

多示例学习的包层次覆盖 k 近邻算法

赵 姝^{1,2}, 芮 辰^{1,2}, 陈 洁^{1,2}, 张燕平^{1,2}

¹ (安徽大学 计算智能与信号处理教育部重点实验室, 合肥 230601)

² (安徽大学 人工智能研究所, 合肥 230601)

E-mail: zhangyp2@gmail.com

摘 要: 多示例学习是一种新型的机器学习框架, 正包中大量的噪声使多示例数据集具有很大的歧义性. 为了排除多示例数据集正包中大量的假正例, 提高分类精度, 结合邻域覆盖算法, 提出一个新的多示例包层次覆盖 k 近邻算法. 覆盖算法的学习结果是一系列的球形邻域, 在每一个球形邻域中只含有同类样本, 本文利用的覆盖算法的这一特性重新组织多示例数据集的包结构. 概括的说, 为了排除正包中大量的假正例, 首先对原有的多示例包结构进行重新构造, 使用覆盖算法生成的球形邻域做为新的包结构, 从而提高多示例样本在新的特征空间中的可分离性. 然后, 使用包层次的 k 近邻算法排除正包中大量的噪声并预测测试包的类别. 实验表明, 多示例学习的包层次覆盖 k 近邻算法具有很好的性能.

关键词: 机器学习; 多示例学习; 覆盖算法; k 近邻算法

中图分类号: TP18

文献标识码: A

文章编号: 1000-4220(2014)11-2511-04

Multi-instance Learning Bag-level Covering-kNN Algorithm

ZHAO Shu^{1,2}, RUI Chen^{1,2}, CHEN Jie^{1,2}, ZHANG Yan-ping^{1,2}

¹ (Key Lab of Intelligent Computing & Signal Processing, Ministry of Education, Anhui University, Hefei 230601, China)

² (Institute of Artificial Intelligence, Anhui University, Hefei 230601, China)

Abstract: Multi-instance learning is a new framework in machine learning. An extensive number of noises in the positive bags is the inherent difficulty of multi-instance learning. In order to improve the classification accuracy, this paper puts forward a novel multi-instance learning bag-level Covering-kNN algorithm to exclude the noises in multi-instance data set. The learning results of Covering algorithm is a set of sphere neighbors and each sphere neighbor only contains patterns belong to the same class. This feature help us re-organize the structure of bags in multi-instance data set. Generally speaking, in order to exclude false positive instances in the positive bags, first, we reconstruct the structure of multi-instance data set by treating the sphere neighbors obtained using Covering algorithm as the new structure of bags. Thus, improving the separable of multi-instance samples in the new feature space. Then, the bag-level kNN algorithm is utilized to exclude the noises in positive bags and predict the labels of test bags. The experiments demonstrate the effectiveness of the proposed multi-instance bag-level Covering-kNN algorithm.

Key words: machine learning; multi-instance learning; covering algorithm; kNN algorithm

1 概述

作为传统监督学习的一个变种, 在多示例学习中, 学习样本是一系列的包的集合, 而每个包又是由多个示例所组成的. 包做为一个整体具有唯一的标记, 而包中的示例是没有标记的. 在传统的监督学习中, 样本与标记之间是一一对应的关系, 即每一个样本都具有唯一明确的标记; 在无监督学习中, 所有的样本都是没有标记的; 而在多示例学习中, 只要在一个包中有一个示例是正示例, 那么这个包就是正包, 当包中所有示例都是反示例时, 这个包才是反包. 也就是说, 在多示例中, 样本整体有标记 (包层次有标记), 而部分无标记 (样本层次无标记). 由于正包中可能存在大量的假正例, 因此, 多示例数据集中的正包具有相当大的歧义性. 多示例学习一经发现, 就引起了国际机器学习界的极大关注, 被认为是一种新型的机器

学习的框架^[1]. 目前, 多示例学习已被广泛的应用于基于内容的图像检索^[2,3], 视觉追踪^[4,5], 物体检测^[6]和计算机安全^[7]等领域.

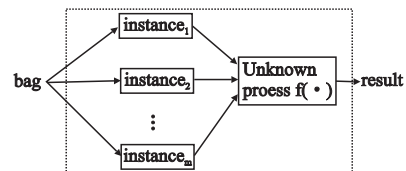


图1 多示例学习问题描述

Fig. 1 Description of multi-instance problem

多示例学习是在药物活性预测问题中被首先被 T. G. Dietterich 等人发现的^[8]. 由于一个特定的药物分子可以通过

收稿日期: 2013-06-18 收修改稿日期: 2013-08-13 基金项目: 国家自然科学基金项目 (61073117, 61175046) 资助; 安徽省教育厅基金项目 (KJ2013A016) 资助; 安徽大学研究生学术创新项目 (10117700183) 资助. 作者简介: 赵 姝, 女, 1979 年生, 博士, 副教授, 主要研究方向为智能计算和机器学习; 芮 辰, 男, 1989 年生, 硕士研究生, 主要研究方向为机器学习; 陈 洁, 女, 1982 年生, 博士研究生, 讲师, 研究方向为智能计算和机器学习; 张燕平 (通信作者), 女, 1962 年生, 博士, 教授, 主要研究方向为智能计算、商空间理论和机器学习.

旋转内部化学键来改变自身的形状. 对于一个特定的分子来说, 只要有一种形状能够使该分子具有制造药物的活性, 那么这个分子就是具有活性的; 当一个分子的所有形状都不能够用来制造药物时, 那么该分子才是无活性的.

为了解决多示例问题, 近年来国内外的研究者提出了许多多示例学习算法. 从整体上来说, 可以将这些算法分为两类: 包层次的算法和示例层次的算法. 大部分的多示例学习算法是示例层次的, 例如, T. G. Dietterich 等人在提出多示例问题时设计了三个轴平等矩形 (Axis-Parallel Rectangles, APR) 学习算法^[8]. 这种方法用一个最小的矩形找到多示例数据集中每一维数据的上限和下限, 从而将所有不在矩形框内的样本排除, 最终落在矩形框内的样本即为真正例. O. Maron 和 T. Lozano-Pérez 提出了多示例学习的第一个框架—Diverse Density (DD) 算法^[9]. DD 算法是一个基于概率的算法, 该算法通过二个梯度下降过程在示例空间中找到一个概念点 p , 使该点距离数据集中的所有的正包更近, 同时远离所有的反包. Q. Zhang 和 S. A. Goldman 将 EM 算法和 DD 算法相结合, 提出了 EM-DD 算法^[10]. 他们在 EM 算法的 E 步骤中估计出决定各训练包概念标记的训练示例, 然后在 M 步骤中对这些训练示例使用多样性密度算法以最大化多样性密度, 迭代执行算法直到 E 步骤和 M 步骤收敛为止. 周志华和张敏灵对经典的 BP 算法进行了拓展, 提出了多示例学习的神经网络算法 BP-MIP^[11]. 最近, C. H. Li 等人提出了多示例学习的并行神经网络算法 MIBP^[12].

在包层次的多示例学习算法方面, 为了解决多示例问题, J. Wang 和 J. D. Zucker 对 k 近邻算法进行了改进, 提出了两个 k 近邻算法的变种, 即 Bayesian- k NN 算法和 Citation- k NN 算法^[13]. 前者是基于统计学习中的贝叶斯方法, 后者利用了信息科学中引用的概念. S. Andrews 等人将支持向量机和核方法引入多示例学习邻域, 提出了包层次的多示例 SVM 算法-MI-SVM^[14].

本文将邻域覆盖算法和包层次的 k 近邻算法相结合, 首先使用邻域覆盖算法重新组织多示例数据集, 用邻域覆盖算法得到的一系列球形邻域作为新的包结构, 然后使用包层次的 k 近邻算法排除正包中大量的假正例. 测试时使用包层次的 k 近邻算法对新的包结构进行学习, 预测测试包的标签. 实验结果表明, 提出的新方法可以有效减低多示例数据集正包中大量噪声所带来的干扰, 所提出的算法 Covering- k NN 可以和已有的最优秀多示例算法相媲美.

2 相关算法

2.1 覆盖算法

覆盖算法^[15-16]是由张铃和张钹在 1999 年提出的, 是一种基于 McCulloch-Pitts 神经网络模型^[17]的构造型监督学习算法. 覆盖算法的基本思想是将样本点映射到空间中的一个超球面 S^n 上后, 进行球形邻域的划分. 假设样本空间中各样本的向量的长度相等 (否则可将样本投影到 S^{n+1} 上), 给定 w , $x \in R^N$, $\theta \in R$, 此时 $w \cdot x - \theta > 0$ 表示球面上由超平面 $w \cdot x - \theta = 0$ 所分割的正半空间的部分, 称为球面上的球形邻域. 若 w 与 x 等长, 则 w 就是这个球形邻域的中心. 以每个球形邻域作为一个神经元, 取 $w \cdot x - \theta = 0$ 为其功能函数. 此时, 构造一个网

络, 对给定的样本集进行分类, 等价于求出一组邻域, 对给定样本集 S 中的点, 按分类的要求将它们分隔开来. 通过覆盖算法, 在样本空间构造覆盖簇 $\{C_i\}$, 使 c_i 只盖住某一类点, 且 $\bigcup_i C_i = S$. 设已求得覆盖组 C_1, C_2, \dots, C_n , 取三层神经网络, 隐层取 n 个神经元, 每个神经元为一个覆盖, 第 i 个神经元的激励函数为 C_i 的特征函数. 输出层取 k 个神经元, 第 i 个神经元的输入为覆盖第 i 类点的覆盖的输出, 激励函数采用或门, 这样的三层神经网络, 就可对样本集 S 进行分类.

目前大多数粗糙集都是建立在不同的邻域基础上的, 覆盖算法对同类型样本的学习结果是一系列的球形邻域, 令 U 为论域, $\mathcal{C} = \{c_1, c_2, \dots, c_n\}$ 是 U 中同类样本的覆盖集. 如果样本 $x \in c_i$, $i \in \{1, 2, \dots, n\}$, 那么 c_i 也是 x 的一个邻域. 因此当 $x \in c_i$ 时, 就可以得到由覆盖导出的邻域粗糙集系统.

2.2 k 近邻算法

k 近邻算法 (k Nearest Neighbors, k NN) 是一个经典的机器学习方法. k 近邻算法用样本之间的距离来度量样本之间的相似度, 认为样本之间的距离越近越相似. k 近邻算法在对样本进行分类时, 数据集中的每个样本都带有标签, 表示样本自身的类别. 实际上, k 近邻算法将训练样本当做模型. 当对一个新样本进行分类, 预测新样本的标记时, 必须首先确定测试样本与训练集中的每个样本之间的相似度, 然后进一步考虑训练集中与新样本相距最近的样本的类别. 将新样本分配到一个类别中的依据是: 这个类别包含了 k 个最近样本中的最多的样本, 即新样本与这个类别的相似度最高. k 近邻算法的时间复杂度为 $O(n)$, 其中 n 是训练集的大小. k 近邻算法优点是并不要求在分类时给定待分样本的分布函数, 因此具有直观、无需先验统计知识等优点, 从而被广泛应用于机器学习和数据挖掘等领域.

3 本文算法

为了便于描述本文提出的 Covering- k NN 算法, 首先给出一些在本文中常用的定义. 给定一个多示例数据集 D . 将数据集 D 中的示例的集合定义为: $\{x_k, k = 1, 2, \dots, m\}$. 不妨设正包中示例组成的集合为 $D^+ = \{x_i^+, i = 1, 2, \dots, p\}$, 反包中的示例组成的集合为 $D^- = \{x_j^-, j = 1, 2, \dots, n\}$. 将 D 中示例 x 的维度记为 d .

本文所提出的 Covering- k NN 算法为了排除正包中大量的反示例, 使用覆盖算法重新组织多示例样本中包的结构. 首先, 用邻域覆盖算法处理多示例数据集, 将得到的一系列球形邻域作为新的包结构; 接下来, 使用包层次的 k 近邻算法排除正包中大量的噪声, 从而降低正包中的噪声对样本分类带来的干扰; 最后, 在预测测试包的类别时, 同样使用包层次的 k 近邻算法对形成的球形邻域进行学习从而得到测试包的标记.

由于在多示例学习中, 包是由多个示例所组成的, 样本可以看作是一些示例所组成的集合. 因此无法直接使用一般的距离公式 (例如欧式距离) 来度量包之间的相似度. 为了能够测量包之间的相似度 (即集合之间的距离), 本文使用豪斯多夫 (Hausdorff) 距离来计算包之间的距离. J. Wang 和 J. D. Zucker 提出了两种豪斯多夫距离来计算给定两个集合之间的距离^[13], 即最大豪斯多夫距离 ($\max \delta_{H(A, B)}$) 和最小豪斯多夫距离 ($\min \delta_{H(A, B)}$). 给定两个集合 $A = \{a_1, a_2, \dots, a_m\}$ 和 B

$= \{b_1, b_2, \dots, b_n\}$,

最小豪斯多夫距离记为:

$$\min \delta_H(A, B) = \min_{a \in A, b \in B} d(a, b), \quad (1)$$

最大豪斯多夫距离记为:

$$\max \delta_H(A, B) = \max(\delta_H(A, B), \delta_H(B, A)), \quad (2)$$

其中 $\delta_H(A, B) = \max_{a \in A} \min_{b \in B} d(a, b)$ 表示示例 a 和

示例 b 之间的欧氏距离。

本文所提出的 Covering-kNN 算法可分为两个过程:覆盖过程和分类过程。算法的具体步骤如下所示。

3.1 覆盖过程

1) 定义一个变换 $T(x) = (x, \sqrt{R^2 - \|x\|^2})$, 其中 $R = \max\{\|x\| | x \in D\}$, R 为球形邻域的半径;

2) 通过变换 $T(x)$ 将多示例数据集 D 中的所有样本映射到球面上, 即 $D \rightarrow S^d$;

3) 初始化覆盖集: $C = \emptyset$;

4) 从数据集 D 中随机选取一个示例 x ;

5) 求 x 和距离 x 最近的异类样本 x' 之间的距离, 记其为 d_1 , $d_1 = \max\{\langle x, x' \rangle | x \in D_i, x' \notin D_i\}$;

6) 求 x 和距离 x 最远的同类样本 x' 之间的距离, 记其为 d_2 , $d_2 = \min\{\langle x, x' \rangle | x, x' \in D_i, \langle x, x' \rangle > d_1\}$;

7) 将覆盖半径设置为 r , $r = (d_1 + d_2) / 2$;

8) 使用示例 x 作为覆盖中心 r 作为覆盖半径作覆盖 c_i , c_i 中只含示例 x 的同类样本;

9) 重复做步骤 4) ~ 步骤 8), 直到 $D = \emptyset$, 即 D 中所有样本均被覆盖时为止, 覆盖过程结束。

注: $\langle x, x' \rangle$ 表示示例 x 和示例 x' 之间的内积。

3.2 分类过程

1) 对于在过程 1 中构造的每一个标记为正的球形邻域 (正覆盖) c_i , 分别计算其到附近最近的 k_1 个球形邻域之间的豪斯多夫距离;

2) 如果 c_i 附近标记为负的球形邻域更多, 那么将球形邻域 c_i 的标记改为负; 否则, 球形邻域 c_i 的标记不变;

3) 重复执行 1) ~ 2), 直到过程 1 中所有标记为正的球形邻域的标记均被更新为止;

4) 对于一个测试包 $testBag_i$, 计算其到附近最近的 k_2 个球形邻域之间的豪斯多夫距离;

5) 如果测试包 $testBag_i$ 附近标记为负的球形邻域更多, 则该测试包 $testBag_i$ 的标记为负; 否则, 标记为正;

6) 重复执行 4) ~ 5), 直到得到所有的测试包的类别为止, 分类过程结束。

4 实验与分析

为了验证 Covering-kNN 算法的有效性, 本文使用麝香分子 (MUSK) 数据集和三个人工合成的多示例数据集 LJ-16.30.2, LJ-80.166.1 和 LJ-160.166.1 进行实验。

4.1 数据集

MUSK 数据集是至今为止唯一真实世界的多示

例学习数据集, 一共有两个麝香分子数据集, 即 MUSK1 和 MUSK2, MUSK 数据集的详细信息如表 1 所示。

表 1 MUSK 数据集的详细信息

Table 1 Detailed information of MUSK data sets

数据集	MUSK1	MUSK2
包总数	92	102
有活性包	47	39
无活性包	45	63
包中平均示例数	5.2	64.7

Dooley 等人根据麝香分子和受体之间的亲和力模型, 提出了一种人工生成麝香分子数据集的方法^[18]。为了验证我们算法的有效性, 在本文的实验中我们使用了 LJ-16.30.2, LJ-80.166.1 和 LJ-160.166.1 三个人工生成的麝香分子数据集进行实验。在人工生成的多示例数据集中, 每个人工生成的包中包含了 3 ~ 5 个示例。数据集的命名规则如下: LJ-r.f.s, 其中 r 表示对分子的活性具有影响的相关属性维数, f 表示数据集中示例的维数, s 表示属性的相关系数, 如果 $s = 1$, 则表示存在 0 和 1 两种相关系数。在人工麝香分子数据集中, 分子的活性是一个介于 0-1 之间的实数。Dooley 等人指定, 当活性大于 0.5 时分子是有活性的。因此, 在人工的麝香分子数据集中, 分子的活性由该分子活性最大的形状所决定。如果一个特定的分子具有活性大于 0.5 的形状, 则该分子是有活性的, 否则该分子是没有活性的。LJ-16.30.2, LJ-80.166.1 和 LJ-160.166.1 的维数分别是 30 维, 166 维和 166 维, 对活性具有影响的相关属性分别是 16 维, 80 维和 160 维。

4.2 实验设置

本文使用 10 倍交叉验证 (10CV) 方法进行实验。在每次实验中, 将多示例样本平均分成 10 份, 其中 9 份用做训练集, 剩下的一份用做测试集, 10 份样本轮流作为测试集, 算法重复执行 10 次, 取 10 次验证的平均值为 10 倍交叉验证的最终精度。

J. Wang 和 J. D. Zucker 的工作^[13]表明, 当 kNN 算法应用于多示例问题时, k 的取值较小时分类精度较高。对于真实的麝香分子数据集和人工的麝香分子数据集, 我们为 k 近邻算法中的 k 值设置了三组不同的参数, 当将 k_1 的值设置为 1, 3, 5 时, 分别令 k_2 的值为 3, 5, 7。目的是通过实验得出分类精度最高时 k 值。当 $k_1 = 1$ 时, 我们考察包层次的最近邻算法应用于多示例学习问题时的表现。

4.3 实验结果和分析

详细的实验结果列在表 2 和下页表 3 中。对 MUSK1 来

表 2 在 MUSK 数据集上得到的分类精度 (%)

Table 2 Classification accuracy obtained in MUSK data sets (%)

数据集	MUSK1		MUSK2	
相似性测度	最小距离 ¹	最大距离 ²	最小距离	最大距离
$k_1 = 1, k_2 = 3$	87.8	81.3	91.2	86.4
$k_1 = 3, k_2 = 5$	92.8	84.3	82.4	77.3
$k_1 = 5, k_2 = 7$	91.3	79.6	84.0	70.1

说, 当 $k_1 = 3, k_2 = 5$ 时, Covering-kNN 算法得到了最高的预

¹ 最小豪斯多夫距离

² 最大豪斯多夫距离

测精度 92.8% ,使用最小豪斯多夫距离做为包之间相似度的度量方式;对 MUSK2 来说,当 $k_1 = 1$ $k_2 = 3$ 时,Covering-kNN 算法得到了最高的预测精度 91.2% ,包间相似度的度量方式同样是最小豪斯多夫距离.实验表明,对于真实世界的麝香分子数据集,用最小豪斯多夫距离度量包之间的相似性的效果明显好于最大豪斯多夫距离.

表 3 在人造麝香分子数据集上得到的分类精度(%)

Table 3 Classification accuracy
obtained in artificial MUSK data sets(%)

数据集	LJ-16.30.2		LJ-80.166.1		LJ-160.166.1	
相似性测度	最大 距离	最小 距离	最大 距离	最小 距离	最大 距离	最小 距离
$k_1 = 1$ $k_2 = 3$	80.6	76.8	89.7	83.1	97.2	92.3
$k_1 = 3$ $k_2 = 5$	84.2	77.6	93.0	92.2	95.6	92.5
$k_1 = 5$ $k_2 = 7$	79.7	74.3	91.3	86.4	92.4	89.4

对于人造多示例数据集 LJ-16.30.2 和 LJ-80.166.1 最高的分类精度均在 $k_1 = 3$ $k_2 = 5$ 时取得;对于 LJ-160.166.1 当 $k_1 = 1$ $k_2 = 3$ 时,得到最高的分类精度.与真实的麝香分子数据集上的实验相同的是,用最小豪斯多夫距离度量包之间的相似性的效果明显优于最大豪斯多夫距离.

表 4 和其他多示例算法的比较

Table 4 Compare with other multi-instance algorithms

算 法	MUSK1	MUSK2	LJ-16.30.2	LJ-80.166.1	LJ-160.166.1
IAPR	92.4	89.2	N/A	N/A	N/A
DD	88.9	82.5	86.7	N/A	88.0
EM-DD	96.8	96.0	83.3	87.0	88.3
B-kNN	90.2	82.4	N/A	N/A	N/A
C-kNN	92.4	86.3	83.3	91.4	97.8
MI-SVM	77.9	84.3	90.0	92.3	99.0
本文算法	92.8	91.2	84.2	93.0	97.2

表 4 对 Covering-kNN 算法和已有的最优秀的多示例学习算法进行了比较.从表 4 中可以明显看出,在两个真实的和三个人造的麝香分子数据集上,本文所提出的算法均可以和已有的经典的多示例学习算法相媲美.

当和示例层次的经典多示例学习算法相比较时,Covering-kNN 算法仅在真实麝香分子数据集上表现落后于 EM-DD 算法.当和包层次的经典多示例学习算法相比较时,Covering-kNN 算法的性能优于已有的两个包层次的 k 近邻算法 Bayesian-kNN 算法和 Citation-kNN 和包层次的多示例支持向量机算法 MI-SVM.

本文认为,Covering-kNN 之所以优于其他包层次多示例学习算法的原因在于:由于正包中大量的假正例所带来的干扰是多示例问题的根本困难所在.而在邻域覆盖算法的作用下,原有的离散的无规律的多示例数据集的包结构被重新组织为一系列的球形邻域,在每个球形邻域中只包含标记相同的示例.用生成的球型邻域作为新的多示例包结构,每个球型邻域类似于一个同类示例所组成的"聚类",从而将原来的无规则的离散包重新组织成为一个个有规律的,同类示例高度聚集的多示例包,便于包层次的 kNN 算法排除大量的假正例所带来的干扰.在大量的假正例被排除后,Covering-kNN 算法取得了比其他包层次多示例算法更高的分类精度.

5 结 论

为了降低多示例数据集的歧义性,排除正包中大量的反

示例,提高分类精度.本文将覆盖算法和包层次的 k 近邻算法相结合提出了多示例学习的 Covering-kNN 算法.首先通过覆盖算法预处理多示例数据集,重新组织多示例数据集中包的结构,排除正包中的反示例所带来的歧义性.实验表明,本文所提出的算法具有很高的分类精度,可以有效降低正包中大量的噪声所带来的分类误差.

此外,本文所提出的算法使用包层次的 kNN 算法来排除正包中的噪声并预测测试包的类别.其他监督学习算法能否更准确的排除正包中的反示例并预测测试包的类别,还有待进一步的研究.

References:

- [1] Maron O. Learning from ambiguity [D]. Massachusetts Institute of Technology, USA, 1998.
- [2] Fu Z Y, Robles-Kelly A, Zhou J. MILIS: multiple instance learning with instance selection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, USA, 2011, 33(5): 958-977.
- [3] Nguyen D T, Nguyen C D, Hargraves R, et al. mi-DS: multiple-instance learning algorithm [J]. IEEE Transactions on Systems, Man, and Cybernetics Society, 2013, 43(1): 143-154.
- [4] Babenko B, Yang M H, Belongie S. Robust object tracking with on-line multiple instance learning [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(8): 1619-1632.
- [5] Xie Y, Qu Y Y, Li C H, et al. Online multiple instance gradient feature selection for robust visual tracking [J]. Pattern Recognition Letters, 2012, 33(9): 1075-1082.
- [6] Qi Z Q, Xu Y T, Wang L S, et al. Online multiple instance boosting for object detection [J]. Neurocomputing, 2011, 74(10): 1769-1775.
- [7] Bellare M, Ristenpart T, Tessaro S. Multi-instance security and its application to password-based cryptography [J]. Advances in Cryptology-CRYPTO, 2012: 312-329.
- [8] Dietterich T G, Lathrop R H, Lozano-Pérez T. Solving the multiple instance problem with axis-parallel rectangles [J]. Artificial Intelligence, 1997, 89(1): 31-71.
- [9] Maron O, Lozano-Pérez T. A framework for multiple-instance learning [J]. Advances in Neural Information Processing Systems, 1998, 10: 570-576.
- [10] Zhang Q, Goldman S A, et al. EM-DD: an improved multiple-instance learning technique [J]. Advances in Neural Information Processing Systems, 2001, 14(2022): 1073-1080.
- [11] Zhou Z H, Zhang M L. Neural networks for multi-instance learning [P]. In Proceedings of the International Conference on Intelligent Information Technology, Beijing, China, 2002: 455-459.
- [12] Li C H, Gondra I, Liu L. An efficient parallel neural network-based multi-instance learning algorithm [J]. The Journal of Supercomputing, 2012, 62(2): 724-740.
- [13] Wang J, Zucker J D. Solving multiple-instance problem: a lazy learning approach [P]. Proceedings of the 17th International Conference on Machine Learning, 2000: 1119-1126.
- [14] Andrews S, Tsochantaridis I, Hofmann T. Support vector machines for multiple-instance learning [P]. In Advances in Neural Information Processing Systems, 2003, 15: 561-568.
- [15] Zhang L, Zhang B. A geometrical representation of mcculloch-pitts neural model and its applications [J]. Neural Networks, IEEE Transactions on, 1999, 10(4): 925-929.
- [16] Zhang Lin, Zhang Bo. A geometrical representation of M-P neural model and its applications [J]. Journal of Software, 1998, 9(5): 334-338.
- [17] McCulloch W S, Pitts W. A logical calculus of the ideas immanent in nervous activity [J]. Bulletin of Mathematical Biology, 1943, 5(4): 115-133.
- [18] Dooly D R, Zhang Q, Goldman S A, et al. A multiple instance learning of real-valued data [J]. The Journal of Machine Learning Research, 2002, 3: 651-678.

附中文参考文献:

- [16] 张 铃, 张 钹. M-P 神经元模型的几何意义及其应用 [J]. 软件学报, 1998, 9(5): 334-338.