互 联 网 金 融 支 付 产 业 安 全 联 盟

支付风险智能风控应用与评估指引

Guidelines for Application and Evaluation of Intelligent Risk Prevention and Control System in Payment Scenario

2020 – 11– 24 发布

互联网金融支付产业安全联盟秘书处

目　　录

前　言

本指引是互联网金融支付产业安全联盟支付风险防控系列指引之一；

本指引由互联网金融支付产业安全联盟秘书处组织编写，同盾科技、邦盛科技、维择科技等机构参与起草编写。其中同盾科技起草第7章节，并对5.1节做出重要贡献；邦盛科技起草第6章节和第9章节，并对7.1节做出重要贡献；维择科技起草第8章节和第10章节，并对7.2节做出重要贡献；秘书处起草第1章至第5章，并负责统稿校对。

引　　言

伴随宏观经济环境变化、支付监管愈趋从严、金融科技不断创新、支付参与主体日趋多元，支付行业正面临着业务发展与合规经营、支付便捷与安全、数据挖掘与隐私保护等诸多挑战，支付风险的复杂性与日俱增，共同建设安全支付生态的必要性不断凸显，行业风险联防联控工作在面临着巨大压力的同时也正孕育着机遇，可以预期风险防控能力将成为支付参与主体获取竞争优势的关键。

人民银行前期已组织编写了《基于大数据的支付风险智能防控技术规范》（征求意见稿），明确利用大数据技术对风控数据进行保护、接入、处理、存储、变量计算等，进而支撑风险控制技术在风险防控策略、风险信息过滤、风险评估模型设计、风险决策及风险运营等风控流程中的应用，实现精准而全面的防控各类支付风险。

为加快推进智能风控技术的应用落地，快速提升产业风控能力，互联网金融支付产业安全联盟秘书处以合作互利共赢的理念为引领，凝聚行业共识，组织产业各方提炼最佳实践经验并编写指引，为支付参与主体提供智能风控体系建设、运营、优化的相关框架及评价方法，引导各支付参与主体优化完善风险防控体系。

支付风险智能风控应用与评估指引

1 范围

本指引适用于商业银行、非银行支付机构、清算机构等各类支付业务参与主体，指导各类参与主体应用智能风控技术防范支付风险、评估智能风控技术应用效果。

2 适用风险类别

本指引所提的支付风险主要包括欺诈风险和合规风险两大类，各类支付风险中具体情形如下：

2.1 欺诈风险

本指引所提的欺诈风险，指涉嫌使用虚假身份获取银行卡（或账户），或冒用他人银行卡（或账户）获取商品或服务的欺骗性交易行为，或指不法分子通过电话、网络和短信方式，编造虚假信息，设置骗局，对受害人实施远程、非接触式诈骗，诱使受害人本人给不法分子打款或转账的犯罪行为。

包括但不限于：

失窃卡欺诈，指冒用或盗用持卡人的银行卡进行欺骗交易，盗取账户内资金，包括丢失卡与被盗卡两种情形。

未达卡欺诈，指截取发行、交付过程中的银行卡并进行的欺诈交易。

虚假申请欺诈，指使用虚假身份或冒用他人身份获取银行卡（或开立账户）进行的欺诈交易。

伪卡欺诈，指非法使用银行卡磁条信息伪造真实有效的银行卡，或通过改造丢失卡、被盗卡、未达卡、过期卡的表明凸印（含全息防伪标识）信息或重新写磁后进行的欺诈交易。伪卡欺诈包括伪造卡、变造卡、白卡欺诈。

账户盗用，指假冒真实持卡人身份或盗用银行卡账户信息后进行的欺诈交易。账户盗用方式包括通过变更持卡人地址、要求换发卡等方式盗用银行卡，假冒真实持卡人身份申请开通手机闪付、二维码支付等，盗用支付应用信息后接管应用服务方账户等。

非面对面欺诈，指窃取或骗取卡片主账号、PIN、有效期、支付短信验证码及其它关键身份验证信息后，通过邮购/电购、互联网、手机等非面对面渠道发起的欺诈交易。

营销欺诈：借助机构推出的优惠活动进行恶意的资源套取和倒卖以达到获利目的，并导致实际优惠无法触达营销目标用户的欺诈交易。

2.2 合规风险

本指引所提的合规风险，指机构、个人在支付业务中，由于未能遵循法律法规、监管要求、行业规定给自身及其他参与方带来经济、声誉损失或影响的风险。

包括但不限于：

1、相关支付业务或产品被个人利用开展网络赌博、洗钱等非法或违规活动的合规风险：

网络赌博，指利用互联网进行的博彩行为。

洗钱风险，指将违法所得及其产生的收益，通过各种手段掩饰、隐瞒其来源和性质，使其在形式上合法化。常见于毒品犯罪、黑社会性质的组织犯罪、恐怖活动犯罪、走私犯罪、贪污贿赂犯罪、破坏金融管理秩序犯罪、金融诈骗犯罪等各类违法犯罪过程。

恶意套现，指持卡人或团伙套取信用卡或消费贷资金时恶意不还，造成相关金融机构发生资金损失的行为。

2、机构自身违反监管政策要求和行业制度规定，开展挪用客户备付金、为无证机构提供支付结算服务、系统化变造交易等违规活动，给其他支付业务参与方带来风险损失及影响的合规风险。

3 主要参考文件

《GB/T 35273-2020 信息安全技术 个人信息安全规范》

《GB/T 37973-2019 信息安全技术 大数据安全管理指南》

《基于大数据的支付风险智能防控技术规范》（征求意见稿）

4 术语及定义

即时查询

根据需要，临时组装查询条件进行数据查询。

风险特征库

风险特征是风险案件的行为的刻画，来源于金融交易数据挖掘与关联的非金融交易数据挖掘，具备描述风险行为的功能。

专家规则

专家规则是指该行业内具有丰富实践经验的风控专家，总结风险行为中某些特征或特征组合，归纳形成具有因果关系的运行法则。

标签画像

抽象分类和概括某一类特定群体或对象的某项特征，形成用户标签；特定群体或对象的多项特征组合构成用户画像。

机器学习

通过算法运用计算机提取风险行为的有用特征，构造特征到标签的映射，建立模型，识别用户风险行为。

关系网络

基于已有数据开展挖掘，利用图的数据结构，表达复杂多层的关联关系，最终识别关联性风险。

无监督聚类

针对已知数据空间范围内的，将一组数据按照相似性和差异性分类，使相似数据集聚在同一类别，并使不同类别间的数据产生足够的差异性。

异常检测

检测具有异常特性的实体或团伙，属于无监督机器学习中的一种方法。

决策树

在已知各种情况发生概率的基础上，通过构成决策树来求取净现值的期望值大于等于零的概率，评价项目风险，判断其可行性的决策分析方法，是直观运用概率分析的一种图解法。由于这种决策分支画成图形很像一棵树的枝干，故称决策树。

随机森林

随机森林是一个包含多个决策树的分类器， 并且其输出的类别是由个别树输出的类别的众数而定。

梯度提升决策树

是一种迭代的决策树算法，由多棵决策树组成，所有树的结论累加起来作为最终答案。

支持向量机

是一类按有监督学习方式对数据进行二元分类的广义线性分类器，其决策边界是对学习样本求解的最大边距超平面。

朴素贝叶斯分类

是一系列以假设特征之间强（朴素）独立下运用贝叶斯定理为基础的简单概率分类。

K-means算法

是一种迭代求解的聚类分析算法，其步骤是随机选取K个对象作为初始的聚类中心，然后计算每个对象与各个种子聚类中心之间的距离，把每个对象分配给距离它最近的聚类中心。

BIRCH算法

是一个综合的层次聚类算法。它用到了聚类特征和聚类特征树两个概念，用于概括聚类描述。聚类特征树概括了聚类的有用信息，并且占用空间较元数据集合小得多，可以存放在内存中，从而可以提高算法在大型数据集合上的聚类速度及可伸缩性。

DBSCAN算法

是一个比较有代表性的基于密度的聚类算法。与划分和层次聚类方法不同，它将簇定义为密度相连的点的最大集合，能够把具有足够高密度的区域划分为簇，并可在噪声的空间数据库中发现任意形状的聚类。

混合模型（mixture models）

包含不同尺度机制的模型，其中宏观层次的变量受到微观层次变量的调制，但是微观变量部分地表现出参数化特征。

Isolation Forest

是一种适用于连续数据的无监督异常检测方法，与其他异常检测算法通过距离，密度等量化指标来刻画样本间的疏离程度不同，孤立森林算法通过对样本点的孤立来检测异常值。

轮廓系数

轮廓系数（Silhouette Coefficient）结合了聚类的凝聚度（Cohesion）和分离度（Separation），用于评估聚类的效果。该值处于[-1,1]之间，值越大，表示聚类效果越好。具体计算方法如下：

假设已经通过一定算法，将待分类数据进行了聚类。常用的比如使用 K-means ，将待分类数据分为了 k 个簇 。对于簇中的每个向量。分别计算它们的轮廓系数。

对于其中的一个点 i 来说：

计算 a(i) = average(i向量到所有它属于的簇中其它点的距离)

计算 b(i) = min (i向量到与它相邻最近的一簇内的所有点的平均距离)

那么 i 向量轮廓系数就为：

$$S(i) = \frac{b(i) - a(i)}{max\{a(i), b(i)\}}$$

将所有点的轮廓系数求平均，就是该聚类结果总的轮廓系数。

SDK

软件开发工具包

宿主APP

集成SDK的APP

AES256

密码学中的高级加密标准（Advanced Encryption Standard，AES），又称Rijndael加密法，是美国联邦政府采用的一种区块加密标准。256表示秘钥长度256位。

SM4

SMS4分组加密算法是中国无线标准中使用的分组加密算法，在2012年已经被国家商用密码管理局确定为国家密码行业标准，标准编号GM/T 0002-2012并且改名为SM4算法，与SM2椭圆曲线公钥密码算法，SM3密码杂凑算法共同作为国家密码的行业标准，在我国密码行业中有着极其重要的位置。

前向安全性

前向安全性或前向保密性（英语：Forward Secrecy，缩写：FS），有时也被称为完美前向安全（英语：Perfect Forward Secrecy，缩写：PFS），是密码学中通讯协议的安全属性，指的是长期使用的主密钥泄漏不会导致过去的会话密钥泄漏。

数据脱敏

数据脱敏，指对某些敏感信息通过脱敏规则进行数据的变形，实现敏感隐私数据的可靠保护。这样就可以在开发、测试和其它非生产环境以及外包环境中安全地使用脱敏后的真实数据集。

5　总体应用框架

智能风控体系包括事前风控数据采集与处理、事中风险监测与预警、事后风险排查处置三个部分，整个体系基于风控指标开展效果评估，并基于效果评估不断优化完善，同时应符合信息安全要求并满足运营性能需求，总体应用框架如图1所示。

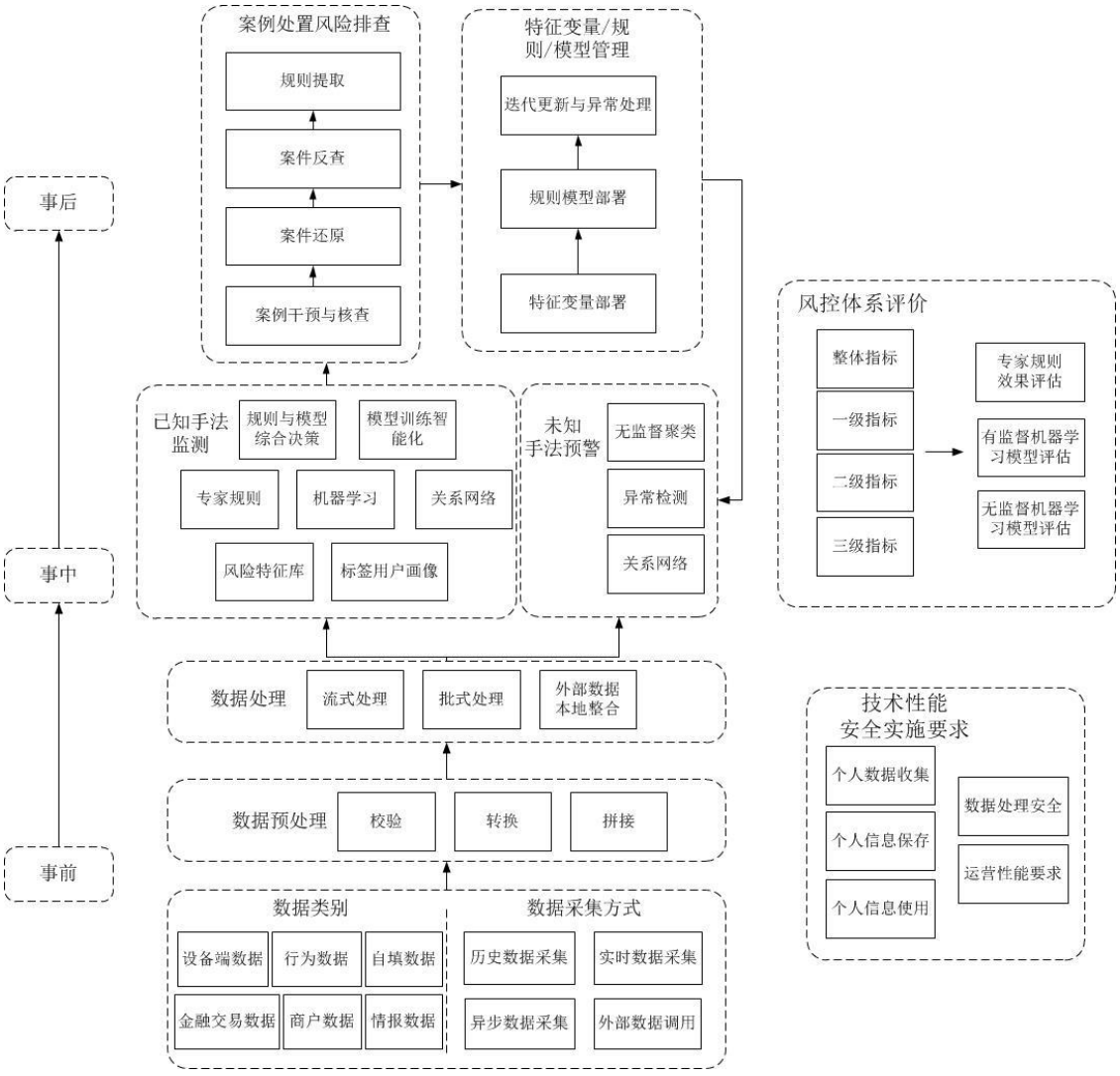

图1　智能风控总体应用框架

5.1　应用环节

5.1.1 事前风控数据采集与处理

风控数据采集是风险监测的前提，通过收集各类信息数据、行为数据、位置数据，全面描述交易主体的行为动作、关系、目的等属性；在数据采集过程中，需要清洗整理形成结构化数据，并抽取其中的重要信息，最终输出各种维度的数据、报表、图像、语音等形式的信息，从而为后续风险交易的风险特征提取作准备。

5.1.2 事中风险监测与预警

事中风险监测与预警是指基于风险数据，对交易行为动作执行过程的疑似风险交易进行实时或非实时性的干预，包括交易阻断、增强认证、挂起等操作其中，对于已知手法的风险监测主要基于风险特征、标签画像和关系网络，通过专家规则+有监督机器学习模型的方式识别风险交易；对于未知手法的预警，主要采用无监督机器学习模型以及等方法，开展异常检测，从而预警新型风险行为。

5.1.3 事后风险排查与处置

事后风险排查目标是基于已知风险特征识别潜在风险，确保整个支付交易流程体系稳定。

风险处置包括资金处置与非资金处置两类。资金处置包括退回拦截资金或赔付资金损失。非资金处置包括止付账户、要求修改密码、智能身份验证[1]等，收单侧非资金处置还包括关停交易、撤机、降低交易额度等。

风险特征/规则/模型管理是智能风控体系运营的必要环节，能够有效保障智能风控体系发挥作用。

5.2 规则/模型智能风控评估

欺诈风险智能风控应用评价的核心指标是欺诈率，即欺诈交易占业务交易总金额的比例，可以最直观的评估风控整体效果。

规则和模型的分类效果评估则主要基于覆盖率和准确率指标。其中一级指标和二级指标用于评估专家规则效果，一级指标包括风险覆盖率、交易报警率、准确率，二级指标包括用户打扰率、误报率、漏报率等；三级指标主要用于评估有

[1] 智能验证常见手段有：图片验证码、语音验证、人脸识别等。

监督机器学习模型和无监督机器学习模型效果，具体包括查准率（即准确率）、召回率（即覆盖率）、PR曲线、ROC 曲线/AUC、F1分数、轮廓系数等。

合规风险智能风控应用评价主要着重于对于重大合规风险事件的覆盖率，同时可引入用户打扰率、准确率等规则和模型一级及二级指标量化评估具体风控效果。

另外，在交易及支付等时效性要求高的场景，还需专门针对性能指标进行评价，包括吞吐量和高可用性等。

5.3 信息安全与运营性能要求

机构开展风险防控时，需在合法依规前提下采集、保存、处理及使用个人信息，遵循相关安全要求；风险监测与数据处理方式需要匹配业务风险防控需求，确保风控运营流程顺畅，不影响业务开展。

6 事前风控数据采集与处理

智能风控应用的前提是获取充分的风控数据，数据来源不同相应采集方式也存在差异，采集后还需进行预处理确保数据完整可用。数据处理环节，根据数据和智能风控应用场景不同特点可采用流处理、批式处理等方式。

6.1 核心风控数据采集与预处理

6.1.1 数据类别

丰富的数据是智能风控系统处理的基础，也是风控效果的有力保证，核心风控所需数据主要包括端数据，非金交易数据，金融交易数据，用户自填数据、外部数据和黑灰产情报数据。

1、设备端数据

指客户端设备数据，涵盖各类非隐私要素信息。典型的端数据包括IMEI，屏幕分辨率等。风控系统主要基于这些数据为设备生成唯一的身份识别码（设备指纹），同时也可以获取设备型号等信息，丰富数据维度。

2、 行为数据

主要是指注册、登录等各种非金融交易的行为数据，用以在金融交易发生之前识别风险，也可以用作金融交易前的行为分析，将其与金融交易的行为相结合，识别潜在的交易风险。典型的非金融交易数据包括登录时间，登录IP等。

3、金融交易数据

主要是指涉及到金融交易的动账类数据，例如充值、转账等。由于金融交易会产生直接的资金损失，所以金融交易行为的异常需要更全面的识别和管控。典型的金融交易数据包括转出账户，交易金额等。

4、商户/用户注册信息数据

主要指商户或用户注册时的信息，包括名称、身份证号、手机号、职业、年龄、商户注册地、注册资金等相关注册信息。欺诈行为经常体现在注册信息的虚假和聚集上，通过注册/入网信息，可以有效提前识别欺诈主体。

5、商户舆情数据

主要指部分商户涉及的企业违规经营、涉诉、司法执行等，此类商户风险较高，需及时预警并执行相应风控策略。

6、自填数据

特指为用户根据风控要求为完成交易（包括金融交易或非金融交易），在客户端补充填入的信息。

7、外部负面数据

主要是不法分子常用的手段或信息相关外部数据，或者非法网站、非法商户等外部信息，用来辅助区分是正常行为（或交易）与异常行为（或交易）。典型外部数据包括代理IP列表，虚假手机号列表等。

在部分合规风险场景下，例如网络赌博，赌博网站等外部信息是一种重要的信息来源，基于赌博网站，可还原赌博收款商户信息及收款账户信息等外部数据，此类数据对于网络赌博交易识别与监控十分重要。

8、黑灰产情报数据

主要是指通过社交软件、网站及相关黑产社区，深入黑灰产内部，获取一手黑产信息，从攻击者的角度了解不法分子的策略、技术、流程，提供实时情报，帮助机构提前感知风险，及时止损。

6.1.2 采集方式

数据采集根据数据来源的不同，可分为机构内数据采集和机构外数据采集。机构内数据采集方式又分历史数据采集、实时数据采集、异步数据采集；机构外数据采集主要是指外部数据采集。

1、历史数据采集

历史数据采集对象，主要是机构内过去一段时间内留存的历史数据，包括客户信息、银行卡信息、转账数据、取现数据、余额查询数据等。采集目的是为支持需要历史数据的风控规则立即生效。

根据具体风控规则，可能需要获取不同时间跨度的数据，例如半年的交易信息、全量的客户信息等。历史数据一般来源于机构内集中的大数据平台，如果缺少此类大数据平台，则可以从各相关系统的备库中采集。

2、实时数据采集

实时数据采集，是指在交易中采集的实时数据，包括当前交易时间、当前交易金额、设备信息等。实时数据的丰富程度越高，风控效果越好，例如对于线上交易若能获取到客户端的GPS信息，则可依据GPS归属地制定规则，从而辅助识别异地操作的风险交易。

一般来说，在数据信息从前端渠道送往后端授权系统处理的过程中，相关信息如IP地址等可能被过滤，导致风控规则能使用的风险特征有所减少，因此实时采集越靠近客户端采集效果越好。

3、异步数据采集

异步数据采集，主要是指非实时数据的采集，例如账户的注册信息等，这类数据通常存在于其他系统或介质中，用以辅助识别事中的风险交易。

当前渠道或系统里没有留存的数据，例如黑名单、内部征信等信息，需要从其他系统抽取，以便在实时风控时应用。为保证交易风控的时效性，此类数据无

法通过实时接口调用的方式获得，所以通常定时批量从其他系统抽取到风控系统，以便在实时风控中及时调用。

4、外部数据采集

外部数据采集后用于辅助识别可疑交易，例如虚假手机号库、代理IP库等。对于数量庞大的外部数据，通常由外部供应商侦测收集，并提供给金融机构使用。

在网络赌博此类合规风险场景下，关于赌博网站中收款商户信息及收款账户信息等外部数据的采集，需通过人工方式或网络爬虫等技术手段抓取官方公开媒体等渠道发布的赌博网站中的信息，如其中收款信息是通过二维码方式展示，还需还原二维码申码时的收款商户或收款账户。

6.1.3 预处理

机构内机构外收集的各类数据，需要进行一系列的预处理才能运用于后续处理计算。

1、数据校验

生产环境采集的数据，由于各种原因，不可避免会出现空值、超范围值、错误类型值等各类异常。在进行数据处理和指标计算之前，数据校验是非常重要的预处理流程，在合规风险场景下，部分数据可能通过网络爬虫等技术手段自动抓取，或涉及二维码收款信息的还原，对于数据有效性的校验不可或缺。

2、数据清洗

经过数据校验后，需要针对异常数据进行清洗。包括但不限于将空值用合适的默认值替代，将超范围的值用正常值修正等操作，剔除无意义的错误值等。数据清洗过程可配置化程度越高，清洗过程越准确高效。

3、数据转换

各类数据源产生的数据采用的数据字典不尽相同，可能会存在同样字段的枚举值不同，例如一份数据使用数字标记该枚举值，另一份数据使用字符串标记枚举值；也可能一份数据的某个字段，在另一份数据中需要由2个字段拼接起来或计算得出。此类转换加工工作，在数据处理和指标计算前必不可少。

4、数据拼接

部分渠道或场景采集到的数据并不完整，需要与不同来源的另一份数据拼接，才能形成一条完整的交易流水，运用于风险交易智能识别。此类数据拼接操作，在数据处理和指标计算前必须完成处理。

6.2 数据处理

风控数据的处理，主要是指内外部数据的整合加工，包括在实时风控时进行流式处理，准实时风控或事后批量分析时进行批式处理，外部数据整合应用等。

6.2.1 流式处理

1、定义及简介

流式处理，是指根据每条数据的流入，结合历史指标处理新的增量数据。流式处理无需针对整个数据集执行操作，仅对于每个数据项执行相应的数据处理。

流式处理中的数据集是“无边界”的，符合下列特征：

a)完整数据集只能代表截至目前已经进入到系统中的数据总量；

b)工作数据集在特定时间只能代表某个单一数据项；

c)数据处理工作是基于事件进行的，除非事件明确停止否则指标计算没有“尽头”。

d)处理结果立刻可用，并会随着新数据的抵达继续更新。

流式处理系统可以处理几乎无限量的数据，但同一时间只能处理一条（真正的流式处理）或很少量（微批处理）数据。

2、复杂算法支持

流式处理需要能够支撑复杂算法，以便应对不法分子不断更新的作案手法。例如在实际应用中，针对机器有规律的间隔尝试登录，或者伪卡盗刷商户金额连续递减等现象，需要相应算法识别规律性。

3、长周期指标支持

部分风控规则需要基于长时间跨度计算，例如某一账户过去24小时转出的对手账户数量可能比较少，但将时间周期扩大到3个月，该账户的对手账户数量可能会存在集中性异常特征。部分规则甚至需要查看过去1年的数据情况，因此在流式处理中支持长周期数据计算显得非常重要。

4、大维度指标支持

部分指标涉及到较多维度，例如一笔交易可能涉及成百上千个指标维度的分解组合计算并最终及时输出结果。大维度指标在风控数据处理中不可避免，所以也需要流式处理引擎有能力处理此类数据。

5、适用场景

流式处理常见于可疑交易的实时监控与拦截，因此需要具有高响应性，以应对业务场景下有可能剧烈变动的流量，例如营销优惠期间的交易量突增。流式处理中运用长周期大维度指标、复杂算法时，对数据处理能力要求较高。

受存储空间和性能限制，单纯流处理对需要保存原始数据的指标（例如“同地址的客户列表”指标需要保存所有客户及地址）的加工处理同样存在瓶颈。风控系统在使用此类指标时，建议采用流处理+批处理相结合的方式。

数据处理需要具备灵活性、实时部署生效的能力，来缩短风控规则、模型从训练到部署的整个研发周期。

6.2.2 批式处理

1、定义及简介

批式处理，是指针对有限且确定的数据集进行批量分析计算，并在计算过程完成后返回结果。批式处理的数据通常具备大容量的特点，经常用于分析大量历史数据。

批式处理中的数据集通常符合下列特征：

a)有界：批式处理数据集是个有限的数据集合；

b)持久：数据通常始终存储在某种类型的持久存储位置中，为提升计算性能也可能加载到内存进行计算；

c)大量：批式处理通常处理极为海量的静态数据集；

相比流式处理而言，批式处理聚焦于针对静态的、现有的数据集计算，每次计算时需要将所需计算的数据全部加载并计算，在支持即时查询上具备较大的灵活性，但基于实时流水数据处理历史数据时，时效性上通常逊于流式处理。

2、适用场景

当数据为日终获得时，或者需要基于历史数据进行即时计算查询时，使用批处理较为有效，常见于针对批量客户（或账户）评级评分。

6.2.3 外部数据整合

外部数据，在实时交易监控中往往具有良好效果，这些数据包括但不限于风险手机号码、风险IP、各类归属地解析、各类黑名单等。由于支付风险对处理时效性要求非常高，所以此类数据需要落地在本地，从而保证尽可能快速完成数据服务的调用。

7 事中风险监测与预警

事中风险监测根据作案手法是否已知可以分类为对已知手法的风险监测和对未知手法的风险预警，通过专家规则、机器学习、聚类分析、关系网络等方法分析样本数据，识别高风险交易，预警异常特征。

7.1 应对已知手法的风险监测

7.1.1 主要目标

应对已知手法的风险监测基于对已知作案手法的调查分析，使用风控数据总结描述具体特征，并与监测对象进行比较，通过计算相似程度识别风险交易。

7.1.2 体系搭建

1、风险特征库

风险特征是对风险行为的刻画，可直接用于与监测对象比较。风险特征的总结除了需要金融交易数据外，还需要与金融交易数据相关联的非金融交易数据等。风险特征的总结计算可能涉及长周期的风险数据挖掘与统计。

风险特征库的建立方法：

1. 构建标准“事件日志”（风险实例）：用于确保风险事件的操作具可还原，主体，场景，动作，对象，结果。

2. 构建标准“案件库”（风险案件）：约定企业内部形成基础的案件梳理办法，至少包含 - 创建时间，业务，涉及场景，风险类型，攻击手法，回溯起

点，涉案金额，涉案个体（账户，金额，设备，IP，POS……），每个“案件库”是多个“风险实例”的总结。

3. 原始特征总结：通过对风险案件中风险实例的时间，空间，动作，主体属性的测量，直接得到的特征；

4. 有效特征筛选：通过一系列的统计计算，对原始特征进行筛选，进而得到特定场景下，特定问题的相对比较敏感的特征。

5. 特征体系构建：由于特征的构造是基于复杂问题的认知，为了便于后期维护，必须对构建的特征进行体系化的分类整理，比如：原始特征，有效特征，静态特征，动态特征，稳定特征，枚举特征 等等，

6. 特征维护：特征的维护方案需要在特征构建初期同步完成（初稿），避免后期的维护出现断层，使得特征失效逐步的影响到上层应用（规则、模型）

2、标签及画像

标签是对某一类特定群体或对象的某项特征抽象分类和概括的结果，标签值具备可分类性。画像由某一特定群体或对象的多项特征构成，可理解为多个标签的集合。

标签分为基础属性标签和业务知识标签：

a) 基础标签用于描述客户特征及行为的客观存在，不需要经过深入的转换和分析就能直接得到，这些属性往往是用户在使用产品时直接反映，或者是用户本身所具有的显性属性，是用户数据转化为信息的表现。

b) 知识标签是具备描述某种业务场景的标签，存在主观的变化性，来源于基础标签与业务场景规则综合提炼，是信息转化为业务知识的表现。

基于标签的用户多维分析是一项重要的数据服务功能，应用标签与画像可以更好的连接数据和业务人员，降低数据获取和操作的难度，赋能业务人员参与到数据分析之中。

3、专家规则

专家规则是指该行业内具有丰富实践经验的风控专家，将已识别风险行为中具有某些特定特征（或特征组合）总结成具有因果关系的运行法则。按照规则类型进行分类，包括但不限于：

a）静态规则，基于某个静态值的限制性规则，比如交易要素验证一致、是否命中黑名单、常用联系人验证等。

b）统计规则，使用统计计算值输入作为验证条件，如来自某个设备或者IP的申请次数大于某个值。

c）关联规则，指用户在进行了某个操作之后，又执行了另一个操作。例如，修改账号密码后进行交易操作。

d）行为规则，指用户的异常行为，例如用户在非常驻地址或者某个特殊的时间（如深夜）进行了交易操作，一般也认为具备一定的风险。

4、机器学习

机器学习可以根据丰富的数据和算法，对数据进行多重处理分析。机器学习根据已有的标签，通过提取风险行为的有用特征，构造特征到标签的映射，建立模型，实时识别用户风险行为。

有监督机器学习在已知手法的风险监测中应用最为广泛，其每组训练数据都有一个明确的标识，如可疑或正常交易，可以明确区分。输入已知风险数据和正常数据做训练集，可以训练出学习模型来填补并增强专家规则无法覆盖的复杂风险行为。

常见的有监督机器学习算法包括决策树、随机森林、梯度提升决策树、支持向量机和朴素贝叶斯分类等。

5、关系网络

关系网络是利用图的数据结构来表达现实问题中复杂的业务关系，并通过大规模的图计算算法，发现关联性风险，识别不法团伙。图的数据结构是由节点和边组成，每个节点代表一个个体，每条边代表个体与个体之间的关系。关系网络把不同的个体按照其关系连接在一起，从而提供了从“关系”的角度分析问题的能力，有利于从正常行为中识别出到异常的团伙性风险行为。

关系网络的结构取决于如何定义个体与个体之间的关系。在解决实际问题的时候，关系的定义需要依据业务需求并且常常极为复杂，因此只有基于大量的业务实践和专家经验，才能设计出一个强大的关系网络，以有效识别团伙性风险行为。

在已知作案手法及特征的情况下，可以基于已知风险样本拓展识别范围，寻找并识别发现其他的异常交易。

根据欺诈风险的实践经验，银行卡和欺诈设备构成的关系网络对于识别团伙性欺诈行为十分重要，尤其在设备唯一性标识可信的前提下，基于一张确认欺诈的银行卡对应的欺诈设备，可排查出此欺诈设备上其他的银行卡信息，再基于这些银行卡信息关联其他的移动设备，可挖掘出更多的欺诈信息。

根据合规风险中网络赌博等风险场景的实践经验，收款卡/商户和付款卡构成的关系网络在违规交易的风险侦测识别中十分重要，主要基于确认存在涉赌等违规行为的收款卡/商户排查对应的付款卡，再基于付款卡关联其他收款卡/商户，挖掘更多的风险信息。对于存在多张付款卡交集的收款卡/商户，以及交易金额、频率等特征异常的收款卡/商户和付款卡，应重点关注和调查。

关系网络单元需支持：

a) 分布式的数据存储和计算能力，能够随着硬件的增加提升计算能力；

b) 多图建设能力，能够在一个平台上构建不同风险场景的关系图；

c) 图规则的识别能力，能够部署图规则，并识别满足图规则的关系图。

7.1.3 计算量与算法要求

有监督机器学习在模型训练环节的计算量与原始数据量级、特征数量及复杂度、特征回溯长度、机器学习算法种类等有关。在模型预测环节，有监督机器学习模型主要消耗的是特征计算。

7.1.4 规则模型综合决策

智能风控体系应能实现规则和模型相结合的决策能力，通过规则决策灵活适应风控场景的变动，通过模型决策增强提升风险防控的准确性和覆盖率。智能风控系统应提供规则引擎和模型引擎相结合的双引擎技术框架，规则引擎和模型引擎需实现无缝对接，以支撑综合决策需求。

规则引擎，模型引擎的结果会被视为整个决策链中的关键节点，智能风控体系中的决策引擎，会将每一个关键节点按照需要进行二次构造（逻辑的与/或/非，数值的六则运算，实体关系的包含/比较 等等），只经过多元叠加的方法，使得决策结果更加的趋向理想结果。

7.1.5 模型训练的智能化

智能风控系统应确保机器学习模型训练和模型决策之间的无缝对接，训练结束的模型应能一键导入避免重新开发，防止部署时指标转译造成逻辑不一致，引起模型决策不稳定。

7.2 应对未知手法的风险预警

7.2.1 主要目标

随着支付产品的创新，不法分子作案手法和违规交易的特征也在不断更新迭代，风险防控需要更加智能化，随着攻击的发生自动检测出未知作案手法。无监督机器学习区别于传统的有监督机器学习，无需依赖于任何标签数据来训练模型，故主要用于应对未知手法的风险预警。

无监督机器学习可以通过关联分析和相似性分析，发现可疑用户行为间的联系，创建群组，并在一个或多个其他群组中发掘新型风险行为和案例，以应对复杂多变交易特征，保障支付安全。

7.2.2 体系搭建

1、无监督聚类

无监督聚类是常见的数据挖掘手段，适用于小额高频特征的交易监控。其主要假设是数据间存在相似性，而相似性存在价值，因此无监督聚类可以用于探索数据中的内在特性。例如在互联网时代，风险行为往往表现为大规模团伙形式，也就是“坏人扎堆，好人分散”，故可以通过无监督聚类的方式抓取异常团伙，例如虚假账户注册等场景。

聚类分析过程是将一组数据按照相似性和差异性分为几个类别，其目的是使得属于同一类别数据间的相似性尽可能大，不同类别数据间的相似性尽可能小。

常见的聚类算法包括基于划分的K-means算法、基于层次的BIRCH算法、基于密度的DBSCAN算法等。

2、无监督异常检测

异常检测是在无监督模型学习中比较有代表性的方法，即在数据中找出具有异常性质的点或团体，常用于检测抓取大额低频交易特征。

与有监督机器学习模型训练所依赖的标签数据相比，异常性质没有标准答案。认定异常的标准可以基于异常数据跟样本中大多数数据不太一样，异常数据在整体数据样本中占比较小等。

异常可以分为点异常、条件异常、集体异常：

a) 点异常：单个数据实例相对于其余数据被视为异常；

b) 条件异常：数据实例在特定条件下是异常的，但在其他情况下不是；

c) 集体异常：单个数据实例本身可能不是异常，但它们作为集合一起出现是异常的。

对点异常、条件异常、集体异常进行检测的方法包括：

a) 基于接近度:假设异常值距离其他数据点很远;与邻居的距离明显偏离大多数其他数据点与其邻居的距离；

b) 基于聚类:假设正常数据属于大而密集的集群,而异常值属于小集群或稀疏集群,或者不属于任何集群；

c) 基于分类：假设可以训练分类模型来描述正常数据。

常见的异常检测算法包括混合模型（mixture models）, Isolation Forest, DBSCAN等。

3、关系网络

关系网络（图计算）是一种灵活的数据模型,能很好的借助多种关联关系，建模具有分层,复杂甚至任意结构的数据集，同时，可以非常有效的查询或分析涉及数据实体之间的多层关联关系。因此可以应用于未知手法的预警，当异常关系聚集出现时，即可识别风险行为。

关系网络不需要数据符合严格的模式，能够简便地创建并动态转换数据,迁移成本低。由于图计算的多种优势和特性,除风险交易监控外被广泛用于社交关系分析、精准营销、舆情及社会化聆听、信息传播等具有丰富关系数据的场景。

7.2.3 计算量与算法要求

在模型训练环节,无监督机器学习模型的计算量要求与有监督机器学习较为类似。而在模型预测环节，与有监督机器学习模型相比，无监督机器学习模型的计算量除特征计算外，通常需要重新执行算法，如关系网络建网、聚类分析等。

与无监督聚类算法一样，图计算算法的计算量同样非常大，需要设计强大有效的并行算法，应用可靠的并行解决方案。

8 事后风险排查与处置

事后风险排查与处置是智能风控应用的重要环节，目标是排查潜在风险，控制损失敞口，迭代优化风控体系。案例处置与风险排查时需要配置相应人力资源开展调查、核实、反查、新规则提取等工作；风险特征/规则/模型开展生命周期管理，目标是防止风控模型效果衰减，确保智能风控体系效果符合预期。

8.1 案例处置与风险排查

8.1.1 基本人员架构与责任分工

案例处置相关人员须具备：

1、对规则/模型侦测发现的交易报警，具有业务理解能力；

2、对发现的可疑案例，具有调查确认能力；

3、对于案例调查落实，具有业务沟通能力；

4、对调查结果的反馈，具有分析研判能力。

按照案例处置各环节工作要求分，可设置以下岗位人员：

1、案例调查岗：负责从侦测系统中提取报警，按调查流程落实调查步骤，并确认调查结果；

2、案例分析岗：负责案例的特征分析，反馈建模人员，协助模型的调优；

3、管理及稽核岗：负责日常人员管理，操作流程制定和更新，以及调查结果的稽核工作。

8.1.2 侦测案例干预与核查（实时、准实时、批量）

侦测系统报警后，可视报警案例的风险程度进行风险决策，包括拦截阻断、挂起确认、批准通过，相应采取实时、准实时或批量方式加以干预，并落实相关核查工作。采取的上述干预方式，可通过系统自动和人工管控或核查的形式落实，取决于本机构的业务流程、系统支持程度以及特定场景要求等因素。

1、实时干预、核查是适用于业务时效性要求较高的场景，如金融交易监控、网银及App端非金融交易监控，其中核查环节多使用与客户互动模式完成（如短信验证、预设密码验证或人脸识别等）；

2、准实时干预、核查是在适度中断业务请求的前提下使用，中断时间长短由机构从业务流程和客户体验角度考虑。适用于时效性要求不高的业务，如修改信息或预设付款人等；

3、批量干预、核查对时效要求偏低，多用在客户准入环节，如贷款审批中侦测发现的部分存在疑点的案例，可根据特征或调查流程分类后，批量处理进入调查环节人工核查。

上述三种干预方式中，实时和准实时干预多为系统管控或核查方式落实，批量干预会由人工管控或核查方式落实。

8.1.3 案件还原

案例还原是对案例分析研究的过程，其目的是研判案例中是否存在作案手法上的变化，以更新、完善现有侦测系统中规则/模型的侦测效果。还原过程负责人员须具备多年案例调查经验，且了解本机构业务流程。

案例还原需从全流程角度，分析研判现有流程中可能存在的不足之处，评估流程中的管控措施是否有效、规则/模型侦测是否准确，总结案例发生原因，特别是因管控不足、侦测未识别而发生的案例。

8.1.4 案件反查

案例反查是以已确认案例为基础，使用部分案例信息（如地址、IP等）关联其他潜在案例，主要运用于识别和调查团体案例，其关联的范围需包含过去一段历史时间（视业务场景，制定关联时间长短）。

案例反查时，机构需配备具有一定经验的分析人员和相应的关联技术工具，如支持中文模糊匹配的搜索引擎工具、图形化展示的调查操作界面等，以提高案例反查的效率和效果。

8.1.5 风险案件库管理

对于日常工作中发现的风险案例，可依据案例的表现特征、是否未知风险、发生渠道以及来源等维度分类汇总，并建立风险案例库进行管理。风险案例库内容可包含对欺诈场景的还原、作案手法和过程总结、揭示风险点等内容，特别是对典型和新型案例的分析研究。风险案例库的建立是为机构内部案例查找、补充、完善提供基础，同时也为同业案例分享提供可能。

8.1.6 新规则提取

新规则的来源分为调查确认案例总结和未知案例特征分析。前者是以案例调查中确认的特征为基础，跟踪已知案例作案手法变化，形成新的规则；后者是使用不依赖于数据标签的无监督机器学习技术，侦测未知作案手法模式，发现未知案例异常特征后形成新规则，丰富规则引擎，扩展规则侦测的识别能力。两者都是对规则引擎的补充和完善，机构可视自身技术力量采取具体方式。

8.2 特征变量/规则/模型的管理

特征变量是规则和模型的重要支持，特征分析建模能力决定了规则和模型的应用广度。规则和模型存在应用生命周期，会随着时间推移效果衰减。为了确保规则、模型上线效果符合预期，需要在多个环节进行管理。

8.2.1 特征变量部署

特征变量计算需要支持多维度多种方式计算，如多业务跨渠道计算、自定义时间窗口、滑动时间窗口、短时间/长时间窗口等。

8.2.2 规则/模型上线部署

规则模型上线部署，需要支持分布式部署，具备保护机制，设置观察和应用区。

分布式部署要求包括支持一定的分布式并发处理及多模型，多规则，多版本部署。

保护机制要求当请求量激增时，为了保证模式实时响应速度，能过滤掉部分请求。

规则/模型应区分应用等级，按照观察区和应用区分类管理。观察区用于观察规则和模型的稳定性；应用区规则/模型参与风险决策；应用区与观察区应支持切换。

8.2.3 迭代更新和异常处理

由于模型会随着时间衰减，所以模型需要定期重新训练，需要考虑模型自动化更新。在应对新的业务形式风险，或者老规则模型失效的情况下，需要调优或部署新的规则模型。

如果监测到规则和模型上线后存在异常，例如触发率过高过低、无响应、处理超时等情况，可采取规则模型回退或规则模型应用等级下调等措施。

9 评价方式

各机构可根据本章所述风险指标评估智能风控的应用效果，或在对外采购风控技术提供方时利用风险指标评估其风控技术水平和能力。

9.1 核心指标分级分类

欺诈风险的核心评估指标为欺诈率，其他指标分为三个等级。一级指标包含风险覆盖率，交易报警率，准确率；二级指标包含用户打扰率，误报率，漏报率；三级指标包含查准率和召回率，PR曲线，ROC 曲线和AUC，F1分数等。一级和二级指标用于专家规则效果评估，三级指标用于评估有监督模型效果和无监督模型效果。另外，在交易及支付等时效性要求高的场景，还需专门针对性能指标进行评价，包括吞吐量和高可用性等。

对于合规风险，除电信网络诈骗风险场景外，其他风险场景下交易因是持卡人（或账户所有人）/机构本人意愿发起的非法或违规交易，对于未侦测出的风险交易，持卡人/机构实际不会投诉反馈，故较难获取此部分数据，主要通过公安、媒体等外部渠道获取重大合规性风险事件信息。故重点关注对于重大风险事件的覆盖率，同时可借鉴欺诈风险，引入用户打扰率、准确率、误报率等一级和二级指标量化评估具体风控效果。对于电信网络诈骗，因受害人是被诱使给不法

分子打款或转账，受害人察觉到被骗后会投诉反馈，故在此类风险场景下的评价指标可参照欺诈风险指标。

9.2 核心指标计算方式

9.2.1 整体指标

欺诈风险中，欺诈率是指风控系统未侦测防控的欺诈交易金额占所有（经风控系统）交易总金额的比例。它的变动能够量化通过智能风控体系控制风险的能力。

公式为：风控系统未侦测防控的欺诈交易金额/交易总金额

合规风险中，重大合规风险事件覆盖率是指风控系统侦测出的重大合规风险事件占所有重大合规风险事件的比例。

公式为：风控系统侦测出的重大合规风险事件/所有重大合规风险事件

9.2.2 一级指标

1、覆盖率

指风控系统侦测出的风险交易占所有风险交易的比例，该指标反映出风控系统能够覆盖风险交易的能力。

公式为：风控系统侦测出的风险交易笔数/所有风险交易笔数

2、报警率

指所有交易中，风控系统侦测出的所有疑似风险交易占所有（经风控系统监测）交易的比例。

公式为：风控系统侦测出的所有疑似风险交易笔数/所有交易笔数

3、准确率

指风控系统侦测出的风险交易占所有疑似风险交易的比例。

公式为：风险系统侦测出的风险交易笔数/所有疑似风险交易笔数

9.2.3 二级指标

1、打扰率

指所有发生交易的客户中，风控系统侦测出的疑似风险用户且对该用户正常的交易流程产生影响占所有（经风控系统）交易用户的比例。打扰率直观的反映了所有用户中，多少用户会被打扰。正常用户经常性被打扰会降低用户体验。

如以用户数作为统计单位，公式为：风控系统侦测出的所有疑似风险交易用户/所有交易用户。

如以卡片数量作为统计单位，公式为：风控系统侦测出的所有疑似风险交易卡片数/所有交易卡片数

一般在同样的覆盖率下准确率越高，那么打扰率就越低。建议统一覆盖率、准确率及打扰率指标统计单位，以便评估风控效果。

2、误报率

指风控系统侦测出但确认非风险交易的正常交易占所有疑似风险交易的比例。误报率=1-准确率。

如以卡片数量作为统计单位，公式为：风控系统侦测出但确认非风险交易的正常交易卡片数/所有疑似风险交易卡片数。

误报率=1-准确率。

或以交易笔数作为统计单位，公式为：风控系统侦测出但确认非风险交易的正常交易笔数/所有疑似风险交易笔数。

具体统计单位由机构根据具体情形及需求合理确定。误报率应和准确率保持同一统计单位。

因合规风险场景下，交易是持卡人（或账户所有人）/机构本人意愿发起的非法或违规交易，较难主动与持卡人/机构联系确认是否为风险交易，在此情形下准确率指标可能无法计算。

故更多关注持卡人主动来电投诉正常交易被误拦截的情况，可计算：持卡人主动来电投诉正常交易被误拦截的卡片数或交易笔数/所有疑似风险交易卡片数或交易笔数，替代对误报率和准确率指标的观察。

3、漏报率

指风控系统未侦测出的风险交易占所有风险交易的比例，所有风险交易即确认为某一风险场景的所有交易。

公式为：风控系统未侦测出的风险交易金额/所有风险交易金额。

漏报率=1-覆盖率。

9.2.4 三级指标

在欺诈风险侦测中，三级指标可以综合反映机器学习模型在准确性、打扰率等方面的效果，但是合规风险由于黑样本的稀缺性，三级指标的意义相对较差。

1、查准率和召回率（查全率）

在反欺诈机器学习模型中，通过训练机器学习模型用于欺诈交易的侦测，一般将欺诈案件作为正样本(或者黑样本)，将正常交易作为负样本(或者白样本)：

		真实类别	
		Positive	Negative
预测类别	Positive	True Positive (TP，真正)	False Positive (FP，假正)
	Negative	False Negative (FN，假负)	True Negative (TN,真负)

TP（True positive）：实际是黑样本，被识别成黑样本(识别正确)。

FP（False positive）：实际是白样本，被识别成黑样本(识别错误)。

FN（False negative）：实际是黑样本，被识别为白样本(识别错误)。

TN（True negative）：实际是白样本，被识别为白样本(识别正确)。

黑样本查准率(Precision，通常简称为查准率)：公式为TP/(TP+FP)，在业务含义中，即指标准确率。

黑样本召回率(Recall，通常简称为召回率或查全率)：公式为TP/(TP+FN)，在业务含义中，即指标覆盖率。

2、PR曲线

PR值是相互矛盾的指标，分别代指准确率(Precision)和覆盖率(Recall)。准确率越高则覆盖率相对会低，同样覆盖率提高，准确率相对会降低。机器学习模型判断样本是否为黑样本会输出一个概率值，若该概率值比设置的阈值大，则

将其判断为黑样本。因此，不同的阈值下，模型判断为黑白样本的数量不同，故也会对应不同的PR值。PR曲线就是不同阈值下不同PR值的点所组成的曲线。

根据实际业务场景，机构可以自由选取相对应的准确率和覆盖率：当市场风险较大时，可使用比较严格的策略，即降低阈值来提升覆盖率；当市场风险较小时，可使用比较宽松的策略，即增加阈值来提升准确率。

3、ROC 曲线和AUC

根据每个测试样本属于黑样本的概率值从大到小排序,当测试样本属于黑样本的概率大于或等于这个阈值时，模型将其预测为黑样本，否则为白样本。每次选取一个不同的阈值，就可以得到一组FPR（公式为FP/(FP+TN)，即白样本误识率）和TPR（公式为TP/(TP+FN)，即覆盖率）。所有(FPR,TPR)对连接起来，就得到了ROC曲线，因此ROC曲线越靠拢(0,1)点，越偏离45度对角线，代表误识率低、覆盖率高，分类效果越好。

AUC（Area Under Curve）被定义为ROC曲线下的面积。AUC值代表当前的分类算法将正样本排在负样本前面的概率，因此AUC值越大分类效果越好，可以用于直观评价分类效果。

4、F1分数

F1分数，又称平衡F分数，它综合了准确率和召回率的指标,定义为准确率和召回率的调和平均数，可用于比较不同模型。

F1=1/(1/Precison+1/Recall)

9.2.5 性能指标

在高时效性要求的风控场景中，性能是智能风控系统及其重要的考量指标，性能不足直接会导致风控系统的不可用。

1、吞吐量

风控系统处理能力，一般使用TPS、并发数、响应时间等作为参考。通常根据业务峰值确定TPS数值及并发数，响应时间需考虑端到端（从业务系统经过一系列网络节点后到达风控系统，风控系统处理完后返回到业务系统）的总时间。

2、高可用

风控系统直接实时对接业务系统，对在办业务有着明显影响。风控系统需要具备高可用性，风控系统的故障直接影响业务的进行，需确保部分组件/服务器节点出现故障，整体系统依然稳定运行，或者及时熔断。

9.3 专家规则监控效果评估

对于欺诈风险，衡量风控效果最关键的指标是欺诈率，能最直观的评估风控的整体效果。规则准确率和覆盖率也是规则监控的核心指标，两个指标相互约束，呈负相关，同时在不同场景下专家规则应用的准确率和覆盖率也存在差异。对于合规风险，衡量风控效果最关键的指标是重大合规风险事件覆盖率。在此基础上，跟踪观测用户打扰率、准确率、误报率等一级和二级指标波动情况。各机构需要根据业务场景、风控业务能力、风控运营支持情况，制定最合适的规则准确率和覆盖率指标。

9.4 有监督机器学习模型效果评估

有监督模型，可以使用查准率、召回率（覆盖率）、F1、AUC等做为评估指标，并且通过ROC曲线和PR曲线辅助判断。

查准率和召回率是负相关的指标，一定情况下互相制约，机构需要根据业务风险策略进行动态调整。

F1指标和AUC是综合类指标，综合体现了查准率和召回率的两方面的因素，一般来说F1指标和AUC指标越高说明模型的区分能力越强，机构可基于综合类指标，在数据准备、数据清洗、特征提取和参数调优等环节不断优化模型以达到更好防控效果。

9.5 无监督机器学习模型效果评估

无监督模型以常见的聚类分析为例，可以使用轮廓系数等指标来评估模型效果，轮廓系数越大代表模型效果越好，即组内的样本更为相似，组间的样本差异更大，原因是大部分的正常支付行为会聚集在一起，小部分的异常支付行为同样

会聚集在一起。如需定量计算三级指标，由于事前没有标签，需要人工介入区分正负样本后再加以计算。

10　安全实施与技术性能要求

智能风控应用应在合法依规前提下采集、保存、使用和处理个人数据，满足相应安全要求，同时风控系统运营性能，例如交易处理能力、响应时效等，也应与机构风险防控需求相匹配，确保运营流程顺畅。

10.1　个人数据采集要求

机构出于风控数据采集需求等各方面的原因，可能会据此获取用户敏感信息，机构应履行个人信息安全保护义务，采取必要措施，保障个人信息安全。

当个人信息主体同意收集某类型的信息时，机构应遵循最小必要原则，且不应在征得个人信息主体授权同意前，产生个人信息收集行为。

支付标记（Payment Token，简称“Token”），是由13至19位数字组成，且符合账号的基本验证规则的数据串，可作为银行卡主账号（PAN）的一个替代值，用于完成特定场景支付的交易。作为一项国际通用性安全技术，支付标记化有效降低了信息泄露风险：支付标记化方案采用支付标记替代了实际卡号，根本上杜绝了卡号信息泄露的可能；另外，在支付标记产生时对标记应用的范围进行了限定，进一步降低了支付标记泄露后的影响范围。

故建议在个人信息采集和交易处理过程中，通过支付标记化技术对于卡号等数据进行替代，并限定支付标记应用范围，从源头上控制信息泄露风险。

10.1.1 敏感信息采集的合法性要求

个人信息采集要求：

1、不得欺诈、诱骗、强迫个人信息主体提供其个人信息；

2、不得隐瞒产品或服务所具有的收集个人信息的功能；

3、不得从非法渠道获取个人信息；

4、不得收集法律法规明令禁止收集的个人信息。

10.1.2 敏感信息采集时的授权同意

收集个人信息前，应向个人信息主体明确告知所提供产品或服务的不同业务功能，分别收集的个人信息类型，以及收集、使用个人信息的规则，如收集目的、存储期限等，并获得个人信息主体的授权同意。

信息采集分两种类型，以APP信息采集为例，一般分为APP自有信息采集方式和嵌入SDK方式进行采集。当用户授权APP采集数据时，该APP集成的SDK则默认获取与APP一致的用户授权。SDK提供方有责任和义务向宿主APP提供SDK所需权限列表及各项权限的使用原因、是否是必要权限。宿主APP有权利和义务检测SDK是否有越权采集用户数据行为。SDK提供方有涉及专利技术、计划申请专利技术、保密技术等核心技术有不告知宿主APP的权利。

10.1.3 无需征得授权同意的例外情形

以下情形中，收集、使用个人信息无需征得个人信息主体的授权同意：

1、与履行法律法规及行业主管部门规定的义务相关的；

2、与国家安全、国防安全直接相关的；

3、与公共安全、公共卫生、重大公共利益直接相关的；

4、与犯罪侦查、起诉、审判和判决执行等直接相关的；

5、出于维护个人信息主体或其他主体的生命、财产等重大合法权益但又很难得到本人同意的；

6、个人金融信息主体自行向社会公众公开的；7、从合法公开披露的信息中收集个人信息的，如合法的新闻报道、政府信息公开等渠道；

8、根据个人信息主体要求签订和履行合同所必需的；

9、用于维护所提供的金融产品或服务的安全稳定运行所必需的，例如识别、处置金融产品或服务中的欺诈或被盗用等。

10.1.4 敏感信息采集的明示同意

采集敏感信息时，应确保个人信息主体在其知情的基础上明确表示同意，如APP弹窗确认等方式。

通过主动提供或自动采集方式收集敏感信息前，应向个人信息主体告知所提供产品或服务的核心业务功能及所必需收集的个人敏感信息，并明确告知拒绝提

供或拒绝同意将带来的影响，应允许个人信息主体选择是否提供或同意自动采集。

产品或服务如提供其他附加功能，需要收集个人敏感信息时，收集前应向个人信息主体逐一说明个人敏感信息为完成何种附加功能所必需，并允许个人信息主体逐项选择是否提供或同意自动采集个人敏感信息。当个人信息主体拒绝时，可不提供相应的附加功能，但不应以此为理由停止提供核心业务功能，并应保障相应的服务质量。

10.1.5 隐私政策的内容和发布

个人隐私信息采集方需明示发布以下信息：

1、采集、使用个人信息的目的，以及目的所涵盖的各个业务功能；

2、各业务功能分别采集的隐私信息，采集方式、频率、存储期限等个人信息处理规则，实际采集的个人信息范围；

3、对外共享、转让、公开披露个人信息的目的，涉及的个人信息类型，接收个人信息的第三方类型，以及所承担的相应法律责任；

4、遵循个人信息安全基本原则，具备的数据安全能力，以及采取的个人信息安全保护措施；

5、隐私政策所告知的信息应真实、准确、完整。

10.2 个人信息保存

机构应采取必要措施，保障个人信息存储安全，合理利用必要的个人信息。个人信息在保存时需要限定保存时间长度，并进行标记化处理；在展示和使用时，需要控制展示和使用范围。

10.2.1 个人信息保存时间

个人信息保存应遵循最小化原则，对个人信息控制者的要求包括：个人信息保存期限应为实现目的所必需的最短时间，超出上述个人信息保存期限后，应对个人信息进行删除或匿名化处理。如用户本人或监护人提出要求删除隐私数据，存储方需要在合理时限内清除隐私数据。

10.2.2 去标识化处理

收集个人信息后，个人信息控制者应立即进行去标识化处理，并采用技术和管理方面的措施，将去标识化后的数据与可用于恢复识别个人的信息分开存储，确保个人信息无法还原。去标识化建立在个体基础上，保留了个体颗粒度，采用假名、加密、哈希函数等技术手段替代标识个人特征信息。

10.2.3 个人敏感信息的传输和存储

传输和存储个人敏感信息时，应采用加密等安全措施。存储个人生物识别信息时，应采用技术措施处理后再进行存储。

10.3 个人信息使用

10.3.1 访问控制

对个人信息使用的访问控制包括：

1、对被授权访问个人信息的内部数据操作人员，应按照最小授权原则，设置其访问职责所需的最小数据操作权限；

2、应对个人信息的重要操作设置内部审批流程，如批量修改、拷贝、下载等；

3、应对安全管理人员、数据操作人员、审计人员的角色进行分离设置；

4、如确因工作需要，需授权特定人员超权限处理个人信息，应由个人信息保护责任人或个人信息保护工作机构进行审批，并记录在册。

5、对个人敏感信息的访问、修改等行为，应在对角色权限控制的基础上根据业务流程需求触发操作授权。

10.3.2 展示控制

对于通过界面展示个人信息的情形，机构应对需展示的个人信息采取去标识化处理等措施，降低个人信息在展示环节的泄露风险。

10.3.3 使用控制

对于个人信息使用的要求包括：

1、除必需情形外，使用个人信息时应消除身份指向性，避免精确定位到特定个人。

2、当所收集的个人信息进行加工处理后，能够单独或与其他信息结合识别自然人个人身份或者反映自然人个人活动情况，应将其认定为个人信息。处理时应遵循处理个人信息时获得的授权同意范围。

3、使用个人信息时，不得超出与收集个人信息时所声称的目的具有直接或合理关联的范围。因业务需要，确需超出上述范围使用个人信息的情形，应再征得个人信息主体的明示同意。

10.4 数据处理安全要求

10.4.1 数据存储安全

App数据存储安全：APP及SDK采集的用户相关数据需要做高强度加密后方可保存在APP本地环境中。加密算法需要高于或等于AES256或SM4。加密密钥在APP运行期间不可以明文的方式存在于磁盘、内存等位置。

对于客户的支付敏感信息（有效期、cvn2等）不得在APP本地环境中存储，使用结束后即需清除。

云端数据存储安全：当云端存储用户数据时，对外需要保证数据库访问隔离，不可外网访问；对内需要添加数据库用户密码等访问限权。

10.4.2 数据传输安全

采用自有协议进行数据传输时，需要满足的安全点包括：验证服务端真实有效、防中间人攻击、发送的数据内容应加密并且对加密数据签名、保证前向安全性。

采用通用协议进行数据传输时，禁用http协议，可以选用https（TLS1.2及以上版本）、http2、http3、grpc等协议。

10.4.3 数据展示脱敏

数据脱敏不仅要执行数据漂白，抹去数据中的敏感内容，同时也需要保持原有的数据特征、业务规则和数据关联性，保证开发、测试、培训以及大数据类业务不会受到脱敏的影响，达成脱敏前后的数据一致性和有效性。

对外展示的过程中需要将用户隐私类数据分级脱敏展示:一级数据包括真实姓名、年龄、电话、身份证号码、住址、即时通信号码、邮箱、聊天记录等；二级数据包括联系人及联系人相关信息、账户ID、订单信息、消费信息、收货地址、地理位置信息（高精度、低精度）、转账记录；三级数据包括用户登录时间、用户APP/WEB操作行为、用户传感器信息、业务日志等。以上未涉及信息则划归二级敏感信息。

1、一级敏感数据需要做不可逆脱敏之后方可提供给内部员工做批量数据查询。系统进入生产阶段可使用原始数据，但是一级敏感数据不可以以日志方式打印、不可发送到其他系统等。脱敏方式如 hmac-sha1、hmac-SHA256等，不可使用md4、md5 等不可逆的脱敏方式。严禁使用base64/urlencode/hexencode等编码方式替代脱敏。

2、二级敏感数据需要做不可逆或加密等级AES256以上的加密之后，方可提供给内部人员做数据查询。例如隐藏部分关键数据，账户id 9001****8934。

3、三级敏感数据可由业务方自行保管，可用于数据分析、生产、对外宣传等。

10.4.4 数据保留时间

机构可根据自身需求选择长期保留用户设备类信息，但信息保存应遵循最小化原则。

10.5 运营性能要求

数据采集：SDK采集需要在200ms内完成数据采集，在60s内完成数据发送与接收。

实时监控与拦截：对于风险交易会立即带来损失的情况下，需要实时返回处理结果，通常时效要求是50ms以内，且交易处理能力要与机构业务风险防控需求相匹配，对资源配置、特征计算方法、模型准确性的要求较高。

准实时监控：在延迟决策或决策时效性要求不太高的情况下，响应时效要求可略微放低至秒级或分钟级，对资源配置要求也相对降低。

批量监控：定期触发，适用于在较长周期内进行回溯排查，对资源的配置要求不高。