

基于改进主动生成式过采样的个人信用风险评估研究

顾哲涵, 黄宝凤

(南京邮电大学经济学院, 江苏南京 210023)

摘要: 针对个人信用风险评估中的样本不均衡和类别重叠问题, 提出一种改进的主动生成式过采样模型。首先, 在辅助分类器生成对抗网络(ACGAN)框架的基础上引入 Wasserstein 距离改善真假判别损失函数, 加入梯度惩罚以防止模式崩溃; 其次, 采用 Focal loss 代替传统交叉熵损失, 以增强对困难样本的识别能力; 最后, 利用所提模型对不平衡数据进行过采样, 以提升分类器性能。针对真实信贷数据的实验表明, 该模型将分类器的分类性能指标 F1、AUC 及 G-means 分别提升 11.2%、1.7%、12.8%, 在增强样本多样性、减少类别重叠及提升分类器针对非平衡数据集的分类效能方面取得了显著成效。

关键词: 深度学习; 不平衡数据; 类重叠; ACGAN; focal loss; Wasserstein 距离

DOI: 10.11907/rjdk.241335

开放科学(资源服务)标识码(OSID):



中图分类号: TP183

文献标识码: A

文章编号: 1672-7800(2024)009-0163-07

Research on Personal Credit Risk Assessment Based on Improved Active Generative Oversampling

GU Zhehan, HUANG Baofeng

(School of Economics, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

Abstract: Aiming at the problems of sample imbalance and category overlap in personal credit risk assessment, an improved active generative oversampling model is proposed. Firstly, based on the auxiliary classifier generative adversarial network (ACGAN) framework, Wasserstein distance is introduced to improve the true false discrimination loss function, and gradient penalty is added to prevent pattern collapse; Secondly, Focal loss is used instead of traditional cross entropy loss to enhance the ability to identify difficult samples; Finally, the proposed model is used to oversample imbalanced data to improve classifier performance. Experiments on real credit data show that the model improves the classifier's classification performance indicators F1, AUC, and G-means by 11.2%, 1.7%, and 12.8%, respectively. It achieves significant results in enhancing sample diversity, reducing class overlap, and improving the classifier's classification performance on imbalanced datasets.

Key Words: deep learning; unbalanced data; class overlap; ACGAN; focal loss; Wasserstein distance

0 引言

信贷评估领域研究主要集中在借款人的贷前信用评估, 通常基于借款人的多维信息构建二分类预测模型, 合理的预测模型为信贷机构的贷前评估提供了极大帮助。由于大多预测模型通常假设不同类别的样本数量相等, 因

此信贷数据特有的不平衡性及类重叠性往往会造成分类模型对违约样本识别能力不足^[1]。具体而言, 不平衡数据可能导致模型偏向于预测多数类别, 忽视少数类别样本, 使风险评估不够准确, 且类重叠现象可能导致模型难以区分不同类别样本, 增加了误判概率, 而合成少数类过采样(SMOTE)、边界合成少数过采样(Borderline-SMOTE)等传统线性插值方法在处理该问题时存在改变数据初始分布

收稿日期: 2024-03-04

扫描二维码阅读全文:



作者简介: 顾哲涵(2000-), 男, 南京邮电大学经济学院硕士研究生, 研究方向为社会经济统计、数据挖掘; 黄宝凤(1964-), 男, 南京邮电大学经济学院教授, 研究方向为社会经济统计、数据挖掘。

等问题^[2]。随着计算机技术快速发展,对抗神经网络(Generative Adversarial Networks, GANs)等深度学习技术被证明可学习到真实样本分布和生成新数据,但处理类重叠数据时仍存在局限性^[3]。因此如何兼顾样本生成和类重叠处理,是当前信用风险评估中亟需解决的问题。

1 相关研究

信用数据的不平衡性是指信用数据中履约客户数据量远多于违约客户数据量。杨莲等^[4]指出不平衡性会导致模型判别结果偏向于多数类,即对少数类样本的识别能力不足,而数据类重叠问题是指特征相似样本具有不同类别标签。Pattaramon等^[5]指出类间重叠区域中的多数类属于优势类,会造成分类模型决策边界倾向于多数类。Zhu等^[6]、Lee等^[7]指出类重叠的不平衡数据会进一步加大分类模型对少数类样本的学习难度。针对该问题,常见方案包括随机过采样、随机欠采样,虽然二者可快速构建平衡数据集,但容易导致过拟合或损失重要分类信息等问题^[8]。为此,Chawla等^[9]提出一种合成少数类的过采样方法(SMOTE),从少数类样本邻域出发,以线性插值法为基础生成样本,如今在多领域得到广泛应用,但仍存在产生噪声和过拟合等问题。Han等^[10]提出 Borderline-SMOTE 算法,基于邻域多数类样本数量减少对非边界样本的处理,不仅提升了算法运行效率还极大保留了数据特征。

针对数据类重叠问题,有学者基于传统 SMOTE 框架提出 Safe-leval-SMOTE 算法,通过设置安全区和噪声区有效降低了生成样本的样本重叠程度^[11]。陶叶辉等^[12]提出一种基于高斯混合模型聚类的 SMOTE 算法,通过预先进行 GMM 聚类并删除冗余样本,然后根据不同分类簇进行过采样,有效提升了分类模型分类不平衡数据的效果。田园等^[13]将改进的 SMOTEENN-XGBoost 信用卡风险预测模型应用于 Kaggle 公开信用数据集中,显著提升了违约预测准确性。

尽管上述方法在不平衡性和类重叠性的数据中均取得了不错效果,但 SMOTE 框架下样本生成方法基于线性插值法难以模拟真实数据的分布特征、易于产生噪声,以致于模型在不同数据集上的适应性较差^[14]。随着深度学习不断发展,Goodfellow 等^[15]提出的生成对抗网络在经济金融领域中得到了广泛运用,例如生成金融记录、市场风险模型优化及资产波动性预测等。该模型在不依赖任何先验假设的情况下,通过自主学习原始数据分布特征生成类同原始数据新样本。马甜等^[16]将该模型应用于中国股票市场因子投资预测,相较于传统线性模型显著提升了预测精度和投资绩效,弥补了以往研究在处理复杂市场数据模式和预测微观企业基本状况方面的不足。姚潇等^[17]利用生成对抗网络对债券违约样本数据进行过采样,显著提升了违约风险预测模型的 AUC、F1 得分,有效解决了非平

衡样本下的预测准确性问题。Oh 等^[18]、Dong^[19]等将 GAN 和条件生成对抗网络(Conditional Generative Adversarial Network, CGAN)应用于高维不平衡数据的过采样,结果表明该模型可更好模拟原始数据分布,生成较线性插值方法更真实的数据,但将生成对抗网络框架应用于非平衡数据的数据过采样时,原始 GAN 通常仅使用少数类数据进行对抗训练,将造成生成模型过分关注少数类别特征表达,忽视了类别间的整体分布关系^[20]。尽管能在一定程度上缓解少数类数据的不足,但模型对少数类样本的过度拟合往往会损害其泛化能力,难以保证生成样本的质量和多样性^[21]。

为此,Odena 等^[22]结合 CGAN 与 DCGAN 提出辅助分类器生成对抗网络(Auxiliary Classifier Generates Adversarial Network, ACGAN),只将类别约束信息加入生成器,并在判别器中额外添加辅助分类器以提升生成样本质量及多样性。张在美等^[23]将该模型应用于两个类重叠性严重的不平衡数据集中,发现模型生成数据与多数类数据存在较明确的分类边界,在平衡数据集的同时可有效缓解数据类重叠性。

综上所述,基于 GAN 的生成式过采样方法相较于 SMOTE 类方法可模拟更真实的数据,且带有辅助分类器的 ACGAN 能在平衡数据集的同时缓解类重叠问题。然而,现有生成类模型同样存在以下问题:①由于生成对抗网络固有的结构模式崩塌问题和模型训练不稳定性难以保证生成样本的质量,甚至会产生噪声污染的数据样本^[24];②面对极度不平衡的数据时,判别器类重叠样本识别不足时将直接影响生成样本质量,而传统交叉熵损失给予两类样本相同权重,使模型难以学习困难样本的分布特征^[25]。为此,本文综合考虑信贷数据集不平衡性及类重叠性特征,分别优化现有模型判别损失函数、分类损失函数,提出改进的主动生成式过采样模型(WACGAN_FL),以生成更高质量的数据样本,提升模型对不平衡数据识别的准确性。

2 主动生成式过采样模型

2.1 辅助分类器生成对抗网络

辅助分类器生成对抗网络是一种生成对抗网络的扩展,在生成模型架构中引入辅助分类器能使网络生成更逼真的数据,还能控制生成数据类别,具体框架如图 1 所示。首先,为生成器 G 输入类别标签 c 来引导 GAN 生成指定类别的样本 $X_{fake} = G(z, c)$;其次,将类别信息与噪声相结合使生成器产生特定类别的合成数据;最后,判别器在判别样本真假的同时预测样本类别。该设置为生成器提供了额外的梯度信息,能帮助其更有效地学习不平衡数据的边界信息。

ACGAN 损失函数可分为传统判别损失和辅助分类损

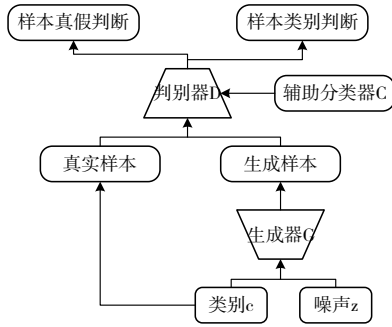


Fig. 1 ACGAN modeling framework
图1 ACGAN模型框架

失, 以反映真假判别和分类判别两方面, 具体损失函数如下:

$$L_S = E_{x \sim P_{data}} [\log p(s = real | x_{real})] + E_{z \sim P_z} [\log p(s = fake | x_{fake})] \quad (1)$$

$$L_C = E_{x \sim P_{data}} [\log p(C = c | x_{real})] + E_{z \sim P_z} [\log p(C = c | x_{fake})] \quad (2)$$

式中: L_S 为真假损失函数, 用于判别样本真、假; L_C 为辅助分类损失, 用于判断生成样本与真实样本标签相似程度, 以优化生成数据类别的准确性。

在训练过程中, 判别器 D 的训练目标为尽可能分清数据来源, 并识别数据的正确类别, 即最大化 $L_S + L_C$; 而生成器 G 的训练目标为尽可能生成被判别器识别为真且被分类器正确分类的样本, 即最大化 $L_C - L_S$ 。由此, 判别器和生成器持续对抗训练直至达到纳什均衡, 使生成器准确学习少数类标签相应的数据分布。针对类重叠性和不平衡性较强的信贷数据, 加入辅助分类器能更好帮助模型理解数据特征与类别间的交互关系, 有效降低生成数据落在类重叠区间的概率, 从而缓解合成样本的类重叠程度, 进而提升后续分类模型的泛化能力。

2.2 改进的主动生成式过采样模型

针对 ACGAN 训练过程不稳定、对不平衡数据识别能力不足的问题, 本文对该模型进行两处优化, 提出 WACGAN_FL (Wasserstein Auxiliary Classifier Generative Adversarial Network with Focal Loss)。具体优化方式如下:

2.2.1 判别函数优化

ACGAN 真假判别函数采用 JS 散度衡量两种不同分布间的差异程度, 然而该方法存在一定缺陷: 当判别器处于最佳状态时, JS 散度难以度量两个没有交集的分布, 从而出现梯度消失问题。Wasserstein 距离可克服这一缺陷, 无论两个分布是否存在交集, 均可通过判别器输出差值衡量分布相似度, 从而使模型训练过程更稳定^[26]。Wasserstein 距离的前提要求判别函数服从 Lipschitz 连续条件, 即导函数不可超过 Lipschitz 常数 K, 常见的做法是通过权重裁剪将判别器权重限定在某个特定范围使其满足 Lipschitz 条件, 然而该方式容易引起梯度消失或梯度爆炸, 从而影响训练过程的稳定性^[27]。因此, 梯度惩罚在嵌入空间中引入一个额外的正则化项, 预先设定梯度上界与 Lipschitz 常数 K 之间的关系, 具体数学表述式如式(3)所示。

$$Relu(\|\nabla_x D(x)\|_2 - K)^2 \quad (3)$$

由于常数 K 的取值并不影响梯度下降方向, 因此本文设 $K=1$ 得到式(4)。

$$L_{CP} = \lambda E_{\hat{x} \sim P_{\hat{x}}} [(\|\nabla_x D(x)\|_2 - 1)^2] \quad (4)$$

式中: λ 为梯度惩罚的权重超参数; ∇ 为梯度; $\|\cdot\|_2$ 为 2 的范数; D 的梯度基于真实样本和生成样本之间的线性插值 $\hat{x}_{P_{\hat{x}}}$ 进行求解; $P_{\hat{x}}$ 为线性插值的采样分布。

由于梯度惩罚方式相较于权重裁剪能直接控制梯度范围, 有效提高训练过程的稳定性。为此, 将式(4)并入初始的 ACGAN 的损失函数中, 即将 L_S 修改至式(5), L_C 保持不变, 进而作为判别器的训练函数。

$$L_S = E_{x \sim P_{data}} [\log p(s = real | x_{real})] + E_{z \sim P_z} [\log p(s = fake | x_{fake})] + L_{CP} \quad (5)$$

2.2.2 分类损失优化

针对类别重叠且类别分布极度不平衡的信贷数据集, 常规的交叉熵损失函数难以驱动模型高效学习到少数类别的特征表达, 这是由于交叉熵损失通常会被大量占优势类别的样本所主导, 进而导致模型在训练过程中过分关注这些类别, 忽视了对少数类别的判别^[28]。为此, 本文采用 Focal Loss 作为 ACGAN 的分类损失优化策略来提高分类性能, 并在训练过程中实现类别平衡的有效学习。

Focal Loss 是一种改进的交叉熵损失函数, 最初提出以解决目标检测领域中的类别不平衡问题, 其核心思想是通过增加一个调节因子减少分类较容易样本(易分类样本)的损失贡献, 而对于难分类样本则增加损失权重^[29]。具体数学表达如式(6)所示。

$$L_{FL}(p_i) = -\alpha_i (1 - p_i)^\gamma \log(p_i) \quad (6)$$

式中: p_i 表示模型对于某一类别的预测概率; α_i 用于调节类别不平衡的权重系数; γ 为聚焦参数, 决定易分类样本的权重下降速率。随着 p_i 增大调制因子 $\alpha_i (1 - p_i)^\gamma$ 减小, 从而降低易分类样本的损失贡献; 而当 p_i 很小即模型对难分类样本的预测概率很低时损失贡献变大, 从而使模型更关注这些样本。

本文将 Focal Loss 引入 ACGAN 架构, 可充分考虑信贷数据集中存在的类别重叠和极度类别不平衡状况, 促使模型捕捉易忽视的少数类别信号。具体而言, 在 ACGAN 原有的损失函数中替换掉分类损失项。通过该方式, WACGAN_FL 不仅能生成高质量样本, 还能在类别不平衡的条件下增强模型对于信贷风险评估中高风险样本的识别能力。具体数学表达式如式(7)所示。

$$L'_{WACGAN} = -L_W + L_{CP} + L_{FL} \quad (7)$$

综上所述, 本文模型架构如图 2 所示。通过对判别器的分类损失及判别损失进行针对优化, 以期稳定模型训练过程, 聚焦生成少数类样本。

2.3 WACGAN_FL 算法流程

基于现有真实数据训练 WACGAN_FL 模型并保存权

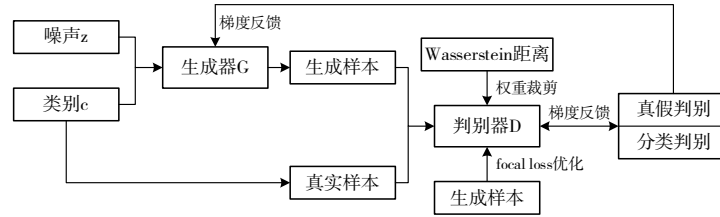


Fig. 2 WACGAN_FL model process

图2 WACGAN_FL模型流程

重,生成信用风险样本,从而获取平衡数据集。首先将原始数据集分为训练集及测试集,统计训练集中多数类与少数类的数量差值 n ;其次将训练集作为WACGAN_FL输入数据,聚焦学习多数类样本与少数类样本的共同特征以及各自分布,进而使用训练好的生成器生成 n 个少数类样本获取平衡数据集 D'_{train} ;最后统一使用 D'_{train} 对GBDT模型进行分类训练,基于分类器在测试集 D_{test} 的表现评估数据质量。具体算法流程如下:

算法1 WACGAN_FL过采样算法

输出:扩充后的训练集 D'_{train}

- 1: 初始化:判别器D的参数 θ_d 和生成器G的参数 θ_g 。
- 2: for 迭代次数 epoch = 1 to epochs do。
- 3: 从 D_{train} 中随机抽取 m 个真实样本 $\{x_1^{real}, \dots, x_m^{real}\}$ 及其标签 $\{y_1, \dots, y_m\}$ 。
- 4: 随机采样 m 个隐向量 $\{z_1, \dots, z_m\}$ 。
- 5: 利用生成器G产生假样本 $\{x_1^{fake}, \dots, x_m^{fake}\}$ 。
- 6: 通过判别器D对 $\{x_1^{real}, \dots, x_m^{real}\}$ 和 $\{x_1^{fake}, \dots, x_m^{fake}\}$ 进行真假以及标签判定得到 $\{s, \dots, s_m\}$ 和 $\{y'_1, \dots, y'_m\}$ 。
- 7: 根据判别结果和真实标签计算Focal Loss作为类条件损失 L_{FL} 。
- 8: 生成混合样本 $x = \epsilon \cdot x^{real} + (1 - \epsilon) \cdot x^{fake}$ 并计算梯度惩罚损失 L_{CPO} 。
- 9: 使用Adam优化器更新判别器D的参数 θ_d 。
- 10: 使用Adam优化器更新生成器G的参数 θ_g 。
- 11: end for。
- 12: 使用训练好的生成器G生成 n 个少数类样本并加入到 D_{train} 中,得到扩充的训练集 D'_{train} 。
- 13: return D'_{train} 。

3 实验结果与分析

3.1 实验设计与数据来源

本文为验证生成式过采样模型的有效性,将实验设计分为两个部分:①评估WACGAN_FL生成样本质量;②将WACGAN_FL生成式过采样与GBDT相结合,以分类器分类效果为基准与其他模型进行比较,从而验证模型的有效性。受限于数据的可得性,个人信用风险评估数据来源于国际Kaggle大数据竞赛平台提供的Lending Club个人信贷数据集,实验数据共有初始特征145项。本文对数据进行

特征编码、特征衍生后,使用组合过滤法筛选特征以选取有效分类特征32项;其次标准化数据,以消除特征间量纲差异。综上,通过预处理后的数据集共包含样本187 339条,其中履约样本171 340条,违约样本15 999条,不平衡率为10.7%。

3.2 实验环境与超参数设置

WACGAN_FL判别器D与生成器G均采用多层感知机结构,隐藏层层数设置为3,交替训练次数设置为5,即训练5次判别器后训练一次生成器。隐向量 z 的维度设置为数据集特征数,D、G隐藏层均采用LeakyReLu激活函数,去除D的真假判断激活函数,类别判断采用Softmax函数,G的输出层使用Tanh激活函数。依据数据复杂程度,将实验数据集所使用的隐藏层节点数设定为256,梯度惩罚系数预设为10,使用Adam优化算法,学习率设为0.000 05。模型采用Python作为编程语言,IDE为Pycharm 2020,代码基于tensorflow框架实现,硬件配置为8 GB内存,Intel(R) Core(TM) i5-7300HQ CPU@2.50 GHz处理器。

3.3 生成样本质量评估

3.3.1 分布相似度评估

本文为衡量不同样本生成方法生成数据的质量差异,基于生成数据与原始数据集样本分布的相似度视角进行评估。其中,以JS散度为度量两个概率分布之间相似性的统计指标,该指标相较于其他度量手段,凭借其对称性和稳定性被公认为是一种用于评价概率分布差异的可信方法。JS散度基于两个概率分布的KL散度,通过对称化并平均两个概率分布来减少KL散度可能产生的偏差。设两个离散概率分布分别为P、Q,则JS散度可表达为:

$$D_{JS}(P // Q) = \frac{1}{2} D_{KL}(P // \frac{P+Q}{2}) + D_{KL}(Q // \frac{P+Q}{2}) \quad (8)$$

由于JS散度反映了两个概率分布平均值情况下的相对熵,值域在0~1之间,当JS散度值趋近于0时两个分布更接近,即生成样本与原始数据集在统计分布上更相似。本文首先对原始数据集进行概率分布估计;然后估计每种样本的生成方法生成样本的概率分布;最后计算每种生成样本概率分布与原始数据集概率分布的JS散度,并以此作为评价标准来衡量样本生成方法的有效性。

本文实验共选用3种SMOTE类过采样方法和4种生成类过采样方法来评估分布相似度,并对生成数据和原始数据的各个特征进行JS散度计算,以均值为评估依据,具体结果如表1所示。由此可知,除了未经优化的AC-

GAN 模型外, GAN 类方法能生成与原始数据集更相似的样本, 表现优于 SMOTE 类过采样方法, 证明了 GAN 类数

据生成方法可更好地模拟原始数据分布, 能生成更真实的样本。

Table 1 Distributional similarity assessment

表 1 分布相似度评估

| 模型 | SMOTE | B_SMOTE | SMOTEENN | GAN | WGAN | ACGAN | WACGAN-FL |
|-------|---------|---------|----------|---------|---------|---------|-----------|
| JS 散度 | 0.087 7 | 0.173 4 | 0.084 2 | 0.072 5 | 0.068 8 | 0.117 6 | 0.082 8 |

3.3.2 类重叠性评估

本文采用 UMAP 算法将多维数据映射至二维空间, 以比较原始数据集与经过不同过采样技术处理后的数据分布, 实验结果图 3 所示。SMOTE 类传统过采样技术生成的样本缺乏多样性且并未解决类别重叠问题, 而未结合辅助分类器结构的 GAN、WGAN 尽管在一定程度上识别了少数类分布, 但未明确区分不同类别, 生成的样本仍聚集于类

别重叠区。

此外, 尽管 ACGAN 能识别类别间差异, 但原始数据集的不均衡性使生成的少数类样本受到多数类影响, 使模型的特征表现不足。相比之下, 采用改进的 WACGAN_FL 算法生成的样本不仅呈现出更高的多样性, 且基于 Focal loss 的损失函数优化有效避免了生成样本落入类重叠区的问题。

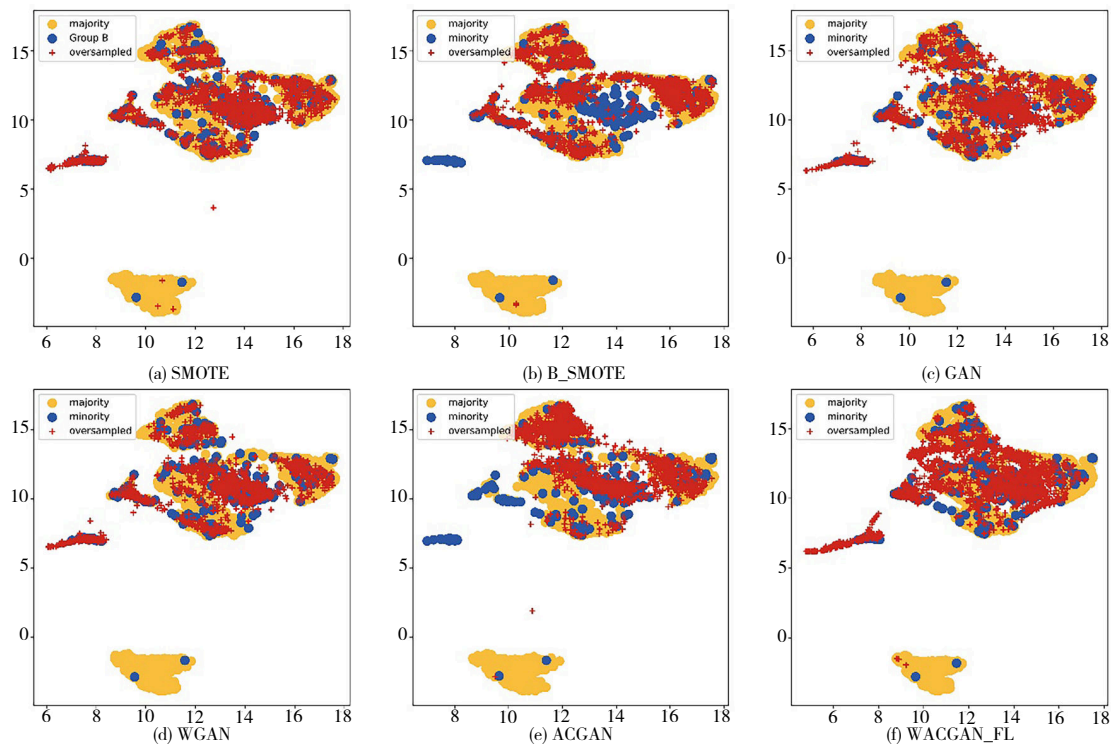


Fig. 3 Visualization results of data distribution

图 3 数据分布可视化结果

3.4 实证结果分析

个人信贷领域中风险评估模型的关键目标是正确识别潜在违约用户, 降低信用风险, 这一任务相较于正确识别履约用户更有意义, 因此本文模型在追求兼顾多数类识别准确率的同时, 会尽可能提高少数类识别准确率。

为检验模型有效性, 本文选取了以下 3 类对照模型用于性能比较: ①未进行数据质量提升处理的 GBDT 分类模型; ②应用广泛的样本不平衡处理方法与 GBDT 的组合模型, 例如插值过采样 SMOTE、边界过采样 B_SMOTE 和联合过采样 SMOTEENN 等; ③GAN 系列样本生成方法与 GBDT 的组合模型, 包括 GAN、WGAN、ACGAN 等。表 2 为所有模型在 4 个评价指标上的实证结果。其中, “G”表示“GBDT”, 本文模型以“WACGAN-FLG”表示; 针对本文不平衡

数据分类的问题背景, 选用 F1、AUC 及 G-mean 作为分类器评估指标。

Table 2 Evaluation metrics for classifiers based on different oversampling methods

表 2 基于不同过采样方法的分类器评价指标

| Model | F1 | AUC | G-mean |
|------------|-------|--------|--------|
| GBDT | 0.746 | 0.937 | 0.775 |
| SMOTE-G | 0.741 | 0.949 | 0.871 |
| BSMOTE-G | 0.602 | 0.946 | 0.865 |
| SMOTEENN-G | 0.781 | 0.927 | 0.801 |
| GAN-G | 0.822 | 0.9515 | 0.844 |
| WGAN-G | 0.834 | 0.952 | 0.856 |
| ACGAN-G | 0.825 | 0.952 | 0.847 |
| WACGAN-FLG | 0.829 | 0.953 | 0.874 |

由表2可知,从F1分数指标而言,WGAN-G模型表现最佳,WACGAN-FLG模型仅与其存在不足0.6%的差距,相较于基线GBDT模型提升11.1%,证明了所提模型在维持对多数类识别准确性的同时,强化了对少数类的检测效果,且在精确率与召回率之间取得了更好的平衡。从AUC指标而言,所有模型的曲线下面积值均较高,总体上具有不错的分类性能,本文模型以0.953的表现领先于其他模型,相较于基线GBDT模型提升1.7%,证明了本文模型对履约和违约两类样本具有良好的分类能力。从G-mean指标而言,WACGAN-FLG模型在所有模型中表现最佳,G-mean为0.874,相较于对照模型提升0.3%~12.7%,并且G-mean指标还考量了类别不平衡情况下的准确性,进一步验证了所提模型在识别少数类即违约用户方面优越的综合能力。

综上,本文模型充分考虑了信用样本不平衡、类间重叠等问题,通过改进原始ACGAN的判别损失和分类损失,并结合GBDT分类模型进行有效分类,达到了改善信用风险评估性能的目的。

3.5 模型稳健性检验

为验证本文模型的有效性及其稳健性,从KEEL数据库中选取6组不平衡数据集进行验证,数据集不平衡率皆在1~10之间,具体信息如表3所示。

Table 3 Basic information of the dataset
表3 数据集基本信息

| 数据集 | 少数类 | 多数类 | 特征数 | 不平衡率 |
|-------------|-----|-------|-----|------|
| page-blocks | 559 | 4 913 | 10 | 8.79 |
| ecoli | 77 | 259 | 7 | 3.36 |
| yeast | 429 | 1 055 | 8 | 2.46 |
| pima | 268 | 500 | 8 | 1.87 |
| haberman | 81 | 225 | 3 | 2.78 |
| wine | 199 | 1 400 | 11 | 7.04 |

接下来,使用SMOTE类与GAN类模型作为比较模型,评估改进的WACGAN_FL生成的样本质量,结果如表4所示。由此可知,本文模型在6个数据集的评估指标均较优,F1值平均排名为1.34,AUC值平均排名为1.67,G-mean平均排名为1.5,在所有模型中排名最高,证明了该模型具有较强的泛化能力,改进的主动生成式过采样具有较强的稳健性。

4 结语

本文为应对个人信用风险评估中普遍存在的样本不均衡和高度类别重叠问题,提出一种改进型WACGAN_FL信贷风险评估模型。该模型借鉴图像生成领域广泛使用的生成对抗网络(GANs)架构,将其优化于生成结构化数据,并针对ACGAN框架进行关键创新。

首先,借鉴WGAN理念使用Wasserstein距离改进原始ACGAN模型的真假判别损失函数,引入梯度惩罚策略稳定模型训练过程以避免模型发生崩溃现象。其次,采用

Table 4 Comparison of GBDT classification results

表4 GBDT分类结果比较

| 指标 | 模型 | 数据集 | | | | | |
|--------|------------|----------------|----------------|----------------|----------------|----------------|----------------|
| | | page-blocks | ecoli | yeast | pima | haberman | wine |
| F1 | GBDT | 0.986 7 | 0.913 6 | 0.840 7 | 0.805 4 | 0.802 8 | 0.974 3 |
| | SMOTE-G | 0.982 0 | 0.930 8 | 0.801 3 | 0.762 2 | 0.769 2 | 0.942 4 |
| | BSMOTE-G | 0.974 0 | 0.904 5 | 0.802 0 | 0.758 0 | 0.763 4 | 0.968 6 |
| | SMOTEENN-G | 0.973 9 | 0.910 3 | 0.768 4 | 0.717 0 | 0.694 9 | 0.905 9 |
| | GAN-G | 0.988 0 | 0.936 7 | 0.827 5 | 0.800 0 | 0.829 2 | 0.956 9 |
| | WGAN-G | 0.987 7 | 0.937 5 | 0.839 4 | 0.797 3 | 0.829 6 | 0.962 9 |
| | ACGAN-G | 0.985 6 | 0.936 7 | 0.835 9 | 0.805 6 | 0.814 8 | 0.962 9 |
| | WACGAN-FLG | 0.988 7 | 0.950 0 | 0.841 1 | 0.824 3 | 0.818 2 | 0.981 9 |
| | GBDT | 0.990 8 | 0.902 8 | 0.762 6 | 0.795 8 | 0.604 9 | 0.716 2 |
| AUC | SMOTE-G | 0.990 8 | 0.902 5 | 0.758 1 | 0.793 5 | 0.677 8 | 0.695 6 |
| | BSMOTE-G | 0.990 2 | 0.899 7 | 0.761 1 | 0.786 1 | 0.665 5 | 0.760 4 |
| | SMOTEENN-G | 0.985 6 | 0.911 7 | 0.756 2 | 0.784 5 | 0.696 1 | 0.789 2 |
| | GAN-G | 0.990 1 | 0.931 2 | 0.761 1 | 0.810 9 | 0.670 7 | 0.797 4 |
| | WGAN-G | 0.990 3 | 0.900 0 | 0.770 0 | 0.803 6 | 0.683 6 | 0.724 6 |
| | ACGAN-G | 0.990 8 | 0.932 7 | 0.767 2 | 0.804 4 | 0.677 7 | 0.804 2 |
| | WACGAN-FLG | 0.991 3 | 0.946 9 | 0.770 0 | 0.828 2 | 0.692 6 | 0.774 0 |
| | GBDT | 0.935 3 | 0.770 6 | 0.612 1 | 0.725 6 | 0.482 2 | 0.242 3 |
| | SMOTE-G | 0.962 2 | 0.854 9 | 0.679 2 | 0.698 0 | 0.591 3 | 0.402 5 |
| G-mean | BSMOTE-G | 0.944 3 | 0.810 8 | 0.707 3 | 0.708 3 | 0.566 1 | 0.413 2 |
| | SMOTEENN-G | 0.954 4 | 0.837 4 | 0.695 2 | 0.692 6 | 0.598 7 | 0.663 4 |
| | GAN-G | 0.949 9 | 0.881 2 | 0.645 5 | 0.758 7 | 0.651 3 | 0.532 4 |
| | WGAN-G | 0.944 3 | 0.860 7 | 0.657 9 | 0.733 3 | 0.651 3 | 0.529 3 |
| | ACGAN-G | 0.950 5 | 0.881 2 | 0.655 5 | 0.752 7 | 0.620 2 | 0.475 0 |
| | WACGAN-FLG | 0.958 8 | 0.893 0 | 0.659 1 | 0.764 3 | 0.663 7 | 0.667 8 |

Focal loss代替传统交叉熵损失函数来强化模型对难分样本的关注能力。实验表明,WACGAN_FL模型不仅在生成多样化和减少类间重叠样本方面表现出色,在分类器评估方面的AUC、G-mean皆优于对照组模型,证明了该模型能在兼顾履约样本识别率的同时,显著提升违约样本识别率。此外,通过有效区分违约与非违约客户,有助于显著降低金融机构的信用风险、强化信贷管理,确保信用体系健康稳定。因此,WACGAN_FL模型的推广及应用对优化、提升金融部门的信贷评估流程具有深远意义。

后续,将进一步优化模型各个部分,以提升模型在各种数据集上的性能。

参考文献:

[1] FAN D X, YE C M. Fusion clustering oversampling algorithm for credit imbalance data classification[J]. Software Guide, 2021, 20(11):70-74. 樊东醒,叶春明.融合聚类过采样算法的信贷不平衡数据分类[J].软件导刊,2021,20(11):70-74.

[2] LI Y X, CHAI Y, HU Y Q, et al. A review of unbalanced data classification methods[J]. Control and Decision Making, 2019, 34(4):673-688. 李艳霞,柴毅,胡友强,等.不平衡数据分类方法综述[J].控制与决策,2019,34(4):673-688.

[3] LIU H L, CAO S J, XU J Y, et al. Progress of anti-fraud research on digital credit transactions[J]. Computer Science and Exploration, 2023, 17(10):2300-2324.

- 刘华玲,曹世杰,许珺怡,等. 数字信用交易反欺诈研究进展[J]. 计算机科学与探索,2023,17(10):2300-2324.
- [4] YANG L, SHI B F, DONG Y Z. Modified cross-entropy unbalanced sample credit risk evaluation model based on class balanced loss [J]. Journal of Systems Management, 2022, 31(2):255-269.
杨莲,石宝峰,董轶哲. 基于Class Balanced Loss修正交叉熵非均衡样本信用风险评估模型[J]. 系统管理学报,2022,31(2):255-269.
- [5] PATTARAMON V, EYAD E. Improved overlap-based undersampling for imbalanced dataset classification with application to epilepsy and parkinson's disease[J]. International Journal of Neural Systems, 2020, 30(8): 2050043.
- [6] ZHU Y, YAN Y, ZHANG Y, et al. EHSO: evolutionary hybrid sampling in overlapping scenarios for imbalanced learning [J]. Neurocomputing, 2020, 417:333-346.
- [7] LEE D, KIM K. An efficient method to determine sample size in oversampling based on classification complexity for imbalanced data [J]. Expert Systems with Applications, 2021, 184:115442.
- [8] LI Y, LIU Z D, ZHANG H J. A review of integrated classification algorithms for unbalanced data[J]. Computer Application Research, 2014, 31(5):1287-1291.
李勇,刘战东,张海军. 不平衡数据的集成分类算法综述[J]. 计算机应用研究,2014,31(5):1287-1291.
- [9] CHAWLA V N, BOWYER W K, HALL O L, et al. SMOTE: synthetic minority over-sampling technique [DB/OL]. <https://arxiv.org/abs/1106.1813>.
- [10] HAN H, WANG W Y, MAO B H. Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning [C]// International Conference on Intelligent Computing, 2005:878-887.
- [11] BUNKHUMPORNPAT C, SINAPIROMSARAN K, LURSINSAP C. Safe-Level-SMOTE: safe-level-synthetic minority over-sampling technique for handling the class imbalanced problem [C]// Pacific-Asia Conference on Advances in Knowledge Discovery & Data Mining, 2009: 475-482.
- [12] TAO Y H, ZHAO S W. An improved SMOTE algorithm based on gaussian hybrid clustering for unbalanced data [J]. Software Guide, 2022, 21(5):110-114.
陶叶辉,赵寿为. 面向不平衡数据基于高斯混合聚类的SMOTE改进算法[J]. 软件导刊,2022,21(5):110-114.
- [13] TIAN Y, GUO H L, JI Q. Credit card risky customer prediction based on SMOTEENN-XGBoost [J]. Software Guide, 2024, 23(7):138-143.
田园,郭红烈,吉倩. 基于SMOTEENN-XGBoost的信用卡风险客户预测[J]. 软件导刊,2024,23(7):138-143.
- [14] FAN D X, YE C M. A credit assessment method for high-dimensional missing unbalanced data [J]. Computer Application Research, 2021, 38(9):2667-2672.
樊东醒,叶春明. 一种面向高维缺失不平衡数据的信用评估方法[J]. 计算机应用研究,2021,38(9):2667-2672.
- [15] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets [DB/OL]. <https://arxiv.org/abs/1406.2661>.
- [16] MA T, JIANG F W, TANG G H. Deep learning and factor investment in Chinese stock market based on generative adversarial network approach [J]. China Economic Quarterly, 2022, 22(3):819-842.
马甜,姜富伟,唐国豪. 深度学习与中国股票市场因子投资——基于生成式对抗网络方法[J]. 经济学(季刊),2022,22(3):819-842.
- [17] YAO X, LI K, YU L A. Research on corporate bond default risk prediction based on generative adversarial network oversampling technique under unbalanced samples [J]. Systems Engineering Theory and Practice, 2022, 42(10):2617-2634.
姚潇,李可,余乐安. 非平衡样本下基于生成对抗网络过采样技术的公司债券违约风险预测研究[J]. 系统工程理论与实践,2022,42(10):2617-2634.
- [18] OH J, HONG Y J, BAEK J. Oversampling method using outlier detectable generative adversarial network [J]. Expert Systems with Applications, 2019, 133:1-8.
- [19] DONG Y F, XIAO H X, DONG Y. SA-CGAN: an oversampling method based on single attribute guided conditional GAN for multi-class imbalanced learning [J]. Neurocomputing, 2021, 472: 326-337.
- [20] WANG X Y, WU X L, FENG Y J. Fusion of over-undersampling and GAN for network intrusion detection [J/OL]. Small Microcomputer Systems, 2024-04-13. <http://kns.cnki.net/kcms/detail/21.1106.TP.20240130.1526.018.html>.
王秀玉,吴晓鸽,冯永晋. 融合过-欠采样与GAN的网络入侵检测方法[J/OL]. 小型微型计算机系统,2024-04-13. <http://kns.cnki.net/kcms/detail/21.1106.TP.20240130.1526.018.html>.
- [21] CHAI M T, ZHU Y P. Advances in generative adversarial network research and application [J]. Computer Engineering, 2019, 45(9): 222-234.
柴梦婷,朱远平. 生成式对抗网络研究与应用进展[J]. 计算机工程, 2019,45(9):222-234.
- [22] ODENA A, OLAH C, SHLENS J. Conditional image synthesis with auxiliary classifier GANs [C]// Proceedings of the 34th International Conference on Machine Learning, 2017:2642-2651.
- [23] ZHANG Z M, LYU J, LIU Y. Research on personal credit risk assessment model based on optimized ACGAN-GBDT [J]. Financial Theory and Practice, 2022, 43(5):84-89.
张在美,吕娟,刘彦. 基于优化ACGAN-GBDT的个人信用风险评估模型研究[J]. 财经理论与实践,2022,43(5):84-89.
- [24] JIANG H X, JIANG J Y, LIANG X. A review of machine learning-based credit card transaction fraud detection research [J]. Computer Engineering and Applications, 2023, 59(21):1-25.
蒋洪迅,江俊毅,梁循. 基于机器学习的信用卡交易欺诈检测研究综述[J]. 计算机工程与应用,2023,59(21):1-25.
- [25] LI H F, XU G H, HUANG S X. Bearing fault diagnosis based on active generative oversampling and deep stacked network [J]. Computer Integrated Manufacturing Systems, 2023, 29(1):146-159.
李慧芳,徐光浩,黄双喜. 基于主动生成式过采样和深度堆叠网络的轴承故障诊断[J]. 计算机集成制造系统,2023,29(1):146-159.
- [26] CHEN F J, ZHU F, WU Q X, et al. A review of research on generative adversarial networks and their applications in image generation [J]. Journal of Computer Science, 2021, 44(2):347-369.
陈佛计,朱枫,吴清潇,等. 生成对抗网络及其在图像生成中的应用研究综述[J]. 计算机学报, 2021, 44(2):347-369.
- [27] GULRAJANI I, AHMED F, ARJOVSKY M, et al. Improved training of Wasserstein GANs [DB/OL]. <https://arxiv.org/abs/1704.00028>.
- [28] ZHU Y M, GUO R Y, JU J J, et al. A boosted tree classification algorithm for unbalanced datasets incorporating focal loss [J]. Software Guide, 2021, 20(11):65-69.
朱翌民,郭茹燕,巨家骥,等. 一种结合Focal Loss的不平衡数据集提升树分类算法[J]. 软件导刊,2021,20(11):65-69.
- [29] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [DB/OL]. <https://arxiv.org/abs/1708.02002>.