



中國農業大學
China Agricultural University

《机器学习》课程——

半监督学习

主讲人： 徐义田教授

学 校： 中国农业大学



- 研究背景
- 定义
- 经典算法
- 总结



➤ 研究背景

一、研究背景



中國農業大學
China Agricultural University

传统机器学习：有监督学习、无监督学习
(有标记) (无标记)

有标记数据



猫



狗

无标记数据



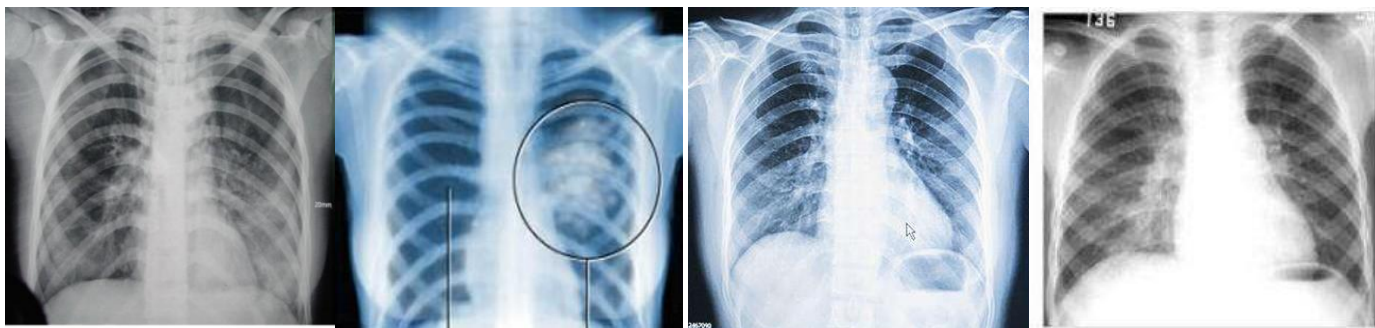
(没有标记的狗和猫的图片)

一、研究背景



在很多实际问题中，只有少量的带有标记的数据，有大量的未标记的数据

- 大量医学影像：只有部分影像被标记出有无病，大部分的影像未被标记



- 医生可以把每张片子上的每个病灶都标出来进行学习（昂贵的）
- 只利用有标记的样本（浪费无标记样本的信息）

如何利用大量未标记的样本改善学习性能成为当前机器学习研究的热点

一、研究背景



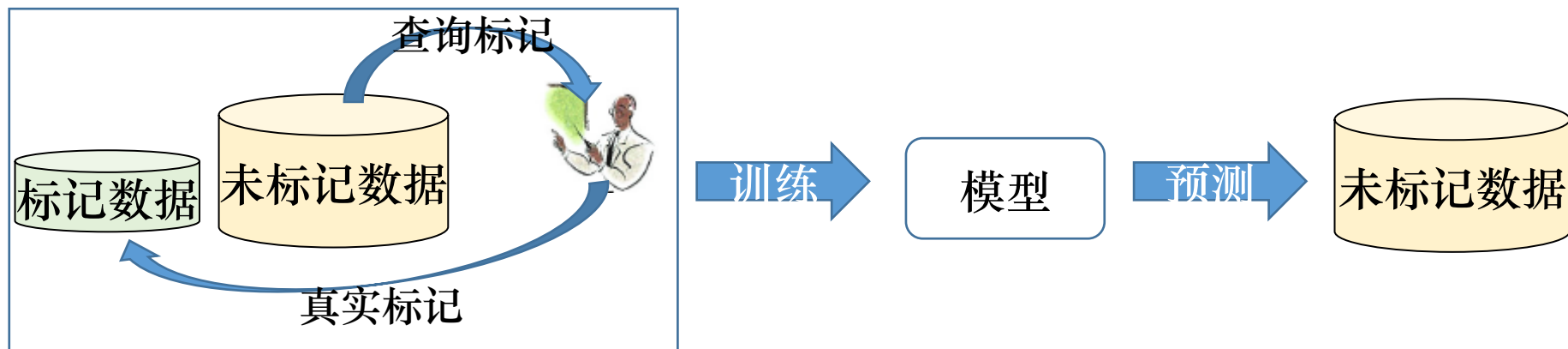
中國農業大學
China Agricultural University

主动学习：使用尽量少的“查询”来获得尽量好的性能

基本过程：

- 先利用有标记的样本训练一个模型，
- 利用这个模型挑选一个最有可能的样本，
- 询问专家（医生）得到一个标记，
- 加入有标记样本进行训练，不断重复

引入了额外的
专家知识



可不可以不借助专家的信息，利用大量未标记的样本？



➤ 定义

二、定义



定义： 在利用有标签的同时，让学习器**不依赖外界交互**、自动地利用**未标记样本**来提升学习性能。

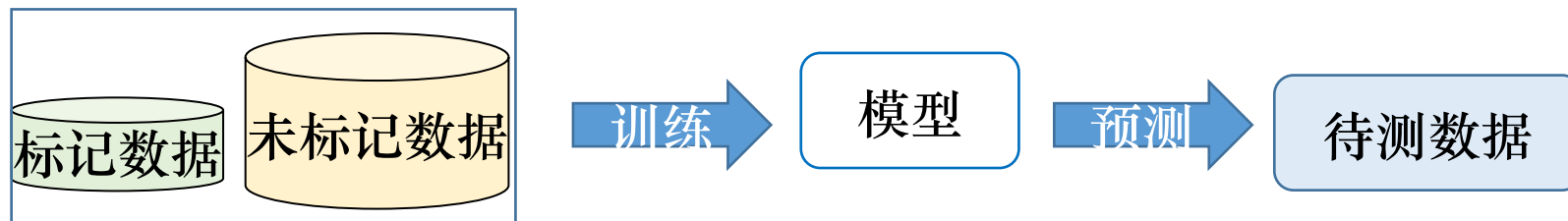
分类： 纯半监督学习和直推学习

- 纯半监督学习：假定训练数据中的未标记样本并非预测的数据，基于“**开放世界**”假设；
- 直推学习：假定学习过程中所考虑的未标记样本恰是待预测数据，基于“**封闭世界**”假设。

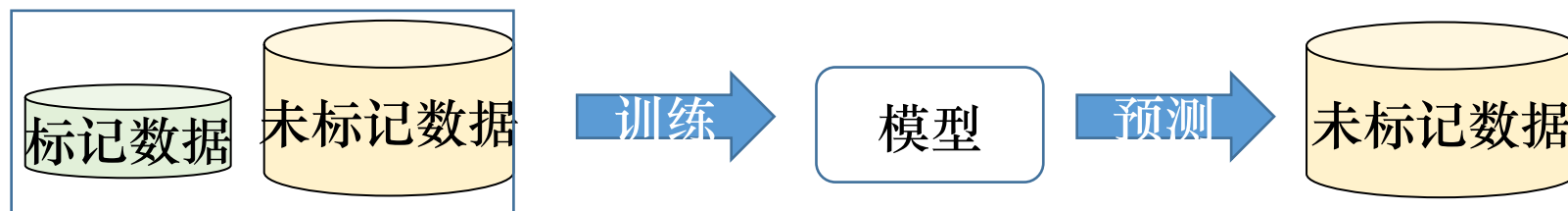
二、定义



纯半监督学习



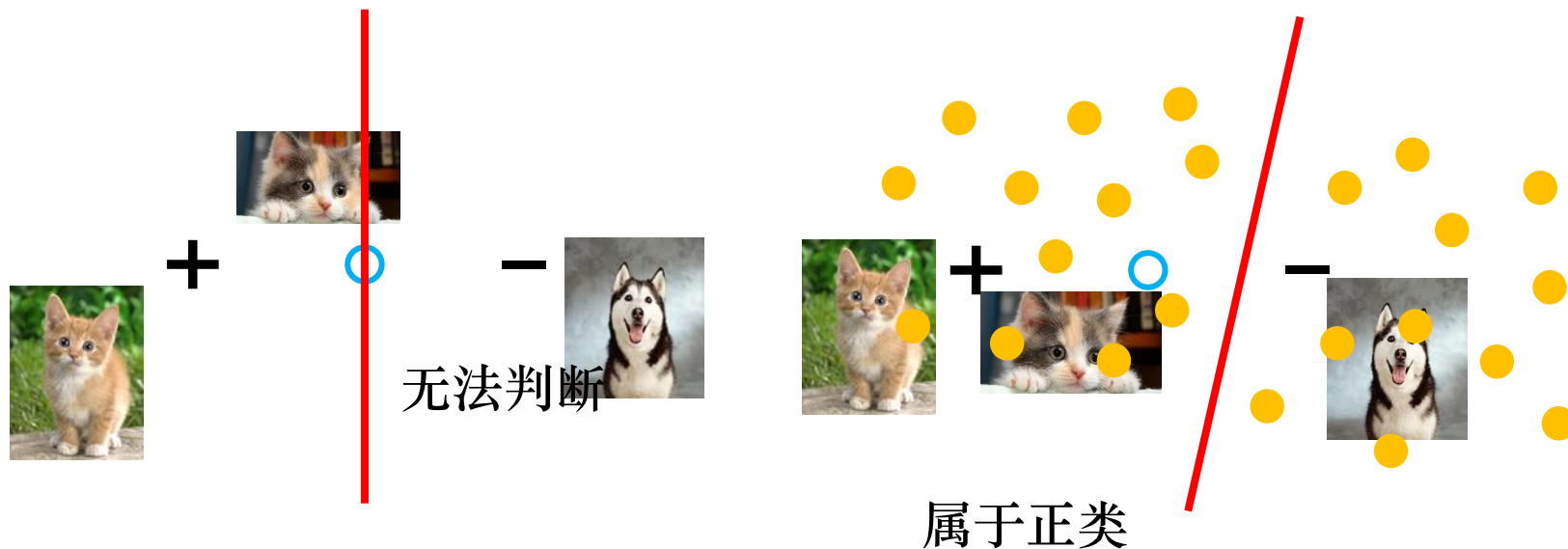
直推学习



二、定义



效果比较:



未标记样本的分布

模型假设

用途: 分类、回归、聚类、降维等问题



➤ 經典算法

三、经典算法



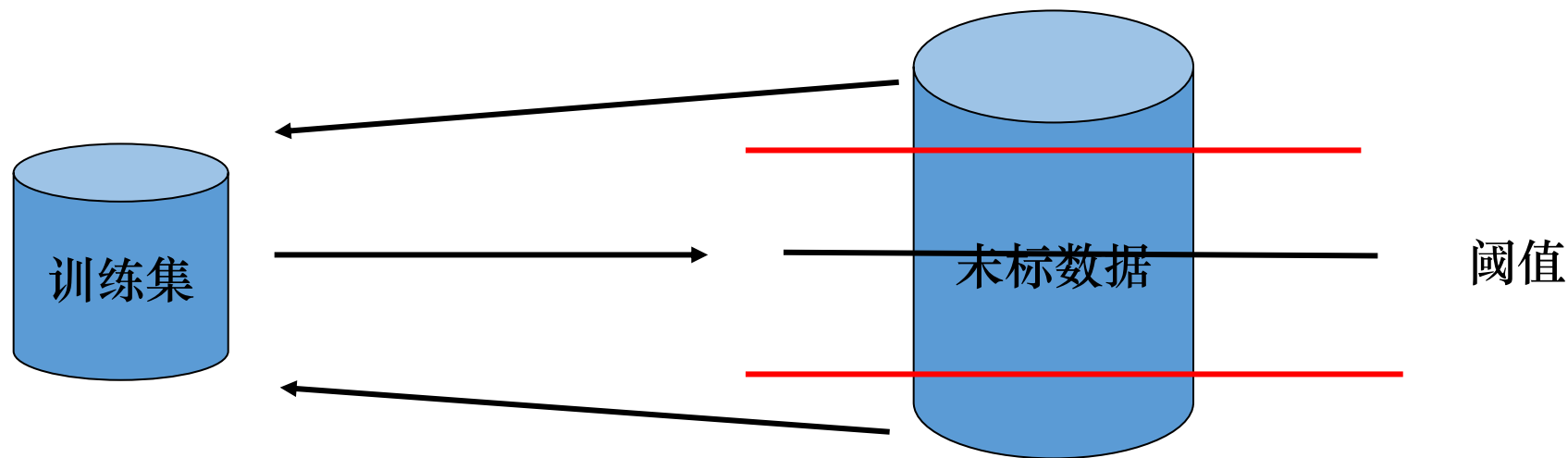
1. 自训练方法
2. 生成式方法
3. 半监督支持向量机
4. 基于图的方法
5. 协同训练方法

三、经典算法



1. 自训练方法

基本思想：将“最确定”的分类结果加入训练集中重新训练



三、经典算法



1. 自训练方法

具体过程:

给定: 有标记的数据 $\{(X_i, y_i)\}_{i=1}^l$ 和未标记数据 $\{X_i\}_{i=l+1}^{l+u}$

Repeat:

任意模型

- 将有标记的数据集作为训练集, 根据训练集训练得到分类器 C_{int} .
- 利用 C_{int} 对无标记的数据集进行分类, 得到: $\{(X_i, y_i)\}_{i=l+1}^{l+u}$
- 计算未标记数据的置信度, 得到置信度比较高的 (X_{conf}, y_{conf})
- 从 X_u 中去掉 (X_{conf}, y_{conf}) , 将 (X_{conf}, y_{conf}) 加入到有标记的数据集中

加权方法

$$(X_{train}, y_{train}) \leftarrow (X_l, y_l) \cup (X_{conf}, y_{conf}),$$

三、经典算法



1. 自训练方法

举例：图像分类问题

- 将月亮和汽车分类



三、经典算法



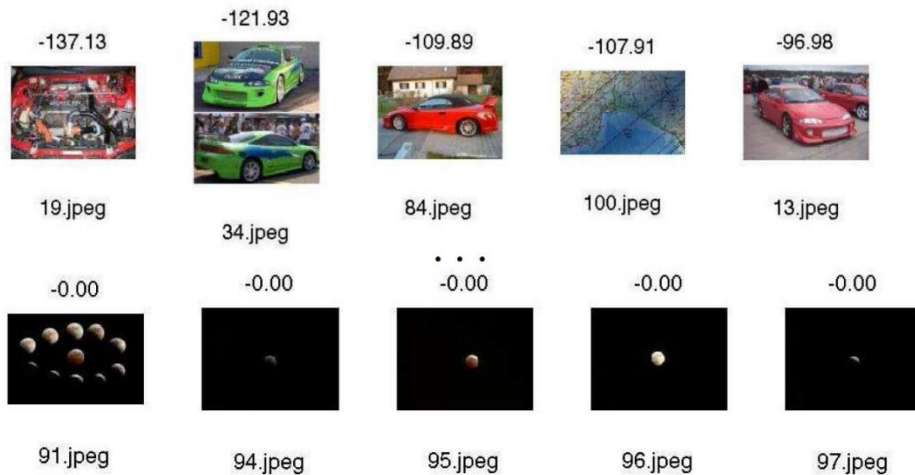
1. 自训练方法

利用self-training训练过程：

① 利用贝叶斯分类器在两个初始的有标记的图像上进行训练



② 对未标记图片进行标记，按置信度排序



三、经典算法

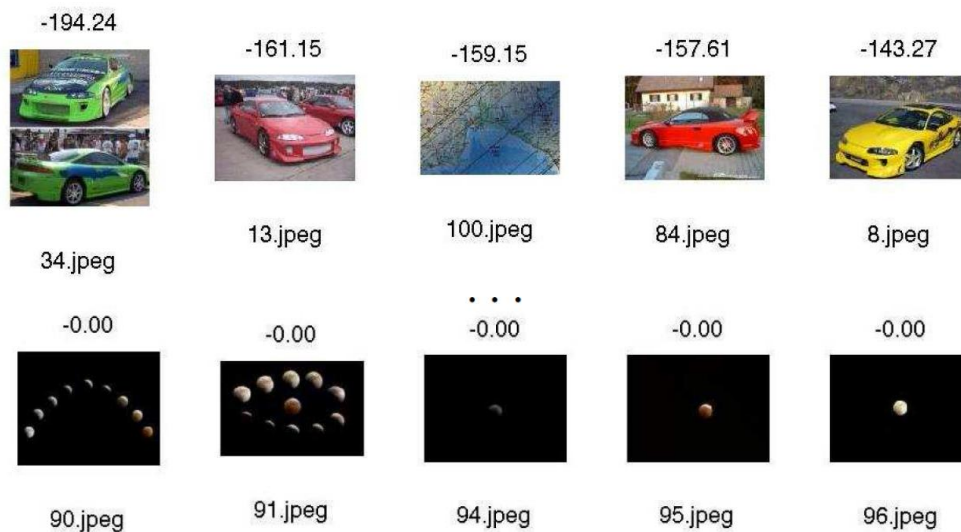


1. 自训练方法

③ 将置信度高的未标记的图片加入训练集



④ 重新训练的分类器，重复以上过程

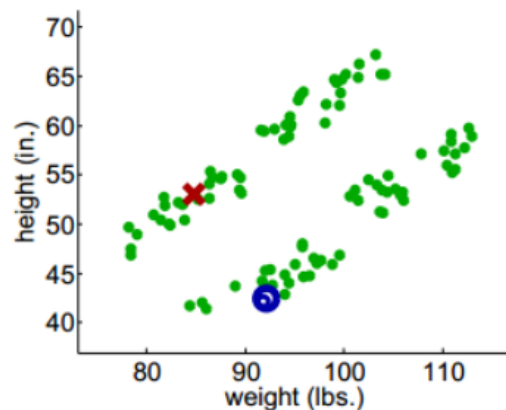


三、经典算法

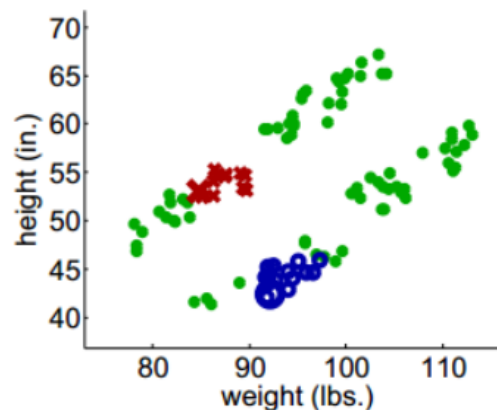
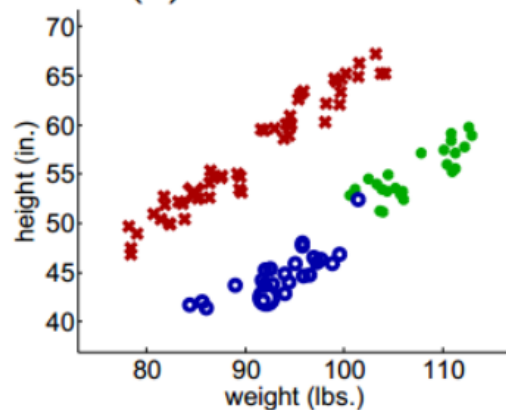


1. 自训练方法

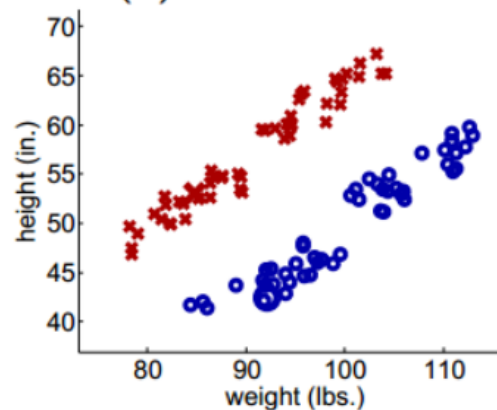
算法效果图:



(a) Iteration 1



(b) Iteration 25

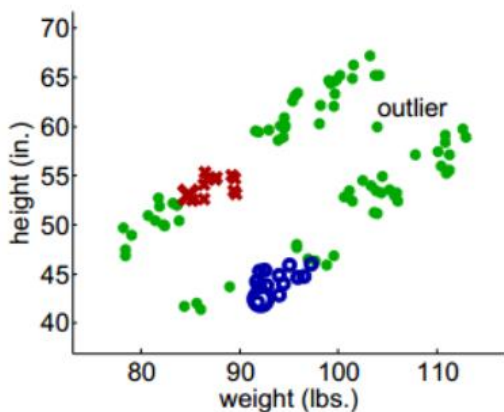


三、经典算法

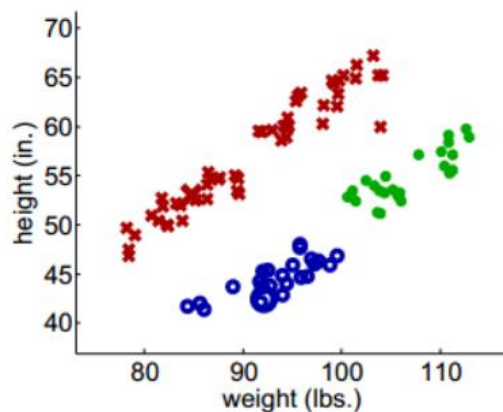


1. 自训练方法

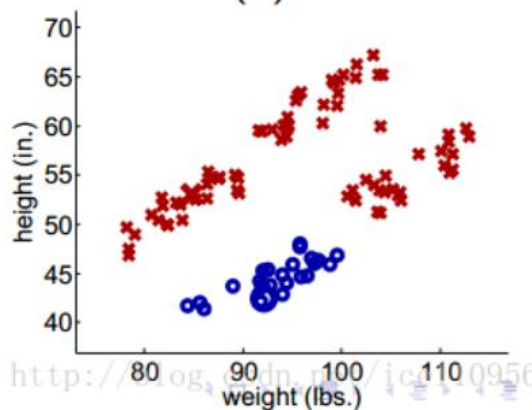
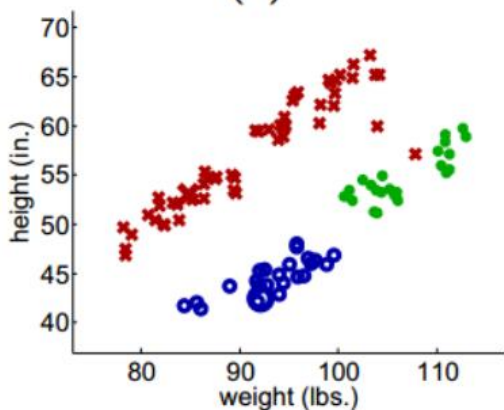
算法效果图:



(a)



(b)



三、经典算法



1. 自训练方法

优点:

- 最简单的半监督学习方法
- 一种包裹式方法，应用于任何存在的分类器
- 经常用于实际问题，如：自然语言处理

缺点:

- 错误会被加强
- 只有少数小规模实验效果有提高，大部分实验效果降低

三、经典算法



2. 生成式方法

生成式方法与判别式方法的区别：

举例：对小狗和小猫进行分类



猫



狗

- 判别式方法：利用训练集得到决策超平面，新来动物时，利用决策线直接将其分开 $\{0, 1\}$ （例如：SVM，逻辑回归）
- 生成式方法：根据狗和猫的特征分别学习一个模型，新来一个动物时，提取其特征，将它分别放入狗和猫的模型中，得到两个概率，哪个概率大就属于哪种类型。

2. 生成式方法

假设所有数据（无论是否有标记）都是由同一潜在模型生成的，分布类型相同

主要的生成式方法：

- 混合高斯分布算法(GMM)
 - ▣ EM算法、图像分类
- 混合多项式分布算法(Naïve Bayes)
 - ▣ 文本分类、EM算法
- 隐马尔可夫模型(HMM)
 - ▣ 语音识别、Baum-Welch算法

三、经典算法



2. 生成式方法

高斯混合模型：

假设样本由高斯混合模型生成，且每个类别对应一个高斯混合成分，即数据样本是基于如下概率密度生成：

选择第 i 个混合模型的概率（先验概率）

$$p_M(x) = \sum_{i=1}^N \alpha_i p(x|\mu_i, \Sigma_i)$$

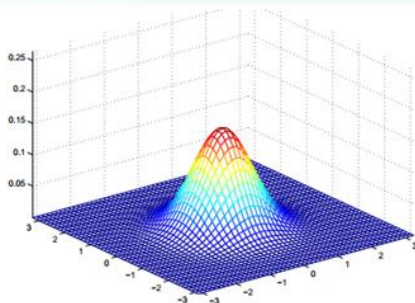
其中概率密度函数： $p(x|\mu, \Sigma) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)^T}$

三、经典算法

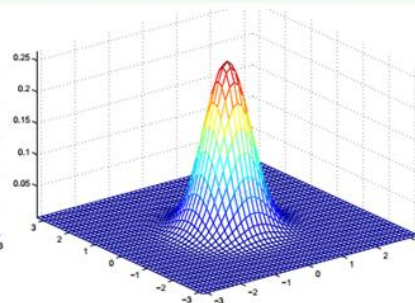


μ 决定中心位置， Σ 决定椭圆的朝向和大小

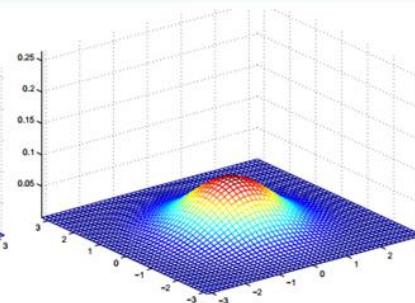
固定 $\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$



$$\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

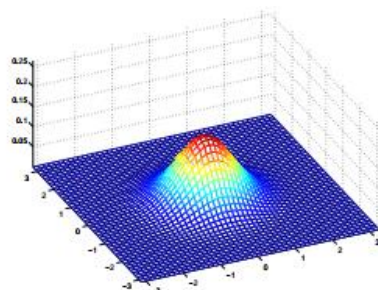


$$\Sigma = \begin{bmatrix} 0.6 & 0 \\ 0 & 0.6 \end{bmatrix}$$

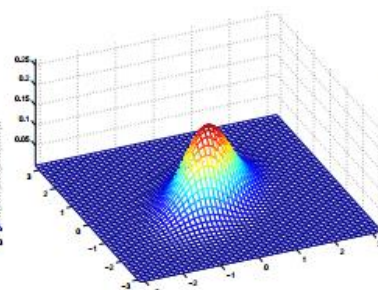


$$\Sigma = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

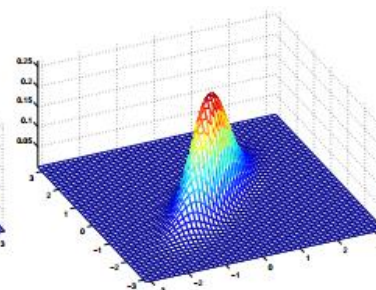
固定 $\mu = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$



$$\Sigma = \begin{bmatrix} 1.5 & 0 \\ 0 & 1.5 \end{bmatrix}$$

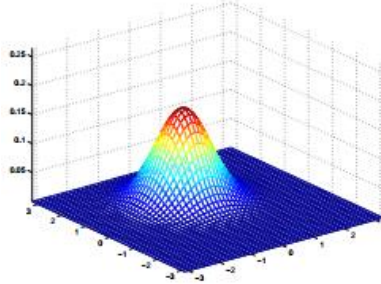


$$\Sigma = \begin{bmatrix} 1.5 & 0.5 \\ 0.5 & 1.5 \end{bmatrix}$$

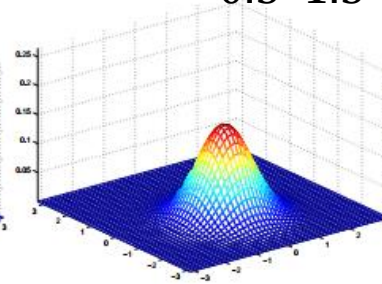


$$\Sigma = \begin{bmatrix} 1.5 & 0.8 \\ 0.8 & 1.5 \end{bmatrix}$$

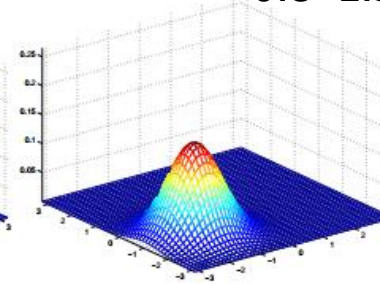
固定 $\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$



$$\mu = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$



$$\mu = \begin{bmatrix} -0.5 \\ 0 \end{bmatrix}$$



$$\mu = \begin{bmatrix} -1 \\ 1.5 \end{bmatrix}$$

三、经典算法



2. 生成式方法

贝叶斯公式 $P(A \cap B) = P(A)P(B|A) = P(B)P(A|B)$

后验概率：未标记样本 x_j 属于第 i 个分布的概率

$$\gamma_{ji} = \frac{\alpha_i \cdot p(x_j | \mu_i, \Sigma_i)}{\sum_{i=1}^N \alpha_i p(x_j | \mu_i, \Sigma_i)}$$

贝叶斯公式

预测：最大化后验概率

$$f(x) = \operatorname{argmax}_{j \in Y} \gamma_{ji}$$

如何确定模型中的参数：
极大似然估计

三、经典算法



2. 生成式方法

极大似然估计:

使数据出现的可能性最大

$$\max LL(D) = \ln\left(\prod_{j=1}^m p(x)\right)$$

$$\begin{aligned} \max LL(D_l \cup D_u) = & \sum_{(x_j, y_j) \in D_l \cap y_j=i} \ln(\alpha_i p(x_j | \mu_i, \Sigma_i)) \\ & + \sum_{x_j \in D_u} \ln\left(\sum_{i=1}^N \alpha_i p(x_j | \mu_i, \Sigma_i)\right) \end{aligned}$$

$$\text{求导得} \begin{cases} \frac{\partial LL}{\partial \mu_i} = 0 \Rightarrow \mu_i = \frac{1}{\sum_{x_j \in D_u} \gamma_{ji} + l_i} \left(\sum_{x_j \in D_u} \gamma_{ji} x_j + \sum_{(x_j, y_j) \in D_l \cap y_j=i} x_j \right) \\ \frac{\partial LL}{\partial \Sigma_i} = 0 \Rightarrow \Sigma_i = \frac{1}{\sum_{x_j \in D_u} \gamma_{ji} + l_i} \left(\sum_{x_j \in D_u} \gamma_{ji} (x_j - \mu_i) (x_j - \mu_i)^T \right) \\ \frac{\partial LL}{\partial \alpha_i} = 0 \Rightarrow \alpha_i = \frac{1}{m} \left(\sum_{x_j \in D_u} \gamma_{ji} + l_i \right) \end{cases}$$

三、经典算法



2. 生成式方法

EM算法求解:

- 初始化参数: $\alpha_1, \mu_1, \Sigma_1$,

E

- Step 1: 根据当前模型参数计算未标记样本的后验概率

$$\gamma_{ji} = \frac{\alpha_i \cdot p(x_j | \mu_i, \Sigma_i)}{\sum_{i=1}^N \alpha_i p(x_j | \mu_i, \Sigma_i)}$$

返回step 1

M

- Step 2: 更新模型参数

$$\mu_i = \frac{1}{\sum_{x_j \in D_u} \gamma_{ji} + l_i} \left(\sum_{x_j \in D_u} \gamma_{ji} x_j + \sum_{(x_j, y_j) \in D_l \cap y_j = i} x_j \right)$$

$$\Sigma_i = \frac{1}{\sum_{x_j \in D_u} \gamma_{ji} + l_i} \left(\sum_{x_j \in D_u} \gamma_{ji} (x_j - \mu_i) (x_j - \mu_i)^T \right)$$

$$\alpha_i = \frac{1}{m} \left(\sum_{x_j \in D_u} \gamma_{ji} + l_i \right)$$

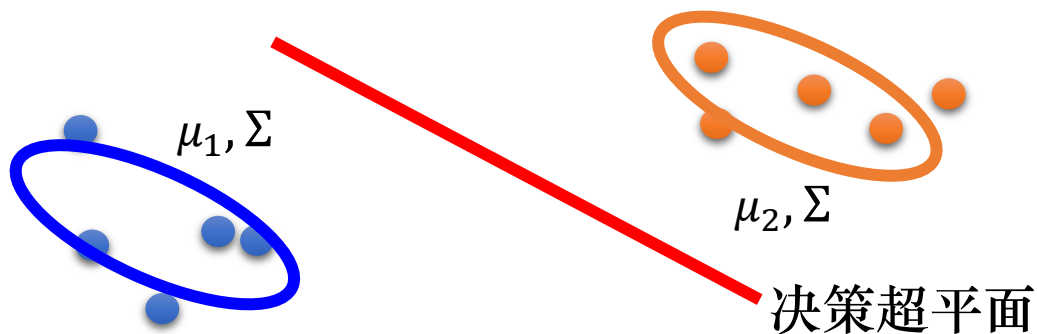
三、经典算法



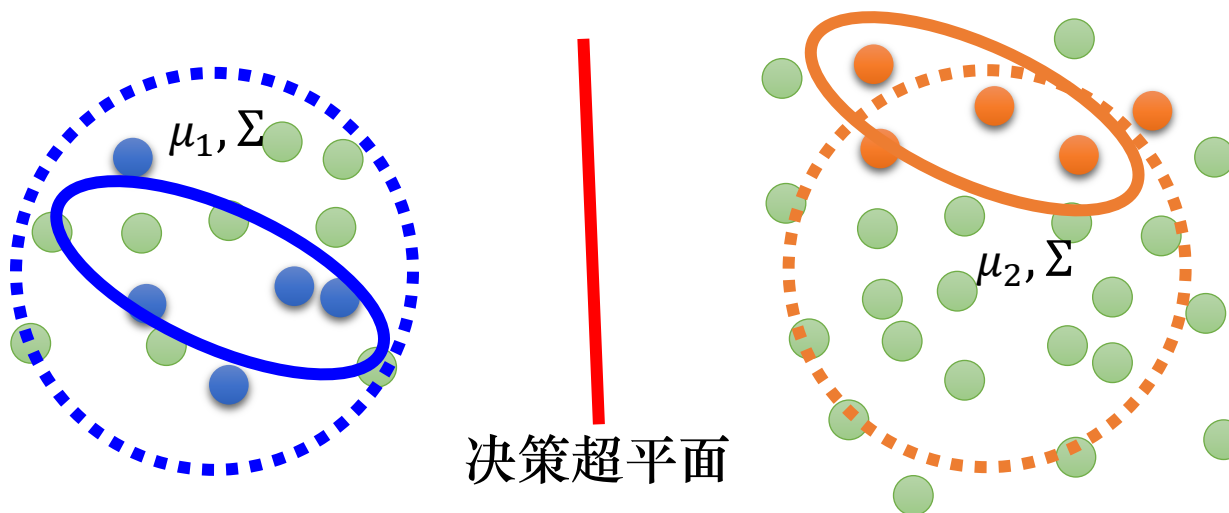
2. 生成式方法

举例：

有监督：



半监督：



三、经典算法



2. 生成式方法

优点:

- 以概率视角求解半监督算法
- 如果数据分布符合模型，效果非常好

缺点:

- 通常很难验证模型的正确性
- EM算法是局部最优的
- 如果生成模型是错误的，未标记样本可能会误导

三、经典算法

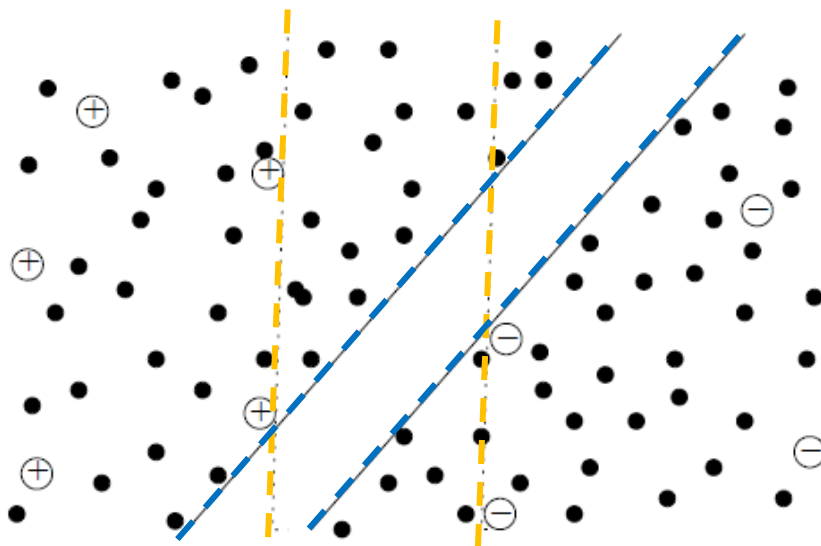


3. 半监督支持向量机

半监督支持向量机是支持向量机在半监督学习上的推广

基本思想：找到能将两类样本点分开，且穿过**低密度**区域的划分超平面

基本假设



3. 半监督支持向量机

最著名的半监督支持向量机 (S^3VM) : 转导支持向量机 (TSVM)

TSVM的基本思想:

假设未标记数据属于不同的类(正类、负类), 可以被超平面分离

- 遍历所有未标记样本 X_u 的标签, 共 2^u 中可能
- 对于每一种标记, 构建一个标准的支持向量机模型
- 从所有可能中筛选出最大间隔的模型

三、经典算法



3. 半监督支持向量机

模型:

$$\min_{w, b, \hat{y}, \xi_i} \frac{1}{2} \|w\|^2 + C_l \sum_{i=1}^l \xi_i + C_u \sum_{i=l+1}^m \xi_i$$

$$s. t. \quad y_i((w \cdot x_i) + b) \geq 1 - \xi_i, \quad i = 1, 2, \dots, l,$$

$$\hat{y}_i((w \cdot x_i) + b) \geq 1 - \xi_i, \quad i = l + 1, 2, \dots, m,$$

$$\{+1, -1\} \quad \xi_i \geq 0, \quad i = 1, 2, \dots, m.$$

求解

- ① 精确解：遍历所有可能的情况，用传统SVM的方法求解（未标记样本少量时有效，否则太耗时）
- ② 非精确解：用局部搜索迭代的方法寻找近似解

三、经典算法



3. 半监督支持向量机

局部搜索迭代寻找近似解

输入：有标记样本集 $D_l = \{(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)\}$;
未标记样本集 $D_u = \{x_{l+1}, x_{l+2}, \dots, x_{l+u}\}$;
折中参数 C_l, C_u ;

过程：

- 1: 用 D_l 训练一个 SVM_l ; 初始的 SVM
- 2: 用 SVM_l 对 D_u 中的样本进行预测, 得到 $\hat{y} = (\hat{y}_{l+1}, \hat{y}_{l+2}, \dots, \hat{y}_{l+n})$; 伪标记
- 3: 初始化 $C_u \ll C_l$;
- 4: **while** $C_u < C_l$ **do**

基于 $D_l, D_u, \hat{y}, C_l, C_u$ 求解模型, 得到 $(w, b), \xi$;
while $\exists \{i, j\} | (\hat{y}_i, \hat{y}_j < 0) \cap (\xi_i > 0) \cap (\xi_j > 0) \cap (\xi_i + \xi_j > 2)$ **do**
 $\hat{y}_i = -\hat{y}_i; \hat{y}_j = -\hat{y}_j$;
 基于 $D_l, D_u, \hat{y}, C_l, C_u$ 求解模型, 得到 $(w, b), \xi$;
end while
 $C_u = \min\{2C_u, C_l\}$

end while

输出：未标记样本的预测结果： $\hat{y} = (\hat{y}_{l+1}, \hat{y}_{l+2}, \dots, \hat{y}_{l+n})$

3. 半监督支持向量机

基于低密度假设得半监督模型：

$$\begin{aligned} \min_{w, b, \xi_i} & \frac{1}{2} ||w||^2 + C_l \sum_{i=1}^l \xi_i + C_u \sum_{i=l+1}^m \xi_i \\ \text{s.t. } & y_i((w \cdot x_i) + b) \geq 1 - \xi_i, \quad i = 1, 2, \dots, l, \\ & |((w \cdot x_i) + b)| \geq 1 - \xi_i, \quad i = l + 1, 2, \dots, m, \\ & \xi_i \geq 0, \quad i = 1, 2, \dots, m. \end{aligned}$$

非凸优化问题



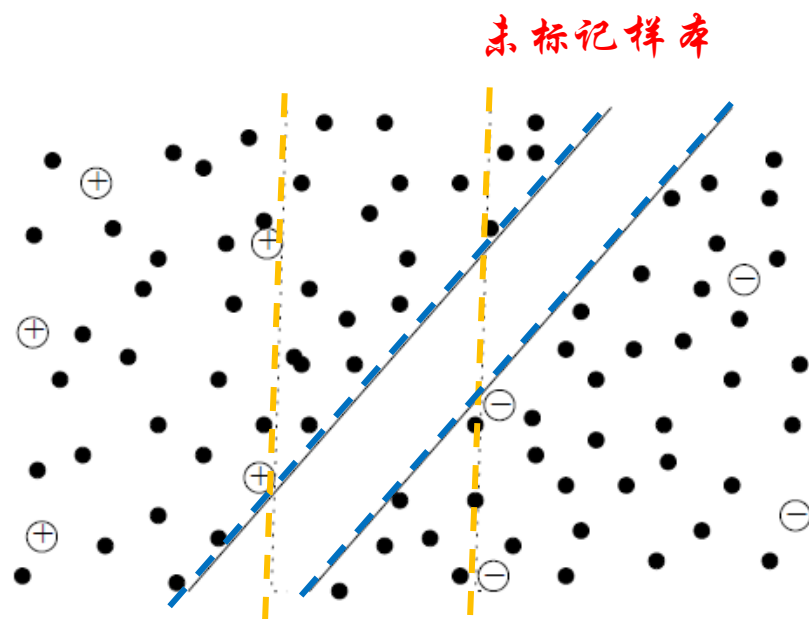
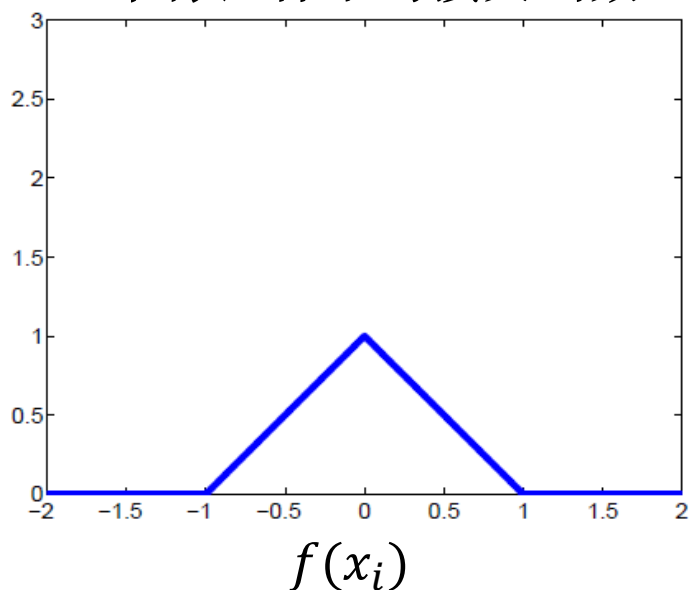
min 正则化项+损失函数项

三、经典算法

3. 半监督支持向量机

模型: $\min_f \sum_{i=1}^l (1 - y_i f(x_i))_+ + \lambda \|w\|^2 + \lambda_2 \sum_{i=l+1}^{l+u} (1 - |f(x_i)|)_+$

未标记样本的损失函数



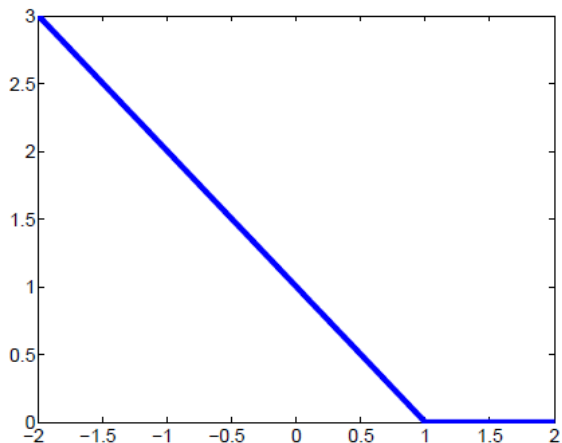
未标记样本在间隔带以外，即未标记样本不在决策超平面 $f(x)$ 周围

三、经典算法

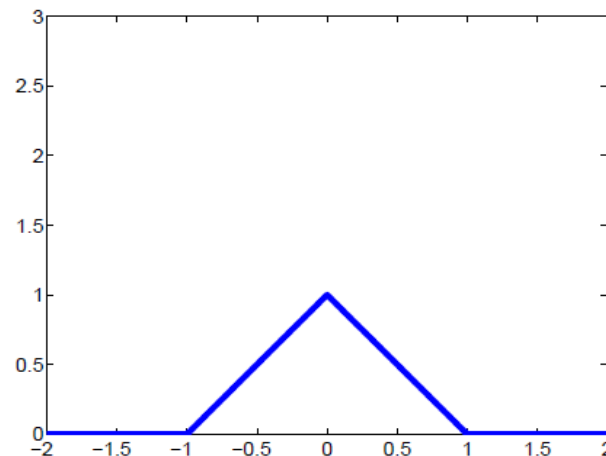


3. 半监督支持向量机

求解：



SVM: 凸



S^3VM : 非凸

求解非凸问题的优化算法：

SVM^{light} , ∇S^3VM , CCCP, deterministic annealing, Branch and Bound, SDP convex relaxation 等等

三、经典算法



3. 半监督支持向量机

优点:

- 适用于任何SVM使用的地方
- 清晰的数学框架

缺点:

- 求解困难
- 会陷入局部最优解
- 比生成式模型和图模型更温和的假设

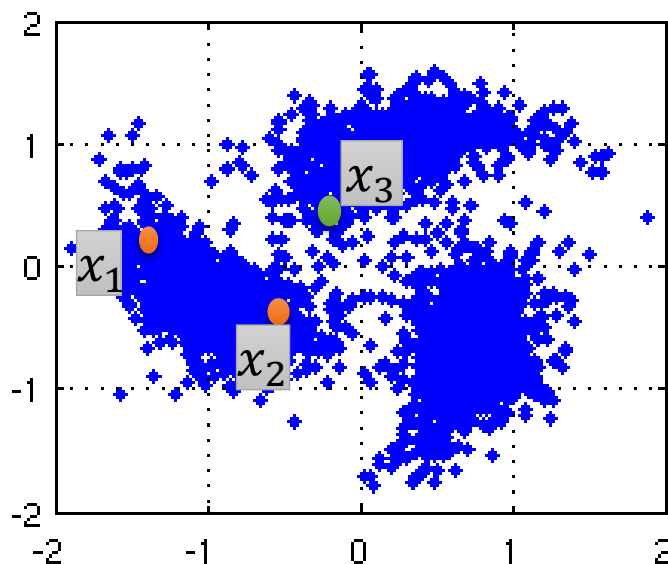
三、经典算法



4. 基于图的算法

光滑性假设：近朱者赤，近墨者黑

- 假设：相似的 x 拥有相同的 y
- 更精确：如果 x_1 和 x_2 在一个高密度区域里相似，则 y_1 和 y_2 相同



x_1 和 x_2 具有相同的标签， x_2 和 x_3 拥有不同的标签

三、经典算法



4. 基于图的算法

基本思想：

假设一个图由标记和未标记数据共同组成，有边连接的样本具有相同的标签

- 将标记样本和未标记样本构成无向图
- 加入正则化项，使的预测函数在整个图上尽可能的光滑
(相似样本具有相同的标签)

三、经典算法



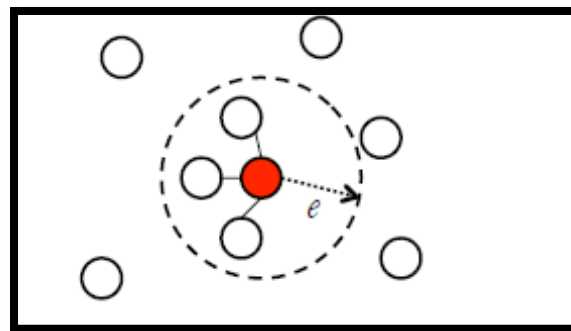
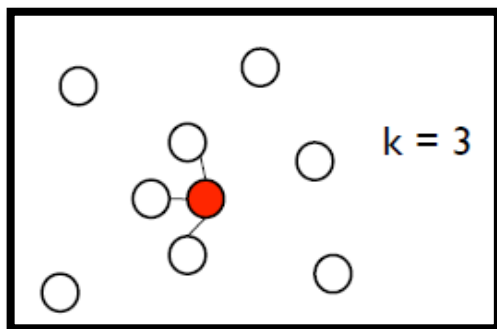
4. 基于图的算法

图的构建:

顶点: $X_l \cup X_u$

边:

- 有无边: k近邻图、e近邻图



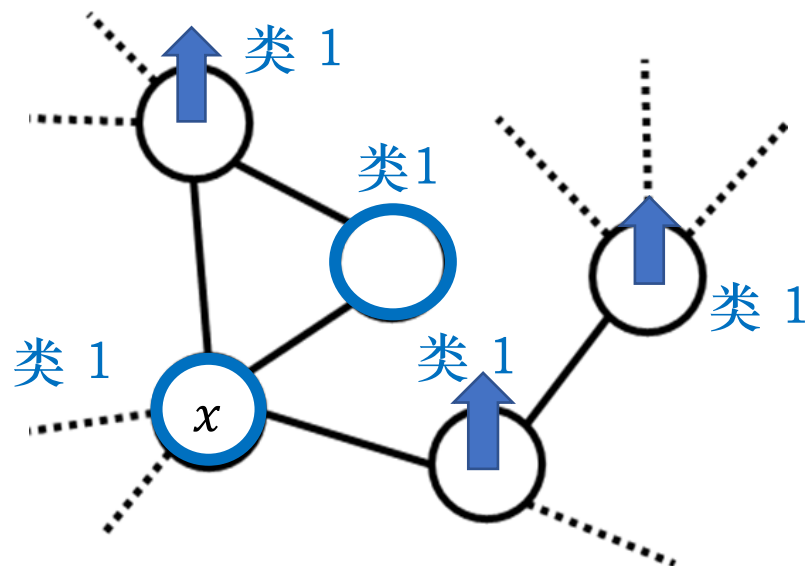
- 边的权重: x_i 和 x_j 的相似性 $s(x_i, x_j)$

高斯径向基函数:

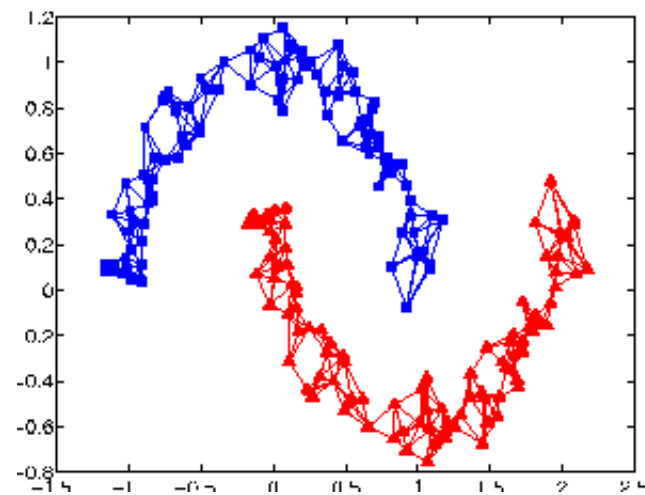
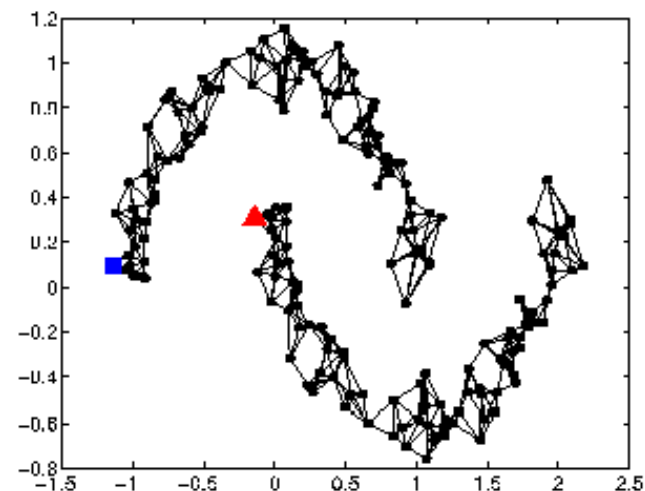
$$s(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{\sigma^2}\right)$$

三、经典算法

4. 基于图的算法



通过图进行传播，标记样本对周围的点产生影响



三、经典算法



4. 基于图的算法

举例：

① 文本分类：

- 将天文学文章和旅游文章分类
- 用内容词重叠测量相似性

	d_1	d_3	d_4	d_2
asteroid	•	•		
bright	•	•		
comet		•		
year				
zodiac				
⋮				
airport				
bike				
camp			•	
yellowstone			•	•
zion				•

有重叠

	d_1	d_3	d_4	d_2
asteroid	•			
bright	•			
comet				
year				
zodiac		•		
⋮				
airport			•	
bike			•	
camp				
yellowstone				•
zion				•

无重叠

三、经典算法



4. 基于图的算法

① 文本分类：

	d_1	d_5	d_6	d_7	d_3	d_4	d_8	d_9	d_2
asteroid	•								
bright	•	•							
comet		•	•						
year			•	•					
zodiac				•	•				
.									
.									
airport						•			
bike						•	•		
camp							•	•	
yellowstone								•	•
zion									•

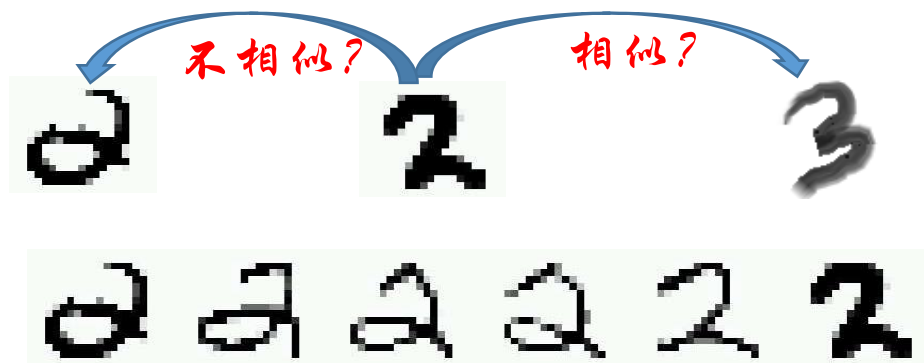
标签通过类似的未贴标签的文章“传播”

三、经典算法

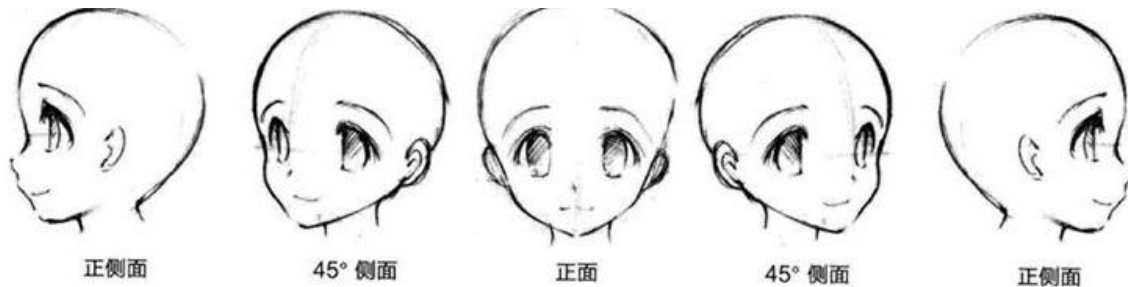


4. 基于图的算法

② 基于像素欧几里得距离的手写数字识别



③ 不同方位的人脸识别



三、经典算法



4. 基于图的算法

具体过程:

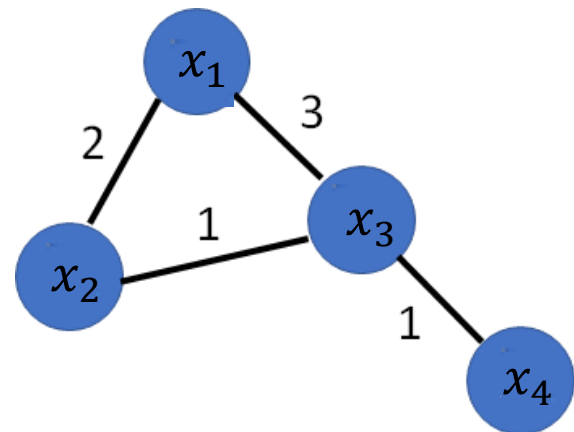
- 能量函数: $S = \frac{1}{2} \sum_{i,j} w_{i,j} (f_i - f_j)^2 = f^T L f$

其中: $f = [\dots f_i \dots f_j \dots]^T$

$L = D - W$ 为图拉普拉斯矩阵,

其中:

$$W = \begin{bmatrix} 0 & 2 & 3 & 0 \\ 2 & 0 & 1 & 0 \\ 3 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad D = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



- 目标: $\min S$

三、经典算法



4. 基于图的算法

以第 l 行与第 l 列为界, 采用分块矩阵表示方式:

$$W = \begin{bmatrix} W_{ll} & W_{lu} \\ W_{ul} & W_{uu} \end{bmatrix}, \quad D = \begin{bmatrix} D_{ll} & D_{lu} \\ D_{ul} & D_{uu} \end{bmatrix},$$

则可得:

$$\begin{aligned} S &= (f_l^T; f_u^T) \left(\begin{bmatrix} D_{ll} & D_{lu} \\ D_{ul} & D_{uu} \end{bmatrix} - \begin{bmatrix} W_{ll} & W_{lu} \\ W_{ul} & W_{uu} \end{bmatrix} \right) \begin{bmatrix} f_l \\ f_u \end{bmatrix} \\ &= f_l^T (D_{ll} - W_{ll}) f_l - 2f_u^T W_{ul} f_l \\ &\quad + f_u^T (D_{uu} - W_{uu}) f_u \end{aligned}$$

由 $\frac{\partial S}{\partial f_u} = 0$ 可得 $f_u = (D_{uu} - W_{uu})^{-1} W_{ul} f_l$

三、经典算法



4. 基于图的算法

$$\text{令 } P = D^{-1}W = \begin{bmatrix} D_{ll}^{-1} & 0_{lu} \\ 0_{ul} & D_{uu}^{-1} \end{bmatrix} \begin{bmatrix} W_{ll} & W_{lu} \\ W_{ul} & W_{uu} \end{bmatrix} = \begin{bmatrix} D_{ll}^{-1}W_{ll} & D_{lu}^{-1}W_{lu} \\ D_{ul}^{-1}W_{ul} & D_{uu}^{-1}W_{uu} \end{bmatrix},$$

即 $P_{uu} = D_{uu}^{-1}W_{uu}$, $P_{ul} = D_{uu}^{-1}W_{ul}$, 则可得:

$$\begin{aligned} f_u &= (D_{uu}(I - D_{uu}^{-1}W_{uu}))^{-1} W_{ul}f_l \\ &= (I - D_{uu}^{-1}W_{uu})^{-1} D_{uu}^{-1}W_{ul}f_l \\ &= (I - P_{uu})^{-1} P_{ul}f_l \end{aligned}$$

于是, 将 D_l 上的标记信息作为 $f_l = (f(x_1); f(x_2); \dots; f(x_l))$ 代入上式, 即可求得 f_u 对未标记样本的预测

对于新样本时, 要么将其加入原数据集中重新构造图进行标记, 要么引入额外的预测机制 (如SVM), 将有标记的样本和伪标记合并为训练集, 学得一个分类器, 对新的样本进行预测.

三、经典算法



4. 基于图的算法

优点:

- 清晰的数学框架
- 如果图的构造合适，效果会非常好

缺点:

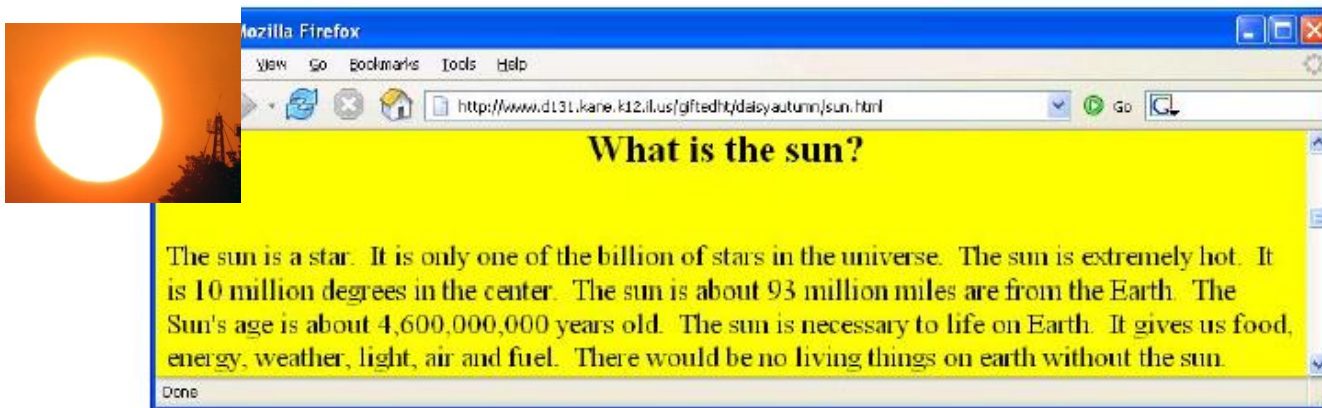
- 如果图的构造不合适，效果会很差
- 对于图的结构和变得权重比较敏感

三、经典算法



5. 协同训练方法

举例：两个视角（图像和网页文本）



三、经典算法

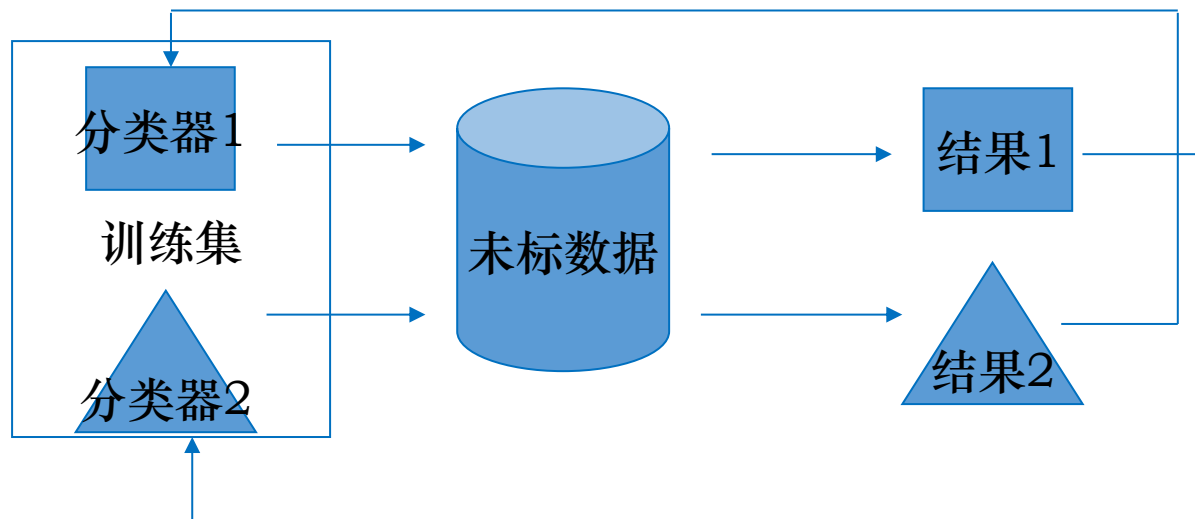


5. 协同训练方法

主要思想：

每一个样本都可以被特征的两个集合表示 $x = [x^1; x^2]$

- x^1 为图像特征； x^2 为文本特征
- 分别训练图像分类和文本分类
- 将置信度高的样本加入到**对方**的训练集中



三、经典算法



输入：有标记样本集 $D_l = \{((x_1^1, x_1^2), y_1), ((x_2^1, x_2^2), y_2), \dots, ((x_l^1, x_l^2), y_l)\}$;
未标记样本集 $D_u = \{(x_{l+1}^1, x_{l+1}^2), (x_{l+2}^1, x_{l+2}^2), \dots, (x_{l+u}^1, x_{l+u}^2)\}$;
缓冲池大小 s ; 每轮挑选的正例数 p ; 每轮挑选的反例数 n ; 基学习算法 \mathcal{L} ; 学习轮数 T ;

过程：

1: 从 D_u 中随机抽取 s 个样本构成缓冲池 D_s ; $D_u = D_u \setminus D_s$;

2: **for** $j=1, 2$ **do**

$D_l^j = \{(x_i^j, y_i) | ((x_i^j, x_i^{3-j}), y_i) \in D_l\}$;

3: **end for**

4: **for** $t=1, 2, \dots, T$ **do**

for $j=1, 2$ **do**

$h_j \leftarrow \mathcal{L}(D_l^j)$;

根据 h_j 在 D_s^j 上的分类置信度，分别挑选 p 和 n 个置信度最高的正例 D_p 和反例 D_n

由 D_p^j 生成伪标记正例 $\check{D}_p^{3-j} = \{(x_i^{3-j}, +1) | x_i^j \in D_p^j\}$;

由 D_n^j 生成伪标记反例 $\check{D}_n^{3-j} = \{(x_i^{3-j}, -1) | x_i^j \in D_n^j\}$;

$D_s = D_s \setminus (D_p \cup D_n)$;

end for

if h_1, h_2 均未发生改变 **then**

break;

else

for $j=1, 2$ **do**

$D_l^j = D_l^j \cup (\check{D}_p^j \cup \check{D}_n^j)$

end for

从 D_u 中随机抽取 $2p + 2n$ 个样本加入 D_s

end if

end for

输出：分类器 h_1, h_2

算法流程

三、经典算法



5. 协同训练方法

优点:

- 实现简单
- 适用于任何分类器
- 比self-training受的噪音干扰小

缺点:

- 很多问题无法分割



➤ 总结

四、总结



1. 研究背景

2. 定义

★ 3. 经典算法

- 自训练方法
- 生成式方法
- 半监督支持向量机
- 基于图的方法
- 协同训练方法



中國農業大學
China Agricultural University

谢 谢 !

