

# 智能拉普拉斯分类器\*

戴宏亮

(广东商学院 数学与计算科学学院, 广东 广州 510320)

**摘要:** 针对拉普拉斯分类器的核参数选择问题, 通过首先假设窗的三个估计中核参数取不同的值, 然后运用智能遗传算法对核参数进行优选, 得到一种新的分类器——智能拉普拉斯分类器。多个基准数据集上的实验结果证明, 智能拉普拉斯分类器相对普通拉普拉斯分类器和支持向量机而言, 具有较高的分类精度和稳定性, 是一种有效的分类方法。

**关键词:** 核方法; 拉普拉斯分类器; 支持向量机; 智能遗传算法

**中图分类号:** TP39; TP18 **文献标识码:** A **文章编号:** 0529-6579(2009)02-0014-06

## The Intelligent Laplacian Classifier

DAI Hongliang

(School of Mathematics and Computational Science, Guangdong University of Business Studies,  
Guangzhou 510320, China)

**Abstract:** Supposing the kernel parameters of the three Parzen window-based estimators are different and using intelligent genetic algorithm (IGA) to optimize these parameters, a novel classifier is developed—the intelligent Laplacian classifier (TILC). The experimental results in several benchmark datasets indicate that the proposed TILC achieve higher classification accuracy and stable than the Laplacian classifier and standard support vector machine (SVM). Consequently, TILC provides a promising alternative for classification.

**Key words:** kernel methods; the Laplacian classifier; support vector machine; intelligent genetic algorithm

分类是机器学习的一个基本问题<sup>[1-2]</sup>。一般的分类规则是构造分类器, 使得分类错误的概率最小化, 贝叶斯分类器就是这样的一个分类器。但是贝叶斯分类器需要知道类概率密度函数(class probability density functions, pdf's), 实际应用中, 类概率密度函数一般是很难获得<sup>[3]</sup>。解决此问题的一个办法是从训练数据中得到 Parzen 窗估计<sup>[4]</sup>, 然后用 Parzen 窗估计代替真正的类概率密度函数。Parzen 窗估计表示成加权 Parzen 窗或者核的和的形式。它是一种核方法, 在对测试数据集进行分类时, 对每一个数据点, 加权和都必须计算, 计算复杂度的阶为  $O(N_{\text{test}}N_{\text{train}})$ 。

在过去十几年里, 机器学习中的核方法已经变

得非常流行。其中, 比较著名的核方法有核 Fisher 判别分析(KFDA)<sup>[5]</sup>、核主成分分析(KPCA)<sup>[6]</sup>和支持向量机(SVM)等<sup>[7-9]</sup>。虽然这些方法的理论基础不同, 但是都被表示成加权核展开的形式。KFDA 基本思想是经过非线性映射将输入数据映射到一个高维特征空间, 在特征空间中进行 Fisher 线性判别分析, 隐含地实现输入空间的非线性判别。经典的主成分分析(PCA)是一种线性算法, 不能抽取数据中非线性的结构。而 KPCA 用非线性变换将输入数据空间映射到高维特征空间, 转化为求核矩阵的特征向量和特征值。而输入数据在特征向量上的投影转换为求核函数的线性组合。SVM 是基于统计学习理论框架的一种通用机器学习方法,

\* 收稿日期: 2008-07-01

基金项目: 国家自然科学基金资助项目(10771220); 教育部高等学校博士点科研基金资助项目(SRFDP-20070558043)

作者简介: 戴宏亮(1978年生), 男, 讲师, 博士研究生; E-mail: daihongliang@tom.com

它是结构风险最小化原则思想的一种较好的实现。SVM 是从线性可分情况下的最优超平面发展起来的, 它通过最大化分类间隔控制学习机器的容量, 实现结构风险最小化原则。SVM 通过核函数将样本从输入空间映射到高维特征空间, 并在这个特征空间中构造最优分类超平面, 实现在高维空间中的推广。SVM 模型的求解最终转化成为一个凸二次优化问题, 理论上存在全局最优点, 避免了局部极值问题。SVM 最终的形式与窗贝叶斯分类器类似, 训练阶段计算复杂度为  $O(N_{\text{train}}^3)$ , 测试过程的计算复杂度为  $O(N_{\text{test}}^3 N_{\text{sv}})$ 。

与 SVM 分类性能相当, 但是不需要规划过程, 计算复杂度较低的一种新型分类器——拉普拉斯分类器被提出。拉普拉斯分类器建立在 Mercer 核特征空间中类平均向量间的测量角度的基础上, 分类损失函数可由 pdf's 的 Cauchy-Schwarz (CS) 信息理论散度的值得到。使用窗进行密度估计, 分类规则最终根据加权核的形式表示。每个数据点的加权和该点的概率密度成反比, 重点强调小概率密度区域。整个过程中没有规划运算, 训练阶段的运算复杂度的阶为  $O(N_{\text{train}}^3)$ , 测试过程的计算复杂度的阶为  $O(N_{\text{test}}^3 N_{\text{sv}})$ 。拉普拉斯分类器根据 Parzen 窗得到, 因此, 理论上进行 Parzen 窗密度估计就可以得到合适的核参数。但是, 实际操作中, 这是非常困难的。文献 [10-13] 对 Parzen 核参数选择进行了比较充分的研究, 但是这些方法比较复杂, 操作难度比较大。本文提出一种新的核参数优选方法, 先假设 Parzen 窗的三个估计中核参数取不同的值, 然后运用智能遗传算法 (Intelligent genetic algorithm, IGA) 对核参数进行参数优选, 得到一种新的分类器——智能拉普拉斯分类器 (TILC)。多个基准数据集上的实验结果证明, 智能拉普拉斯分类器相对普通拉普拉斯分类器、支持向量机而言, 具有较高的分类精度和稳定性, 是一种有效的分类方法。

## 1 拉普拉斯分类器

拉普拉斯分类器是一种新型的分类器, 它定义在核特征空间上, 与拉普拉斯数据矩阵特征谱紧密相关。其分类规则建立在核诱导的特征空间中, 通过比较测试数据点和类平均向量的角度进行分类。文献 [10] 中已经验证了拉普拉斯分类器具有与支持向量机相当的分类性能, 并且具有更低阶的计算复杂度, 可以用于大规模数据。拉普拉斯分类器分类规则根据加权核展开表示, 每个数据点加权与

该点的概率密度成反比, 因此拉普拉斯分类器对小概率类别具有更高的分类精度。

首先考虑两类情形。假设两类分别为  $\omega_1$  和  $\omega_2$ , 两类的条件概率密度函数分别为  $p_1(x)$  和  $p_2(x)$ 。分类损失函数定义为这两个条件概率密度函数间的 Cauchy-Schwarz 距离, 即

$$D(p_1, p_2) = -\log \frac{\langle p_1, p_2 \rangle_f}{\sqrt{\langle p_1, p_1 \rangle_f \langle p_2, p_2 \rangle_f}} \quad (1)$$

式中,  $\langle p_i, p_j \rangle_f \equiv \int p_i(x) p_j(x) f^{-1}(x) dx, i, j = 1, 2$ 。  $P_1$  和  $P_2$  分别是两个类的类先验概率,  $f(x) = P_1 p_1(x) + P_2 p_2(x)$  是数据集的全部概率密度函数。令  $h(x) = f^{-\frac{1}{2}}(x) p_1(x)$ ,  $g(x) = f^{-\frac{1}{2}}(x) p_2(x)$ 。将 (1) 式中对数的真数记为

$$P_f = \frac{\int h(x) g(x) dx}{\sqrt{\int h^2(x) dx \int g^2(x) dx}} \quad (2)$$

损失函数能使用类似于 Parzen 窗密度估计技术进行估计。给定一个数据集  $x_i, i = 1, \dots, N$ 。该数据集由两类数据组成, 一类点  $x_i, i = 1, \dots, N_1$ ; 二类点  $x_j, j = 1, \dots, N_2$ 。建立在数据抽样的基础上, 定义 Parzen 窗估计

$$\begin{aligned} \hat{f}(x) &= \frac{1}{N} \sum_{i=1}^N W_\sigma(x, x_i) \\ \hat{h}_1(x) &= \frac{1}{N_1} \sum_{i=1}^{N_1} f^{-\frac{1}{2}}(x_i) W_\sigma(x, x_i) \\ \hat{h}_2(x) &= \frac{1}{N_2} \sum_{j=1}^{N_2} f^{-\frac{1}{2}}(x_j) W_\sigma(x, x_j) \end{aligned} \quad (3)$$

式中,  $W_\sigma(\cdot, \cdot)$  是 Parzen 窗。

给定一个单峰高斯密度函数

$$W_\sigma(x, x_i) = \frac{1}{(2\pi\sigma^2)^{\frac{d}{2}}} \exp\left\{-\frac{\|x - x_i\|^2}{2\sigma^2}\right\} \quad (4)$$

使用这些估计, 然后利用卷积定理核 Mercer 定理, 可以得到

$$\int \hat{h}_1(x) \hat{h}_2(x) dx = \langle m_{1f}, m_{2f} \rangle \quad (5)$$

式中,  $m_{1f}$  和  $m_{2f}$  是数据映射到 Mercer 核诱导特征空间后的类平均向量,  $m_{1f} = \frac{1}{N_1} \sum_{i=1}^{N_1} \Phi_f(x_i), m_{2f} =$

$\frac{1}{N_2} \sum_{j=1}^{N_2} \Phi_f(x_j), \Phi_f(\cdot)$  是一个非线性映射。

随之,  $P_f$  的 Parzen 窗估计能被表示为

$$\hat{P}_f \equiv \text{LIC} = \cos \angle(m_{1f}, m_{2f}) \quad (6)$$

式中, LIC 表示拉普拉斯信息截取。

Cauchy-Schwarz 散度很容易被推广到多于两类

的情形, 给定

$$D(p_1, \dots, p_c) = -\log \sum_{i=1}^{c-1} \sum_{j>i} \frac{\langle p_i, p_j \rangle_f}{k \sqrt{\langle p_i, p_i \rangle_f \langle p_j, p_j \rangle_f}} \quad (7)$$

式中,  $k = \sum_{c=1}^{c-1} c$ 。在 Mercer 核特征空间中, 对应于一种多路拉普拉斯信息截取

$$\text{LIC} = \sum_{i=1}^{c-1} \sum_{j>i} \frac{1}{k} \cos \angle (m_{if}, m_{jf})$$

建立在训练集的基础上, 可以定义相应于类  $\omega_1, \dots, \omega_c$  的类平均向量  $m_{1f}, \dots, m_{cf}$ 。通过计算  $\Phi_f(x_i)$  与每个均值向量间角度的余弦值, 然后把测试点归为角度最小的那一类, 分类规则即

$$x_i \rightarrow \omega_c: \max_{1 \leq c \leq C} \cos \angle (\Phi_f(x_i), m_{cf}) \quad (8)$$

通过推导, 可以得到

$$\begin{aligned} \cos \angle (\Phi_f(x_i), m_{cf}) &= \\ &= \frac{\sum_{i=1}^{N_c} \hat{f}^{-\frac{1}{2}}(x_i) W_\sigma(x_i, x_i) \hat{f}^{-\frac{1}{2}}(x_i)}{\sqrt{\sum_{i,j=1}^{N_c, N_c} \hat{f}^{-\frac{1}{2}}(x_j) W_\sigma(x_j, x_j) \hat{f}^{-\frac{1}{2}}(x_j)}} \end{aligned} \quad (9)$$

然后定义检验统计量

$$\gamma_c = \frac{\sum_{i=1}^{N_c} \hat{f}^{-\frac{1}{2}}(x_i) W_{\sqrt{2}\sigma}(x_i, x_i)}{V_c} \quad (10)$$

式中,  $V_c = \sqrt{\sum_{i,j=1}^{N_c, N_c} \hat{f}^{-\frac{1}{2}}(x_j) W_\sigma(x_j, x_j) \hat{f}^{-\frac{1}{2}}(x_j)}$  为常量。因此, 分类规则 (8) 变成

$$x_i \rightarrow \omega_c: \max_{1 \leq c \leq C} \cos \gamma_c \quad (11)$$

上述分类规则就称为拉普拉斯分类器。

由 (10) 式可知, 测试点  $x_i$  处的核函数  $W_{\sqrt{2}\sigma}(x_i, x_i)$  有一个加权  $\hat{f}^{-\frac{1}{2}}(x_i)$ , 因此, 加权因子与测试点  $x_i$  处所估计的全部概率密度函数成反比。这意味着在分类器设计中低概率密度区域赋予更高的加权。

## 2 智能遗传算法

文献 [10] 中已经证明拉普拉斯分类器具备良好的分类性能, 但是其核参数选择确没有太有效的办法。文献 [10] 中为了计算的简单, 将 (3) 式中的三个核参数当成同样的。但是正样本、负样本和全部样本的概率密度是不一样的, 所以把三个参数作为同样对待是不合理的。所以本文首先假设 (3) 式中三个不同估计中的核参数分别取不同的

值, 然后利用智能遗传算<sup>[14]</sup>对这三个核参数进行优选。IGA 良好的性能主要基于有效的遗传算法染色体编码和智能交叉操作。智能遗传操作建立在正交实验设计的基础上, 使用 “divide-and-conquer strategy” 来解决具有大量参数难以处理的优化问题。而且, 智能交叉采用 “systematic reasoning method” 而不是传统遗传算法的 “generate-and-go method” 来加速搜索。

### 2.1 适应度函数

运用 IGA 设计 TILC 的主要目的是分类精度  $C_a$  达到最大。适应度函数可以被定义为

$$\max y(S) = C_a(S) \quad (12)$$

在这里  $S$  表示需要被 IGA 优化的参数集。

为了减轻 IGA 的搜索负担, 对 TLC 参数设置整数个值构成搜索空间

$$\sigma_1, \sigma_2, \sigma_3 \in \{0.0001 \times 2^d, 0.001 \times 2^d, 0.01 \times 2^d, 0.1 \times 2^d, 1 \times 2^d\}, d = 0, 1, 2, 3, 4$$

### 2.2 正交实验设计

正交实验设计综合正交列 (orthogonal array, OA) 和因子分析 (factor analysis) 的优点。因子可以看作变量或者参数, 一个集合中的一个元素可以看作因子中的一个阶层。一个完全的因子实验将考虑因子中所有的阶层组合。但是, 在实际问题中, 由于所有阶层组合数目太大, 进行完全的因子实验是不可行的。因此, 比较明智的做法是, 考虑子集的阶层组合, 进行部分因子实验。正交实验设计正是利用部分因子实验的良好特性来选取最优组合, 达到解决问题的目的。下面对智能遗传算法中的二阶层正交阵列作一个简单的描述。令有  $N$  个因子, 每个因子有两个阶层, 一个完全的因子实验有  $2^N$  个组合。令  $M = 2^{\lceil \log_2(N+1) \rceil}$ , 在这里  $[M]$  表示不大于  $M$  的最大整数。构造一个  $M$  行,  $M-1$  列的正交阵列  $L_M(2^{M-1})$ , 只用考虑前面的  $N$  列。正交阵列能大大减少因子分析中阶层组合的数目。正交阵列中需要分析的所有因子组合数目仅仅是  $M = O(N)$ , 在这里  $N+1 \leq M \leq 2N$ 。再令  $y_t$  表示组合  $t$  的适应度函数值,  $t = 1, \dots, M$ 。定义因子  $i$  阶层  $k$  的主要效应为  $S_{ik}$ ,  $i = 1, \dots, N$ :

$$S_{ik} = \sum_{t=1}^M y_t W_t \quad (13)$$

在这里如果因子  $i$  的阶层的组合  $t$  的值是  $k$ , 则  $W_t = 1$ , 否则  $W_t = 0$ 。当  $S_{i1} > S_{i2}$  时, 阶层 1 更好; 当  $S_{i1} < S_{i2}$  时, 阶层 2 更好; 当  $S_{i1} = S_{i2}$ , 阶层 1 和 2 有同样的贡献。最有效的因子  $i$  具有最大的主要效应差  $MED_i = |S_{i1} - S_{i2}|$ 。在每个因子两个阶

层中更好一个被确定好之后, 一个更有效的组合即可以得到。

### 2.3 智能交叉操作

类似传统 GA, 在一次交叉操作中, 两个父代个体  $P_1$  和  $P_2$  产生两个子代个体  $C_1$  和  $C_2$ 。下面给出其具体操作步骤。

Step 1: 使用一个正交阵列  $L_M(2^{M-1})$  的前  $N$  列;

Step 2: 令第  $i$  个因子的阶层 1 和阶层 2 分别表示父代个体  $P_1$  和  $P_2$  产生的第  $i$  个参数种群;

Step 3: 计算第  $t$  次试验的适应度值  $y_t$ ,  $t = 2, \dots, M$ 。  $y_1$  是  $P_1$  的适应度值;

Step 4: 计算主要效应差  $S_{ik}$ ,  $i = 1, \dots, N, k = 1, 2$ ;

Step 5: 决定每个因子中更好的一个阶层;

Step 6: 从导出的相应父代个体中组合更好的 GA 基因构成子代个体  $C_1$  的 GA 染色体;

Step 7: 子代个体  $C_2$  的 GA 染色体构成和  $C_1$  类似, 但是有一点不同的是具有最小主要效应差的因子选自其它阶层;

Step 8: 在  $P_1, P_2, C_1, C_2$  和  $M - 1$  个正交阵列组合中的最好的两个个体作为最后的优秀子代个体  $G_1$  和  $G_2$ 。

一个智能交叉操作需要通过计算  $M + 1$  个适应度值来搜索  $2^N$  个组合,  $N + 1 \leq M \leq 2N$ 。

### 2.4 智能遗传算法

由上面的介绍, 可以得到具体的 IGA 算法如下:

Step 1: 初始化。随机生成一个初始种群, 该种群具有  $N_{pop}$  个可行个体, 并且在一个遗传算法染色体中每一个基因  $g_i$  是唯一的。

Step 2: 评价。评价种群中所有个体的适应度值。令  $I_{best}$  为种群中最优的个体。

Step 3: 使用简单的截断选择用最好的  $P_s \cdot N_{pop}$  个个体来替换最坏的  $P_s \cdot N_{pop}$  个个体, 在这里  $P_s$  是一个选择概率。

Step 4: 随机选择  $P_c \cdot N_{pop}$  个个体 (包括  $I_{best}$ ), 进行交叉操作, 在这里  $P_c$  是一个交叉概率。

Step 5: 使用变异概率  $P_m$  对种群进行变异操作。为了防止最优个体恶变, 不对最优个体应用变异操作。

Step 6: 终止测试。如果提前设定的终止条件满足, 停止算法, 否则, 转到步骤 2。

## 2.5 TILC 模型流程图

TILC 模型流程图见图 1。

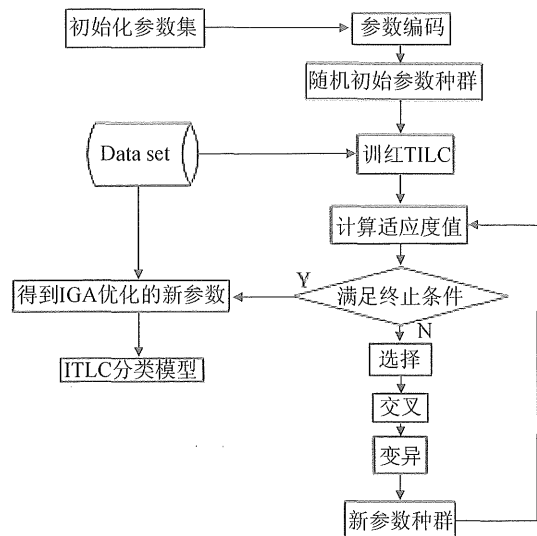


图 1 TILC 模型流程图

Fig. 1 Flow chart of TILC

## 3 实验结果及分析

### 3.1 实验 1

本实验数据集是从 UCI 选取的一个超声波心动图数据集<sup>[15]</sup>。该数据集收集的是病人最近遭受一次心脏侵袭的信息, 判别病人在遭受袭击后能否活过 1 年。选取 20 个数据组成训练集, 其中有 6 个数据点存活期超过一年。共有 8 个特征, 分别是年龄、性别、各种心脏条件等。并且类与类之间有较大的交叉。测试数据集共选取 84 个数据。本文选取标准支持向量机和拉普拉斯分类器作为对比方法。这两种分类器的分类性能已经得到证实。其中标准支持向量机参数采用 5 - fold 交叉验证确定, 拉普拉斯分类器核参数采用 Silverman's 规则确定。拉普拉斯分类器和智能拉普拉斯分类器采用 MATLAB7.2 编程, 标准支持向量机采用 Libsvm 软件<sup>[17]</sup>。最终得到智能拉普拉斯分类器的 3 个核参数分别为:  $\sigma_1 = 0.02$ ;  $\sigma_1 = 0.4$ ,  $\sigma_3 = 0.8$ 。其它具体结果如表 1 - 3 所示。

表 1 标准支持向量机对超声波心动图数据集分类的混淆矩阵

	$\hat{\omega}_1$	$\hat{\omega}_2$
$\omega_1$	4	20
$\omega_2$	4	56

表 2 拉普拉斯分类器对超声波心动图数据集分类的混淆矩阵

Table 2 Confusion matrix for standard TLC on Echocardiogram data

	$\hat{\omega}_1$	$\hat{\omega}_2$
$\omega_1$	18	6
$\omega_2$	20	40

表 3 智能拉普拉斯分类器对超声波心动图数据集分类的混淆矩阵

Table 3 Confusion matrix for ITLC on Echocardiogram data

	$\hat{\omega}_1$	$\hat{\omega}_2$
$\omega_1$	19	5
$\omega_2$	22	38

从上面实验结果看出, 3 种分类器分类精度大致相当, 智能拉普拉斯分类器分类精度比拉普拉斯分类器有明显提升, 但是拉普拉斯分类器和智能拉普拉斯分类器对小概率类别  $\omega_2$  分类更准确, 智能拉普拉斯分类器保留了拉普拉斯分类器重点强调小概率区域的特性, 并且分类精度有明显提高。

### 3.2 实验 2

本实验数据集选用 *Ratsch* 数据集中的 4 个著名的基准数据集<sup>[16]</sup>。如表 4 所示。

表 4 4 个著名基准数据集

Table 4 Four famous benchmark data

名称	维数	训练数据个数	测试数据个数
Banana	2	400	4 800
Thyroid	5	140	75
limage	18	1 300	1 010
Twonorm	20	400	4 600

本实验中依然选取标准支持向量机和拉普拉斯分类器作比较。参数选取和软件采用方法与实验 1 相同。实验结果如表 5。

表 5 3 种分类器在 *Ratsch* 数据集分类结果比较

Table 5 Comparison of classification results for three classifiers on *Ratsch* data

数据	SVM	TLC	ITLC
Banana	89.2 ± 0.7	88.4 ± 0.6	90.3 ± 0.5
Thyroid	95.2 ± 2.2	94.4 ± 2.6	95.1 ± 2.3
limage	97.0 ± 0.6	96.2 ± 0.7	97.2 ± 0.2
Twonorm	97.0 ± 0.2	97.3 ± 0.3	98.2 ± 0.2

实验结果可以看出, 智能拉普拉斯分类器分类精度和稳定性明显优于拉普拉斯分类器, 和标准支持向量机分类器相比, 在 3 个数据集上分类性能高于标准支持向量机, 另外一个数据集上的性能和标准支持向量机相当。

### 3.3 实验 3

本实验数据集选用 UCI 中的 5 个著名的标准数据集。其维数和数目如表 6。

表 6 5 个 UCI 标准数据集

Table 6 Five UCI standard data

名称	Wine	Ionos	Iris	Wisconsin	Pima
维数	13	34	4	9	8
数目	178	351	150	683	768

在这个实验中, 每个数据集进行 50 次实验, 每次实验, 随机将数据分为两部分, 一部分进行训练, 一部分进行测试。本实验仍然选取标准支持向量机和拉普拉斯分类器作为对比方法。其中标准支持向量机参数采用 5-fold 交叉验证确定, 拉普拉斯分类器核参数采用 Silverman's 规则确定或者通过 5-fold 交叉验证确定, 下面结果都取 50 次实验中最好的 (表 7)。

表 7 3 种分类器在 UCI 标准数据集分类结果比较

Table 7 Comparison of classification results for three classifiers on UCI standard data

数据	SVM	TLC	ITLC
Wine	97.5 ± 1.4	97.3 ± 1.4	97.6 ± 1.3
Ionos	94.1 ± 1.2	92.5 ± 1.7	94.5 ± 0.9
Iris	95.7 ± 2.0	94.5 ± 2.1	94.8 ± 2.0
Wisconsin	96.9 ± 0.7	97.1 ± 0.7	97.8 ± 0.5
Pima	76.8 ± 1.5	73.9 ± 1.7	74.2 ± 1.2

实验结果可以看出, 智能拉普拉斯分类器相对拉普拉斯分类器而言, 分类性能和稳定性都有明显的提升, 在 3 个数据集上分类精度好于标准支持向量机, 另外两个稍逊于标准支持向量机。

## 4 结 论

(1) 拉普拉斯分类器是一种核方法, 具备分类性能好、计算复杂度低、重点强调小概率区域等优良特性;

(2) 智能拉普拉斯分类器不仅保留了拉普拉斯分类器计算复杂度低、重点强调小概率区域等优良特性, 而且在分类性能和稳定性方面都有明显的

提升;

(3) 智能遗传算法建立在正交实验的基础上, 具备优良的特性。利用智能遗传算法优选拉普拉斯分类器核参数非常成功, 是一种有效的核参数选择方法。

#### 参考文献:

- [1] PIERRE B, SOREN B. Bioinformatics: The machine learning approach[M]. 2nd ed. USA: arrangement with MIT through Arts & Licensing International, Inc, 2001.
- [2] WEBB A R. Statistical pattern recognition[M]. John Wiley and sons, LTD, 2002.
- [3] 杨淑莹. 模式识别与智能计算—Matlab 技术实现[M]. 北京: 电子工业出版社, 2008.
- [4] PARZEN E, On the estimation of a probability density function and the mode[J]. Annal Math Stat, 1962, 32: 1065 - 1076.
- [5] MIKA S, RATSCH G, WESTON J, et al. Fisher discriminant analysis with kernels [C] // Proc IEEE Workshop Neural Netw. Signal Process. Madison, WI, 1999: 41 - 48.
- [6] YANG J, FRANGI A F, YANG J Y, et al. KPCA plus LDA: a complete kernel Fisher discriminant framework for feature extraction and recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(2): 230 - 244.
- [7] NELLO C, JOHN S T. An Introduction to support vector machines and other Kernel-based learning methods[M]. China Machine Press, 2005.
- [8] VAPNIK V N. Statistical learning theory[M]. New York: Wiley, 1998.
- [9] VAPNIK V N. The Nature of Statistical Learning Theory [M]. New York: Springer-Verlag, 1995.
- [10] JENSSEN R, ERDOGMUS D. et al. The laplacian classifier [J]. IEEE Transactions on Signal Processing, 2007, 55(7): 3262 - 3271.
- [11] MIKA S, RATSCH G, WESTON J, et al. Fisher discriminant analysis with kernels [C] // Proc IEEE Workshop Neural Netw Signal Process, 1999: 41 - 48.
- [12] WILLIAMS C K, BARBER D. Bayesian classification with Gaussian processes [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(12): 1342 - 1351.
- [13] JENSSEN R, ERDOGMUS D, PRINCIPE J C, et al. The Laplacian PDF distance: A cost function for clustering in a kernel feature space [C] // Proc Adv Neural Inf Process Syst, 2005: 625 - 632.
- [14] HO S Y, SHU L S, CHEN J H. Intelligent evolutionary algorithms for parameter optimization problems [J]. IEEE Trans, Evolutionary Comput, 2004, 8(6): 522 - 541.
- [15] MURPHY R, ADA D. UCI repository of machine learning Databases [J]. Tech Rep Dept Comput Sci, Univ California, Irvine, 1994.
- [16] RAETSCH G, ONODA T, MÜLLER K R. Soft margins for adaboost [J]. Mach Learn, 2001, 42: 287 - 320.
- [17] CHANG C C, LIN C J. LIBSVM: a library for support vector machines [EB/OL]. 2001, software available at: <http://www.csie.ntu.edu.tw/~ejlin/libsvm>.

(上接第 13 页)

- [8] ZORKO C T. Morrey space [J]. Proc Amer Math Soc, 1986, 98: 586 - 592.
- [9] WU Z J, XIE C P. Q spaces and Morrey spaces [J]. J Funct Anal, 2003, 201: 282 - 297.
- [10] MEYER Y, COIFMANN R. Wavelets Calderon-Zygmund and multilinear operators [M]. Translated by David Salinger, Cambridge University Press, 1997.
- [11] DAUBECHIES I. Ten lectures on wavelets [M]. SIAM, Philadelphia, 1992.
- [12] DUONG X T, XIAO J, YAN L X. Old and new Morrey space with heat kernel bounds [J]. J Four Anal Appl, 2007, 13(1): 87 - 111.