

基于机器学习的银行个人信用风险评估研究

薛琦, 罗鄂湘

上海理工大学管理学院, 上海

收稿日期: 2023年4月24日; 录用日期: 2023年7月11日; 发布日期: 2023年7月18日

摘要

本文运用CCF竞赛提供的中原银行个人信用贷款违约数据, 进行了数据清洗和特征工程的工作, 从初始的38个特征缩减到18个特征, 结合5C理论和预期收入理论探究了影响银行个人信用风险的重要因素, 经过特征重要性排序排名前五的因素是: 信贷周转余额合计、贷款发放日期距初始日期天数、借款人贷款评分平均分、当前贷款利率和匿名变量f0。为提升银行对个人信用风险评估的准确率, 本文基于随机森林模型比较了SMOTE、随机欠采样和SMOTEENN三种非平衡数据的处理方法进行实验, SMOTEENN组合采样的效果最好; 然后建立了决策树、随机森林、AdaBoost和LightGBM共4个机器学习模型, 结果表明平衡后LightGBM的准确率最高, 达到了96.1%。

关键词

机器学习算法, 个人信用风险评估, 特征工程, SMOTEENN

Research on Bank Personal Credit Risk Assessment Based on Machine Learning

Qi Xue, Exiang Luo

Business School, University of Shanghai for Science and Technology, Shanghai

Received: Apr. 24th, 2023; accepted: Jul. 11th, 2023; published: Jul. 18th, 2023

Abstract

In this paper, using the personal credit loan default data of Zhongyuan Bank provided by the CCF competition, the data cleaning and feature engineering was carried out and the initial 38 features were reduced to 18 features. Then the important factors affecting the bank personal credit risk were explored by combining the 5C theory and expected income theory, and the top five factors ranked by feature importance were: total credit working balance, loan disbursement date according to the initial date days, borrower's average loan score, current loan interest rate and anonymous

variable f0. In order to improve the accuracy of bank personal credit risk assessment, this paper compared three methods of processing unbalanced data, SMOTE, random under sampling and SMOTEENN, based on the random forest model, and SMOTEENN combined sampling had the best effect; then a total of four machine learning models, decision tree, random forest, AdaBoost and LightGBM, were established and it's showed that LightGBM had the highest accuracy rate after balancing, reaching 96.1%.

Keywords

Machine Learning Algorithms, Personal Credit Risk Assessment, Feature Engineering, SMOTEENN

Copyright © 2023 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

随着中国经济的发展,人们开始追求更有品质的生活。政府为了刺激经济和鼓励金融创新,也逐渐放宽了金融监管政策,使得近年来银行的贷款总额始终呈现上升趋势。根据中国银保监会统计数据,2022年末的贷款总额达到了1,827,315亿元,同比年初增长了179,152亿元。但与此同时,银行也愈发受到信用风险的威胁,商业银行的不良贷款余额始终呈现上升趋势。同时数据也显示,2022年末的不良贷款余额达到了29,829亿元,较年初增长了1359亿元。为了提升个人信用风险评估模型准确度,降低银行面临的个人信用风险,专家学者主要在探究个人信用风险影响因素和改进评估方法两个方面做出了许多努力。

在对影响因素的探究上,顾洲一等(2022)发现客户忠诚度、客户历史信贷数据、信贷合同特征是和个人信用风险评估高度相关的三个重要特征[1]。Wang等(2022)发现首付款金额、担保余额、抵押物价值(元)、担保方式、房屋数量、月房产成本、家庭月收入、分期还款方式、审批期限、是否欠息、账户是否贬值、安全系数是相对重要的因素[2]。张丽颖等(2022)基于Kaggle的Loan Defaulter数据集,发现性别、家庭人数以及借款人所在城市、住房类型、总收入、所属行业、职业类型、工作年限、受教育程度、消费贷款额度、贷款金额、贷款年金等都是影响客户违约的重要因素[3]。

针对评估方法,Durand(1941)最早使用了线性判别法评估个人信用风险[4],Davis(1992)发现分类树机器学习算法优于传统算法[5]。后续越来越多的专家开始使用机器学习算法进行个人信用风险评估,到今天,各种集成算法逐渐被广泛应用。张承钿等(2022)结合集成算法的思想和XGBoost低偏差的特性设计了一种异构集成模型(HEM)来评估个人信用,通过UCI德国个人信用数据集验证,模型相对于传统的LR、DT、SVM和XGBoost集成模型的AUC值平均提升了7.38% [6]。深度学习算法也逐渐走入大家的视线,Dastile等(2021)将表格数据集转换为图像,使用2D卷积神经网络(CNNs)进行了信用评估,结果优于其他信用评分方法[7]。

本文根据CCF竞赛提供的中原银行个人信用贷款违约数据,基于5C理论和预期收入理论分析了影响个人信用风险评估的主要因素。同时使用多种平衡数据集的采样方法,基于多种机器学习算法对银行个人信用风险进行评估,通过比较模型精度,选择准确度更高的模型。

2. 银行个人信用风险影响因素分析

信用风险(Credit Risk)是指由于借款人在偿还借款或债务时可能无法按时、足额偿还本金和利息,而

对另一方造成财务损失的可能性或经济损害的风险, 是商业银行是最为关键和复杂的风险。银行个人信用风险指银行在向个人提供信贷时, 可能遭受到的潜在损失的风险, 即个人贷款违约的风险。这种风险是由于借款人从银行等金融机构获得信用贷款后, 没有按照规定时间还款而导致的[8]。

银行个人信用风险会给银行带来损失, 因此银行会在发放贷款之前进行信用评估。常用的评估理论是“5C”理论, 即针对借款人的品格(Character)、能力(Capacity)、资本(Capital)、担保(Collateral)和条件(Conditions) 5个方面进行信用评估。同时, 预期收入理论由 H.V. Prochnow 在 1945 年首次提出, 该理论认为, 借款人的预期收入能够被作为偿还贷款能力的衡量标准。如果预期收入水平高, 借款人则会倾向于提前还款以减少利息支出, 银行也很乐意向其提供与预期收入相匹配的贷款金额。

本文所用数据来自 CCF 大数据与计算智能大赛提供的, 中原银行实际业务场景下的个人信用贷款违约数据。该数据集包含了 1 万条脱敏记录, 共有 38 个特征变量和 1 个标签, 包括用户基本属性信息和借贷相关信息。由于数据来自实际场景, 数据集保证了样本的真实性和客观性。

根据 5C 理论和预期收入理论将 38 个特征变量进行分类, 如表 1 所示。同时, 由于本文使用的是非抵押信用贷款数据集, 不存在担保(Collateral)这一维度的数据。

Table 1. Feature classification based on 5C and expected income theory

表 1. 基于 5C 和预期收入理论的特征分类

类别	因素
品格 Character	del_in_18month 借款人过去 18 个月逾期 30 天以上违约事件数、scoring_low 借款人在贷款评分中所属的下限范围、scoring_high 借款人在贷款评分中所属的上限范围、known_outstanding_loan 借款人档案中未结信用额度的数量、known_dero 贬损公共记录数量、pub_dero_bankrup 公开记录清除数量
能力 Capacity	employment_type 所在公司类型(世界五百强、国有企业、普通企业)、industry 工作领域(传统工业、商业、互联网、金融)、work_year 工作年限、debt_loan_ratio 债务收入比
资本 Capital	home_exist 房屋状态
条件 Conditions	total_loan 贷款数额、year_of_loan 贷款年份、interest 当前贷款利率、monthly_payment 分期付款金额、class 贷款级别、censor_status 审核情况、issue_date 贷款发放的日期、use 贷款用途类别、recircle_b 信贷周转余额合计、recircle_u 循环额度利用率
预期收入	early_return 借款人提前还款次数、early_return_amount 借款人提前还款累积金额、early_return_amount_3mon 近 3 个月内提前还款金额
其他	loan_id 贷款记录唯一标识、user_id 借款人唯一标识、initial_list_statu 贷款的初始列表状态、app_type 是否个人申请、earlies_credit_mon 借款人最早报告的信用额度开立的月份、title 借款人提供的贷款名称、policy_code 是否公开可用策略、post_code 借款人申请时邮政编码、region 地区编码、f 系列匿名特征匿名特征 f0~f4, 为一些借款人行为计数特征的处理

3. 建模过程

本文使用了 Python 语言作为分析和建模的工具, 建模流程如图 1 所示。

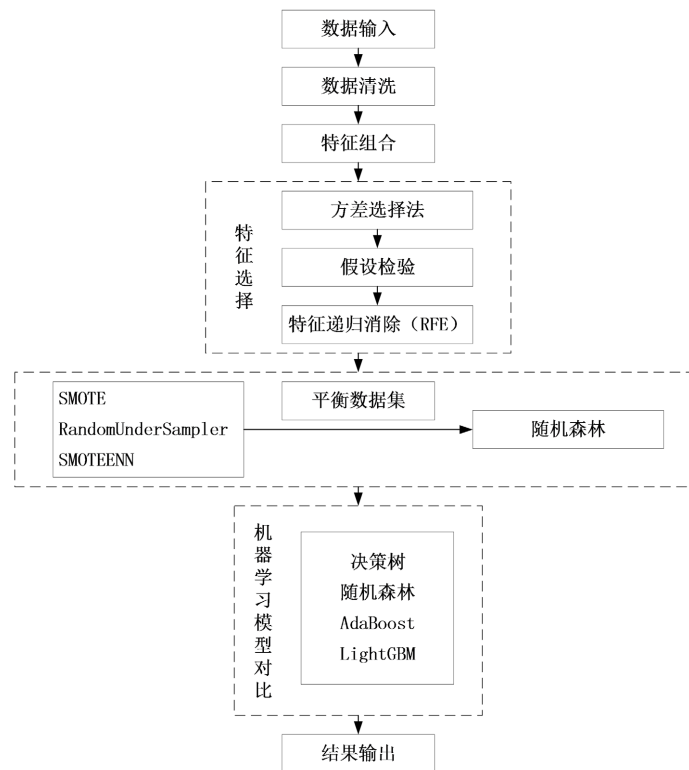


Figure 1. Modeling process

图 1. 建模流程图

3.1. 数据清洗

数据清洗(Data Cleaning)是指检测和纠正数据中存在的缺失、异常值、重复值等。它的主要目的是提高数据的质量,确保数据准确、完整、一致。只有在充分检查和清理数据后,才能提高数据的质量和可靠性,以便于后续的数据分析和建模。本文的数据清洗主要进行了以下步骤。

3.1.1. 不相关的变量处理、数据类型的转换

本文首先删除了不相关的变量。贷款记录唯一标识 `loan_id` 和借款人唯一标识 `user_id` 无实际意义。信用额度开立时间 `earliest_credit_mon` 仅为额度开立的时间,与是否贷款、何时贷款没有直接关系,且该数据存在一定的年份缺失问题。

然后进行了数据类型的转换,通过构建工作年限 `work_year`、贷款等级 `class`、工作类型 `employer_type` 和所在行业 `industry` 的字典,结合 `map` 函数的使用进行数值化处理。并将贷款发放时间 `issue_date` 使用 `to_datetime` 函数转换成标准日期格式,然后将其分成年份 `issue_date_year` 和月份 `issue_date_mon`; 将贷款发放时间 `issue_date` 和设定的初始 `base_time` 相减,将日期转化数值型数据天数 `issue_date_diff`; 在完成上述两步后,删除原始字段 `issue_date`。

在完成上述处理后,共剔除了 4 个特征,增加了 3 个特征,得到了 37 个特征。

3.1.2. 缺失值、异常值处理

本文数据集中有 7 列存在缺失值,根据不同特征的特点,我们对工作年限 `work_year` 使用众数进行填充,对公开记录中清除数量 `pub_dero_bankrup` 和匿名变量 `f0~f4` 使用中位数来补齐。由于匿名变量 `f0~f4` 的缺失表现出了极强的相关性,说明 `f0~f4` 的缺失存在某种关系,具体将在特征工程部分进行分析。

而针对离群值, 在信用评估领域中, 它们的存在通常具有意义, 可能会包含其他有用的信息, 比如一些潜在的风险。因此, 我们不会轻易对其进行处理。而对于方差明显过小的特征, 例如公开可用的策略 `policy_code`、匿名变量 `f1` 等, 在后续的特征选择中一并进行处理。同时针对不一致的数据, 即样本中存在提前还款次数 `early_return` 为 0, 但提前还款金额 `early_return_amount` 却不为 0 的情况。鉴于这些样本中提前还款金额具体可见, 我们猜测使该部分提前还款次数填写错误。因此, 我们将这部分提前还款次数由 0 更正为 1。

本文缺失值和异常值的处理仅进行了填补和修改, 此时仍剩余 10000 个样本和 37 个特征。

3.2. 特征组合

有时候将不同的特征组合在一起, 可以得到更有信息量和预测能力的特征变量, 进而改善模型的性能。结合前文的分析, 本文针对该数据集作如下几点特征组合:

① 评分上下限两个特征 `scoring_low` 和 `scoring_high` 相关性较高, 但代表程度却较低。因此我们将评分上下限组合起来, 通过加权平均值构造特征平均分 `scoring_ave`。

② 在缺失值处理时, 我们看出 `f0~f4` 的缺失之间似乎存在某种联系。我们尝试以 `f0` 和 `f1` 为例构建了特征 `f_miss`。具体规则如下: 当取值为 0 时表示 `f1` 和 `f0` 均不缺失, 当取值为 1 时表示 `f1` 和 `f0` 均缺失, 而当取值为 2 时表示 `f1` 缺失而 `f0` 不缺失。

③ 为了更好地衡量“5C”理论条件(Condition)因素中外部市场环境变化的影响, 我们考虑构建特征贷款结束年份 `end_year`, 用贷款结束的时间的来表示不同时期的外部市场环境。将贷款发放时间 `issue_date_year` 和贷款年限 `year_of_loan` 相加来表示, 即 $end_year = issue_date_year + year_of_loan$ 。

④ 通过观察 `early_return` 与 `isDefault` 的柱状分布图, 发现 `early_return` 是否大于 0 对违约的区分作用十分明显。即有提前还款行为的客户违约率远低于没有提前还款行为的客户, 因此本文在提前还款次数 `early_return` 基础上新增是否提前还款 `is_early_return` 这一特征。

经过特征组合, 增加了 4 个特征并删除了 2 个特征, 得到了 39 个特征。

3.3. 特征选择

在进行数据清洗和特征组合后, 此时还剩下 39 个特征。过多的特征中往往存在与预测无关的特征。根据问题的不同, 从原始数据中挑选出最具有判别性、与目标变量相关的特征, 同时滤除不相关或者噪声特征, 以提高模型的泛化能力和处理效率。具体的特征选择步骤如下。

3.3.1. 方差选择法(Variance-Based Method)

方差选择法根据特征的方差大小排序, 选择方差较大的特征。因为方差较小的特征说明其数据集中度较高, 缺乏有效信息, 而方差较大的特征则可能有更好的区分性。

此处筛选掉了方差小于 0.1 的特征, 包括公开可用的策略 `policy_code`, 匿名特征 `f1` 和是否个人申请 `app_type` 三个特征。其中 `policy_code` 具有唯一取值 1, 因此方差为 0, `f1` 的方差为 0.001, `app_type` 的方差为 0.020。方差太小说明它们过于稳定, 对结果的预测不起作用因此删除, 此时剩余 36 个特征。

3.3.2. 假设检验(Hypothesis Testing)

假设检验(Hypothesis Testing)分为 t 检验、F 检验、卡方检验等, 将计算得到的统计量与置信区间和显著性水平进行比较, 显著性水平是指原假设被拒绝的概率, 通常取值为 0.05。即当 p 值低于 0.05, 拒绝原假设, 说明该特征和目标变量显著相关。

本文对数值型变量进行了 F 检验, 从 26 个数值型特征中删除了所有 p 值小于 0.05 的特征, 共计 9

个变量。包括了贷款数额 `total_loan`, 分期付款金额 `monthly_payment`, 借款人申请时邮政编码 `post_code`, 地区编码 `region`, 借款人过去 18 个月逾期 30 天以上的违约事件数 `del_in_18month`, 借款人提供的贷款名称 `title`, 匿名特征 `f2`, `f4`, 贷款发放月份 `issue_date_mon`, 选择了剩余 17 个数值型变量。

然后对类别变量使用 `scipy` 库中的 `stats.chi2_contingency` 函数进行卡方检验, 从 10 个特征删除了所有 `p` 值小于 0.05 的特征, 共计 5 个变量。包括了所在公司类型 `employer_type`, 工作领域 `industry`, 工作年限 `work_year`, 贷款用途类别 `use`, 贷款的初始列表状态 `initial_list_status`, 选择了剩余 5 个类别变量。

3.3.3. 特征递归消除

递归特征消除(Recursive Feature Elimination, 简称 RFE)是一种基于机器学习的特征选择方法。该方法通过不断地训练模型, 逐步删除权值较低的特征, 得到新的特征子集, 直到剩下指定数量或达到指定阈值的特征个数为止。

本文使用 `sklearn.ensemble` 库中的 `RandomForestClassifier` 函数, 以随机森林作为基分类器进行多次训练。每进行一次训练, 删除其中最不重要的特征, 并利用决策树算法计算准确率。通过准确率的比较, 确定了选择 18 个特征如表 2 所示, 此时准确率 `Accuracy` 达到了 81.3%。

Table 2. Important factors affecting bank personal credit risk after feature selection

表 2. 特征选择后影响银行个人信用风险的重要因素

类别	因素
品格 Character	<code>Scoring_ave</code> 借款人贷款评分平均分、 <code>known_outstanding_loan</code> 借款人档案中未结信用额度的数量
能力 Capacity	<code>debt_loan_ratio</code> 债务收入比
资本 Capital	<code>home_exist</code> 房屋状态
条件 Conditions	<code>interest</code> 当前贷款利率、 <code>class</code> 贷款级别、 <code>tensor_status</code> 审核情况、 <code>issue_date_year</code> 贷款发放年份、 <code>issue_date_diff</code> 贷款发放日期距初始日期天数 <code>end_year</code> 贷款结束的年份、 <code>recircle_b</code> 信贷周转余额合计、 <code>recircle_u</code> 循环额度利用率
预期收入	<code>early_return</code> 借款人提前还款次数、 <code>early_return_amount</code> 借款人提前还款累积金额、 <code>early_return_amount_3mon</code> 近 3 个月内提前还款金额 <code>is_early_return</code> 是否提前还款
其他	匿名变量 <code>f0</code> 、 <code>f3</code>

3.4. 非平衡数据集处理

3.4.1. 数据非平衡性判断

针对本文的标签“是否违约”, 在数据集中没有发生违约行为的样本有 8317 条, 发生违约的仅为 1683 条, 样本比例接近 5:1, 存在一定不平衡问题。在这种情况下, 由于多数类样本过多, 使得分类器不能很好地学习少数类中重要特征, 而更加倾向于将所有样本都分为多数类。这会影响模型的准确率和泛化能力, 需要在后续进行非平衡数据处理。

3.4.2. 平衡前后模型结果对比

本文利用了 SMOTE 过采样、随机欠采样和 SMOTEENN 组合采样三种方式来对数据集进行不平衡

性的处理。

SMOTE (Synthetic Minority Over-Sampling Technique)是一种过采样技术, 它通过少数类样本之间的距离来创建合成样本, 来增加其数量, 从而使数据更加平衡; 随机欠采样技术(Random Under-Sampling)则是通过从多数类中随机选择一个子集来减少多数类样本数的一种方法; 组合采样(Combining Sampling)本质上是将欠采样和过采样方法的优势结合起来, 通过从多数类别中欠采样, 从少数类别中过采样, 生成一组新的平衡数据集, SMOTEENN 就是一个很好的例子, 它结合了过采样(SMOTE)和欠采样(ENN)两种技术, 提供了更加平衡的数据集。

将没有平衡处理的样本以及三种采样方法后得到的样本, 按照 6:4 划分为训练集和测试集, 把训练集放到随机森林模型中进行拟合, 用测试集进行模型评估, 表现如表 3 所示。

Table 3. Performance of different sampling methods on Random Forests

表 3. 不同采样方法在随机森林上的表现

数据集	Accuracy	F1-Score	AUC
没有经过平衡处理	0.849	0.452	0.658
SMOTE 过采样	0.876	0.882	0.876
随机欠采样	0.795	0.700	0.776
SMOTEENN 组合采样	0.959	0.964	0.956

表 3 中数据表明, 在随机森林模型上 SMOTEENN 组合采样后的三个评估指标均达到了 0.9 以上, 效果远好于 SMOTE 采样和随机欠采样。其中随机欠采样虽在 F1-Score 和 AUC 上略有提升, 但准确率低于 0.8, 甚至比不处理更差。可能是由于违约样本的数量比较少, 欠采样会使整体数据量减少, 因此丢失了许多有价值的样本。所以最终选择了 SMOTEENN 方法来平衡本文的数据集,

3.5. 分类器对比分析

本节将经过 SMOTEENN 组合采样得到的样本, 放到决策树、随机森林、AdaBoost 和 LightGBM 模型中进行拟合。

分类与回归树算法(Classification and Regression Tree, CART)最早由 Breiman 等人提出[9], 采用基尼指数(Gini Index)来度量信息纯度, 其值越小代表纯度越高, CART 在每个内部节点上进行二分法划分, 直到达到预定条件; 随机森林(Random Forest)通过自助采样(Bagging)和随机子空间法(Random Subspace Method)生成多棵决策树, 采用投票法集成多个决策树来提高预测准确率和稳定性; AdaBoost (Adaptive Boosting)针对多个分类器训练数据是否正确进行加权重处理, 通过不断对错误样本提升权重的的方式, 使得分类器对该错误分类样本的重视程度不断提高, 直到达到预定的迭代次数或误差率满足某一阈值; LightGBM (LightGradient Boosting Machine)使用了“leaf-wise”的决策树生长策略, 在进行梯度下降的过程中采用了直方图加速的方式。此外, LightGBM 还支持类别特征和缺失值的自动处理, 同时支持分布式训练, 具有效率高、精准度高和可扩展性强的优点。

将四种分类器的预测精确度、F1-Score 和 AUC 值进行汇总比较, 结果如表 4 所示。同时绘制了如图 2 所示的 ROC 曲线。

根据表 4 和图 2 中的信息, 我们可以得出以下结论:

首先, 我们可以看到所有模型的准确率 Accuracy 得分都很高, 超过了 0.92。其中, LightGBM 模型具有最高的准确率得分, 为 0.961。

Table 4. Performance of different classifiers based on SMOTEENN
表 4. 基于 SMOTEENN 的不同分类器表现

分类器	Accuracy	F1-Score	AUC
决策树	0.927	0.935	0.927
随机森林	0.959	0.964	0.956
AdaBoost	0.958	0.962	0.956
LightGBM	0.961	0.966	0.959

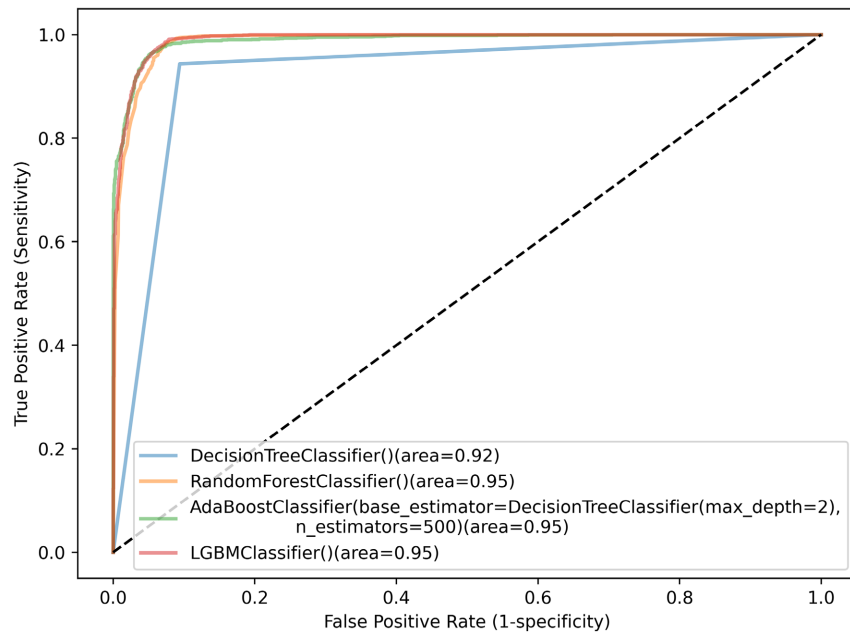


Figure 2. ROC curves for different classifiers based on SMOTEENN
图 2. 基于 SMOTEENN 的不同分类器 ROC 曲线

其次, 各模型在 F1-Score 上也表现出色, 都在 0.93 以上, 说明这些模型较好的平衡了准确率和召回率。具体而言, LightGBM 模型在 F1-Score 上表现最好, 达到了 0.966。

最后, 我们考虑 AUC 分数。图 2 中的 ROC 曲线可以通过可视化的方式, 简单的了解各模型间的效果差异, 因此可以首先看出决策树模型的效果最差, 但由于随机森林、AdaBoost、LightGBM 的结果比较相近, 因此需要通过表格中具体数值判断。从表格中可以看出, 在 AUC 得分方面, LightGBM 模型又一次表现出色, 达到了 0.959, 略高于其他三个模型。

因此综合来看, 我们认为在这四个模型当中, LightGBM 模型是表现最佳的。它在准确率、F1-Score 和 AUC 得分方面都表现出非常出色的表现, 同时拥有最高的得分。在实际应用中, 我们可以借鉴 LightGBM 模型的算法思路和特点, 并将该模型应用于银行个人信用风险评估领域中, 以提高分类器性能和分类效果。

3.6. 特征重要性排序

经过前两节的对比, 我们使用效果好的 SMOTEENN 组合采样和 LightGBM 分类器建立模型, 使用 plot_importance 函数得到特征排序如图 3 所示。

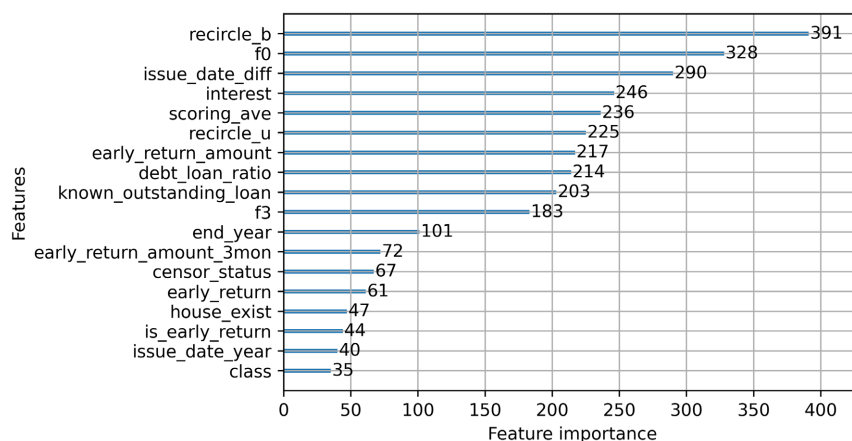


Figure 3. Feature Importance Ranking

图 3. 特征重要性排序

4. 结论

经过以上研究, 本文得出以下结论:

1) 针对银行个人信用风险的影响因素, 本文通过数据清洗和特征工程, 从 38 个特征里选出了 18 个特征如表 4 所示。除了匿名变量 f0 和 f3 由于表意不清无法分类外, 被选中的 18 个特征在 5C (除担保) 和预期收入理论中的类别中均有分布, 再次验证了结合 5C 理论和预期收入理论的科学性。同时根据图 3 特征重要性排序, 我们可以看出排在前五的因素为信贷周转余额合计、贷款发放日期据初始日期天数、借款人贷款评分平均分、当前贷款利率和匿名变量 f0, 说明在银行个人信用评估中, 借款本身的具体条件和借款人的品格是决定是否违约的较重要的因素, 应当给予更高的权重。

2) 对于非平衡数据处理, SMOTEENN 组合采样在准确率、F1-Score 和 AUC 值上的表现均达到了 0.9 以上, 效果远好于 SMOTE 采样和随机欠采样, 是效果最优的非平衡数据处理方式。模型选择上, LightGBM 的效果最佳, 准确率达到 96% 以上, 准确性较高; F1-Score 为 0.966, 说明模型对准确率和召回率的平衡比较好; AUC 值为 0.959, 说明对是否违约具有很好的判断识别能力。综上, 本文选择了经过 SMOTEENN 组合采样后的样本构建了 LightGBM 模型, 具有很好的准确性和泛化能力。

参考文献

- [1] 顾洲一, 胡丽娟. 机器学习视角下商业银行客户信用风险评估研究[J]. 金融发展研究, 2022(1): 79-84.
- [2] Wang, T., Liu, R. and Qi, G. (2022) Multi-Classification Assessment of Bank Personal Credit Risk Based on Multi-Source Information Fusion. *Expert Systems with Applications*, **191**, Article ID: 116236. <https://doi.org/10.1016/j.eswa.2021.116236>
- [3] 张丽颖, 杨若瑾. 基于机器学习的个人贷款违约预测模型的应用研究[J]. 金融监管研究, 2022(6): 46-59.
- [4] Risk, D.D. (1941) *Elements in Consumer Instalment Financing*. National Bureau of Economic Research, New York.
- [5] Davis, R.H., Edelman, D.B. and Gamberman, A.J. (1992) Machine-Learning Algorithms for Credit-Card Applications. *IMA Journal of Management Mathematics*, **4**, 43-51. <https://doi.org/10.1093/imaman/4.1.43>
- [6] 张承钊, 何浩龙, 许建龙. 基于异构集成模型的个人信用评估研究[J]. 计算机仿真, 2022, 39(3): 485-489.
- [7] Dastile, X. and Celik, T. (2021) Making Deep Learning-Based Predictions for Credit Scoring Explainable. *IEEE Access*, **9**, 50426-50440. <https://doi.org/10.1109/ACCESS.2021.3068854>
- [8] 罗方科, 陈晓红. 基于 Logistic 回归模型的个人小额贷款信用风险评估及应用[J]. 财经理论与实践, 2017, 38(1): 30-35.
- [9] Breiman, L., Friedman, J., Olshen, R., et al. (1984) Classification and Regression Trees (CART). *Biometrics*, **40**, 358-361. <https://doi.org/10.2307/2530946>