ICS 35.240.40
CCS A 11

JR

中华人民共和国金融行业标准

JR/T 0263—2022

机器学习金融应用技术指南

Technical guidance on application of machine learning in financial services

2022－11－25 发布　　2022－11－25 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科技司提出。

本文件由全国金融标准化技术委员会（SAC/TC 180）归口。

机器学习金融应用技术指南

1 范围

本文件提供了金融业开展机器学习应用涉及的体系框架、计算资源、数据资源、机器学习引擎、机器学习服务、安全管理、内控管理等方面的建议。

本文件适用于开展机器学习金融应用的金融机构、技术服务商、第三方安全评估机构等。

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（包括所有修改单）适用于本文件。

GB/T 27910—2011 金融服务 信息安全指南

JR/T 0071.2—2020 金融行业网络安全等级保护实施指引 第2部分：基本要求

JR/T 0071.5—2020 金融行业网络安全等级保护实施指引 第5部分：审计要求

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

机器学习 machine learning

功能单元获取新知识或者技能，或通过已有的知识或技能改进其性能的过程。

[来源：GB/T 5271.31—2006，31.01.02]

3.2

监督学习 supervised learning

基于外部知识源的反馈以测试获得知识的正确性的学习策略。

[来源：GB/T 5271.31—2006，31.03.08，有修改]

3.3

无监督学习 unsupervised learning

无需基于外部知识源的反馈以测试获得知识的正确性的学习策略。

[来源：GB/T 5271.31—2006，31.03.09，有修改]

4 总体原则

机器学习在金融服务中应用的原则一般包括以下内容。

a) 普遍性。确认目标群中所有主体均能成功且一致地使用机器学习进行服务。

b) 不可否认性。涉及机器学习且已经发生的活动或事件不可被否认。

c) 可控性。主体对机器学习应用的使用范围、运行状态、访问权限等具备主动控制的能力。

5 体系框架

机器学习能为金融应用系统提供更丰富、更便利、更通用的智能化支撑服务，例如智能语音、自然语言处理、计算机视觉、生物特征识别、知识图谱等。基于机器学习的金融应用系统一般性体系框架见下图。

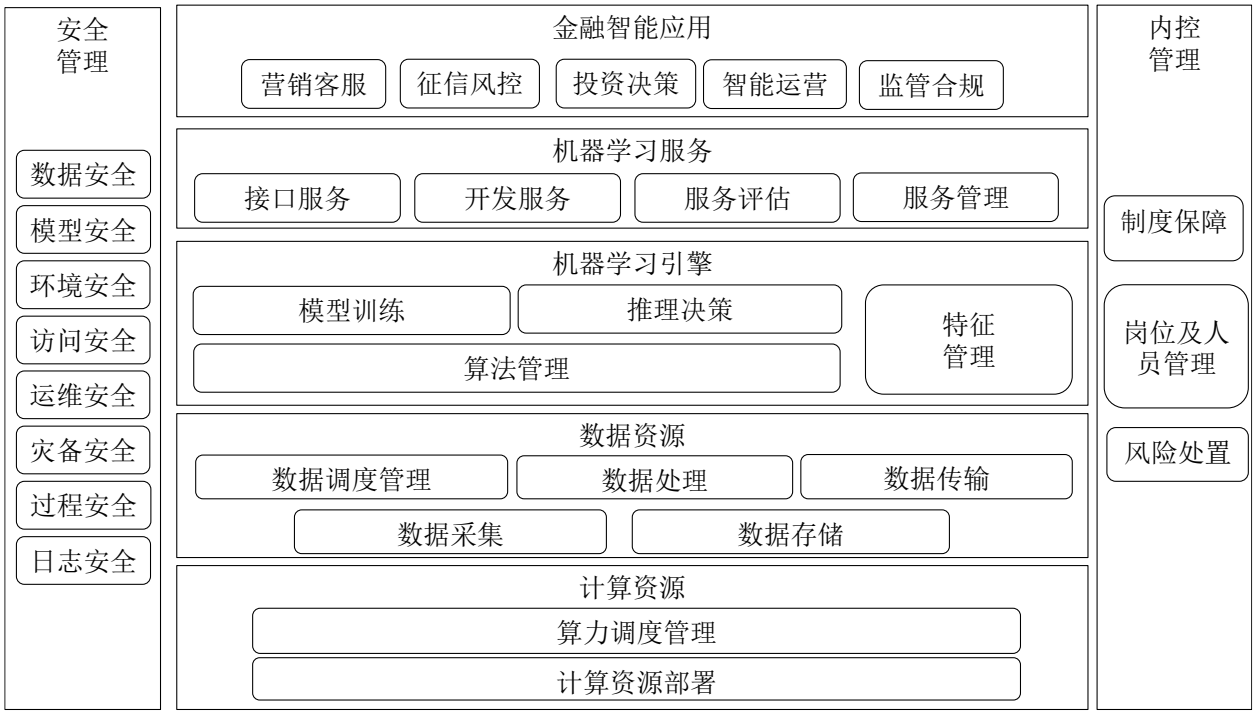

图　基于机器学习的金融应用系统一般性体系框架

基于机器学习的金融应用系统一般性体系框架包括计算资源、数据资源、机器学习引擎、机器学习服务、金融智能应用、安全管理与内控管理等。整个框架表明了机器学习金融应用可遵循的、适用于不同规模的一般性层次化抽象、结构及依赖关系。体系框架中包含如下部分。

a) 计算资源部分统筹管理多种可选用的系统底层计算硬件，例如中央处理器（CPU）、图像处理器（GPU）、张量处理器（TPU）、神经网络处理器（NPU）、可编程逻辑阵列（FPGA）和专用集成电路（ASIC）等，通过部署并调度管理各类计算资源，提供机器学习过程所需的算力支持。

b) 数据资源部分负责为机器学习的数据使用需求提供支撑，包含数据采集、数据存储、数据调度管理、数据处理、数据传输等过程。

c) 机器学习引擎部分为上层机器学习服务提供特征管理、算法管理、模型训练、推理决策等支撑。

d) 机器学习服务部分提供不同场景应用所需的基础通用服务框架，包含接口服务、开发服务、服务评估和服务管理等功能。

e) 金融智能应用部分提供面向最终用户、针对不同金融领域应用场景的综合服务，包含营销客服、征信风控、投资决策、智能运营、监管合规等，具体应用场景见附录。

f) 安全管理部分提供计算资源、数据资源、机器学习引擎、机器学习服务及金融智能应用整个体系的安全保障，包含数据安全、模型安全、环境安全、访问安全、运维安全、灾备安全、过程安全和日志安全等。

g) 内控管理部分为机器学习在金融场景的有效应用提供组织、方法和流程上的保障，包含制度保障、岗位及人员管理和风险处置等。

6 计算资源

6.1 计算资源部署

设计和部署底层计算资源时宜充分考虑其安全性、可靠性及可扩展性，需考虑的内容如下。

a) 宜整体考虑硬件、操作系统、上层计算服务组件的安全，确保底层算力的安全性。

b) 宜保证计算资源启动过程的安全性及上层应用的安全性，例如可通过逐级进行安全签名校验等方式确保固件的完整性和可信度。

c) 在实时性较强的机器学习金融应用系统中，宜保证计算资源的冗余、异构性和任务的可恢复性，从而保证业务连续性。

d) 宜保证架构的可扩展性，对计算资源进行虚拟化、池化管理，支持水平扩展，能提供按需（数据量、算法类型、实时性偏好等）分配计算资源的能力。

e) 宜定期开展底层计算资源安全性的审核工作，监控设备运行状态、资源使用情况，发生异常情况及时告警。

f) 宜根据数据密集程度、格式、数据量方面的差异合理分配计算资源，以达到理想的实时处理效果。

g) 宜具备对不同硬件资源的兼容能力，例如服务器、一体机、边缘计算节点、计算集群和云基础设施等。

h) 宜提供异构资源管理调度功能，屏蔽下层硬件的异构复杂性，提供统一的计算资源服务接口。

6.2 算力调度管理

算力调度管理过程涉及多模块间的协同工作，宜重点关注过程中的数据隐私、数据安全等问题，需考虑的内容如下。

a) 在不同任务或计算节点间切换时保证数据处理过程的保密性，避免因调度过程导致数据泄露等问题。

b) 在人机协同模式中，宜保证主机安全并使用完整的加固流程，例如严格裁剪服务、进行网络端口扫描、开启主机入侵检测和漏洞管理服务。

c) 在云化协同模式中，宜开展周期性安全指数评估工作，部署信息防泄露监控工具，保证金融行业数据中心运行环境位于高安全区域。

d) 宜定期检查或评估数据连接通道的安全性和可靠性。

e) 宜根据场景和运算需求分解应用负载，并分配任务执行的优先级。

f) 宜按需使用多处理单元，实现多芯片、多核、多机、多上下文并行。

7 数据资源

7.1 数据采集

采用机器学习的金融应用系统的数据采集过程需考虑的内容如下。

a) 对数据采集的来源和方式、数据范围、数据形式、数据状态等方面进行说明。

b) 对数据采集的环境、工具、技术以及校验方法等采取必要的管控措施，保证所采集数据及其标记的完整性、真实性和一致性。
c) 数据采集后，宜对采集数据进行初步清洗，包括去除重复及错误数据、补充残缺数据等。
d) 宜记录数据采集过程中的活动及其责任人，保证采集活动的可追溯性和可审计性。
e) 个人金融信息数据和金融业务敏感数据的采集过程宜符合 JR/T 0171—2020 相关要求。

7.2 数据存储

采用机器学习的金融应用系统的数据存储过程需考虑的内容如下。
a) 在必要时，宜对采集或应用生成的数据进行持久化存储，以便后续用于模型的训练及校验。
b) 数据的存储宜设置存储时限，并满足时间最小化原则，即为满足机器学习目的所必需的最短时间。
c) 在超出数据存储时限后，宜立即对所存储的数据进行处理（例如删除、销毁、脱敏等）。
d) 宜支持结构化存储方式（例如关系型数据库）和非结构化存储方式（例如键值数据库、图数据库等）。
e) 个人金融信息数据和金融业务敏感数据的数据存储过程宜符合 JR/T 0171—2020 相关要求。

7.3 数据处理

采用机器学习的金融应用系统的数据处理过程需考虑的内容如下。
a) 宜采用规范化的数据处理过程和数据校验机制，保证数据处理结果的一致性和准确性。
b) 宜支持数据聚合功能，即对分散的数据进行融合或拼接。
c) 宜对数据处理过程进行管理和记录，保证数据处理活动的可追溯性和可审计性。
d) 个人金融信息数据和金融业务敏感数据的委托处理和加工处理过程宜符合 JR/T 0171—2020 相关要求。

7.4 数据传输

采用机器学习的金融应用系统的数据传输过程中，对适用于 JR/T 0171—2020 相关要求的个人隐私及业务敏感数据，宜建立相应的数据传输安全策略和规程并采用有效的技术手段和控制措施，确保数据在传输过程中的保密性、完整性和可用性。

7.5 数据调度管理

采用机器学习的金融应用系统的数据调度管理过程需考虑的内容如下。
a) 宜支持数据宿主选择，即选择数据服务过程中承接数据或驱动调度的节点。
b) 对于智能金融应用场景中高频次的实时推理决策需求（例如事中不当操作检查），宜使用适当拥塞控制策略以确保整个系统处理链条的健康状态。
c) 宜提供弹性扩展功能以增强系统整体吞吐量并分摊数据服务压力。
d) 宜支持执行算子抽象技术，将数据流控制算法、数据获取算法等功能抽象为统一数据服务接口，并遵循标准通信协议，避免跨机构金融协作活动场景中出现数据格式、协议不一致等情况。
e) 宜支持数据虚拟化和基于规则的报警机制（例如日志审计）等，确保数据服务的整体健壮性。

8 机器学习引擎

8.1 特征管理

金融应用系统中机器学习引擎的特征管理需考虑的内容如下。

a) 宜保证特征的保密性，使特征在产生、传输、处理和存储的各个环节中不被泄露给未授权的个人和实体。
b) 宜保证特征的完整性，即特征在传输过程中不被篡改、破坏和插入，不发生延迟、乱序和丢失的情况，特别是采用分布式存储方式时，确保内容的一致性。
c) 宜保证特征的可用性，使特征可被已授权的其他机器学习子系统或实体使用。
d) 根据算法的需要，宜支持自动及手动选择与构建特征。
e) 宜支持离散特征、连续特征、时序特征、组合特征的抽取。
f) 宜支持显式特征构造方法，以增强特征的可解释性。
g) 宜保证离线特征与在线特征的一致性。

8.2 算法管理

在基于机器学习的金融系统核心业务中，例如核心交易系统、清结算系统、量化交易系统等，算法设计者宜提高算法的鲁棒性，增强安全性，需考虑的内容如下。

a) 当训练数据中有恶意数据破坏原有训练数据分布时，宜能区分和识别恶意数据，防止模型精度降低。具体对策可包括增加算法参数、丰富特征库、优化权重比例等。
b) 在算法设计过程中，宜综合考虑安全性、泛化性能和算法开销，合理权衡算法的安全性、实用性和性能，以更好地满足实际应用需求。
c) 宜采用集成学习、模型融合等提升手段集成不同的决策方法，以增强模型泛化能力，更好地适应未知数据分布情况。

8.3 模型训练

8.3.1 训练准备

金融应用系统中机器学习引擎的训练准备需考虑的内容如下。

a) 宜根据金融应用场景需求建立模型选择策略，策略选择依据包括但不限于模型复杂度和可解释性、原始数据及中间数据规模和维度、存储与计算资源成本、金融业务诉求精度和准确率、模型结果的时效性以及潜在的外围干扰项。
b) 宜根据金融应用场景需求定义约束条件和评价指标，并将其转换为对应监督学习或无监督学习的性能指标和评价函数，具体内容如下。
   ——监督学习评价指标，例如混淆矩阵、查准率、查全率、准确率、置信度、误差平方和、决定系数、对数似然损失函数、受试者工作特征曲线（ROC）的曲线下面积（AUC）值等。
   ——无监督学习评价指标，例如方差、还原误差、置信度、困惑度、类间距离、类内距离等。
c) 当数据量相对较少时，宜采用交叉验证，例如简单交叉验证、留一交叉验证等。

8.3.2 训练过程

金融应用系统中机器学习引擎的模型训练过程宜根据金融应用场景需求选择合适的训练方法，需考虑的内容如下。

a) 宜使用特征工程方法提升模型效能，具体内容如下。
   ——引入行业专家筛选优选特征，借助特征可视化提升场景描述效率。
   ——强化原始特征数学统计描述并提升特征表达能力，并通过特征组合或拆解提取特征的相关性，调整数据特征以增强特征可分析性。
   ——构造适用于训练学习过程和结构的特征以便机器学习算法使用。
   ——通过坏样本发现易被忽视的特征。

——采用上下采样或数据权重调整等方法进行数据均衡化处理。

——挖掘数据异常点和孤立点并采用数据清洗和数据变换等方法降低数据噪声。

b) 宜根据场景需要提升训练模型的稳定性，具体内容如下。

——可通过数据增强技术和数据扩展技术，提高模型泛化、抗噪能力。

——对于模型训练过程本身，宜使用剪枝、部分激活、批正则化或权重衰减等方式降低算法对训练数据的过拟合风险。

——在模型参数设置阶段宜根据数据分布特点进行合理初始化，特别是对于无监督学习方法，以提升模型稳定性、加快收敛速度。

——根据应用场景特点，对模型进行定期评估更新以适应用场景的变化，并对模型更新过程作出规范，使用模型再训练、在线学习、迁移学习等方式保证模型持续有效。

c) 宜根据应用场景数据规模优化训练效果，具体内容如下。

——对指标不敏感的场景，可通过调整学习率以加快模型训练收敛。

——对于较复杂或规模较大的业务网络，可使用残差网络等方式解决梯度衰减问题。

——针对小批量样本，可扩充数据集，保证训练集的数量级不小于模型的复杂度。

——通过调整模型结构和模型正则化，降低模型复杂度，保证模型的复杂度不大于训练集的数量级。

——在监督学习训练数据集中引入对抗样本，通过对抗攻击构建高容忍模型，提高鲁棒性。

8.4 推理决策

推理决策宜满足金融应用场景需求，需考虑的内容如下。

a) 宜保证模型的可用性，对样本数据有较好的容忍度，在极端情况下，保证模型可以正常返回结果，供系统进行决策处理。

b) 宜保证输入样本数据及输出返回结果的保密性，确保不被未授权用户非法获取。

c) 宜保证输入样本数据及输出返回结果的完整性，确保不被非法篡改。

d) 宜保证关键性场景中模型的可解释性，从业务建模、参数设置和样本选择等多方面提升模型可解释性。

e) 宜使用模型压缩或者剪裁提升推理速度，并对使用的压缩、蒸馏、裁剪方法记录备案。

f) 从多方面提升推理决策服务的可用性和效率，主要内容如下。

——具备任务按需调度能力，根据模型、业务特点确定计算节点和存储节点。

——统一管理中心集群与边缘节点，使边缘业务就近调度到边缘设备执行。

——具备任务跨集群调度与多级资源拉通共享能力，可将本地任务调度到另一个集群中计算。

9 机器学习服务

9.1 接口服务

金融应用系统中机器学习服务的接口服务需考虑的内容如下。

a) 宜制定机器学习服务接口调用的技术规范，并确保外界无法通过接口获取相关隐私或者敏感数据。

b) 对于机器学习服务接口的定义，宜尽量缩小输入输出的数据范围，保证输入参数和输出数据没有冗余，防止额外的数据通过接口泄露。

c) 机器学习服务接口调用宜做好数据参数检查，防范数据库注入攻击等风险。

d) 算法接口宜支持对算法的能力信息进行展示，包含算法名称、算法运行环境信息、算法运行能力。

e) 宜支持对算法资源进行管理，包含算法资源增加、算法资源修改、算法资源删除、算法资源查询、算法资源下载等接口。
f) 宜对算法接口处理结果返回状态进行明确定义，例如处理成功、处理失败、未匹配等。
g) 宜对接口返回内容进行明确定义，例如分类标签范围、概率值区段等。
h) 算法接口宜支持多种市场主流工具和模型文件格式。

9.2 开发服务

金融应用系统中机器学习服务的开发服务需考虑的内容如下。
a) 宜提供数据集管理、半自动化标注、模型超参数自动调优、自动机器学习、可视化建模与评估等功能，减少外接组件数目和配置工作量，降低兼容性风险。
b) 宜提供图形化工作流编辑界面，快速构建应用处理逻辑，减少金融系统复杂场景应用的开发维护工作量。
c) 宜在训练过程中，通过适当的自动化模型质量评估手段（含可视化），来反馈模型的质量，提升模型选择及构建的效率。
d) 宜兼容不同架构的机器学习框架及不同数据源。
e) 宜在模型开发、训练、构建、数据处理等过程中，为开发人员或生产运营人员提供机器学习服务、各类资源的使用情况和可用性的监控手段。
f) 宜支持学习训练、数据处理等过程中各类计算资源的扩容，以适应不同数据规模和模型复杂度的计算需求。
g) 宜支持模型训练和模型预测服务中所需要的各种周边辅助算子，包括数据分析、隐私集合求交等多种数据预处理算子。

9.3 服务评估

金融应用系统中机器学习服务的服务评估需考虑的内容如下。
a) 宜具备对机器学习的任务进行状态跟踪与记录的功能。
b) 宜构建完整的机器学习金融应用模型上线与评估机制，包含效果评估、性能评估、安全评估、对比测试、上线流量控制等机制。
c) 机器学习金融应用新模型上线后，宜定期观测新模型在线上的效果、检测模型在测试与线上环境下的效果和性能差距并进行迭代训练。
d) 机器学习金融应用模型上线宜采用对比测试和灰度发布机制，使得上线过程能够平滑过渡,且具备版本回溯功能。

9.4 服务管理

金融应用系统中机器学习服务的服务管理需考虑的内容如下。
a) 机器学习金融应用系统开发宜做好模块分离，包括但不限于降低应用层模块、算法层模块之间的耦合。
b) 宜对机器学习服务进行封装，降低使用方的接入成本。
c) 宜支持数据源水平切分及垂直切分、模型训练和模型预测等服务。
d) 宜支持对机器学习分布式任务的分解与调度。

10 安全管理

10.1 数据安全

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：https://d.book118.com/398110025102006043