

数字金融时代 机器学习模型在实时反欺诈中的应用与实践

曹汉平, 张晓晶, 祝睿杰, 黄潇拉

(中国银行总行个人数字金融部, 北京 100818)

摘要: 近年来, 数字金融蓬勃发展, 金融科技日趋成熟, 信息技术的发展对社会产生巨大积极作用的同时也带来了新型风险, 网络黑产呈爆发式增长, 电信网络诈骗给人民群众造成了巨大的财产损失。在数字金融时代, 商业银行既迎来了新的机遇与动力, 又面临着新的挑战 and 数字化变革的要求, 线上金融业务已经成为新的主战场。在此背景下, 基于 RFM 高维衍生特征和对机器学习算法的研究, 构建了基于高维交易行为画像的事中反欺诈机器学习模型。依托大数据、流计算等技术, 通过在系统化部署、应用策略以及模型迭代优化等环节的实践, 形成了一套基于机器学习模型的事中风控方案。实践证实, 该模型的 AUC 达到了 0.972, 可以敏锐洞察欺诈风险, 实现毫秒级的欺诈交易识别, 对于提升线上数字金融业务的事中风控能力具有一定的推广价值和借鉴意义。

关键词: 数字金融; 机器学习; 反欺诈; RFM; 流计算

中图分类号: TP391

文献标识码: A

doi: 10.11959/j.issn.2096-6652.201939

Application and practice of machine learning model in real-time anti-fraud in the era of digital finance

CAO Hanping, ZHANG Xiaojing, ZHU Ruijie, HUANG Xiaola

Digital Personal Banking Department, Bank of China, Beijing 100818, China

Abstract: In recent years, with the rapid development of FinTech, digital finance has flourished and brought huge positive effect on society. Meanwhile, new risks have been introduced into banks. For example, the black production related to network security has experienced explosive growth, and telecommunication network fraud has caused property losses to the public. In the era of digital finance, the commercial banks have not only ushered in new opportunities and dynamics, but also faced new challenges and requirements for digital transformation. As a result, e-finance has become a new battlefield. With this context, a real-time anti-fraud machine learning model based on high-dimensional transaction behavior portrait through enhanced RFM feature-derivation and machine learning modeling was established in this paper. Relying on the new technologies such as big data, stream computing, a model application solution to real-time risk control was formed including systematic deployment, model application strategies and iterative model optimization. Through practical observation, the AUC of the model reaches 0.972, which provides a keen insight into fraud risk, realizes millisecond-level risk identification, and promotes risk control ability of e-finance significantly.

Key words: digital finance, machine learning, anti-fraud, RFM, stream computing

收稿日期: 2019-10-10; 修回日期: 2019-11-12

通信作者: 祝睿杰, wljrzhuruijie_hq@mail.notes.bank-of-china.com

1 引言

1.1 背景及意义

在新一轮科技革命和产业变革的背景下，银行业数字化浪潮蓬勃兴起，大数据、人工智能、云计算等新技术与金融业务深度融合，成为推动金融转型升级的新引擎、服务实体经济的新途径、防范化解金融风险的新利器^[1]，数字化转型已成为商业银行提高服务质量和竞争力的共同选择。

随着商业银行运用新科技开展创新转型的加速，网上银行、手机银行、移动支付等线上数字金融业务在给银行带来更大创新空间，为客户带来更加高效、优质的金融服务的同时，也给反欺诈带来了严峻挑战。在巨大经济利益的驱动下，不法分子利用钓鱼链接、木马、电信诈骗等各种手段盗取、骗取客户资金的案件层出不穷。针对线上转账、支付等环节的欺诈，已经发展成为组织严密、分工明确的黑色产业链条，给客户和银行都造成了严重的损失。

保护客户资金安全是银行的目标和使命。如何在互联网中向客户提供更加安全的服务，让客户免受诈骗、盗刷等互联网黑产的侵害，对银行的反欺诈能力提出了更高的要求，也成为数字金融时代银行必须打造的核心能力。

1.2 交易反欺诈的难点

线上数字金融业务的特点加大了银行在交易环节进行用户身份真实性识别的难度。经过多年的积累，各大银行根据线上业务的开展情况逐步建立了配套的反欺诈体系，目前主要以风控业务专家设计的规则模型为核心，有效保护了客户资金安全，并有力打击了犯罪分子。

在互联网、移动互联网技术的助力下，犯罪分子的欺诈手法也在不断升级和变化，银行与犯罪分子间的对抗持续升级。相对静态的专家规则模型在欺诈案件的准确识别和新型案件的适应能力上均存在一定的局限性，主要表现在3个方面：一是专家规则是针对欺诈案件进行总结和归纳形成的，而非基于欺诈交易与正常交易的行为差异，经验的不足导致难以避免“管窥蠡测”的局限；二是专家规则是刚性的，一套标准适用所有客户，无法“因人施策”地进行欺诈交易识别，在漏掉欺诈交易的同时会误将正常客户的正常交易判定为欺诈交易，影响客户体验；三是在与网络黑产的快速对抗中，专

家规则对新型案件的防控存在一定滞后性，需要风控专家不断总结案件规律，才能对规则模型进行调整和优化。

此外，线上数字金融业务会产生高并发、海量、多维度的数据，如何有效使用线上业务产生的数据，如何利用新型技术架构解决时效性难题，如何基于大数据方法提升线上交易的欺诈防范能力，都是交易反欺诈领域迫切需要解决的问题。

1.3 基于大数据、人工智能的风控新方向

近年来，以大数据、人工智能等为代表的先进技术为商业银行提高风控能力提供了新的方向。以大数据的“流”技术为依托，利用 Spark 分布式集群计算引擎，可以快速、高效地处理实时交易环节产生的海量数据^[2]；基于大数据平台的历史数据存储和计算能力，能够灵活加工长期、多维度的历史交易行为特征，为利用海量数据开展实时的事中风控工作奠定了坚实的技术基础。

在金融科技浪潮的推动下，部分商业银行积极探索大数据、流计算、人工智能等新技术的应用，采用先进的分布式大数据系统架构和流处理引擎，建立了新型的智能反欺诈风控系统。该系统具备高并发、低时延、高精度、高可靠等特点，可进行毫秒级的风险判定，并能够支撑实时的机器学习模型和智能风险决策。

依托智能反欺诈系统的流处理能力、大数据运算能力、人工智能支持能力和实时决策能力，本文开展了线上场景交易维度的反欺诈机器学习算法的模型研究与实践应用工作。利用机器学习等人工智能算法对欺诈交易进行建模，并采用高维特征对每笔交易进行画像，通过对每笔交易的细致刻画和历史交易行为的全面“了解”，敏锐感知不符合历史画像的交易。向每一个客户提供柔性的安全服务，在更加准确、全面地识别欺诈交易的情况下，及时发现新型欺诈行为，最大限度地防范欺诈行为的发生，有效提升线上数字金融业务的风险防控能力。

2 基于交易行为画像的建模方案

2.1 算法选择

通过大量调研与实践验证，选择了改良版梯度提升（extreme gradient boosting, XGBoost）决策树机器学习算法建模。该算法以树模型为基分类器，在分类效果、业务解释性、建模效率等方面获得了

业界广泛的认可，是当前应用广泛的机器学习模型之一。

XGBoost 是在传统梯度提升决策树 (GBDT) 模型的基础上改良而来的，主要对目标函数和缺失值样本的处理方式进行了优化。

XGBoost 目标函数为^[3-4]：

$$\lambda^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_i(x_i)) + \Omega(f_t) \quad (1)$$

通过二阶泰勒展开后改写目标函数为：

$$\lambda^{(t)} \cong \sum_{i=1}^n [l(y_i, \hat{y}_i^{(t-1)}) + g_i f_i(x_i) + \frac{1}{2} h_i f_i^2(x_i)] + \Omega(f_t) \quad (2)$$

代入传统 GBDT 训练过程，得到 XGBoost 目标函数：

$$\begin{aligned} \text{Obj}^{(t)} = & \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)}) + g_i f_i(x_i) + \\ & \frac{1}{2} h_i f_i^2(x_i) + \Omega(f_t) + \text{constant} \end{aligned} \quad (3)$$

通过泰勒展开，得出一阶项（通常称为残差）和二阶项：

$$g_i = \partial_{\hat{y}_i^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)}) \quad (4)$$

$$h_i = \partial_{\hat{y}_i^{(t-1)}}^2 l(y_i, \hat{y}_i^{(t-1)}) \quad (5)$$

从目标函数的角度来看，XGBoost 对目标函数的损失项进行二阶泰勒展开，保留了更多有关目标函数的信息。同时，XGBoost 在目标函数中新增了分支节点权重的正则化项，使得模型有更低的方差^[5]，提升了模型的性能和效果。

从缺失值处理的角度来看，XGBoost 增加了自动处理缺失值的策略，通过计算缺失值样本处于不同分支状态下的目标函数，并根据目标函数的优劣自动进行缺失值样本分类^[6]，能够非常好地解决反欺诈场景中设备、场所、IP 等要素信息缺失的问题。

与逻辑回归 (logistic regression, LR)、随机森林 (random forest, RF)、支持向量机 (support vector machines, SVM)、人工神经网络等人工智能算法进行深入比较可知，XGBoost 具有分类效果出众、业务解释性好和在线预测速度快等优点。主流人工智能模型算法比较见表 1。

基于线上数字金融业务交易量大、正样本稀疏等特点，以及实时决策响应速度高的要求，尤其当欺诈模式或样本变化时，XGBoost 算法在预防速度和泛化能力方面的优势使其能够更好地适应实时交易反欺诈应用场景，故将 XGBoost 算法选作本次应用实践的机器模型算法。

2.2 特征构建方案

特征工程对于机器学习模型的构建异常重要，特征的质量直接决定了机器学习模型的性能上限^[7]。为寻找潜藏的欺诈规律、准确感知风险，本文设计了基于 RFM (recency, frequency, monetary) 的特征衍生方案，以精细刻画每一笔交易为目标，通过特征爆炸的方式构建了十几万维交易行为特征，并在此基础上探索了一套特征筛选方案，从海量特征中去劣选优，控制最终的入模特征规模，在不降低模型性能的基础上保证实时预测的可行性。

表 1 主流人工智能模型算法比较

算法	优点	缺点
XGBoost	分类效果出众； 业务解释性好； 预测速度快	模型复杂度高，容易过拟合
逻辑回归	实现简单； 预测速度快； 业务解释性好	当特征维度很高时，逻辑回归的数据处理过程工作量大； 不能考虑特征之间的相关性； 容易欠拟合、过拟合
随机森林	可以有效运行在大数据集上； 模型泛化能力强； 一般不需要额外做剪枝	容易过拟合； 取值划分比较多的特征容易对决策产生更大的影响，进而影响拟合的模型效果
支持向量机	可处理非线性问题； 能避免“维数灾难”； 具有一定的稳健性	难以实施大规模训练样本； 对缺失数据敏感； 对参数和核函数的选择敏感
人工神经网络	分类准确度高； 噪声神经有较强的稳健性和容错能力； 具备联想记忆的功能	需要调整大量参数，如网络拓扑结构、权值和阈值等； 输出结果难以解释； 学习时间长，甚至可能达不到学习的目的

2.2.1 RFM 特征衍生方案

RFM 是客户关系管理场景中常用的方案,可以通过对客户消费行为的刻画实现客户价值的细分。在具体的应用实践中,根据交易实时拦截的业务场景和线上交易流水的特点对基本的 RFM 框架进行了拓展,定义了以客户、账号、设备为核心的特征衍生主体(统计对象),并将 R 拓展为时间窗口,将 F 拓展为聚合函数,将 M 拓展为数据变量,具体衍生方法见表 2。

通过对 RFM 方案的拓展,从交易流水的统计变量出发,利用灵活多变的统计窗口和种类丰富的聚合函数,将相互独立的交易流水分别从时间维度、价值维度、空间维度进行衍生,构造出富含时空信息的特征,满足对交易行为多角度、细粒度、深景深画像的要求,为模型提供了充足的特征样本。

2.2.2 特征筛选方案

特征筛选是从衍生的样本空间中挑选有益特征、摒弃无关或冗余特征的过程,核心目的是提升模型预测效率及模型在未知数据上的泛化能力。

为保证降维后的特征具备良好的分类能力,并满足事中风控场景对时效性的严苛要求,通过应用实践,探索了一套“四步特征筛选方案”:第一步,基于 AUC 的单特征粗粒度筛选;第二步,基于 XGB 特征重要性的筛选;第三步,基于递归特征消除(recursive feature elimination, RFE)的迭代特征筛选;第四步,基于深度业务理解的特征筛选。在算法和业务专家的支撑下,解决了特征爆炸式衍生后的维度灾难,能有效剔除冗余信息与噪声特征,筛选出实时计算速度快、业务可解释性好、模型泛化能力强、预测结果稳定的入模特征。

3 特征构建及特征选择

3.1 数据准备

为深入挖掘欺诈交易与正常交易的区别,在数据准备环节抽取了 14 个月包含动账、登录、设置、贷款等交易在内的几十亿条线上交易流水,包括客户信息、账户信息、交易时间、交易金额、交易场所、交易耗时等内容。

通过数据质量评估和数据字段业务理解,在缺失数据填充、字段编码的基础上,剔除缺失率较高、唯一取值和业务价值低的字段,建立基础数据宽表。

最后,基于欺诈样本的积累,对欺诈交易逐条打标,形成包含正/负样本标记的建模数据宽表。

3.2 数据下采样及特征衍生

由于银行线上交易量巨大,欺诈交易在每天千万级的流水数据面前,显得非常稀疏,正/负样本比例极端不平衡。为了能够精准感知欺诈交易与正常交易的区别,避免欺诈交易的特征被正常交易淹没,同时为了提升建模效率,降低特征衍生环节的计算资源和存储资源开销,根据正样本的数量对负样本进行了下采样,正/负样本的采样比为 1:100。

采样后,通过对性别、年龄、金额、时间、活跃度等进行平衡性分析,确保样本能够代表总体数据的分布特性,避免因采样偏差导致的模型偏差问题。

最后,基于 RFM 衍生方案对采样后的样本进行特征衍生计算,每笔交易都获得了 19.8 万维的特征。

3.3 特征选择

根据“四步特征筛选方案”进行特征选择。

表 2

RFM 特征衍生方法

RFM 拓展方案	描述	要素
统计对象	进行多维度交易行为刻画的主体	客户、账户、交易设备
统计窗口	根据当前交易向前追溯的历史交易范围(包括时间窗和频次窗),可分别描述当前交易主体的长期与短期行为画像	时间窗:短期窗以小时为单位,如 1 小时、24 小时等,最长不超过 48 小时; 长期窗则以月为单位,如 1 个月、3 个月,最长不超过 6 个月; 频次窗:3 次、5 次、10 次
聚合函数	对统计变量进行深度特征挖掘的统计分析函数	连续型函数:最大值、最小值、和、均值等; 离散型函数:计数、频次、比值等
统计变量	聚合函数的加工对象,即各类原始交易字段及衍生特征字段,根据字段属性可分为连续型变量和离散型变量	连续型变量是指在一定区间内可以任意取值的变量,如交易金额等; 离散型变量是指有限固定取值、可列举的变量,如交易类型等

一是基于 AUC 的单特征粗粒度筛选。基于决策树算法计算单特征 AUC,剔除没有区分能力和区分能力差的特征,选出正/负样本区分能力较强的 3 万维特征。

二是基于 XGB 特征重要性的筛选。将 3 万维特征送入 XGBoost 模型,由算法根据特征重要性(weight)进行粗筛,选出 1 303 维特征。

三是基于 RFE 的迭代特征筛选。基于 XGBoost 模型进行递归特征删除,每次迭代删除 10%贡献度最低的特征,经过 29 次迭代,精细筛选 61 维特征。

四是基于深度业务理解的特征筛选。结合数据稳定性、工程实现难度和特征的业务可解释性,由风控专家对特征进行业务可解释性分析和分类结果评估,精选出业务解释性好、区分能力好、模型泛化能力强的 55 维入模特征。

特征筛选进程如图 1 所示。

在特征选择过程中,主要评价指标为模型在验证集上的 AUC。从 3 万维特征到 55 维特征的验证集 AUC 如图 2 所示。在特征选择的第 1 轮,3 万维入模特征的 AUC 为 0.991,此时 XGBoost 模型实际仅使用了 3 万维特征中的 1 303 维;经过 29 轮的 RFE 迭代,剔除冗余特征的同时也降低了过拟合的风险,61 维特征在验证集上的 AUC 小幅上升至 0.994 7;第 31 轮,经过业务理解的 55 维特征 AUC 为 0.994 8,基本与 61 维时保持同一水平。

从 19.8 万维到 55 维,特征压缩比高达 3 600:1。在“四步特征筛选方案”的“算法+业务”支撑下,模型在验证集上的 AUC 未降低,并在评估集上表现出了较强的泛化能力:从 AUC 指标看,评估集新数据 AUC 仅比验证集降低了 0.025,未出现显著下降;从业务角度看,模型查全率、查准率均优于专家规则,并达到了行业较好水平。

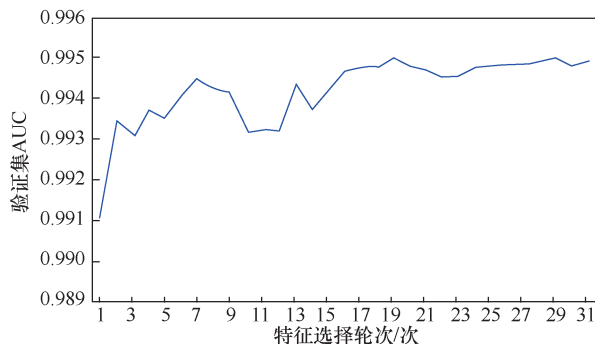


图2 特征选择验证集 AUC 曲线

3.4 入模特征评价

完成入模特征筛选后,从特征相关性和特征重要性等角度进行了进一步的评价和分析。

从入模特征相关性矩阵看,55×55 的特征相关性矩阵如图 3 所示。该图中每个小正方形代表 1 对特征的相关程度,白色方块表示线性正相关、黑色方块表示线性负相关,黑色、白色方块越少,特征间的相关性越小。从图 3 可以看出入模特征间的相关性程度低、独立性较高,这体现了特征对交易刻画维度的“多样性”。

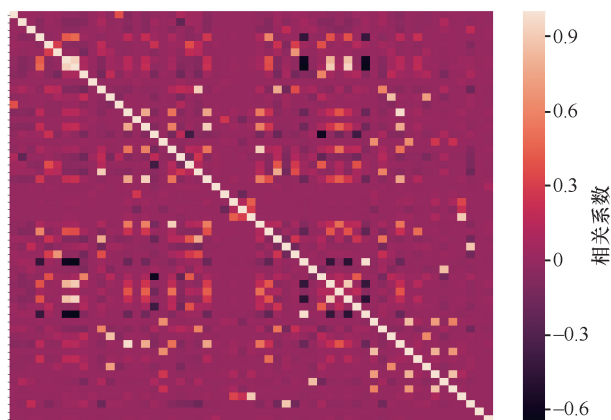


图3 55 维特征相关性矩阵

从入模特征重要性的角度看,入模特征包含时间、金额、场所、频率、距离、短期、长期、汇总、

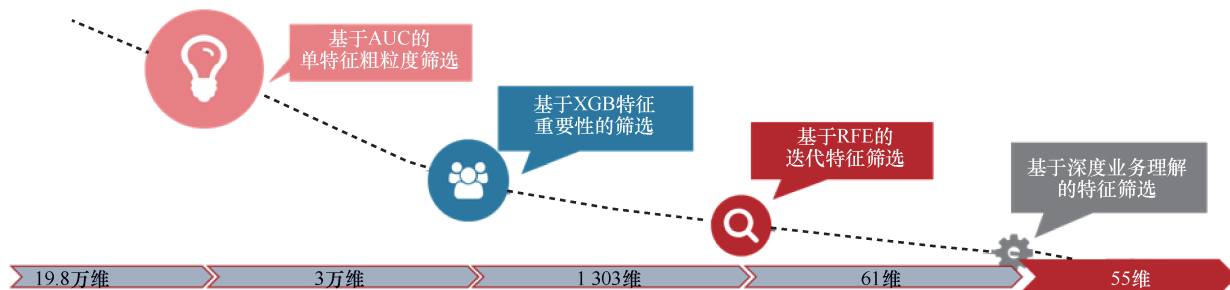


图1 特征筛选进程

加权等，表明了刻画角度的“丰富性”。特征重要性排序如图4所示。

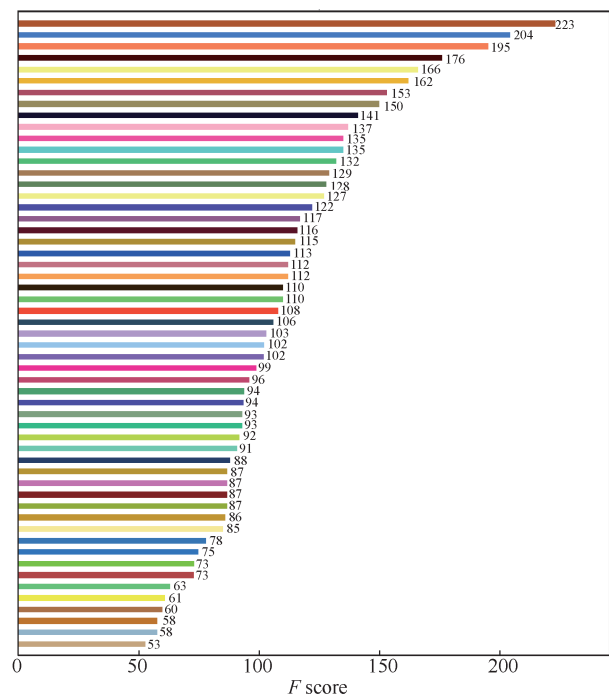


图4 特征重要性排序

综合特征相关性和重要性评价，55 维入模特征可从多维度、多角度对每一笔交易进行精细刻画，实现了交易欺诈风险识别的“千人千面”。

4 试验及结果

4.1 数据集划分

14 个月的整体数据按 14:3:3 的比例划分为训练集、验证集、评估集，训练集用于模型训练，验证集用于模型优化，评估集用于模型效果检验。为了能准确评价模型对“未来”的预测能力，数据集以自然月进行切分，10 个月用于模型训练、2 个月用于模型验证、2 个月用于模型评估。

前 12 个月的训练集、验证集根据正样本的数量进行了随机下采样，采样后数据规模降为 81.8 万笔。将第 13、14 个月的全量数据留作评估集，避免了数据穿越对模型泛化能力的影响，也能更准确地评价模型的真实表现。

4.2 模型训练及超参调优

根据欺诈交易识别准确率、召回率“双提升”的业务目标，以及实时决策的工程可实现性要求，进行了 21 次模型训练和优化迭代。根据每次迭代

模型效果，分别对正负样本权重、学习率、最大树深度、最大树棵数等参数进行了大量调整和优化，最终应用 55 维入模特征，训练了一个由 949 棵决策树（每颗树深度为 3）组成的 XGBoost 机器学习模型。

4.3 评价指标

4.3.1 技术评价指标

针对交易反欺诈场景正负样本极端不平衡的特点，将对数据不均衡敏感程度低的 AUC 与 KS 作为技术评价指标。两项评价指标均可由真阳率（TPR）和假阳率（FPR）计算。

表3 分类结果混淆矩阵

实际情况	预测结果	
	正	负
正	true positive (TP)	false negative (FN)
负	false positive (FP)	true negative (TN)

$$TPR = \frac{TP}{TP+FN} \quad FPR = \frac{FP}{FP+TN}$$

AUC 值表示不同阈值下 TPR 和 FPR 形成的 ROC 曲线下方面积，反映的是模型对正负样本的区分能力，AUC 值越接近 1，模型分类能力越强；KS 值为 TPR 与 FPR 差值的最大值，可用于量化模型的最优区分效果，也可用于指导制定模型输出概率值的最佳阈值切分点。

4.3.2 业务评价指标

从提升风控能力的业务目标出发，在具体的应用实践中将召回率和误报比作为评价模型效果的核心指标。

召回率也称查全率，表示在指定概率阈值下，模型识别出来的欺诈交易占全部欺诈交易的百分比，召回率越高说明模型识别的准确率越高。

误报比是查准率的倒数，表征模型正确识别一笔欺诈交易所需要执行的判定次数，该值越低意味着对正常客户的打扰程度越低。

$$\text{召回率} = \frac{\text{预测为正且实际为正的样本数量}}{\text{实际为正的样本数量}}$$

$$\text{误报比} = \frac{\text{预测为正的样本数量}}{\text{预测为正且实际为正的样本数量}}$$

召回率代表了欺诈交易的识别能力，误报比代表了欺诈交易识别的准确度，是风控能力最直接的评价标准。

4.4 模型性能

在本试验中,通过多次迭代优化后,模型各方面指标均达到较好水平,验证集模型 AUC 达到 0.997。应用第 13、14 个月的全量交易进行模型泛化能力评估,模型 AUC 达到 0.972,KS 值达到最大值(0.83)。

如图 5 所示,模型在评估集上的 AUC 与训练集的 AUC 相比下降程度较小,表明模型泛化能力强,在新数据上仍具有较好的性能。

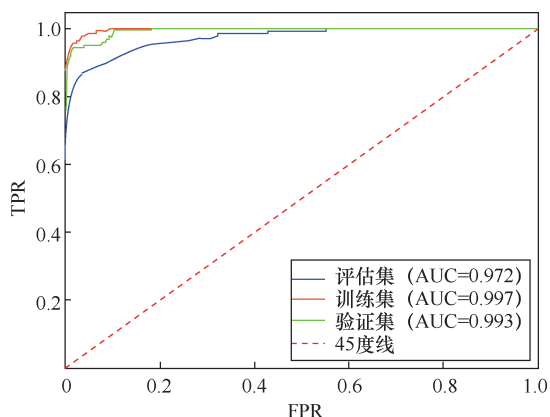


图 5 ROC 曲线

KS 曲线如图 6 所示,该模型的 KS 值为 0.83,此时,TPR=0.87,FPR=0.05,判定正负的评分阈值为 2.25。

误报比曲线衡量的是业务层面,随着误报比(成本)的上升,模型召回率(收益)逐渐增长,如图 7 所示。曲线越陡峭,说明模型识别欺诈交易的能力越强。同时,可参考该曲线,依据系统运营能力、业务目标以及欺诈形势对阈值进行灵活调整,建立柔性的欺诈防范体系。

5 模型应用方案

5.1 流式与批量结合的系统部署方案

机器学习模型的系统化部署主要涉及实时特

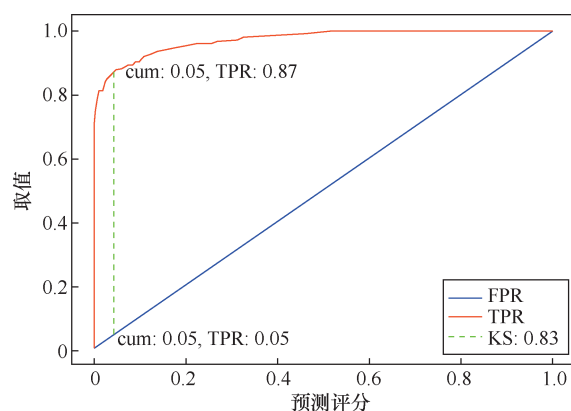


图 6 KS 曲线

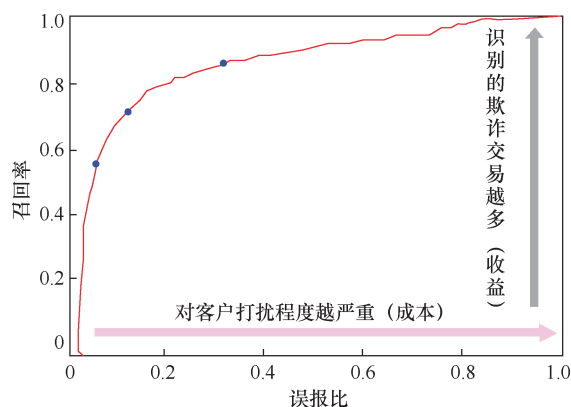


图 7 召回-误报曲线

征加工、历史特征加工、模型计算评分、风控系统决策和业务系统处置 5 个环节。为提升模型的运行效率,满足毫秒级的实时决策响应要求,本文在模型的系统化部署上,提出了基于智能风控系统和机器学习平台的“流式+批量”“规则+模型”的应用方案,发挥了企业级平台资源的优势。系统部署方案如图 8 所示。

通过智能风控系统,可实现基于流式数据的实时特征加工、历史特征合并,基于模型评分的决策

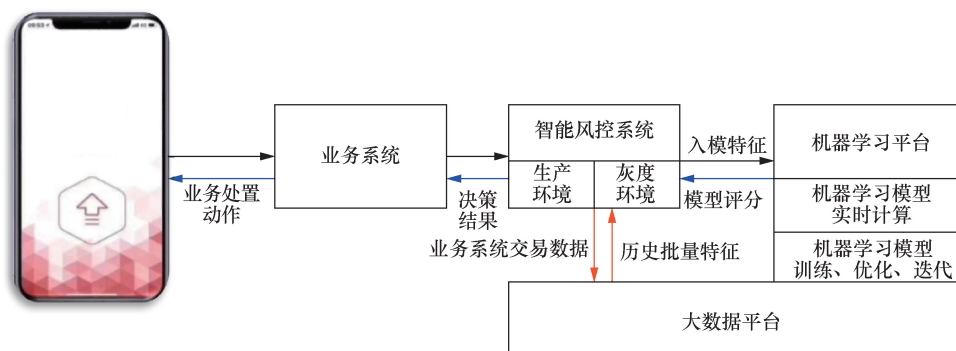


图 8 系统部署方案

结果生成、决策结果推送业务系统等功能。

通过大数据平台，运用海量数据存储优势和分布式计算资源优势，以每日跑批的方式对历史行为数据进行加工计算，生成历史特征。

通过机器学习平台，基于对模型运行计算能力的支撑，实时计算每笔交易的模型评分。

通过业务系统，根据智能风控系统的决策结果执行放行、拦截、挑战、增强挑战等处置动作。

5.2 模型与规则组合的业务应用方案

为有效发挥模型效用，通过对规则模型预警数据与机器学习模型预警数据进行综合分析，发现将机器学习模型与现有专家规则模型进行组合应用，发挥各自的长板，可有效提升风险查控能力，切实保障风险防控效果。

如图9所示，根据机器学习模型对每笔交易计算的欺诈评分，将全体交易划分为高风险头部区域、中风险中部区域和低风险尾部区域3类，不同类的交易与规则采用不同的组合应用方式。当交易属于头部区域时，无论专家规则判定结果如何，均进行增强认证；当交易位于尾部区域时，无论专家规则判定结果如何，均视为正常交易直接放行；当交易处于中部区域时，则采用以专家规则判定为主、模型评分为辅的方式，根据事中风控运营实时可用能力进行选择性的增强认证。

采用表4所述组合方式可以获得两大收益：一

是针对高风险的头部区域，可以有效挖掘专家规则之外的欺诈模式，提升召回率，实现“控风险”的目标；二是针对低风险的尾部区域，可大幅减少对正常客户的打扰，提升拦截的准确率，达到“升体验”的效果。

表4 规则/模型组合应用策略

专家规则判定	机器学习模型概率评分		
	头部区域	中部区域	尾部区域
高风险	增强认证	增强认证	放行
低风险	增强认证	选择性增强认证	放行

5.3 模型运行保障措施

在嵌入高频交易、实时决策处置的机器学习模型应用场景中，为确保模型高效运行、控制模型风险、保障客户体验，设计了超时、容灾两套应急方案，建立模型异常监控与自动处理机制，确保模型运行安全、稳定、可控。

一是建立超时处置机制。若超过30ms还未计算出模型评分，则模型自动停用，将规则判定结果作为交易挑战依据。

二是建立异常熔断机制。从交易类型、客群、特殊日期等维度考虑，分组设定每日挑战交易笔数阈值，若超过阈值，则模型评分不再触发交易决策；设定每日高风险交易笔数阈值，若超过阈值，则模型评分不再触发交易决策。

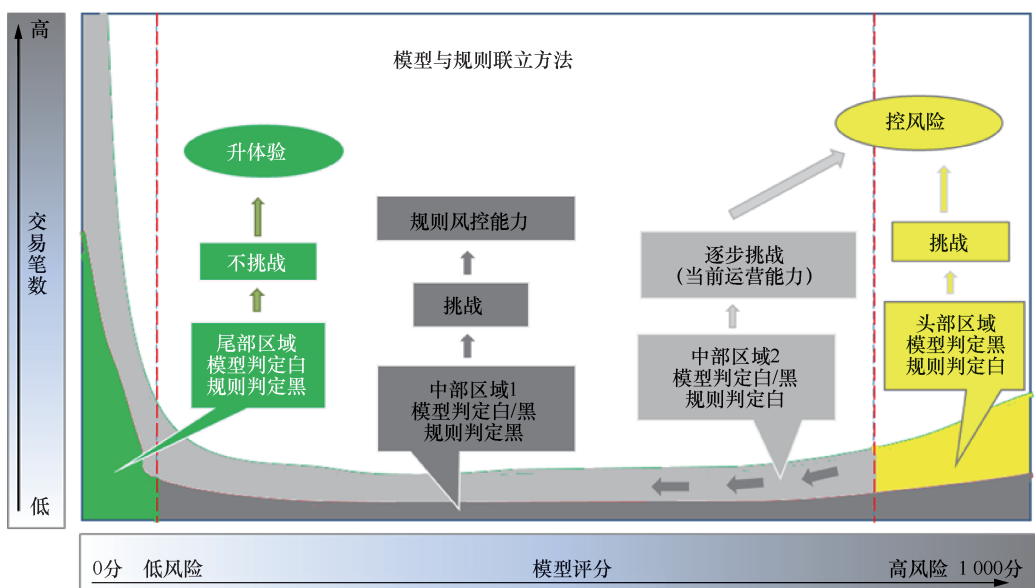


图9 规则/模型预警交易分布

6 模型应用效果及优化迭代

6.1 模型对规则的优化与提升

在当前专家规则模型日趋复杂、提升空间有限的情况下,引入机器学习模型可敏锐洞察用户的异常行为。从高维的行为特征角度看,客户的交易习惯是相对稳定的,因此在互联网黑产的作案手法发生变化时,也能感知到异于客户本人操作的欺诈交易。

根据模型对生产数据的应用与监控效果,将规则与机器学习模型进行组合应用后,在误报比降低40%的情况下,召回率提升了20%,实现了召回率和准确率“双提升”的目标,有效提升了系统识别能力,并降低了运营成本投入。

6.2 模型的优化迭代

6.2.1 监控与迭代机制

模型的效果提升无法一蹴而就,随着模型的系统上线运营和欺诈模式的变化,后续还需专业的数据分析师队伍进行不断的迭代、优化,持续提升效果,让模型发挥更大的价值和作用。为此,在具体的应用实践中建立了模型日常监控与定期优化迭代的机制,形成了模型监控-优化-迭代-灰度运行-发布生产的生命周期管理体系。

(1) 模型监控

在系统中固化模型监控报表,重点监控关键字段缺失率、特征 PSI 指标等模型稳定性监控指标, AUC 值、KS 值等模型性能指标, 召回率、误报比等业务评价指标, 定期重检模型运行效果。

(2) 模型优化与迭代

- “小”迭代。将运行监控效果与刚上线时的基准进行比较,在出现模型效果持续弱化、运行异常(如触发熔断机制、告警量过大或过小)等情况时,对模型进行“小”迭代,即不调整特征加工方案,只更新模型文件,在无须进行系统改造的前提下,快速迭代、提升模型效果。

- “大”迭代。为保证模型长期稳定运行,每年/每两年对整个风控模型进行大的演进,可能涉及入模特征加工方案变化。

6.2.2 优化迭代方案

根据模型上线后的运行效果、欺诈数据的积累,并结合人工智能技术的发展,将模型的优化迭代方案分为以下4类。

- 引入新的数据样本:模型的效果和预测能力高度依赖训练样本,持续引入新的欺诈样本对模型进行迭代,让模型学习最新的欺诈交易模式和最新的正常交易模式,这是保证模型的泛化能力的重要手段。

- 模型调参: XGBoost 机器学习模型有数十个超级参数,在模型效果不理想、无法达到预期业务目标时,可以结合欺诈交易特征,综合考虑调整学习率、正/负样本权重等关键参数,改善模型效果。

- 引入新的特征:随着业务的发展以及互联网黑产的对抗升级,引入新的特征将能提升模型对交易的刻画能力,这是模型在长周期运行后提升效果的有效手段。

- 引入新的算法:人工智能技术的发展日新月异,保持对人工智能技术的追踪与研究,积极探索和实现其他人工智能技术在反欺诈领域的应用,这是模型持续优化演进的必然方向。

7 结束语

在黑色利益的驱动下,银行与欺诈犯罪分子的对抗将是长期和持续性的,外部交易欺诈仍是银行开展线上金融业务的最主要威胁。本文从特征工程方案、模型训练优化到系统化部署、模型应用策略、迭代优化等方面开展了全面深入的研究,对利用大数据和人工智能提升欺诈风控能力进行了有益的探索和实践,能够在一定程度上解决交易欺诈模式变化快、识别难的风控难题,对于数字金融时代商业银行的事中风险防控体系的建设具有一定的推广价值和借鉴意义。

参考文献:

- [1] 中国人民银行. 金融科技(FinTech)发展规划(2019-2021年)[R]. 2019.
Bank of China. Development plan of FinTech (2019-2021) [R]. 2019.
- [2] 张旭, 赵奕. “在线实时决策+离线机器学习”实现高效实时反欺诈[J]. 金融电子化, 2016(12): 89.
ZHANG X, ZHAO Y. “Online real-time decision-making + offline machine learning” for efficient real-time anti-fraud[J]. Financial Computerizing, 2016(12): 89.
- [3] FRIEDMAN J H. Greedy function approximation: a gradient boosting machine[J]. The Annals of Statistics, 2001, 29(5): 1189-1232.
- [4] CHEN T Q, GUESTRIN C. XGB: a scalable tree boosting system[C]// The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, August 13- 17, 2016, San Francisco, USA. New York: ACM Press, 2016: 785-794.

- [5] 刘戡祺, 张智斌, 陈昊昱, 等. 基于 XGBoost 集成的可解释信用评分模型[J]. 数据通信, 2019(3): 27-32.

LIU Y Q, ZHANG Z B, CHEN H Y, et al. Interpretable credit scoring model based on XGBoost integration[J]. Shuju Tongxin, 2019(3): 27-32.

- [6] 成晨, 程新洲, 张恒, 等. 智能化调参的 XGBoost 算法及其在电信营销中的应用[J]. 邮电设计技术, 2018, 512(10): 26-30.

CHENG C, CHENG X Z, ZHANG H, et al. Intelligent parameter adjustment XGBoost and its application in telecom marketing[J]. Designing Techniques of Posts and Telecommunications, 2018, 512(10): 26-30.

- [7] 仵伟强, 后其林. 基于机器学习模型的消费金融反欺诈模型与方法[J]. 现代管理科学, 2018(10): 51-54.

WU W Q, HOU Q L. Anti-fraud models and methods for consumer finance based on machine learning models[J]. Modern Management Science, 2018(10): 51-54.

[作者简介]



曹汉平(1976—), 男, 博士, 中国银行总行个人数字金融部副总经理, 主要研究方向为宏观经济、大数据系统架构与应用、新型渠道、数字金融风险防控等。



张晓晶(1980—), 女, 中国银行总行个人数字金融部高级经理, 主要研究方向为机器学习建模、智能风控系统建设、数字金融业务风险防控等。



祝睿杰(1986—), 男, 中国银行总行个人数字金融部业务经理, 主要研究方向为数据挖掘、机器学习建模、数字金融业务风险防控等。



黄潇拉(1988—), 女, 中国银行总行个人数字金融部业务经理, 主要业务方向为数据分析、智能风控系统建设、数字金融风险防控等。