

# 互联网金融背景下的信用卡审批的预测

彭 宇

贵州大学数学与统计学院, 贵州 贵阳

收稿日期: 2024年9月4日; 录用日期: 2024年9月29日; 发布日期: 2024年11月29日

## 摘 要

本文基于UCI机器学习库中的一个信用卡审批的数据, 以是否同意审批为响应变量, 以15个描述客户信息的离散和连续指标作为解释变量建立预测模型。提出了一种改进的弹性网损失支持向量机(QCaenSVM)预测模型, 旨在提高信用卡审批行业中审批周期长且结果不一致以及数据利用不充分的不足之处。QCaenSVM模型通过融合弹性网损失函数和分位数的概念, 优化了传统支持向量机的性能。该模型在含噪声数据环境下具有较好的表现性能, 并有效处理了数据中的不确定性。在应用于信用卡的预测实践中, QCaenSVM成功识别出更可能选择审批的客户, 明显提高了预测效果, 为相关部门和客户群体提供了有力工具。

## 关键词

分位数盖帽非对称弹性网支持向量机, 信用卡审批, 预测模型

# Prediction of Credit Card Approval in the Context of Internet Finance

Yu Peng

School of Mathematics and Statistics, Guizhou University, Guiyang Guizhou

Received: Sep. 4<sup>th</sup>, 2024; accepted: Sep. 29<sup>th</sup>, 2024; published: Nov. 29<sup>th</sup>, 2024

## Abstract

In this paper, we build a predictive model based on data from a credit card approval in the UCI Machine Learning Library, with whether or not to agree to the approval as the response variable, and 15 discrete and continuous metrics describing the customer information as the explanatory variables. An improved Quantile-Capped Asymmetric Elastic Net Support Vector Machine (QCaenSVM) prediction model is proposed to improve the shortcomings of the credit card approval industry in terms of long approval cycles with inconsistent results and underutilization of data. The QCaenSVM

文章引用: 彭宇. 互联网金融背景下的信用卡审批的预测[J]. 电子商务评论, 2024, 13(4): 6340-6349.

DOI: 10.12677/ec.2024.1341873

model optimizes the performance of the traditional support vector machine by incorporating the concepts of the elastic net loss function and quartiles. The model has better performance in noisy data-containing environments and effectively handles uncertainties in the data. In the prediction practice applied to credit cards, QCaenSVM successfully identifies customers who are more likely to choose approval, significantly improves the prediction effect, and provides a powerful tool for relevant departments and customer groups.

## Keywords

Quantile-Capped Asymmetric Elastic Net Support Vector Machines, Credit Card Application, Predictive Models

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

在信用卡申请审批领域，审批部门的征信审核岗位员工起着关键作用。他们通过评估申请者的信用风险水平，旨在拒绝高风险客户的申请，同时批准低风险高回报客户的信用卡申请，以维护信用卡部门的盈利能力，并降低持卡人违约风险。然而，随着信用卡市场竞争的激烈化，各银行信用卡中心普遍将规模 and 市场份额作为业绩考核指标。在这一背景下，审批环节的效率和质量显得尤为重要。目前，依赖征信审核岗位员工的经验来判断申请者是否符合信用卡申请条件存在较大主观性和不稳定性，审批效率和质量无法充分保障。随着信用卡业务的蓬勃发展、同业竞争的日益激烈以及信息技术在银行业中的广泛应用，银行业务的经营理念 and 方式正在发生重大变革。网络技术和电子商务的快速发展促使银行业开始注重数据挖掘、机器学习等审批方法的应用。在对信用卡申请的监督方面，为了降低信用卡管理风险，国内许多学者进行了大量研究。刘继海，陈晓剑[1]通过引入 SVM 的理论，建立了信用卡申请管理的评分模型用于挑选优质的客户，大大节约了人力资源。随着更多外资银行大量进入中国市场，各家银行之间对优质客户的争夺以及对市场份额的拓展也愈演愈烈，各家银行该如何赢得竞争，依赖于能够从海量的顾客数据中发掘出潜在的宝贵信息，增强顾客事务处理能力深层次地发掘顾客和合理定位顾客，从而达到高效管理的目的。顾客是商业银行争夺的重点，银行利润归根到底来自于顾客，为了增强盈利能力就必须重视对顾客价值的挖掘。从顾客需求角度看，不同种类的顾客需求存在差异，要使不同顾客对同一家银行均满意，需要银行有针对性地提供满足顾客需要的产品与服务，而要想满足这一多样化异质性需求，必须根据不同标准细分客户群体；在顾客价值层面上，不同顾客能给银行带来的价值也不一样，银行需要了解什么顾客对银行有高价值，什么顾客对银行忠诚，银行潜在客户有哪些、成长性最强、最易流失的客户有哪些，银行必须将其客户细分；从银行资源与能力角度出发，如何优化配置不同客户的有限资源是每一家银行必须思考的问题，而商业银行的服务资源是有限的，顾客的数量是很多的，因此在顾客管理中十分需要进行顾客的统计与分析。如何建立合理的模型预测客户是否认购定期存款十分重要。

Vapnik 和 Cortes 基于结构风险最小原则提出的一种全新的分类方法[2]：支持向量机(Support vector machine, SVM)。随着 SVM 理论与算法日趋成熟，在图像分类，文本分类和医疗疾病都有很好的应用，更多相关应用场景可以参考相关文献[3]-[5]。Huang 等人在结合分位数思想下，基于最小分位数距离最大的情况下提出了弹球损失函数，将弹球损失函数应用于 SVM，提出了基于弹球损失函数的支持向量机

(PinSVM) [6]。同时研究了 PinSVM 的性质, 包括噪声不敏感性, 鲁棒性和误分类误差最小化。弹球损失会对分类正确的点进行惩罚, 因而弹球损失对决策边界附近的噪声具有不敏感性。由于弹球损失在零点不可微而导致计算量增加, Huang 等人[7]考虑了非对称的最小二乘损失, 提出了非对称最小二乘支持向量机(aLSSVM), 非对称最小二乘损失和弹球损失的关系就如同分位数回归和期望回归一样。Qi 和 Yang [8]结合非对称最小二乘函数, 提出一种非对称几何双参数边缘支持向量机, 对于标签和特征噪声鲁棒性较好, 同时模型对离群点不敏感以及费歇尔一致的。

本文中, 结合分位数和截断损失函数的思想, 在 Qi [9]的基础上提出了一种分位数盖帽非对称弹性网支持向量机(QCaenSVM)模型应用于信用卡审批的预测模型, 提高预测精度, 减少噪声对决策超平面的影响, 使 QCaenSVM 在处理噪声数据时具有稳健性。所提的分位数盖帽非对称弹性网支持向量机对异常值具有鲁棒性, 为信用卡审批提供更加精确的判断, 减少的审批所需的人工时间。

## 2. 预备知识

在本章中, 对 SVM、PinSVM、aLSSVM 的理论进行简要的介绍, 在一个二分类问题中, 假设输入空间中训练样本集为  $\{x_i, y_i\}_{i=1}^n$ , 其中  $x_i \in \mathbb{R}^m, y_i \in \{-1, 1\}$ 。

### 2.1. 基于弹球损失函数的支持向量机

弹球损失在回归分析中得到了深入的研究和广泛的应用, 但尚未用于分类, Huang 等人[6]为了克服 SVM 对决策超平面附近的特征噪声敏感提出了基于弹球损失函数的支持向量机。首先简单介绍弹球损失 (Pinball loss) 函数:

$$l_{pin}(u) = \begin{cases} u, & u \geq 0 \\ -\tau u, & u < 0 \end{cases} \quad (1)$$

其中,  $0 \leq \tau \leq 1$  为自定义参数, 它控制重采样的稳定性水平, 图 1 为参数  $\tau = 0.3, 0.5, 0.7$  时的函数图像。

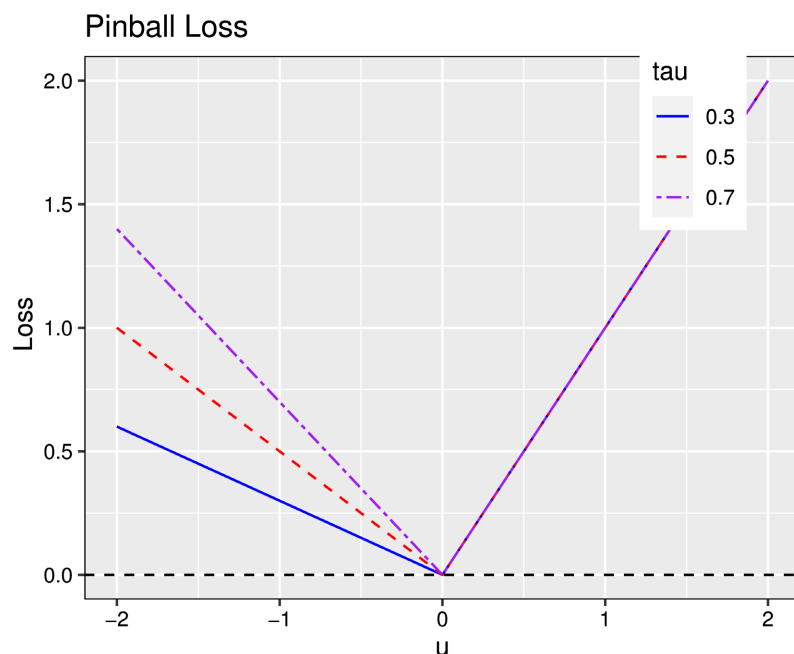


Figure 1. Image of the pinball loss function when  $\tau$  taking different values

图 1. 不同的  $\tau$  取值时弹球损失函数图像

基于弹球损失函数的支持向量机(PinSVM)模型定义如下:

$$\min_{w,b} \frac{1}{2} \|w\|^2 + c \sum_{i=1}^n l_{pin} (1 - y_i (w^T x_i + b)). \quad (2)$$

对于问题(2), 可改写为如下形式:

$$\begin{aligned} \min_{w,b,\xi} \quad & \frac{1}{2} \|w\|^2 + c \sum_{i=1}^n \xi_i \\ \text{s.t.} \quad & y_i (w^T x_i + b) \geq 1 - \xi_i \\ & y_i (w^T x_i + b) \leq 1 + \frac{1}{\tau} \xi_i \\ & \xi_i \geq 0, i = 1, \dots, n \end{aligned} \quad (3)$$

随着参数  $\tau$  的值增加, 正确分类的样本点的权重变大, 因此边距宽度变大。所以类边界附近的点在决定最优决策超平面中变得不那么重要。对于 PinSVM, 在正确分类的样本点上给予惩罚带来了对决策边界周围的噪声的不敏感性, 特征噪声的影响被削弱, 使得 PinSVM 对特征噪声不敏感。当  $\tau = 0$  时, (3) 的约束条件变为和标准的支持向量机一样, 这说明弹球损失是合页损失的一种特例。

## 2.2. 基于非对称最小二乘损失的支持向量机

非对称最小二乘损失函数结合了分位数的思想, 分位数就是将一个概率分布切分为相同概率的连续区间的切分点, 公式如下:

$$P(X < x_a) = a, (0 < a < 1) \quad (4)$$

由于弹球损失是非光滑的, 优化它比平滑损失需要更多的时间, Newey 和 Efron [10] 在对期望回归的统计性质探讨下, 得到了如下的非对称最小二乘损失函数, 其定义如下:

$$l_{aLS}(u) = \begin{cases} \tau u^2, & u \geq 0 \\ (1-\tau)u^2, & u < 0 \end{cases} \quad (5)$$

当  $\tau = 1$  时, 上式变为合页损失的平方形式, 也就是为  $L_2$  损失。

把它与支持向量机结合, 即为非对称最小二乘支持向量机(aLSSVM):

$$\min_{w,b} \frac{1}{2} \|w\|^2 + \frac{c}{2} \sum_{i=1}^n l_{aLS} (1 - y_i (w^T x_i + b)) \quad (6)$$

问题(6)可以改为如下形式,

$$\begin{aligned} \min_{w,b,\xi} \quad & \frac{1}{2} \|w\|^2 + \frac{c}{2} \sum_{i=1}^n \xi_i^2 \\ \text{s.t.} \quad & y_i (w^T x_i + b) \geq 1 - \frac{1}{\tau} \xi_i \\ & y_i (w^T x_i + b) \leq 1 + \frac{1}{1-\tau} \xi_i \\ & \xi_i \geq 0, i = 1, \dots, n \end{aligned} \quad (7)$$

当  $\tau = 1$  约束变为和合页支持向量机的一样, 但损失为合页损失的平方, 这也说明, 非对称最小二乘支持向量机为合页支持向量机的特例。

### 3. 分位数盖帽非对称弹性网支持向量机(QCaenSVM)

#### 3.1. 分位数盖帽非对称弹性网支持向量机介绍

结合分位数和截断损失函数的优点下, 本节提出了分位数盖帽非对称弹性网损失, 把它与支持向量机结合, 提出了 QCaenSVM。损失函数如下:

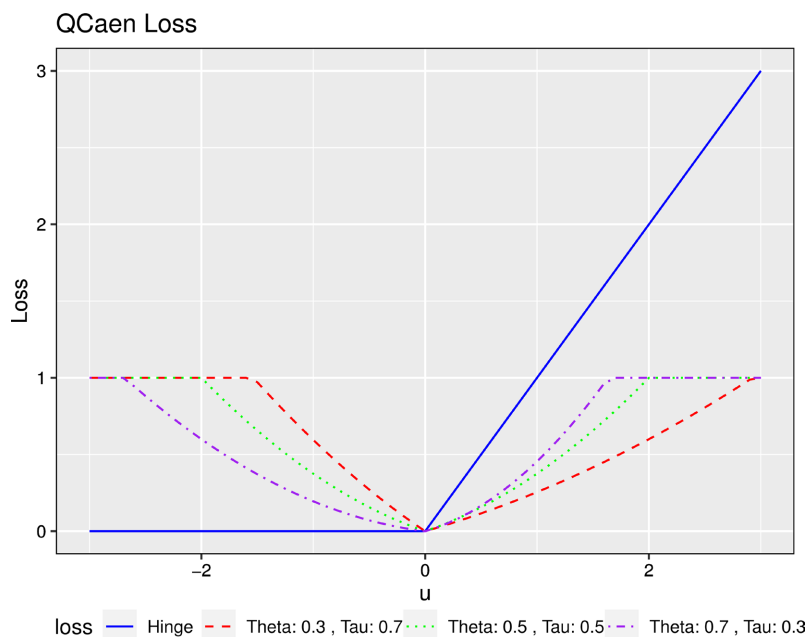
$$l_{QCaen}(u) = \begin{cases} s, & u > s \\ \tau \left( \frac{\theta}{2} u^2 + (1-\theta)u \right), & 0 \leq u \leq s \\ (1-\tau) \left( \frac{\theta}{2} u^2 - (1-\theta)u \right), & -s \leq u < 0 \\ s, & u < -s \end{cases} \quad (8)$$

$\theta \in [0, 1]$ , 它是控制损失函数在  $L_1$  范数和  $L_2$  范数的调整参数, 当  $\theta = 1$ , 函数类似为  $L_2$  范数, 当  $\theta = 0$ , 函数类似为  $L_1$  范数。  $s > 0$  是一个阈值参数。分位数盖帽非对称弹性网支持向量机(QCaenSVM)如下:

$$\min_{w, b} \frac{1}{2} \|w\|^2 + \frac{c}{n} \sum_{i=1}^n l_{QCaen}(1 - y_i (w^T x_i + b)) \quad (9)$$

其中  $C > 0$  为自定义参数。

由于所提损失函数的非凸, 不对称和有界性, 它可以减少标签噪声的干扰以及对特征噪声重采样的稳定性, 对于异常值和有缺失数据的数据集有很好的鲁棒性, 所以所提的 QCaenSVM 对于信用卡审批有很好的表现性能。QCaen 损失函数的图像见图 2, 它的有界性对于噪声很好的鲁棒性, 非对称可以对不同的结果给予不同的惩罚。



**Figure 2.** Image of the quantile-capped asymmetric elastic net loss function when  $\tau$  and  $\theta$  taking different values  
**图 2.** 不同的  $\tau$  和  $\theta$  取值时分位数盖帽非对称弹性网损失函数图像

对于非线性的问题, 我们可以将样本从原始空间映射到高维空间, 即加入核函数, 由核函数构成的

超平面为:  $w^T \phi(x) = 0$ 。其中  $\phi(x)$  是将  $x$  映射到高维的希尔伯特空间中的特征向量。在特征空间上的模型:

$$f(x) = w^T \phi(x) + b \quad (10)$$

在实际中, 由于特征空间的维数可能很高, 导致计算  $\phi(x_i)^T \phi(x_j)$  十分困难, 因而我们通常考虑如下形式, 可得:

$$\kappa(x_i, x_j) = \langle \phi(x_i), \phi(x_j) \rangle = \phi(x_i)^T \phi(x_j) \quad (11)$$

即在  $x_i, x_j$  特征空间上的内积等于它们在原始空间中经过  $\kappa(.,.)$  计算的结果。

### 3.2. 算法

由于损失函数是非凸的, 优化问题为一个非凸优化, 由于非凸优化不太容易解决, 本文主要采用 DC 算法进行求解。DC 算法是将一个函数拆为两个凸函数之差, 进而转化为一列凸子问题。原问题的优化形式如下式:

$$l_{QCaen}(u) = l_{QCaen1}(u) - l_{QCaen2}(u) \quad (12)$$

$$\text{其中 } l_{QCaen2}(u) = \max(l_{QCaen1}(u) - s, 0), \quad l_{QCaen1}(u) = \begin{cases} \tau \left( \frac{\theta}{2} u^2 + (1-\theta)u \right), & u \geq 0 \\ (1-\tau) \left( \frac{\theta}{2} u^2 - (1-\theta)u \right), & u < 0 \end{cases}, \text{把上面 2 式带入(9)式,}$$

可以化解为如下形式:

$$\min_{w,b} \frac{1}{2} \|w\|^2 + \underbrace{\frac{c}{n} \sum_{i=1}^n l_{AQCl} \left( 1 - y_i (w^T x_i + b) \right)}_g - \underbrace{\frac{c}{n} \sum_{i=1}^n l_{AQCl} \left( 1 - y_i (w^T x_i + b) \right)}_h \quad (13)$$

DC 算法需要计算微分, 但由于函数不是处处可微的, 便用次微分来替代, 设  $u^* = 1 - (y_i w^T x_i + b)$ :

$$\nabla l_{QCaen2}(u) = \begin{cases} (1-\tau) (\theta x_i x_i^T w + (1-2\theta) y_i x_i), & u^* < u_1 \\ 0, & u_1 < u^* < u_2 \\ \tau (\theta x_i x_i^T w - y_i x_i), & u^* > u_2 \end{cases} \quad (14)$$

根据  $l_{QCaen1}(u)$  和  $l_{QCaen2}(u)$ , 可以推出  $u_1, u_2$ :

$$u_1 = \frac{\theta - 1 + \sqrt{(1-\theta)^2 + \frac{2\theta s}{\tau}}}{\theta}, u_2 = \frac{1 - \theta + \sqrt{(1-\theta)^2 + \frac{2\theta s}{1-\tau}}}{\theta} \quad (15)$$

依次进行下面的计算, 直到收敛:

$$X^{(k+1)} = \underset{X}{\operatorname{argmin}} g(X) - \nabla h(X^{(k)})^T (X - X^{(k)}) \quad (16)$$

由(16)式得到  $w^{(k)}$  后, 我们采用随机梯度下降算法来提高算法的拓展性。令  $A_t \subset \{1, 2, \dots, n\}$ ,  $|A_t| = h$  是  $k$  个样本的子集, 我们可以得到如下的近似目标函数:

$$F(v; A_t) = \frac{1}{2} \|v\|_2^2 + \frac{c}{h} \sum_{i \in A_t} l_{QCaen1} \left( 1 - y_i v^T x_i \right) - \frac{c}{h} \sum_{i \in A_t} \nabla l_{QCaen2} \left( w^{(k)} \right)_i^T v, v \in \mathbb{R}^m \quad (17)$$

关于  $v$  在  $v^{(t)}$  处的次梯度  $F(v; A_t)$  有如下形式:

$$\nabla F\left(v^{(t)}\right)=v^{(t)}+\frac{c}{h} \sum_{i \in A_t} \sigma\left(v^{(t)}\right)_i-\frac{c}{h} \sum_{i \in A_t} \nabla l_{A Q C 2}\left(w^{(k)}\right)_i \quad (18)$$

$v^{(t)}$  是第  $t$  次的最优值,  $\sigma\left(v^{(t)}\right)_i$  有如下形式:

$$\sigma\left(v^{(t)}\right)_i=\left\{\begin{array}{ll} \tau\left(\theta x_i x_i^T w-y_i x_i\right), & u^* \geq 0 \\ (1-\tau)\left(\theta x_i x_i^T w+(1-2 \theta) y_i x_i\right), & u^* < 0 \end{array}\right. \quad (19)$$

$v^{(t)}$  的更新有如下形式:

$$v^{(t+1)}=v^{(t)}-\eta_t \nabla F\left(v^{(t)}\right), \eta_t=\frac{c}{t} \quad (20)$$

其中  $\eta_t$  为步长。

其中  $k\left(x_i, x_j\right)=\phi\left(x_i\right)^T \phi\left(x_j\right)$  是核函数, 本文所用的核函数为高斯核函数:

$$k\left(x_i, x_j\right)=\exp \left(-\frac{\left\|x_i-x_j\right\|^2}{2 \sigma^2}\right), \sigma>0.$$

算法的具体步骤见表 1。

**Table 1.** The process of solving model parameters

**表 1.** 求解模型参数过程

**算法:** 基于随机梯度下降的 DC 算法

**输入:**  $T_1, T_2, eps, w^0, c, s, \tau, \theta, \left\{\left(x_i^T, y_i\right)\right\}_{i=1}^n$

**输出:**  $w^{(k+1)}$

1: 设  $t=0, v^0=w^0$ ;

2: 当  $t < T_2$  时;

3: 随机选择  $A_t \subset \{1, 2, \dots, n\}$ , 计算  $\nabla F\left(v^{(t)}\right)$ ;

4: 设  $v^{(t+1)}=v^{(t)}-\eta_t \nabla F\left(v^{(t)}\right), \eta_t=\frac{c}{t}$ ;

5: 结束;

6: 设  $k=0, w^{(1)}=v^{(t+1)}$ ;

7: 当  $k \leq T_1$  or  $\left\|w^{(k+1)}-w^{(k)}\right\| \geq eps$  时;

8: 设  $t=0, v^{(0)}=w^{(0)}$ ;

9: 当  $t < T_2$  时;

10: 随机选择  $A_t \subset \{1, 2, \dots, n\}$ , 计算  $\nabla F\left(v^{(t)}\right)$ ;

11: 设  $v^{(t+1)}=v^{(t)}-\eta_t \nabla F\left(v^{(t)}\right), \eta_t=\frac{c}{t}$ ;

12: 结束;

13: 设  $w^{(k+1)}=v^{(t+1)}$ ;

14: 结束;

15: 返回  $w^{(k+1)}$ ;

4. 基于 QCaenSVM 的信用卡审批预测模型

4.1. 数据来源与说明

本文的数据来自 UCI 机器学习库(<https://archive.ics.uci.edu/>)公开的 Credit Approval Data Set 数据。本数据聚焦于一个包含 690 位信用卡申请客户详细资料的综合性数据集,该数据集精心设计了 16 个特征变量以全面刻画每位申请者的多维度属性。鉴于数据隐私及商业敏感性的考量,部分变量采用了匿名化处理,即以名义型标签(如字母或特定字符)替代其具体含义,这既保护了个人隐私与企业机密,也确保了数据分析的合规性。尤为关键的是,该数据集的最终响应变量(即目标变量)为第 16 个变量,它明确标识了申请者对于信用卡申请的最终态度——“+”代表积极同意申请,而“-”则表明申请者拒绝了该申请。这一二元分类的设置,为研究信用卡申请审批过程中的决策机制、客户行为模式以及潜在影响因素提供了宝贵的实证基础。数据变量的形式包括有分类、整数和实数。变量具体说明如表 2,表 3 为信用卡审批类别的分布,表 4 是信用卡审批中的缺失数据的统计表格。

Table 2. Data indicator variables  
表 2. 数据指标变量

变量类型	变量名	类型	取值
因变量	V16	分类	-: 不认购 +: 认购
解释变量	V1	分类	a, b
	V2	数值	13.75~80.25
	V3	数值	0~28
	V4	分类	u, y, l, t
	V5	分类	g, p, gg
	V6	分类	c, d, cc, i, j, k, m, r, q, w, x, e, aa, ff
	V7	分类	v, h, bb, j, n, z, dd, ff, o
	V8	数值	0~28.5
	V9	分类	t, f
	V10	分类	t, f
	V11	数值	0~67
	V12	分类	t, f
	V13	分类	g, p, s
	V14	数值	0~2000
	V15	数值	0~100,000

Table 3. Distribution of credit card approval categories  
表 3. 信用卡审批类别分布

类别	样本量
+	307 (44.5%)
-	383 (55.5%)



**Table 4.** Credit card approval missing data statistics  
**表 4.** 信用卡审批缺失数据统计

变量	缺失个数
V1	12
V2	12
V4	6
V5	6
V6	9
V7	9
V14	13

**4.2. 定期存款预测模型的建立**

在本节中进行实证分析，将提出的 QCaenSVM 应用于信用卡审批的 UCI 数据集上进行预测模型，实验都在 R4.3.1 中进行，在基于系统配置 Intel(R) Core(TM) i7-10700 CPU(2.90GHz)和 16 GB RAM 的 PC 上运行，使用基于梯度下降的 DC 算法进行求解模型。并把所提模型同 SVM，PinSVM 和 aLSSVM 进行比较。

关于模型参数的设置，对于 SVM、PinSVM、aLSSVM 和 QCaenSVM，设置  $C = \{2^{-8}, 2^{-7}, \dots, 2^7, 2^8\}$ ；设置 PinSVM、aLSSVM 和 QCaenSVM 中  $\tau$  的取值范围为  $\tau = \{0.3, 0.5, 0.7, 0.9\}$ ，QCaenSVM 的  $s \in \{1, 1.5, 2\}$ 。非线性情况，考虑高斯核函数，评价指标为准确率(Acc)。对于所有模型参数的选择，本文使用五折交叉验证法来搜索模型的最优参数。为了消除实验随机性的影响，进行 10 次数值实验，计算评价指标的平均值(Acc)和标准差(sd)。考虑实际数据为非线性可分，所以建立高斯核函数下非线性可分模型，在基于梯度下降的 DC 中，我们把 eps 的值固定为  $10^{-3}$ ， $T_1 = 10$ ， $T_2 = 500$ 。线性情况下的模型结果如表 5 所示，高斯情况下的模型结果如表 6 所示。提出的模型提高了预测准确率，将其应用于银行是否同意给客人提高信用卡的审批有更好的表现效果。

**Table 5.** Prediction results of linear models  
**表 5.** 线性下模型的预测结果

	SVM Acc. $\pm$ sd	aLSSVM Acc. $\pm$ sd	PinSVM Acc. $\pm$ sd	QCaenSVM Acc. $\pm$ sd
数据集	$0.795 \pm 0.025$	$0.873 \pm 0.016$	$0.801 \pm 0.031$	$0.937 \pm 0.037$

**Table 6.** Prediction results of Gaussian models  
**表 6.** 高斯下模型的预测结果

	SVM Acc. $\pm$ sd	aLSSVM Acc. $\pm$ sd	PinSVM Acc. $\pm$ sd	QCaenSVM Acc. $\pm$ sd
数据集	$0.763 \pm 0.065$	$0.856 \pm 0.031$	$0.821 \pm 0.028$	$0.907 \pm 0.021$

**5. 结论**

在本论文中，通过对信用卡审批的深入比较与分析，我们聚焦于 QCaenSVM 模型在信用卡审批应用中的性能表现。实验结果显示，无论是在线性核函数还是高斯核函数场景下，QCaenSVM 模型均展现出相较于传统 SVM、PinSVM 及 aLSSVM 模型的显著优势，具体表现为更高的预测精度与更强的泛化能

力。表明了 QCaenSVM 模型在复杂信贷风险评估领域的独特优势。首先, QCaenSVM 通过其创新的算法设计与优化策略, 有效提升了模型在处理高维、非线性及非平衡数据集时的效率与准确性, 这对于信用卡审批中常见的大量且多变的客户数据尤为关键, 同时极大地提高了审批效率, 通过自动化的风险评估流程减少了人工审核的时间与成本, 加速了信贷服务的响应速度。其次, 该模型在保持高精度的同时, 也展现出了良好的鲁棒性, 能够抵御一定程度的噪声干扰和异常值影响, 确保审批决策的稳健性, 降低坏账风险, 同时也有助于保护消费者权益, 避免因不恰当的审批决策而导致的信用损失。

综上所述, QCaenSVM 模型在信用卡审批领域的应用不仅显著优于传统 SVM 模型, 更以其卓越的性能为信用卡审批提供了强有力的技术支持, 推动了信贷审批流程的优化与升级, 具有重要的实践意义与广泛的应用前景。

## 参考文献

- [1] 刘继海, 陈晓剑. SVM 模型在信用卡申请管理中的创新应用[J]. 哈尔滨工业大学学报(社会科学版), 2007, 9(4): 133-136.
- [2] Cortes, C. and Vapnik, V. (1995) Support-Vector Networks. *Machine Learning*, **20**, 273-297. <https://doi.org/10.1007/bf00994018>
- [3] Wu, W., Xu, Y. and Pang, X. (2021) A Hybrid Acceleration Strategy for Nonparallel Support Vector Machine. *Information Sciences*, **546**, 543-558. <https://doi.org/10.1016/j.ins.2020.08.067>
- [4] Pang, X., Zhang, Y. and Xu, Y. (2022) A Novel Multi-Task Twin-Hypersphere Support Vector Machine for Classification. *Information Sciences*, **598**, 37-56. <https://doi.org/10.1016/j.ins.2022.03.067>
- [5] Wang, Z., Shao, Y., Bai, L., Li, C., Liu, L. and Deng, N. (2018) Insensitive Stochastic Gradient Twin Support Vector Machines for Large Scale Problems. *Information Sciences*, **462**, 114-131. <https://doi.org/10.1016/j.ins.2018.06.007>
- [6] Huang, X.L., Shi, L. and Suykens, J.A.K. (2014) Support Vector Machine Classifier with Pinball Loss. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **36**, 984-997. <https://doi.org/10.1109/tpami.2013.178>
- [7] Huang, X., Shi, L. and Suykens, J.A. (2014) Asymmetric Least Squares Support Vector Machine Classifiers. *Computational Statistics & Data Analysis*, **70**, 395-405. <https://doi.org/10.1016/j.csda.2013.09.015>
- [8] Qi, K. and Yang, H. (2022) Joint Rescaled Asymmetric Least Squared Nonparallel Support Vector Machine with a Stochastic Quasi-Newton Based Algorithm. *Applied Intelligence*, **52**, 14387-14405. <https://doi.org/10.1007/s10489-022-03183-2>
- [9] Qi, K. and Yang, H. (2023) Capped Asymmetric Elastic Net Support Vector Machine for Robust Binary Classification. *International Journal of Intelligent Systems*, **2023**, Article ID: 2201330. <https://doi.org/10.1155/2023/2201330>
- [10] Efron, B. (1991) Regression Percentiles Using Asymmetric Squared Error Loss. *Statistica Sinica*, **1**, 93-125.