基于 Tri-training 算法的多分类信用评级方法

曹欣妍，周　杰
(四川大学数学学院，成都 610064)

摘　要：随着经济的快速发展，信用贷款在企业资金周转中的作用越来越重要. 信用评级是信用贷款发放的基本依据之一. 本文针对实际信用评级中有标签样本数量不足的问题，提出一种基于 Tri-training 算法的多分类信用评级方法，该方法选择支持向量机、决策树和最大熵模型作为基分类器组合. 最后，本文使用真实的信用数据集验证了该方法的实际效果.

关键词：多分类信用评级；半监督学习；Tri-training
中图分类号：O175.5　　**文献标识码**：A　　DOI：10.19907/j.0490-6756.2023.021001

Multi-class credit rating method based on Tri-training algorithm

CAO Xin-Yan, *ZHOU Jie*
(School of Mathematics, Sichuan University, Chengdu 610064, China)

Abstract: Credit loans become more and more important in the capital turnover of corporations with the rapid development of economy. Credit rating is the base of credit loan. In this paper we focus on the problem of insufficient number of label samples in actual credit rating and propose a multi-class credit rating method based on the Tri-training algorithm, which selects the support vector machine, the decision tree and the maximum entropy model as the base classifiers combination. Finally, the performance of the method is verified by using some real credit datasets.

Keywords: Multi-class credit rating; Semi-supervised learning; Tri-training algorithm

1　引　言

随着国民经济的快速发展，信用贷款在企业资金周转中的作用越来越大. 信用评级是金融机构发放信用贷款的一个基本依据. 传统的信用评级方法是专家通过对贷款申请者的各方面要素进行分析评定，例如借款用途、经济能力、偿债记录等. 这种方法依靠人工经验，具有效率低、结果不稳定等缺点. 随着信贷行业的飞速发展，贷款业务种类越来越多、贷款申请者数量越来越大、信用评级的准确率要求越来越高，传统的专家要素分析法已经不再适用.

另一方面，随着数据科学的发展和大数据时代的到来，数据分析与数据挖掘方法被广泛应用于信用评级问题中. 此类信用评级方法需要大量有标签的样本数据，以便对分类器进行训练，保证其准确率和泛化性. 但是，在实际应用场景中，多数贷款申请者没有信用标签，如果仅使用少量有标签申请者信息作为样本数据，就可能因样本量不足导致分类器性能不佳，导致分类效果变差. 因此，如何利用大量无信用标签申请者的信息来提高分类器的性能是一个值得研究的问题.

信用评级方法的研究经历了从基于传统统计学习方法到基于机器学习方法的转变. 1941 年，

收稿日期：2022-04-18
基金项目：国家自然科学基金 (11871357)
作者简介：曹欣妍(1997－)，女，四川成都人，硕士研究生，主要研究方向为大数据分析. E-mail：1216510013@qq.com
通讯作者：周杰. E-mail：jzhou@scu.edu.cn

Durand[1]提出了基于使用数理统计的模型,将线性判别分析法(Linear Discriminant Analysis, LDA)应用于个人信用风险评估中.这是最早将信用评估问题从定性分析转变为定量分析的方法.1970 年,Orgler[2]首先将回归分析应用于消费者贷款的信用评级问题,通过建立多元线性回归模型(Multiple Linear Regression, MLR),对尚未偿还的贷款进行分数评级,以预测风险.同一时期,运筹学的相关方法也被广泛用于信用评估领域.例如,1965 年,Mangasarian[3]第一个提出线性规划(Linear Programming,LP)方法可用于信用评估领域.1984 年,Breiman[4]提出了分类树(Classification Tree)与回归树(Regression Tree)算法.此后,决策树(Decision Tree)被广泛用于信用评估领域.1987 年,Carter 和 Catlett[5]最早将决策树方法应用于信用卡申请者的信用风险评估,得到了决策树方法的分类效果优于简单线性回归模型的结论.1992 年,Salchenberger 等[6]用神经网络预测了储贷危机.1994 年,Rosenberg 和 Gleit[7]分析探讨了神经网络在信用决策中的实际应用和效果.1995 年,Vapnik[8]提出支持向量机(Support Vector Machine, SVM)算法,该算法具有适用于小样本、存在理论全局最优点、泛化能力强等优点,成为个人信用风险评估中研究最多的模型[9].

近年来,更多的半监督学习模型和集成学习模型被应用到信用评估领域.例如,2017 年,Xia 等[10]使用极限梯度提升(Extreme Gradient Boosting,XGBoost)算法完整展示了一套个人风险评估模型建立的流程.2018 年,Tounsi 等[11]比较了 7 种监督分类方法和 5 种集成学习方法,结果表明集成学习方法普遍优于监督学习方法.2019 年,Wei 等[12]通过半监督文本挖掘方法从银行财务报表中提取信息,对银行进行了风险信用评级.2020 年,Li 等[13]使用极限梯度提升法解决信用评估问题,通过实证研究得到极限梯度提升法与逻辑回归等传统模型相比分类效果更好的结论.2020 年,Xiao 等[14]将半监督学习、成本敏感学习、数据分组处理方法和集成学习相结合,提出了基于数据分组处理的成本敏感半监督的选择集合模型.

值得注意的是,多数文献仅研究了二分类信用评级问题.本文则基于分歧的思想,结合信用评级问题数据的特征,选择支持向量机、决策树和最大熵模型作为基分类器,提出一种基于多分类的 Tri-training 算法的多分类信用评级方法,然后使用真实信用数据集验证了该方法的效果.

2　预备知识

2.1　支持向量机

传统的支持向量机分类方法是一种二分类方法,其基本思想为在训练数据集间找到可以有效划分不同类别数据且几何间隔最大的分离超平面.记有标签的信用训练数据集 $D=\{(x_1,y_1),(x_2,y_2),\cdots,(x_N,y_N)\}$,其中 $x_i\in\mathbb{R}^n$ 为第 i 个样本的特征向量,$y_i\in\{-1,+1\}$ 为信用标签.记 $\langle\cdot,\cdot\rangle$ 为向量间内积.假设信用数据集 D 线性可分,使用拉格朗日乘子法求解最大间隔超平面问题约束最优化问题的对偶:

$$\max\sum_{i=1}^{N}\alpha_i-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j\langle x_i,x_j\rangle \tag{1}$$

$$\text{s.t.}\sum_{i=1}^{N}\alpha_i y_i=0,\alpha_i\geqslant 0,i=1,\cdots,N \tag{2}$$

其中,$\alpha=(\alpha_1,\cdots,\alpha_N)$,$\alpha_i\geqslant 0$ 为拉格朗日乘子.对该凸二次规划问题可依次求得最优解 α^*,w^*,b^*,则可得信用等级分类决策函数为:

$$f(x)=\text{sign}\Big(\sum_{i=1}^{N}\alpha_i^*\ y_i\langle x_i,x\rangle+b^*\Big) \tag{3}$$

传统的支持向量机方法只适用于解决二分类问题.本文采用一对一法将支持向量机方法推广到多分类问题,即将多分类问题转化为多个二分类问题.对 M 个信用等级类别,每两类之间都构造一个二分类支持向量机,即有 $M(M-1)/2$ 个二分类模型,采用投票策略进行分类决策.

2.2　决策树

决策树(Decision Tree)是一种监督的学习方法,可用于回归与分类.决策树模型每个叶子节点表示一个类别,非叶子节点是对一个特征的属性判断.从根节点开始,自上而下得到需判断的特征属性,对包含多类样本的节点进行分割,直至某节点仅包含一个类别的样本,即形成最终的叶节点.该叶节点表示类别即为输出的决策结果.

分叉时选择不同特征的顺序会生成不同结构的树.因此,如何确定特征的划分选择十分重要.根据不同的选取准则,主要包含 ID3、C4.5 和 CART 三种生成决策树的算法,分别对应信息增益、信息增益率、基尼指数三种准则.本文使用基尼指数准则.

记信用训练数据集 D 的基尼值为

$$\text{Gini}(D)=\sum_{k=1}^{M}\sum_{k'\neq k}p_k\,p_{k'}=1-\sum_{k=1}^{M}p_k^2 \tag{4}$$

其中，p_k表示第 k 个类别的信用等级在训练集 D 中所占比例.

由表达式可知，基尼值越小，信用数据集纯度越高. 因此，定义特征 A 的基尼指数为：

$$\text{Gini\_Index}(D,A)=\sum_{v=1}^{V}\frac{|D^v|}{|D|}\text{Gini}(D^v) \tag{5}$$

其中，V 为特征 A 可能取值的个数. 由定义可知，在进行划分时，应选择基尼指数较小的特征.

2.3　最大熵模型

最大熵模型（Maximum Entropy Model）基于最大熵原理，可被用于解决多分类问题. 在模型的训练过程中，最大熵模型对有标签的信用样本进行学习，得到模型的参数的估计. 在对无标签样本分类阶段，用已知参数估计的最大熵模型计算无标签信用样本属于不同类别的概率，取概率最大的类别为该样本的信用等级.

假设 $D=(x_1,y_1),(x_2,y_2),\cdots,(x_N,y_N)$为信用训练数据集，其中$x_i\in\mathbb{R}^n$为第 i 个样本的特征向量，n 为特征个数，各样本间相互独立. 记共有 M 个信用等级，向量 $y_i=(y_{i1},\cdots,y_{iM})^{\mathrm{T}}\in\mathbb{R}^M$表示第 i 个样本的信用等级 标签. 第 i 个样本的信用等级为第 p 类可表示为 $y_{ip}=1,p\in1,\cdots,M$,且 $y_{i1}+\cdots+y_{iM}=1$. 记 $W\in\mathbb{R}^{M\times n}$为特征权重矩阵，$b$ 为偏置项，则基于最大熵模型可得第 i 个样本 的信用等级为第 p 个信用类别的概率为：

$$P(y_{ip}=1\mid x_i,W,b)=\frac{\exp(w_p x_i+b_p)}{\sum_{m=1}^{M}\exp(w_m x_i+b_m)} \tag{6}$$

其中，

$$x_i=(x_{i1},\cdots,x_{in})^T \tag{7}$$

$$b=(b_1,\cdots,b_M)^T \tag{8}$$

$$W=\begin{bmatrix}w_1\\ \vdots\\ w_M\end{bmatrix}=\begin{bmatrix}w_{11} & \cdots & w_{1n}\\ \vdots & \ddots & \vdots\\ w_{M1} & \cdots & w_{Mn}\end{bmatrix} \tag{9}$$

在训练过程中可使用极大似然、损失函数梯度下降等方法得到矩阵 W 和向量 b 的估计.

3　基于 Tri-training 算法的多分类信用评级方法

传统的 Tri-training 算法使用的三个基分类器为同一种分类器通过三次自助采样训练生成. 这样得到的三个分类器可能会存在差异较小的问题. 分类器间差异过小时，当存在某个样本被错误分类后，会对后续样本的分类较大的影响，从而影响整个算法的分类性能. 本文基于分歧的思想，结合信用评级问题数据的特征，选择三个不同类别分类器作为基分类器的组合，分别为支持向量机、决策树和最大熵模型，选择原因如下.

（1）信用数据集一般维数较高，即具有较多特征. 支持向量机能较好地处理大型特征空间，同时适用于非线性问题，具有泛化能力强的优点.

（2）信用数据集涉及较多的分类型数据特征，决策树可同时处理分类型数据和数值型数据，且运行速度较快.

（3）最大熵分类器不依赖于对样本分布的主观假设，可避免因对样本数据分布假设错误而导致分类准确率低的问题，使得分类效果稳健.

假设有标签样本集 $L=\{(x_i,y_i)\}_{i=1}^{l}$和无标签样本集 $U=\{x_j\}_{j=l+1}^{l+u}$，其中 x_i为样本的信用特征，$y_i\in label$ 为 第 i 个样本的信用等级标签，$label$ 为信用标签类别集合. 信用集 L 和 U 的样本数分别为 $|L|$ 和 $|U|$. 记三个基分类器h_1，h_2和h_3分别为支持向量机、决策树和最大熵模型，均为从原始有标签数据集 L 中训练得到.

基于 Tri-training 算法的多分类信用评级方法将三个基分类器h_1，h_2和h_3轮流作为主分类器. 在每一轮训练中，使用除主分类器外的另两个分类器对无信用等级标签集的样本进行分类. 将分类结果一致的无标签样本及其伪标签加入主分类器的训练集中，使用扩充的训练集对主分类器进行更新训练. 如此重复迭代，直到三个分类器均不再发生变化即可停止训练. 最后根据投票集成分类器得到最终的信用等级分类结果.

在迭代训练的标记过程中，产生的错误标记样本会对分类结果造成影响. 为了减少分类器扩充训练集中噪声样本的数量，对于标记过程中得到的伪标签样本，考虑根据分类噪声率和每轮训练的分类错误率决定其是否可以加入有标签样本集中用于分类器的更新训练.

由 Angluin 和 Laird[15] 的结果，对于包含 m 个样本的序列σ，样本大小 m 满足不等式

$$m\geqslant\frac{2}{\varepsilon^2(1-2\eta)}\ln\left(\frac{2N}{\delta}\right) \tag{10}$$

其中，ε 为分类错误率的上界；$\eta(<0.5)$为分类噪声率的上界；N 为假设个数；δ 为置信度.

为方便计算，令 $c=2\mu\ln\left(\frac{2N}{\delta}\right)$，$u=\frac{c}{\varepsilon^2}$. 其中 μ

使(10)式等号成立,则有

$$m=\frac{c}{\varepsilon^2(1-2\eta)} \tag{11}$$

$$u=\frac{c}{\varepsilon^2}=m(1-2\eta^2) \tag{12}$$

在分类器的更新训练中,假设L^{t-1}和L^t分别表示第$t-1$轮迭代训练和第t轮迭代训练中获得伪标签并被添加到主分类器中的信用样本,则主分类器第$t-1$轮迭代和第t轮迭代对应的训练集样本数分别为$|L\cup L^{t-1}|$和$|L\cup L^t|$. 假设η_L为初始有信用标签集L的分类噪声率,第t轮训练的主分类器为h_i,$\hat{e}_i$表示第t轮训练分类器组h_j和h_k分类错误率的上界,其中$j,k\neq i$. 由此可得第t轮训练的分类错误率为

$$\eta^t=\frac{\eta_L|L|+\hat{e}_i^t|L^t|}{|L\cup L^t|} \tag{13}$$

由(12)式可知u与$1/\varepsilon^2$成正比. 假设$\hat{e}_i^t,\hat{e}_i^{t-1}<0.5$,同时考虑到一般情况下$\eta_L$较小,可知在更新训练中,若$L^t$和$\hat{e}_i^t$满足下式,则主分类器可通过更新训练提高分类性能:

$$0<\frac{\hat{e}_i^t}{\hat{e}_i^{t-1}}<\frac{|L^{t-1}|}{|L^t|}<1 \tag{14}$$

即可由条件(14)式判断在第t轮训练中由分类器h_j和h_k给出相同信用等级标签的无标签样本是否应放入主分类器的扩充训练集中.

由上所述,以支持向量机分类器h_1作为主分类器为例,算法第t轮训练的具体步骤如下.

(1) 使用经过$t-1$轮更新后的决策树分类器h_2和最大熵分类器h_3对无标签数据集U中样本x的信用等级进行标注,将信用等级标签一致的样本加入集合L_i中,即$L_1^t=\{(x,y)|h_2(x)=h_3(x)=y,x\in U\}$.

(2) 因假设有标签样本与无标签样本服从同一概率分布,故以分类器h_2和h_3在第t轮对原始有标签集L分类错误率估算分类器h_2和h_3对集合L_1^t的分类错误率$e_1^t=ME(h_2\&h_3)$,具体为

$$ME(h_2\&h_3)=\frac{|\{(x,y)\in L|h_2(x)=h_3(x)\neq y\}|}{|\{(x,y)\in L|h_2(x)=h_3(x)\}|} \tag{15}$$

(3) 由(14)式计算得到的e_1^t,若满足条件(14)式,即$e_1^t|L_1^t|<e_1^{t-1}|L_1^{t-1}|$,则$L_1^t$中样本可加入主分类器的更新训练集中,即支持向量机分类器h_1第t轮的训练集为$L\cup L_1^t$. 需要注意的是,如果从U中得到的$|L_1^t|$过大会导致(14)式不成立. 因此考虑对集合L_1^t下采样,得到大小为s的集合使得

$$e_1^t|L_1^t|<e_1^{t-1}|L_1^{t-1}|$$

成立,且$s>|L_1^{t-1}|$. 记$\langle\cdot\rangle$表示向上取整,$\langle\cdot\rangle$表示向下取整. s可由下式计算得到:

$$s=\langle\frac{e_1^{t-1}|L_1^{t-1}|}{e_1^t}-1\rangle \tag{16}$$

其中

$$|L_1^{t-1}|>\frac{e_1^t}{e_1^{t-1}-e_1^t} \tag{17}$$

记下采样得到集合为 Subsample(L_1^t,s),则该情况下支持向量机分类器h_1第t轮的训练集为$L\cup$Subsample(L_1^t,s).

(4) 支持向量机分类器h_1使用所得扩充训练集进行更新训练. 完成第t轮训练后,将$|L_1^t|$中样本放回无信用等级标签集U中. 判断进入下一轮更新或输出信用等级分类结果.

完整算法如算法 1 所示.

算法 1　基于 Tri-training 的多分类信用评级算法

输入: L:有标签样本集; U:无标签样本集;
$Learn^1$:支持向量机; $Learn^2$:决策树;
$Learn^3$:最大熵模型

1: for $i\in\{1,2,3\}$ do
2:　对 L 自主采样得到 S_i
3:　由 $Learn^i(S_i)$ 训练得到分类器 h_i
4:　初始化e_i'为 0.5,l_i'为 0
5: end for
6: while $h_i,i\in\{1,2,3\}$ 都不再改变 do
7:　for $i\in\{1,2,3\}$ do
8:　　$L_i\leftarrow\varnothing$
9:　　$update_i\leftarrow FALSE$
10:　　$e_i\leftarrow ME(h_2\&h_3),j,k\neq i$
11:　　if $e_i<e_i'$ then
12:　　for 每个 $x\in U$ do
13:　　if $h_j(x)=h_k(x),j,k\neq i$ then
14:　　将$\{x,h_j(x)\}$加入L_i
15:　　end if
16:　　end for
17:　　if $l_i'=0$ then
18:　　$l_i'\leftarrow\langle\frac{e_i}{e_i'-e_i}+1\rangle$
19:　　end if
20:　　if $l_i'<|L_i|$ then
21:　　if $e_i|L_i|<e_i'l_i'$ then
22:　　$update_i\leftarrow TRUE$

```
        else if l'_i > e_i/(e'_i − e_i) then
          计算 s
          L_i ← Subsample(L_i, s)
          update_i ← TRUE
        end if
      end if
    end if
  end for
  for i ∈ {1,2,3} do
    if update_i = TRUE then
      使用 L ∪ L_i 更新训练分类器 h_i
      更新 e'_i ← e_i; l'_i ← |L_i|
    end if
  end for
end while
```

输出：$h(x) \leftarrow argma\,x_{y\in label}\sum_{i=1}^{3}1\{h_{i(x)}=y\}$

4　结果与分析

4.1　实验数据集

本文使用数据集来源于 2020 年全国大学生数学建模竞赛（http://www.mcm.edu.cn/）. 数据集为某银行的贷款申请企业的进项发票信息和销项发票信息，每条发票数据有发票号码、开票日期、销方单位代号、金额、税额、价税合计及发票状态七项信息. 数据集共有 425 家企业的发票信息，其中包括 123 家已经评定信用等级的企业和 302 家无信用等级企业，分别为有标签数据集和无标签数据集，信用等级分为 A，B，C，D 四个类别.

记第 i 个企业的进项发票总个数为α_i，其中作废发票个数为a_i；销项发票总个数为β_i，其中作废发票个数为b_i. 第 i 个企业进项有效发票总金额为φ_i，第 k 个月进项有效发票金额为$\varphi_i(k)$；销项有效发票总金额为ψ_i，第 k 个月销项有效发票金额为$\psi_i(k)$，正销项有效发票金额为γ_i，负销项有效发票金额为ρ_i. 第 i 个企业有发票记录的月份总数为λ_i.

根据以上数据信息计算得到六个信用评级特征，记第 i 个企业的特征向量为$x_i=(x_{i1},x_{i2},x_{i3},x_{i4},x_{i5},x_{i6})$，其中$x_{ij}$表示第 i 个企业的第 j 个特征. 各特征x_{ij}的含义及计算方法分别如下.

（1）第 i 个企业进项发票作废率

$$x_{i1}=\frac{a_i}{\alpha_i} \tag{18}$$

（2）第 i 个企业销项发票作废率

$$x_{i2}=\frac{b_i}{\beta_i} \tag{19}$$

（3）第 i 个企业的月平均营业额

$$x_{i3}=\frac{\psi_i}{\lambda_i} \tag{20}$$

（4）第 i 个企业的月平均利润

$$x_{i4}=\frac{\psi_i-\varphi_i}{\lambda_i} \tag{21}$$

（5）第 i 个企业的月平均净利润率

$$x_{i5}=\frac{1}{\lambda_i}\sum_{k=1}^{\lambda_i}\frac{\psi_i(k)-\varphi_i(k)}{\varphi_i(k)} \tag{22}$$

（6）第 i 个企业的销售退回率

$$x_{i6}=\frac{\rho_i}{\gamma_i} \tag{23}$$

计算得到各特征量后，对数据进行归一化，记归一化后第 i 个企业的第 j 个特征为x'_{ij}，采用最大最小归一化公式

$$x'_{ij}=\frac{2(x_{ij}-x_j^{\min})}{x_j^{\max}-x_j^{\min}}-1 \tag{24}$$

其中，$x'_{ij}\in[-1,1]$，$x_j^{\min}$和$x_j^{\max}$分别表示所有企业第 j 个特征中的最小值和最大值.

4.2　评价指标

本文对实验结果的评估采用最直观的评价指标：分类准确率（Accuracy），计算公式如下.

$$\text{Accuracy}=\frac{T_a}{T} \tag{25}$$

其中，T_a为测试集中分类正确的样本数量；T 为测试集样本总数.

4.3　实验结果分析

在数据集中取 73 个样本为有标签训练集，50 个样本为测试集，实验结果如表 1 所示. 由实验结果对比可知，半监督算法在数据集上的分类正确率较监督算法均有提高，且本文提出的使用决策树、支持向量机和最大熵分类器作为基分类器组合的 Tri-training 信用评级算法与使用单一种类基分类器的 Tri-training 算法相比，准确率都有较明显的提升.

在该多分类数据集上，本文提出的基于 Tri-training 算法在的信用评级方法相比于监督算法分类准确率平均提升了 18.74%，且可达到近 90% 的分类正确率，表明所选的分类器组合在信用评级问题上有较好的效果.

表 1　算法在数据集上的分类准确率

Tab. 1　Classification accuracy of the method on the dataset

分类模型	分类准确率/%
决策树	70.0
支持向量机	64.0
最大熵	72.0
Tri-training-DT	78.0
Tri-training-SVC	68.0
Tri-training-MaxEnt	80.0
Tri-training-3	89.0

5　结　论

本文在仅有少量有信用标签样本的应用背景下,提出了 基于 Tri-training 算法的多分类信用评级方法,采用三种不同类别分类器的组合作为基分类器,分别为决策树、支持向量机、最大熵模型. 实验结果表明,本文提出的信用评级方法可有效提高分类性能.

多分类的信用评级可用于更多的实际应用中,例如贷款方可根据信用等级决定发放贷款的额度、最大放贷金额及贷款利率等.

参考文献:

[1] Durand D. Risk elements in consumer installment financing [M]. Cambridge: National Bureau of Economic Research, 1941.

[2] Eisenbeis R A. Problems in applying discriminant analysis in credit scoring models [J]. J Bank Finan, 1978, 2: 206.

[3] Mangasarian O L. Linear, nonlinear separation of patterns by linear programming [J]. Oper Res, 1965, 13: 444.

[4] Breiman L. Classification and regression trees [M]. Boca Raton: CRC Press, 1984.

[5] Carter C, Catlett J. Assessing credit card applications using machine learning [J]. IEEE Intell Syst, 1987, 2: 71.

[6] Salchenberger L M, Cinar E M, Lash N A. Neural networks: a new tool for predicting thrift failures [J]. Decis Sci, 1992, 23: 899.

[7] Rosenberg E, Gleit A. Quantitative methods in credit management: a survey [J]. Oper Res, 1994, 42: 589.

[8] Vapnik V. The nature of statistical learning theory [M]. Berlin: Springer, 1999.

[9] Louzada F, Ara A, Fernandes G B. Classification methods applied to credit scoring: systematic review and overall comparison [J]. Surv Oper Res Manag Sci, 2016, 21: 117.

[10] Xia Y, Liu C, Li Y. A boosted decision tree approach using Bayesian hyper-parameter optimization for credit scoring [J]. Expert Syst Appl, 2017, 78: 225.

[11] Tounsi Y, Hassouni L, Anoun H. An enhanced comparative assessment of ensemble learning for credit scoring [J]. Int J Mach Learn Comput, 2018, 8: 409.

[12] Wei L, Li G, Zhu X, *et al*. Discovering bank risk factors from financial statements based on a new semi-supervised text mining algorithm [J]. Account Finan, 2019, 59: 1519.

[13] Li H, Cao Y, Li S. XGBoost model and its application to personal credit evaluation [J]. IEEE Intell Syst, 2020, 35: 52.

[14] Xiao J, Zhou X, Zhong Y, *et al*. Cost-sensitivesemisupervised selective ensemble model for customer credit scoring [J]. Knowl Based Syst, 2020, 189: 105118.

引用本文格式:

中　文: 曹欣妍, 周杰. 基于 Tri-training 算法的多分类信用评级方法[J]. 四川大学学报: 自然科学版, 2023, 60: 021001.

英　文: Cao X Y, Zhou J. Multi-class credit rating method based on Tri-training algorithm [J]. J Sichuan Univ: Nat Sci Ed, 2023, 60: 021001.