

基于可解释集成学习的信贷违约预测^①



蔡青松, 吴金迪, 白宸宇

(北京工商大学 计算机学院, 北京 100048)
通讯作者: 蔡青松, E-mail: caiqs@btbu.edu.cn

摘要: 人工智能促进了风控行业的发展, 智能风控的核心在于风险控制, 信贷违约预测模型是解决这一问题必须倚靠的手段. 传统的解决方案是基于人工和广义线性模型建立的, 然而现在通过网络完成的交易数据, 具有高维性和多重来源等特点, 远远超出了现有模型的处理能力, 对于传统风控提出了巨大的挑战. 因此, 本文提出一种基于融合方法的可解释信贷违约预测模型, 首先选取 LightGBM、DeepFM 和 CatBoost 作为基模型, CatBoost 作为次模型, 通过模型融合提升预测结果的准确性, 然后引入基于局部的、与模型无关的可解释性方法 LIME, 解释融合模型的预测结果. 基于真实数据集的实验结果显示, 该模型在信贷违约预测任务上具有较好的精确性和可解释性.

关键词: 金融风险; 违约预测; 信贷风险; 可解释性; 集成模型

引用格式: 蔡青松, 吴金迪, 白宸宇. 基于可解释集成学习的信贷违约预测. 计算机系统应用, 2021, 30(12): 194-201. <http://www.c-s-a.org.cn/1003-3254/8220.html>

Prediction of Credit Default Based on Interpretable Integration Learning

CAI Qing-Song, WU Jin-Di, BAI Chen-Yu

(School of Computer Science and Engineering, Beijing Technology and Business University, Beijing 100048, China)

Abstract: Artificial intelligence accelerates the development of the risk control industry. Undoubtedly, risk control is the core of intelligent risk control, and a credit default prediction model is its essential means. The traditional access to risk control is based on artificial and generalized linear models. However, the data of transactions completed on the Internet are characterized by high dimensions and multiple sources, which cannot be processed by existing models. This poses a great challenge to traditional risk control. In view of this, this study proposes an interpretable credit default model based on the fusion method. To be specific, the accuracy of the prediction results is first enhanced through the fusion of base models (LightGBM, DeepFM, and CatBoost) and secondary model (CatBoost). Then, the prediction result of the fusion model is interpreted by the introduced local-based interpretability method LIME that is independent of the model. According to the experimental result of a real dataset, the satisfactory accuracy and interpretability of the model can be witnessed on the task of credit default prediction.

Key words: financial risk control; prediction of default; credit risk; interpretability; integration model

近年来, 关于信用贷款的违约预测已成为高校科研人员和金融机构重点研究的内容^[1]. 信用贷款是根据借款人的综合信用发放的一笔借款资金, 其不需要固

定资产抵押, 但是对借款人征信有着严格要求. 网贷公司、互联网公司等机构的信贷业务, 也伴随着互联网的高速发展而飞速前进, 与此同时, 这些贷款机构面临

① 基金项目: 国家自然科学基金 (61702020); 北京市自然科学基金-海淀原始创新联合基金 (L182007)

Foundation item: National Natural Science Foundation of China (61702020); Joint Fund of Natural Science Foundation of Beijing Municipality and Haidian Original Innovation Fund (L182007)

收稿时间: 2021-03-02; 修改时间: 2021-03-29; 采用时间: 2021-04-06

着部分客户无法按时还款的问题, 违约现象愈发显著^[2]. 对于贷款申请是否通过, 以前是凭借人工处理与机器审核相结合的方式来做决策, 在大数据时代, 这种方式已经不再适合市场发展的需要. 要解决此类问题, 必须构建一个可靠的预测模型^[3]. 信贷违约预测模型是金融大数据风控领域必不可少的一部分, 能够帮助审批人员准确高效地识别客户是否具有偿还贷款的能力, 在一定程度上能够减轻贷款审批人员的工作压力, 同时提高审批通过的准确率. 因此, 对信贷违约预测模型进行研究具有很大的现实意义^[4].

近年来, 国内外关于信贷违约预测的研究都取得了一定的成果. 章宁等人针对借贷过程中的信息不对称问题, 为更有效地整合不同的数据源和贷款违约预测模型, 提出一种基于 AUC 和 Q 统计值的学习方法^[5]. Deng 等人使用 LendingClub 平台提供的数据进行研究, 选出影响最大的前 20 个因素, 使用 Logistic 回归模型, 确定了影响违约风险的主要因素, 并通过可视化一些变量之间的关系, 得到定性分析结果^[6]. Kim 等人提出了一种半监督学习方法, 通过考虑社会借贷数据的特点, 利用社会贷款的大量未标记数据来强化边界, 利用标签传播、转换支持向量机 (TSVM) 与 Dempster-Shafer 理论相结合的方法, 实现对社交借贷的准确预测^[7]. 魏力等人针对 P2P 信贷项目如何减少借贷过程中的风险冲击, 建立由规则推导而成的标签模型, 从多角度描述用户, 构建客户画像, 提升分类任务的效果^[8].

上述信贷违约预测模型主要是以简单模型为主. 采用这种模型虽然保证了一定的可解释性, 但也出现了预测精确率低的情况. 因此, 为了避免使用简单模型存在的问题, 一些研究人员将目光投向了使用多模型的方式来构建信贷违约预测模型. Tong 等人针对英国的贷款违约问题, 利用来自英国银行的大量房屋抵押贷款违约数据集, 针对拖欠贷款产生的总损失金额建立了混合离散连续模型, 从而提高预测方面的性能^[9]. 马晓君等人运用 CatBoost 算法对于 P2P 网络借贷中的违约预测情况展开了研究, 通过对违约影响因素进行综合分析, 使得模型出错所导致的损失成本进行有效控制^[10]. Ma 等人又使用了多角度和多维的数据清洗方法, 将 XGBoost 算法应用于真实数据, 其使用基于分布式的模式结合弱学习器, 提高了模型的训练效率^[11].

受上述研究启发, 本文基于 Stacking^[12] 策略, 构建了基于 LightGBM-DeepFM-Catboost 的信贷违约预测

模型^[13-15]. 此外, 本文新颖性地引入了基于局部可解释的、与模型无关的 LIME 方法^[16], 能够对于复杂模型的预测结果做出值得人们信赖的解释, 在提升违约预测精确度的同时, 进一步提升本文所提出模型的可解释性.

1 相关技术研究

1.1 LightGBM 算法

LightGBM 是对 GBDT 算法进行改进和提升后的优化版本, 属于集成模型, 主要采用了基于 leaf-wise 的决策树生长策略和基于直方图的决策树算法. 基于 leaf-wise 的决策树生长策略每次从当前所有叶子中找到分裂增益最大的一个叶子, 然后分裂, 如此循环, 因此同 level-wise 相比, 在分裂次数相同的情况下, leaf-wise 可以降低更多误差, 得到更好的精度, 在保证高效率的同时防止过拟合. 同时基于直方图的决策树算法不需要额外存储预排序的结果, 而且可以只保存特征离散化后的值, 明显降低了内存消耗, 从而对于内存的利用率会更低, 在二叉树中可以通过利用叶节点的父结点和相邻结点的直方图的相减来获得该叶节点的直方图, 所以只要为一个叶节点建立直方图, 就可以通过直方图的相减来获得相邻结点的直方图, 因此花费的代价较小, 有助于提高融合模型的效率.

1.2 DeepFM 算法

DeepFM 是一个集成了 DeepNN 和 FM 的神经网络框架, 由 DeepNN 和 FM 通过并行的方式组合而成. DeepNN 是用来做特征之间的高阶组合, 所有的特征都会被转化成 Embedding 向量作为 Deep 部分的输入, 在 DNN 层利用深度学习数据在网络间的传播性得到高阶特征. FM 是用于特征之间的低阶组合, 在 Dense Embedding Layer 的基础上, 做了二项交叉特征, 能够自动识别特征组合, 来减少手工提取特征的工作量, 且计算过程中不需要用户干预. 这两个部分分配同样的特征输入, 在经过模型后, 将他们的输出汇总成为最终的预测结果. 既考虑低阶又考虑高阶特征的方式在违约预测模型领域具有较好的覆盖性.

1.3 CatBoost 算法

CatBoost 能够很好地处理类别型特征, 对于每个样本, 都单独构建一个利用该样本之前的样本点的梯度估计得到的模型. 针对这些模型, 估计该样本的梯度,

然后利用新模型重新对样本打分. 与原有的梯度提升方法不同, 它具有有序提升的特点, 通过在每个训练步骤中, 使用独立置换的历史样本进行无偏差的提升, 相比其他提升算法能够更好地响应过拟合问题. CatBoost 还可以提供类似于基于决策树的特征重要性, 是在训练过程中处理类别型特征, 而无需进行包括调整超参数在内的额外预处理. 因为其简单性和高性能的特性, 所以该算法已被越来越多的应用在各领域中.

1.4 Stacking 策略

集成方法可分为两类, 分别是同质集成和异质集成, 同质集成是使用同一种机器学习算法来构建集成模型, 异质集成是使用不同的机器学习算法来构建集成模型. 传统上构建集成模型时, 每个模型都具有相同的权重, 有时可能认为某些模型更好或更准确, 希望手动为其分配更高的权重. 有一种方法是通过使用另一层学习算法来自动地估计每个模型的权重, 这种方法称为 Stacking 策略, 具体流程如图 1 所示.

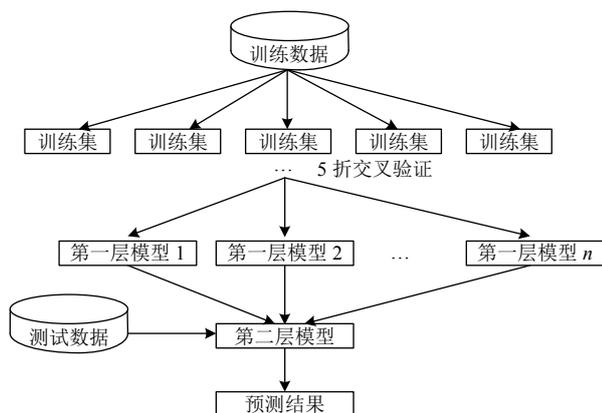


图 1 Stacking 流程图

由于每个基模型都做出了重大贡献, 而且每个模型各有优劣, 使用了 Stacking 策略后, 通过集成的方式使得其他模型的优势弥补了某个模型的弱点和不足, 提高了整个预测模型的准确性. 将 Stacking 策略应用于现实世界中的大数据问题, 也可以产生更高的预测准确性. 其主要思想是将原始数据集划分为训练集和测试集, 采用交叉验证的方式对第一层的各种模型进行训练, 使训练后的模型在验证集和测试集上进行预测, 将验证集上的预测结果叠加作为新的特征, 用作第二层学习模型的输入, 将测试集上的预测结果的平均作为第二层模型的测试集. 该第二层算法经过训练, 从

而做出最终预测结果. 算法流程如算法 1 所示.

算法 1. Stacking 算法

输入: 训练集 $D(x_1, y_1, \dots, x_n, y_n)$; 测试数据 x'
输出: 预测结果 y'

```

1. for  $i=1 \rightarrow m$  do
2.    $h_i \leftarrow c_i(D)$ 
3. end for
4.  $D' \leftarrow \emptyset$ 
5. for  $j=1 \rightarrow n$  do
6.   for  $i=1 \rightarrow m$  do
7.      $z_{ji} \leftarrow h_i(x_j)$ 
8.   end for
9.    $D' \leftarrow D' \cup ((z_{j1}, \dots, z_{jm}), y_j)$ 
10. end for
11.  $h' \leftarrow c'(D')$ 
12.  $y' \leftarrow h'(h_1(x'), \dots, h_m(x'))$ 

```

1.5 模型解释机制

针对模型的可解释性分为全局解释和局部解释, 预测结果可解释性对于本文研究问题的意义有以下 4 点: 首先针对于信贷领域, 可解释性能够将不可信任的预测结果转换为值得信任的, 让用户对于预测结果产生充分信任. 其次提供了观察信贷违约预测模型的新视角, 以便应对来自银保监会和人行的监管压力, 同时保证了信贷领域的公平性. 再次可以帮助金融机构决定是否部署融合模型和进一步改善模型. 最后可以消除风控模型在金融机构部署时的潜在威胁, 建立起机构、用户与信贷违约预测模型之间的信任关系.

全局可解释性是指能够基于全部数据集上的响应变量和输入特征之间的相互作用来解释模型的决策. 对于本文所研究的信贷违约预测模型, 全局可解释性可以在模块级别上帮助我们理解模型, 比如说针对不同特征, 目标变量的分布是什么, 这对于本文的研究具有一定的意义, 但是对于参数级层面的解释, 在信贷违约预测模型具有大量参数的情况下, 人们很难想象特征之间是如何通过相互作用, 才能获得这样的预测结果.

LIME 是一种利用简单模型来解释任何复杂模型的局部解释技术, 与被解释的模型无关, 不需要进行模型适配, 也不会深入模型内部, 它用人们理解的方式来解释模型. LIME 解释技术对于数据集中特征相关与否并无明确要求, 只是复杂模型的预测精度会随着不相关特征的剔除而不断上升, 其主要是通过扰动局部的人类容易理解的特征来改变输入, 通过对复杂模型的局部近似, 观察预测结果的变化来理解模型内部的行

为,生成人们可以信赖的解释.对于模型预测结果进行事后解释,我们能拥有更多的信息来决定是否信任信贷违约预测模型以及模型的预测结果.

我们将 g 作为预测模型 f 的解释: $g \in G$, 定义一个目标函数 ξ , 同时,为了测量 g 在局部如何逼近 f , 我们定义了一个损失函数 L , 在保持模型复杂度 $\Omega(g)$ 足够低的同时,我们通过最小化损失函数 L 得到目标函数 ξ 的最优解.

目标函数定义为:

$$\xi(x) = \arg \min_{g \in G} L(f, g, \pi_x) + \Omega(g) \quad (1)$$

损失函数定义为:

$$L(f, g, \pi_x) = \sum_{z, z' \in Z} \pi_x(z) (f(z) - g(z'))^2 \quad (2)$$

其中, z 是原始数据集中一个被扰动的样本点, z' 是通过对该样本随机扰动产生的新样本点的集合, 通过 LIME 的可视化机制, 给出模型预测结果的局部解释.

2 基于 Stacking 与 LIME 信贷违约预测模型

信贷违约预测模型通过在脱敏数据集上完成训练, 来预测客户后期产生违约的概率, 目的是帮助机构人员对客户的资质完成审核, 预测客户的还款能力, 提前发现贷款的潜在损失, 避免产生不必要的纠纷, 降低借贷风险. 本文采用 Stacking 策略, 使用单个模型来学习如何最佳地组合互补模型, 并将第一层模型的输出作为标签附加到我们的样本中, 还需要使用交叉验证来比较模型的泛化性能.

由于不同的算法模型具有不同的性能特征, 所以在综合比较了常用的模型后, 我们选择了 LightGBM、DeepFM、CatBoost 作为第一层模型, 将第一层基模型的输出作为第二层模型的输入. 由于 CatBoost 模型在 Stacking 中的应用对于模型融合更为有效, 所以我们使用 CatBoost 模型作为第二层的次模型. CatBoost 模型学习第一层模型中子模型的性能, 并根据子模型的误差给予模型赋予不同的权重后进行训练, 以进行最终预测. 因此, 如果单个模型表现不佳, 不一定会对结果产生不利影响. 两层模型融合后的新模型优于第一层的单个子模型, 这也表明多模型融合比单一模型具有更好的预测准确度.

基于 Stacking 与 LIME 信贷违约预测模型, 其预测的流程如图 2 所示.

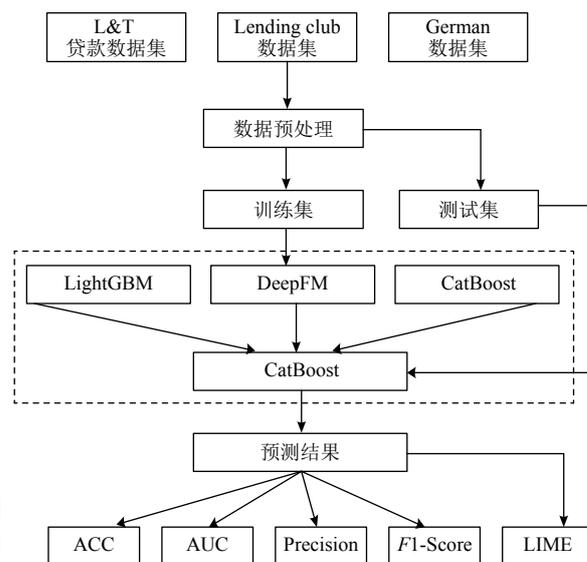


图 2 信贷违约预测模型

基于 Stacking 策略的融合模型, 通过引入 LIME 方法, 构造了具备可解释性的异质集成模型. LIME 为了搞清楚哪一部分输入对预测结果产生贡献, 在要解释的样本周围做微小的扰动, 然后我们根据这些扰动的数据点距离原始数据的距离分配权重, 基于学习得到一个可解释的稀疏线性模型, 对于融合模型最终的预测结果, 事后解释机制 LIME 通过可视化的方式在局部进行了解释.

3 仿真实验

3.1 实验环境

在实验过程中, 我们使用开源语言 Python 3 编写代码, 利用 Python 的 NumPy 和 Pandas 等进行数据预处理和特征工程. 操作系统为 Windows 10, 机器配置为 2.40 GHz Intel i5-9300H GTX1650.

3.2 数据描述

本文使用的数据是来自 Kaggle 的公开数据集, 包含了贷款客户脱敏后的信用交易数据, 并带有一个标签, 指示客户是否会违约. 3 个数据集中分别存储了不同用户的基本信息数据, 每一行代表一个样本, 来自 3 个数据集的主要特征的含义均展示在如表 1 所示的主要特征含义表中.

每笔借贷样本属性主要包括年龄、借款利率、借款用途、借款目的、借贷持续时间、信用账户额度、借款时长、活期存款、借款人现在拖欠账目的逾期情况、过去 12 个月的征信查询次数、过去 6 个月的贷

款违约次数、房屋状况、教育程度、性别、工作等信息。

表1 主要特征含义表

特征表示	含义
Checking_account	活期存款
Credit_amount	信用账户额度
Duration	借贷持续时间
Saving_accounts	储蓄账户
Age	年龄
Housing	房屋状况
Purpose	目的
Age_Group	年龄组
Sex	性别
Job	工作
collection_12_mths_ex_md	12个月内购买的收藏品数目
int_rate	借款利率
purpose	借款目的
Deline_amnt	借款人拖欠账目的逾期情况
Chargeoff_wihtin_12_mths	12个月内贷款冲销总数
Mo_sin_old_rev_tl_op	首个账户开户时长
Inq_last_12m	12个月内征信查询次数
lt_v	贷款资产的价值
NO. OF INQUIROES	调查人数
ASSET_COST	资产成本
TOTAL_OVERDUE_ACCOUNTS	逾期账户数量
CREDIT.HISTORY.LENGTH	自首贷至今的时间
DELINQUENT.ACCTS.IN. LAST. SIX. MONTHS	过去6个月违约次数
AVERAGE. ACCT. AGE	平均贷款期限

3.3 实验流程

(1) 数据预处理

由于原始数据集具有缺失值、类别不平衡和高维特征等问题。在预测模型中，类别不平衡会导致预测不准确，而高维稀疏数据也会产生维数问题，为了解决上述问题，我们进行了以下工作。首先关注类别不平衡问题，经过分析表明，由于模型中正负样本比例过大，从而导致过拟合，以至于训练后的模型预测效果差。导致训练后的模型精度虽然高，但是却没有价值，通过下采样对这一问题进行了控制。

对于应用了级联原理模型的正负样本阈值的确定，我们将数据集分出一部分作为验证集，对于级联中选取的每个基模型进行训练后去不断的动态调整阈值，通过在实验中调整参数，找到在验证集上效果最好的阈值。部分样本的缺失值过多，会影响模型训练效果。但是缺失数据的比例占总体样本的比例较少，而且缺

失数据是随机出现的，所以我们从训练数据中直接删除了这部分缺失值过高的数据，这样删除缺失数据，对后续的分析结果影响不大。数据集中还有一些类别特征样本数过低的以及为0的特征，由于它们对最终的预测结果影响不大，因此，我们将其进行直接删除。

(2) 特征工程

原始数据集中有近百维的数值特征变量，拥有过多的维数容易导致过拟合问题，所以我们需要对数据进行降维。除了平常使用的PCA和t-SNE等方法外，我们还可以选择克拉默相关系数、最大信息系数、皮尔逊相关系数等方法。我们选择使用皮尔逊相关系数来进行连续特征的筛选。皮尔逊相关系数是一种线性相关系数，是最常用的一种相关系数，记为*r*，用来反映两个变量*X*和*Y*的线性相关程度，值介于-1到1之间。相关系数越接近于1或-1，相关性越强，相关系数越接近于0，相关度越弱。通过计算特征间的皮尔森系数进行特征降维，两个特征间皮尔逊相关系数高于0.95已经是属于极强相关，可以随机选择删去一个。基于上述的方法，可以用减少特征数量的方式实现对数据的降维。

3.4 评价指标

为了验证该模型的性能，选择*ACC*、*F1-Score*、*Precision*和*AUC*作为评价指标，其中*AUC*表示预测的正样本排在负样本前面的概率，相较于*ROC*能够更好的评价模型性能。混淆矩阵如表2所示。

$$ACC = \frac{Tp + Tn}{Tp + Tn + Fp + Fn} \tag{3}$$

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

$$Precision = \frac{Tp}{Tp + Fp} \tag{5}$$

表2 混淆矩阵

类别	正例	反例
正例	<i>Tp</i>	<i>Fn</i>
反例	<i>Fp</i>	<i>Tn</i>

3.5 结果分析

将我们所提出的融合模型与其他3个模型进行比较，通过在3个数据集上进行的实验可以看出，本文所提出的模型较*LightGBM*、*DeepFM*、*CatBoost*来说，在*AUC*、*ACC*、*Precision*和*F1-Score*这些指标上面均有明显提升，说明本文所提出的模型有效的提升了

信贷违约预测准确度.

从表3中我们可以得到如下结论:我们提出的模型,在4项指标上对比LightGBM、DeepFM和CatBoost模型均有提高,其中AUC最高提升了12.47%,F1-Score最高提升了2.8%,ACC最高提升了9%,Precision最高提升了9.07%.通过对表4和表5中实验指标的观察分析,可以看出我们提出的模型也优于其他模型,证实了我们所提出的模型有效提升了预测精确度.

表3 German数据集实验结果对照表

序号	指标	LightGBM	DeepFM	CatBoost	本文
0	AUC	0.7395	0.6819	0.7561	0.8066
1	ACC	0.7400	0.7300	0.7700	0.8200
2	Precision	0.7820	0.7263	0.7701	0.8170
3	F1-Score	0.8243	0.8363	0.8535	0.8815

表4 Lending club数据集实验结果对照表

序号	指标	LightGBM	DeepFM	CatBoost	本文
0	AUC	0.7207	0.6765	0.7042	0.7310
1	ACC	0.8356	0.8328	0.8356	0.8384
2	Precision	0.8362	0.8339	0.8356	0.8425
3	F1-Score	0.9099	0.9085	0.9100	0.9107

表5 L&T数据集实验结果对照表

序号	指标	LightGBM	DeepFM	CatBoost	本文
0	AUC	0.6414	0.6387	0.6434	0.6486
1	ACC	0.7810	0.7810	0.7811	0.7813
2	Precision	0.5714	0.5454	0.5060	0.5454
3	F1-Score	0.6431	0.6488	0.6370	0.6521

3.6 预测结果解释

我们用信贷违约预测模型的目的是判断一个客户

在未来产生违约的可能性.通过前边的实验分析,可以说该模型适用于信贷机构的需求,但是想要理解一些客户为什么会被预测为违约,而其他客户不会被预测为违约,同时风控部门也需要评估模型所作出的预测是否值得相信.因此,本文认为能解释每个模型的预测是获取信任的一种重要方式.

为了实现这一目标,通过LIME方法来研究预测结果中不同特征的贡献,可以证明我们的模型是值得信任的.这将极大地改善风控部门原有复杂模型解释性不足的问题.

对于本文所提出的融合模型在公开数据集上的预测结果(图3-图5),利用事后解释机制LIME进行可视化解释,输出了模型对这3个数据集中单个样本的解释结果,分别显示影响样本预测结果重要性前10的特征,及其相应贡献度和特征对应的值.第1列表示从预测模型得到的负面结果和正面结果的预测概率,第2列显示某个特征对正面或负面结果的贡献,为了清楚起见,只显示了前10个特征.第3列显示特征的原始数据值或将离散变量经Label编码后的值.

接下来,如图3所示,可以看出模型对该客户有72%的概率相信其不会违约,那么模型预测该客户大概率不会违约的理由就是基于该客户信用卡额度、借贷持续时间、储蓄账户、年龄这些因素,另外活期存款、房屋状况、目的,成为了预测该用户不会违约的一个干扰.这种解释结果在现实世界中是有意义的,并揭示了我们可以在一定程度上信任我们的预测模型.

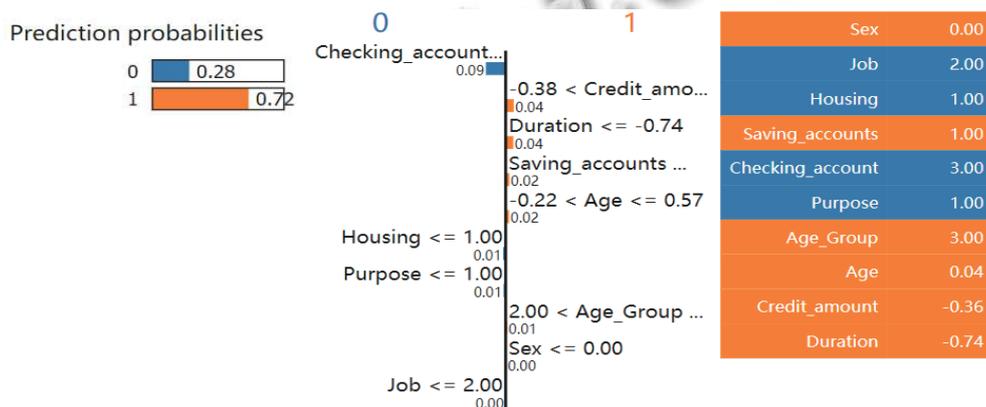


图3 German数据集样本预测结果解释

在图4中,对于该客户模型有74%的概率相信他不会产生违约行为,根据解释结果可以得出,该客户在

12个月中除了医疗收藏品以外的收藏品数目、借款利率、借款目的、12个月内的贷款冲销总数这些方面信

用是十分好的,但是该客户现在拖欠账目的逾期情况、过去12个月的征信查询次数,会对系统的解释结

果产生负面影响.用户通过可解释性结果了解自己接下来需要在哪方面进行改进.



图4 Lending club 数据集样本预测结果解释

从图5看出,我们的融合模型判断该客户违约的概率是72%,可解释机制给我们的解释就是基于该客户在过去6个月的贷款违约次数、贷款资产的价值这

些方面信用不好,导致了该预测结果,另外该客户在资产成本、年龄这些方面的表现还是不错的,接下来可以继续加强,以提升自己的信用度.

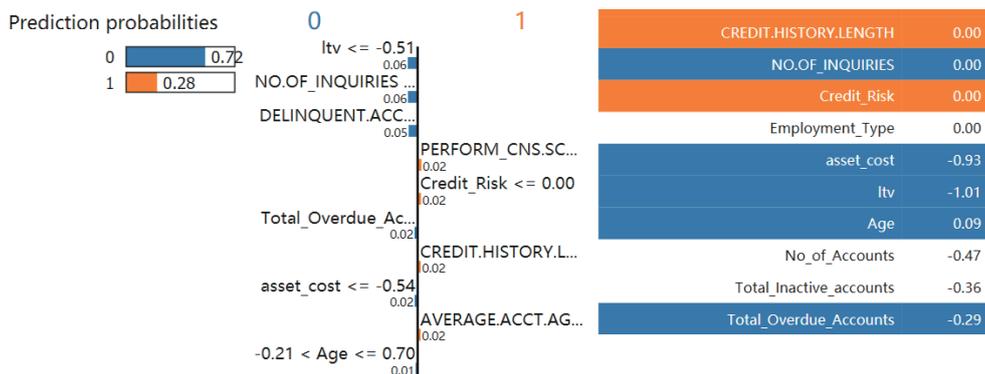


图5 L&T 数据集样本预测结果解释

4 结论与展望

本文为了获得具备可解释性的信贷违约预测模型,通过不同模型间的异质集成,以融合的方式构建了一个由 LightGBM、DeepFM 和 CatBoost 算法组成的异质融合模型,用于信贷审批时预测申请人的违约风险.在基于 Stacking 策略的融合模型中,采用 LightGBM、DeepFM 和 CatBoost 三个不同的模型作为第一层的预测模型,然后我们使用 CatBoost 作为第二层模型,在预测精确度、泛化性能等方面都有较好的表现.实现了信贷违约预测模型预测精确度的提高.此外,通过引入事后解释机制 LIME,也显著的提升了模型的可解释性.通过在3个公开数据集上的实验表明,本文建立的

融合模型可以在准确预测信贷违约的同时,又兼具可解释性.对于具备解释性的黑盒信贷违约模型的研究有一定的现实意义.此外,本文采用融合模型在提升准确度的同时,还存在的时间效率较低的问题,在后续的研究中,我们希望能够改进特征选择方法,进一步降低原始特征中存在的冗余和噪声,将有意义的特征输入机器学习模型中进行训练,同时通过网络计算实现资源共享,使用多线程模式结合基模型,合理利用硬件资源,以提高基于融合方法模型的时间效率.

参考文献

- 程大伟,牛志彬,张丽清.大规模不均衡担保网络贷款的风

- 险研究. 计算机学报, 2020, 43(4): 668–682. [doi: 10.11897/SP.J.1016.2020.00668]
- 2 庞素琳. 违约风险下的信贷决策模型与机制. 管理科学学报, 2012, 15(4): 58–70. [doi: 10.3969/j.issn.1007-9807.2012.04.008]
- 3 韦璠, 宋云飞, 邵明莉, 等. 利用特征融合和整体多样性提升单模型鲁棒性. 软件学报, 2020, 31(9): 2756–2769. [doi: 10.13328/j.cnki.jos.005943]
- 4 Chen JD, Tao Y, Wang HR, *et al.* Big data based fraud risk management at Alibaba. *The Journal of Finance and Data Science*, 2015, 1(1): 1–10. [doi: 10.1016/j.jfds.2015.03.001]
- 5 章宁, 陈钦. 基于 AUC 及 Q 统计值的集成学习训练方法. 计算机应用, 2019, 39(4): 935–939.
- 6 Deng TN. Study of the prediction of micro-loan default based on logit model. 2019 International Conference on Economic Management and Model Engineering (ICEMME). Malacca: IEEE, 2019. 260–264.
- 7 Kim A, Cho SB. An ensemble semi-supervised learning method for predicting defaults in social lending. *Engineering Applications of Artificial Intelligence*, 2019, 81: 193–199. [doi: 10.1016/j.engappai.2019.02.014]
- 8 魏力, 王子炫. 结合标签规则的 P2P 网贷风控模型. 计算机与数字工程, 2020, 48(7): 1687–1692.
- 9 Tong ENC, Mues C, Thomas L. A zero-adjusted gamma model for mortgage loan loss given default. *International Journal of Forecasting*, 2013, 29(4): 548–562. [doi: 10.1016/j.ijforecast.2013.03.003]
- 10 马晓君, 宋嫣琦, 常百舒, 等. 基于 CatBoost 算法的 P2P 违约预测模型应用研究. 统计与信息论坛, 2020, 35(7): 9–17. [doi: 10.3969/j.issn.1007-3116.2020.07.002]
- 11 Ma XJ, Sha JL, Wang DH, *et al.* Study on a prediction of P2P network loan default based on the machine learning LightGBM and XGboost algorithms according to different high dimensional data cleaning. *Electronic Commerce Research and Applications*, 2018, 31: 24–39. [doi: 10.1016/j.elerap.2018.08.002]
- 12 盛杰, 刘岳, 尹成语. 基于多特征和 Stacking 算法的 Android 恶意软件检测方法. 计算机系统应用, 2018, 27(2): 197–201. [doi: 10.3969/j.issn.1003-3254.2018.02.033]
- 13 徐磊, 孙朝云, 李伟, 等. 基于 SSA-LightGBM 的交通流量调查数据趋势预测. 计算机系统应用, 2021, 30(1): 243–249. [doi: 10.15888/j.cnki.csa.007750]
- 14 王美, 龙华, 邵玉斌, 等. 基于 FM 与 DeepFM 模型对 GTD 特征的研究. 通信技术, 2019, 52(6): 1495–1499. [doi: 10.3969/j.issn.1002-0802.2019.06.033]
- 15 党存禄, 武文成, 李超锋, 等. 基于 CatBoost 算法的电力短期负荷预测研究. 电气工程学报, 2020, 15(1): 76–82.
- 16 Ribeiro MT, Singh S, Guestrin C. “Why should I trust you?”: Explaining the predictions of any classifier. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco: ACM, 2016. 1135–1144.