

基于概率神经网络和 K-L 散度的样例选择*

翟俊海, 李 畅, 李 塔, 王熙照

(河北大学 数学与计算机学院 河北省机器学习与计算智能重点实验室, 河北 保定 071002)

摘要: 提出了一种基于概率神经网络和 K-L 散度的样例选择算法。该算法利用概率神经网络估计训练样例的概率分布, 利用 K-L 散度作为启发式来进行样例选择, 用该方法选出的样例大多分布在分类边界附近。与五个著名的样例选择算法 CNN、ENN、RNN、MCS 和 ICF 进行了实验比较, 实验结果显示, 算法的选择比更低, 训练出分类器具有更好的泛化能力, 提出的方法是有效的。

关键词: 概率神经网络; 样例选择; K-L 散度; 最近邻分类

中图分类号: TP181 **文献标志码:** A **文章编号:** 1001-3695(2014)01-0063-03

doi: 10.3969/j.issn.1001-3695.2014.01.014

Instance selection based on probabilistic neural network and K-L divergence

ZHAI Jun-hai, LI Chang, LI Ta, WANG Xi-zhao

(Hebei Province Key Laboratory of Machine Learning & Computational Intelligence, College of Mathematics & Computer Science, Hebei University, Baoding Hebei 071002, China)

Abstract: This paper proposed an instance selection method based on probabilistic neural network and K-L divergence. Firstly, it employed the probabilistic neural network to estimate the probabilistic distribution of training samples, used the K-L divergence as heuristic to select instances, and distributed most of the selected instances with the proposed method near the class boundary. It experimentally compared the proposed method with five famous instance algorithms which were CNN, ENN, RNN, MCS and ICF, much lower selection ratios could be achieved and better generalization ability could be obtained with the classifier trained with the selected instances. The experimental results show that the proposed method is effective and efficient.

Key words: probabilistic neural network; instance selection; K-L divergence; nearest neighbor classification

0 引言

样例选择是从数据集中选择一个样例子集,使得在该子集上训练出的分类器,其性能与在原数据集上训练出的分类器的性能大致相同。从没有类标的样例集中进行样例选择称为主动学习。主动学习的目的是为了减少标注样例的代价。文献[1,2]对基于主动学习样例选择进行了很好的综述。本文所指的样例选择是指从有类标的样例集中选择样例。样例选择算法根据性能的要求大致可分为能力增强型、能力保持型和混合类型^[3]三类。能力增强型算法代表性的工作包括编辑近邻算法 ENN (edited nearest neighbor)^[4]、RENN (repeated edited nearest neighbor)^[5]算法和 All kNN 算法^[5]。混合型算法代表性的工作包括 ICF (iterative case filtering)^[3]、基于样例 (instance based) 学习的算法 IBx ($x = 1 \sim 3$)^[6]和基于递减约简优化过程 (decremental reduction optimization procedure) 的算法 DROPx ($x = 1 \sim 3$)^[7]。这两类算法近几年研究得相对较少,而能力保持型算法是研究的热点。能力保持型算法代表性的工作包括压缩近邻算法 CNN (condensed nearest neighbor)^[8]、约简近邻算法 RNN (reduced nearest neighbor)^[9]、选择近邻算法 SNN (selective nearest neighbor)^[10]、最小一致子集算法 MCS (minimal consistent subset)^[11]和合并样例算法 MI (merging in-

stances)^[12]等。CNN 算法是 Hart 于 1968 年提出的第一个样例选择算法,它的核心概念是一致子集(即能正确分类训练集中每一个样例的子集)。CNN 试图寻找训练集的最小一致子集,但 CNN 算法选择出的子集往往不是最小一致子集。另外, CNN 算法对噪声和样例出现的顺序都非常敏感,选择出的子集仍可能包含冗余样例。为了克服 CNN 算法的不足, Gates 提出了 RNN 算法。该算法首先用 CNN 算法选出一个子集,然后去除冗余的样例。若最小一致子集包含于 CNN 算法选出的子集中,则 RNN 算法可以得到最小一致子集。SNN 算法是对 RNN 的进一步改进,该算法能够保证选择出最小一致子集,但 SNN 算法思想复杂,且计算复杂度高。MCS 算法的核心概念是最近异类近邻子集 NUNS (nearest unlike neighbor subset),它根据 NUNS 定义样例的重要度,并根据重要度选择样例。该算法能够找到训练集的最小一致子集。MI 算法的思想是通过反复尝试将两个样例合并为一个样例,直到压缩后的数据集达成一致。近几年提出的有代表性的能力保持型样例选择算法包括 Angiulli^[13]于 2007 年提出的基于沃罗诺伊图 (Voronoi diagram) 的快速压缩近邻 (FCNN) 算法, Fayed 等人^[14]于 2009 年提出的模板约简算法, Nikolaidis 等人^[15]于 2011 年提出的类边界保护的数据压缩算法, Nicolás 等人提出的基于民主距离的样例约简算法等^[16,17]。

在诸多的样例选择算法中,大都倾向于选取靠近分类边界

收稿日期: 2013-04-12; **修回日期:** 2013-05-29 **基金项目:** 国家自然科学基金资助项目(61170040);河北省自然科学基金资助项目(F2013201220, F2013201110);河北大学自然科学基金资助项目(2011-228043);河北大学教育教学改革研究项目(JX07-Y-27)

作者简介: 翟俊海(1964-),男,教授,博士,主要研究方向为机器学习(mczjh@hbu.cn);李畅(1987-),女,硕士研究生,主要研究方向为机器学习;李塔(1985-),男,硕士研究生,主要研究方向为机器学习;王熙照(1963-),男,教授,博导,主要研究方向为机器学习。

的样例。这是因为这些样例的分类不确定性大,它们所提供的分类信息量也大,对分类起着重要的作用。另外,大多数样例选择算法只适用于近邻分类算法。本文提出了一种基于概率神经网络和 K-L 散度的样例选择算法。训练好的概率神经网络可用于计算样例属于每一类的概率分布,而 K-L 散度可度量两个分布之间的距离,这样概率神经网络与 K-L 散度结合可用于估计样例在空间的分布信息,这些分布信息对选取重要的样例非常有用。本文提出的算法利用概率神经网络估计训练样例的概率分布,利用 K-L 散度作为启发式来进行样例选择,并与五个著名的样例选择算法 CNN、ENN、RNN、MCS 和 ICF 进行了实验比较。实验结果显示,本文算法的压缩比更高,训练出的分类器具有更好的泛化能力,本文提出的方法是行之有效的。

1 概率神经网络

本章简单介绍概率神经网络^[18, 19]。概率神经网络具有很多成功的应用,如叶志锋等人^[20]将概率神经网络应用于发动机故障诊断,杨鼎强等人^[21]将差异演化概率神经网络应用于纹理图像识别,程智辉等人^[22]将概率神经网络应用于计算机辅助医疗诊断等。概率神经网络(probabilistic neural networks, PNN)是 Specht^[18]于 1990 年提出的。PNN 由输入层、模式层和类别层构成。对于给定的训练集 $T = \{(x_j, y_j) | j = 1, 2, \dots, n\}$, 其中 $x_j \in \mathbb{R}^d, y_j \in \{1, 2, \dots, c\}$, 即训练集中的样例分为 c 类 $\omega_1, \omega_2, \dots, \omega_c$ 。 T 所对应的 PNN 如图 1 所示。

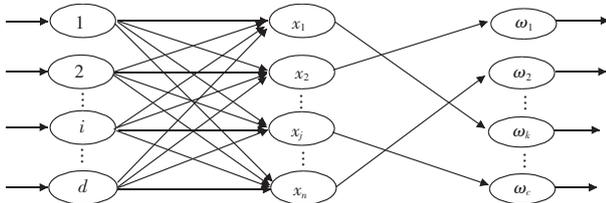


图 1 概率神经网络

图中输入层有 d 个节点,模式层有 n 个节点,类别层有 c 个节点。PNN 的输入层与模式层之间采用全连接方式,连接到模式层第 j 个节点的 d 个连接权分别是模式向量 x_j 长度归一化后的各个分量,即 $w_{jk} = x_{jk} (j = 1, 2, \dots, n; k = 1, 2, \dots, d)$ 。模式层到类别层采用稀疏连接方式,模式层的第 j 个节点只与其对应的类别节点有连接,与其他节点无连接,连接权为 1。PNN 的训练过程实际上是模式向量长度归一化的过程,训练算法^[19]如算法 1 所示。

算法 1 PNN 训练

```

let  $j = 0, a_{ji} = 0$ 
for each  $j$ 
  for each  $k$ 
    let  $j = j + 1$ 
  calculate  $x_{ji} = \frac{x_{ji}}{\sqrt{\sum_{s=1}^d x_{js}^2}}$ 
  let  $w_{ji} = x_{ji}$ 
  if  $x \in \omega_k$  then  $a_{jk} = 1$ 
    
```

训练好的 PNN 用于分类归一化的测试样例 $x \in \mathbb{R}^d$, 这一过程也称为分类过程。对于给定的测试样例 x , 首先计算模式层节点 $j (j = 1, 2, \dots, n)$ 的激活值 $t_j = w_j \times x$, 其输出为 $o_j = \exp[(t_j - 1)/\sigma^2]$, 其中 σ 是用户定义的参数。PNN 的分类过程^[19]如算法 2 所示。

算法 2 PNN 分类

```

let  $j = 0$ 
do
    
```

```

let  $j = j + 1$ 
let  $t_j = w_j \cdot x$ 
for each  $k$ 
  if  $a_{jk} = 1$  then  $p_k(x) = p_k(x) + \exp[(t_j - 1)/\sigma^2]$ 
until  $j = n$ 
let  $\omega = \arg \max_k \{p_k(x)\}$ 
    
```

2 基于概率神经网络和 K-L 散度的样例选择

K-L 散度也称为相对熵,是两个概率分布之间距离的一种度量。对于离散型随机变量 x , 设它的所有可能取值的集合为 V , 它的两个概率密度函数为 $p(x)$ 和 $q(x)$, 则这两个概率分布之间的 K-L 散度定义为

$$D(p|q) = \sum_{x \in V} p(x) \log \frac{p(x)}{q(x)}$$

本文算法的基本思想如图 2 所示。算法首先将数据集划分成 $n + 1$ 个子集,对于任意子集 i ,用其他 n 个子集训练出的 PNN 分类器组成一个样例选择委员会 $B^{[23]}$, 然后根据 K-L 散度选择子集 i 中的样例,算法以迭代的方式选出最终的样例子集。

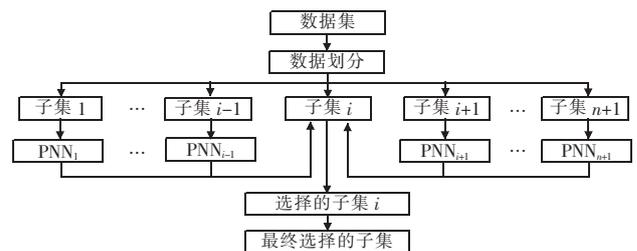


图 2 基于 PNN 的样例选择方法

设委员会 B 的 n 个成员为 $PNN_1, PNN_2, \dots, PNN_n$, 每个成员是一个概率神经网络,基于 K-L 散度按下面的标准选择样例:

$$x^* = \arg \max_x \left\{ \frac{1}{n} \sum_{i=1}^n D(P_{PNN_i} | P_B) \right\}$$

其中: $D(P_{PNN_i} | P_B) = \sum_{k=1}^c P_{PNN_i}(\omega_k | x) \log_2 \frac{P_{PNN_i}(\omega_k | x)}{P_B(\omega_k | x)}$

$$P_B(\omega_k | x) = \frac{1}{n} \sum_{i=1}^n P_{PNN_i}(\omega_k | x)$$

算法描述如下:

- a) 训练一个 PNN 委员会 B , 它含有 n 个概率神经网络;
- b) 对任意测试样例 x , 对委员会中的每一个成员 $PNN_i (1 \leq i \leq n)$, 计算 x 属于每一类 $\omega_k (1 \leq k \leq c)$ 的后验概率 $P_{PNN_i}(\omega_k | x) (1 \leq i \leq n; 1 \leq k \leq c)$;
- c) 计算委员会 B 关于样例 x 的平均后验概率 $P_B(\omega_k | x) = \frac{1}{n} \sum_{i=1}^n P_{PNN_i}(\omega_k | x)$;

d) 计算 K-L 散度:

$$D(P_{PNN_i} | P_B) = \sum_{k=1}^c P_{PNN_i}(\omega_k | x) \log_2 \frac{P_{PNN_i}(\omega_k | x)}{P_B(\omega_k | x)} = \sum_{k=1}^c P_{PNN_i}(\omega_k | x) \log_2 \frac{P_{PNN_i}(\omega_k | x)}{\frac{1}{n} \sum_{i=1}^n P_{PNN_i}(\omega_k | x)}$$

e) 计算委员会的平均 K-L 散度:

$$\text{avg}(D(P_{PNN_i} | P_B)) = \frac{1}{n} \sum_{i=1}^n D(P_{PNN_i} | P_B)$$

f) 选择样例:

$$x_{KL}^* = \arg \max_x \left\{ \frac{1}{n} \sum_{i=1}^n D(P_{PNN_i} | P_B) \right\}$$

3 实验结果

为了进一步验证本文算法的有效性,采用十折交叉验证的

方法,分别在两个人工数据集六个 UCI 数据集^[24]上与五个著名的样例选择算法进行了实验比较,这五种算法分别是 CNN、ENN、RNN、MCS 和 ICF。为描述方便,本文算法简记为 PNN-KL。实验环境是 PC 机,双核 1.86 GHz CPU,2 GB 内存,Windows XP 操作系统,MATLAB 7.1 实验平台。

两个人工数据集分别是两类二维云数据^[25]和两类二维高斯数据。两类二维云数据包含 10 000 个样本点,每类 5 000 个点。第一类数据服从的概率分布为

$$p(\mathbf{x}|\omega_1) = \frac{1}{2} \left(\frac{p_1(\mathbf{x})}{2} + \frac{p_2(\mathbf{x})}{2} + p_3(\mathbf{x}) \right)$$

其中: $\mathbf{x} = (x_1, x_2)$, $p_i(\mathbf{x}) = \frac{1}{2\pi\sigma_{ix_1}\sigma_{ix_2}} \exp\left(-\frac{(x_1 - \mu_{ix_1})^2}{2\sigma_{ix_1}^2} - \frac{(x_2 - \mu_{ix_2})^2}{2\sigma_{ix_2}^2}\right)$ 。

第二类数据服从的概率分布为

$$p(\mathbf{x}|\omega_2) = \frac{1}{2\pi} \exp\left(-\frac{x_1^2 + x_2^2}{2}\right)$$

二维云数据的分布如图 3 所示(为了显示清楚,图中只取了 15% 的数据点)。用本文方法选出的样例分布情况如图 4 所示。

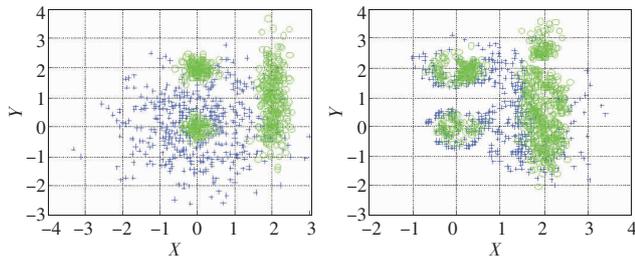


图 3 二维云数据的分布 图 4 选出的二维云数据的分布

两类二维高斯数据包含 1 000 个样本点,每类 500 个点, $p(\mathbf{x}|\omega_i) \sim N(\mu_i, \Sigma_i)$, $i = 1, 2$, 其中, $\mu_1 = (1.0, 1.0)^T$, $\mu_2 = (2.5, 2.5)^T$, $\Sigma_1 = \begin{pmatrix} 0.6 & -0.2 \\ -0.2 & 0.6 \end{pmatrix}$, $\Sigma_2 = \begin{pmatrix} 0.2 & -0.1 \\ -0.1 & 0.2 \end{pmatrix}$ 。两类二维高斯数据的分布如图 5 所示(取 50% 的点显示),用本文方法选出的样例分布如图 6 所示。

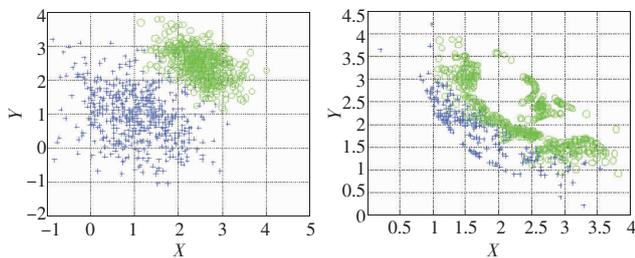


图 5 二维高斯数据的分布 图 6 选出的二维高斯数据的分布

从图 4 和 6 可以看出,利用本文算法选择出的样例大多靠近分类边界,对于 UCI 数据集也有类似结论。

六个 UCI 数据集的基本信息如表 1 所示。

表 1 实验用的六个 UCI 数据集的基本信息

datasets	# training instances	# testing instances	# attributes	# classes
banana	3 533	1 767	2	2
sensor	3 637	1 819	2	4
mushroom	3 762	1 882	22	2
satImage	4 290	2 145	36	6
USPS	6 198	3 100	257	10
nursery	8 640	4 320	8	5

实验中,根据数据集的规模人工确定划分子集的大小,与五个著名的样例选择算法从平均测试精度和选择比两方面进

行了比较,实验结果列于表 2 中。表 2 中,aver 表示平均测试精度,ratio 表示选择比例。从实验结果可以看出,本文算法的平均测试精度优于其他五种方法,选择比例仅次于 MCS 算法,优于其他四种算法,实验结果证实了本文方法的有效性。

表 2 实验结果

datasets	CNN		ENN		RNN		MCS		ICF		PNN-KL	
	aver	ratio	aver	ratio								
cloud	0.84	0.26	0.89	0.85	0.84	0.26	0.87	0.15	0.88	0.12	0.90	0.20
Gaussian	0.98	0.08	0.99	0.98	0.98	0.14	0.98	0.03	0.98	0.01	0.99	0.06
banana	0.86	0.23	0.88	0.87	0.86	0.23	0.88	0.14	0.88	0.14	0.90	0.08
sensor	0.98	0.05	0.99	0.99	0.98	0.14	0.98	0.05	0.99	0.14	0.95	0.07
mushroom	1.00	0.11	1.00	1.00	1.00	0.11	1.00	0.01	1.00	0.21	1.00	0.16
satImage	0.88	0.20	0.91	0.90	0.88	0.20	0.90	0.13	0.89	0.18	0.89	0.17
USPS	0.94	0.13	0.98	0.97	0.94	0.13	0.96	0.10	0.95	0.15	0.94	0.23
nursery	0.85	0.29	0.74	0.78	0.85	0.29	0.79	0.21	0.73	0.48	0.84	0.27
平均性能	0.916	0.169	0.921	0.917	0.915	0.188	0.920	0.102	0.912	0.178	0.925	0.155

4 结束语

大多数样例选择算法只适用于近邻分类算法,针对这一问题,本文提出了一种基于概率神经网络和 K-L 散度的样例选择算法。该算法利用概率神经网络估计训练样例的概率分布,利用 K-L 散度作为启发式来进行样例选择,并与五个著名的样例选择算法 CNN、ENN、RNN、MCS 和 ICF 进行了实验比较。实验结果显示,本文算法的平均测试精度优于这五种算法,选择比例仅次于 MCS 算法,基于本文算法选择出的样例大多分布在分类边界附近;该方法行之有效。

参考文献:

- [1] 龙军,殷建平,祝恩,等. 主动学习研究综述[J]. 计算机研究与发展, 2008, 45(增刊): 300-304.
- [2] 吴伟宁,刘扬,郭茂祖,等. 基于采样策略的主动学习算法研究进展[J]. 计算机研究与发展, 2012, 49(6): 1162-1173.
- [3] BRIGHTON H, MELLISH C. Advances in instance selection for instance-based learning algorithms [J]. Data Mining and Knowledge Discovery, 2002, 6(2): 153-172.
- [4] WILSON D L. Asymptotic properties of nearest neighbor rules using edited data [J]. IEEE Trans on Systems, Man and Cybernetics, 1972, 2(3): 408-421.
- [5] TOMK I. An experiment with the edited nearest-neighbor rule [J]. IEEE Trans on Systems, Man and Cybernetics, 1976, 6(6): 448-452.
- [6] AHA D W, KIBLER D, ALBERT M K. Instance based learning algorithms [J]. Machine Learning, 1991, 6(1): 37-66.
- [7] WILSON D R, MARTINEZ A R. Instance pruning techniques [C]// Proc of the 14th International Conference on Machine Learning. San Francisco: Morgan Kaufmann, 1997: 403-411.
- [8] HART P. The condensed nearest neighbor rule [J]. IEEE Trans on Information Theory, 1968, 14(5): 515-516.
- [9] GATES G W. The reduced nearest neighbor rule [J]. IEEE Trans on Information Theory, 1972, 18(3): 431-433.
- [10] RITTER G L, WOODRUFF H B, LOWRY S R, et al. An algorithm for the selective nearest neighbour decision rule [J]. IEEE Trans on Information Theory, 1975, 21(6): 665-669.
- [11] DASARATHY B V. Minimal consistent set identification for optimal nearest neighbor decision systems design [J]. IEEE Trans on Systems, Man, and Cybernetics, 1994, 24(1): 511-517.
- [12] CHANG C L. Finding prototypes for nearest neighbor classifiers [J]. IEEE Trans on Computers, 1974, 23(11): 1179-1184.

果。从表 3 中可以看出,本文算法与 VPR 时序驱动布线算法相比,运行时间增加比例最大的是 diffeq 电路,增加了 13.7%;比例最小的是 clma 电路,增加了 0.4%;运行时间平均增加了 6.8%。从运行时间的比较结果中可看出,修复的时序违反路径越多,增加的时间比例也就越大。实验证明,本文算法只需要少量的运行时间,就可以达到修复短时序违反路径的目的。

表 2 短时序约束违反路径个数的比较

电路	VPR 时序布线	文献[10]	本文算法		
			违反路径 个数	比 VPR 减少/%	比文献[10] 减少/%
diffeq	37	26	2	94.6	92.3
dsip	5	7	0	100.0	100.0
frisc	17	12	2	88.2	83.3
s38417	7	3	0	100.0	100.0
tseng	2	1	0	100.0	100.0
elliptic	8	6	0	100.0	100.0
clma	2	1	0	100.0	100.0
s298	3	1	0	100.0	100.0
bigkey	4	1	0	100.0	100.0
s38584.1	9	2	1	88.9	50.0
平均	9.4	6	0.5	94.7	91.7

表 3 运行时间的比较

电路	VPR 时序布 线算法/ms	本文算 法/ms	比 VPR 增加%	电路	VPR 时序布 线算法/ms	本文算 法/ms	比 VPR 增加%
diffeq	3 766	516	13.7	clma	26 656	109	0.4
dsip	2 125	63	3.0	s298	3 156	16	0.5
frisc	8 438	875	10.4	bigkey	2 953	78	2.6
s38417	11 391	1 500	13.2	s38584.1	10 516	1 266	12.0
tseng	1 688	94	5.6	平均	7 931	543	6.8
elliptic	8 625	922	10.7				

4 结束语

本文提出一种自动修复短时序违反路径的 FPGA 布线算法。该算法首先调用 VPR 时序算法进行整体布线,在整体布线布通之后,调用短路径时序分析来获取违反短路径时序约束的布线连接,然后对每条违反短路径时序约束的连接进行增量布线。在增量布线时通过修改代价函数,使得每条连接的路径延时尽可能达到满足时序约束所需要的延时。实验结果表明,该算法与 VPR 时序布线算法相比,在增加少量运行时间的情况下,能够修复绝大部分的短时序约束违反路径。

参考文献:

- [1] SHENOY N, BRAYTON R, SANGIOVANNI-VINCENTELLI A. Minimum padding to satisfy short path constraints [C]//Proc of IEEE/ACM International Conference on Computer-Aided Design. Santa Clara: ACM Press, 1993:156-161.
- [2] FRANKLE J. Iterative and adaptive slack allocation for performance-driven layout and FPGA routing [C]//Proc of the 29th ACM/IEEE Design Automation Conference. Los Alamitos: IEEE Computer Society, 1992:536-542.
- [3] McMURCHIE L, EBELING C. PathFinder: a negotiation-based performance-driven router for FPGAs [C]// Proc of the 3rd International ACM Symposium on Field-Programmable Gate Arrays. New York: ACM Press, 1995:111-117.
- [4] MLEXANDER M, ROBINS G. New performance-driven FPGA routing algorithms [J]. IEEE Trans on Computer-Aided Design of Integrated Circuits and Systems, 1996, 15(12):1505-1517.
- [5] BETZ V, ROSE J. VPR: a new packing, placement and routing tool for FPGA research [C]//Proc of the 7th International Workshop on Field Programmable Logic and Application. London: Springer, 1997: 213-222.
- [6] BETZ V, ROSE J, MARQUARDT A. Architecture and CAD for deep-submicron FPGAs [M]. Norwell: Kluwer Academic Publishers, 1999.
- [7] RUBIN R, De HON A. Timing-driven pathfinder pathology and remediation: quantifying and reducing delay noise in VPR-pathfinder [C]//Proc of the 19th International ACM Symposium on Field-Programmable Gate Arrays. New York: ACM Press, 2011:173-176.
- [8] ROSE J, LUU J, YU Chi-wai, et al. The VTR project: architecture and CAD for FPGAs from verilog to routing [C]//Proc of the 20th International ACM Symposium on Field-Programmable Gate Arrays. New York: ACM Press, 2012:77-86.
- [9] FUNG R, BETZ V, CHOW W. Simultaneous short-path and long-path timing optimization for FPGAs [C]//Proc of IEEE/ACM International Conference on Computer-Aided Design. Washington DC: IEEE Computer Society, 2004:838-845.
- [10] FUNG R, BETZ V, CHOW W. Slack allocation and routing to improve FPGA timing while repairing short-path violations [J]. IEEE Trans on Computer-Aided Design of Integrated Circuits and Systems, 2008, 27(4):686-697.
- [11] 黄娟, 杨海钢, 李威, 等. 可编程逻辑阵列减少毛刺的低功耗布线算法 [J]. 计算机辅助设计与图形学报, 2010, 22(10):1664-1670.
- [12] 1990, 3(1):109-118.
- [13] ANGIULLI F. Fast nearest neighbor condensation for large datasets classification [J]. IEEE Trans on Knowledge and Data Engineering, 2007, 19(11):1450-1464.
- [14] FAYED H A, ATIYA A F. A novel template reduction approach for the K-nearest neighbor method [J]. IEEE Trans on Neural Networks, 2009, 20(5):890-896.
- [15] NIKOLAIDIS K, GOULERMAS J Y, WU Q H. A class boundary preserving algorithm for data condensation [J]. Pattern Recognition, 2011, 44(3):704-715.
- [16] De HARO-GARCIA A, GARCÍA-PEDRAJAS, Del CASTILLO A R. Large scale instance selection by means of federal instance selection [J]. Data & Knowledge Engineering, 2012, 75:58-77.
- [17] CÉSAR G O, De AIDA H G, NICOLÁS G P. Democratic instance selection: a linear complexity instance selection algorithm based on classifier ensemble concepts [J]. Artificial Intelligence, 2010, 174(5-6):410-441.
- [18] SPECHT D F. Probabilistic neural networks [J]. Neural Networks, 1990, 3(1):109-118.
- [19] DUDA R O, HART P E, STORK D G. Pattern classification [M]. 2nd ed. Beijing: China Machine Press, 2005.
- [20] 叶志锋, 孙健国. 基于概率神经网络的发动机故障诊断 [J]. 航空学报, 2002, 23(2):155-157.
- [21] 杨鼎强, 肖淑苹, 蒋加伏. 基于差异演化概率神经网络的纹理图像识别 [J]. 计算机工程与应用, 2008, 44(11):179-181, 198.
- [22] 程智辉, 陈将宏. 基于概率神经网络的乳腺癌计算机辅助诊断 [J]. 计算机仿真, 2012, 29(9):166-169.
- [23] SEUNG H S, OPPER M, SOMPOLINSKY H. Query by committee [C]//Proc of ACM Workshop on Computational Learning Theory. 1992:287-294.
- [24] FRANK A, ASUNCION A. UCI machine learning repository [EB/OL]. (2010). <http://archive.ics.uci.edu/ml>.
- [25] VERIKAS A, LIPNICKAS A, MALMQVIST K, et al. Soft combination of neural classifiers: a comparative study [J]. Pattern Recognition Letters, 1999, 20(4):429-444.

(上接第 65 页)