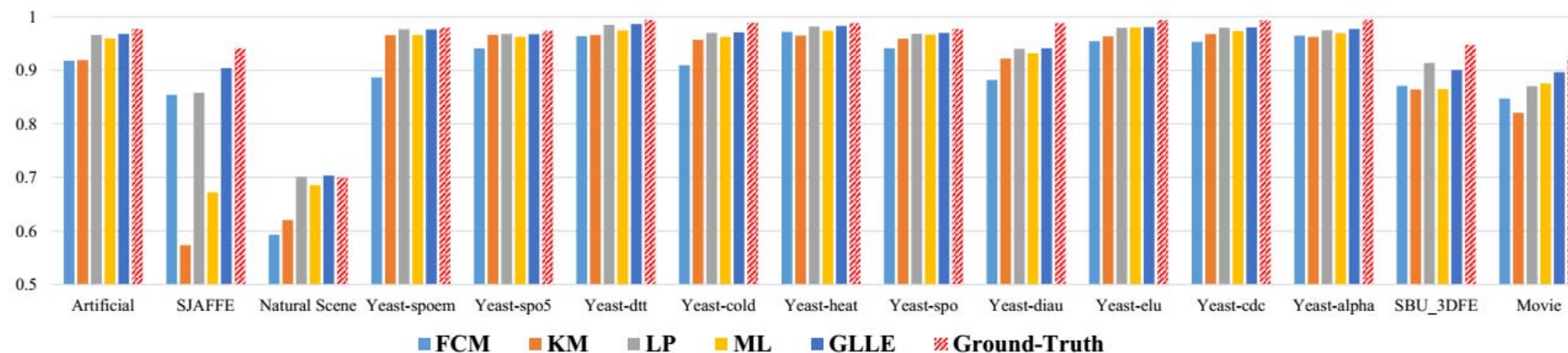
标记分布学习预测实验



实验结果



Comparison of the LDL after the LE pre-process against the direct LDL measured by Cheb ↓.



Comparison of the LDL after the LE pre-process against the direct LDL measured by Cosine ↑.

实验结果

The Average Ranks of Five Algorithms on Six Measures

Criterion	FCM	KM	LP	ML	GLLE
Cheb	4.40	4.20	2.00	3.13	1.27
Clark	4.33	4.07	2.27	3.07	1.27
Canber	4.20	4.13	2.27	3.13	1.27
KL	4.37	4.30	2.00	3.13	1.20
Cosine	4.53	4.27	1.93	3.07	1.20
Intersec	4.40	4.20	1.93	3.13	1.33

报告内容

- 研究背景
- 标记增强
 - 概念定义
 - 学习算法
 - 实验
- 结论



结论

- **标记分布学习**

- 是一种比传统单标记和多标记学习更为泛化的学习范式
- 能够处理标记的不同重要程度（描述度）
- 对许多实际问题本质上具有普适性
- 缺乏标记分布标注限制了其适用范围

- **标记增强**

- 将简单逻辑标记增强为标记分布
- 显式利用特征向量间拓扑关系或标记间相关性
- 是标记分布学习普适性的前提

对标记分布学习&标记增强感兴趣?

所有相关**论文**、**算法代码**和**数据集**可从如下地址下载:

<http://cse.seu.edu.cn/PersonalPage/xgeng/LDL>

[LDL](#) [Home](#) [Download](#) [Contact](#)

Label Distribution Learning

For real applications where the overall distribution of the importance of the labels matters.
A more general learning framework which includes both single-label and multi-label learning as its special cases.

Introduction

Label Distribution Learning is a novel machine learning paradigm. A label distribution covers a certain number of labels, representing the degree to which each label describes the instance. LDL is a general learning framework which includes both single-label and multi-label learning as its special cases.

Further details about LDL can be found in the following paper:

X. Geng, **Label Distribution Learning**, *IEEE Transactions on Knowledge and Data Engineering (IEEE TKDE)*, 2016, in press.

Our algorithms can be used freely for academic, non-profit purposes. If you intend to use it for commercial development, please contact us.

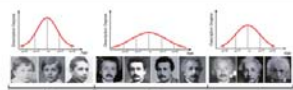
In academic papers using our codes and data, the following references will be appreciated:

[1] X. Geng, **Label Distribution Learning**, *IEEE Transactions on Knowledge and Data Engineering (IEEE TKDE)*, 2016, in press.

[2] X. Geng, C. Yin, and Z.-H. Zhou, **Facial Age Estimation by Learning from Label Distributions**, *IEEE Transactions on Pattern Analysis and Machine Intelligence (IEEE TPAMI)*, 2013, 35(10): 2401-2412.

[Introduction](#)
[Applications of LDL](#)
[Matlab Code](#)
[Data Sets](#)
[Back to top](#)

Applications of LDL



Facial Age Estimation

- X. Geng, Q. Wang, and Y. Xia, **Facial Age Estimation by Adaptive Label Distribution Learning**, In: *Proceedings of the 22nd International Conference on Pattern Recognition (ICPR'14)*, Stockholm, Sweden, 2014, pp. 4465 - 4470.
- X. Geng, C. Yin, and Z.-H. Zhou, **Facial Age Estimation by Learning from Label Distributions**, *IEEE Transactions on Pattern Analysis and Machine Intelligence (IEEE TPAMI)*, 2013, 35(10): 2401-2412.
- X. Geng, K. Smith-Miles, Z.-H. Zhou, **Facial Age Estimation by Learning from Label Distributions**, In: *Proceedings of the 24th AAAI Conference on Artificial Intelligence (AAAI'10)*, Atlanta, GA, 2010, pp. 451-456.



参考文献

1. 耿新, 徐宁, 邵瑞枫. 面向标记分布学习的标记增强. 《计算机研究与发展》, 2017, 54(6): 1171-1184.
2. X. Geng and M. Ling. Soft Video Parsing by Label Distribution Learning. In: Proceedings of the 31th AAAI Conference on Artificial Intelligence (*AAAI'17*), San Francisco, CA, 2017, 1331-1337.
3. X. Geng. Label Distribution Learning. *IEEE Transactions on Knowledge and Data Engineering (IEEE TKDE)*, 2016, 28(7): 1734-1748.
4. D. Zhou, Y. Zhou, X. Zhang, Q. Zhao and X. Geng. Emotion Distribution Learning from Texts. In: *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP'16)*, Austin, TX, 2016, 638-647.
5. P. Hou, X. Geng and M.-L. Zhang. Multi-Label Manifold Learning. In: Proceedings of the 30th AAAI Conference on Artificial Intelligence (*AAAI'16*), Phoenix, AZ, 2016, 1680-1686.
6. X. Yang, X. Geng and D.-Y. Zhou. Sparsity Conditional Energy Label Distribution Learning for Age Estimation. In: *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI'16)*, New York City, NY, 2016, 2259-2265.
7. C. Xing and X. Geng. Logistic Boosting Regression for Label Distribution Learning. In: *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition (CVPR'16)*, Las Vegas, NV, 2016, 4489-4497.
8. X. Geng and P. Hou. Pre-release Prediction of Crowd Opinion on Movies by Label Distribution Learning. In: Proceedings of the International Joint Conference on Artificial Intelligence (*IJCAI'15*), Buenos Aires, Argentina, 2015, 3511-3517.
9. Y. Zhou, H. Xue and X. Geng. Emotion Distribution Recognition from Facial Expressions. In: Proceedings of the 23rd ACM International Conference on Multimedia (*ACM MM'15*), Brisbane, Australia, 2015, 1247-1250.
10. Y.-K. Li, M.-L. Zhang and X. Geng. Leveraging Implicit Relative Labeling-importance Information for Effective Multi-label Learning. In: Proceedings of the 15th IEEE International Conference on Data Mining (*ICDM'15*), Atlantic City, NJ, 2015, 251-260.
11. X. Geng and Y. Xia. Head Pose Estimation Based on Multivariate Label Distribution. In: Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition (*CVPR'14*), Columbus, OH, 2014, pp. 1837-1842.
12. X. Geng and L. Luo. Multilabel Ranking with Inconsistent Rankers. In: Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition (*CVPR'14*), Columbus, OH, 2014, pp. 3742-3747.
13. X. Geng, C. Yin, and Z.-H. Zhou. Facial Age Estimation by Learning from Label Distributions. *IEEE Transactions on Pattern Analysis and Machine Intelligence (IEEE TPAMI)*, 2013, 35(10): 2401-2412.
14. X. Geng, K. Smith-Miles, Z.-H. Zhou. Facial Age Estimation by Learning from Label Distributions. In: Proceedings of the 24th AAAI Conference on Artificial Intelligence (*AAAI'10*), Atlanta, GA, 2010, pp. 451-456.

致谢

- 导师与合作者



Prof. Zhi-Hua Zhou



Prof. Kate Smith-Miles



Prof. Min-Ling Zhang



Prof. Deyu Zhou



Dr. Hui Xue

- 学生

Chao Yin

Peng Hou

Rongzi Ji

Miaogen Ling

Ning Xu

Yu-Kun Li

Yu Xia

Chao Xing

Quan Zhao

Ruifeng Shao

Ying Zhou

Longrun Luo

Xu Yang

Xuan Zhang



谢谢!



[http:// palm.seu.edu.cn](http://palm.seu.edu.cn)