ICS 35.240.40
CCS A 11

JR

中华人民共和国金融行业标准

JR/T 0202—2020

基于大数据的支付风险智能防控技术规范

Big data based intelligent payment risk control technical specification

2020-12-03 发布　　2020-12-03 实施

中国人民银行　发布

目　　次

前　言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由中国人民银行提出。

本文件由全国金融标准化技术委员会（SAC/TC 180）归口。

引　言

随着大数据、移动互联、人工智能、生物特征识别等技术的快速发展，支付方式正在发生着巨大而深刻的变革，新技术在丰富支付手段、提高支付效率的同时，带来了新的隐患，也对从业机构的支付风险防控能力提出了更高的要求。

为规范大数据与人工智能技术在支付风险防控领域的应用，提高支付风险防控技术的针对性和有效性，切实保障人民群众信息和资金安全，编制本文件。

基于大数据的支付风险智能防控技术规范

1 范围

本文件规定了基于大数据、人工智能等技术开展支付风险防控所需的技术框架和系统实现的安全要求。

本文件适用于与支付相关的商业银行、非银行支付机构和清算机构等开展支付风险防控体系建设、运用智能防控技术搭建风险智能防控系统、提供支付风险防控服务等工作。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

JR/T 0071—2020 金融行业网络安全等级保护实施指引

JR/T 0171—2020 个人金融信息保护技术规范

3 术语和定义

下列术语和定义适用于本文件。

3.1

风险 risk

不确定性对目标的影响。

注：1.影响是指偏离预期，可以是正面的和/或负面的。

2.目标可以是不同方面（如财务、健康与安全、环境等）和层面（如战略、组织、项目、产品和过程等）的目标。

3.通常用潜在事件、后果或者两者的组合来区分风险。

4.通常用事件后果（包括情形的变化）和事件发生可能性的组合来表示风险。

5.不确定性是指对事件及其后果或可能性的信息缺失或了解片面的状态。

[来源：GB/T 23694—2013,2.1]

3.2

支付风险 payment risk

从业机构在开展支付业务时所面临的各类风险。

3.3

风险智能防控 intelligent risk control

通过采用大数据和人工智能等相关技术提升对风险的识别、评估和应对等能力的一种风险防控方式。

3.4

大数据 big data

具有体量巨大、来源多样、生成极快且多变等特征，并且难以用传统数据体系结构有效处理的包含大量数据集的数据。

注：国际上，大数据的4个特征普遍不加修饰地直接用volume、variety、velocity和variability予以表达，并分别赋予其大数据语境下的定义：

a）体量 volume：构成大数据的数据集的规模。

b）多样性 variety：数据可能来自多个数据仓库、数据领域或多种数据类型。

c）速度 velocity：单位时间的数据流量。

d）多变性 variability：大数据其他特征，即体量、速度和多样性等特征都处于多变状态。

[来源：GB/T 35295—2017,2.1.1]

3.5

大数据技术 big data technology

对大数据进行采集、处理、存储、分析、挖掘、管理，从中发现新知识、创造新价值、提升新能力的新一代信息技术。

3.6

机器学习 machine learning

在历史数据中自动发现规律并利用规律对未知数据进行应用（预测）的算法（技术）。

3.7

监督学习 supervised learning

利用已标记的有限训练数据集，通过某种学习策略、方法建立模型，实现对新数据或实例的标记（分类）或映射。

注：最典型的监督学习算法包括回归和分类。

3.8

半监督学习 semi-supervised learning

在训练过程中利用小部分的标记数据，以及大部分的非标记数据进行训练学习，介于监督学习（3.7）和无监督学习（3.9）之间的1种学习方法。

3.9

无监督学习 unsupervised learning

利用无标记的有限数据描述隐藏在未标记数据中的结构和规律。

注：最典型的无监督学习算法包括单类密度估计、单类数据降维、聚类等。

3.10

设备指纹 device fingerprint

可以用于唯一标识出该设备的设备特征或者独特的设备标识。

3.11

流处理 stream processing

针对处理高并发且对时效性有较高要求的大规模计算场景，能够对具有实时、高速、无边界、瞬时性等特性的流式数据进行实时处理的技术。

注：流处理具备低时延、高可用、高扩展等特性。

3.12

图计算 graph processing

以“图论”为基础的对数据的1种“图”结构的抽象表达，以及在这种数据结构上的计算模式。

注：在图计算中，基本的数据结构表达包括：节点、边、权重等。

3.13

内存计算 in-memory processing

优先使用内存对数据进行存储、计算、分析的1种数据处理技术。

3.14

批处理 batch processing

将一个大型作业分解为多个任务，交由多个节点分别处理，并将分解后多个任务处理的结果进行汇总，得出最终分析结果的计算框架。

注：批处理具备高可用、高扩展、高并发等特性。

4 技术框架

4.1 概述

基于大数据的支付风险智能防控技术，主要包含大数据技术、风险防控技术、风险类型3部分内容。通过大数据技术利用在线和离线方式分析数据，为智能化的风险防控技术提供技术支撑，使得机构可利用人工智能等技术不断迭代风控模型，主动识别和防控支付业务过程中的风险，框架见图1（图中虚线框表示具体的分类和方法由业务单位根据实际进行划分）：

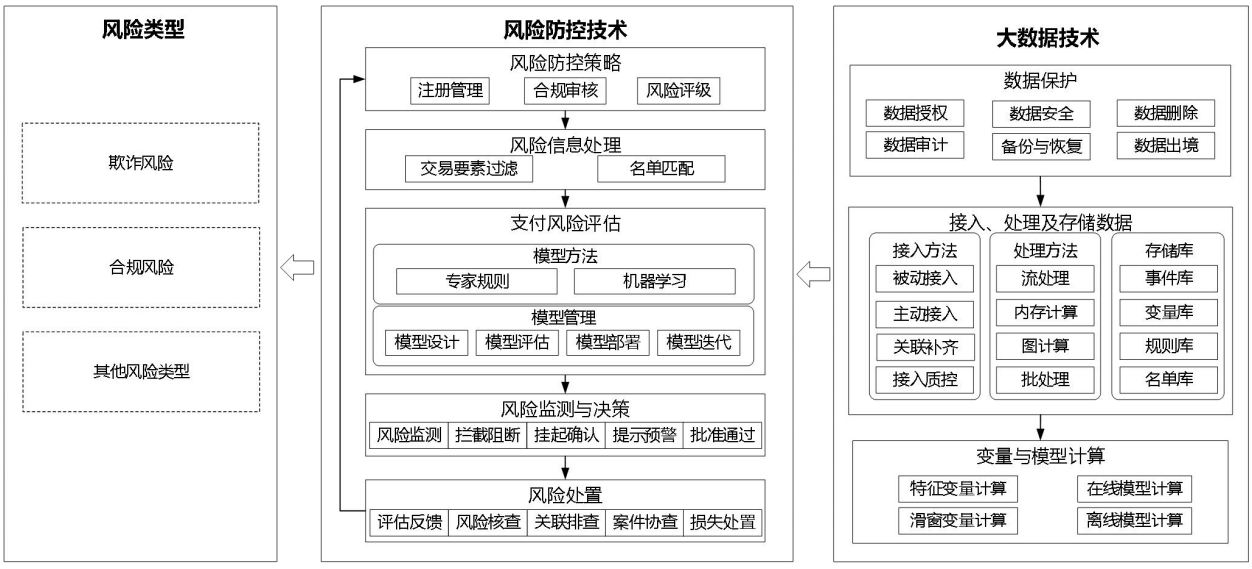

图1 基于大数据的支付风险智能防控技术框架

4.2 风险类型

4.2.1 概述

根据支付产业中存在的各种风险的特点，本文件将支付风险划分为欺诈风险、合规风险和其他风险等类型，同时也存在多种风险交织并存的情况。

4.2.2 欺诈风险

欺诈风险指不法分子利用虚假申请、伪造或变造银行卡、盗用账户等手段盗取银行卡（或账户）交易资金的风险，或者不法分子勾结持卡人通过虚构交易等方式，造成发卡银行或第三方机构资金、权益等方面损失的风险，包括但不限于：

a) 伪卡欺诈，包括芯片交易方式伪卡欺诈、降级使用交易方式伪卡欺诈和磁条交易方式伪卡欺诈，具体内容如下：
——芯片交易方式伪卡欺诈是指伪卡交易通过芯片交易方式完成。
——降级使用交易方式伪卡欺诈（Fall Back）是指伪卡交易通过 Fall Back 降级交易完成。
——磁条交易方式伪卡欺诈是指伪卡交易通过磁条交易完成。
b) 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，包括丢失卡欺诈与被盗卡欺诈两种情形。
c) 非面欺诈，指欺诈分子窃取或骗取卡片主账号、有效期、支付短信验证码及其他关键身份验证信息后，通过邮购、互联网、手机等非面对面渠道进行欺诈冒用。
d) 账户盗用，指欺诈分子冒充真实持卡人或者账户所有人的身份，通过修改账单地址、虚假挂失等一系列手段获取重制卡片或者账户信息进行的欺诈交易。
e) 伪冒申请，又称虚假申请，指使用虚假身份或冒用他人身份申领银行卡（或开立账户）完成欺诈交易。
f) 商户合谋，指特约商户在受理支付交易时，违规操作、蓄意进行欺诈交易或纵容、包庇、协助持卡人开展欺诈交易的行为。
g) 营销欺诈，指不法分子利用营销主办方的营销漏洞，与商家勾结、虚构交易，骗取营销活动主办机构的营销费用，获得不正当收益。
h) 套现风险，指持卡人未通过正常合法手续（如 ATM 或柜台等）提取现金，与商户或其他第三方合谋方式，以虚构交易、虚开价格、现金退货等手段将账户中信用额度内的资金以现金的方式套取，同时又不支付银行提现费用的行为。

4.2.3 合规风险

合规风险指银行、非银行支付机构、特约商户及第三方专业化服务机构等支付业务参与方因未能遵循法律法规、监管要求、业务规则及内部规范等，可能遭受法律制裁、监管处罚、违规约束进而引发财务或声誉损失的风险，包括但不限于：

a) 洗钱风险，指将通过各种手段掩饰违法所得，隐瞒违法来源，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。
b) 电信诈骗，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汇款、转账、购物或代付等的犯罪行为，从而给受害人造成资金和权益方面的损失。
c) 挪用客户备付金风险，指不法分子通过网络攻击，编造虚假交易和信息，伪冒商户等方式挪用、占用、借用客户备付金而造成持卡人或账户所有人、商户和机构资金和声誉损失的风险。
d) 非法集资，指单位或者个人未依照法定程序经有关部门批准，以发行股票、债券、彩票、投资基

金证券或者其他债权凭证的方式向社会公众筹集资金，并承诺在一定期限内以货币、实物以及其他方式向出资人还本付息或给予回报的行为。

4.2.4 其他风险类型

除了欺诈风险和合规风险之外，不同机构、不同业务场景可能存在其他的风险类型，如资金清算风险、用户道德风险等。

4.3 风险防控技术

4.3.1 概述

风险防控技术是基于大数据的支付风险智能防控技术框架的核心组成部分，包含风险防控策略、风险信息处理、支付风险评估、风险监测与决策、风险处置等五个模块。一是通过大数据、机器学习等技术建立满足要求的风控模型，进一步加强对风险的事前预测和事中识别的能力。二是通过合理引入多个模型、强鲁棒性的模型、抗AI攻击的模型等方式，提高模型评分的稳定性。三是支持多渠道、多维度的数据整合，形成机构内统一的风控系统。

风险防控策略作为风险防控的第一道屏障，通过注册管理、合规审核、风险评级等方面控制，对潜在风险进行初步分辨。风险信息处理包含交易要素过滤和名单匹配等，将过滤所得信息输出到支付风险评估模型中。支付风险评估从模型方法、模型管理等方面设计模型，完成对潜在风险的识别、分析和评价。根据风险模型的计算结果，结合业务要求，采取阻断、挂起、预警、批准等不同的决策行为。最后，在决策的基础上，开展风险调查、关联排查、案件协查等，其结果可以优化风险防控策略。

4.3.2 风险防控策略

风险防控策略主要是指用户在注册等环节实行的审核与风险防控措施，主要功能应至少包含注册管理、合规审核、用户风险评级等方面。机构应根据自身的业务场景、风险类型、风险防控需求等，对风险强度进行级别划分，并实施相应的风险防控策略，具体要求如下：

a) 注册管理方面应满足以下要求：
   ——建立完备的注册管理机制，在用户注册阶段，对其身份进行核验和管理。
   ——用户提交必要的身份核验资料，企业用户提交法人身份信息、企业经营信息（如企业名称、负责人、联系方式、所属行业、经营状况等）资料，个人用户提交身份信息、联系方式等资料。
   ——对用户所提交资料的有效性、完整性、真实性进行审核。
   ——采用多因子的验证方式对用户身份进行核验。
   ——对用户提交的信息进行定期的复审和更新。
b) 合规审核方面应满足以下要求：
   ——制定审核机制，对用户的业务目的、业务性质以及交易来源等业务信息进行验证。
   ——识别与确认用户业务权限。
   ——判断并标记用户（信息或行为）的类别、级别、名单归属（如是否在黑名单内），并根据标记对当前用户请求进行响应（通过、拒绝、关注等）。
   ——定期审核和更新用户信息，及时发现潜在风险。
c) 用户风险评级方面：
   ——评级指标包括但不限于：
   - 交易属性；
   - 经营信息；

- 资金往来；
- 信用状况；
- 设备指纹。

——评级规则：根据评级指标，设计相应的评级规则。

——评级频次：定期或者不定期评级。

4.3.3 风险信息处理

风险信息处理是根据具体的业务场景采集数据要素用于风控模型计算，至少包括交易要素过滤和名单匹配两个环节，具体包括：

a) 交易要素过滤方面，具体要求如下：

——应根据业务场景识别所面临的风险类型，业务场景包括但不限于注册、登录、支付等。

——应根据风险类型确定过滤的交易要素，包括要素名称、要素格式、要素条件等。

——交易要素包括但不限于交易信息、账户信息、设备信息、交易方信息，具体要求为：

- 交易信息包括但不限于卡号（账号）、手机号码、交易时间、交易金额、交易地区等；
- 账户信息包括但不限于账户开立时间、账户可用额度等；
- 设备信息包括但不限于 IP 地址、设备指纹、经纬度信息等；
- 交易方信息包括但不限于用户 ID、名称信息等。

——应定义唯一的主键索引每笔交易。

——宜满足支付风险评估模型计算的要求。

——宜在过滤信息时进行衍生变量的计算，如根据手机号码计算归属地，根据经纬度信息计算所在地。

b) 名单匹配方面，具体要求如下：

——名单包括但不限于卡号（账号）、手机号码、设备指纹、IP 地址等。

——应匹配公安、司法机关公布的具有明确业务含义的名单。

——宜匹配自有的黑、灰、白名单库，对交易真实性和合法性进行初步识别。

——可匹配外部第三方平台的黑、灰、白名单库，实现风险联防联控。

4.3.4 支付风险评估

4.3.4.1 概述

支付风险评估是 1 种对支付业务中的风险进行分析、识别和评价的技术手段，主要包括模型方法和模型管理两个方面。

4.3.4.2 模型方法

模型方法指基于黑样本等已知风险和历史数据表现，根据机构自身的风险偏好，在各业务场景及环节中，将数据变量通过运算逻辑关系自由组合，设置实时、准实时、批量、验证类等模型和规则，以实现对风险的识别判断，包括专家规则、机器学习等。

a) 专家规则应符合以下要求：

——专家规则包含但不限于以下信息：

- 商户（用户）身份信息；
- 操作设备信息；
- 地址位置信息；
- 交易信息；

- 营销活动信息。

——形成文档保存或在风控后台中记录，保留每次规则设置和更新时间及内容。

——依据业务情况设置查看、新增、修改、删除、复核等管理权限。

——在新增和修改前进行充分有效的评估。

——宜有量化评价规则效能的指标，如触发率、准确率、覆盖率等。

b) 机器学习主要分为监督学习、半监督学习、无监督学习（参见附录），针对不同的风险类型，可采用多步骤式建模方法，挖掘并提取海量特征对事件或者主体进行描述，也可采用端到端建模方法，直接基于数据构建模型，具体要求如下：

——监督学习宜符合以下要求：

- 有足够数量的高风险、低（无）风险样本；
- 有样本选取、特征计算、训练、预测和效果评价等必备的环节；
- 有特征列表及模型可解释性指标，如特征重要度列表等（端到端建模除外）；
- 有量化评价效果的数据指标，如精确率、召回率等。

——半监督学习宜符合以下要求：

- 有一定数量的高风险样本和更多的未标记样本；
- 计算效率能够满足业务的性能要求；
- 有量化评价效果的数据指标，如准确率等。

——无监督学习宜符合以下要求：

- 有一定数量的数据集；
- 计算效率能够满足业务的性能要求；
- 有量化评价效果的数据指标，如准确率等。

此外，关系网络也是机器学习常用的模型方法。关系网络是以图论为基础，旨在描述真实世界中存在的各种实体或概念及其相互之间的关系，并以此构成关系网络图，它提供更好地组织、管理和理解海量信息的能力（参见附录），具体要求如下：

——宜有关系网络的可视化展现。

——宜有量化评价效果的数据指标，如准确率等。

4.3.4.3 模型管理

模型管理是指对模型的全生命周期进行管理，包括模型设计、模型评估、模型部署和模型迭代 4 个阶段，具体要求如下：

a) 模型设计是基于不同的业务场景和风险类型，结合实际需求选择合适的变量和模型方法，包括但不限于应用特征工程、端到端建模等方法清洗和加工数据，对样本组成的训练集和测试集用各类模型算法进行训练，基于模型指标选择最优模型等，具体要求如下：

——应有模型设计文档。

——宜有模型训练或模拟过程。

b) 模型评估是指对于模型设计产生的模型进行效果和安全等方面的评估，以作为选择模型或者调整模型设计的依据，具体要求如下：

——宜有模型评估流程。

——宜根据业务制定模型评估方法。

——宜对模型结果制定量化评价指标。

c) 模型部署是将已完成训练或模拟的模型结果上线部署到生产环境中，用于识别各类风险，具体要求如下：

——应建立模型部署审批流程（含业务评审），遵循规范步骤。

——应在部署前完成模型性能测试。
——应在模型部署后对模型效果进行监控。

d) 模型迭代是针对生产环境中表现不佳或出现效能衰退的模型进行调整优化的过程，具体要求如下：
——模型上线后应定期观测模型表现。
——应对已衰退的或存在明显问题的模型进行及时调优和更新。
——宜有模型迭代上线流程。
——宜有模型迭代文档和更新记录。

4.3.5 风险监测与决策

风险监测是对各种已识别或关注的风险以及整体风险情况，表征风险的指标等进行监控和测算。

风险决策是指对当前业务经过风险信息处理和风险评估处理后，判断其是否具有风险。根据自身业务要求及风险分析结果，按照不同的风险级别采取不同的决策行为，主要包括拦截阻断、挂起确认、提示预警、批准通过等。

a) 拦截阻断将使得当笔交易失败，具体要求如下：
——交易授权系统宜支持对交易进行拦截阻断。
——宜在阻断拦截后开展事后调查分析，评估确认是否为欺诈交易，必要时与持卡人或账户所有人核实确认是否为欺诈交易，具体要求如下：
- 如确认欺诈，可采取进一步措施控制风险，如加入卡号（账户）黑名单、限制 IP、向公安或司法机关报送欺诈等；
- 如确认非欺诈，宜重新批准通过，并加入白名单。

b) 挂起确认将延长当笔交易的授权时间，具体要求如下：
——宜在满足用户体验的情况下，对中等级的欺诈风险交易执行挂起确认。
——交易挂起后，宜对用户身份进行二次验证，验证通过可批注交易，否则宜阻断交易。
——二次验证包括但不限于短信、电话、网页、APP、生物识别等一种或多种方式向用户进行提示。

c) 提示预警不影响当笔交易的授权，具体要求如下：
——应至少对低等级的欺诈风险交易执行提示预警。
——应对合规风险交易执行提示预警。
——应对无风险交易执行批准通过。
——对于欺诈风险交易，宜在提示预警后开展事后调查分析，必要时与持卡人或账户所有人核实确认是否为欺诈交易。
——对于合规风险交易，宜在提示预警后开展调查分析。
——提示预警可包括准实时和批量两种方式。

除上述情形之外，在出现人为误操作、突发事件引起规则模型异常时，应由系统或人工监控上述异常，必要时采取人工干预或通过风险控制系统熔断机制等方式，避免因风控系统的决策失控造成大面积误杀等错误处置。

4.3.6 风险处置

风险处置是在风险决策结束后进行的评估反馈、风险核查、关联排查、案件协查和损失处置的相关后续活动。风险处置的结果旨在完善现有风险防控的策略、风险信息处理的内容与支付风险评估的能力，实现风险防控流程的闭环反馈优化。

a) 评估反馈。

评估反馈是指对于风险监测和决策输出的结果进一步进行复核、分析和反馈，以确保风险监测和决策的系统、流程、操作正确，并初步确认风险评估和决策结果是否合适，以作为后续相关系统、流程、风险评估模型和风险决策改进的依据，具体要求如下：

——宜确认相关系统正常工作、相关操作和流程符合要求。

——宜对风险评估和决策的结果进行初步分析和确认，并将分析和确认的结果向风险评估、风险监测和决策进行反馈。

b） 风险核查。

风险核查指在评估反馈的基础上，对于已识别有风险的业务进行调查，分析原因与风险特征，以确认当时的决策是否准确恰当。核实无误的，宜将相关信息录入黑名单，作为后续风险决策的依据，核实确认为无风险的，宜作为后续模型优化和风险处置的依据。风险核查包括但不限于以下方式开展：

——对于拦截阻断的交易，具体要求如下：

- 应配套后续调查流程完善防控手段；
- 宜与业务方进行确认，判断拦截阻断的准确性；
- 经调查拦截无误的，相关信息应纳入黑灰名单和负面样本，作为后续优化事中监测依据；
- 经调查拦截不准确的，宜恢复交易权限，及时调整事中监测策略。

——对于挂起确认和提示预警的交易，具体要求如下：

- 应配套调查处置流程，并借鉴简化后续类似情况下的处理流程；
- 宜事后统计分析存在的可疑点，集中与业务方沟通确认，回溯挂起确认和提示预警的必要性与准确性，并判断下一次类似条件的业务风险处理方式。

——对于批准通过的交易，具体要求如下：

- 如发生用户投诉，应配套相应的处置流程进行风险分析和处置，并将此作为后续优化风险监测，完善相关规则和模型的依据；
- 如未发生用户投诉，但通过关联排查能够识别的可疑交易，应和业务方沟通确认疑点，条件允许的情况下宜与用户进行沟通，进一步确认或排除风险，并将其中确认的风险交易录入为负面样本。

c） 关联排查。

关联排查指对于有风险的业务相关元素，基于潜在关系进行关联分析，以挖掘是否存在同类风险或衍生风险，弥补事中监测决策可能未识别的潜在风险敞口。关联排查包括但不限于以下方式开展：

——应对存在风险交易的同卡片或账户关联交易进行分析。

——应对存在信息泄露风险的商户在一段时间内有交易的卡片或账户进行分析。

——宜对存在虚假申请风险的卡片或账户关联的设备信息进行分析。

——宜对存在风险交易的卡片或账户的位置信息进行分析，或者对存在风险交易的持卡人或账户所有人的位置信息进行分析。

——宜对存在风险交易的手机号码进行分析，包括验证手机号码、注册手机号码等。

d） 案件协查。

案件协查主要是指配合公安、司法机关开展的风险案件协查，包括但不限于以下方式开展：

——应提供必要的交易明细。

——应提供必要的商户开立获批和持卡人或账户所有人开户获批的相关信息；根据公安、司法机关的指令冻结账户和资金。

——宜提供已采集的交易信息、账户信息以外的风险案件行为特征，例如 IP、MAC 等。

e） 损失处置。

损失处置主要指对于明确产生的风险损失，通过快速挽损、风险责任认定，将风险化解、转移或者赔偿的处置方式，并控制后续风险损失敞口。

——损失处置遵循的主要原则如下：

- 应制定明确的风险责任认定标准；
- 应按照监管部门要求及时通报风险损失情况；
- 宜保存风险损失的电子资料；
- 宜保护持卡人或账户所有人的正当权益。

——损失处置采取的主要方法包括但不限于：

- 延迟结算：收单机构针对商户合谋支付风险，快速采取结算资金延迟到账方式挽回损失；
- 货物拦截：针对互联网渠道实物类商品销售付款与收货存在较长时间的特性，在风险识别后及时控制在途货物，采取退款措施，控制风险损失敞口；
- 追偿结算：通过事后损失追偿，转移化解已有风险损失；
- 限制功能：针对识别的具有高风险特征的交易行为，对相关银行卡账户、支付账户、商户终端采取限制交易权限措施；
- 关闭通道：关闭支付通道，防范新增损失；
- 保险赔付：通过事先投保、事后理赔方式，分散化解风险损失；
- 法律诉讼：通过提起法律诉讼方式，解决风险责任认定和损失处置争端，转移化解风险损失。

4.4 大数据技术

4.4.1 概述

大数据技术主要为风险智能防控提供基础的数据处理支撑，对数据保护、数据接入、数据处理与存储、变量与模型计算等提出了技术和安全要求。

4.4.2 数据保护

应建立符合《中华人民共和国个人信息保护法》、JR/T 0171—2020等相关法律、法规和标准的个人信息和业务数据保护策略、管理规范、管理制度等数据保护机制，在确保业务数据安全性的同时，加强个人信息保护，具体要求如下：

a) 数据授权方面：

——应符合数据所有者和相关人授权原则，涉及采集个人信息，应遵循最少够用原则，并在收集前给予当事个人声明数据采集内容、使用用途及保护措施，并征得当事人同意，采集的信息不能超范围使用。

——应明确并严格执行身份权限管理机制。

——应基于权限最小化等安全原则，制定支付风险防控相关数据访问控制管理机制。

b) 数据安全方面：

——个人金融信息、支付敏感信息等数据的存储应符合 JR/T 0171—2020 的相关规定。

——应采用满足数据传输安全策略相应的安全控制措施，如安全通道、可信通道、数据加密等。

——应依据数据资产和数据主体建立相应的数据脱敏安全机制与管控措施。

——应采用校验技术或密码技术保证支付风险防控相关数据传输过程中的完整性。

——应在发生个人信息泄露时立即采取补救措施，按照规定及时告知用户。

c) 数据删除方面：

——数据删除后，应确保数据及其副本不可检索、不可访问。

——根据不同的存储方式，如网络存储数据和闪存、硬盘、磁带、光盘等存储数据，分别明确相应的删除方法和技术。

——应确保个人信息、重要数据等敏感信息的删除符合国家相关法律、法规和标准。
——宜配置必要的数据删除工具，并对数据删除效果进行核验。

d) 数据审计方面：
——应对数据全生命周期中的采集、处理、销毁等操作行为进行记录，包括且不限于时间、操作方式、数据类型、操作结果等。
——应支持对数据操作行为的审计、追溯。
——宜定期对影响业务连续性的风险进行评估，并将相关的风险信息告知客户。

e) 数据备份与恢复方面：
——应具备支付风险防控相关数据的本地备份与恢复能力。
——宜具备支付风险防控相关数据的异地实时备份能力。

f) 数据出境方面：
——涉及数据出境的，应符合国家相关法律、法规和标准对出境数据处理流程的要求。

4.4.3 接入、处理与存储数据

4.4.3.1 概述

在大数据技术中，对于数据接入、数据处理和数据存储等各环节应支持存储结构化、半结构化及非结构化数据，提供丰富的 API 接口和 SDK 开发包，支持分布式计算、内存计算技术、流处理技术实现稳定的大数据处理能力。

4.4.3.2 数据接入

数据接入包括但不限于被动接入、主动接入、关联补齐和接入质控，具体要求如下：

a) 被动接入方面：
——宜支持被动获取方式，如提供对外接口，由其他系统调用传输数据。
——宜明确外部接口的服务窗口、接口标准，并明确异常补录流程。

b) 主动接入方面：
——宜支持主动获取方式，如订阅消息队列、定期抽取关系型数据或非关系型数据。
——宜支持关系型数据库、非关系型数据库、消息队列、批数据文件、ftp、接口等数据源。

c) 关联补齐方面：
——宜具备对相似数据的关联去重能力。
——对于数据源的缺失信息，宜支持自动补齐处理。

d) 接入质控方面：
——应支持对采集数据的数据质量进行监测，宜建立实时监测机制。
——宜具备对不同数据类型下数据质量的评价标准，如数据完整度、数据有效性。

4.4.3.3 处理方法

各类型处理方法的相关要求如下：

a) 流处理方面具体要求如下：
——应支持数据的实时获取、处理、输出和持久化。
——应支持对消息处理任务进行全生命周期管理，包括创建、浏览、中止、激活、去激活等。
——应支持事件驱动的流处理，降低处理延迟。
——应支持处理乱序事件流、窗口计算、复杂事件处理（CEP）等。
——应支持出现故障情况下使用容错机制处理事件。

——应支持提供 SQL 或者类 SQL 的数据操作接口。
——宜支持采用滑动窗口方式的实时分析任务，其时间窗口大小应可调，支持短窗口和长窗口。
——宜支持提供用户级别的访问控制功能。

b) 内存计算方面具体要求如下：
——应支持内存计算操作符，如聚集操作、转换操作等功能。
——应支持用高度抽象算子构建分布式的数据处理应用。
——宜支持标准 SQL 语法。
——宜支持读取非关系型数据库数据的能力。
——宜支持负载均衡和水平扩展能力。

c) 图计算方面具体要求如下：
——应支持多种数据导入方式，包括：全量导入、增量导入以及自定义导入。
——应支持图的基本操作，包含定义图、图的基础操作、图中点数据集和边数据集的相关操作。
——应支持同步计算模型或异步计算模型编写迭代算法。
——应支持单节点、多节点多层关系的分布式图分析和查询。
——应支持主流开发接口，如 RESTful 等。
——宜支持图关系的实时、可视化呈现。

d) 批处理方面具体要求如下：
——应支持从多种数据源读取数据，包括分布式文件系统、分布式列式存储等多种格式的数据源。
——应支持自定义的数据处理操作。
——应支持批处理任务的创建、配置等。
——应支持多节点离线任务联动执行。
——应支持离线计算任务进度与状态的实时上报。
——宜支持多种语言分析任务的开发接口。

4.4.3.4 存储库

系统应包含对采集的原始数据及分析处理之后生成的结构化数据进行持久化。

a) 对风控事件发生时的流水或快照建立的事件库应满足以下要求：
——支持实时、单次、批量事件的导入。
——支持事件查询、添加、删除、修改。
——包含事件对象的时间、地点、客户、卡号、类型等描述信息。

b) 对数据处理、变量统计、模型结果、关联特征等构建的风险变量库应满足以下要求：
——支持实时变量、批次变量、事件变量。
——支持单次、批量变量的导入。
——支持变量的查询、添加、删除、修改。
——支持多业务场景的变量共享。

c) 对具体风险特征和策略建立的规则库应满足以下要求：
——支持自定义规则和模型的创建。
——支持规则和模型的查询、添加、删除、修改。
——支持多业务场景的规则和模型共享。
——支持规则和模型结果编排。
——支持规则灰度发布。
——支持多维度组合规则。

d) 对支付及相关流程的数据要素建立不同类型的名单库应满足以下要求：

——支持第三方黑、白、灰名单导入。
——支持黑、白、灰名单的查询、添加、删除、修改。
——支持多业务场景的黑、白、灰名单共享。
——支持多维度的黑、白、灰名单整合。

4.4.4 变量与模型计算

4.4.4.1 特征变量计算

特征变量计算是对于模型所需的特征变量进行计算。宜满足以下要求：

a) 支持多业务、多渠道的变量计算。
b) 支持多维度（持卡人、卡片、商户、IP、MAC）变量的组合计算。
c) 支持对变量的自定义及动态扩展。
d) 支持对变量重要程度和变量筛选的计算。

4.4.4.2 滑窗变量计算

滑窗变量计算是基于一定的时间窗口向前滑动的变量计算，宜满足以下要求：

a) 支持按多维度进行变量定义。
b) 支持自定义时间窗口的变量计算。
c) 支持滑窗时间的动态调整。

4.4.4.3 在线模型计算

在线模型计算是在业务执行过程中实时进行的模型计算，应满足以下要求：

a) 支持灵活的模型编排设定，对模型的执行做好优先级控制，并有效控制模型的串执行关系，根据系统资源使用情况合理设定并行执行个数。
b) 支持在线模型的灵活下线。
c) 对模型预估时间具备有效的管理机制，避免模型异常影响正常交易。

4.4.4.4 离线模型计算

离线模型计算是在业务执行后非实时的模型计算，具体要求如下：

a) 应对离线计算资源做好隔离，避免影响在线业务运行。
b) 宜支持对海量数据的训练。
c) 宜支持用户侧对训练结果的评估。
d) 宜支持对离线模型多版本管理。

5 风险防控系统安全要求

5.1 总体要求

基于大数据的支付风险防控系统的安全控制措施部署应符合JR/T 0071、国家网络安全等级保护有关标准，并根据风险防控所面向的业务系统级别设定相应等级。

5.2 安全规划

基于大数据、人工智能的支付风险防控安全规划至少满足以下要求：

a) 总体规划：应将安全规划纳入系统总体规划中，制定相应的安全规划，包含数据治理、数据质量、

元数据、授权管理等方面的安全策略，并对其进行评估，确保规划内容的合规性。

b) 需求分析的具体要求如下：

——应建立安全需求分析和评审机制，识别并分析威胁、脆弱性等安全风险及其应对措施需求。

——宜使用数据驱动分析方法或安全需求工程思想进行安全需求分析，确保安全需求的有效制定和规范化表达。

c) 方案评估的具体要求如下：

——方案实施及重大业务变更前，均应对其进行评估和检查，明确评估要素和内容，形成评估报告。

——宜对安全方案的执行情况进行跟踪和评估，并对所用开源软件进行安全管理。

5.3 开发部署

基于大数据、人工智能的支付风险防控开发部署至少满足以下要求：

a) 设计安全：安全架构应与安全规划保持一致性，并论证其有效性，明确安全功能和服务接口，包括接口参数等，并编制安全功能设计文档。

b) 开发安全：应制定并严格遵循的编码规范，建立适宜的源代码管控机制、开发外包安全管控及软件安全测试规程，并加强开发和交付人员权限管理，确保开发过程的安全性。

c) 部署安全：应依据授权最小化原则，明确安装部署过程的角色职责及其权限，并制定相应的授权策略，及时清除安装部署过程中产生的中间文件，避免中间文件引起的数据泄露。

d) 边界安全的具体要求如下：

——应规划业务控制、应用隔离相关的安全域，制定边界安全控制策略和管理规则。

——宜具备安全域间数据隔离机制和访问控制机制。

e) 接口安全的具体要求如下：

——应制定接口安全控制策略，如身份鉴别、授权策略、访问控制机制等，提供接口异常处理能力，如对接口非法输入参数进行限制或过滤，并具备接口访问的审计能力。

——跨安全域的接口调用应采用安全通道、加密传输等安全机制。

f) 文档安全：应建立文档安全管理机制，并明确访问权限及安全责任，定期对文档进行评审、更新、批准和发布。

5.4 应用安全

基于大数据、人工智能的支付风险应用安全至少满足以下要求：

a) 应用程序管理方面，具体要求如下：

——应建立组件管理规程及运行环境安全评估策略与规程，建立安装包及升级包管理机制和安全检查机制，并明确应用访问权限。

——宜定期对已部署的应用程序进行安全风险评估。

b) 应用终端安全方面，应严格管控应用终端的数据访问权限，建立终端输入约束规范和安全防护机制，并建立数据采集、监控与审计系统，追踪、分析和记录终端用户行为以识别异常操作。

c) 身份鉴别方面，具体要求如下：

——应建立基于多因素鉴别技术的身份标识和鉴别机制。

——应采用密码技术和访问控制技术等对鉴别凭证信息的传输和存储进行保护。

——应定期对账号使用情况进行安全性分析（如登录时间、登录位置、访问时长、访问模块等），评估账号安全风险。

d) 授权与访问控制方面，依据角色控制和最小授权原则，建立访问授权机制，并制定信息流安全控制策略和机制，对数据导入、导出、迁移、发布等信息流动进行控制。

e) 租户数据安全方面，应制定多租户应用程序及其服务数据资源的隔离策略与规程，并建立多租户应用可用性保障策略和机制。

f) 应用行为监测方面，具体要求如下：

——应建立大数据应用行为及其数据使用监测策略和规程，具备异常行为记录、统计分析和告警能力。

——宜支持自定义行为监测规则。

5.5 安全运维

基于大数据、人工智能的支付风险安全运维至少满足以下要求：

a) 配置与变更管理方面，具体要求如下：

——应制定管理规程，确定安全基线配置清单，在实施配置或变更前，对受控配置项和变更项进行测评，并定期对配置项进行安全审查。

——宜定期或在业务、系统架构发生重大变更时，开展配置管理效果风险评估。

b) 补丁管理方面：应建立漏洞、脆弱性等补丁管理规程，补丁部署安装前应经过兼容性测试。

c) 系统与数据迁移方面，应建立迁移策略与规范，配置必要的迁移工具，记录迁移过程并确保可溯源性，具备安全风险分析能力及迁移完整性和一致性检测能力。

d) 第三方服务管理方面，具体要求如下：

——应建立外部服务组件合作方安全管理制度，通过合作协议等方式明确其义务和责任。

——宜对其资质和安全能力进行评估，明确外部组件访问权限，并与其形成应急联动机制。

e) 运维监控方面，具体要求如下：

——应建立安全监控架构，具备安全漏洞库、漏洞扫描工具等，支持分布式节点统一监控，进行报警并生成状态分析报告。

——宜具备大数据服务安全能力检测和安全势态分析能力。

f) 安全风险评估方面，建立安全风险评估机制，定期或系统运行环境等发生重大变更时开展风险评估，并定期开展安全评估情况抽查。

g) 灾备及恢复方面，根据业务目标和安全策略，建立系统灾备及恢复机制，明确需求并划分灾难恢复能力等级，制定相应预案。宜定期进行预案演练，根据演练情况修订预案，确保备份系统与数据的有效性。

h) 系统应急响应方面，制定大数据服务应急处理机制、应急响应预案及应急响应定期演练计划，记录并保存演练记录及总结报告。

i) 业务连续性计划方面，应建立并执行业务连续性计划及其定期演练计划，验证业务连续性及数据与系统资产的可用性，并定期对连续性风险进行评估。

5.6 安全审计

基于大数据、人工智能的支付风险安全审计至少满足以下要求：

a) 审计策略管理方面，具体要求如下：

——应制定大数据服务行为和数据活动的审计策略，明确审计规程。

——宜定期对审计策略与规程的实施情况进行检查和评价。

b) 审计内容要求方面，应明确与行为（如用户登录、账号管理等）和数据活动（如数据采集、数据访问等）相关的可审计事件。

c) 审计数据保护方面，具体要求如下：

——应加强审计数据的访问控制，记录对审计数据的所有操作。

——应对导出的数据进行脱敏处理。

——宜采用密码技术等安全技术保障审计数据的完整性和抗抵赖性。

d) 安全审计方面，宜定期通过内部审计或委托外部审计机构的方式对支付风险防控系统安全情况进行审计，并及时将审计结果按照监管要求进行相应的报送。

附　录
（资料性）
机器学习

1　监督学习

通过对已知的输入和输出数据的分析学习建立预测能力，从而可以对新的输入数据预测结果，这就是监督学习。用于进行模型训练的数据就是训练集。监督学习的训练集要求包括输入和输出，也称为特征和目标。训练集中的输出数据是已知的结果，称为标记数据。学习过程中，对训练样本集的每个输入训练样本数据，都提供对应期望输出结果的标记数据，在选定算法模型和训练集后，通过训练过程中不断对比标记结果，自动反馈调整算法模型参数，得到最优的 1 组模型参数的学习过程，见图 1：

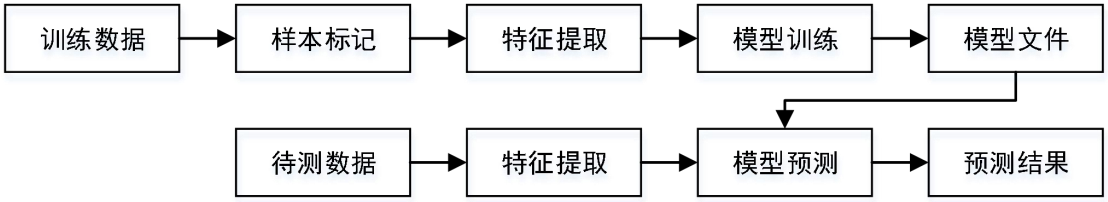

图1　监督学习的典型流程

监督学习由训练和预测2部分构成：训练环节针对带标签的样本进行特征提取，然后采用一定的监督学习算法进行训练并生成模型文件，同时应进行模型效果评价；预测环节首先针对待测数据进行特征提取，然后进行模型预测，最后生成预测结果文件。

在有足够数量的高风险、低（无）风险样本的场景下，宜使用监督学习技术。在监督学习中，每个样本都由一个输入对象和一个期望的输出值（标记）组成。

监督学习技术可以从有标记的样本中自动学习风险场景对应的行为模式。针对不同的数据类型和业务场景，可以选取不同的监督学习模型有针对性地识别多维度特征空间的风险模式，如逻辑回归、随机森林、梯度提升决策树、可扩展梯度提升以及深度学习等模型。

2　半监督学习

在进行模型训练时，如果训练集的输入训练样本数据对应的期望结果标记数据较少，大部分都是无标记结果的数据，学习器不依赖外界交互、自动地利用未标记样本来提升学习性能就是半监督学习。

半监督学习的典型流程见图2：

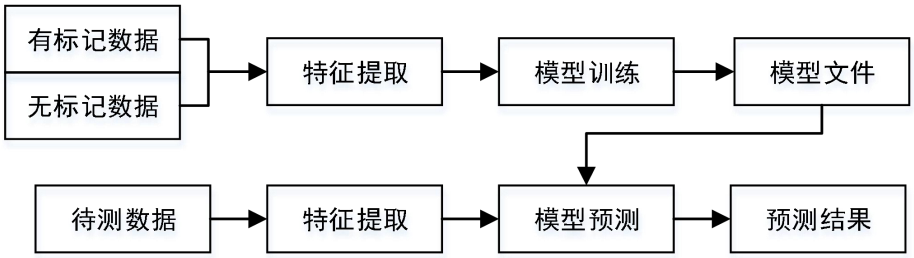

图2　半监督学习的典型流程

半监督学习由训练和预测2部分构成：训练环节针对有标记样本和未标记样本同时进行特征提取，然后采用一定的半监督学习算法进行训练并生成模型文件；预测环节针对待测数据进行特征提取，然后进行模型预测，最后生成预测结果文件。

半监督学习可分为纯半监督学习和直推学习，两者的区别在于如何看待未标记的样本。前者假定训练数据中的未标记样本并非待预测的数据，而后者则假定学习过程中所考虑的未标记样本恰是待预测数据，学习的目的就是在这些未标记样本上获得最优泛化性能。纯半监督学习希望学得模型能适用于训练过程中未观察到的数据，而直推学习试图对学习过程中观察到的未标记数据进行预测。半监督学习技术自动地利用未标记样本以提升学习性能。

在风控检测任务中，当未标记样本多、高风险标记的样本少时，宜使用半监督学习技术。半监督学习技术自动地利用未标记样本以提升学习性能，可以同时从有标记样本和无标记样本中学习风险模式。此时，未标记样本可以作为风险模式对应分布特征，提供某种正则化，从而使模型学到的信息更为全面。典型的半监督学习包括标签传播算法、直推式支持向量机等。以标签传播算法为例，该算法可以基于少量的风险样本进行标签传播，从而发现更多的风险样本。半监督学习的典型应用场景有团伙挖掘、生物数据分析等。半监督学习同时使用未标记样本和有标记样本，降低了样本标记的工作量，同时又能够带来比较高的准确性。

3　无监督学习

无监督学习区别于监督学习和半监督学习的一个重要特点是，训练样本集的每个输入训练样本数据，没有对应的期望结果标记数据，而是对训练样本集本质特征的归纳抽取信息，如样本的分布、样本距离、关系、密度、相似性等度量特征，并在训练过程中调整算法模型参数，直到满足训练目标。当前在无监督学习中应用最广的就是聚类。

聚类是按照数据内部之间的分布结构，将数据划分成多个没有交集的子集（每个子集被称为簇）。通常这些数据划分的逻辑和意义应通过人为分析、总结去进行定义。通过这样的数据划分和分析定义，簇就可能对应一些实际的概念和意义。聚类可以用来寻找数据分布的潜在特点，还可以用来作为其他学习任务的前置任务，先进行划分后再进行其他后续分析学习。

无监督学习的典型流程见图3：

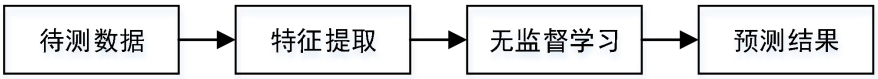

图3　无监督学习的典型流程

无监督学习首先针对样本进行特征提取，然后采用一定的无监督学习算法进行学习，最后生成结果文件。

无监督学习技术可以从未标记样本分布中自动计算出可能的风险模式。在支付风险防控场景常用的无监督机器学习技术主要包括K均值聚类、层次聚类等各种聚类算法、主成分因子分析等降维技术等。无监督学习的典型应用场景有相似用户挖掘、文本聚类等。无监督学习的优势在于在缺少标记样本的情况下，可以进行有效的样本挖掘。无监督学习的劣势在于针对大数据的聚类等处理的计算复杂度较高。

4　关系网络

传统通用的风险识别，更多的是从个体的指标角度识别。针对群体作案形式，风险识别不能局限于个体特征，宜使用关系网络来识别团体性风险活动。关系网络通过个体之间的行为等信息建立全局的关系图，进而在全局关系图上，通过图算法，发现具有一定行为模式的团体。关系网络可以通过关系识别、特征挖掘、构建网络等方式识别团体性风险活动。关系网络作为风险识别的1种手段，可以在事前、事中、事后的各个阶段进行风险防控。

关系网络图算法，利用图结构挖掘相关特征的算法，常见的有社交关系网络图，个体社会信息异构图等。算法包括构建图，识别图形异常结构，标签风险传播等。

关系网络的典型处理流程见图4：

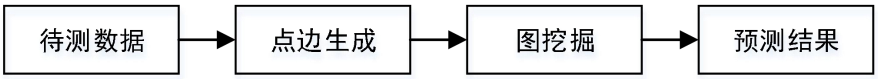

图4　关系网络的典型处理流程

关系网络图算法首先针对数据集计算生成点和边，然后采用社区挖掘、标签传播等不同的图挖掘算法进行计算并生成结果数据文件。

参 考 文 献

[1] GB/T 5271.31—2006 信息技术 词汇 第31部分：人工智能 机器学习

[2] GB/T 23694—2013 风险管理 术语

[3] GB/T 35295—2017 信息技术 大数据 术语

[4] GB/T 37721—2019 信息技术 大数据分析系统功能要求

[5] GB/T 37722—2019 信息技术 大数据存储与处理系统功能要求

[6] 银监会.中国银监会关于印发银行业金融机构全面风险管理指引的通知（银监发〔2016〕44号），2016年09月27日