

融合时间特征增强与定向GAN过采样的XGBoost信用卡欺诈检测研究

宋蔓蔓¹, 郭禹辰¹, 刘梦², 佟磊³

¹河北金融学院河北省科技金融重点实验室, 河北 保定

²河北金融学院金融博物馆, 河北 保定

³河北软件职业技术学院, 河北 保定

收稿日期: 2025年10月29日; 录用日期: 2025年11月27日; 发布日期: 2025年12月4日

摘要

针对信用卡欺诈检测中样本极度不平衡导致模型偏向正常样本、传统特征缺乏时间关联性导致区分度不足、过采样生成样本无针对性导致误判率高等核心问题, 本文提出一套融合时间特征增强与定向GAN过采样的XGBoost检测方案。首先, 基于公开信用卡欺诈数据集构建“交易小时 - 临时用户标识 - 交易间隔 - 间隔统计”四层时间特征体系, 补充欺诈行为的时序关联信息; 其次, 针对小额漏检与大额欺诈两类核心漏检模式, 设计定向GAN过采样模型, 通过特征约束生成目标样本, 并以余弦相似度筛选高质量样本; 最后, 优化XGBoost关键参数构建高精度分类模型。实验结果表明, 该方案在测试集上欺诈类精确率达98.81%、召回率达84.69%、F1值达91.21%; 消融实验验证, 时间特征增强可使F1值提升4.25个百分点, 定向GAN过采样可提升0.99个百分点。

关键词

信用卡欺诈检测, 时间特征增强, 定向GAN过采样, XGBoost

Study on XGBoost-Based Credit Card Fraud Detection Integrating Temporal Feature Enhancement and Targeted GAN Oversampling

Manman Song¹, Yuchen Guo¹, Meng Liu², Lei Tong³

¹Hebei Key Laboratory of Science and Technology Finance, Hebei Finance University, Baoding Hebei

²Financial Museum, Hebei Finance University, Baoding Hebei

³Hebei Software Institute, Baoding Hebei

Received: October 29, 2025; accepted: November 27, 2025; published: December 4, 2025

文章引用: 宋蔓蔓, 郭禹辰, 刘梦, 佟磊. 融合时间特征增强与定向GAN过采样的XGBoost信用卡欺诈检测研究[J].
计算机科学与应用, 2025, 15(12): 58-65. DOI: [10.12677/csa.2025.1512321](https://doi.org/10.12677/csa.2025.1512321)

Abstract

Aiming at the core issues in credit card fraud detection, such as extreme class imbalance leading models to favor normal samples, insufficient discriminative power of traditional features due to the lack of temporal correlation, and high misjudgment rate caused by the lack of targeting in over-sampled generated samples, this paper proposes a set of XGBoost detection schemes integrating temporal feature enhancement and targeted GAN oversampling. Firstly, based on the public credit card fraud dataset, a four-layer temporal feature system of “transaction hour—temporary user ID—transaction interval—interval statistics” is constructed to supplement the temporal correlation information of fraudulent behaviors. Secondly, targeting the two core missed detection patterns (small-amount missed detection and large-amount fraud), a targeted GAN oversampling model is designed. Target samples are generated through feature constraints, and high-quality samples are screened using cosine similarity. Finally, the key parameters of XGBoost are optimized to construct a high-precision classification model. Experimental results show that the scheme achieves a fraud class precision of 98.81%, recall of 84.69%, and F1-score of 91.21% on the test set. Ablation experiments verify that temporal feature enhancement can increase the F1-score by 4.25 percentage points, and targeted GAN oversampling can increase it by 0.99 percentage points.

Keywords

Credit Card Fraud Detection, Temporal Feature Enhancement, Targeted GAN Oversampling, XGBoost

Copyright © 2025 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

随着数字支付的普及，全球信用卡交易规模与欺诈损失同步增长。2023 年全球支付卡欺诈损失已达 338.3 亿美元，较 2022 年上升 1.1%；预计到 2033 年，全球支付卡交易规模将增至 81.8 万亿美元，欺诈损失将突破 485.1 亿美元，每 100 美元交易额中就有 5.93 美分因欺诈损失[1]。此外，2024 年《AFP 支付欺诈与控制调查报告》显示，80% 的支付机构在 2023 年遭遇过欺诈攻击，其中 30% 的机构无法追回损失资金[1]，这凸显了信用卡欺诈检测对金融安全的重要性。传统欺诈检测方法多依赖专家规则引擎[2]，但此类方法对新型欺诈模式适应性差，且易产生大量误判，机器学习模型在样本极度不平衡场景下，常因欺诈样本不足导致模型偏向正常样本，出现“高漏检率”问题[3]。

信用卡欺诈检测数据一个很重要的特点是样本不平衡，公开数据集中欺诈样本占比普遍低于 0.2%，模型易过拟合正常样本，针对样本不平衡问题，可以采用过采样，通过增加少数类或减少多数类平衡数据分布。传统过采样方法 SMOTE 通过插值生成虚拟欺诈样本，但易产生“边界模糊”问题，导致模型泛化能力下降[4]；ADASYN 通过加权生成困难样本，虽改善 SMOTE 缺陷，但对噪声敏感[5] [6]。通过调整模型损失函数或参数，提升对少数类的关注，极端梯度提升算法(XGBoost)的参数可设置少数类权重，但仅靠参数调整难以完全解决样本不平衡[7]；Focal Loss 通过降低易分类样本的权重，增强模型对困难样本的学习，但在小样本场景下效果有限[8]。另外，可以利用深度学习生成高质量少数类样本，Goodfellow 等[9]提出的 GAN 通过生成器与判别器的对抗训练，可生成贴近真实分布的样本。此后，针对 GAN 的改

进算法开始应用到信用卡欺诈样本生成中[10][11]。

信用卡欺诈检测常用模型可分为传统机器学习与深度学习两类。在机器学习方面, XGBoost 等集成模型因强特征学习能力和可解释性, 成为欺诈检测的主流选择。户媛姣等[12]使用 XGBoost 作为分类算法, 有效改善了类别不平衡场景下信用卡欺诈的精准检测问题; 张海洋等[13]将 XGBoost 与逻辑回归结合, 提高了信用卡欺诈检测的准确率。在深度学习方面, CNN、LSTM 等模型可挖掘时序特征, 但需大量数据支撑, 且可解释性差。丁卫星等[14]构建基于深度置信网络的交易欺诈侦测原型系统, 确认了基于深度学习技术训练欺诈评分模型的可行性和有效性。

本研究针对信用卡欺诈检测中样本极度不平衡与特征区分度不足的问题, 以提升检测精度为目标展开研究: 首先, 基于公开信用卡欺诈数据集开展数据预处理与特征工程, 提出“交易小时 + 交易间隔统计”的精细化时序特征组合, 通过刻画欺诈行为的时间关联性弥补原始特征的信息缺口, 增强欺诈与正常交易的区分度; 其次, 为缓解样本不平衡问题, 构建定向生成对抗网络生成欺诈样本, 通过特征约束定向生成“小额漏检”与“大额欺诈”样本, 并使用余弦相似度进行高质量样本筛选, 显著提升过采样效果; 接着, 优化极端梯度提升树的关键参数, 构建高精度分类模型, 满足实际金融场景的实时性需求; 最后, 通过多组对比实验证所提模型的有效性, 并深入分析关键特征对检测性能的影响, 形成一套兼顾精度与实用性的信用卡欺诈检测方案。

2. 相关工作

2.1. 生成对抗网络

生成对抗网络(GAN, Generative Adversarial Networks)由 Goodfellow 等[8]于 2014 年提出, 核心思想是通过生成器(Generator)与判别器(Discriminator)的对抗训练, 生成贴近真实数据分布的样本。生成器定义输入噪声变量 z 的先验分布 $p_z(z)$, 再将数据空间的映射表示为 $G(z; \theta_g)$, 其中 G 是由含参数 θ_g 的多层感知机构成的可微函数。判别器定义多层感知机 $D(x; \theta_d)$, 其输出为单个标量。 $D(x)$ 代表 x 来自真实数据而非生成分布 p_g 的概率。

GAN 的训练目标是最小化生成器损失与最大化判别器损失, 形成极小极大博弈, 其损失函数为:

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log (1 - D(G(z)))]$$

其中, E 表示期望, $p_{data}(x)$ 为真实样本分布, $p_z(z)$ 为噪声分布。实际训练中, 采用二元交叉熵计算损失, 通过 Adam 优化器交替更新 G 和 D 的参数。

2.2. 极端梯度提升树

极端梯度提升树(XGBoost)是 Chen 等[6]提出的集成学习算法, 基于梯度提升决策树改进, 通过正则化、并行计算提升模型性能与效率, 在分类任务中表现优异。XGBoost 通过迭代训练弱分类器, 每次训练都拟合前一轮模型的残差。其目标函数包含损失项与正则项:

$$L(\Phi) = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{j=1}^m \Omega(f_j)$$

其中, $\sum_{i=1}^n l(y_i, \hat{y}_i)$ 为损失函数, $\sum_{j=1}^m \Omega(f_j)$ 为正则项, 控制树的复杂度避免过拟合。

3. 检测模型

3.1. 模型整体框架

本文提出的信用卡欺诈检测模型分为 4 个核心模块: 数据预处理与特征工程、定向 GAN 过采样、

XGBoost 分类、模型评估。整体框架如图 1 所示, 各模块协同解决“样本不平衡”与“特征区分度不足”问题。

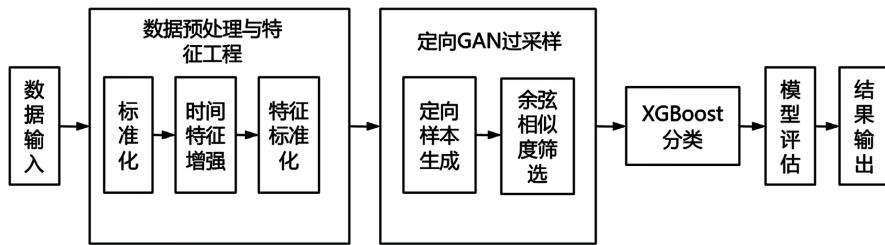


Figure 1. Detection model framework

图 1. 检测模型框架

3.2. 数据预处理与特征工程

本文采用 Kaggle 公开 Credit Card Fraud Detection 数据集[15], 该数据集包含 2013 年 9 月欧洲信用卡交易记录, 共 284,807 条样本, 每条样本含 30 个特征。其中, 匿名特征 V1~V28 通过 PCA 降维得到, 保护用户隐私; 原始特征 Time 为交易时间, 从首笔交易开始的秒数; Amount 为交易金额。样本中, Class 为标签, 1 为欺诈, 0 为正常。本文为了消除量纲影响对 Time、Amount 及 V1~V28 进行了标准化处理。

针对原始特征缺乏时序信息的问题, 设计时间特征增强, 新增 4 个精细化时序特征, 刻画欺诈行为的时间关联性:

(1) 交易小时(hour): 将 Time 列使用公式转换为小时, 捕捉欺诈的时段分布;

$$\text{hour} = \frac{\text{Time}}{3600} \bmod 24$$

(2) 临时用户标识(temp_user): 通过“hour + Amount 区间”构建临时用户, 模拟用户交易行为的连续性, 刻画欺诈的时段偏好;

(3) 交易间隔(tx_interval): 按 temp_user 分组, 计算当前交易与前一笔交易的时间差, 首笔交易间隔设为 3600 秒, 捕捉短间隔连续交易;

(4) 交易间隔统计(tx_interval_mean、tx_interval_std): 按 temp_user 分组, 计算最近 3 笔交易的间隔均值与标准差, 捕捉短间隔连续交易的欺诈模式。

3.3. 定向 GAN 过采样模型

3.3.1. GAN 网络设计

针对样本不平衡, 设计 GAN 生成高质量欺诈样本, 生成器和判别器网络结构如下图 2:

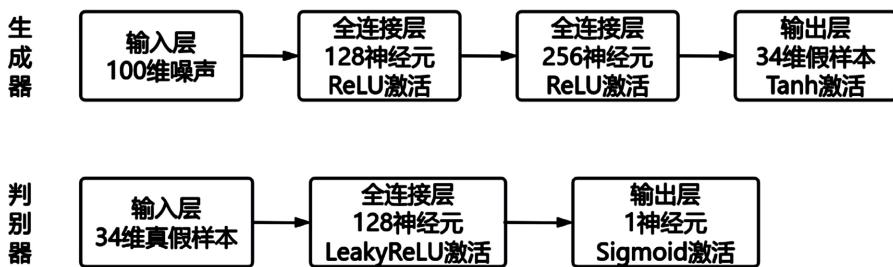


Figure 2. Generator and discriminator network structure

图 2. 生成器与判别器网络结构

3.3.2. 定向生成策略

使用 XGBoost 采用平衡策略对数据集进行测试后发现, 漏检样本存在“小额漏检”与“大额欺诈”两类关键模式, 见图 3 和表 1。

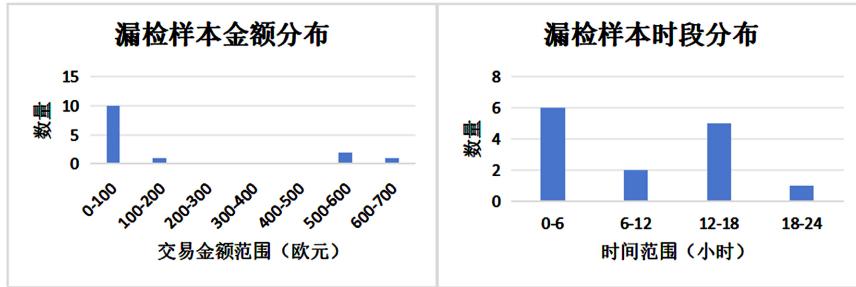


Figure 3. Missed detection samples statistical chart
图 3. 漏检样本统计图

Table 1. Missed detection samples statistical table
表 1. 漏检样本统计表

模式类型	Amount (元)	hour (时段)	漏检率占比
小额漏检样本	0~100	6~23	57.1%
大额欺诈样本	>500	0~6, 13~23	14.3%

从图 3 和表 1 中可以看出, 大部分漏检是小额交易。结合 hour 统计, 这些小额交易集中在 6~13 点(早高峰)、13~23 点(消费高峰), 属于用户非典型小额高频交易场景, 模型易漏检。而大额交易集中在 0~6 点(凌晨)、13~23 点(消费高峰但金额异常), 属于大额时段异常欺诈模式, 模型同样易漏检。

对于传统 GAN 生成样本随机, 无法覆盖“小额漏检”与“大额欺诈”两类关键模式, 本文通过特征约束实现定向生成。基于实验中对漏检样本的特征分布分析, 定向样本生成策略针对小额漏检模式, 重点生成金额处于 0~100 元、交易时段覆盖 6~13 点与 13~23 点、交易间隔匹配漏检样本集中分布的短间隔特征的样本, 以补充模型对小额高频活跃时段交易这类低特征显著性模式的学习; 针对大额欺诈模式, 定向生成金额落在 500~1000 元区间、时段聚焦 0~6 点与 13~23 点、交易间隔贴合漏检样本中观分布特征的样本, 以强化模型对大额高风险时段交易这类异常组合模式的捕捉能力。该策略通过精准覆盖漏检样本的核心特征区间, 避免传统过采样对非关键模式的冗余生成, 有效提升模型对两类核心欺诈模式的识别精度。

3.3.3. 样本筛选

为排除低质量伪样本, 采用余弦相似度筛选生成样本, 计算真实欺诈样本的均值向量 \bar{x}_{fraud} , 即

$$\bar{x}_{\text{fraud}} = \frac{1}{n} \sum_{i=1}^n x_{\text{fraud},i}$$

其中, n 为真实欺诈样本数量, $x_{\text{fraud},i}$ 为第 i 个真实欺诈样本的特征向量, \bar{x}_{fraud} 代表真实欺诈模式的特征中心。计算每个生成样本 x_{fake} 与 \bar{x}_{fraud} 的余弦相似度, 保留相似度大于等于 0.8 的生成样本, 确保生成样本的欺诈特征显著。

3.4. XGBoost 分类模型

考虑到信用卡欺诈检测的实时性需求, 且 XGBoost 在小样本不平衡数据上具有训练快、泛化能力强、

可解释性高的优势，本文采用 XGBoost 作为分类器，适配经过定向 GAN 过采样的平衡训练集。

为提升模型对欺诈样本的检测精度，针对学习率、树深度、正则化系数等关键参数，通过网格搜索确定最优组合，平衡模型复杂度与过拟合风险；考虑到正负样本成本差异，不采用默认阈值，通过 F1 值最大化确定最优预测阈值，提升欺诈样本的召回率。

4. 实验设计与结果分析

4.1. 实验环境与评价指标

本文的实验采用 Windows 11 系统，Intel(R) Core(TM) i5-11400H CPU，内存 16 GB，显卡 NVIDIA GeForce RTX 3050，python 采用 3.7 版本，scikit-learn 采用 1.0.2 版本，pytorch 采用 1.13.1 版本。

针对信用卡欺诈检测的不平衡特性，选用 4 个核心评价指标，避免单一准确率(*Accuracy*)的误导，包含精确率(*Precision*)、召回率(*Recall*)、F1 值(*F1-Score*)、AUPRC (Area Under the Precision-Recall Curve)。

精确率判定为欺诈的样本中，实际为欺诈的比例，反映模型的误判率，计算方法如下：

$$Precision = \frac{TP}{TP + FP}$$

召回率实际为欺诈的样本中，被正确判定的比例，反映模型的漏检率：

$$Recall = \frac{TP}{TP + FN}$$

F1 值精确率与召回率的调和平均，平衡两者关系，适合不平衡数据：

$$F_1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}$$

AUPRC 是二分类模型中用于评估性能的重要指标，尤其适用于类别不平衡的场景。其核心是通过计算不同召回率下的精度 - 召回率曲线下的面积，综合反映模型的预测能力。

其中，*TP* (True Positive) 为正确检测的欺诈样本，*FN* (False Negative) 为漏检的欺诈样本，*FP* (False Positive) 为误判的正常样本，*TN* (True Negative) 为正确检测的正常样本。

4.2. 对比实验设计

为验证本文方案的优势，设计 4 组对比模型，所有模型均基于相同的训练集与测试集，仅在过采样方法和时间特征利用上存在差异，确保对比公平性：

- (1) Base 模型：无过采样的 XGBoost，含/不含时间特征增强，设置 *scale_pos_weight* 参数平衡类别权重；
- (2) SMOTE-XGBoost：SMOTE 过采样 + XGBoost，含/不含时间特征增强，过采样后正负样本比例为 1:1；
- (3) GAN-XGBoost：传统随机 GAN 过采样 + XGBoost，含/不含时间特征增强，生成样本无定向约束；
- (4) 本文方案：时间特征增强 + 定向 GAN 过采样 + XGBoost，含/不含时间特征增强。

所有模型的测试集性能结果见表 2 所示，本文方案在精确率、F1 值、误判数上均显著优于对比模型，且召回率保持稳定：

本文方案在综合性能与时间特征利用上展现出显著优势。其精确率达 98.81%，F1 值达 91.21%，较 Base 模型(含时间特征增强)提升 10.26 个百分点，较 SMOTE-XGBoost (含时间特征增强)提升 38.58 个百分点，充分证明方案的综合性能优势；召回率稳定在 84.69%，在低误判的同时仍能覆盖 85% 左右的欺诈

交易,实现了用户体验与风险控制的业务平衡。此外,时间特征增强对性能提升的作用至关重要:所有含时间特征增强的模型F1值均高于无时间特征版本,其中Base模型提升15.58个百分点、SMOTE-XGBoost提升15.48个百分点、本文方案提升4.25个百分点,数据说明时间特征增强有效捕捉了欺诈行为的时间异常模式。

Table 2. Test results of comparative experiments
表2. 对比实验测试结果

模型类型	精确率(%)	召回率(%)	F1值(%)	AUPRC
Base模型(含时间特征增强)	75.89	86.73	80.95	0.8720
Base模型(无时间特征增强)	52.83	85.71	65.37	0.8507
SMOTE-XGBoost(含时间特征增强)	37.78	86.73	52.63	0.8677
SMOTE-XGBoost(无时间特征增强)	23.56	87.76	37.15	0.8388
GAN-XGBoost(含时间特征增强)	96.51	84.69	90.22	0.8681
GAN-XGBoost(无时间特征增强)	92.05	82.65	87.10	0.8795
本文方案(含时间特征增强)	98.81	84.69	91.21	0.8681
本文方案(无时间特征增强)	93.02	81.63	86.96	0.8793

通过采样方法的对比结果凸显了定向GAN的技术价值。SMOTE过采样因生成边界噪声,精确率仅23.56%~37.78%,难以满足金融场景对低误判的需求;而GAN类过采样通过生成器与判别器的对抗训练生成高质量样本,精确率达92.05%~98.81%,显著优于SMOTE。其中,本文定向GAN较传统GAN进一步提升精确率2.3个百分点,证明定向生成策略能精准覆盖小额漏检与大额欺诈两类核心漏检模式,为极不平衡数据下的样本增强提供了高效且具针对性的技术路径。

4.3. 消融实验

为明确时间特征增强和定向GAN过采样两个核心模块的贡献,设计消融实验,结果如表3所示:

Table 3. Results of ablation experiments
表3. 消融实验的结果

实验方案	精确率(%)	召回率(%)	F1值(%)	F1值变化(百分点)	核心结论
本文方案	98.81	84.69	91.21	-	全模块协同最优
去掉“时间特征增强”	93.02	81.63	86.96	-4.25	时间特征增强提升F1值4.25个百分点
去掉“定向GAN”	96.51	84.69	90.22	-0.99	定向GAN提升F1值0.99个百分点

消融实验表明:时间特征增强贡献最大,去除后F1值下降4.25个百分点。定向GAN不可或缺,去除后F1值下降0.99个百分点。

5. 结论与展望

本文围绕信用卡欺诈检测的样本不平衡与特征区分度不足问题展开研究,提出融合时间特征增强与定向GAN过采样的XGBoost信用卡欺诈检测方案,并通过实验证其有效性。核心结论如下:针对数据层面的样本失衡与特征缺口,构建的四层时间特征增强体系,使短间隔连续交易和凌晨时段欺诈等模

式的识别率得到提升；针对过采样质量不足，定向 GAN 过采样通过特征约束精准覆盖小额漏检与大额欺诈模式，较传统随机 GAN 有所提升；方案在 Kaggle 数据集上实现 98.81% 的欺诈类精确率、84.69% 的召回率及 91.21% 的 F1 值，显著优于 Base、SMOTE-XGBoost 等对比模型；消融实验进一步证实，时间特征增强对性能贡献最大，定向 GAN 过采样为关键补充。

尽管本文方案在信用卡欺诈检测中表现优异，但仍存在可深化与扩展的方向：其一，可引入用户画像特征(如常用商户类型、设备指纹、信用评分)，构建“时序特征 + 用户行为特征”的多维度特征体系，进一步提升对非本人异常交易的区分度；其二，可构建 XGBoost + LightGBM + 随机森林的多模型集成框架，通过 Stacking 融合多模型预测结果，降低单一模型的过拟合风险，提升复杂欺诈模式的泛化检测能力。

基金项目

河北省社会科学基金项目“转型金融对河北省高碳企业低碳技术创新的驱动机制研究”(HB24ERJ027)。

参考文献

- [1] CardInfoLink (2025) How Does Risk Control Reshape Security Boundaries in the Face of Payment Fraud? <https://www.cardinfolink.com/en-US/newsroom/article/20>
- [2] Milo, T., Novgorodov, S. and Tan, W. (2016) Rudolf: Interactive Rule Refinement System for Fraud Detection. *Proceedings of the VLDB Endowment*, **9**, 1465-1468. <https://doi.org/10.14778/3007263.3007285>
- [3] Chandola, V., Banerjee, A. and Kumar, V. (2009) Anomaly Detection: A Survey. *ACM Computing Surveys*, **41**, 1-58. <https://doi.org/10.1145/1541880.1541882>
- [4] Chawla, N.V., Bowyer, K.W., Hall, L.O. and Kegelmeyer, W.P. (2002) SMOTE: Synthetic Minority Over-Sampling Technique. *Journal of Artificial Intelligence Research*, **16**, 321-357. <https://doi.org/10.1613/jair.953>
- [5] He, H.B., Bai, Y., Garcia, E.A. and Li, S.T. (2008) ADASYN: Adaptive Synthetic Sampling Approach for Imbalanced Learning. 2008 *IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence)*, Hong Kong, 1-8 June 2008, 1322-1328. <https://doi.org/10.1109/ijcnn.2008.4633969>
- [6] 顾明, 李飞凤, 王晓勇, 等. 基于改进 SMOTE 算法和深度学习集成框架的信用卡欺诈检测[J]. 贵阳学院学报(自然科学版), 2024, 19(2): 99-104.
- [7] Chen, T. and Guestrin, C. (2016) XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, 13-17 August 2016, 785-794. <https://doi.org/10.1145/2939672.2939785>
- [8] Lin, T., Goyal, P., Girshick, R., He, K. and Dollar, P. (2017) Focal Loss for Dense Object Detection. 2017 *IEEE International Conference on Computer Vision (ICCV)*, Venice, 22-29 October 2017, 2999-3007. <https://doi.org/10.1109/iccv.2017.324>
- [9] Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al. (2014) Generative Adversarial Nets. *Advances in Neural Information Processing Systems 2014*, Montreal, 8-13 December 2014, 2672-2680.
- [10] 刘汝欣, 徐洪珍. 基于注意力机制优化的 WGAN-BiLSTM 信用卡欺诈检测方法[J]. 现代电子技术, 2024, 47(10): 73-78.
- [11] 李占利, 唐成, 靳红梅. aLMGAN-信用卡欺诈检测方法[J]. 计算机工程与设计, 2024, 45(3): 830-836.
- [12] 户媛姣, 牛梦园, 张留美. 融合蝙蝠优化与 XGBoost 的信用卡欺诈检测方法[J]. 西安邮电大学学报, 2025, 30(5): 51-60.
- [13] 张海洋, 陈玉明, 曾念峰, 等. 基于 XGBoost 与 LR 融合模型的信用卡欺诈检测[J]. 重庆理工大学学报(自然科学), 2024, 38(3): 195-200.
- [14] 丁卫星. 基于深度学习技术的信用卡交易欺诈侦测研究[D]: [硕士学位论文]. 上海: 上海交通大学, 2015.
- [15] Kaggle: Credit Card Fraud Detection. <https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud>