

非平衡样本分类的集成迁移学习算法

于重重^{1,2}, 田蕊¹, 谭励¹, 涂序彦²

(1.北京工商大学计算机与信息工程学院,北京 100048; 2.北京科技大学信息计算机与通信工程学院,北京 100089)

摘要: 针对冗余数据量大且正负样本不平衡的辅助训练数据,提出了一种改进集成迁移学习算法,利用这些辅助训练数据迁移帮助目标数据进行分类.新的样本初始权重分配及调整策略,突出了对负样本的识别能力.通过动态调整辅助训练集,根据设定好的权重阈值下限适时地淘汰冗余数据,降低了冗余数据对分类器性能的影响,提升了迁移学习对非平衡样本的学习能力.本文利用桥梁实际监测数据进行的实验表明了该算法较 TrAdaboost 算法的有效性.

关键词: 迁移学习; 分类器集成; 冗余数据淘汰; 权重分配

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2012) 07-1358-06

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2012.07.012

Integrated Transfer Learning Algorithmic for Unbalanced Samples Classification

YU Chong-chong^{1,2}, TIAN Rui¹, TAN Li¹, TU Xu-yan²

(1. School of Computer & Information Engineering, Beijing Technology and Business University, Beijing 100048, China;

2. School of Computer and Communication Engineering, University of Science and Technology, Beijing 100083, China)

Abstract: According to the auxiliary training data with large redundancy and imbalance between positive and negative samples, an improved integrated transfer learning algorithmic -The Unbalanced Integrated Transfer Learning Algorithmic is proposed. Applied these auxiliary training data to transfer and help classifying on target data. New sample initialization and regulation weight method highlighted negative sample identification ability. Through dynamic adjusting auxiliary training set, eliminated redundant data according to the weight lower threshold, reduced their influence on the classifier and improved the transfer learning's performance. Experimental results on the actual bridge monitoring data show that this algorithmic is advanced than TrAdaboost.

Key words: transfer learning; classifier integration; redundant data elimination; weight distribution

1 引言

迁移学习是近年来机器学习领域研究的热门,它针对新任务中标记数据量小的特点,提出有效地利用已过时的数据迁移应用到新的任务中:虽然大量过时数据与待解决问题领域有所差异,但是其中肯定会存在一些对新的分类问题有所帮助的数据.为了能够找到这些有用的数据,利用少量已被分类的新数据,来挖掘旧数据中的有价值信息.最后根据两部分数据中所有的有用信息来训练一个较高效的分类模型,实现旧数据到新数据的知识迁移^[1~5].

目前,针对不同迁移学习任务已经提出多种解决方法, Rigutini 等人^[1]针对跨特征空间的迁移学习问题,提

出了一种基于 EM 的跨语言文本分类算法. Raina 等人^[6]提出了自学习的概念,让机器从一些无标注的数据中学习出需要的知识,来帮助目标学习任务. Q Yang 等人提出将朴素贝叶斯分类器 (Naive Bayes Classifier) 推广成一个支持跨领域文本分类的分类器^[7], 实现了不同领域文本之间知识的迁移. Dai 等人^[4]提出将集成学习应用到迁移学习中,通过 boosting 技术将弱学习算法“提升”为强学习的算法 TrAdaboost, 该算法将迁移辅助数据和目标数据这两部分数据集直接组合在一起, 组成一个混合数据集作为训练集, 然后在这个数据集上利用 Adaboost 算法训练分类模型.

将集成学习算法应用到迁移学习中,可以在不改变弱分类器分类精度的情况下,通过集成将弱学习算法

“提升”为强学习算法^[6],从而有效提升迁移学习效果,然而该方法也存在一些问题:首先,TrAdaboost 算法适用于解决基于对称的二分类问题,正负数据同等看待.然而,在现实世界中表征两类不同类别的样本分布上有可能是极度不平衡的,并且重要性也存在很大的差异.另外,辅助数据中往往存在大量冗余数据,这些数据可能与目标数据集很不相似,它们的存在不但会影响模型的训练速度,还会导致分类精度的下降.本文根据这些问题提出改进迁移学习算法—非平衡集成迁移学习算法(UBITLA),通过优化样本权重分配及调整策略,提高数据量小、信息量大的一类样本(负样本)的贡献率;并在训练过程中动态剔除“不相关”数据,根据设定好的样本阈值下限,淘汰权重值过小的那部分数据,经过 N 轮的迭代训练,辅助训练数据将不断趋于优化.

本文将迁移学习引入到桥梁结构健康监测领域,在该领域中,对于长期监测得到的数据如何有效利用,将已有的大量数据帮助新桥建立评估模型是迁移学习可以解决的问题.另外对于桥梁实际监测数据,表征结构健康的正数据和表征结构损伤的负数据在分布上是极度不平衡的,负样本数量远远小于正样本,但却具有很高的的重要性,这使得桥梁结构损伤判定可以看作是一个非对称的二分类问题,本文提出的非平衡集成迁移学习算法适用于解决此类问题.

2 算法理论基础与改进

2.1 TrAdaboost 算法流程

TrAdaboost 算法是 Dai 等人^[8]提出基于实例的迁移学习算法,其基本思想是利用 boosting 技术来过滤掉辅助数据中那些与目标数据最不像的数据,通过自动调整权重的机制找到最有用的样本.该算法中,AdaBoost^[9]被用在了目标数据中,以保证分类模型在目标数据上的准确性.

TrAdaboost 算法描述如下:

输入:迁移辅助数据集 A 与目标数据集 O 根据实际数据量按比例混合成训练数据集 $C = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$, 即 C 中前 n 个样本为 A 中数据,剩下的 m 个样本为 O 中的数据.其中 $x_i \in X$ 是样本的特征矢量, $y_i \in \{0, +1\}$ 是样本的标签, $i = 1, 2, \dots, N$. 预定的迭代次数为 T .

初始化样本权值:

$$\omega_i^1 = \frac{1}{n} \text{ 或 } \frac{1}{m}, i = 1, 2, 3, \dots, N \quad (1)$$

DO FOR $t = 1, \dots, T$

用弱学习算法 P, 训练得到弱分类器:

$$h_t(x): x \rightarrow \{0, +1\}$$

计算 $h_t(x)$ 的错误率:

$$\epsilon_t = \sum_{i=n+1}^{n+m} \frac{\omega_i^t |h_t(x_i) - y_i|}{\sum_{i=n+1}^{n+m} \omega_i^t} \quad (2)$$

$$\beta_t = \frac{1}{1 + \sqrt{2 \ln(n/T)}}$$

$$\beta_t = \frac{\epsilon_t}{1 - \epsilon_t}, \text{ 其中 } \epsilon_t < \frac{1}{2}$$

调整样本权值:

$$\omega_i^{t+1} = \begin{cases} \omega_i^t \beta_t^{1 - |h_t(x_i) - c(x_i)|}, & 1 \leq i \leq n \\ \omega_i^t \beta_t^{-1 - |h_t(x_i) - c(x_i)|}, & n+1 \leq i \leq n+m \end{cases} \quad (3)$$

输出: 最终的分类器集

$$H(x) = \begin{cases} 1, & \sum_{t=\frac{T}{2}}^T \ln\left(\frac{1}{\beta_t}\right) h_t(x) \geq \frac{1}{2} \sum_{t=\frac{T}{2}}^T \ln\left(\frac{1}{\beta_t}\right) \\ 0, & \text{其他} \end{cases} \quad (4)$$

2.2 冗余数据动态剔除

上述 TrAdaboost 算法在整个训练过程中一直都在使用没有改变过的训练集 C , 然而在很多迁移学习的情况下,用于迁移的辅助训练数据量往往很大,而且其中不少数据与目标数据集很不相似,如果在训练过程中一直保留这些冗余数据,一定会影响分类器效率和精度,甚至可能引起“负迁移”现象^[10]. 本文提出采用了一种循环式的冗余数据淘汰方法来解决这个问题.由于每次迭代训练后,分类器会对迁移辅助数据集进行测试和样本权值的修改:冗余的样本的权值会随之减小,有用样本的权值会相应增大,并将样本权重与设定好的阈值下限相比较,低于此下限的部分样本会被剔除.实现了训练数据集的动态优化,提高了训练效率和精度.具体调整流程如下:

冗余数据动态剔除算法描述:

输入: 目标数据集 O 与迁移辅助数据集 A 根据实际数据量按比例抽取混合训练数据集 C . A 中样本初始权值 ω_1 , ω_{\min} 为辅助训练样本权重阈值下限 ($0 \leq \omega_{\min} \leq \omega_1$), A_{\min} 为训练数据中辅助数据量的最小比例,预定的迭代次数为 T .

Do For $t = 1, 2, \dots, T$

① 用弱学习算法 P, 训练得到弱分类器: $h_t(x): x \rightarrow \{0, +1\}$

② 返回到在迁移辅助数据集上利用上一步得到的集成分类器 h_t 进行测试,对于 A 上每个样本,如果 h_t 预测的结果与其真实类标签相同,则加大其对应的权值,否则减小其权值^[11].

③ 判断 U 中所有样本权值 ω_i 是否小于 ω_{\min} , 将权值小于 ω_{\min} 的样本剔除,得到裁剪后的新的迁移辅助数据集 A ; 如果经过裁剪的 A 剩下的样本占 C 总数的比例小于 A_{\min} , 停止循环.

输出: $C = O + A$

2.3 优化样本权重分配及调整

由于在很多领域中大量存在正负样本数量不平衡的现象,如果将初始权重平均分配给所有样本,经过集成训练后,对占多数的那一类样本分类能力强的分类器产生的概率大,这类样本的权重经过调整后会相应增大,而样本权重与被抽中概率是成正比的,这样容易导致被抽中的全部是一类样本,从而造成分类器的训练失败^[12,13].为了确保训练成功,需要提高占少数的一类样本(这里指负样本)的贡献率,因此对 TrAdaboost 算法做出如下改进:

①将负样本赋予较大的初始权重,保证了这类样本被抽中的概率更大.并在每轮训练后,通过给被分错的负样本增加一个衰减项,即 $dr \times \omega_i^t (0 \leq dr \leq 1)$,使被分错的负样本权重调整具有记忆功能,保证其不致迅速变小.

②在抽取训练样本子集后,判断该子集中是否含有两种类别样本,如果只包含一类样本,则直接在缺少的那类样本集中抽取部分样本放入训练子集,确保训练子集中含有两类样本.

3 UBITLA 算法思想及流程

本文提出的 UBITLA 算法是充分考虑了现实世界中存在大量正负样本不平衡现象,运用集成学习方法,结合上述改进策略解决非对称二分类的迁移学习问题.

3.1 UBITLA 算法思想

UBITLA 的基本思想是将数据向量的每维特征属性分别应用弱学习算法,在初始化时,将正负样本赋予不同的权重,保证占总样本比例小但信息量大的负样本初始权重较大.每轮训练过程中按比例抽取部分样本作为训练子集进行训练,训练结束后,从若干个简单分类器中选择误差最小的那个,作为一个弱分类器 h_t ,并按照冗余数据动态剔除算法调整训练数据集.这样, t 轮

迭代后就可以得到一个弱分类器序列 (h_1, h_2, \dots, h_t) ,其中,分类效果比较好的权重较大.最终的分类函数 $f(x)$ 采用一种有权重的投票方式产生,即将多个弱分类器通过一定的方法叠加(boost)起来组合成一个强分类器.算法流程如图 1 所示.

3.2 UBITLA 算法数学描述

基于以上思想,设计的非平衡集成迁移学习算法 UBITLA 流程如下:

输入:迁移辅助数据集 A 与目标数据集 O 按比例混合成训练数据集 $C = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$,即 C 中前 n 个样本为 A 中数据, C 剩下的 m 个样本为 O 中的数据.预定的迭代次数为 T .其中 $x_i \in X$ 是样本的特征属性向量,维数为 q , $y_i \in \{0, +1\}$ 是样本的标签, $i = 1, 2, \dots, N$.训练子集大小为 $N/2$.

初始化样本权重:

$$\omega_{i_0}^1 = \begin{cases} 1/d & A \text{ 中正样本} \\ 1/n-d & A \text{ 中负样本} \\ 1/l & O \text{ 中正样本} \\ 1/m-l & O \text{ 中负样本} \end{cases} \quad (5)$$

其中 $i = 1, 2, \dots, N$. d, l 分别为 A 和 O 中正样本数量.

归一化样本权重:

$$\omega_i^1 = \frac{\omega_{i_0}^1}{\sum_{i=1}^N \omega_{i_0}^1} \quad (6)$$

Do For $t = 1, 2, \dots, T$

①抽取训练子集 D (其中 $D = 1/2C$)

②判断训练子集中是否含有两类样本.如果只含一类样本,则直接在另一类中抽取并填入训练子集.

③在训练子集上,用弱学习算法 P ,训练基分类器 $h_j^t, j = 1, 2, \dots, q$,得到弱分类器

$$h_t = \sum_{j=1}^q h_j^t \quad (7)$$

④计算弱分类器 h_t 在 V 上的训练错误率:

$$\epsilon_t = \frac{\sum_{i=n+1}^{n+m} \omega_i^t |h_t(x_i) - y_i|}{\sum_{i=n+1}^{n+m} \omega_i^t} \quad (8)$$

$$\text{令 } \beta = \frac{1}{1 + \sqrt{2 \ln \frac{n}{T}}}, \beta_t = \frac{\epsilon_t}{1 - \epsilon_t}$$

其中 $\epsilon_t < \frac{1}{2}$

⑤调整样本权重:

if $y_i = 0$ and $h_t(x) \neq y_i$, 其中 $1 \leq i \leq n$

$$\omega_i^{t+1} = \omega_i^t \beta^{|h_t(x_i) - c(x_i)|} + dr \times \omega_i^t (0 \leq dr \leq 1) \quad (9)$$

else

$$\omega_i^{t+1} = \begin{cases} \omega_i^t \beta^{|h_t(x_i) - c(x_i)|}, & 1 \leq i \leq n \\ \omega_i^t \beta_t^{|h_t(x_i) - c(x_i)|}, & n+1 \leq i \leq m+n \end{cases} \quad (10)$$

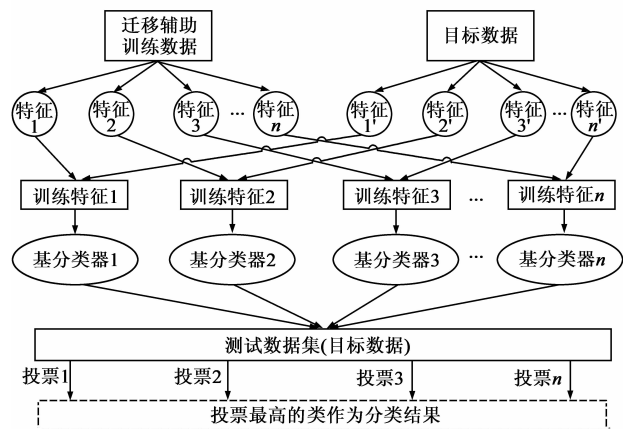


图1 UBITLA算法步骤框图

⑥应用冗余数据动态剔除算法

输出:最终的分类器集:

$$H(x) = \begin{cases} 1, \sum_{t=\frac{T}{2}}^T \ln(\frac{1}{\beta_t}) h_t(x) \geq \frac{1}{2} \sum_{t=\frac{T}{2}}^T \ln(\frac{1}{\beta_t}) \\ 0, \text{其他} \end{cases} \quad (11)$$

4 桥梁结构健康监测领域迁移学习可行性分析

采用迁移学习判定桥梁结构异常情况的可行性是基于对实际监测数据的观察. 尽管不同时间或空间的监测数据形成了不同的分布, 但随着交通高峰时段到低峰时段的交替, 不同监测点数据的变化规律会呈现一定的相似性, 尤其相同构建上的测点在同时段的监测数据的概率分布更加相似. 这些相似的规律往往蕴含着可以被迁移利用的信息, 从而使得在不同监测点相似分布数据上的迁移变得可行.

5 实验数据及结果分析

5.1 实验数据

本实验采用一座已有两年监测历史的跨海大桥监测数据集 (DataS1) 和一座新建公路桥监测数据集 (DataS2) 作为研究对象, 从两座桥的实际监测应变数据中按照早, 中, 晚高峰时段和上下午低峰时段分别抽取一定比例的数据分别作为迁移辅助训练数据集和目标数据集, 数据组成成分如表 1 所示. 其中正负样本比例为 5:1, 分别采用两座桥的桥板上 14 个关键监测点静应变数据, 作为 14 维输入数据. 输出数据: 1 代表正常, 0 代表损伤. 在每轮的训练过程中随机抽取 1/2 的数据作为训练数据子集, 同样随机抽取一部分目标数据作为测试数据子集.

表 1 训练数据组成成分

数据描述	早高峰	上午低峰	午高峰	下午低峰	晚高峰
time	7:00-9:00	9:00-11:00	11:00-13:00	13:00-17:00	17:00-19:00
Data-Ratio	30%	5%	30%	5%	30%
Data-S1	1500	250	1500	250	1500
Data-S2	300	50	300	50	300

5.2 实验结果分析

5.2.1 UBITLEA 算法在训练集及测试集上的测试

图 2 是 UBITLEA 算法在训练集和测试集上的分类相对误差, 从图 2 可知, 在输入数据量固定的情况下, 随着迭代次数的增加, UBITLEA 算法在训练集和测试集上的总体分类误差呈下降趋势, 在两个数据集上的分类错误率接近, 说明迁移数据辅助建立分类模型的有效性. 另外, UBITLEA 算法利用较少的分类器 (迭代次数) 错误率较高, 随着分类器数量的增加, 错误率单调降低, 从

图 2 可以得出, 集成学习算法在不改变单个分类器精度的情况下, 可以通过弱分类器个数的增加提高分类器性能.

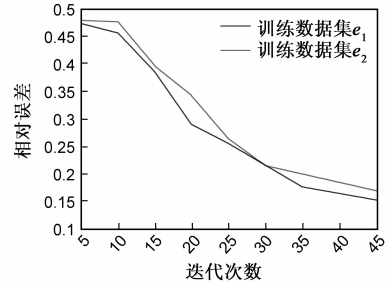


图 2 UBITLEA 算法在训练数据及测试数据上的错误率

5.2.2 UBITLEA 与 TrAdaboost 算法的比较

考虑到训练样本中正负样本的不平衡, 算法具有一定的随机性, 同时做了多次试验, 设定输入为测试数据集, 迭代次数 50. TrAdaboost 算法在试验中出现了多次失败现象, 表 2 是 10 次训练成功的 TrAdaboost 算法与 UBITLEA 算法测试结果的对照表. 从表中可以看出, TrAdaboost 算法只有 2 次达到低错误率, 而 UBITLEA 算法则几乎全部达到较低错误率. 另外, 在训练数据集上按 10% 的交叉变异后的测试结果表明: UBITLEA 算法相比传统的 TrAdaboost 算法具有更高的容错率及稳定性.

表 2 TrAdaboost 算法与 UBITLEA 算法的测试结果

训练次数	传统 Adaboost 算法		UBITLEA 算法	
	相对误差 e	交叉变异后相对误差 e'	相对误差 e	交叉变异后相对误差 e'
1	0.336	0.391	0.183	0.192
2	0.258	0.265	0.185	0.112
3	0.232	0.247	0.196	0.102
4	0.265	0.331	0.165	0.197
5	0.286	0.312	0.189	0.196
6	0.182	0.234	0.129	0.132
7	0.401	0.389	0.216	0.217
8	0.302	0.320	0.174	0.198
9	0.175	0.179	0.189	0.205
10	0.378	0.396	0.155	0.212

图 3, 图 4 分别是 TrAdaboost 算法及 UBITLEA 算法在迭代次数固定为 50 次, 训练子集数据量从 500 逐渐增加到 3000 的实验结果. 从图 3 可知, TrAdaboost 算法的分类精度不是单调变化的, 由于迭代过程中迁移训练集中冗余数据的始终存在, 会有分类精度低的弱分类器降低集成分类器精度, 随着训练样本的增加并不能提高最终分类器的性能. 从图 4 可以看出, UBITLEA 算法的分类错误率曲线是单调下降的, 说明通过优化辅助训练样本集可以提高最终分类器精度. 训练数据集按 10% 交叉变异后单调性也较好, 其性能明显优于 TrAdaboost 算法.

一方面, UBITLEA 算法区分正样本和负样本的初始

化权重及调整策略,突出了分类器对负样本的识别能力,并保证了每轮抽样训练子集中含有两类样本,避免了由于正负样本不平衡造成的训练失败.另一方面,UBITLA 算法按照设定好的权重阈值下限淘汰权重小的部分样本,减小了迁移数据集中冗余数据对分类模型的影响,随着训练数据的增加,最终分类器的分类能力不断增强.实验结果表明,在不平衡样本分类中,UBITLA 算法较 TrAdaboost 算法具有一定的优势.

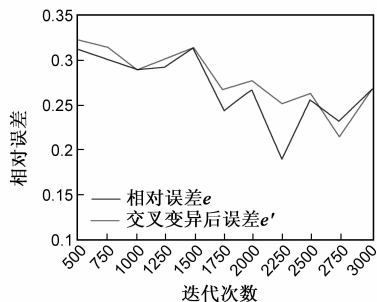


图3 TrAdaboost随训练数据增加相对误差

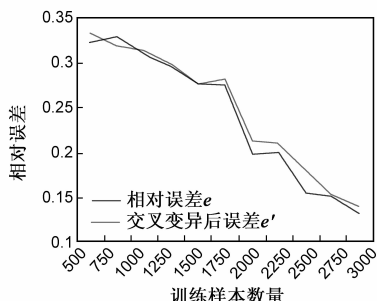


图4 UBITLA随训练数据增加相对误差

6 总结

迁移学习是近来机器学习领域研究的新热门,这种在给定大量已有标记数据以及少量新标记数据的情况下,是一个重要例子.本文针对冗余数据量大,正负样本不平衡的特点,合理的优化了正负样本的初始化权重及调整策略,提高了对迁移学习过程起重要作用的样本的贡献率.并提出冗余数据动态剔除算法,不断优化辅助训练数据集,从而达到效率和精度的双赢,提升了迁移学习效果,最后将该算法应用到桥梁结构健康监测领域,在实际监测数据上的实验表明该算法较 TrAdaboost 算法表现出更高的分类精度,通过交叉验证具有更好的稳定性.

参考文献

[1] L Rigutini, M Maggini, B Liu. An EM based training algorithm for cross-language text categorization [A]. IEEE International Conference on Web Intelligence [C]. University of Technology of Compiegne, France, 2005. 282 - 287.

[2] W Dai, Q Yang, G-R Xue, Y Yu. Boosting for transfer learning

[A]. Proceedings of the Twenty-Fourth International Conference on Machine Learning [C]. Corvallis, Oregon, USA, 2007. 193 - 200.

[3] W Dai, Y Chen, G-R Xue, Q Yang, Y Yu. Translated learning: Transfer learning across different feature spaces [A]. Advances in Neural Information Processing Systems 21 [C]. Vancouver, British Columbia, Canada, 2009. 786 - 791.

[4] Y Liu, P Stone. Value-function-based transfer for reinforcement learning using structure map-ping [A]. Proceedings of the Twenty-First National Conference on Artificial Intelligence [C]. Boston, Massachusetts, 2006. 877 - 882.

[5] Sinno J Pan, Qiang Yang. A survey on transfer learning [A]. IEEE Transactions on Knowledge and Data Engineering [C]. Los Alamitos, CA, USA, 2009. 556 - 562.

[6] R Raina, A Battle, H Lee, B Packer, A Y Ng. Self-taught learning: Transfer learning from unlabeled data [A]. The Twenty-fourth International Conference on Machine Learning [C]. Corvallis, Oregon, USA, 2007. 759 - 766.

[7] W Dai, G.-R. Xue, Q Yang, Y Yu. Transferring naive bayes classifiers for text classification [A]. The Twenty-Second National Conference on Artificial Intelligence [C]. Corvallis, Oregon, USA, 2007. 540 - 545.

[8] 戴文渊. 基于实例和特征的迁移学习算法研究 [D]. 上海: 上海交通大学, 2010. 6.

Dai Wen-yuan. Transfer learning algorithm based on the examples and characteristics [D]. Shanghai: Shanghai Jiaotong University, 2010. 6. (in Chinese)

[9] 张彦峰, 何佩琨. 一种改进的 Adaboost 算法—M - Asy Adaboost [J]. 北京理工大学学报, 2011, 31(1): 65 - 68.

Zhang Yan-feng, He Pei-kun. An improved Adaboost algorithm—M - Asy Adaboost [J]. Beijing University of Science and Technology Journal, 2011, 31(1): 65 - 68. (in Chinese)

[10] 付忠良, 赵向辉, 苗青, 姚宇. Adaboost 算法推广——组集成学习算法 [J]. 四川大学学报, 2010, 42(6): 91 - 98.

Fu Zhong-liang, Zhao Xixang-hui, Miao Qing, Yao Yu. Adaboost algorithm promotion- A set of integrated learning algorithm [J]. Sichuan University Journal, 2010, 42(6): 91 - 98. (in Chinese)

[11] Yoav Freund, Robert E. Schapire. A decision- -theoretic generalization of on-line learning and an application to boosting [J]. Journal of Computer and System Sciences, 1997, 55(1): 119 - 139.

[12] 欧阳震铮, 罗建书, 胡东敏, 吴泉源. 一种不平衡数据流集成分类模型 [J]. 电子学报, 2010, 38(1): 184 - 189.

OuYang Zhen-zheng, Luo Jian-shu, Hu Dong-min, Wu Quan-yuan. An unbalanced data flow integrated classification model [J]. Acta Electronica Sinica, 2010, 38(1): 184 - 189. (in Chinese)

[13] 刘伟, 张华祥. 数据集动态重构的集成迁移学习 [J]. 计

计算机工程与应用,2010,46(12):126-128.

Liu Wei, Zhang Hua-xiang. Integrated transfer learning based on data sets dynamic reconfiguration[J]. The Computer Engi-

neering and Application, 2010, 46(12): 126-128. (in Chinese)

作者简介



于重重 女,1971 年生于重庆,北京工商大学计算机与信息工程学院教授,硕士研究生导师,研究领域:智能信息处理、计算机网络.

E-mail: chongzhy@vip.sina.com



谭励 女,1980 年生于广西南宁,北京工商大学计算机与信息工程学院副教授,研究领域:机器学习、多模态信息处理.



田蕊 女,1986 年生于北京,现为北京工商大学计算机与信息工程学院在读硕士研究生,研究领域:机器学习、人工智能.



涂序彦 男,1935 年生于江西南昌,中国人工智能学会荣誉理事长、北京科技大学信息工程学院特聘教授、北京科技大学计算机与系统科学研究所所长,研究领域:人工智能、专家系统、智能控制.