ICS 03.060
CCS A 11

Q/HSQH

# 徽商期货有限责任公司企业标准

Q/HSQH 001—2026

# 基于大模型的客户风险画像应用建设指南

Guidelines for Large-Model-Based Customer Risk Profiling Application

2026-03-02 发布　　2026-03-02 实施

徽商期货有限责任公司　　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范》的规定起草。

本文件由徽商期货有限责任公司提出。

本文件由徽商期货有限责任公司归口。

本文件起草单位：徽商期货有限责任公司。

本文件主要起草人：常飞、胡毅、叶晗、范涛、乔焰、李敏悦、王俊杰。

# 引　言

客户风险画像作为风险管理的重要组成部分，是实现差异化风控、精准服务和合规监管的关键技术手段。然而，在实践过程中也面临诸多挑战，如客户数据来源复杂、结构异构，数据缺失与质量问题频发，风险指标体系缺乏统一规范，不同机构所采用的大模型性能标准不一致，模型结果可解释性和评估机制尚未形成共识等。这些问题限制了客户风险画像的应用效果与跨机构协同能力。

基于大模型的客户风险画像评估需要充分发挥大模型在多模态数据处理、自然语言理解与生成等方面的技术优势，并构建标准统一的数据输入规范、科学合理的用户画像评估指标体系，以及多维度、多层次的模型性能评估体系。通过统一的数据结构、指标分类和评估维度，有望有效提升画像结果的准确性、稳定性和可解释性，并逐步实现客户风险的动态评估与穿透式监管。

在推进客户风险画像体系建设的过程中，安全问题不容忽视。一方面，画像所依赖的客户数据往往包含大量敏感个人信息和金融行为数据，需严格落实数据最小化采集、分级分类保护、访问权限控制和全生命周期安全管理；另一方面，大模型本身可能存在提示注入、对抗攻击、隐私泄露等新型安全风险，亟需建立覆盖模型训练、推理、部署各环节的安全防护机制与审计追踪能力。此外，画像结果若被误用或滥用，可能引发歧视性决策、算法偏见等伦理与合规问题，因此必须嵌入公平性约束、可解释性保障与人工复核机制，确保风险评估过程安全、可信、可控。

模型所用的各类数据元应建立定义清晰、属性明确、约束规范的数据标准体系，有助于业务、数据、技术等各相关方对数据的统一理解和使用，确保数据在采集、加工、交换和共享等环节的一致性和准确性，减少系统数据转换和清洗工作复杂程度，降低数据处理和应用成本，促进数据规范融合、提升数据质量水平、提高数据交换效率、发挥数据应用价值等。

本指南对数据规范、用户画像评估指标体系与大模型性能评估指标进行了统一规范，明确了的结构设计与分配要求。

# 基于大模型的客户风险画像应用建设指南

## 1 范围

本文件规定了基于大模型的客户风险画像系统的数据规范、风险评估指标体系、大模型性能评估指标等内容的指示。

本文件适用于各类金融机构、监管机构、技术服务商等单位在设计、建设、运维客户风险画像系统。

## 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。

GB/T 18391.1 信息技术 元数据注册系统（MDR） 第1部分：框架（ISO/IEC 11791-1:2004,IDT）

JR/T 0319—2024 证券期货业数据标准属性框架

JR/T 0304.1—2024 证券期货业基础数据元规范 第1部分：基础数据元

JR/T 0304.2—2024 证券期货业基础数据元规范 第2部分：基础代码

JR/T 0276—2023 证券期货业信息系统渗透测试指南

JR/T 0292—2023 证券公司核心交易系统技术指标

JR/T 0176.1—2019 证券期货业数据模型 第1部分：抽象模型设计方法

Q/ JFY 087-2023 金融智能合规算法模型技术标准

## 3 术语和定义

下列术语和定义适用于本文件。

3.1

**客户风险画像 customer risk profiling**

指利用多源数据与智能分析模型，对客户的风险倾向、行为特征、异常行为模式等进行刻画、量化与分类的过程。

3.2

**大语言模型 large language model**

以大规模文本语料训练，具备上下文语义理解与生成能力的深度学习模型。

3.3

**异构 heterogeneous**

指系统、组件或数据在结构、类型、协议或技术实现上存在差异，不具备统一性或互操作性的特性。

3.4

**提示词 prompt**

使用大模型进行微调或下游任务处理时，插入到输入样本中的指令或信息对象。

[来源：GB/T 45288.1-2025]

3.5

**用户数据 Customer Data：**

用户上传或系统调用的客户相关数据。

3.6

**数据 data**

信息的可再解释的形式化表示，以适用于通信、解释或处理。

**注：**数据可以由人工或自动的方式加工、处理。

[来源：GB/T 5271.1-2000，01.01.02]

3.7

**值域 value domain**

允许值的集合。

[来源：GB/T 18391.1—2009，3.3.38]

3.8

**模型微调 model fine-tuning**

指在预训练模型的基础上，使用特定任务或领域数据对其进行二次训练，以提升模型在特定应用场景下的表现。

3.9

**测试 test**

在规定的条件下，对系统、组件或产品的特定功能与性能进行运行，并将其结果与预期要求进行比较验证的活动。

3.10

**模型推理 model inference**

指利用已训练好的机器学习模型对新的输入数据进行分析，并生成预测或分类结果的过程。

3.11

**风险等级 risk level**

指依据风险评估结果所划分的，用于标识风险严重程度或优先级的等级序列。

3.12

**数据元 data element**

由一组属性规定其定义、标识、表示和允许值的数据单元。

[来源：GB/T18391.1-2022,3.3.8]

3.13

**时序建模 time series modeling**

指对按时间顺序排列的数据序列进行统计分析与建模，以揭示其内在规律、趋势或周期性，并用于预测未来值或检测异常的过程。

3.14

**模型调用协议 Model Calling Protocol（MCP）**

大模型在推理过程中，用于动态调用外部工具、服务或计算模块的一组标准化接口规范。

## 4 制定原则

### 4.1 整体性原则

客户风险画像应用由业务、应用、数据、技术等多个模块构成。其管理宜覆盖需求分析、系统设计、开发实施、运行维护及更新迭代等全生命周期环节，以实现各模块与各阶段的协同管理，并符合整体管控的需要。

### 4.2 稳定性原则

客户风险画像应用宜具备持续稳定运行的能力。系统的容量与性能可结合实际业务规模和技术应用水平进行规划，以支持服务的可靠性和连续性。

### 4.3 可拓展性原则

客户风险画像应用宜具备良好的可扩展性，能够适应不断变化的业务需求，并可融合新技术的应用，以支持未来的业务发展和技术演进。

## 5 需求分析

### 5.1 企业需求分析

企业开展基于大模型的客户风险画像应用建设，宜综合考虑业务发展、技术演进与管理效能之间的协同关系。在当前期货市场复杂多变的背景下，传统客户风险管理方式面临评估维度单一、响应滞后、智能化水平不足等问题。企业宜构建覆盖客户全生命周期的风险管理体系，提升对异常交易行为、潜在违约风险和跨市场风险传染的识别能力。通过引入大模型技术，企业可实现对结构化客户行为数据（如近半年日均期末权益、累计手续费、强平次数、异常交易行为次数、日均交易所风险度等）与外部动态舆情信息（通过 MCP 工具实时获取的监管处罚、市场传闻等）的深度融合与语义关联，推动风险识别从依赖单一阈值规则向基于多维证据链的动态智能评估演进。同时，企业宜关注系统建设的可扩展性与适配性，确保技术方案能够适应不同规模机构的算力条件与业务场景，支持后续迭代优化与跨机构推广应用。

### 5.2 用户需求分析

客户风险画像系统的最终服务对象包括风控管理人员、合规岗位人员及业务决策层，其核心需求在于获取及时、准确、可解释的风险评估结果。用户宜能够通过系统直观了解客户的整体风险等级（C1－C4）、各维度得分（贡献能力、风控能力、风险偏好、盈利能力）及关键风险来源（如“强平争议”“异常交易行为”“近期涉及负面舆情”），并支持基于风险变化趋势的分级预警与差异化管理策略制定（如重点维护高贡献低风险客户、限制高杠杆异常交易客户）。在交互层面，用户可期望系统提供可视化分析界面、风险归因报告及总体评价与业务建议等，提升决策效率。此外，用户对模型输出的透明性与可解释性有较高期待，宜通过量化评分体系、特征贡献度展示等方式增强模型可信度。系统还宜支持个性化配置功能，允许用户根据业务重点调整风险维度权重或设置特定监控规则，以满足不同岗位和管理场景下的差异化使用需求。

### 5.3 市场安全需求分析

随着金融科技的深入应用，金融市场的系统性风险识别与监管科技能力建设日益重要。基于大模型的客户风险画像应用宜服务于行业整体风险防控体系的完善，支持监管部门和行业机构实现更高效的风险监测与早期预警。该类应用宜具备对新型或隐蔽风险模式的潜在识别能力——例如，通过“异常交易行为次数”“强平投诉记录”“低保证金上调频次”等指标组合，辅助发现疑似操纵市场、频繁对倒或情绪驱动型交易行为；同时，融合客户微观行为（如胜算率、盈亏比、持仓集中度）与宏观舆情环境，

构建覆盖资金实力、操作稳定性、合规表现与外部声誉的多层次、动态化风险评估网络，提升对“表面正常但实质高危”客户的甄别能力。在安全合规方面，系统建设宜遵循数据隐私保护相关法律法规，采用去标识化、联邦学习等技术手段保障客户信息安全。同时，模型训练与推理过程宜建立审计机制，确保算法公平性、防止歧视性输出，防范技术滥用带来的新型合规风险。通过标准化建设，宜推动形成统一的技术接口、数据格式与风险分类标准，促进跨机构风险信息共享与协同防控机制的建立。

## 6 建设指南

### 6.1 建设架构

本节描述基于大模型的客户风险画像应用的典型系统架构，旨在为金融机构在系统设计与实施过程中提供参考。实际建设中，机构可根据自身业务规模、技术基础和安全要求对架构进行适应性调整。

客户风险画像应用宜采用分层、模块化的系统架构，主要包括前端交互层、数据处理层、提示工程层、大模型服务层和结果生成层。系统通过多层协同，实现从用户输入到风险评估结果输出的全流程支持。

典型架构流程见图6.1，其中包括：

a) 用户通过 PC 端或 App 端访问前端界面，提交待评估的问题描述及相关的客户数据信息；
b) 系统根据用户选择的功能，结合预设场景库，动态选取或生成适配的提示词；
c) 输入数据与提示词组合为标准化格式，提交至大模型进行推理；
d) 大模型根据输入，通过 AI Agent 的功能，调用 MCP 工具返回所需数据作为大模型输入补充；
e) 大模型输出初步结果后，系统对其进行结构化解析与后处理，生成风险解释分析与综合报告；
f) 最终结果以可视化方式返回前端，供用户查阅与决策参考。

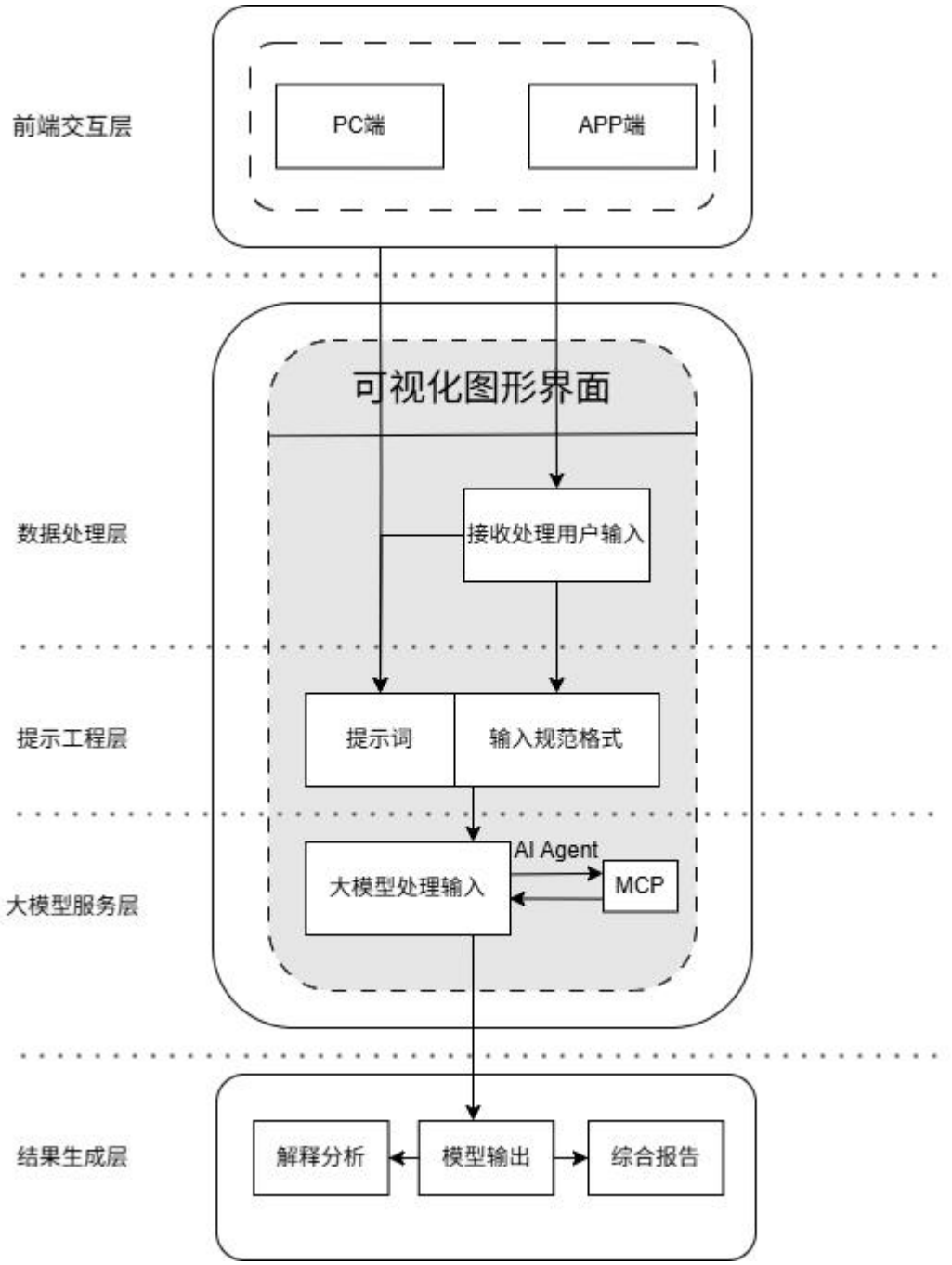


图 6.1　架构流程图

该架构支持人机协同、可解释性强，并具备良好的扩展性，可适配不同大模型服务部署模式（如本地部署、云服务调用等）。

## 6.2　基于大模型的技术选型与适配

### 6.2.1　大模型技术选型适配的一般性原则

在构建基于大模型的客户风险画像应用时，合理选择并适配大语言模型是保障系统有效性与实用性的关键环节。由于不同模型在性能表现、响应效率、输出质量等方面存在差异，机构宜结合自身业务需求、技术能力与部署环境，采用系统化的方法进行模型选型与适应性优化。

大模型的选型宜遵循的一般性原则包括：

a) 目标导向性：选型宜围绕客户风险画像的核心任务展开，如风险识别、归因分析、综合评级等，优先考虑在语义理解、逻辑推理和结构化输出方面表现良好的模型；

b) 多维评估：宜从多个维度对候选模型进行综合评估，包括但不限于：

1)响应速度：反映模型在实际业务场景中的实时性支持能力，指从发送完整提示词（含客户结构化数据与 MCP 工具调用指令）至接收到完整模型输出的时间（单位：秒），测试环境为统一 GPU，排除网络波动影响；

2)准确度：衡量模型输出结果与真实风险水平的一致性，基于人工标注的基准答案，计算模型输出风险等级（C1－C4）与基准一致的样本占比，以百分比表示；

3)回答质量：包括输出的完整性、可解释性、逻辑连贯性及是否包含指标关联分析。

c) 场景匹配性：根据应用场景（如实时预警、批量评估、人工辅助决策）选择合适类型的模型。例如，高并发场景宜优先考虑轻量级、低延迟模型；复杂分析任务可选用参数规模较大、推理能力更强的模型；

d) 可维护性与更新机制：宜关注模型的技术支持情况、版本迭代频率及社区活跃度，确保长期可用性。

### 6.2.2 大模型技术选型与适配举例

大模型的技术选型与适配评测举例见表6.1，该表提供了部分主流大模型在客户风险画像任务中的1000次测试表现示例，供参考。该数据基于特定测试集和提示词设计得出，实际效果可能因输入数据、部署方式和应用场景而异。

表 6.1大模型的技术选型与适配评测举例表

| 模型 | 响应速度/s | 准确度/% | 回答质量 |
|---|---|---|---|
| DeepSeek V3 | 33.85 | 54.2 | 回答质量一般，稳定性一般。评分过程中仅呈现各维度计算结果及简要评价，较少分析维度间的内在关联。评分标准存在波动现象。 |
| DeepSeek R1 | 285.98 | 64.8 | 回答质量与稳定性优于V3版本，输出中可体现一定的推理过程和解释说明。评分标准相对稳定，总分与最终评级较接近真实情况。 |
| QWEN3 | 151.48 | 76.3 | 回答质量与稳定性良好，具备推理痕迹、解释说明及综合评级能力。但在部分案例中倾向于将低风险客户误判为高风险。参数规模较大时响应较慢。 |
| Kimi K2 | 23.19 | 49.9 | 回答质量较低，稳定性差。通常仅输出最终评分结果，缺乏评分依据和指标分 |

| | | | |
|---|---|---|---|
| | | | 析，难以支持风险决策参考。 |
| Gemini 2.5 Flash | 30.92 | 80.7 | 回答质量高，稳定性良好。能够提供评分依据、指标关联性分析，并给出客户综合评价与管理建议，具备较强的应用参考价值。 |
| Gemini 2.0 Flash Lite | 8.18 | 67.5 | 回答质量中等，响应速度快。输出中包含一定解释内容，但多次测试下评分结果存在一定浮动。个别情况下可能出现超出预设评分范围的情况。 |

## 6.3 期货证券数据集构建

### 6.3.1 近半年日均期末权益

中文名称：近半年日均期末权益
英文名称：AVG_DAILY_EQUITY_LAST_6M
业务含义：客户在过去六个月中每日结算后的平均期末权益（单位：万元）。
制定依据：【内部标准】
代码值域：
实际数值>=0

### 6.3.2 近半年累计利息贡献

中文名称：近半年累计利息贡献
英文名称：TOTAL_INTEREST_CONTRIBUTION_LAST_6M
业务含义：客户在过去六个月中产生的利息收入总和（单位：万元）。
制定依据：【内部标准】
代码值域：
1 极高，10000<=半年累计利息贡献
2 较高，5000<=半年累计利息贡献<10000
3 中等，1000<=半年累计利息贡献<5000
4 较低，300<=半年累计利息贡献<1000
5 极低，0<=半年累计利息贡献<300

### 6.3.3 近半年累计总手续费

中文名称：近半年累计总手续费
英文名称：TOTAL_COMMISSION_LAST_6M
业务含义：客户在过去六个月中支付的交易、结算等各类手续费总额（单位：万元）。
制定依据：【内部标准】
代码值域：

1 极高，15000<=近一年累计留存手续费
2 较高，10000<=近一年累计留存手续费<15000
3 中等，5000<=近一年累计留存手续费<10000
4 较低，2500<=近一年累计留存手续费<5000
5 极低，0<=近一年累计留存手续费<2500

#### 6.3.4 结算后差交易所资金超10万次数（多头）

中文名称：结算后差交易所资金超10万次数（多头）
英文名称：COUNT_EXCESS_MARGIN_LONG
业务含义：客户在结算后，其账户资金超出交易所最低要求10万元以上的情形发生次数（多头方向）。
制定依据：【内部标准】
代码值域：
整数>=0

#### 6.3.5 结算后差交易所资金超10万次数（空头）

中文名称：结算后差交易所资金超10万次数（空头）
英文名称：COUNT_EXCESS_MARGIN_SHORT
业务含义：客户在结算后，其账户资金超出交易所最低要求10万元以上的情形发生次数（空头方向）。
制定依据：【内部标准】
代码值域：
整数>=0

#### 6.3.6 强平次数

中文名称：强平次数
英文名称：COUNT_FORCE_LIQUIDATION
业务含义：客户因保证金不足等原因被系统强制平仓的总次数。
制定依据：【内部标准】
代码值域：
整数>=0

#### 6.3.7 对合理强平处理有争议和投诉的次数

中文名称：对合理强平处理有争议和投诉的次数
英文名称：COUNT_DISPUTES_ON_FORCE_LIQ
业务含义：客户对系统执行的合理强平操作提出异议或正式投诉的次数。
制定依据：【内部标准】
代码值域：
整数>=0

#### 6.3.8 出现异常交易行为次数

中文名称：出现异常交易行为次数

英文名称：COUNT_ABNORMAL_TRADING
业务含义：客户被风控系统识别为存在频繁撤单、对倒、拉抬打压等异常交易行为的次数。
制定依据：【内部标准】
代码值域：
整数>=0

### 6.3.9 低保证金被上调的次数

中文名称：低保证金被上调的次数
英文名称：COUNT_MARGIN_INCREASE_DUE_TO_LOW
业务含义：因客户长期维持较低保证金水平，被系统或人工上调保证金比例的次数。
制定依据：【内部标准】
代码值域：
整数>=0

### 6.3.10 日均交易所风险度

中文名称：日均交易所风险度
英文名称：AVG_DAILY_EXCHANGE_RISK_RATIO
业务含义：客户在过去一段时间内，每日交易所计算的风险度（如持仓/保证金比）的平均值，反映整体杠杆水平。
制定依据：【内部标准】
代码值域：
1 高，95%<=风险度
2 较高，90%<=风险度<95%
3 中等，80%<=风险度<90%
4 较低，风险度<80%

### 6.3.11 近一个月盈亏比

中文名称：近一个月盈亏比
英文名称：PROFIT_LOSS_RATIO_LAST_MONTH
业务含义：客户在过去一个月中盈利还是亏损交易总额的百分比，用于衡量收益。
制定依据：【内部标准】
代码值域：
1 极高盈利，资产收益率(%)>=50
2 大幅盈利，资产收益率(%)>=20 且 资产收益率(%)<50
3 小幅盈利，资产收益率(%)>=5 且 资产收益率(%)<20
4 基本持平，资产收益率(%)>=-5 且 资产收益率(%)<5
5 小幅亏损，资产收益率(%)>=-20 且 资产收益率(%)<-5
6 大幅亏损，资产收益率(%) >= -50 且 资产收益率(%)<-20
7 极高亏损，资产收益率(%) < -50

### 6.3.12 近一年的胜算率

中文名称：近一年的胜算率
英文名称：WIN_RATE_LAST_YEAR

业务含义：近一年数据统计周期内浮动平仓盈亏>0的交易次数相对总交易次数的占比。
制定依据：【内部标准】
代码值域：
1 胜算率极高，60%<=胜算率
2 胜算率较高，55%<=胜算率<60%
3 胜算率高，50%<=胜算率<55%
4 胜算率较低，45%<=胜算率<50%
5 胜算率低，胜算率<45%

#### 6.3.13 持仓品种数

中文名称：持仓品种数
英文名称：COUNT_HOLDING_PRODUCTS
业务含义：客户当前或近期持有的不同交易品种（如期货合约、期权标的等）数量。
制定依据：【内部标准】
代码值域：
整数>=0

#### 6.3.14 客户年限

中文名称：客户年限
英文名称：CLIENT_TENURE
业务含义：客户自开户以来的存续年数，反映客户成熟度与忠诚度。
制定依据：【内部标准】
代码值域：
1 客户年限较长，五年及以上
2 客户年限一般，一年至五年
3 客户年限较短，一年以下

#### 6.3.15 客户属性

中文名称：客户属性
英文名称：CLIENT_TYPE
业务含义：客户在监管和业务分类下的主体类型。
制定依据：【内部标准】
代码值域：
1 natural_person（普通自然人客户）
2 corporate_client（法人）
3 special_legal_entity（特殊法人）
4 asset_management（资管户）
5 investment_fund（投资基金类客户）
6 bank_account（银行托管户）

### 6.4 数据安全

#### 6.4.1 数据分类与分级管理

基于大模型的客户风险画像应用在数据处理过程中，应严格遵循国家关于数据安全、个人信息保护及金融行业监管的相关法律法规，建立健全覆盖数据全生命周期的安全保障机制，确保客户数据的机密性、完整性与可用性。宜依据数据敏感程度、业务重要性及潜在风险影响，对用于客户风险画像的数据进行分类分级管理。

客户身份信息、交易明细、行为日志等属于高敏感数据，宜纳入重点保护范围；经脱敏或聚合处理后的衍生特征数据，可根据实际用途适当降低保护等级。数据分级结果应作为访问控制、加密策略和审计频次设定的依据。

#### 6.4.2 数据最小化与目的限定

数据采集宜遵循“最小必要”原则，仅收集与客户风险评估直接相关的字段，避免过度采集。所有数据的使用宜严格限定于风险画像建模、评估与预警等既定业务目的，不得用于无关场景或未经授权的二次利用。

#### 6.4.3 数据脱敏与匿名化处理

在将原始客户数据输入大模型前，宜实施有效的脱敏或匿名化处理，包括：

a) 对客户姓名、身份证号、手机号、账户号等直接标识符进行不可逆哈希、掩码或替换；

b) 对交易金额、持仓数量等敏感数值字段，在不影响风险判断的前提下，采用泛化、扰动或区间化处理；

c) 在批量处理或模型训练场景中，优先采用 k-匿名、差分隐私或联邦学习等技术，防止个体信息被反推或关联识别。

脱敏规则宜可配置、可审计，并定期评估其有效性。

#### 6.4.4 访问控制与权限管理

宜建立基于角色的访问控制机制，确保仅有授权人员或系统组件可访问特定级别的客户数据。风控人员、合规人员与系统管理员的权限宜分离，关键操作（如数据导出、模型调用日志查询）宜记录完整审计日志，并保留不少于6个月。

#### 6.4.5 第三方数据合作安全

若涉及外部数据源（如舆情数据、工商信息、征信接口等），应通过合同明确数据提供方的安全责任，并对其数据合法性、来源合规性进行审核。对外提供风险画像结果或特征数据时，应确保不包含原始个人身份信息，且符合监管关于数据出境与共享的规定。

### 6.5 大模型输入标准化与输出安全

#### 6.5.1 大模型输入标准化说明

在基于大模型的客户风险画像应用中，输入内容的质量与结构直接影响模型输出的准确性与一致性。为提升系统稳定性和可维护性，宜对大模型的输入进行标准化处理。标准化的核心在于将用户提交的问题与客户数据信息，与适配的提示词进行有效整合，形成结构清晰、语义明确的标准输入格式。

#### 6.5.2 输入标准化方法

为实现输入的标准化，建议采用“模板化拼接 + 结构化封装”的方法。

模板化拼接是将提示词作为基础模板，动态嵌入用户问题和客户数据信息。拼接顺序宜遵循“任务指令，输入数据，输出要求”的逻辑结构，确保模型优先理解任务背景。

结构化封装是在系统实现层面，宜采用结构化数据格式（如 JSON）对输入内容进行封装，便于系统解析、日志记录与后续审计。

### 6.5.3 输入内容构成

本应用宜采用非对话式交互模式，用户通过选择特定客户或填写必要业务参数后触发风险评估操作，系统自动生成完整输入并提交至大模型。该输入由以下两部分构成：

a) 提示词部分：用于引导模型执行特定任务，宜包含：

1)任务指令：明确说明本次推理目标为“客户风险画像”；

2)约束条件：包括风险评估规则（如评分维度、阈值定义）、业务侧重点（如侧重操作风险或盈利能力）以及输出格式要求（如结构化 JSON、包含归因说明等）。

b) 数据部分：以结构化形式提供客户相关特征信息，宜采用 JSON 格式，包含但不限于以下字段：

1)客户贡献类指标；

2)风控行为类指标；

3)风险偏好与盈利表现类指标；

4)其他辅助特征。

所有原始客户标识信息（如姓名、身份证号、账户号等）在输入生成前应已完成脱敏或匿名化处理，确保提交至大模型的数据不包含可直接识别个人身份的信息。

### 6.5.4 输出安全要求

本应用宜采用非对话式、任务驱动的交互模式：用户通过前端界面选择特定客户或输入必要业务参数后，点击“风险评估”按钮，系统即自动聚合脱敏后的客户特征数据，并结合预设提示词模板生成结构化输入，提交至大模型进行推理。整个过程无需人工构造自然语言提问，亦不依赖开放式对话上下文。

由于输入内容由系统自动生成且已对原始客户身份信息（如姓名、证件号、账户标识等）完成脱敏处理，大模型仅基于匿名化或泛化的风险特征进行分析，其输出结果天然不包含可识别个人身份的信息。因此，模型返回内容仅限于风险等级、维度评分、归因解释及管理建议等业务相关结论，不具备泄露客户隐私的技术条件。

## 6.6 风险评价指标体系构建

### 6.6.1 评估体系

客户用户画像风险评估体系划包括：

a) 客户贡献能力，反映客户为平台带来的经济价值，评估指标包括期末权益、利息收入与手续费贡献。该维度在综合评分中可参考赋予权重$w_1$。

b) 客户风控能力，用于识别客户操作风险水平，重点考察历史强平、异常交易行为及结算资金偏差等。该维度可参考赋予权重$w_2$。

c) 客户风险偏好，通过日均留仓风险度衡量客户杠杆使用与风险容忍程度。该维度可参考赋予权重$w_3$。

d) 客户盈利能力，通过盈亏比与胜算率评估其交易有效性与长期表现。该维度可参考赋予权重$w_4$。

e) 加分项，包括客户持仓偏好、客户属性及开户时间，用于在评分临界时进行辅助调整，不纳入基础评分，反映客户的潜在价值与平台粘性。

为支持不同机构根据自身业务重点灵活配置评估体系，各权重建议取值范围如下：

1) $w_1$（客户贡献能力）宜在 0.40～0.60 区间内取值，适用于以客户价值为导向的机构；
2) $w_2$（客户风控能力）宜在 0.20～0.40 区间内取值，风险敏感型机构可适当提高该权重；
3) $w_3$（客户风险偏好）与 $w_4$（客户盈利能力）宜分别控制在 0.05～0.15 区间内，二者之和不宜超过 0.25；

机构在实际应用中，可根据监管要求、业务发展阶段、客户结构特征等因素，在上述范围内动态调整权重，使权重之和1，并通过回溯测试验证权重配置对风险识别有效性的支撑作用。建议定期评审权重合理性，确保评估结果与风险管理目标一致。

### 6.6.2 用户画像等级划分

本标准依据多维度评分体系，将客户从低到高划分为四个画像等级：

a) 一级：高风险或低价值群体，行为不稳定，需重点预警管理并限制敏感交易权限。
b) 二级：风险偏好较高或历史存在异常行为，需纳入日常监测并设置响应机制。
c) 三级：整体表现良好，但在部分维度存在波动，仍具备较高发展潜力，适宜进行策略引导与信用提升。
d) 四级：代表低风险、高价值群体，贡献大、行为稳健，盈利能力与交易规范性俱佳，是重点服务与资源配置对象。

## 6.7 技术实现路径

### 6.7.1 实现路径示例

图6.2围绕客户风险画像的落地应用，给出一套端到端技术实现路径：首先基于 6.3 节“期货证券数据集构建”中系统定义的15项核心客户数据元（如近半年日均期末权益、强平次数、客户属性等），形成结构清晰、语义明确的内部特征输入；在此基础上，筛选并部署专用的风险识别大模型；同时集成一组面向金融风控任务的MCP（Model Calling Protocol）工具，实现动态工具调用，并开发配套前后端系统；最终通过标准化提示词工程与输出模板，生成具备业务可解释性的风险评估结果。

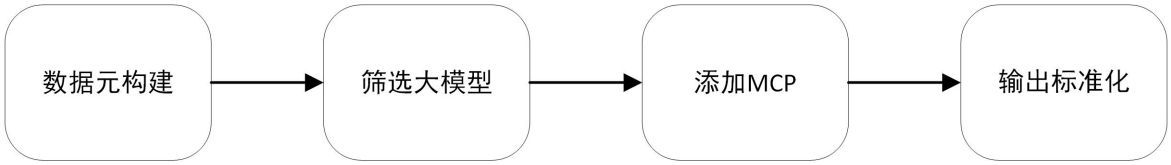


图 6.2 技术路径图

### 6.7.2 构建结构化客户特征数据集与大模型输入

以内部业务标准为基础，系统梳理并定义了15项核心客户数据元，涵盖资金实力（如“近半年日均期末权益”“近半年累计利息贡献”）、交易行为（如“异常交易行为次数”“强平次数”）、风险承受能力（如“日均交易所风险度”“低保证金被上调次数”）及客户属性（如“客户年限”“客户属性”分类编码）等多个维度。所有字段均采用统一单位（万元或次数）和明确业务含义，确保输入数据的一致性与可计算性，为模型提供高质量、高信噪比的结构化输入基础。

优化后大模型输入的提示词样例如表6.2所示：

**表 6.2 大模型提示词样例**

| |
|---|
| **【任务描述与调用工具】**你是一个证券期货领域的客户评估专家 第一步，你需要读取文件，地址为：xxx 第二步，你需要先使用风险等级工具得出初评风险并输出， 第三步，获取30天内的金融舆情信息，然后利用舆情信息和以下 |

| |
|---|
| 维度给客户评分得出最终评分数据，再根据累计总分划分为C1（[0,30)分）、C2（[30,50)分）、C3（[50,80)分）、C4四个等级（[80,115]分），并简要给出你的具体判断依据。 |
| **【划分原则】**其中，C4的客户质量最高，其余等级优质度逐级递减： 1.客户贡献能力：客户近半年的日均期末权益、累计利息贡献、累计总手续费，最高贡献客户的判断阈值分别为300、1、15万元，未达阈值按比例换算，主要看日均期末权益和累计总手续费，总分在0-50分。 2.客户风控能力：包括结算后差交易所资金、强平次数、对合理强平处理有争议和投诉的次数、出现异常交易行为和低保证金被上调的次数，只要出现上述行为就要酌情扣分，总分在0-30分。 3.客户风险偏好：包括客户的日均交易所风险度，超过55%酌情扣分，总分在0-10分。 4.客户盈利能力：包括客户近一个月盈亏比和近一年的胜算率，总分在0-10分。 5.客户加分项：额外加分项，包括客户的持仓品种数，客户属性，客户年限等，酌情加分，老客户判断阈值为5年以上，只加5分，最多15分。 用户信息你需要读取文件中的以下内容： 近半年日均期末权益：{qmqy}万元，近半年累计利息贡献：{lxgx}万元，近半年累计总手续费:{sxf}万元，结算后差交易所资金超过10万次数:{num_cs}，结算后差交易所资金超过10万次数:{num_ls}，强平次数：{num_qp}，对合理强平处理有争议和投诉的次数：{num_ts}，出现异常交易行为次数:{num_yc}，低保证金被上调的次数:{num_st}，日均交易所风险度:{fxd}，近一个月盈亏比:{ykb}，近一年的胜算率:{ssl}，持仓品种数:{num_cc}，客户属性:{sf}，客户年限:{sc}年。 |
| **【用户选择】**你需要评判的用户为：{用户id} |

### 6.7.3 大模型选型与适配

基于课题前期对主流大模型的对比测试（详见第6.2节），本示例在风险画像任务中优先选用具备强推理能力、高输出稳定性且支持MCP工具的开源模型作为基座。选型遵循“目标导向、场景匹配、可维护性”原则，重点考察模型在金融术语理解、数值敏感性、多因子归因分析等方面的表现。实测表明，依赖高质量提示工程（如结构化输入模板、风险维度显式引导）即可满足多数客户风险等级判定需求，且输出格式高度符合后续自动化解析要求，有效支撑了端到端流程的稳定运行。

### 6.7.4 集成 MCP 工具实现实时舆情融合与风险初评

为增强大模型对动态业务环境的感知与计算能力，本示例在技术实现中引入 MCP工具机制——即一组可被大模型按需调用的标准化功能模块，用于读取外部数据、执行领域计算或接入第三方服务。在客户风险画像生成流程中，我们重点集成了三类 MCP 工具：

一是文件解析 MCP，支持自动读取客户上传的 csv 或 Excel 格式的尽调报告、财务报表等文件，提取关键字段（如净资产、历史违约记录）并结构化输出，作为模型输入的补充；

二是TrendRadar 舆情 MCP，在推理前自动调用 TrendRadar 金融舆情 API，基于客户名称、法人代表或持仓品种，检索近30天内的负面事件（如监管处罚、诉讼、异常交易传闻等），并返回结构化的事件摘要与情感倾向标签；

三是风险初评 MCP（已封装为内部 PyPI 包），根据 6.3 节定义的15项客户指标（如强平次数、日均风险度、胜算率等），按预设规则引擎初步计算四维风险分值并给出初始风险等级建议。

上述 MCP 工具的输出结果统一转换为自然语言片段，嵌入大模型提示词上下文。例如：“【舆情】期货市场成交量创历史新高，机构资金流入明显；【初评】风控能力维度得分62，建议等级C3”。通过这种“内部数据 + 外部信号 + 规则初判”的多源融合方式，模型不仅能识别显性高风险行为，还能有效捕捉“交易正常但舆情恶化”或“指标临界但历史稳定”等复杂情形，显著提升风险画像的全面性与前瞻性。

### 6.7.5 输出标准化、可审计的风险评估结果

为确保风险画像结果具备业务可用性、逻辑透明性与监管合规性，本研究设计并实现了结构化、分层次的模型输出机制。最终评估报告严格遵循统一模板，包含以下核心组成部分：

一是过程可追溯的执行日志，清晰记录 MCP 工具调用序列（如文件读取、舆情查询、风险初评等），体现“数据从何而来、计算如何触发”的操作链路；

二是双版本评估结果：

首先，基于预置规则引擎（risk_judge MCP）生成初步风险评分与等级（C1－C4，C4为最优）；

其次，展开四维量化评分体系（客户贡献能力、风控能力、风险偏好、盈利能力）及加分项（持仓多样性、客户属性、客户年限），每项均列明原始指标值、评分规则、得分依据与小计结果；

三是综合判断摘要，以自然语言形式归纳客户优势与需关注点，并给出明确的总体评价与业务建议（如“重点维护”“加强盈利指导”等）；

四是全链路审计支持，所有输入数据（包括 6.3 节定义的15项客户指标）、外部舆情快照、MCP调用响应及模型推理中间结果均被完整记录，确保任意评估结论均可回溯验证。

模型输出样例如表6.3所示：

**表 6.3 大模型输出样例表**

| |
|---|
| **【工具调用】**我将按照您的要求逐步完成客户评估任务。首先读取文件数据。<br>调用filesystem：read_text_file的MCP成功<br>现在我来分析第一位用户的数据。从文件中可以看到第一位用户的信息：<br>调用客户风险计算：risk_judge的MCP成功<br>现在我需要获取30天内的金融舆情信息来完成风险最终评估。<br>调用trendradar：search_news的MCP成功<br>让我尝试搜索更近期的金融相关舆情信息：<br>调用trendradar：get_latest_news的MCP成功<br>让我尝试触发一次爬取来获取最新的金融舆情数据：<br>调用trendradar：trigger_crawl的MCP成功 |
| **【评判依据】**现在我基于获取到的舆情信息和用户数据来完成风险最终评估。根据当前舆情环境（整体较为稳定，没有重大金融风险事件），我将对第一位用户进行详细评分：<br>风险初评结果（基于风险等级MCP）<br>风险评分：81.98分<br>风险等级：C4（客户质量最高）<br>风险最终评估结果（基于多维度评分）<br>1. 客户贡献能力评分（0-50分）<br>近半年日均期末权益：214.78万元，达到最高贡献阈值300万元的71.6%，得分：35.8分<br>累计利息贡献：0.69万元，达到最高贡献阈值1万元的69%，得分：3.45分<br>累计总手续费：10.82万元，达到最高贡献阈值15万元的72.1%，得分：7.21分<br>小计：46.46分<br>2. 客户风控能力评分（0-30分）<br>结算后差交易所资金超过10万次数：0次（优秀），得分：10分<br>结算后差交易所资金少于10万次数：1次（轻微问题），扣1分，得分：9分<br>强平次数：1次（有风险），扣5分，得分：5分<br>对合理强平处理有争议和投诉次数：1次（有争议），扣3分，得分：7分 |

| |
|---|
| 出现异常交易行为次数：0次（优秀），得分：5分<br>低保证金被上调次数：2次（有风险），扣2分，得分：8分<br>小计：34分，调整为满分30分<br>3. 客户风险偏好评分（0-10分）<br>日均交易所风险度：39%（低于55%阈值，表现良好），得分：9分<br>4. 客户盈利能力评分（0-10分）<br>近一个月盈亏比：0.33（表现一般），得分：4分<br>近一年的胜算率：55%（表现良好），得分：6分<br>小计：10分<br>5. 客户加分项评分（0-15分）<br>持仓品种数：7个（品种丰富），得分：5分<br>客户属性：bank_account（银行账户，信誉较好），得分：5分<br>客户年限：3.72年（未达到5年老客户标准），得分：0分<br>小计：10分 |
| **【结果与评价】**最终评分汇总<br>总分：46.46 + 30 + 9 + 10 + 10 = 105.46分<br>最终等级：C4级（客户质量最高）<br>判断依据总结<br>优势方面：<br>贡献能力突出：日均权益214.78万元，手续费贡献良好<br>风控表现优秀：结算风险控制很好，无大额资金缺口<br>风险偏好适中：交易所风险度39%，风险控制得当<br>品种配置丰富：持仓7个品种，分散投资<br>账户类型优质：银行账户，信誉度高<br>需关注方面：<br>强平记录：有1次强平记录和1次争议投诉<br>盈利波动：近期盈利表现一般，盈亏比较低<br>客户年限：3.72年，尚未达到老客户标准<br>总体评价：该客户属于高质量的C4级别客户，具有较强的资金实力和较好的风险控制能力，是值得重点维护的优质客户。建议加强盈利指导，提升交易策略水平。” |

## 7 系统测试

### 7.1 业务功能测试

业务功能测试宜覆盖从前端用户交互到后端模型推理及结果生成的全流程，重点评估系统在真实业务场景下的功能完整性与逻辑合理性。宜检验用户通过界面提交客户数据和评估问题的能力，系统对输入信息的解析与处理准确性，提示词的匹配或生成是否适配任务需求，以及输入内容能否正确组合并传递至大模型。模型输出宜包含风险评分、关键影响因素分析、综合评级结论及管理建议等必要内容，且结果宜符合预设的评分逻辑与风险等级划分标准。同时，应评估生成的综合报告在结构化程度、语言可读性和业务贴合度方面的表现，确保其具备辅助决策的价值。测试过程中宜设计覆盖不同

客户类型、风险特征和数据质量水平的用例，验证系统在多样场景下的适应性与输出一致性，保障其在实际应用中的可用性与可信度。

### 7.2 安全测试

为确保客户信息与系统运行的安全性，宜重点关注数据生命周期各环节的安全防护能力，包括数据在传输和存储过程中的加密机制是否健全，客户敏感信息是否在进入模型前完成有效脱敏，防止隐私泄露。系统宜具备完善的访问控制策略，确保仅授权用户和系统组件可调用核心功能，防止越权操作。宜检查API接口是否存在注入攻击、跨站脚本等常见安全漏洞，并评估大模型服务调用过程中对提示词注入、越狱攻击等对抗性行为的防御能力。同时，宜验证系统日志是否完整记录关键操作，是否支持对异常行为的追溯与审计。输出内容也宜纳入安全审查范围，防止生成包含歧视性、误导性或泄露训练数据的信息。安全测试宜结合自动化工具与人工检测手段，定期执行并形成闭环整改机制，提升系统的整体安全防护水平。

### 7.3 压力测试

压力测试宜模拟多用户并发提交风险评估请求、批量导入大规模客户数据等典型高负载场景，观察系统整体响应能力与资源使用情况。重点监测模型推理服务的平均响应时间、请求成功率、吞吐量及错误率等关键指标，识别可能存在的性能瓶颈。同时，宜评估系统在持续高负载运行下的稳定性，检验其是否出现响应延迟加剧、服务中断或资源耗尽等情况，并验证系统是否具备良好的容错与自我恢复能力。测试过程中宜监控服务器的CPU、内存、GPU等资源占用状态，为后续的容量规划、资源扩容或架构优化提供依据。

## 8 系统上线

### 8.1 上线前置条件

在启动系统上线程序前，宜确保各项技术、功能与组织准备已基本就绪。系统应已完成与现有业务平台的集成对接，支持客户数据的安全接入与结果回传。若涉及加密传输或敏感信息处理，相关安全机制宜已部署到位，并通过内部或第三方机构的安全合规性评估。系统核心功能模块，包括客户信息解析、提示词生成、大模型推理、风险评分计算与报告输出等，宜通过完整的业务功能测试、安全测试和压力测试，验证其准确性、一致性和性能满足实际业务需求。用于调用系统的前端界面或API接口宜完成基础功能验证，确保交互正常、无阻塞性缺陷。相关技术运维人员宜完成岗位培训，熟悉系统架构、操作流程与应急处置规范，明确职责分工。同时，宜对业务管理、客户服务等相关人员开展必要的培训，使其了解系统功能边界、输出含义及对客服务中的使用注意事项，具备应对客户咨询与反馈的基本能力。

### 8.2 上线前准备

系统上线前宜制定周密的准备计划，以降低实施风险。宜明确参与上线的技术、业务及管理团队成员，合理安排实施操作与上线后监控值守任务，并建立向决策层报告进展的机制。宜编制详细的上线实施方案，涵盖具体的实施步骤、时间安排、数据迁移或初始化策略、业务验证方法及测试用例等内容，并提前完成关键路径的模拟验证。同时，宜制定应急回退方案，明确触发回退的条件（如模型输出异常、服务不可用、数据错误等）、回退操作流程、责任主体及回退后的状态检查方式。

### 8.3 上线实施

上线实施宜严格按照既定方案有序推进。由指定人员按步骤执行系统切换、配置更新或服务发布操作，在预定窗口期内完成部署。操作完成后，宜及时调整监控策略，确保关键指标（如请求响应时间、错误率、模型调用成功率、资源利用率等）被有效采集。宜立即开展基础业务验证，通过测试数据调用系统全流程，确认输入解析、模型推理、结果生成与展示等功能正常，输出内容符合预期逻辑且数据准确。验证通过后方可开放正式业务流量。整个实施过程宜保持沟通畅通，实时记录操作日志与异常情况，确保过程可追溯。

### 8.4 上线跟踪

系统上线后，宜安排专人持续监控其运行状态与输出质量。监控范围宜包括服务可用性、响应性能、模型输出的一致性与合理性，以及与其他系统的交互稳定性。宜主动收集业务人员和客户的使用反馈，重点关注输出结果的可解释性、业务贴合度及潜在偏差问题，并据此制定优化改进措施。当系统出现异常（如输出偏离阈值、服务中断、性能劣化等）并达到预设告警指标时，宜启动应急评估流程，及时向管理团队报告。由管理团队综合评估影响范围与持续时间，决策是否继续运行或启动回退程序。问题处置后，宜组织复盘分析，深入排查根本原因，必要时引入外部专家支持，形成整改闭环，防止同类问题重复发生。

## 9 运行保障

### 9.1 日常操作

为保障基于大模型的客户风险画像系统稳定运行，宜建立日常运维机制。相关服务组件宜纳入常规巡检范围，确保功能可用。模型服务、数据接口及核心处理模块的启停操作宜遵循标准化流程。客户数据输入、模型推理记录等关键运行数据宜定期备份，并按安全要求妥善存储。宜定期开展系统安全检查，及时识别并修复潜在漏洞。运维人员宜严格执行操作规程，完整记录系统日志与操作行为，并定期分析运行数据，发现异常趋势。运行过程中如发生故障或异常，宜详细记录时间、现象、处理过程及相关原始信息，留存备查。

### 9.2 系统监控

系统宜纳入统一监控体系，覆盖基础设施、应用性能与业务指标。重点监控模型调用响应时间、请求成功率、输出合规性、数据处理吞吐量等关键参数，设置阈值告警机制，及时发现并响应异常。宜定期分析监控数据，评估系统健康状态与业务适配性。

### 9.3 升级变更

系统升级或变更宜遵循规范流程，涵盖模型更新、提示词调整、接口优化等场景。变更前宜充分测试验证，重大变更需制定回退方案，确保不影响业务连续性与输出稳定性。

## 10 应急管理

### 10.1 应急准备

宜建立应急管理体系，明确突发事件的组织职责与响应机制，定期更新关键人员联系方式。针对模型失效、服务中断、数据异常、安全攻击等典型场景制定应急预案，并定期开展演练与培训，提升处置能力。

### 10.2 应急处置

当发现可能影响风险评估准确性的隐患时，宜立即核实并采取防范措施。发生严重故障或安全事件时，宜迅速启动应急预案，优先恢复核心功能，控制影响范围。事件处置情况宜按规定及时、准确上报。事件结束后，宜组织复盘分析，查明原因，落实整改措施，并形成总结报告，持续完善系统韧性。