



2025 AI+ Development
Digital Summit

AI+ 研发数字峰会

拥抱AI 重塑研发

05/23-24 | 上海站



2025 AI+研发数字峰会

拥抱AI 重塑研发 AI+ Development Digital Summit

下一站预告

08/08-09 | 北京站

11/14-15 | 深圳站



查看会议详情

北京站论坛设置

大模型和 AI 应用评测

智能存储与检索技术

下一代知识工程

AI+ 金融业务创新

智能需求工程

智能体与研发效率工具

AI 产品运营与出海策略

大模型安全与对齐

大模型应用开发框架与实践

智能体经济 (Agentic Economy)

智能测试工具的开发与应用

具身智能与机器人

代码生成及其改进

AI+ 新能源汽车

AI 前沿技术探索与实践

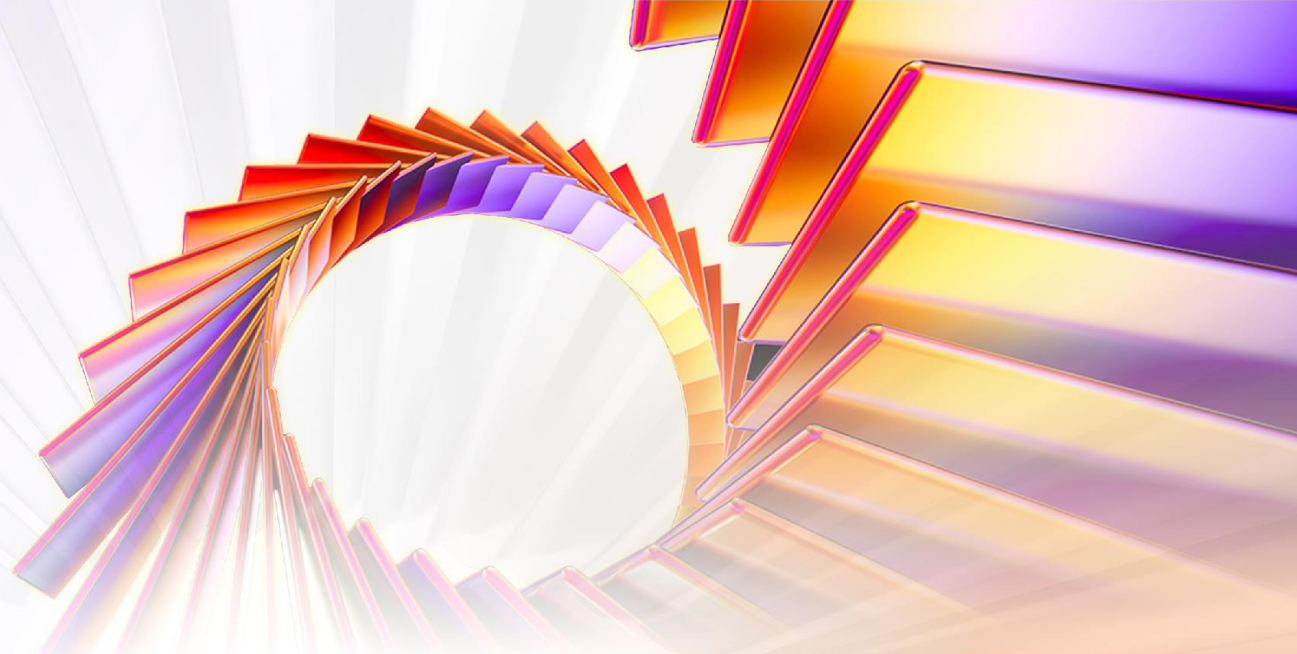


| 05/23-24 | 上海站

2025 AI+ Development
Digital Summit

AI+研发数字峰会

拥抱AI 重塑研发



线上质量保障之 客诉处理体系的智能化实践

吕冬冬 | 快手



吕冬冬

快手商业化线索营销方向测试负责人

南京邮电大学大学计算机系专业硕士毕业，毕业后入职华为、百度、贝壳等公司，现任快手商业化线索营销方向测试负责人。现负责服务端质量保障体系、FT效能提升等专项建设，同时负责AI智能测试技术在质量保障侧的探索、设计与落地，包括智能测试用例生成、需求风险智能分级、AI数字员工、客诉&舆情数据智能挖掘、模型评测等课题，助力产研流程优化、效率提升。

目录

CONTENTS

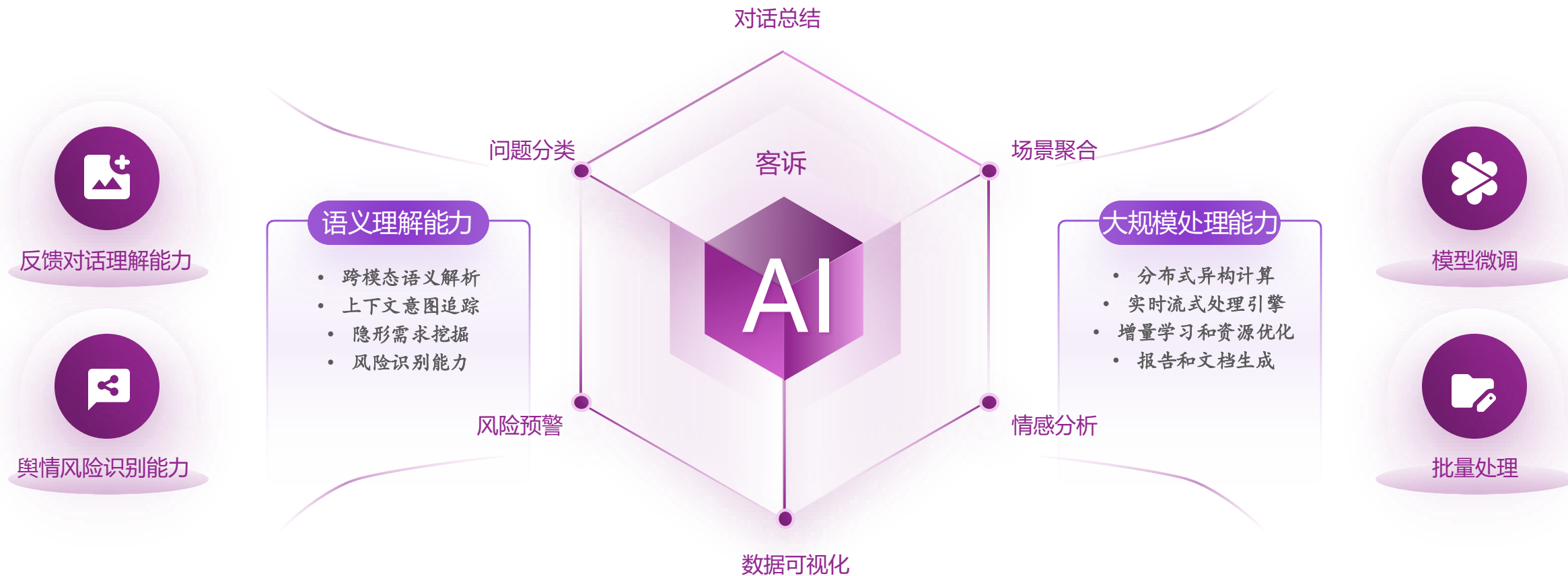
- I. 背景
- II. 问题/痛点
- III. 解决思路/整体方案
- IV. 具体实现/技术实践
- V. 总结与展望

PART 01

AI大模型潜力、客诉背景

▶▶ 1.1 背景-AI大模型潜力

- 大模型在客诉方向能做什么，帮助我们解决哪些问题
- 自然语言理解、多模态融合、智能决策能力为质量保障提供新路径。



▶ 1.2 背景-客诉背景

问题类型

产品问题

在功能设计、用户体验、以及内容生态等方面存在不足，导致用户体验或业务目标下降。

例如：内容质量差、产品使用逻辑复杂、互动功能不完善等

技术问题

代码缺陷或系统故障导致的应用运行异常，直接影响用户使用稳定性和流畅性。

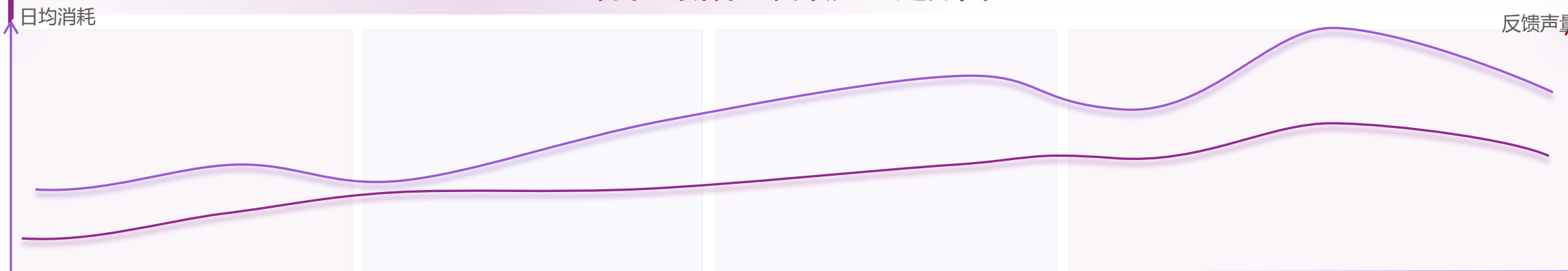
例如：应用崩溃、异常弹窗、性能问题、功能逻辑错误等

风控问题

在风险识别、内容审核、用户行为治理等方面的漏洞，可能会导致法律风险或用户利益受损。

例如：内容安全风险、用户违规行为、技术风控漏洞等等

业务日均消耗-反馈声量趋势图

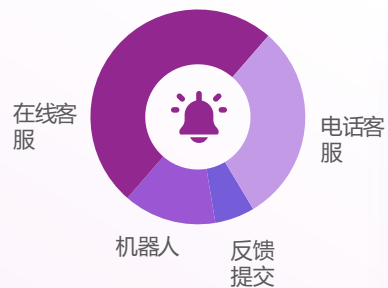


1.3 背景-客诉背景

客诉现状

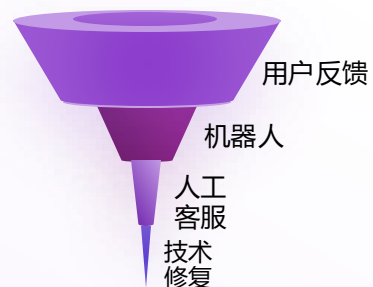
反馈问题

客诉声量大



用户反馈量级：大型业务每天有大量用户反馈。

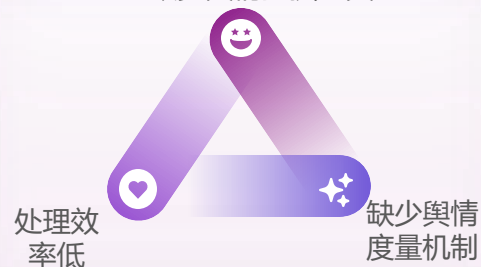
客诉解决率低



客诉解决率：机器人回复准确率较低，人工回复准确率受专业性和疲劳程度影响。

舆情问题

缺少舆情挖掘工具



商业化重大活动缺少站外舆情检测手段，并且噪音大，识别准确率低。

工单问题

工单达成率低



工单量级随业务发展而同步增长

核心问题

不关注

各FT主要依赖人工对接、KIM、工单等方式处理用户反馈，缺乏对用户反馈的主动关注

无感知

缺少主动挖掘反馈的能力，问题感知的时效性较差

不准确

客服对业务缺乏了解，无法精准分类识别商业化业务的问题，用户的真实情感和诉求没有展现

少工具

人工处理用户反馈的效率较低，人工处理的方式不仅耗时，而且容易出现遗漏和错误



可支持大型运营活动舆情，缺少对站外舆情的关注，有舆情发酵风险，可能会导致公司信誉和用户体验下降等连锁反应

1 工单跟进Oncall人力多

在客诉工单流转PUSH动作上需要oncall同学轮换，需要大量精力进行盯屏、PUSH流转。

2 客诉问题归因慢

每周会人工根据客诉底表数据，进行归因统计和分析，每次都需要耗费较多人力，而且归因分类不深入。

3 产运答疑不及时

工单分配依赖人工，并且无有效异常工单感知能力，客诉达成率低。



PART 02

业务问题和痛点

▶ 2.1 业务问题和痛点-客诉处理效率低

01

人工处理慢

问题表现：传统客诉处理依赖人工分类，需手动筛选投诉类型、优先级及责任部门，导致响应延迟。例如，客服需从海量投诉中逐条判断是否为技术问题或服务问题，耗时费力

原因：缺乏智能分类工具，分类标准模糊（如“系统卡顿”可能涉及研发或运维部门），人工判断易出错且效率低下

02

排查根因耗时长

问题表现：技术类客诉需跨部门协作，信息传递断层导致根因定位困难。

原因：缺乏统一数据平台与自动化工具，依赖人工排查日志或复现问题，且部门间权责划分不清

03

机器客服处理不准确

问题表现：机器人客服受限于知识库更新滞后与语义理解能力不足，常误判用户诉求。

原因：传统规则引擎无法处理模糊描述（如“卡顿”“闪退”），且缺乏多模态数据分析能力（如未关联用户操作日志）



▶ 2.2 业务问题和痛点-客诉数据分析困难



PART 03

解決思路

▶▶ 3.1 核心目标

实现“客诉-分析-修复”全链路自动化，全面提升质量保障效率

实现路径

通过融合大模型、自然语言处理（NLP）等前沿技术，构建智能评测体系，实现对内外部舆情反馈的全流程自动化处理，从数据采集、分析预警到问题修复，持续改善用户体验、降低舆情风险，推动业务价值稳步增长

全

全面覆盖

- 利用多渠道采集用户反馈，整合多维度数据，确保数据完整性与实时性。

快

快速响应

- 借助 AI 技术提升问题处理效率，建立实时预警机制，实现对舆情事件的即时响应。

准

准确分析

- 应用大模型和 NLP 技术，优化问题分类和情感分析，确保评测结果精准可靠。

稳

稳定运行

- 通过技术架构优化与持续监控，确保平台长期高效、稳定、可持续运行。



3.2 解决方案



多渠道接入

整合用户反馈、社交媒体、投诉举报等多渠道数据，确保反馈信息来源全面。

大模型优化

利用大模型进行智能总结、分类和聚类，自动提取关键问题并提高识别准确性。

多维度告警

设置高危问题预警、日环比预警、新增聚类预警等多层次告警策略，实现潜在舆情风险的早发现、早预警。



3.3 整体架构

展示层

通过可视化界面展示舆情分析结果。

服务层

利用大模型、情感分析等技术对反馈问题进行智能分类和场景聚合。

数据层

通过采买数据、反馈挖掘等多渠道实时获取和清洗数据。

存储层

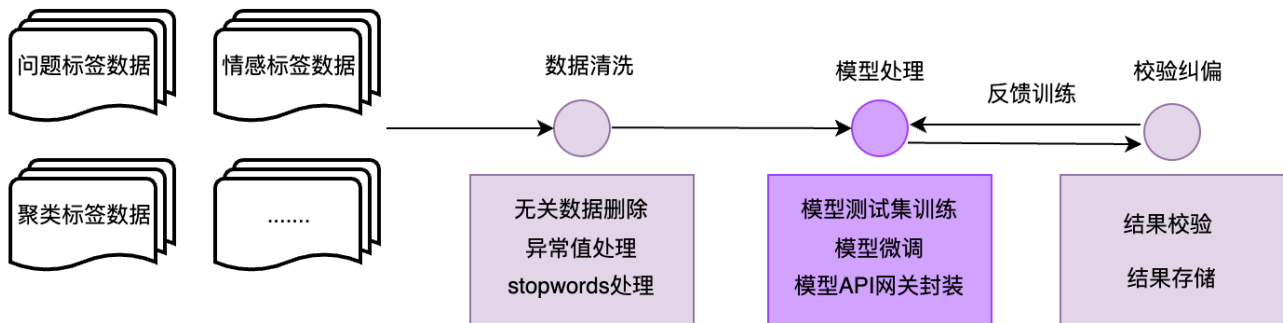
使用MySQL数据库存储用户反馈、分析结果等信息。



3.4 架构设计-模型层

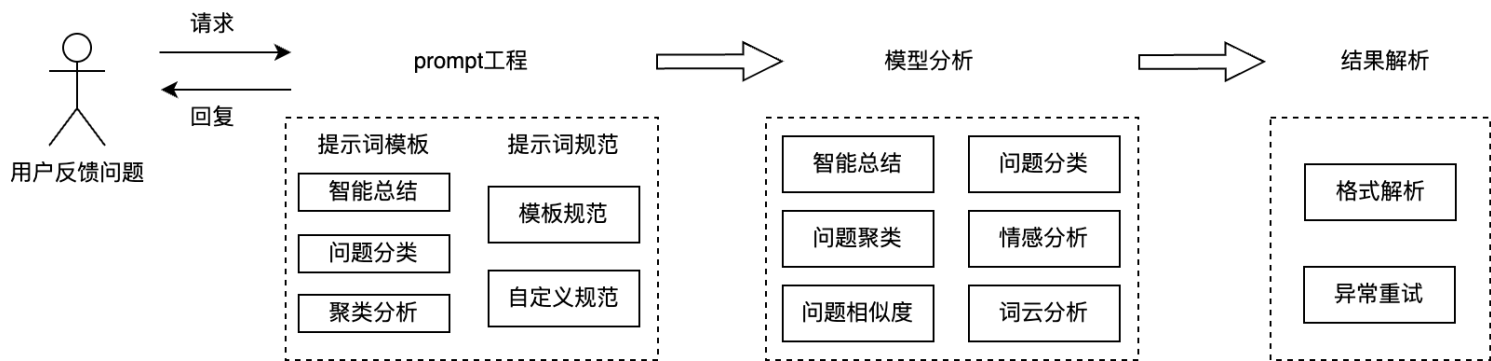
➤ 模型侧主要完成基于历史数据的**离线训练**和实时问题的**在线调用**

离线训练流程



离线流程

在线流程



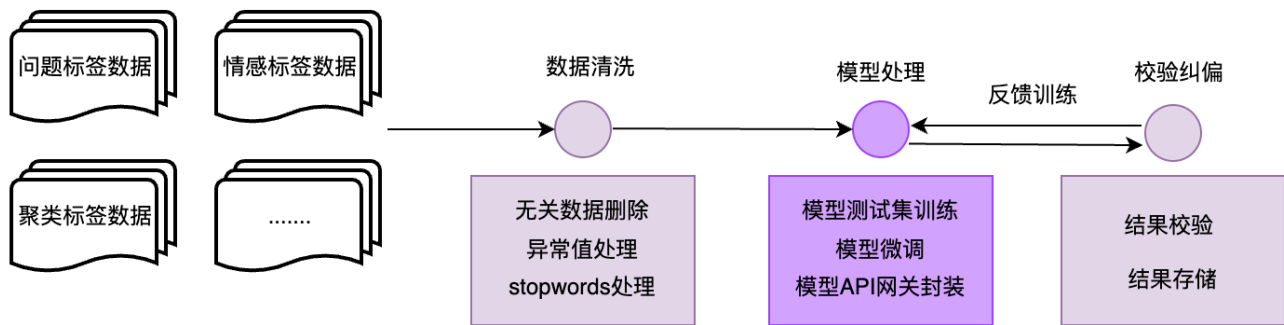
- **数据清洗**: 采集数据，完成对应标签确认；缺失值、异常值处理；数据特征处理
- **模型处理**: 对分类和聚类两个特定任务进行模型微调
- **校验纠偏**: 采用点赞、点踩及人工校验等方式，检查分类结果，识别并纠正潜在错误分类
- **反馈训练**: 根据纠偏的正确数据，反哺于微调模型，对模型再次调优



▶ 3.5 架构设计-模型层

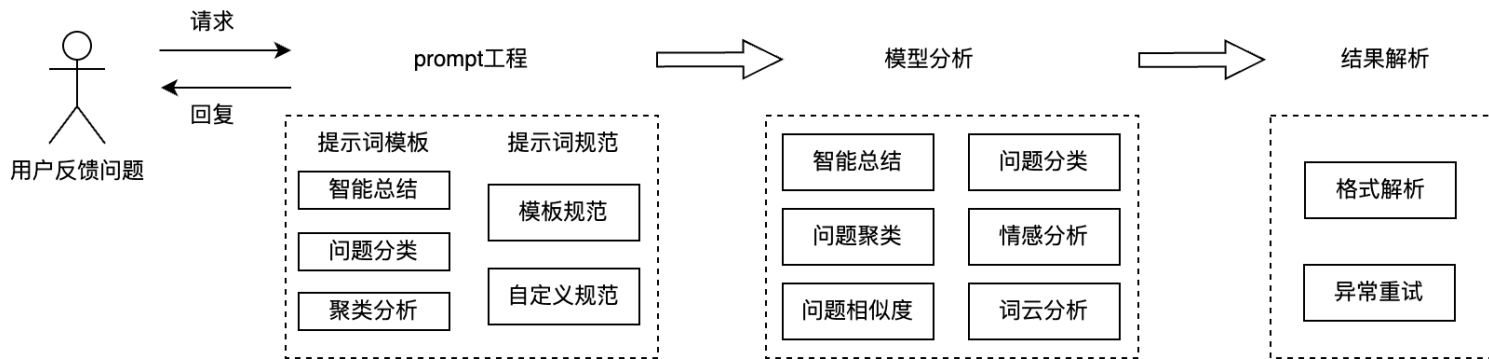
➤ 模型侧主要完成基于历史数据的**离线训练**和实时问题的**在线调用**

在线调用流程



离线流程

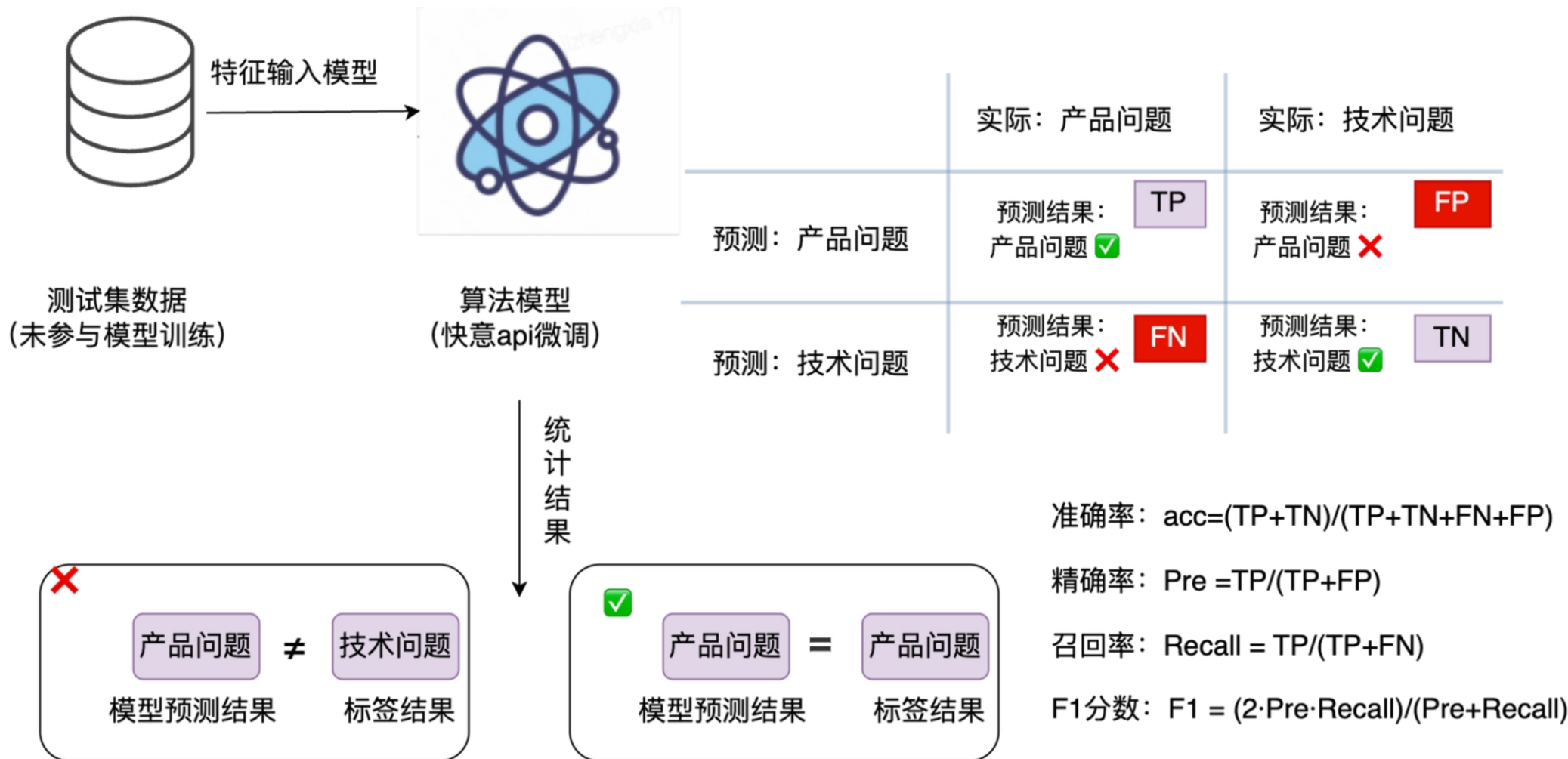
在线流程



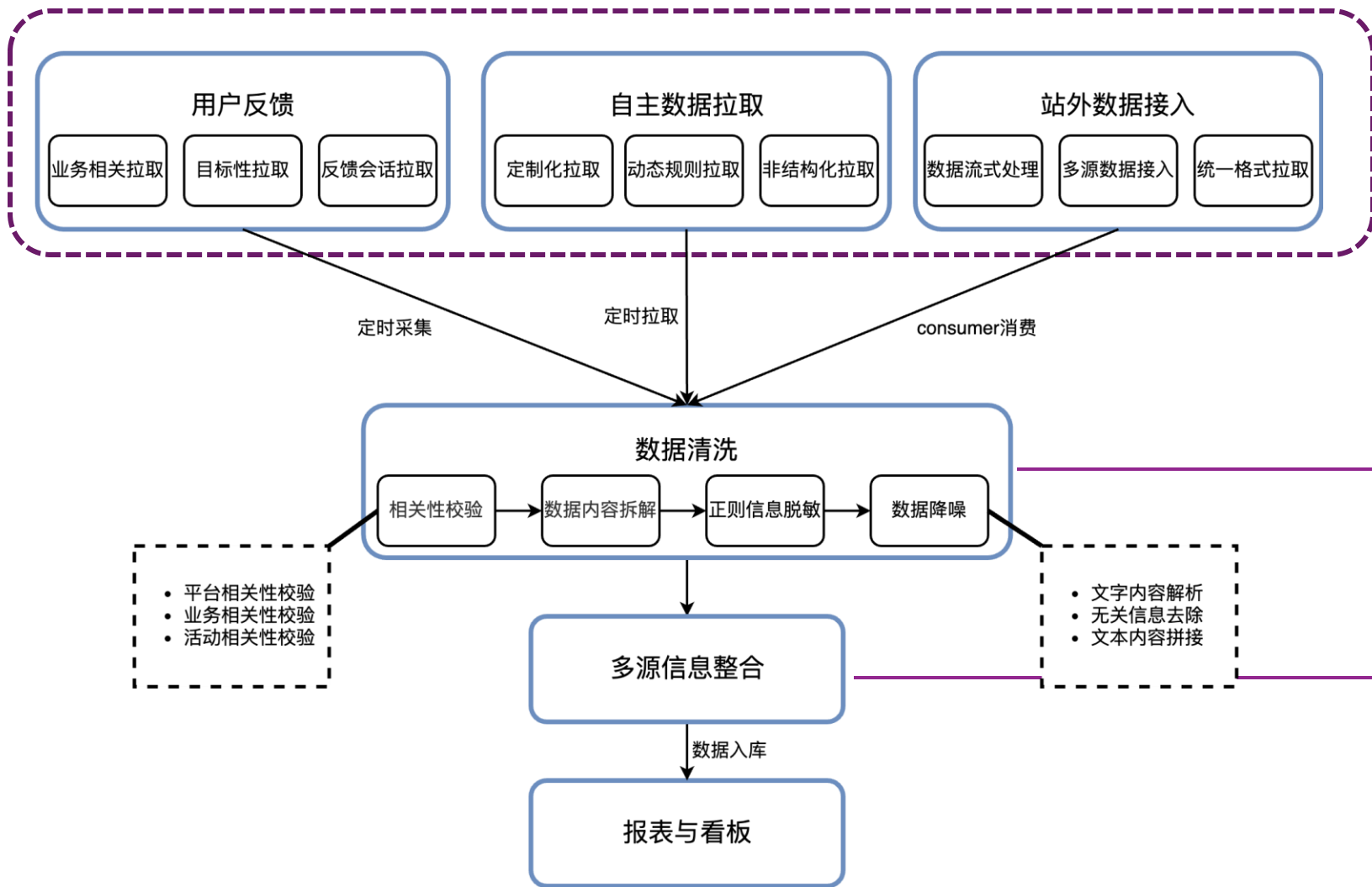
- **prompt工程:** 制定提示词模板、生成常用prompt, 可在线修改
- **模型分析:** 完成用户情感分析、对话内容智能总结、问题分类并对相似问题聚类
- **结果解析:** 多模型输出进行格式解析, 模型输出及解析失败时完成异常重试



▶▶ 3.6 模型衡量指标



3.7 架构设计-数据层



接入多源数据，提升有效性

实现了规则配置化，支持后续的业务、活动的变更。

用户反馈：关注目标导向数据

站外数据：丰富数据源、降低接入成本

数据清洗逻辑收敛

统一、收敛的数据校验、清洗、加密、对象构建规则，清洗、构建方法抽象，方便后续其他来源、方式的数据接入。

数据整合、库表涉及

不同来源的数据的统一整合后给到下游工程、模型测处理。。

▶▶ 3.8 架构设计-数据配置

业务范围配置化

支持拉取的数据业务范围配置化，
实时更新拉取数据的业务范围，降
低业务接入成本

活动范围配置化

针对平台参与的大型活动的数据进行配置化的拉取，实时更新活动内容
和数据来源



清洗规则配置化

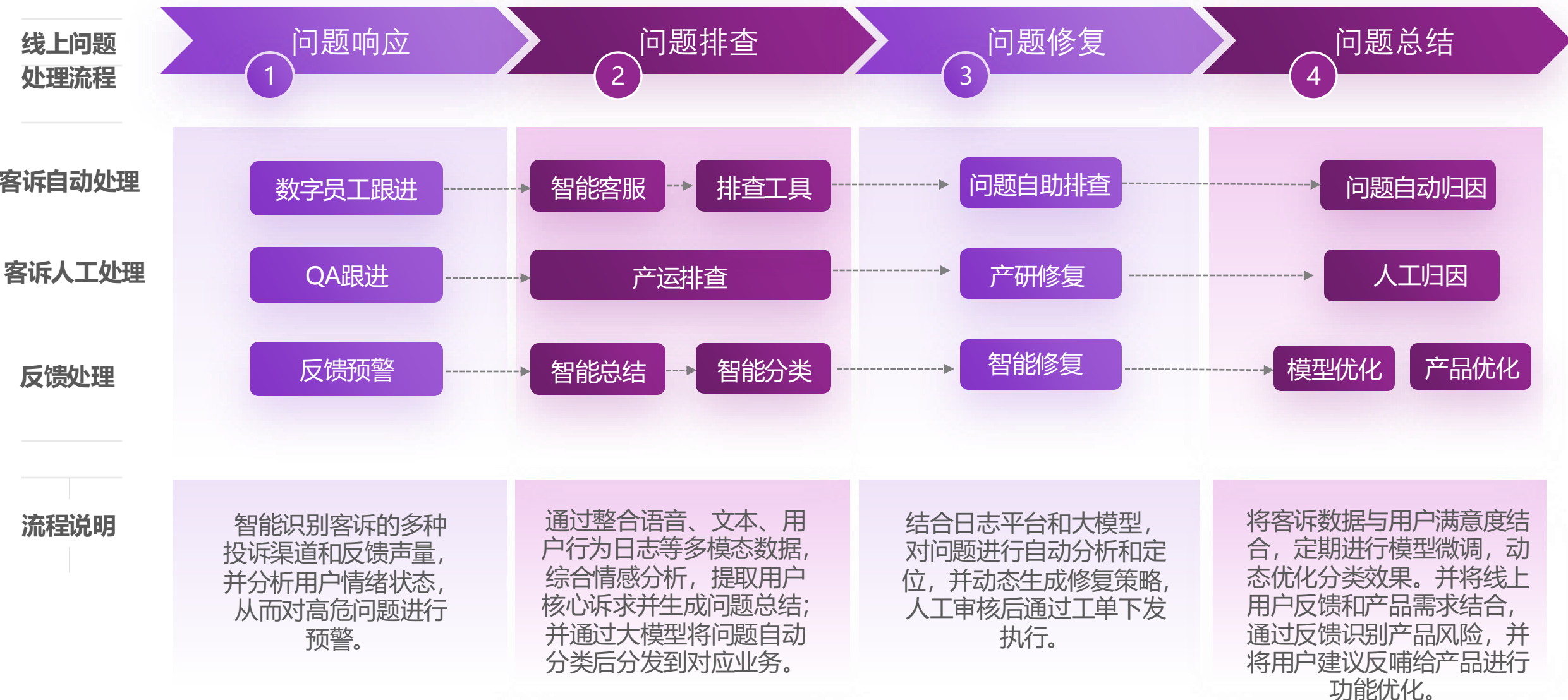
数据的清洗规则、关键词匹配、正则处理配置化，根据不同的需求动态变更数据清洗规则。



PART 04

实践效果

4.1 智能化客诉处理流程



▶▶ 4.2 项目收益

用户反馈

反馈处理量级提升，转化率提升

(一) 反馈处理效率提升

- 用户反馈处理量级提升
- 有效问题转化量提升，共转化有效问题
- 人工处理时长缩短

工单处理

人力释放，客诉达成率提升

(二) 助力客诉问题达成

- 自动跟进工单反馈，释放值周人力
- 客诉问题自动承接，问题达成率提升
- 客诉问题感知时间显著降低

舆情挖掘

大型活动站外舆情保障

(三) 大型活动舆情挖掘

- 有效处理数量提升
- 有效问题挖掘
- 客诉人数降低

4.3 踩坑指南之数据清洗

1. 用户与客服对话噪音大

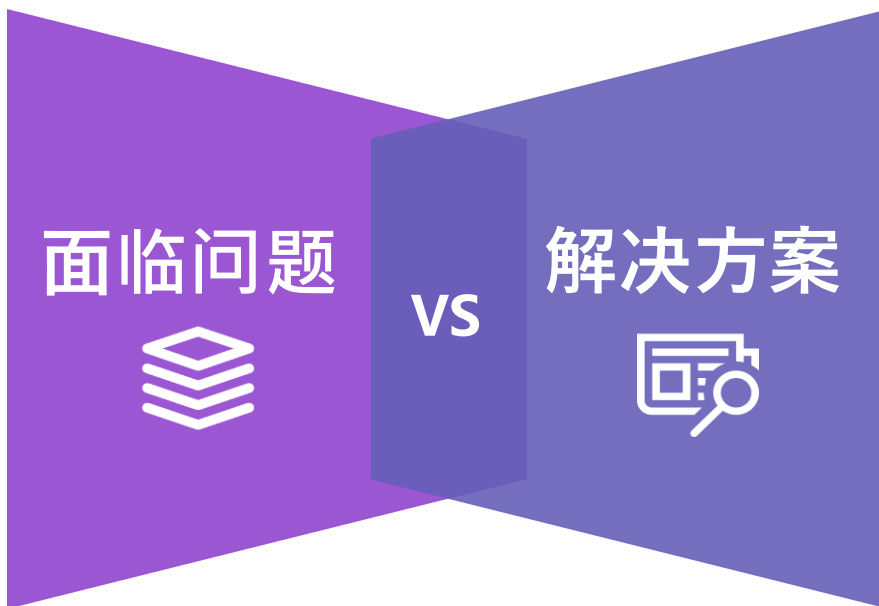
- 对话文本中存在口语化、错别字、语气词、重复内容等噪音；
- 多轮对话中信息散乱、上下文缺失；
- 不同渠道格式不统一。

2. 图片样式多，解析难

- 图片内容多样，存在低质量、重复或无关图片；
- 图片中信息难以直接解析为文字。

3. 语音质量不稳定

- 语音录音质量不稳定，存在背景噪音、断句不清等问题；
- 多语种混杂，方言、口音影响识别准确率。



1. 文本格式统一

- 文本预处理：对原始对话进行分词、去除停用词和噪音词；
 - 标准化格式：统一各渠道文本格式，建立统一对话格式模板；
- 上下文关联：对多轮对话建立上下文关系，以便后续准确分类和分析。

2. 图片内容文本化

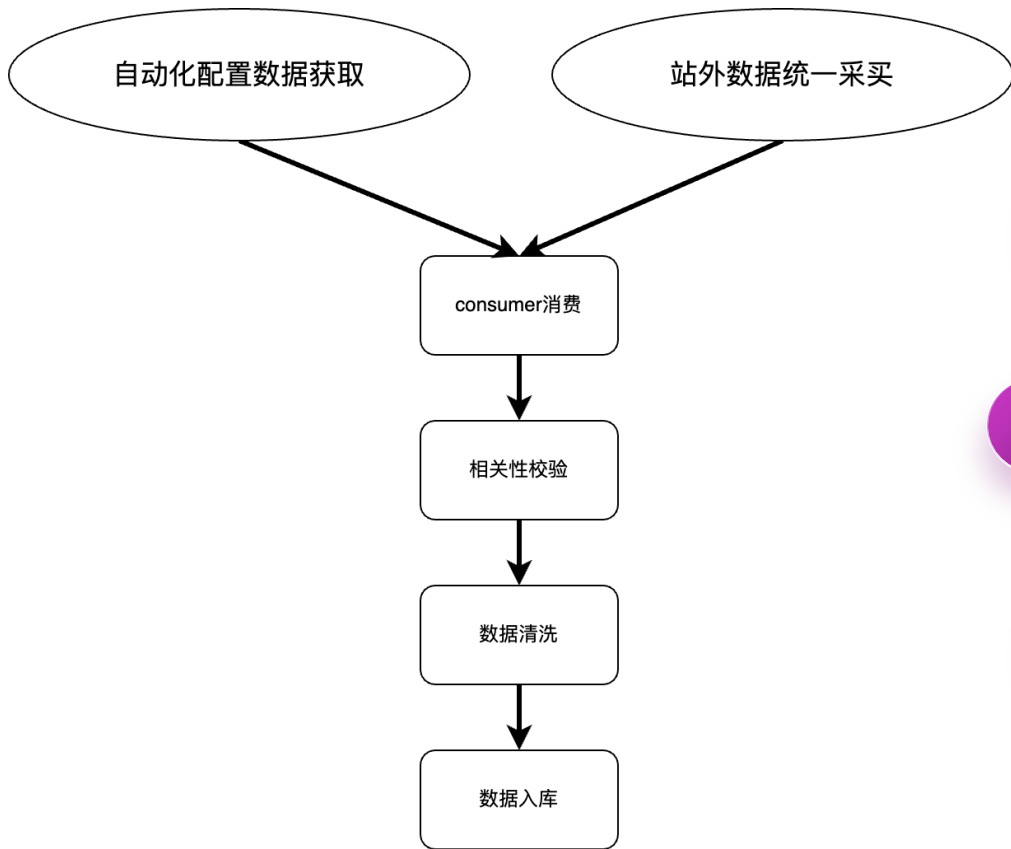
- 图像预处理：采用图像清晰度评估、重复检测等技术过滤噪音图片；
- OCR转换：利用 OCR 工具将图片内容转换为文本，针对不同语言和字体进行优化；
- 分类处理：对转换结果进行分类，单独处理与文本相关的反馈信息。

3. 语音降噪、转写、纠错

- 音频预处理：进行降噪、分割、音频增强等处理，提高语音识别准确率；
- 语音转文字：采用先进的语音识别模型（如深度学习模型）将语音转换为文本；
- 后续纠错：结合上下文和语言模型对转写结果进行纠错和标准化。



▶▶ 4.4 踩坑指南之数据采集



4.5 踩坑指南之模型效果调优

➤ prompt定义了模型应该如何回应问题或者执行任务的方向，那怎样通过prompt提高模型的效果？

完整的角色、任务描述

要尽量完整的描述任务和具体要求，对当前query个性化的信息，那么要尽量加入到prompt中

风格指导

希望模型关注的专有名词、关键词等巧用特殊符号，例如<>等，提高模型的关注度

少样本提示Few-shot

在模型中增加部分样例，以启用上下文学习，以引导模型实现更好的性能

CoT 思维链

可以尝试将“让我们逐步思考”添加到原始提示中，语料中有讨论、教学类场景，内容质量相对更高

调优前

请你帮我判断一下，以下对话属于所提供类别中的哪一类,问题类别和定义分别为：
 产品问题 指用户在使用产品过程中遇到的体验不佳、功能缺失、流程不合理等与产品设计相关的问题。
 技术问题 指产品或者服务在技术实现上存在的缺陷或故障,如系统崩溃、页面无响应、功能不生效等，导致产品功能不可用。
 风控问题 指为防范风险而采取的安全管控措施给用户带来的不便,如审核失败、对未成年人的管控、交易被拦截、对话被屏蔽、证件审核繁琐等。
 咨询问题 指用户对于产品的一些功能使用、操作流程等存在疑问，需要寻求解答和指导。这类问题不涉及产品缺陷，主要是用户自身对产品理解存在偏差或者疑惑。
 回答有如下要求，请仔细阅读：
 1、 深度理解这些分类的意思，并正确分类。如果你觉得该对话不属于上述四类问题，请归类到【其他问题】类。
 2、 要求返回类型为标准的json格式，键值为:问题类型。
 对话内容为：

调优后

```
< background >
对话的问题分类以及问题定义为：
产品问题：指用户在使用产品过程中遇到的体验不佳、功能缺失、流程不合理等与产品设计相关的问题。
技术问题：指产品或者服务在技术实现上存在的缺陷或故障,如系统崩溃、页面无响应、功能不生效等，导致产品功能不可用。
风控问题：指为防范风险而采取的安全管控措施给用户带来的不便,如审核失败、对未成年人的管控、交易被拦截、对话被屏蔽、证件审核繁琐等。
咨询问题：指用户对于产品的一些功能使用、操作流程等存在疑问，需要寻求解答和指导。这类问题不涉及产品缺陷，主要是用户自身对产品理解存在偏差或者疑惑。
< instruction >
请阅读以下用户客服的对话，并进行分析。对话中包含了用户（U）和客服（C）的交互。根据对话内容，总结对话的主要问题和用户的意图，并判断这个问题属于哪个分类。
对话内容：
[ 在此处插入对话内容 ]
< steps >
阅读对话：首先，仔细阅读整个对话，注意用户提出的问题和客服的回应。
识别主要问题：从对话中识别用户的主要问题和关注点。
判断意图：理解用户的意图，即他们希望解决什么问题或达到什么目的。
分类匹配：将识别出的问题和意图与给定的问题分类选项进行匹配
< results >
基于对话内容，我认为这个问题属于[ 问题分类选项 ]。
解释：[ 在此处解释为什么选择这个分类，并可能包括对话中特定部分的引用或总结 ]
< user input >
现在你可以将实际的对话内容插入到[在此处插入对话内容]的位置，并根据实际情况填写[问题分类选项]和[解释]部分。
```

prompt调优



prompt常用模板



立角色



述问题



定目标



补要求



4.6 踩坑指南之模型效果调优

新增问题聚类效果差?

聚类流程图

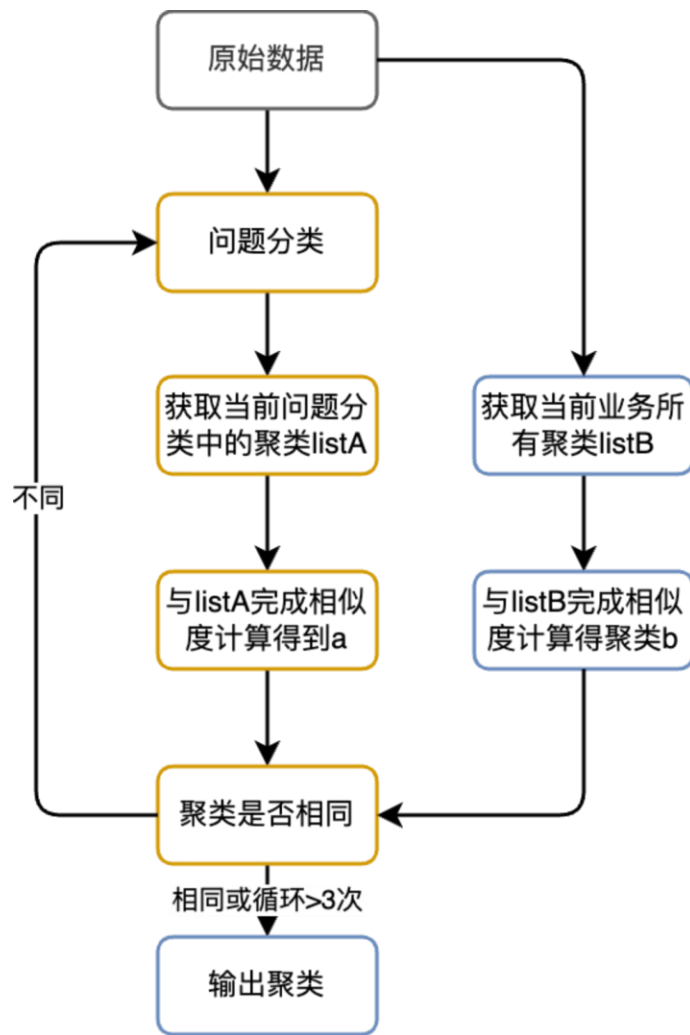
黄色链路：基于问题分类完成相似度聚类的精细化处理，确保聚类列表精准映射当前问题至相应类型。

蓝色链路：直接对原始数据执行相似度评估，打破问题分类限制，全面捕捉问题间相似性。对两次结果进行对比，确保分类和聚类准确率

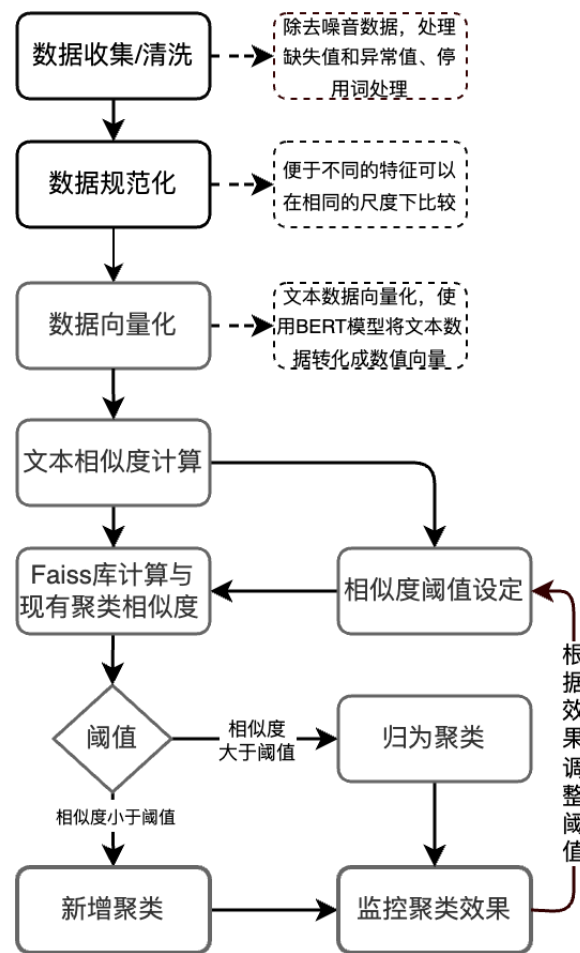
相似度计算流程图

右图相似度计算的详细流程包含文本向量化、搜索匹配、相似度计算、聚类问题选取

聚类流程图



相似度计算流程

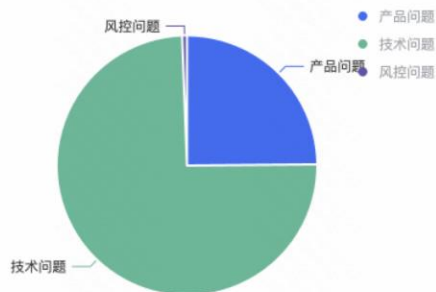


4.7 产品功能展示

1.

问题分类

问题分类占比



2.

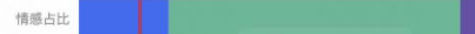
AI分析

ID	问题标题	状态	优先级
1
2
3

3.

情感分析

情感分析 (负向占比)



4.

风险预警

编号	AI分析	反馈时间	原声链接
1
2	...	20:...	...
3	...	20:...	...
4	...	20:...	...
5	...	20:16:25	...

今日命中关键词问题汇总

编号	问题等级	关键词	出现次数	最新反馈时间	最新原声链接
1	P0	点不了	1	2025-09-30 09:39:45	...

PART 05

未来展望

总结与展望：未来趋势与挑战

•未来方向

- **全链路AI闭环**：AI聚类—>人工确认->问题修复->线上效果观测。
- **人机协同深化**：AI处理常规问题，人工聚焦复杂场景。

•挑战与应对

- **数据隐私与合规**：确保用户信息脱敏与合规使用。
- **AI决策可解释性**：提升模型透明度和可信度。



参与调研您将优先获得



AiDD定制版
《AI+软件研发精选案例》



专属学习顾问
1对1需求对接

AiDD会后小调研

AiDD峰会致力于协助企业利用AI技术深化计算机对现实世界的理解，推动研发进入智能化和数字化的新时代。作为峰会的重要共建者，您的真知灼见对我们至关重要。衷心感谢您的参与支持！

2025 AI+研发数字峰会

拥抱 AI 重塑研发



扫码参与调研

科技生态圈峰会 + 深度研习

—1000+ 技术团队的选择



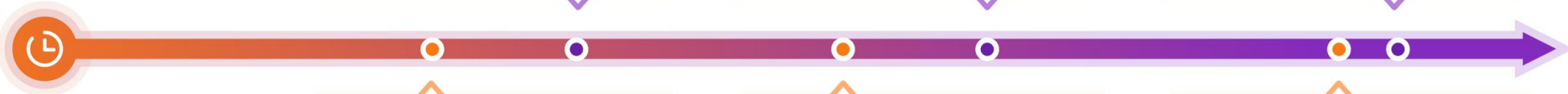
K+峰会 **敦煌站**
K+ 思考周®研习社
时间: 2025.08.29-30

K+峰会 **上海站**
K+ 金融专场
时间: 2025.09.26-27

K+峰会 **香港站**
K+ 思考周®研习社
时间: 2025.11.17-18



K+峰会详情



AIDD峰会 **上海站**
AI+研发数字峰会
时间: 2025.05.23-24

AIDD峰会 **北京站**
AI+研发数字峰会
时间: 2025.08.08-09

AIDD峰会 **深圳站**
AI+研发数字峰会
时间: 2025.11.14-15



AIDD峰会详情



2025 AI+研发数字峰会

AI+ Development Digital Summit

感谢聆听!

扫码领取会议PPT资料

