

# 基于结构学习的 KNN 分类算法<sup>\*</sup>

孙岩<sup>1,2</sup> 吕世聘<sup>3</sup> 王秀坤<sup>3</sup> 唐一源<sup>2</sup>

(辽宁师范大学计算机与信息技术学院 大连 116029)<sup>1</sup> (大连理工大学神经信息学研究所 大连 116023)<sup>2</sup>  
(大连理工大学计算机科学与工程系 大连 116023)<sup>3</sup>

**摘要** KNN(K-Nearest Neighbor)算法和贝叶斯网络分类算法(Bayesian Network, BN)都是目前应用非常广泛的分类算法。本文首先分析了 KNN 和 BN 的分类特点,然后在保留了两个算法在分类问题中优点的基础上,提出了基于贝叶斯网络结构学习的 KNN 算法(BN-KNN)。实验结果表明, BN-KNN 算法能够有效地提高分类的正确率。

**关键词** 贝叶斯网络, K-近邻算法, 距离加权

## K-Nearest Neighbor Algorithm Based on Learning Structure

SUN Yan<sup>1,2</sup> LV Shi-Pin<sup>3</sup> WANG Xiu-Kun<sup>3</sup> TANG Yi-Yuan<sup>2</sup>

(Computer Science Department of Liaoning Normal University, Dalian 116029)<sup>1</sup>

(Neuroinformatics Institute of Dalian University of Technology, Dalian 116023)<sup>2</sup>

(Computer Science Department of Dalian University of Technology, Dalian 116023)<sup>3</sup>

**Abstract** K-Nearest Neighbor algorithm(KNN) and Bayesian network classification algorithm(BN) are currently widely used classification algorithms. At first, this paper analyzes the KNN and BN classified features, and then retains the merits of two classification algorithm. At last, a KNN algorithm based on learning the Bayesian network structure (BN-KNN) is presented. Experimental results show that BN-KNN algorithm can be used to improve the classification accuracy.

**Keywords** Bayesian network, K-nearest neighbor algorithm, Distance-weighted

## 1 KNN 分类算法

KNN 分类算法是一种非常有效的非参数分类算法,已经广泛应用于分类、回归和模式识别等领域中,是一种传统的基于统计的模式识别方法。该算法的主要优点是学习过程只是简单地存储已知的训练数据,当遇到新的查询样本时,取出一系列相似的样本,用来分类新的查询样本。

该算法的基本思想<sup>[1]</sup>是:首先,把训练样本作为欧氏空间的点存放,所有样本对应于  $n$  维空间的点,根据客观的事实或专家的经验,将一定数量的训练数据设定好分组;然后,给定一个待分类数据库样本,计算训练数据和它的距离(一般采用欧几里德距离),根据这个距离,确定在训练样本集中与该新样本距离最近(最相似)的  $k$  个样本,使用这  $k$  个邻居的类别作为新样本的候选类别。最后,根据这  $k$  个样本所属的类别判定待分类样本所属的类别。该样本与  $k$  个邻居之间的相似度按类别分别求和,减去一个预先得到的截尾阈值,就得到新样本的类别测度。

实现该算法的具体步骤如下:

① 构建训练样本集合  $D$ 。

② 设定  $k$  的初始值。 $k$  值的确定目前并没有很好的方法,根据具体应用领域的不同,  $k$  的取值有很大差别,一般采用先设定一个初始值,然后根据实验测试的结果来不断调整  $k$  的取值,使结果达到最优。比如:在脑成像实验中,一般  $k$  初始

值定为 20,30 或更多一些,但对于网页的文本分类可能一般要定为几百到几千之间。

③ 在训练样本集中选出与新样本最相似的  $k$  个样本。假定所有的样本对应于  $n$  维空间  $R^n$  中的点,一个样本的最近邻可以根据标准的欧式距离定义。任意的样本  $x$  表示为特征向量  $x = (x_1, x_2, \dots, x_m)$ ,  $x_i$  表示样本  $x$  的第  $i$  个特征的值。那么两个样本  $x_i$  和  $x_j$  的距离定义为  $d(x_i, x_j)$ , 计算公式为

$$d(x_i, x_j) = \sqrt{\sum_{i=1}^k (x_i^i - x_j^i)^2} \quad (1)$$

④ 给定一个待分类的样本  $x_q$ ,  $x_1, \dots, x_k$  表示训练样本数据库  $D$  中与  $x_q$  距离最近的  $k$  个样本,设离散的目标函数为  $F: R^n \rightarrow v_i, V = \{v_1, \dots, v_k\}, v_i \in V$ ,

$$F(x_q) = \arg \max_{i \in V} \sum_{i=1}^k \delta(v, f(x_i)) \quad (2)$$

当  $a=b$  时,  $\delta(a, b) = 1$ ; 否则,  $\delta(a, b) = 0$ 。

⑤ 使用最近邻的大多数投票法作为待查询样本的预测值。

作为一个被广泛使用的传统 KNN 算法,有以下主要优点:

(1) 对训练数据中存在噪声或有缺失数据的数据库,算法有很好的健壮性,不影响算法结果的正确性;

(2) 当给定足够大的训练集合时,或对于海量的数据库,

<sup>\*</sup> 国家自然科学基金项目(60472017, 30670699)资助课题。孙岩 讲师,博士生,研究方向:人工智能、计算机图形学;吕世聘 博士生,研究方向:数据挖掘、决策支持系统;王秀坤 教授,博士生导师,研究方向:数据库系统、决策支持系统;唐一源 教授,博士生导师,研究方向:神经信息学。

该算法仍然是非常有效的;

(3)不需要存储过多的数据,只需存储训练样本的数据。

算法的不足主要表现在:

(1)需要主观决定最近邻的参数  $k$ 。考虑一种最极端的情况:如果参数  $k$  取数据库中所有样本的个数,则新样本的分类结果为全局最优解,否则分类结果为局部最优解。所以,参数  $k$  可以看作是全局最优解和局部最优解的一个折衷;

(2)应用 KNN 算法的一个实践问题是,样本的距离是根据样本所有特征计算的。在这些特征中,有些特征与分类是强相关,有些特征与分类弱相关,还有一些特征(可能是大部分特征)是不相关的。这样,近邻间的距离会被大量的不相关特征所支配,这种情况也称为维度灾难,最近邻方法对这个问题特别敏感。

针对以上不足,有许多文献对此进行改进<sup>[2-4]</sup>,文[3]认为:传统的 KNN 算法中,所有邻域对于决定分类变量的值有平等贡献,据此提出一个扩展的距离加权 KNN 算法。根据待查询的新样本与训练集的距离多少,赋予不同的权重,该距离越近,赋予的权重越大。相反,距离越远,所赋予的权重越小,并设权重为两个样本间距离平方的倒数,作为该距离的权重。

因为近邻之间的距离是由样本的所有特征按照相同度量确定的,这就很可能造成维度灾难的问题(近邻间的距离会被大量的不相关特征所支配)。文[1]提出的解决办法是对分类强相关的特征赋较高的权重,与分类弱相关的特征赋较低的权重,与分类不相关的特征赋 0 权重。接下来的关键问题是如何确定哪些变量是强相关,哪些变量是弱相关及不相关。

本文对 KNN 算法的一个明显的改进是在相似度评价中,不仅考虑了查询样本与训练样本的数值特征,而且加入了结构的特征,考虑了变量之间的因果关系,用贝叶斯网络的结构学习所建立的网络结构,作为评价其结构特征的基础,在贝叶斯的分层结构中,对功能变量有直接影响的结构因子,赋给较大的权重,影响较小的因子,可以赋给较小的权重。这样,可以提高分类结果的正确性。

## 2 结构学习算法

贝叶斯网络是一种将贝叶斯概率方法和有向无环图的网络拓扑结构有机结合的表示模型。描述了数据库中数据项及其相互之间的依赖关系。贝叶斯网络就是根据各个变量之间的概率关系建立起来的图论模型,随着近年来数据库规模的不断扩大,研究发现应用贝叶斯方法进行数据库的知识发现要优于目前应用在数据库知识发现的方法<sup>[5-9]</sup>。

贝叶斯网络是根据马尔可夫假设寻找的满足条件独立性限制的模型结构。举例来说,对于给定两个变量域  $X$  和  $Y$ ,判断它们是否绝对独立,可以用它们之间的概率来反映其独立性。而且,由于贝叶斯方法用的是概率网络,无需对变量域中的个别变量的独立性进行判定,只需根据变量间的概率分布来计算整个网络模型的独立性。

**定义** 给定一个随机变量集

$$x = \{X_1, X_2, \dots, X_n\}$$

其中  $X_i$  是一个  $m$  维向量。贝叶斯信念网络说明了  $\chi$  上的一个联合条件概率分布。贝叶斯信念网络定义如下:

$$B = \langle G, \theta \rangle$$

第一部分  $G$  是一个有向无环图,其顶点对应于有限集  $\chi$  中的随机变量  $X_1, X_2, \dots, X_n$ ,其弧代表一个函数依赖关系。如果有一条弧由变量  $Y$  到  $X$ ,则  $Y$  是  $X$  的父结点或者直接前

驱,而  $X$  则是  $Y$  的后继。如果一个结点没有父结点,则该结点的一个先验概率函数即为  $\pi_i$ ,一旦给定其双亲,图中的每个变量结点独立于该结点的非后继。在图中, $X_i$  的所有双亲变量用集合  $\pi_i$  表示。

第二部分  $\theta$  代表用于量化网络的一组参数。对于每一个  $X_i, \pi_i$  的取值  $x_i, \theta$  表示在给定  $\pi_i$  发生的情况下  $x_i$  事件发生的条件概率。因此实际上,一个贝叶斯网络给定了变量集合  $x$  上的联合条件概率分布:

$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P(X_i | \pi_i) \quad (3)$$

贝叶斯网络结构学习算法可以表示为:给定一组训练样本  $D = \{x_1, x_2, \dots, x_n\}$ ,  $x_i$  是  $X_i$  的实例,根据文[5]中对贝叶斯网络概率的推导可得

$$P(B_s, D) = P(B_s) \prod_{i=1}^n \prod_{j=1}^{q_i} \frac{(r_i - 1)!}{(N_{ij} + r_i - 1)!} \prod_{k=1}^{r_i} a_{ijk}! \quad (4)$$

其中,  $N_{ij} = \sum_{k=1}^{r_i} a_{ijk}$ ,  $\Phi_i$  表示变量  $x_i$  的父结点在训练样本集  $D$  中出现的变量及相应的取值,如果变量  $x_i$  没有父结点,则  $\Phi_i = \phi_i$ ;  $i$  表示训练样本集中变量的个数,  $V_i$  是  $x_i$  可能的取值列表  $(v_{i_1}, \dots, v_{i_{r_i}})$  中的值,  $q_i = |\Phi_i|$ ,  $j$  表示该结点的父结点的取值,  $r_i = |V_i|$ ;  $k$  表示当前处理结点所具有的可能取值,  $a_{ijk}$  表示在训练样本集  $D$  中,  $x_i = v_{i_k}, \pi_i = \Phi_i$  中出现的样本个数。

这样,我们就可以通过式(4)及训练样本集  $D$  来获得结点所能组成的不同网络结构的概率,可记为  $P(B_s, D)$  和  $P(B_j, D)$ ,其中  $i$  和  $j$  为结点变量组成不同的网络结构的个数。

接下来,我们应该在这些不同网络结构的概率中,寻找一个最匹配该样本的贝叶斯网络结构。常用的学习算法通常是引入一个评估函数,  $S(B|D)$ :

$$\frac{P(B_s | D)}{P(B_j | D)} = \frac{\frac{P(B_s, D)}{P(D)}}{\frac{P(B_j, D)}{P(D)}} = \frac{P(B_s, D)}{P(B_j, D)} \quad (5)$$

其中,  $B_s$  表示任意一个包含训练样本集中变量的网络结构。通过式(5)得到一对网络结构的比率,对这一系列结构的后验概率进行排序,概率越大,其相应的网络结构越优。这样,我们就可以通过一个训练样本集及结点的先验顺序,推导出一个最优的网络结构,从而完成网络结构的学习。

## 3 基于结构学习的 KNN 分类算法

### 3.1 贝叶斯网络结构的学习

在进行贝叶斯网络结构的学习过程中,把结构变量和功能变量一同作为贝叶斯网络中的结点变量,通过第二部分的学习,得到了包含有结构和功能结点变量的贝叶斯最优网络结构。这样,我们就知道了哪些结构变量对功能变量所起的作用比较大,哪些变量所起的作用比较小。

在 KNN 分类算法中计算结点变量相似度的时候,把样本集中的所有变量结点等同对待,即权重相同。也就是说,各特征向量在进行相似度计算时,认为所有特征向量对分类结果所起的作用是相同的。但在解决实际问题时,我们往往还不清楚哪些结构变量对最终的功能变量(分类变量)所起的作用比较大,哪些变量所起的作用比较小,所以在实验的过程中,会取尽可能多的结构变量作为参考,来确定其分类结果。因此,如果把所有的结构变量对结果变量所起的作用都等同起来,则一定影响结果的正确性,并且随着测试数据的增多,算法的计算量将迅速增加。

本文提出：先用贝叶斯网络结构学习算法，找到最优的网络结构，即找到了网络中待分类变量的结构特征。这样，根据各结点变量在贝叶斯网络结构中位置的不同，可以判断出该结点变量在相应的分类结构中所起作用的不同，从而根据广度优先搜索，确定在 KNN 算法中计算样本相似度时各结点变量的权重，并且可以根据具体应用领域的不同，交互式地计算结点变量的权重。

进行贝叶斯网络结构学习时，我们采用启发式搜索，算法的复杂度为  $O(mu^2 n^2 r)$ ，其中  $m$  为训练样本集中所包含的样本数， $n$  为结点变量的个数， $u$  为一个结点可以有父结点的最大个数。 $V_i$  是  $x_i$  可能的取值列表  $(v_{i_1}, \dots, v_{i_n})$  中的值， $r_i = |V_i|$ 。因为我们进行贝叶斯网络结构学习的目的是找到各结点变量对分类结果所起作用的不同，所以为了降低贝叶斯网络结构学习的复杂度，在进行结构学习时，规定每个结点的父结点最多有 2 个，把每个结点变量的数据进行二值化处理，这样网络学习的复杂度为： $O(mn^2)$ 。

把对功能变量的结果所起的作用比较大的结构变量，在 KNN 算法中计算结点变量的相似度时，赋予较大的权重。反之，那些对功能变量所起的作用比较小或比较间接，在计算 KNN 算法中的相似度时赋予较小的权重。

经过贝叶斯网络的学习之后，我们就得到一个结点一父结点的有向无环图，对该图进行广度优先搜索，确定结点变量的遍历顺序，该顺序即为各结点变量与根结点之间因果关系的距离。把结点的顺序号平方的倒数作为该结点变量结构特征权重。

### 3.2 BN-KNN 算法的流程

输入： $n$  个结点变量的集合，一个结点可以有父结点的最大个数为  $u$ ，一个包含  $m$  个案例的训练样本集  $D$ 。

输出：每一个样本所属的类别。

1) set  $k$ ;

2) 训练贝叶斯网络结构;

For ( $i=1; i \leq n; i++$ ) // 启发式搜索

```

{
     $\pi_i = \phi$ ;
     $Pold = f(i, \pi_i)$ ;
    Flag = true;
    While (flag &&  $|\pi_i| < u$ )
    {
         $x = Pred(x_i) - \pi_i$ ; //  $x$  是结点  $x_i$  的所有
        先验结点集中去掉已经成为  $x_i$  父结点的集合。
         $Pnew = f(i, \pi_i \cup \{x\})$ ;
        If ( $Pnew > Pold$ )
        {
             $Pold = Pnew$ ;
             $\pi_i = \pi_i \cup \{x\}$ ;
        }
        else flag = false;
    }
}

```

printf("  $x_i$ 's parent node is:",  $x_i, \pi_i$ );

3) 广度优先搜索，得到访问结点变量的顺序，计算特征向量的权重  $w_d = \frac{1}{(\text{order}[i])^3}$ 。

4) 从训练集中找出最靠近待分类样本的  $k$  个样本  $x_1, x_2, \dots, x_k$ ;

5) 设， $F(x_q) = \arg \max_{v \in V} \sum_{i=1}^k w_i w_d \delta(v, f(x_i))$ ,  $w_i = \frac{1}{d(x_q, x_i)^2} = \frac{1}{\sqrt{\sum_{i,j=1}^m (x_i - x_j)^2}}$ , 当  $a=b$  时,  $\delta(a, b) = 1$ ; 否则,  $\delta(a, b) = 0$

6) 采用投票法，把与分类变量最近邻的样本数，作为待查询样本的预测值。比较待分类样本与各类的相似性程度，通过对候选类相似性程度的排序，然后给出一个阈值，就可以判定待分类样本所在的类别。

## 4 实验结果及分析

我们已经把本文提出的 BN-KNN 算法应用于老年型痴呆的早期诊断和防治系统中，利用 Matlab6.5 实现了该算法。为了使数据特征之间具有可比性，我们已经对数据集进行了标准化处理，并且用 KNN 算法对同一数据库进行了分类，实验结果如下：

图 1 是文[6]中的脑成像研究的数据产生的贝叶斯网络结构，其中网络中的各结点变量都是脑结构中的结构和功能变量，MCI (Mild Cognitive Impairment) 为轻度认知障碍功能变量，其余都为脑结构变量。我们采用该结构作为 KNN 分类的结构特征权重，并把实验结果与 KNN 算法的实验结果进行了比较，其比较的结果如图 2 和图 3 所示。

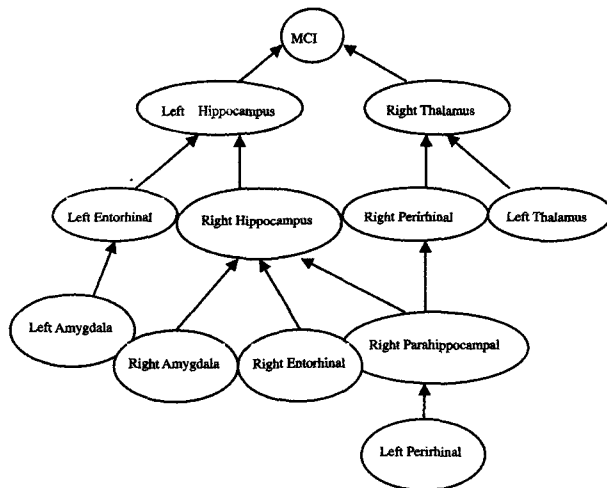


图 1 脑成像研究中的贝叶斯网络结构

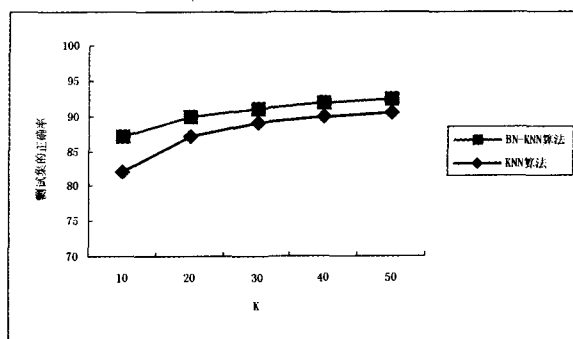


图 2 k 值对测试集正确率的影响

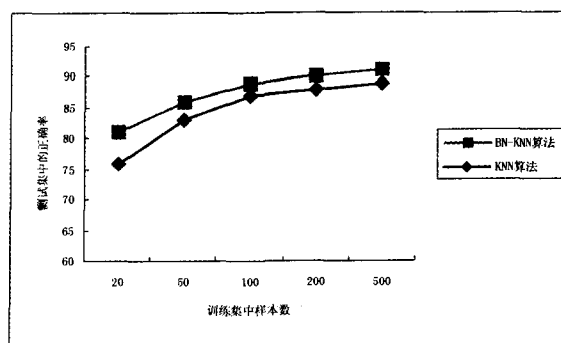


图 3 训练集中的样本数对测试集正确率的影响

表2 图3~5中各种融合方法所得结果图像的相似性度量值

图 像	本文提出的方法	Haar小波融合 <sup>[1]</sup>	形态学小波融合 <sup>[8]</sup>
图3 (Disk)	0.89096 ( $T=0.023441$ )	0.8316	0.8152
图4 (Lab)	0.88490 ( $T=0.068020$ )	0.8386	0.8430
图5 (Pepsi)	0.92620 ( $T=0.019089$ )	0.8997	0.9003

**结论** 本文提出了一种基于区域检测的多焦点图像融合方法,并使用小波算法、形态学运算和 GA 提取原始图像中的聚焦区域,最后将在原始图像中检测到的聚焦区域融合成为结果图像。这种方法使用类似于剪切-粘贴的操作将原始图像中的聚焦区域组合在一起,得到与参考图像相近的各处聚焦的结果图像。大量实验结果证明本文的方法明显优于 Haar 小波融合方法和形态学小波融合方法。特别是在原始图像没有完全配准的情况下,本文方法具有明显优势。

### 参 考 文 献

- 1 覃征,鲍复民,等. 数字图像融合. 西安交通大学
- 2 Pajares G, de la Cruz J M. A wavelet-based image fusion tutori-

(上接第 186 页)

从图 2 和图 3 的结果,可以得到以下结论:

①本算法与 KNN 算法相比,在相同条件下,测试集的正确率有明显提高。

②算法的性能随着训练样本数的增加而有所提高,并且,与近邻数  $k$  的选择也有直接的关系, $k$  的取值越大,正确率越高;反之,取值越小,正确率越低。当然,随着  $k$  的取值逐渐增大,算法的计算量也越来越大。下一步,我们将考虑用决策树来降低算法的计算量。

③根据具体应用问题的不同,可以适当调整待分类变量的结构特征在分类过程中所占的权重,使分类结果更加合理和优化。

**结论** 本文首先讨论了 KNN 和 BN 算法的基本思想及其改进算法,然后分析了目前两个算法存在的主要问题以及各自的优点,最后提出了基于贝叶斯网络结构学习的 KNN 算法 (BN-KNN)。实验结果表明,在相同条件下,与 KNN 算法相比,BN-KNN 算法具有更好的分类正确率。而且,随着与分类变量结果弱相关变量的增多,新算法的优势将越明显。另外,新算法可以根据用户所关注的分类目标自动地调整在

(上接第 215 页)

用不同的机器学习方法;二是选择不同语言处理粒度角度(字、词、语块、篇章小句)进行学习。目的都是有效消除单一方法的不足,尽可能增强整体处理的健壮性和可移植性。

**总结与展望** 指代消解是一项困难的任务。迄今为止,尚未有较好的全自动的指代消解技术和方法。可行的思路是由单一的基于规则方法向结合机器学习与统计方法的多策略技术方向发展,特别是基于小规模语料训练的无监督机器学习方法。借鉴国外的成功经验,将研究内容更好地同国际评测(如 ACE)结合是促进汉语指代消解长足发展的一条可行之策;国内几乎没有用于研究目的开放的汉语指代消解数据集,这方面可开展的研究工作也很多。目前指代消解在计算内容上逐渐由基于领域受限文本向领域无关自由文本方向发展;处理对象由普通文本向 Web 文本发展,因此笔者认为在汉语指代现象的形式化研究方面(如代词所指、省略),尽管已取得了一定成果,仍无法满足计算机自动化机处理,进一步挖掘和

- al. Pattern Recognition, 2004. 1855~1872
- 3 Piella G. A region-based multiresolution image fusion algorithm. In: ISIF Fusion 2002 Conference
- 4 Wen C Y, Chen J K. Multi-resolution image fusion technique and its application to forensic science. Forensic Science International, 2004, 140: 217~232
- 5 Li Min, Cai Wei, Tan Zheng. A region-based multi-sensor image fusion scheme using pulse-coupled neural network. Pattern Recognition Letters, 2006, 27: 1948~1956
- 6 Jiang Zhiguo, Han Dongbing, Chen Jin, et al. A wavelet based algorithm for multi-focus micro-image fusion. In: Proc. eedings of the Third International Conference on Image and Graphics, 2004
- 7 Yang X, Yang W, Pei J. Different focus points images fusion based on wavelet decomposition. In: Preceeding of Third International Conference on Information Fusion, vol 1. 2000. 3~8
- 8 De I, Chanda B. A simple and efficient algorithm for multifocus image fusion using morphological wavelets. Signal Processing, 2006, 86: 924~936
- 9 De I, Chanda B. Enhancing effective depth-of-field by image fusion using mathematical morphology. Image and Vision Computing, 2006, 24: 1278~1287
- 10 Li Shutao, Kwok J T, Wang Yaonan. Combination of images with diverse focuses using the spatial frequency. Information Fusion, 2001, 2: 169~176
- 11 Ardeshir A. Goshtasby. Fusion of multi-exposure images. Image and Vision Computing, 2005, 23: 611~618

分类过程中所占的权重,使分类结果更加合理和优化。

### 参 考 文 献

- 1 Teknomo K. What is K Nearest Neighbors Algorithm? [Z] <http://people.revoledu.com/kardi/tutorial/KNN/Contents.htm>
- 2 陈振洲,李磊,姚正安. 基于 SVM 的特征加权 KNN 算法[J]. 中山大学学报(自然科学版), 2005, 44(1): 17~20
- 3 D'Amato C, Malerba D, Esposito F, et al. Extending the K-Nearest Neighbour classification algorithm to symbolic objects [C]. Convegno Scientifico Intermedio SIS, 9-11 Giugno 2003, Università degli Studi di Napoli "Federico II"
- 4 Vincent P, Bengio Y. K-Local Hyperplane and Convex Distance Nearest Neighbor Algorithms [R]. [Technical Report]. <http://www.iro.umontreal.ca/~lisa/pointeurs/TR1197.pdf>, 2001
- 5 Cooper G F, Herskovits E. A Bayesian Method for the Induction of Probabilistic Networks from Data [J]. Machine Learning, 1992 (9): 309~347
- 6 Chen R, Herskovits E H. Network analysis of mild cognitive impairment [J]. NeuroImage, 2006, 29: 1252~1259
- 7 Yager R R. An extension of the naive Bayesian classifier [J]. Information Sciences, 2006, 176: 577~588
- 8 Yang T Y, Lee J C. Bayesian nearest-neighbor analysis via record value statistics and nonhomogeneous spatial Poisson processes [J]. Computational Statistics & Data Analysis, 2006
- 9 Frey B J, Jojic N. A Comparison of Algorithms for Inference and Learning in Probabilistic Graphical Models [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(9): 1392~1416

设计出对指代消解更有效的属性特征也是关键的突破口。

### 参 考 文 献

- 1 Lappin S, Leass H. An algorithm for pronominal anaphora resolution. Computational Linguistics, 1994, 20(4): 535~561
- 2 MaCarthy J F, Lehnert W G. Using decision trees for coreference resolution. In: Proceedings of 14th International Joint Conference on Artificial Intelligence Montreal, 1995. 1050~1055.
- 3 张威,周昌乐. 汉语语篇理解中元指代消解初步. 软件学报, 2002, 13(4): 732~738
- 4 王晓斌,周昌乐. 基于篇章表述理论的汉语人称代词的消解研究. 厦门大学学报(自然科学版), 2004, 43(1): 31~35
- 5 Ya-qian Z, Chang-ning H, et al. Transformation Based Chinese Entity Detection and Tracking. In: Proc. of the Second International Joint Conference on Natural Language Processing, 2005
- 6 钱伟,郭以昆,周雅倩,等. 基于最大熵模型的英文名词短语指代消解. 计算机研究与发展, 2003, 40(9): 1337~1343
- 7 Claire C, Wagstaff K. Noun phrase coreference as clustering. In: Proc. of the 1999 Joint SIGDAT Conference on Empirical Methods in NLP and Very Large Corpora
- 8 Chi-shing W. A Clustering Approach for Unsupervised Chinese Coreference Resolution. In: Proc. eedings of the 5th SIGHAN Workshop on Chinese Language Processing Sydney, 2006. 40~47