

## 代价敏感学习中的损失函数设计

李秋洁<sup>1†</sup>, 赵亚琴<sup>1</sup>, 顾 洲<sup>1,2</sup>

(1. 南京林业大学 机械电子工程学院, 江苏 南京 210037; 2. 东南大学 自动化学院, 江苏 南京 210096)

**摘要:** 一般的学习算法通过最小化分类损失使分类错误率最小化, 而代价敏感学习则以最小化分类代价为目标, 需构造代价敏感损失. 本文探讨代价敏感损失的设计准则, 首先介绍基于代价敏感风险优化的代价敏感学习方法, 然后在Bayes最优分类理论框架下, 提出两条代价敏感损失设计准则. 接着采用两种常用代价敏感损失生成方法构造平方损失、指数损失、对数损失、支持向量机损失等经典损失函数的代价敏感扩展形式. 根据所提出的设计准则, 从理论上分析这些代价敏感损失的性能. 最后通过实验表明, 同时满足两条设计准则的代价敏感损失能有效降低分类代价, 从而证明了本文提出的代价敏感损失设计准则的合理性.

**关键词:** 学习算法; 代价敏感学习; 损失函数; Bayes最优决策; 代价敏感损失

**中图分类号:** TP273      **文献标识码:** A

## Design of loss function for cost-sensitive learning

LI Qiu-jie<sup>1†</sup>, ZHAO Ya-qin<sup>1</sup>, GU Zhou<sup>1,2</sup>

(1. College of Mechanical and Electronic Engineering, Nanjing Forestry University, Nanjing Jiangsu 210037, China;

2. School of Automation, Southeast University, Nanjing Jiangsu 210096, China)

**Abstract:** Conventional learning algorithms minimize the classification error through minimizing the classification loss. However, the cost-sensitive learning minimizes the classification cost; thus, cost-sensitive losses have to be constructed. This paper studies the design criteria for cost-sensitive loss functions. Firstly, cost-sensitive learning methods based on cost-sensitive risk minimization are briefly introduced. Then, under the theory framework of Bayes optimal classification, two design guidelines of cost-sensitive loss function are proposed. The cost-sensitive extensions of several classic loss functions (e.g., square loss, exponential loss, log loss and support vector machine (SVM) loss) are generated via two most popular construction methods of cost-sensitive loss. The performances of these cost-sensitive losses are theoretically analyzed based on the proposed two design guidelines. Experimental results have shown that those cost-sensitive losses that satisfy both of the two design criteria significantly reduce classification costs, demonstrating the rationality of the proposed design criteria of cost-sensitive loss.

**Key words:** learning algorithms; cost-sensitive learning; loss function; Bayes optimal decision; cost-sensitive risk

### 1 引言(Introduction)

分类问题是模式识别、机器学习、数据挖掘领域的重要研究内容. 经典的最小错误率(minimum-error-rate)学习中, 学习算法最小化分类器对未知样本错误分类的概率, 由于不同类别的样本具有相等的分类代价, 又称为代价不敏感学习(cost-insensitive learning). 现实世界中, 不同类别常具有不同的分类代价, 例如, 欺诈检测系统中, 对欺诈行的漏检将带来更高的损失; 医疗诊断中, 误诊癌症患者付出的代价远高于将健康人误判为病患的代价; 目标检测系统中, 对目标漏检带来的损失远大于虚警. 针对上述问题, 代价敏感学习(cost-sensitive learning)赋予不同类别不等的误分

代价, 学习算法最小化分类器对未知样本的分类代价<sup>[1]</sup>.

代价敏感分类问题具有普遍存在性, 这使得学习算法的代价敏感扩展成为了一个研究热点. 现有研究大致分为两类: 1) 通过优化代价敏感风险最小化分类代价. 2) 在学习过程中引入分类代价, 使得生成的分类器侧重对误分代价较高的类别正确分类. 目前, 分类决策树<sup>[2]</sup>、神经网络<sup>[3-4]</sup>、支持向量机(SVM)<sup>[5]</sup>、boosting<sup>[6-9]</sup>等常用学习算法都有其对应的代价敏感扩展算法. 文献[2]在决策树学习过程中以最小化误分代价作为选择分裂属性的依据, 算法成功应用于欺诈检测. 文献[3]研究应用于邮政编码识别的代价敏感神

收稿日期: 2014-06-05; 录用日期: 2015-01-13.

<sup>†</sup>通信作者. E-mail: liqiujiel@163.com; Tel.: +86 13951810230.

国家自然科学基金项目(31200496, 61473156), 中国博士后基金项目(2014M551487), 江苏省博士后基金项目(1301009A)资助.

Supported by National Natural Science Foundations of China (31200496, 61473156), National Science Foundation for Post-doctoral Scientist of China (2014M551487) and Postdoctoral Science Foundation of Jiangsu Province (1301009A).

神经网络分类器,分别采用代价敏感采样、代价收敛、学习速率调整、判决阈值移动等技术修改原始神经网络算法,实验结果表明4种方法中代价敏感采样取得更好的结果.Cao等人<sup>[4]</sup>将不平衡分类算法评估度量(如G均值)引入代价敏感神经网络的目标函数,采用粒子群优化方法同时优化误分代价、特征子集和网络内部结构参数,改进后的算法在不平衡数据集上取得较好的分类结果.Cao等人<sup>[5]</sup>还将这一技术应用于SVM,训练适用于不平衡数据分类的代价敏感SVM.文献[7,9]提出指数损失、对数损失的代价敏感扩展形式,进而推导出代价敏感Discrete AdaBoost,代价敏感Real AdaBoost,代价敏感Gentle AdaBoost和代价敏感Logit Boost.针对代价敏感多分类问题,Lozano等人<sup>[6]</sup>提出基于p范数损失函数的代价敏感boosting算法,付忠良等人<sup>[8]</sup>提出多分类代价敏感Real AdaBoost算法.

本文研究直接优化代价敏感风险的代价敏感学习算法.此类方法中,代价敏感损失函数是算法有效的理论保证.现有方法仅针对具体学习算法采用的损失函数构造其代价敏感形式,本文研究代价敏感损失函数设计的一般准则,在此基础上,研究两类代价敏感损失函数,分析其性能,并给出实验比较结果.本文剩余部分结构如下:第2节介绍基于代价敏感风险优化的代价敏感学习方法,第3节根据Bayes最优决策提出代价敏感损失函数设计的一般准则,第4节根据提出准则,分析两类代价敏感损失函数的性能,并在第5节给出比较实验结果,最后总结全文.

## 2 基于代价敏感风险优化的代价敏感学习方法(Cost-sensitive learning methods based on cost-sensitive risk minimization)

最小错误率分类的目标是产生分类误差最小的分类器,由于分类误差不可导,这是一个NP-hard问题.一个解决办法是构造分类误差的可导凸上界,即损失函数(loss function),通过优化训练集上的期望损失最小化分类误差.常用损失函数有平方损失、指数损失、对数损失、SVM损失等.同理,针对代价敏感分类问题,可构造代价敏感损失函数,通过优化代价敏感风险(cost-sensitive risk)最小化分类代价.具体表述如下:

已知样本空间 $X$ ,类别空间 $Y = \{\pm 1\}$ ,给定样本 $(x, y)$ ,  $x \in X$ ,  $y \in Y$ ,代价敏感损失函数 $L_C(y, F(x))$ 度量 $F$ 对 $(x, y)$ 的预测损失,其中 $C$ 是代价矩阵,定义见式(3).学习算法优化整个样本空间的代价敏感损失 $E_{X,Y}[L_C(y, F(x))]$ (代价敏感风险),求取代价敏感决策函数 $F_C^*$ ,

$$F_C^* = \arg \min_F E_{X,Y}[L_C(y, F(x))]. \quad (1)$$

给定 $x$ ,式(1)可转化为优化条件代价敏感风险

$E_Y[L_C(y, F(x))|x]$ ,即采用 $F$ 对 $x$ 类别进行预测产生的风险,

$$F_C^*(x) = \arg \min_F E_Y[L_C(y, F(x))|x]. \quad (2)$$

## 3 代价敏感损失函数设计准则(Design criteria of cost-sensitive loss function)

依据Bayes最优分类,本节提出代价敏感损失函数的一般设计准则.

### 3.1 Bayes最优分类(Bayes optimal classification)

对于二分类问题,定义代价矩阵(cost matrix)

$$C = \begin{bmatrix} c_{++} & c_{+-} \\ c_{-+} & c_{--} \end{bmatrix}, \quad (3)$$

其中 $c_{ab}$ 是将 $b$ 类样本预测为 $a$ 类样本的分类代价.根据Bayes决策理论,最优决策应最小化期望分类代价,即给定样本 $x$ ,若式(4)成立,则预测其类别为正,否则,预测其类别为负.

$$\begin{aligned} c_{++}p(x) + c_{+-}(1-p(x)) &\leq \\ c_{-+}p(x) + c_{--}(1-p(x)), \end{aligned} \quad (4)$$

其中: $p(x)$ 为后验概率,  $p(x) = P(y = +1|x)$ ,  $1-p(x) = P(y = -1|x)$ .式(4)可写作

$$\begin{cases} (c_{-+} - c_{++})p(x) \geq (c_{+-} - c_{--})(1-p(x)), \\ c_{+}p(x) \geq c_{-}(1-p(x)), \end{cases} \quad (5)$$

其中: $c_{+} = c_{-+} - c_{++}$ ,  $c_{-} = c_{+-} - c_{--}$ 分别表示正、负类样本的分类代价.因此,Bayes分类器为

$$x \text{ 类别} = \begin{cases} 1, & p(x) \geq \frac{c_{-}}{c_{+} + c_{-}}, \\ -1, & p(x) < \frac{c_{-}}{c_{+} + c_{-}}. \end{cases} \quad (6)$$

Bayes分类器也可写作 $\text{sgn}(p(x) - \frac{c_{-}}{c_{+} + c_{-}})$ ,当正负类别分类代价相等时,即 $c_{+} = c_{-}$ ,Bayes分类器为 $\text{sgn}(p(x) - \frac{1}{2})$ .

### 3.2 代价敏感损失设计准则(Design criteria of cost-sensitive loss)

**准则1** 代价敏感损失 $L_C(y, F(x))$ 需满足Bayes一致性.

即代价敏感损失能产生Bayes分类器,  $\text{sgn}(F_C^*(x))$ 能产生与式(6)一致的分类结果.

**准则2**  $F_C^*$ 对应的条件代价敏感风险 $E_Y[L_C(y, F_C^*(x))|x]$ 在Bayes分类边界 $\{x|p(x) = \frac{c_{-}}{c_{+} + c_{-}}\}$ 处取得最大值.

Bayes分类器(式(6))的期望分类代价为

$$\text{分类代价} = \begin{cases} c_{-}(1-p(x)), & p(x) \geq \frac{c_{-}}{c_{+} + c_{-}}, \\ c_{+}p(x), & p(x) < \frac{c_{-}}{c_{+} + c_{-}}. \end{cases} \quad (7)$$

在分类边界 $\{x|p(x) = \frac{c_{-}}{c_{+} + c_{-}}\}$ 处,将样本 $x$ 判为

正、负类别的分类代价相等并达到最大值  $\frac{c_+ + c_-}{c_+ + c_-}$ , 此时最难预测样本类别. 根据Bayes最优分类, 给定  $x$ ,  $F_C^*$  对其预测产生的风险  $E_Y[L_C(y, F_C^*(x))|x]$  应同样在Bayes分类边界处取得最大值.

准则1保证优化代价敏感风险得到的分类器  $F_C^*$  能最小化期望分类代价, 实现最优代价敏感分类. 准则2保证代价敏感损失能更有效地近似分类代价, 在最优决策处的代价敏感损失和分类代价有一样的全局性质.

#### 4 代价敏感损失函数(Cost-sensitive loss functions)

本节根据所提代价敏感损失设计准则, 以几种常用的损失函数为例, 分析在代价敏感学习算法中得到较多应用的两类代价敏感损失函数的性能.

具体探讨以下损失函数——平方损失、指数损失、对数损失、SVM损失. 平方损失在神经网络、回归中得到广泛应用, 指数损失成功应用于boosting算法<sup>[10-11]</sup>, 如Discrete AdaBoost, Real AdaBoost, Gentle AdaBoost等. 对数损失为LogitBoost采用的损失函数, 其相近形式  $\ln(1 + e^{-yF(x)})$  应用于Logistic回归分析. 这几种损失函数均可表示为间隔  $yF(x)$  的函数, 即  $L(y, F(x)) = L(yF(x))$ , 如表1所示.

表1 几种常用损失函数

Table 1 Several popular loss functions

名称	形式 $L(yF(x))$	最优解 $F^*(x)$
平方损失	$(y - F(x))^2 = (1 - yF(x))^2$	$2p(x) - 1$
指数损失	$e^{-yF(x)}$	$\frac{1}{2} \ln \frac{p(x)}{1 - p(x)}$
对数损失	$\ln(1 + e^{-2yF(x)})$	$\frac{1}{2} \ln \frac{p(x)}{1 - p(x)}$
SVM损失	$\max((1 - yF(x)), 0)$	$\text{sgn}(2p(x) - 1)$

图1以间隔  $yF(x)$  为横坐标绘制了分类误差和几种损失函数, 为便于显示, 对数损失上移了  $(1 - \ln 2)$ .

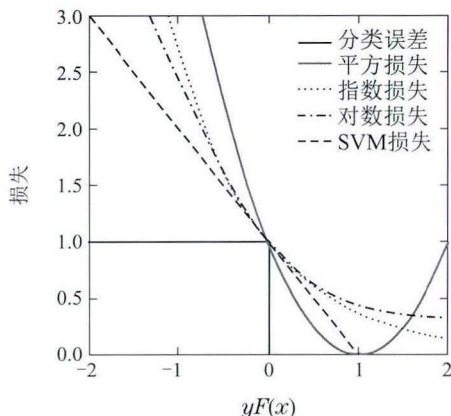


图1 分类误差和几种损失函数

Fig. 1 Classification error and several loss functions

表1给出了各个损失函数对应最优决策函数  $F^*(x)$  及其条件风险  $E_Y[L(yF^*(x))|x]$ . 图2给出了各个损失函数的最优解和最优解处的条件风险, 可看出,  $\text{sgn}(F^*(x))$  与Bayes分类器  $\text{sgn}(p(x) - \frac{1}{2})$  等价, 最优解处的条件风险与最小分类误差一致, 在  $p(x) = \frac{1}{2}$  处取得最大值.

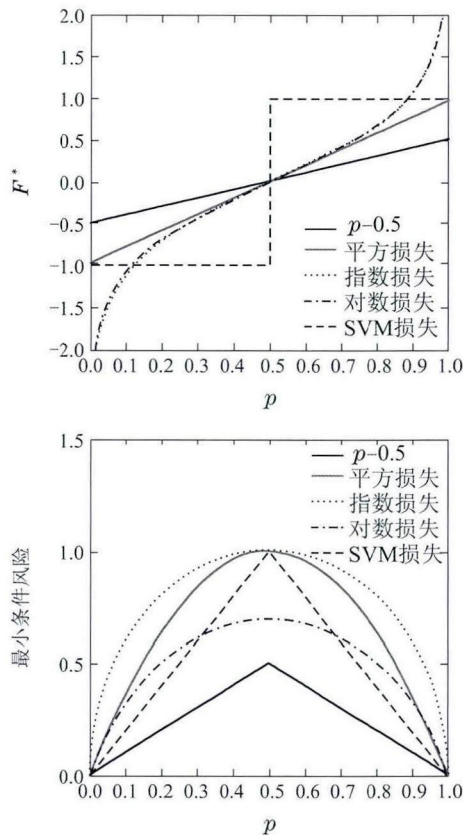


图2 原始损失  $L(yF(x))$  对应的最优解  $F^*(x)$  和最优解处的条件风险  $E_Y[L(yF^*(x))|x]$

Fig. 2 Optimal solution  $F^*(x)$  and its conditional risk  $E_Y[L(yF^*(x))|x]$  with original loss function  $L(yF(x))$

样本类别为  $y$  时, 记其分类代价为  $c_y$ :

$$c_y = \begin{cases} c_+, & y = 1, \\ c_-, & y = -1. \end{cases} \quad (8)$$

现有算法有以下两类常用分类代价引入策略: 1) 分类代价在损失函数外  $c_y L(yF(x))$ ; 2) 分类代价在损失函数内  $L(y c_y F(x))$ .

#### 4.1 分类代价在损失函数外(Classification cost outside of loss function)

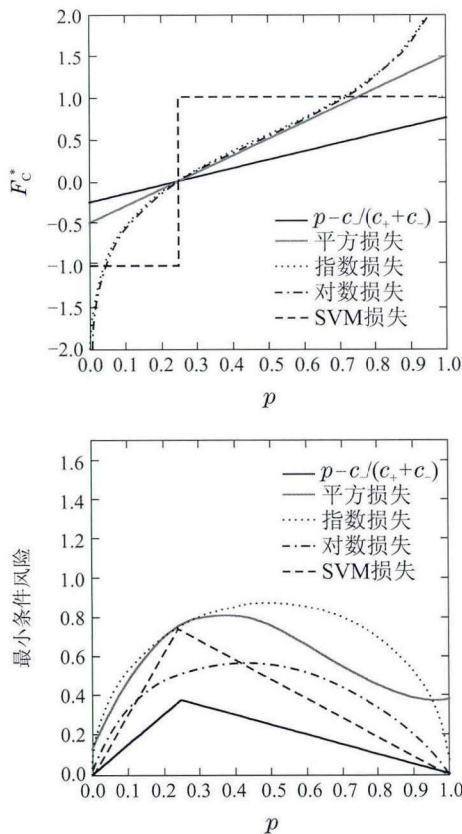
这种类型的代价敏感损失将分类代价与原始损失相乘, 采用这种策略的算法有文献 [3, 5, 9, 13] 等. 表2给出了各个代价敏感损失此时的最优决策  $F_C^*(x)$  和最优决策处的条件代价敏感风险  $E_Y[c_y L(yF_C^*(x))|x]$ , 并绘制为图3(取  $c_+ = 1.5$ ,  $c_- = 0.5$ ). 此时, Bayes分类器为  $\text{sgn}(p(x) - \frac{c_-}{c_+ + c_-}) = \text{sgn}(p(x) - \frac{1}{4})$ . 从图



中可看出,4种代价敏感损失均满足准则1,即最优分类器 $\text{sgn}(F_C^*(x))$ 为Bayes分类器.除代价敏感SVM损失外,其余损失均不满足准则2,最优解处的条件代价敏感风险没有在Bayes分类边界 $\{x|p(x) = \frac{1}{4}\}$ 处取得最大值,而是有不同程度的偏移.

表2 代价敏感损失 $c_y L(yF(x))$ Table 2 Cost-sensitive loss  $c_y L(yF(x))$ 

形式 $c_y L(yF(x))$	最优解 $F_C^*(x)$
$c_y(1-yF(x))^2$	$(c_+ + c_-)p(x) - c_-$
$c_y e^{-yF(x)}$	$\frac{1}{2} \ln \frac{c_+ p(x)}{c_- (1-p(x))}$
$c_y \ln(1 + e^{-2yF(x)})$	$\frac{1}{2} \ln \frac{c_+ p(x)}{c_- (1-p(x))}$
$c_y \max((1-yF(x)), 0)$	$\text{sgn}((c_+ + c_-)p(x) - c_-)$

图3 代价敏感损失 $c_y L(yF(x))$ 对应的最优解 $F_C^*(x)$ 和最优解处的条件风险 $E_Y[c_y L(yF_C^*(x))|x]$ Fig. 3 Optimal solution  $F_C^*(x)$  and its conditional risk

$E_Y[c_y L(yF_C^*(x))|x]$  with cost-sensitive loss function  $c_y L(yF(x))$

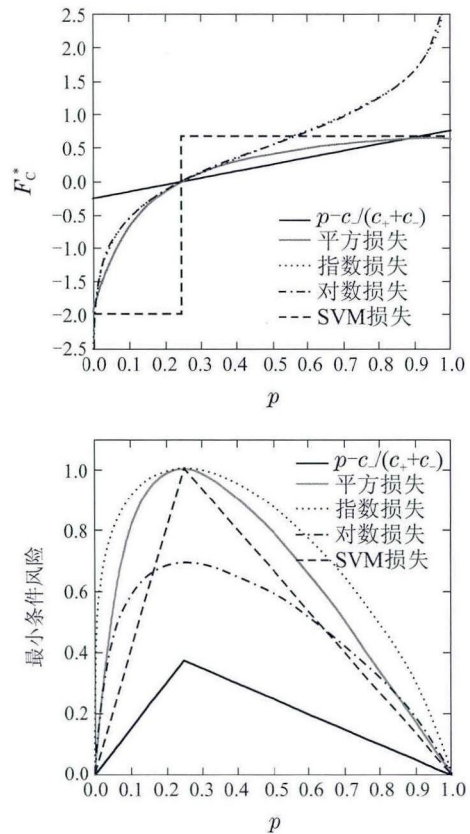
#### 4.2 分类代价在损失函数内(Classification cost inside of loss function)

这种类型的代价敏感损失将分类代价引入损失函数内部,将其与判决函数 $F$ 相乘,采用这种策略的算法较少,典型代表有文献[7,13]等.表3给出了各个代价敏感损失此时的最优决策 $F_C^*(x)$ 和最优决策处的条件代价敏感风险 $E_Y[L(yc_y F_C^*(x))|x]$ ,并绘制为图4(取

$c_+ = 1.5, c_- = 0.5$ ).可看出,4种代价敏感损失均满足准则1和准则2,即最优分类器 $\text{sgn}(F_C^*(x))$ 为Bayes分类器,最优解处的条件代价敏感风险在Bayes分类边界 $\{x|p(x) = \frac{1}{4}\}$ 处取得最大值.

表3 代价敏感损失 $L(yc_y F(x))$ Table 3 Cost-sensitive loss  $L(yc_y F(x))$ 

形式 $L(yc_y F(x))$	最优解 $F_C^*(x)$
$(1 - yc_y F(x))^2$	$\frac{(c_+ + c_-)p(x) - c_-}{(c_+^2 - c_-^2)p(x) + c_-^2}$
$e^{-yc_y F(x)}$	$\frac{1}{c_+ + c_-} \ln \frac{c_+ p(x)}{c_- (1-p(x))}$
$\ln(1 + e^{-2yc_y F(x)})$	$\frac{1}{c_+ + c_-} \ln \frac{c_+ p(x)}{c_- (1-p(x))}$
$\max((1 - yc_y F(x)), 0)$	$\begin{cases} \frac{1}{c_+}, & p(x) \geq \frac{c_-}{c_+ + c_-}, \\ -\frac{1}{c_-}, & \text{其他.} \end{cases}$

图4 代价敏感损失 $L(yc_y F(x))$ 对应的最优解 $F_C^*(x)$ 和最优解处的条件风险 $E_Y[L(yc_y F_C^*(x))|x]$ Fig. 4 Optimal solution  $F_C^*(x)$  and its conditional risk

$E_Y[L(yc_y F_C^*(x))|x]$  with cost-sensitive loss function  $L(yc_y F(x))$

## 5 实验(Experiments)

### 5.1 学习算法(Learning method)

为比较不同损失函数的性能,采用泛函空间的梯度下降法拟合判决函数 $F$ .构造弱分类器集合,每次

迭代从中选择出与前条件风险负梯度拟合最好的弱分类器, 计算其最优步长, 最终判决函数是所有弱分类器的加权组合<sup>[4]</sup>. 算法具体流程如下:

输入: 训练集 $\{(x_i, y_i)\}_{i=1}^N$ , 损失函数 $L(y, F(x))$ , 弱分类器集 $\{h_j(x)\}_{j=1}^M$ , 迭代次数 $T$ .

初始化:  $F(x_i) = 0, i = 1, \dots, N$ .

学习:  $t = 1 : T$ ,

1) 计算所有样本的当前损失负梯度:

$$g_i = -\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}, i = 1, \dots, N.$$

2) 寻找与当前损失负梯度拟合最好的弱分类器:

$$h_t = \arg \min_{h_j} \sum_{i=1}^N g_i \times h_j(x_i).$$

3) 计算最佳迭代步长:

$$\rho_t = \arg \min_{\rho} \sum_{i=1}^N L(y_i, F(x_i) + \rho h_t(x_i)).$$

4) 更新判决函数:

$$F = F + \rho_t h_t.$$

输出: 判决函数 $F^*(x) = \sum_{t=1}^T \rho_t h_t(x)$ .

## 5.2 实验设置(Experiment settings)

实验数据为人工生成的二维数据, 正样本在直径为0.5的圆内随机生成, 负样本在直径为[0.5, 1]的圆环内随机生成, 训练集和测试集均包含300个样本, 其中正样本100个, 负样本200个. 由于数据随机生成, 对于每种算法, 记录在10个不同数据集上的平均结果.

此外, 弱分类器为二维线性分类器, 迭代次数 $T = 50$ , 分类代价设置为 $c_+ = 2, c_- = 1$ .

## 5.3 实验结果(Experimental results)

表4记录不同损失对应的算法在测试集上的分类代价, 其中, ‘N’代表原始损失函数 $L(yF)$ , ‘CS1’和‘CS2’分别为代价敏感损失 $c_y L(yF)$ 和 $L(y c_y F)$ .

表 4 分类代价

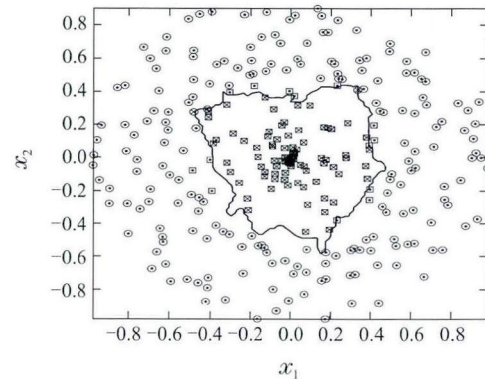
Table 4 Classification cost

损失函数		分类代价/%
平方损失	N	4.43
	CS1	4.38
	CS2	3.91
指数损失	N	3.21
	CS1	3.20
	CS2	3.12
对数损失	N	3.29
	CS1	3.20
	CS2	3.12
SVM损失	N	13.08
	CS1	5.86
	CS2	5.79

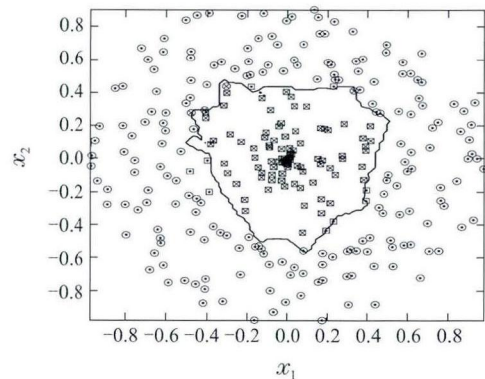
从表4中可看出, 采用代价敏感扩展策略能获得比原始损失更小的分类代价. 对于平方损失、指数损

失、对数损失来说, 代价敏感损失 $c_y L(yF)$ 仅满足第1条设计准则, 代价敏感损失 $L(y c_y F)$ 同时满足两条设计准则, 结果显示后者的分类代价更低. 对于SVM损失来说, 两类代价敏感损失均同时满足两条设计准则, 结果显示两者性能相近, 都能显著降低分类代价.

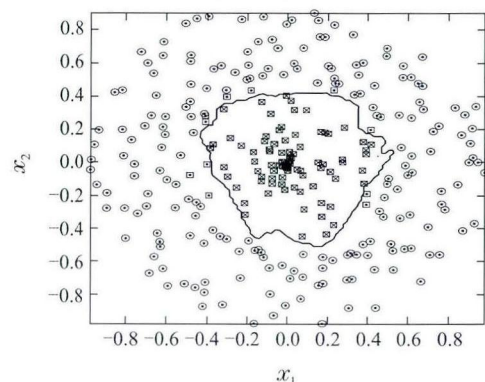
图5以平方损失为例, 直观显示了算法分类结果. 在图5中所示数据集上, 平方损失和其两种代价敏感扩展对应的错误分类的样本个数分别为17, 13, 11, 累积分类代价分别为31, 22, 22.



(a)  $(1 - yF)^2$



(b)  $c_y(1 - yF)^2$



(c)  $(1 - y c_y F)^2$

图 5 平方损失分类结果

Fig. 5 Classification result of square loss

## 6 结论(Conclusions)

本文依据Bayes最优分类提出代价敏感损失函数设计的一般准则,在此准则下,评估两类损失函数代价敏感扩展策略的性能.实验结果表明,较之原始损失函数,满足准则1的代价敏感损失能有效减少分类代价,进一步地,若代价敏感损失同时满足准则2,算法能获得更好的性能.

## 参考文献(References):

- [1] ZHOU Z Z. Cost-sensitive learning [C] // *Modeling Decision for Artificial Intelligence Lecture Notes in Computer Science*. Berlin: Springer, 2011, 6820: 17 – 18.
- [2] SAHIN Y, BULKAN S, DUMAN E. A cost-sensitive decision tree approach for fraud detection [J]. *Expert Systems with Applications*, 2013, 40(15): 5916 – 5923.
- [3] LU S J, LIU L, LU Y, et al. Cost-sensitive neural network classifiers for postcode recognition [J]. *International Journal of Pattern Recognition and Artificial Intelligence*, 2012, 26(5): 1 – 14.
- [4] CAO P, ZHAO D Z, ZAIAANE O. A PSO-based cost-sensitive neural network for imbalanced data classification [C] // *Trends and Applications in Knowledge Discovery and Data Mining Lecture Notes in Computer Science*. Berlin: Springer, 2013, 7867: 452 – 463.
- [5] CAO P, ZHAO D Z, ZAIAANE O. An optimized cost-sensitive SVM for imbalanced data learning [C] // *Advances in Knowledge Discovery and Data Mining Lecture Notes in Computer Science*. Berlin: Springer, 2013, 7819: 280 – 292.
- [6] LOZANO A C, ABE N. Multi-class cost-sensitive boosting with p-norm loss functions [C] // *Proceeding of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: ACM, 2008: 506 – 514.
- [7] MASNADI-SHIRAZI H, VASCONCELOS N. Cost-sensitive boosting [J]. *Pattern Analysis and Machine Intelligence*, 2010, 33(2): 294 – 309.
- [8] 付忠良. 多分类问题代价敏感AdaBoost算法 [J]. *自动化学报*, 2011, 37(8): 973 – 983.  
(FU Zhongliang. Cost-sensitive AdaBoost algorithm for multi-class classification problems [J]. *Acta Automatica Sinica*, 2011, 37(8): 973 – 983.)
- [9] 李秋洁, 茅耀斌, 叶曙光, 等. 代价敏感Boosting算法研究 [J]. *南京理工大学学报*, 2013, 37(1): 19 – 24, 31.  
(LI Qiujie, MAO Yaobin, YE Shuguang, et al. Cost-sensitive boosting algorithms [J]. *Journal of Nanjing University of Science and Technology*, 2013, 37(1): 19 – 24, 31.)
- [10] 韩敏, 穆大芸. 多基于Adaboost算法的回声状态网络预报器 [J]. *控制理论与应用*, 2011, 28(4): 601 – 604.  
(HAN Min, MU Dayun. Improvement of echo state network accuracy with Adaboost [J]. *Control Theory & Applications*, 2011, 28(4): 601 – 604.)
- [11] KINTALI S. Review of boosting: foundations and algorithms by Robert E. Schapire and Yoav Freund [J]. *Sigact News*, 2014, 45(1): 41 – 43.
- [12] ZHANG T. Statistical behavior and consistency of classification methods based on convex risk minimization [J]. *The Annals of Statistics*, 2004, 32(1): 56 – 85.
- [13] KUKAR M Z, KONONENKO L. Cost-sensitive learning with neural networks [C] // *Proceeding of European Conference of Artificial Intelligence*. Amsterdam: IOS Press, 1998: 445 – 449.
- [14] FRIEDMAN. Greedy function approximation: a gradient boosting machine [J]. *The Annals of Statistics*, 2001, 29(5): 1189 – 1232.

## 作者简介:

- 李秋洁 (1983–), 女, 博士, 讲师, 研究方向为不平衡数据分类、视觉目标检测、林业对靶喷雾, E-mail: liqiujie.1@163.com;
- 赵亚琴 (1973–), 女, 博士, 副教授, 研究方向为基于内容的视频检索、林火视频识别, E-mail: yaqinzha@163.com;
- 顾洲 (1973–), 男, 博士, 副教授, 研究方向为网络控制系统、故障检测与容错控制, E-mail: gzh1808@163.com.