第34卷 第10期 计 算 机 工 程 2008年5月
Vol.34 No.10 Computer Engineering May 2008
·博士论文· 文章编号：1000—3428(2008)10—0028—02 文献标识码：A 中图分类号：TP312

权值自适应调整的多分类器融合算法

张冬慧[1]，孙 波[1]，王 鹏[2]，程显毅[2]
(1. 北京师范大学教育技术学院，北京 100875；2. 江苏大学计算机科学与通信工程学院，镇江 212013)

摘 要： 针对度量层输出的多分类器融合，该文提出一种基于Multi-agent思想的融合算法。该算法给出样本集在多分类器下的偏好判断矩阵概念，可以根据各个样本的具体情况自适应地为各分类器赋予权值。实验证明，该算法可得到比其他方法更低的分类错误率。
关键词： 分类器融合；偏好判断矩阵；权值

Multi-classifiers Fusion Algorithm of Adaptive Weight Adjustment

ZHANG Dong-hui[1], SUN Bo[1], WANG Peng[2], CHENG Xian-yi[2]
(1. School of Education Technology, Beijing Normal University, Beijing 100875;
2. School of Computer Science and Telecommunication Engineering, Jiangsu University, Zhenjiang 212013)

【Abstract】 Aiming at the problem of measurement level output, an information fusion algorithm based on multi-agent theory is presented. The concept of fancy judgment matrix is given, and an integration method of multiple classifiers based on adaptive weight adjusting is presented. Adaptive weight adjusting fusion method adaptively assigns weights to classifiers based on the sample. According to the experiments on standard database, this algorithm leads to less error than other methods and individual classifier. Experiments show that the algorithm is convergent.
【Key words】 classifier fusion; fancy judgment matrix; weight

1 概述

多分类器融合方法可以分为3类：(1)决策层融合，即输出为某个确定的类号；(2)排序层融合，即输出为给定测试样本属于各类的可能性的一个排序列表；(3)度量层融合，即输出为样本对于各类的度量值。度量层的融合比前2种方法拥有更丰富的信息量，相应的组合方法也更多[1]。

传统的度量层融合方法是先提出某一计算模型，然后根据训练样本估计模型中的参数，使该模型在统计意义上达到最优。这种方法存在一个明显的缺陷，它只是根据分类器的统计性能赋予该分类器一个权值，而没有考虑各个样本的具体情况，即不论各分类器关于某一样本的置信度是多少，这些分类器的输出结果总是按固定的权值被融合在一起。显然，这是不合理的。针对以上问题，本文提出一种基于Multi-Agent权值自适应调整的多分类器融合算法，该算法能够自适应地确定其权值，而且能够在Agent自身权值的作用下对各类别方案进行分类。

2 Multi-Agent融合系统

图1显示了Multi-Agent自适应权值的多分类器系统。

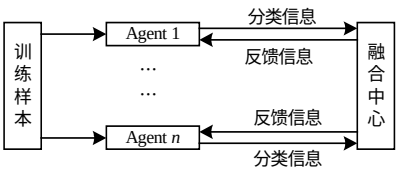

图1 Multi-Agent自适应权值的多分类器系统

每个Agent对训练样本进行分类，得到分类信息。融合中心将这些分类信息合成，得到类别的综合信息，然后将信息反馈给各个Agent，各个Agent综合自身信息和融合信息，对自身的权值进行修改，得到最终权值。

3 自适应权值的确定

3.1 识别性能矩阵

设$D=\{D_1, D_2, \cdots, D_J\}$是一组Agent分类器，训练样本$X$由$C_i, \forall i\in\wedge=\{1, 2, \cdots, M\}$类数据组成。当将一个特征向量 $\boldsymbol{x}\in R^n$输入到Agent D_j中时，Agent会判定其所属类别，即$D_j(\boldsymbol{x})\in$C。然后利用样本集统计每个Agent的性能，得到混淆矩阵：

$$\boldsymbol{CM}_j=(n_{km}^j)_{M\times M}\qquad k,m=1,2,\cdots,M$$

其中，$n_{il}^{(j)}$是D_j将第i类样本判别成第l类样本的数量，当$i=l$时，表示D_j正确判别第i类数据的数量，否则，表示D_j错误判别第i类数据的数量。对Agent D_j而言，识别结果为$l=D_j(x)$的样本总数是：$n_l^{(j)}=\sum_{i=1}^{M}n_{il}^{(j)}, l=1,2,\cdots,M$，在$D_j$的判定结果为$l$的条件下，样本来自$C_i$类的概率如下：

$$p_{il}^{(j)}=p(X\in C_i\,|\,D_j(X)=l)=\frac{n_{il}^{(j)}}{n_l^{(j)}}, l=1,2,\cdots,M \tag{1}$$

对于J个不同的Agent，可以得到J个混淆矩阵$\boldsymbol{CM}_1$，$\boldsymbol{CM}_2, \cdots, \boldsymbol{CM}_J$。当用这$J$个Agent对一个样本分类时，将得到$J$个分类结果，即$D_j(x)=l_j(j=1, 2, \cdots, J,\ \ l_j\in\wedge)$。结合式(1)得到Agent的识别性能矩阵：

$$\boldsymbol{PM}_j=(p_{km}^j)_{M\times M}\qquad k,m=1,2,\cdots,M$$

对于样本X，L个Agent的判定类别分别是$B_1, B_2, \cdots, B_l$ ($B_i\in C, i=\{1, 2, \cdots, M\}$)，根据Agent的识别性能矩阵，得到如下

基金项目： 国家自然科学基金资助项目(60702056)
作者简介： 张冬慧(1969－)，女，博士研究生，主研方向：模式识别，计算机应用，自然语言处理；孙 波，教授、博士生导师；王 鹏，硕士研究生；程显毅，教授、博士生导师
收稿日期： 2007-05-25 **E-mail：** zhdh1997@163.com

矩阵：$\boldsymbol{DP}=(p_{B_jk}^{(j)})_{L\times M}$　$k=1,2,\cdots,M$, $j=1,2,\cdots,L$，$p_{B_j}^{(j)}$ 表示第j个Agent对第B_j类数据的识别性能。

3.2 偏好判断矩阵

定义 1 设有论域 $S=\{S_i \mid i\in I, i=1,2,\cdots,n\}$ 为决策问题的方案集，二元比较矩阵$\boldsymbol{A}=(a_{ij})_{n\times n}$为方案直积$S\times S$上的模糊子集，$0\le a_{ij}\le 1$，表示方案$S_i$优于$S_j$的程度，若满足下列性质：

$a_{ij}=0.5$, $\forall i\in i$

$a_{ij}+a_{ji}=1$, $\forall i,j\in I$, $i\ne j$

则称 $\boldsymbol{A}$ 是偏好判断矩阵。

通过使用训练样本，得到 Agent 识别性能矩阵，也就可以得到偏好判断矩阵。例如，设 3 个 Agent 对由 2 个类别组成的数据进行分类，得到的 $\boldsymbol{DP}$ 矩阵和偏好判断矩阵如下：

$$\boldsymbol{DP}=\begin{bmatrix}0.85 & 0.15\\ 0.70 & 0.30\\ 0.50 & 0.50\end{bmatrix}$$

$$\boldsymbol{A}_1=\begin{bmatrix}0.50 & 0.85\\ 0.15 & 0.50\end{bmatrix},\ \boldsymbol{A}_2=\begin{bmatrix}0.50 & 0.70\\ 0.30 & 0.50\end{bmatrix},\ \boldsymbol{A}_3=\begin{bmatrix}0.50 & 0.50\\ 0.50 & 0.50\end{bmatrix}$$

3.3 Multi-Agent 系统权值的确定

利用每个 Agent 所给出的偏好矩阵与群体的偏好矩阵间的距离度量个体与群体之间的“相近”程度。利用矩阵范数 $\left\|\boldsymbol{A}_K-\bar{\boldsymbol{A}}\right\|=\sum_{i=1}^{n}\sum_{j=1}^{n}\left|a_{ij}^{(k)}-\bar{a}_{ij}^{(k)}\right|$ 表示 $\bar{\boldsymbol{A}}$ 与 $\boldsymbol{A}_K$ 间的距离 $d(\boldsymbol{A}_K,\boldsymbol{A})$，$K\in M$，简记为 $d^{(k)}$，根据 $d(\boldsymbol{A}_K,\boldsymbol{A})$ 度量 $\boldsymbol{A}_K$ 与 $\boldsymbol{A}$ 的相容性程度，Agent k 与群体判断越近说明 Agent k 越能代表所有 Agent 的分类，应该对其赋予越大的权值。在无法确认各 Agent 权值的情况下，假设各 Agent 具有相同的权值，即 $w_k=\frac{1}{m}, k\in M$，于是得到群体综合判断矩阵 $\boldsymbol{A}=(\bar{a}_{ij})_{n\times n}$，其中，

$$\bar{a}_{ij}=\frac{1}{m}\sum_{l=1}^{m}a_{ij}^{(l)}, i,j\in I \tag{2}$$

3.4 自适应权值的多分类器融合算法

假设矩阵 $\boldsymbol{A}_K$ 是第 k 个 Agent 对分类类别的偏好判断矩阵，$\boldsymbol{\lambda}_{(K)}=(\lambda_{1(K)},\lambda_{2(K)},\cdots,\lambda_{n(K)})^{\mathrm{T}}$ 为第 k 个 Agent 对类别给出的权重向量，β_k 为第 K 个 Agent 判断矩阵的可信度权重。用

$$e_{ij}=\cos\theta_{ij}=\frac{(\gamma^{(i)},\gamma^{(j)})}{\left|\gamma^{(i)}\right|\cdot\left|\gamma^{(j)}\right|}$$

来表示判断矩阵 $\boldsymbol{A}^{(i)}$ 和判断矩阵 $\boldsymbol{A}^{(j)}$ 之间的一致性程度。显然，e_{ij} 越大则矩阵 $\boldsymbol{A}^{(i)}$ 和判断矩阵 $\boldsymbol{A}^{(j)}$ 间的一致性程度越高。

考虑Agent的权值问题，定义判断矩阵 $\boldsymbol{A}^{(i)}$ 的平均一致性程度 e_i 为：$e_i=\frac{1}{m-1}\cdot\sum_{j=1,j\ne i}^{m}w_d^{\{j\}}\cdot e_{ij}$，$e_i$ 表示第i个判断矩阵和所有其他判断矩阵之间一致性程度的加权平均值。然后将所有判断矩阵的平均一致性程度 e_i 进行归一化，就得到各个判断矩阵的相对一致性程度 e_i^*。在归一化的过程中，同样要考虑Agent自身权值的存在，归一化公式[2]可表示为

$$e_i^*=\left(w_d^{\{i\}}\cdot e_i\right)/\sum_{j=1}^{m}w_d^{(j)}\cdot e_j \tag{3}$$

第 i 个判断矩阵的相对一致性程度越高，说明它与其他所有判断矩阵的一致性程度越高，则它也就越能代表大多数 Agent 的分类意见。所以可以选取相对一致性程度值 e_i^* 作为第 i 个判断矩阵的可信度权值 β_i，$\beta_i=e_i^*$，合并的算式为

$$\lambda_i=\prod_{k=1}^{m}[\lambda_i^{(k)}]^{\beta_k}, i=1,2,\cdots,n \tag{4}$$

进行归一化处理[3]：$\lambda_i^*=\lambda_i/\sum_{j=1}^{n}\lambda_j$ (5)

其中，$\boldsymbol{\lambda}_i^*$ 是 m 个 Agent 判断矩阵合并后得出的第 i 个类别的度量值。融合算法如下：

Step 1 对融合训练集上各样本的决策结果进行统计，得到识别性能矩阵 $\boldsymbol{DP}$ 和偏好判断矩阵 $\boldsymbol{A}$；

Step 2 由式(2)得到群体综合判断矩阵 $\bar{\boldsymbol{A}}$，计算 $d^{(k)}$；

Step 3 对得到的权值进行归一化处理，得到最终权值；

Step 4 设置一个阈值 label，如果某 Agent 的权值小于该阈值，则该 Agent 权值为 0，即在这次融合过程中不考虑该 Agent，转向 Step 3，否则转 Step 5；

Step 5 用变量 *value* 表示第 i 个判断矩阵的相对一致性程度 e_i^*，若 *value* 大于某阈值 t，则表示第 i 个 Agent 可以代表大多数 Agent 的分类判断，转 Step 7，否则转 Step 6；

Step 6 根据式(4)、式(5)得出各个类别的度量值；

Step 7 得出最终的分类器的融合结果。

算法根据多 Agent 各个判断矩阵之间的一致性，得到了分类结果的集成判决，并自适应地选择了分类器组合及分类器权重。在 Agent 行为分析中，对于待定样本，如果第 i 个 Agent 判断矩阵的相对一致性程度大于某阈值 t，则表示第 i 个 Agent 可以代表大多数 Agent 的分类判断，此时可以直接输出该 Agent 类别作为样本的类别，而无须进行后续的分析，从而减少计算时间。

4 实验结果及分析

下面通过实验来验证本文融合算法的性能，并将该算法的融合结果同单个分类器及加权投票法、加法规则法、K 近邻法、Decision Templates 法的结果作比较。实验选择 UCI 机器学习数据库(UCI machine learning repository)。从这个数据库中选择 Iris 数据集，50 个样本作为训练样本，100 个样本作为测试样本。归纳后的数据集如下：

数据	类数	样本数	特征数	特征分割形式
Iris	3	150	4	[2, 1, 1]

4.1 实验方法

本文采用特征分割的方法来获得分类器，Iris 数据集的 4 维原始特征按顺序被分割为 3 个长度分别为 2 维、1 维和 1 维的特征子集，这些特征子集分别作为分类器的输入。用 Matlab 的模式识别标准工具箱 prtools 进行分类，共得到随机选择的 11 个分类器，各分类器在训练集和测试集上得到的错误率如下：

分类器号	1	2	3	4	5	6
识别率	77%	74%	70%	71%	82%	75%
分类器号	7	8	9	10	11	
识别率	72%	74%	81%	78%	82%	

把分割后的特征子集分别输入各自的分类器，再对分类器的输出进行组合，获得实验结果。在实验中，对错误率的估计使用了 10 次交叉验证方法[4]，m次交叉验证是指将总的n个样本集划分为m个不相交的数据子集，每个子集都有n/m个样本。每次从所有数据中取出一个子集i，以剩下的m-1 个子集来进行训练，然后以i作验证，计算出错误率。重复计算m次以后，估计出的推广错误率是m个错误率的平均值。

4.2 实验结果

图 2 为本文融合算法对 Iris 数据分类的结果，图 3 为本

(下转第 32 页)