

# 一种数值属性的深度置信网络分类方法

孙劲光<sup>1</sup>, 蒋金叶<sup>2</sup>, 孟祥福<sup>1</sup>, 李秀娟<sup>2</sup>

SUN Jinguang<sup>1</sup>, JIANG Jinye<sup>2</sup>, MENG Xiangfu<sup>1</sup>, LI Xiujuan<sup>2</sup>

1. 辽宁工程技术大学 电子与信息工程学院, 辽宁 葫芦岛 125105

2. 辽宁工程技术大学 研究生学院, 辽宁 葫芦岛 125105

1. School of Electronics and Information Engineering, Liaoning Technical University, Huludao, Liaoning 125105, China

2. Institute of Graduate, Liaoning Technical University, Huludao, Liaoning 125105, China

SUN Jinguang, JIANG Jinye, MENG Xiangfu, et al. DBN classification algorithm for numerical attribute. *Computer Engineering and Applications*, 2014, 50(2): 112-115.

**Abstract:** Deep Belief Network (DBN) is a deep architecture that consists of several Restricted Boltzmann Machines (RBM). Generally the inputs of RBM are binary vector which leads to the information loss and in turn degrades the performance of classification. For the problem above, a DBN classification algorithm for numerical attribute is proposed through scaling the input into the interval between 0 and 1 with adding noise to sigmoid units, and achieving classification with one Gaussian hidden node on the top-level RBM. DBN can be used as feature extraction method as well as neural network with initially learned weights. DBN should have a better performance than the traditional neural network due to the initialization of the connecting weights rather than just using random weights in neural network. Experiments conducted on the dataset from UCI show that the proposed algorithm has a better accuracy than the traditional classification algorithm like SVM.

**Key words:** numerical attribute; classification; Deep Belief Network (DBN); associate memory

**摘 要:** 深度置信网络是个包含多个受限玻尔兹曼机的深层架构。针对深度置信网络分类时由于受限玻尔兹曼机的输入一般是二值向量而造成的信息的丢失从而使分类效果降低的问题, 提出了通过在sigmoid单元中增加噪声来将输入缩放到[0, 1]区间, 使用带有一个高斯隐藏节点的顶层受限玻尔兹曼机实现分类功能的一种数值属性深度置信网络分类方法。深度置信网络和受限玻尔兹曼机可以作为特征提取方法也可以认为是带有训练的初始权值的神经网络。由于连接权值的初始化而不仅仅是神经网络的随机权值, 深度置信网络分类应该比原有的传统的神经网络分类拥有更好的性能。在UCI的多个数据集上进行对比验证, 实验结果表明深度置信网络分类方法比传统的SVM算法拥有更高的准确性。

**关键词:** 数值属性; 分类; 深度置信网络; 联想记忆

**文献标志码:** A **中图分类号:** TP391 **doi:** 10.3778/j.issn.1002-8331.1308-0377

## 1 引言

当前多数分类的学习方法为多浅层结构算法, 例如支持向量机(Support Vector Machine, SVM)、Boosting<sup>[1]</sup>和Logistic Regression<sup>[2]</sup>, 其局限性在于有限样本和计算单元情况下对复杂函数的表示能力有限, 针对复杂分类问题其泛化能力受到一定制约。BP算法作为传统训练

多层网络的典型算法, 而实际上对于仅包含几层网络, 该训练方法就已很不理想<sup>[3]</sup>。深度学习通过组合低层特征形成更加抽象的高层表示(属性类别或特征), 以发现数据的分布式特征表示<sup>[4]</sup>。深度置信网络(Deep Belief Network, DBN)是由多层受限玻尔兹曼机(Restricted Boltzmann Machines, RBM)组成的深层网络结构。

**基金项目:** 国家青年科学基金(No.61003162); 国家科技支撑计划(No.2013bah12f01)。

**作者简介:** 孙劲光(1962—), 女, 教授, 博士生导师, 主研方向为数据挖掘、图形和图像处理与人脸识别; 蒋金叶(1989—), 女, 通讯作者, 硕士研究生, 主要研究方向为数据挖掘; 孟祥福(1981—), 男, 博士, 副教授, 主要研究方向为Web数据库柔性查询与优化、信息安全; 李秀娟(1988—), 硕士研究生, 主要研究方向为无线通信与数据处理。E-mail: jiangjinye2012@126.com

**收稿日期:** 2013-08-29 **修回日期:** 2013-10-21 **文章编号:** 1002-8331(2014)02-0112-04

Hinton 等人<sup>[5-7]</sup>基于 DBN 提出的贪心逐层非监督学习过程解决传统 BP 算法训练多层神经网络的难题: (1) 需要大量含标签训练样本集; (2) 较慢的收敛速度; (3) 因不合适的参数选择陷入局部最优。RBM 的训练初始化深度置信网络的参数。

RBM 中包含可视层和隐含层, 底层是可视的且使用二值特征向量的一个原因是 DBNs 上的大多数可用信息假设为二值特征(0 或 1), 并且尝试使用连续特征的数值属性还是实验阶段。数值属性(numerical attribute)是定量的, 即它是可度量的量, 用整数或实数值表示<sup>[8]</sup>。使用二值输入将由于信息丢失而降低网络的分类性能, 那么急需可以将非监督学习算法与输入数据的本质相适应<sup>[9]</sup>的分类方法。通过在 sigmoid 单元增加噪音将连续值缩放到 [0, 1] 区间来解决连续值的输入, 以及带有高斯隐藏节点的顶层 RBM 提出一种连续数据的 DBN 分类方法。训练是无监督的, 但是它产生了有用的特征, 然后通过分类或者维度约简进行调整。网络训练的关键点: (1) 预训练: 以贪婪方式每次只训练一层; (2) 为了保留输入信息在每一层使用非监督学习; (3) 微调整个网络使其到达最优。将解决了连续值输入的分类方法使用 DBN 结构, 应用到著名的 Iris 数据(一个连续数据集)。

## 2 深度置信网络

深度置信网络是个包含多层神经网络的生成模型, 并且可以有效地表示、训练非线性数据。典型的, RBMs 有效的训练过程使它适于作为 DBN 的构成模块。DBN 的核心思想是自底向上每一层 RBM 对输入数据进行提取、抽象, 尽可能保留重要信息。

### 2.1 受限玻尔兹曼机

玻尔兹曼机(无约束的 RBM)是个包括一层可视神经元和一层隐含神经元的网络。它的神经元是二值随机神经元, 意思是它只可以处于两种状态, “开”或者“关”。神经元处于“开”的概率是它的偏置量、网络中其他节点的联接权值以及网络中所有神经元的状态的 sigmoid 函数。一个玻尔兹曼机在可视和隐含神经元的能量分配条件下定义, 分配条件下仅定义了每个神经元的状态。为了简化训练时间复杂度, DBN 使用的是受限玻尔兹曼机, 它由可视变量集合  $V=\{v_i\}$  和隐含变量集合  $H=\{h_j\}$ ,  $i$  是可视层节点,  $j$  是隐含层节点。它只允许隐含神经元和可视神经元之间的连接, 可视神经元之间和隐含神经元之间没有连接。在 RBM 中, 能量定义为:

$$Energy(v, h) = b^T v + c^T h + h^T W v \quad (1)$$

其中,  $W$  是权重矩阵,  $b$  和  $c$  分别是可视层和隐含层的偏置向量。在训练 RBM 时, 本文通过可视神经元将训练数据作为网络的输入, 目的是更新权重和偏置量来最

小化网络分配的能量。分配的概率是由分配的能量决定的, 其中较小的能量分配对应较大概率的发生:

$$p(v, h) = \frac{1}{z} e^{-Energy} = \frac{e^{-Energy}}{\sum_i e^{-Energy(v_i, h)}} \quad (2)$$

其中,  $z$  是归一化常数, 通过可视和隐含单元的所有可能分配的能量来计算。训练 RBM 时由于同一层中的条件独立性(无连接), 条件分布如下:

$$p(h_j = 1 | v) = f(b_j + \sum_i w_{ij} v_i) \quad (3)$$

$$p(h_j = 0 | v) = 1 - p(h_j = 1 | v) \quad (4)$$

$$p(v_i = 1 | h) = f(c_i + \sum_j w_{ij} h_j) \quad (5)$$

$$p(v_i = 0 | h) = 1 - p(v_i = 1 | h) \quad (6)$$

其中, 函数  $f$  是 sigmoid 函数,  $\sigma(z) = 1/(1 + e^{-z})$ 。

RBM 参数的更新公式如下:

$$\Delta w_{ij} = \varepsilon (\delta \ln p(v) / \delta w_{ij}) = \varepsilon (\langle v_i h_j \rangle_{\text{data}} - \langle v_i h_j \rangle_{\text{model}}) \quad (7)$$

$$\Delta v_i = \varepsilon (\langle v_i^2 \rangle_{\text{data}} - \langle v_i^2 \rangle_{\text{model}}) \quad (8)$$

$$\Delta h_i = \varepsilon (\langle h_i^2 \rangle_{\text{data}} - \langle h_i^2 \rangle_{\text{model}}) \quad (9)$$

其中,  $\langle \rangle_{\text{model}}$  需要花费指数级时间来计算, 那么使用逼近梯度的对比差异算法(CD)来代替之<sup>[10]</sup>。对比差异是依赖逼近的方法, 运行吉布斯迭代进行采样, 而不是直到其达到收敛。这种情况下, 那么  $\langle \rangle_{\text{model}}$  将被使用来代表关于样本分布的期望, 该样本是一步吉布斯采样初始化数据得到的。新的更新准则如下:

$$\Delta w_{ij} = \varepsilon (\langle v_i h_j \rangle_{\text{data}} - \langle v_i h_j \rangle_1) \quad (10)$$

$$\Delta v_i = \varepsilon (\langle v_i^2 \rangle_{\text{data}} - \langle v_i^2 \rangle_1) \quad (11)$$

$$\Delta h_i = \varepsilon (\langle h_i^2 \rangle_{\text{data}} - \langle h_i^2 \rangle_1) \quad (12)$$

### 2.2 深度置信网络架构

每一层 RBM 的隐含单元学习、表达原始输入数据的更高程度关系的特征。DBN 可以认为是带有已训练的初始权值的神经网络。已有工作证明了下面的三个规则: (1) 顶层单元个数超越阈值, 准确性在一定水平保持稳定; (2) 层数增多, 计算性能趋于下降; (3) RBM 训练随着迭代次数增长, 性能增长。

当在 DBN 结构(非监督训练)中不使用类别标签和后向传播, DBN 可以用来作为维度约简的特征提取方法。另一方面, 当将类别标签和特征向量联系起来, DBN 被用来进行分类。两种 DBN 分类器架构, 它们是(BP-DBN)以及联想记忆 DBN(AM-DBN)。这两种架构都使用贪婪逐层 CD 准则来初始化权值。

#### 2.2.1 后向传播 DBN

增加一个代表期望的输出( $k$  个输出)变量的最后一层, 然后使用 BP 实现区分性的微调阶段。使用后向传播来微调生成模型学习得到的特征比使用随机初始值

的传统神经网络进行后向传播更有优势。

### 2.2.2 联想记忆 DBN

DBN 的顶层 RBM 可以很好地学习实验数据的抽象特征。然而,这些抽象特征仍然需要分类并与特定的标签相联系。换句话说,顶层 RBM 为输入和相关联的目标类别的联合概率分布建模。当可能的类别较多,不同类的频率分布远非一致,为目标类别使用不同的编码将比标准的  $1-k$  softmax 编码<sup>[11]</sup>更有优势。为每类使用一个任意的固定二值编码是相当直接的。假设用  $q$  维编码向量来表示表达每个类,如此,模型将得到  $q$  个输出单元;令  $y$  为对应特定输入得到的、网络生成的  $q$  维列向量;  $y$  需要被转化成类别前向概率的一个向量。如果  $C_j$  是为类别  $j$  保留编码的行向量,那么模型在已知  $y$  时,分配给类别  $t$  概率表达为:

$$p(t|\theta, y) = \frac{e^{C_y}}{\sum_j e^{C_j}} \quad (13)$$

## 3 深度置信网络分类

原始 RBM 使用二值随机单元为可视层和隐含层的输入,在连续值输入的 RBMs,提出将 sigmoid 单元中增加噪声<sup>[12]</sup>使得输入缩放到  $[0, 1]$  区间。Harmonium RBM 不同于原始的 RBM,它带有高斯连续隐含节点<sup>[10]</sup>,作为网络的最后一层,完成离散输出的分类。其中,  $f$  是标准分布函数,形式如下面的公式:

$$p(h_j = 1|x) = N(c_j + w_j x, 1) \quad (14)$$

DBN 包括三层 RBM, 第一层 RBM 将输入视为可视层,并且隐含层是第二层 RBM 的可视层。最后一层 RBM 的隐含层将是 DBN 的输出,且它只包括一个单元。DBN 分类方法可以总结如下;第一个 RBM 从变量接收输入,并且在隐含层上对其建模。然后隐含层的输入传递给第二层可视节点。这个建模和传递方法延续到只包含一个可视节点的最后一层。同时,网络最后一层只取一个节点,对输出的结果进行区间的编号,并将区间编号与类别匹配,从而实现分类功能。

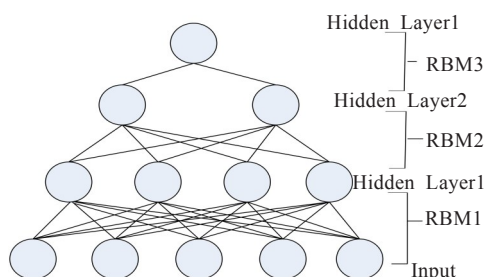


图1 DBN 网络的结构

#### 算法1 DBN 分类算法

- 步骤1 初始化吉布斯方法执行次数  $gn$ , 初始化  $\epsilon$ 。  
步骤2 读取并缩放输入到  $[0, 1]$  区间并赋值给两维

数组  $v[NI][NF]$ , 其中  $NI, NF$  分别为输入, 特征的个数。

步骤3 选取最具有区分性的特征。

步骤4 初始化 RBMs 的数量  $n=3$ , 初始化 RBM 隐含层单元的个数。

步骤5 随机初始化 DBN 网络的权值数组化  $W=[W_1, W_2, W_3]$ 。

步骤6 定义 DBN 训练权值  $W'=[W'_1, W'_2, W'_3]$ 。

步骤7 调用 DBN 训练算法  $DBN\_train(n, v, gn)$ 。

步骤8 聚类输出, 将类别标签赋值给每个聚类簇。

步骤9 用训练过的 DBN 运行测试数据集中的样本, 根据 DBN 层输出的聚类簇的范围, 定义样本的类别。

#### 算法2 DBN 训练算法

$DBN\_train(n, v, gn)$  //训练输入, 训练的结果是个长度为  $NI$  的一维数组

$W_1, b_1 \leftarrow RBM\_Alg(v, \epsilon, W'_1, b, c, n/2, gn)$

For(训练数据集中每个输入  $k$ )

for(所有的隐含单元  $i$ )

$v1[k][i] = P(v1[k][i] = 1 | v[k]) =$

$\text{sigm}(b_1[i] + \text{sum}_j(W_1[i][j] * v1[k][j]))$

$W_2, b_2 \leftarrow RBM\_Alg(v1, \epsilon, W'_2, b, c, n/4, gn)$

训练数据集中每个输入  $k$

For(训练数据集中每个输入  $k$ )

for(所有的隐含单元  $i$ )

$v2[k][i] = P(v2[k][i] = 1 | v1[k]) =$

$\text{sigm}(b_2[i] + \text{sum}_j(W_2[i][j] * v1[k][j]))$

$W_3, b_3 \leftarrow RBM\_Alg(v2, \epsilon, W'_3, b, c, 1, gn)$

For(训练数据集中每个输入  $k$ )

for(所有的隐含单元  $i$ )

$v3[k][0] = P(v3[k][0] = 1 | v2[k]) =$

$\text{sigm}(b_3[0] + \text{sum}_j(W_3[0][j] * v2[k][j]))$

Return  $v_3$  //DBN 网络的输出;

#### 算法3 RBM 训练算法

$RBM\_Alg(v, \epsilon, W, b, c, l, gn)$  //  $l$  是隐含单元的个数  
重复执行  $gn$  次

For(训练数据集中每个输入  $k$ )

for(所有的隐含单元  $i$ )

$h[i] = P(h[i] = 1 | v[k]) =$

$\text{sigm}(b[i] + \text{sum}_j(W[i][j] * v[k][j]))$

for(所有的可视单元  $j$ )

$v^T[k][j] = P(v[k][j] = 1 | h) =$

$\text{sigm}(c[j] + \text{sum}_i(W[i][j] * h[i]))$

for(所有的隐含单元  $i$ )

$h^T[i] = P(h[i] = 1 | v^T[k]) =$

$\text{sigm}(b[i] + \text{sum}_j(W[i][j] * v^T[k][j]))$

$W += \epsilon * (h * v[k] - h^T * v^T[k])$

$b += \epsilon * (h - h^T)$

$c += \epsilon * (v[0] - v[1])$

return  $W, b, c$



4 实验效果与分析

4.1 迭代次数的确定

为了测试训练每个RBM迭代的次数是如何影响整个DBN的性能表现,这里在Iris数据集上使用3层的DBN测试了训练迭代的次数从10到10 000时网络的表示能力。由图2可知,当迭代次数大于1 000时,其准确性基本保持不变。

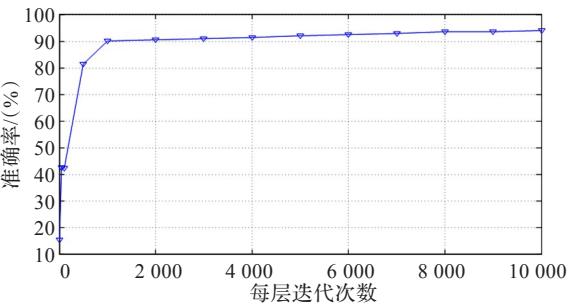


图2 RBM训练迭代次数

4.2 分类效果分析

以DBN分类应用到UCI数据库(<http://archive.ics.uci.edu/ml/datasets.html>)中著名的Iris数据集为例,该数据集包含150个样本和3类。分别使用90%的输入来进行训练,10%进行测试。DBN网络的输出分成3个不同的区间,输出的结果显示在图3中,分别是区间1[0,0.266],区间2[0.267,0.785 658],区间3[0.786,1]。样本的标签显示类2,1,0分别对应区间1,2,3。那么将数据集的10%应用到训练过的DBN并且找到每个样本的输出所在的范围。根据划分的区间,决定每个样本的类别,然后将结果和已与样本相关联的类标签比较。分类的性能不通过增加RBM层数或者吉布斯方法的数量来改变。

表1 数据集描述

DataSet	Object	Attribute	Class
Iris	150	4	3
Flag	194	27	6
Glass	214	9	6
Segment	2 310	19	7
Vehicle	846	18	4
Wine	178	13	3

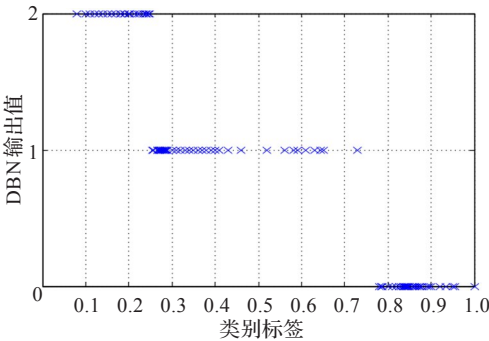


图3 DBN输出结果

为了进一步验证本文方法的有效性,在UCI的6个数据集上进行了实验,每次训练和测试采用5次交叉验证<sup>[13-14]</sup>,即将样本分为5等份,分别选择其中一份作为测试集,其他4份作为训练集。相同数据集上同样将5次交叉实验应用到KNN、SVM<sup>[15]</sup>分类方法中。其中,KNN分类方法设置 $k=3$ ,采用欧式距离计算公式;SVM分类方法中设置 $g$ 搜索值在不同数量级上等公差搜索,SVM训练中时 $C$ 取值为10 000。如表2所示,实验结果表明本文分类方法在相同的数据集情况下,准确性高于其他两种方法;在不同数据集上的分类性能均有较好的表现。

表2 分类算法准确性比较 (%)

DataSet	DBN	KNN	SVM
Iris	98.5	97.3	98.0
Flag	88.9	81.1	86.7
Glass	92.4	89.1	92.1
Segment	99.5	98.7	99.0
Vehicle	93.1	83.1	89.5
Wine	99.2	80.1	91.9

5 结束语

现有分类方法多为浅层学习结构,针对其有限样本和计算单元情况下对复杂函数的表示能力有限,复杂分类问题时其泛化能力受到一定制约等局限性,本文提出一种深度学习结构的数值属性的深度置信网络分类方法,它通过将二值输入缩放到[0,1]区间,解决连续属性输入的问题。本文研究中已经阐明DBN的非监督学习可以导致数据的维度约简以及之后的数据的分类。这种方法已经在UCI中著名的数据集上进行测试,并将其与其他分类方法进行比较,验证了提出的分类方法的有效性。在以后的学习中将继续DBN分类存在问题的改进研究。

参考文献:

[1] Freund Y, Schapire R E.Experiments with a new boosting algorithm[C]//ICML,1996,96:148-156.

[2] Jordan A.On discriminative vs.generative classifiers: a comparison of logistic regression and Naive Bayes[J]. Advances in Neural Information Processing Systems, 2002,14.

[3] Bengio Y.Learning deep architectures for AI[J].Foundations and Trends in Machine Learning,2009,2(1).

[4] Bengio Y, Delalleau O.On the expressive power of deep architectures[C]//Proc of the 14th International Conference on Discovery Science.Berlin: Springer-Verlag, 2011:18-36.

[5] Hinton G, Osindero S, Teh Y.A fast learning algorithm for deep belief nets[J].Neural Computation,2006,18(7): 1527-1554.

(下转 174 页)

- [3] Latecki L J, Lakamper R. Convexity rule for shape decomposition based on discrete contour evolution[J]. Computer Vision and Image Understanding, 1999, 73(3): 441-454.
- [4] Zhang D S, Lu G J. Study and evaluation of different Fourier methods for image retrieval[J]. Image and Vision Computing, 2005, 23(1): 33-49.
- [5] Elghazal A, Basir O, Belkasim S. A novel curvature-based shape Fourier descriptor[C]//15th IEEE International Conference on Image Processing, 2008: 953-956.
- [6] Zhuo Yang, Kamata S. Fast polar and spherical Fourier descriptors for feature extraction[C]//20th International Conference on Pattern Recognition, 2010: 975-978.
- [7] Kunttu I, Lepisto L. Shape-based retrieval of industrial surface defects using angular radius Fourier descriptor[J]. IET Image Processing, 2007, 1(2): 231-236.
- [8] 范春年, 陈建坤, 傅德胜. 一种二维极坐标傅里叶描述子在图像检索中的应用[J]. 计算机工程与应用, 2004, 40(24): 77-79.
- [9] 赵涛, 邓伟. 结合 PCA 和 FPD 的傅里叶形状描述子[J]. 计算机工程, 2011, 37(21): 149-151.
- [10] Zhao Lele, Wang Bing, Shu Huazhong. Multi-scale Fourier descriptor with phase information for image retrieval[C]//1st International Conference on Energy Systems and Electrical Power, 2011, 13: 5068-5075.
- [11] 郑楚君, 杨志勇, 何惠玲. 傅里叶描述子和 Hough 变换检测封闭边界运动[J]. 计算机工程与应用, 2005, 41(28): 68-70.
- [12] 万玮, 冯学智, 肖鹏峰. 基于傅里叶描述子的高分辨率遥感图像地物形状特征表达[J]. 遥感学报, 2011, 15(1): 73-87.
- [13] Huang D Y, Hu W C, Chuang C W. Gait recognition of different people groups based on Fourier descriptor and support vector machine[J]. Hybrid Intelligent Systems, 2011, 11: 601-604.
- [14] 李向阳, 潘云鹤. 基于椭圆的形状描述与检索方法[J]. 中国科学: E 辑, 2002, 32(5): 674-684.

(上接 115 页)

- [6] Bengio Y, Lamblin P, Popvici D, et al. Greedy layer-wise training of deep networks[C]//Proc of the 12th Annual Conference on Neural Information Processing System, 2006: 153-160.
- [7] Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [8] 范明, 孟晓峰. 数据挖掘概念与技术[M]. 北京: 机械工业出版社, 2010: 28-29.
- [9] Larochelle H, Bengio Y, Louradour J, et al. Exploring strategies for training deep neural networks[J]. The Journal of Machine Learning Research, 2009, 10: 1-40.
- [10] Noulas A K, Kröse B J A. Deep belief networks for dimensionality reduction[C]//Proceedings of the 20th Belgian-Dutch Conference on Artificial Intelligence, 2008: 185-191.
- [11] Dietterich T G, Bakiri G. Solving multiclass learning problems via error-correcting output codes[J]. Journal of Artificial Intelligence Research, 1994, 2(1): 263-286.
- [12] Chen H, Murray A F. Continuous restricted Boltzmann machine with an implementable training algorithm[C]//IEE Proceedings Vision, Image and Signal Processing, 2003, 150(3): 153-158.
- [13] Grandvalet Y, Bengio Y. Hypothesis testing for cross-validation[J]. Montreal Université de Montreal, Operationnelle DdLeR, 2006: 1-10.
- [14] Anguita D, Ghio A, Ridella S, et al. K-fold cross validation for error rate estimate in support vector machines[C]//DMIN, 2009: 291-297.
- [15] 肖汉光, 蔡从中. 特征向量的归一化比较性研究[J]. 计算机工程与应用, 2009, 45(22): 117-119.

(上接 155 页)

- [5] Lazic N, Frey B J, Aarabi P. Solving the uncapacitated facility location problem using messages passing algorithms[C]//Proceedings of 13th International Conference on Artificial Intelligence and Statistics(AISTATS), 2010: 429-436.
- [6] Dueck D, Frey B J. Non-metric affinity propagation for unsupervised image categorization[C]//Proceedings of the 11th International Conference on Computer Vision(ICCV). Riode Janeniro: IEEE Press, 2007: 1-8.
- [7] Xi X, Keogh E, Shelton C, et al. Fast time series classification using numerosity reduction[C]//Proceedings of the 23rd International Conference on Machine Learning, 2006: 1033-1040.
- [8] Zhang Dengsheng, Lu Guojun. Review of shape representation and description techniques[J]. Pattern Recognition, 2004, 37(1): 1-19.
- [9] 宋坤, 李丽娟, 赵英英. 基于 PCA 的仿射传播聚类算法[J]. 计算机工程与应用, 2011, 47(34): 212-214.
- [10] Gandhi A. Content-based image retrieval: plant species identification[EB/OL]. (2002-09-27). <http://web.engr.Oregonstate.edu/%7Eetgd/leaves/dataset.htm>.
- [11] UCR time series classification/clustering[EB/OL]. [2011-10-27]. [http://www.cs.ucr.edu/~eamonn/time\\_series\\_data/](http://www.cs.ucr.edu/~eamonn/time_series_data/).
- [12] Tan Pang-Ning, Steinbach M, Kumar V. 数据挖掘导论[M]. 北京: 人民邮电出版社, 2006.
- [13] Liao T W. Clustering of time series data-a survey[J]. Pattern Recognition, 2005, 38(11): 1857-1874.