

2024浦源大模型挑战赛作品：生成式基金诊断话术

行业应用赛道竞赛选题：金融 - 生成式基金诊断话术应用

一、生成式基金诊断话术应用介绍

市场与行业背景

在现今财富管理市场中，金融产品数量急剧增加。以公募基金为例，自2019年以来，产品数量已从约6000只翻番至约12000只，同时市场规模也从1.2万亿增长到3万亿。在这一庞大的市场背景下，投资者面临着如何寻找、持有金融产品、并实现财富长期保值增值的挑战，随之而来投资者对于持有的基金业绩是否符合预期、市场上有没有更合适的基金成为投资后普遍的关注点。

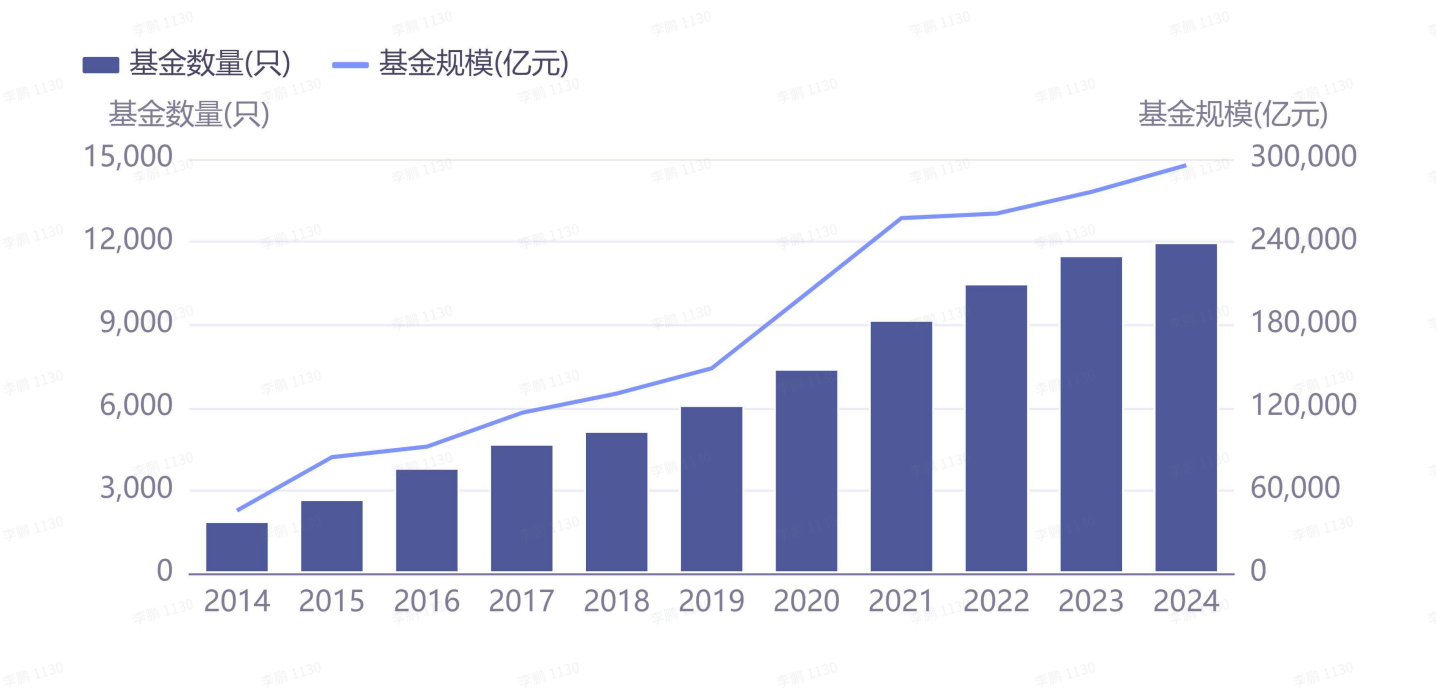


图1-1：基金规模走势

券商作为金融机构，充当着社会财富的管理者，致力于为投资者提供专业的产品研究及咨询服务，助力客户财富的稳健增长。理财师作为业务的重要触点，单体服务范围相对有限（通常一位理财师仅能有效服务约200位客户），其专业能力往往依赖于个人经验的积累，而且随着客户服务数量的增加，服务质量可能会出现波动。为了推进普惠金融，让更多投资者享受到稳定优质的金融服务，人工智能大模型技术的引入或能显著提升理财师在服务广度与响应质量。

业务痛点与应用场景

本应用场景提取自理财师日常投后服务沟通工作。业务痛点在于投后服务的工作偏个性化，每位客户的情况皆有不同，因此个体效能提升难度大，不易普及。详细分析来看，投后服务工作中的难点在于持仓数据与基金数据所需的案头工作准备时间较长、业绩数据对齐过程繁琐、多客户沟通不易复用、理财师沟通技巧又需长期积累，这些因素共同导致了投后服务效率的低下。

解决方案与应用价值

为此，我们聚焦在如何提升理财师沟通效率上，将优秀理财师沟通经验与内容进行了建模，让理财师在与客户沟通前能有一个全面且有效的案头话术，大幅降低服务门槛。同时，我们还加入理财师个性化的惯用素材和语料，让资讯数据通过大模型技术融入沟通话术中，使不同理财师与每位客户都能生成个性化内容。最后，在话术生成前，我们还对敏感词进行监控与拦截，确保生成式内容的合法合规。

总而言之，本应用实现了个性化基金诊断话术的生成能力，在降低理财师与客户沟通的难度和成本的同时，让每一位理财师都能够使用优秀理财师的服务技巧，从而提升业务整体服务效率和服务质量。

二、应用场景设计

产品设计目标

基于上述应用场景，我们希望能激活【员工-下挂客户-持仓产品】上下游全链条数据。通过大数据及大模型能力综合处理评估后，推荐对比的产品，并生成相应的基金诊断话术。全流程有望实现客户千人千面精准化投后营销支持，赋能业务。

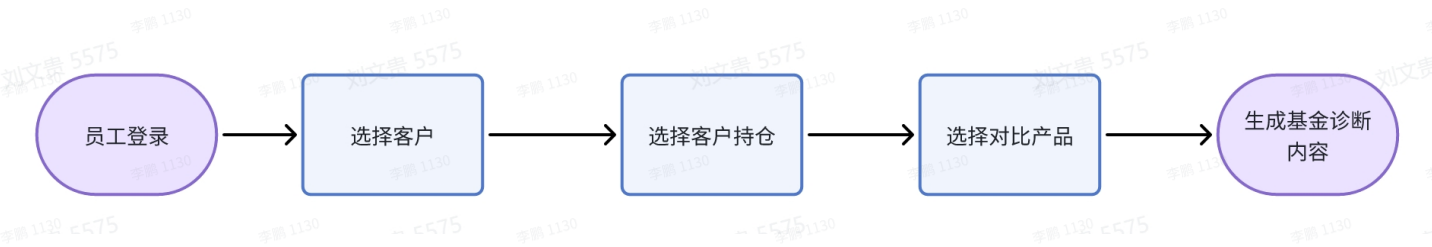


图2-1：用户路径

场景产品示例

产品的使用流程如下图所示。（由于合规要求，部分数据隐藏处理）

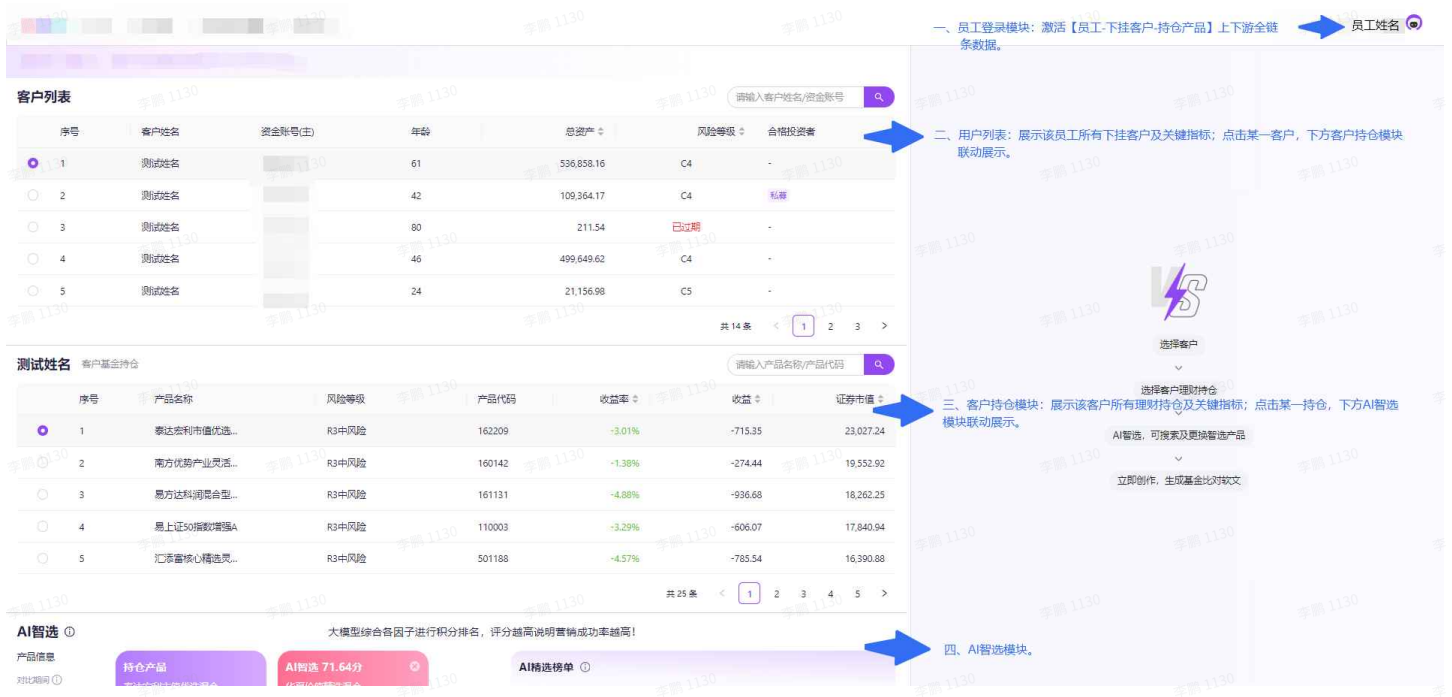


图2-2：产品界面示例（1）

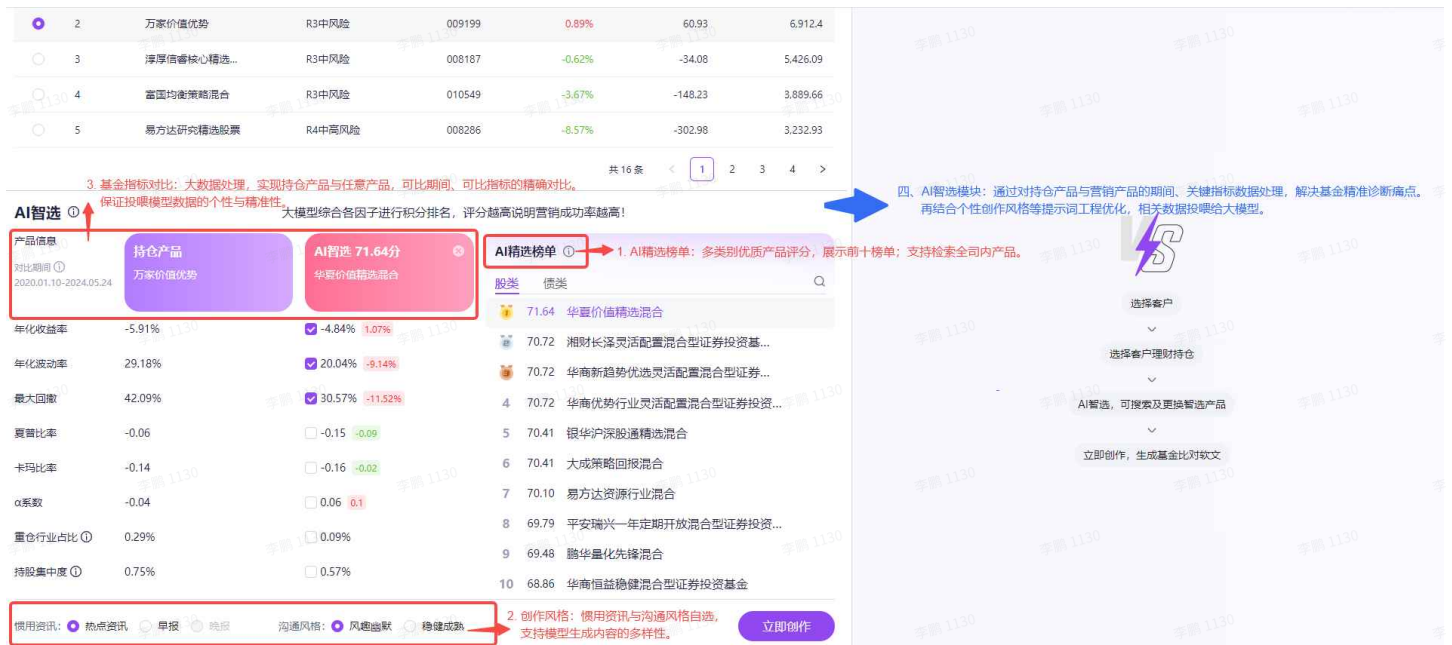


图2-3：产品界面示例（2）

五、内容生成模块：包括多个段落；敏感数据分流处理，不同段落采用系统拼接、模型生成及综合方式处理。

客户列表

序号	客户姓名	资金账号(主)
01	赵小峰	
02	林小武	
03	赵精	
04	赵精	
05	赵精	

林小武客户理财持仓

序号	产品名称	产品代码
01	南方内需增长两年持有期股票A	123456
02	景顺长城沪港深精选	123456
03	南方内需增长两年持有期股票A	123456
04	南方内需增长两年持有期股票A	123456
05	南方内需增长两年持有期股票A	123456

AI智选

产品信息

对比期间 2023.06.06 - 2024.05.05

年化收益率 -111.50%

年化波动率 2.00%

持仓产品

南方内需增长两年持有期股票A

南方内需增长两年持有期股票A

景顺长城沪港深精选

开场白

尊敬的【测试姓名】，您好！我是湘财证券陆家嘴营业部的投顾！

慎用资讯

我一直关注市场最新热点资讯，当前特别留意到芝加哥商业交易所确认未产生任何指数定价的情况，这反映出市场在数据依赖性上的暂时空白，我们需耐心观察后续动态，确保投资决策基于充分信息。您的持仓属于债类基金，这类配置在当前环境下通常注重稳定收益与风险控制，与我们的稳健投资理念相契合。

客户持仓观察

您持仓的华夏现金增利基金，已经陪伴您有一段时间了。尽管它在流动性与稳定性方面表现出色，但我们也不可忽视其在追求更高收益率上的局限性。作为您的投资伙伴，我始终在这儿，关心您的每一分投资，随时准备帮您探讨如何优化配置，让您的财富之路走得更稳健。

持仓基金表现

您持仓的华夏现金增利基金，近年来表现不尽人意，年化收益率2.63%与市场平均水平相近，但考虑到较低的风险水平，如夏普比率仅0.20%，表明其单位风险所获得的收益相对有限。卡玛比率负值及显著低于标准的a比率揭示了基金在控制风险的同时，未能有效捕捉超额收益，这可能是由于投资组合中较高比例的现金及低风险资产配置，如大量持有买入返售金融资产和银行存款，而股票资产占比极低，影响了潜在的增值空间。基金资产主要分布在安全性较高的金融工具上，虽保障了资金的安全性，但也限制了收益提升的可能性。综上，基金在追求稳健的同时牺牲了增长潜力，建议评估是否符合您的风险偏好与收益期望。

目标基金引导

考虑到您期望在保持较高收益水平的同时优化投资组合，可以考虑向稳定性更强的方向调整。而提及的华夏价值精选混合型证券投资基金【基金代码：007592】，尽管其在盈利能力与选股能力上表现出色，择时能力一般，但提供了不同于您当前持仓的配置思路。这只基金尤其适合寻求长期稳健增值的投资者，它虽以较低的收益率为特点，但在市场波动中展现出较好的风险控制能力，是分散投资风险的一个不错选择。由于您的现有持仓与本产品不属于同类，纳入该基金可进一步多元化您的资产配置，提升整体投资组合的均衡性与抵御市场不确定性能力。

目标基金优势

近年来，您持有的华夏现金增利基金在流动性和稳定性上确实展现了其优势，但面对收益提升的需求，我们或许需要更广阔的视角。对比之下，华夏价值精选混合基金虽在同类1215只基金中评级4星，盈利能力与选股能力脱颖而出，其年化收益率2.63%与华夏现金增利相比，在稳健基础上寻求了适度增长。值得注意的是，华夏价值精选的夏普比率与华夏现金增利相近，显示出在承担每单位风险下，两者收益水平相当；而卡玛比率略低，反映在控制下行风险的同时获取收益的能力需进一步观察。特别提到的是，该基金a比率显著，表明其主动管理带来的超额收益潜力较大，尽管近期数据呈现为负，这可能是市场短期波动影响的结果。

华夏价值精选混合基金在资产配置上偏向股票，特别是电气机械、食品制造等行业，持股较为分散，如海尔智家、伊利股份等优质企业均在其重仓之列，这种布局有利于分散风险并捕捉多领域增长机会。与您当前持仓相比，华夏价值精选在行业选择与资产配置上更为积极，旨在长期增值，适合寻求在稳健中求发展的投资者。

综上所述，华夏价值精选混合基金以其独特的资产配置策略和良好的市场适应性，在同类产品中展现出较为平衡的风险收益特征，尤其适合期待在保持资金相对安全的同时，探索更高收益可能的投资者。不过，任何投资决策都应基于个人风险承受能力和投资目标，细致考量后再做决定。

优势指标解析

一、夏普比率

夏普比率，简单来说，就是衡量您的基金收益与所承担风险的比例。想象一下，这是计算您每承受一份风险，能获得多少额外收益的指标。比率越高，意味着基金在单位风险上给予的回报越丰厚，表明基金在控制风险的同时，获取收益的能力较强。

二、卡玛比率

卡玛比率，又称为收益回撤比，它关注的是基金收益与最大回撤之间的关系。这个指标帮助我们了解，基金每赚取一单位的收益，最多可能会损失多少。比率越高，表示基金在获取收益的同时，对下跌风险

图2-4：产品界面示例（3）

三、技术方案设计

整体设计方案

基于大模型的AI系统的工作流程，应用分成了用户端、后端、推理服务和大模型四个模块。

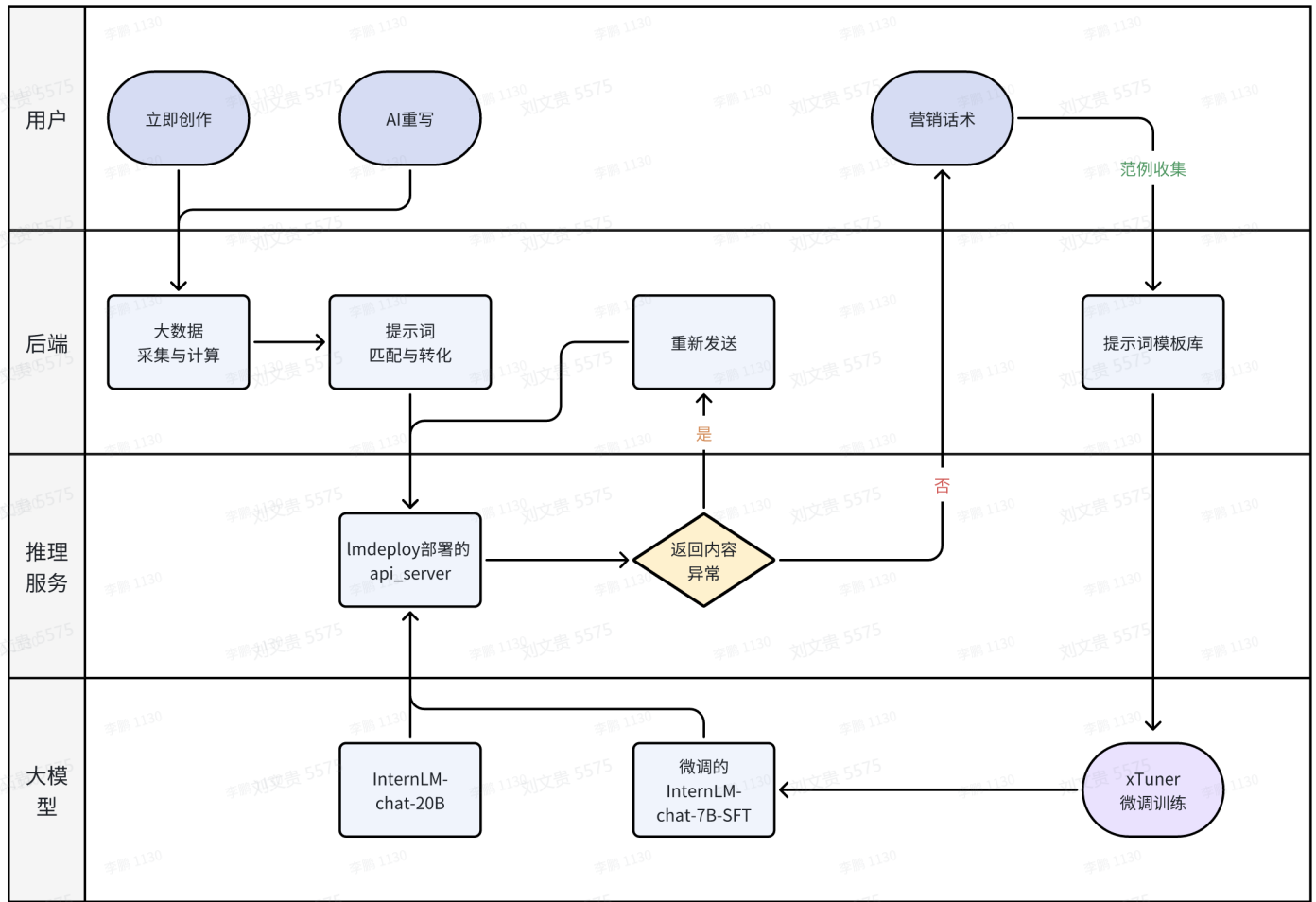


图3-1：生成式基金诊断话术数据流程图

系统有如下特色：

- 集成大模型**：系统集成了InternLM-chat-20B和微调的InternLM-chat-7B-SFT，实现了高效的自然语言处理和内容生成。且不超过20B的参数数量保证了模型能在大多数硬件条件上运行。
- 灵活的用户交互**：提供“立即创作”和“AI重写”两种用户交互方式，满足不同需求。
- 实时异常处理**：在推理服务中，设有返回内容异常检测机制，保证系统输出的准确性和可靠性。
- 微调优化**：通过XTuner微调训练，不断优化大模型，提高其处理特定任务的能力，提供更私有化的升级迭代。
- 数据驱动**：系统利用大数据采集与计算，确保提示词匹配和数据转换的智能化和个性化。
- 数据合规**：通过提示词工程的设计与优化，应用保证了输出文本的合规性以及敏感数据的保护。

提示词工程设计

提示词工程确保大语言模型有效地理解任务并执行。本应用构建9条业务提示词评测标准，从提示词产生到提示词转换全流程地评估提示词的效能。另外，完成超20+个投顾专家提示词优秀模板建模，创新实现基于上下文多轮对话的提示词TOKENS节省方案，单次任务TOKENS节省超73%。

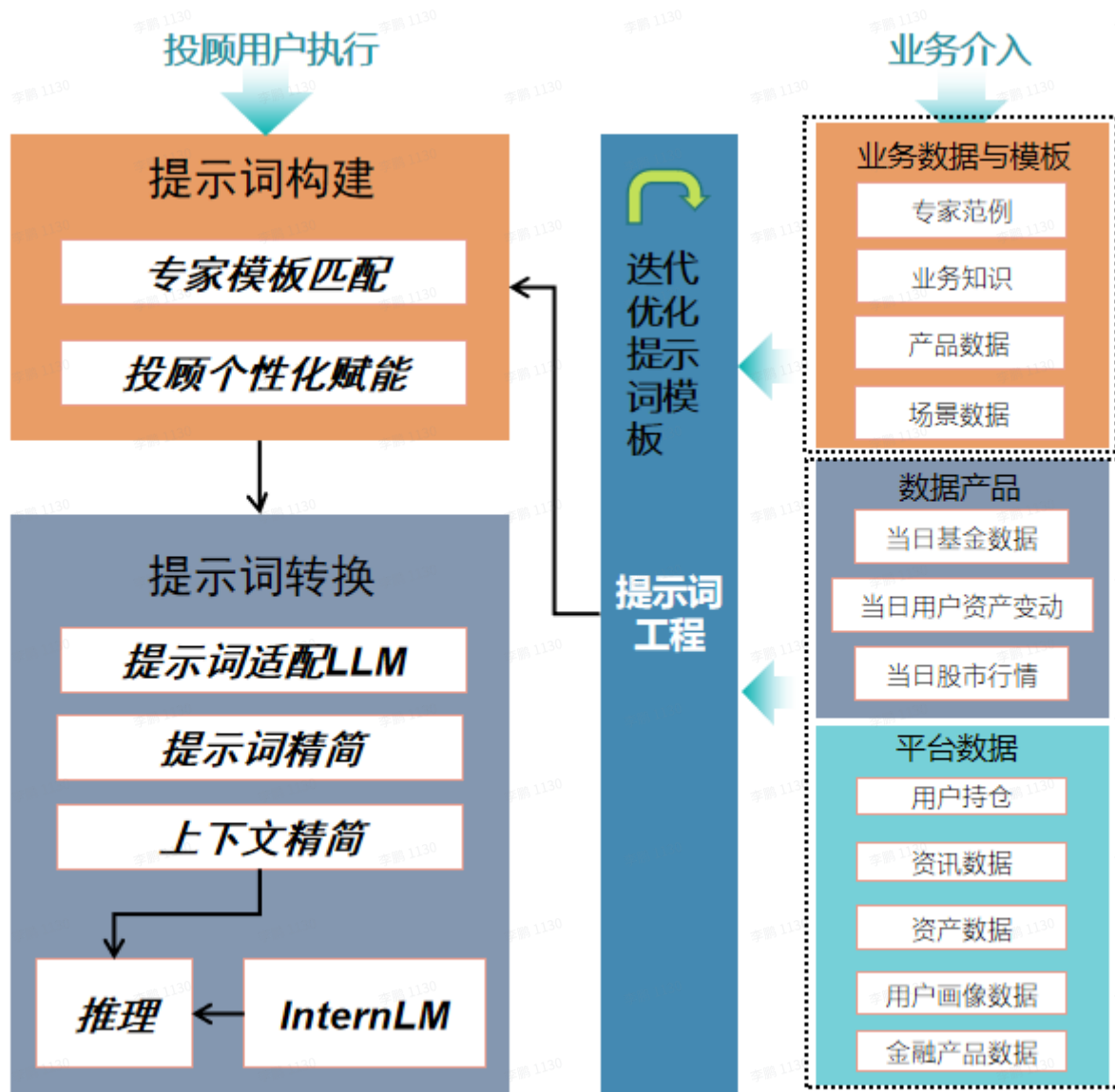


图3-2: 提示词工程设计

数据应用

行业数据、司内数据及部分特殊数据是行业垂类大语言模型应用不可或缺的基石。基于丰富客户画像、产品画像以及资讯中心数据，我们精心打造了AI算法，旨在为员工提供精准的市场营销产品建议。员工根据这些建议选择相应的产品，InternLM2模型便会发挥其强大功能，输出详尽的基金指标、资讯数据、产品持仓分析等关键信息，并自动生成“专家级经验”的营销话术，从而显著提升营销效率。

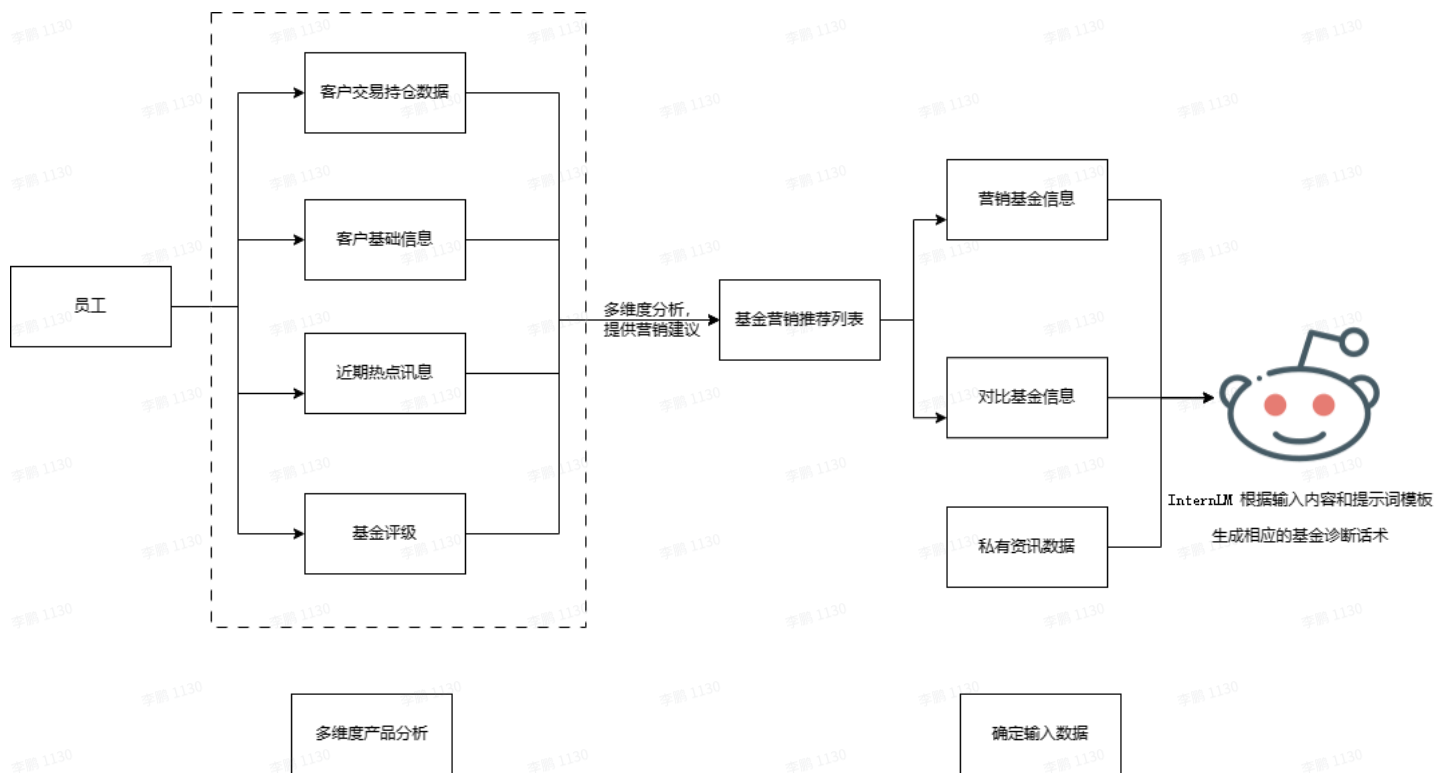


图3-3：数据流转流程

数据合规

鉴于大语言模型作为新兴技术在金融领域的广泛应用，以及国家出台的相关法律法规如《生成式人工智能服务管理暂行办法》和《国家互联网信息办公室关于发布生成式人工智能服务已备案信息的公告》等对其进行的严格限制，我们结合金融应用场景的特性，特别是敏感数据的处理、行业监管及合规要求，以及用户、资讯、产品等数据量大的情况，设计了以下流程：

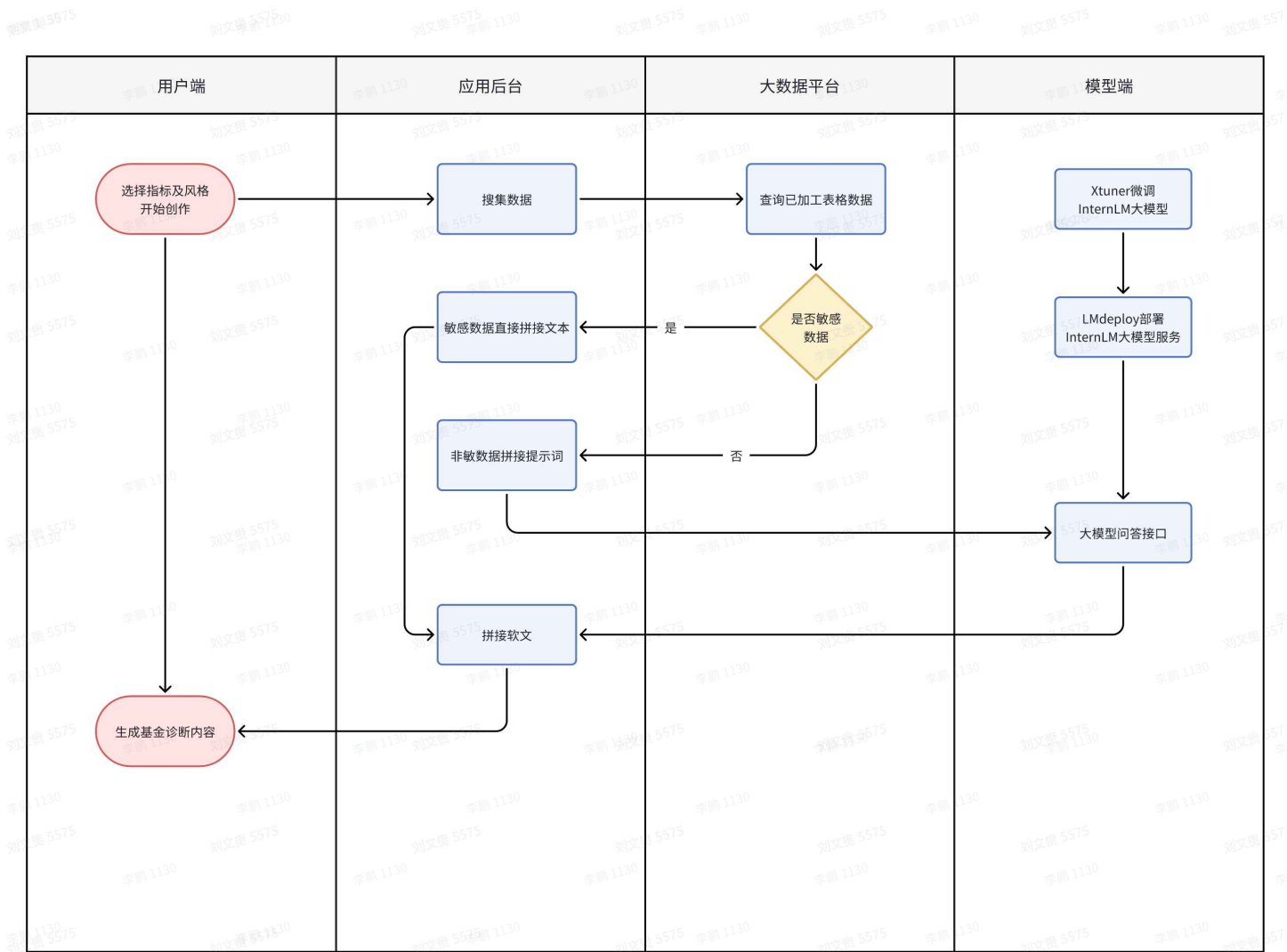


图3-4：敏感数据分流处理

在这个流程中，我们做了如下处理：

- 敏感词策略：**输入、输出双控。由于不同阶段，敏感词范围有一定差异性，投喂模型时候明确敏感词范围，同时数据阶段再次进行敏感词过滤，保证内容的合法合规。
- 敏感数据分流处理：**针对不同类别数据，在系统功能实现过程中，进行专门的处理。重点规避敏感信息的外发，将敏感信息拆分，并利用后端拼接的方式替代模型生成文本。

这些流程确保了在保证数据能被安全合规使用的同时，大语言模型也能有效地输出基金诊断文本。

模型微调

通过提示词工程的持续优化，我们成功引导大语言模型高效、合理地利用各类数据资源，输出“专家经验的基金诊断文本”。在此基础上，我们采取了双管齐下的策略来优化模型的性能。一方面，我们整合了丰富的金融基础数据和金融公告，构建了一个全面且深入的金融数据集。通过对此数据集的精细微调训练，我们显著提升了InternLM-chat-20B模型在金融领域的专业能力和理解深度。这使得模型能够更准确地捕捉金融市场的动态，为用户提供更加专业、精准的金融分析和建议。另一方面，我们特别搜集了专家投顾的优质话术数据，这些数据不仅体现了专业的金融知识，还蕴含了丰富的沟通技巧和策略。我们对这些数据进行了分割、格式化和标注，用于对模型进行二次微调训练。这一举措极

大地提升了模型在话术生成方面的能力，使其在与用户的交流中能够更加自然、流畅地表达专业观点和建议，从而为用户提供更加贴心、高效的金融服务。

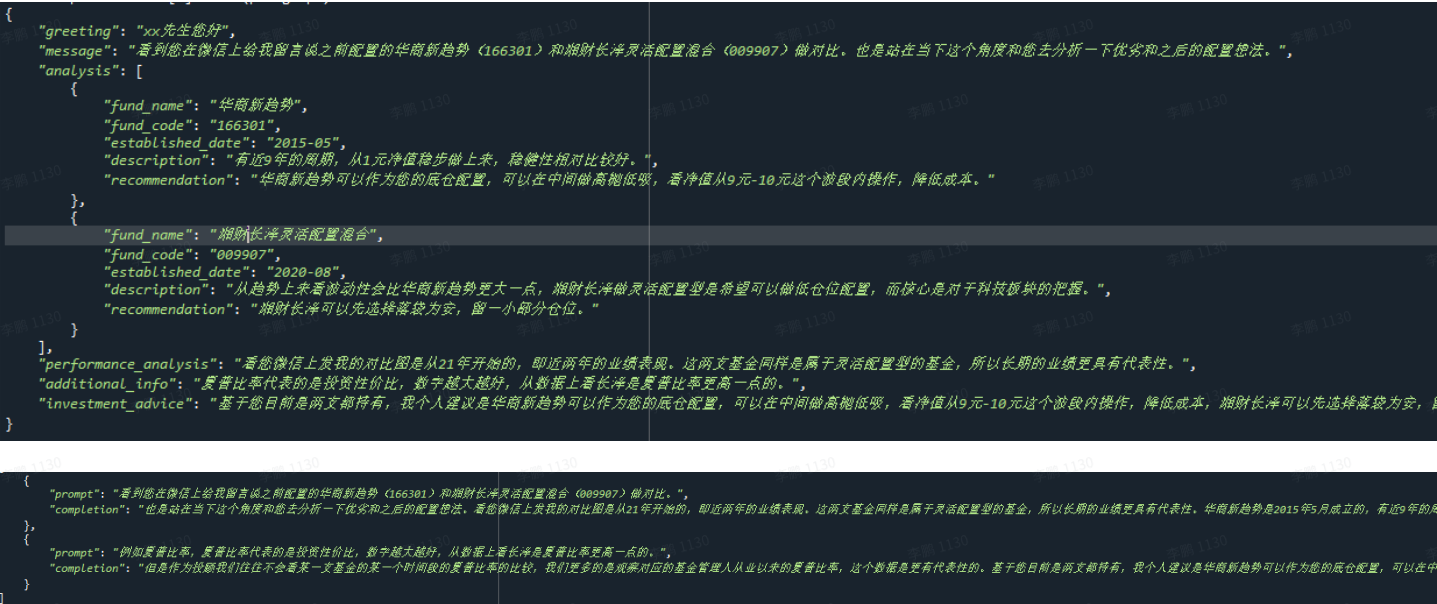


图3-5：模型微调

四、测试与验证

业务效果验证

在评估本应用的实际效果时，我们采用了三个核心维度来确保评测的全面性和准确性。首先，我们关注理财师日均客户沟通次数的变化，以此作为衡量本应用是否切实提升了理财师沟通效率的直观指标。通过对比使用前后沟通次数的变化，我们能够清晰地看到应用在实际操作中的效果。

其次，我们关注生成式内容的使用比例。这一指标不仅反映了理财师对生成内容的接受程度，更重要的是，它验证了生成内容在实际业务中的有效性。当生成式内容的使用比例持续上升时，意味着这些内容为理财师提供了实质性的帮助，进而提高了他们的工作效率。

最后，我们深入分析业务转化的成交归因。通过追踪和分析业务转化的过程，我们能够明确本应用在理财师展业过程中的服务赋能价值。这一维度不仅展示了应用对业务转化的直接贡献，更体现了其在提升理财师服务质量和客户满意度方面的潜在价值。

综上所述，通过这三个维度的综合考量，我们能够全面、准确地评估本应用的实际效果，为后续的优化和改进提供有力的数据支持。

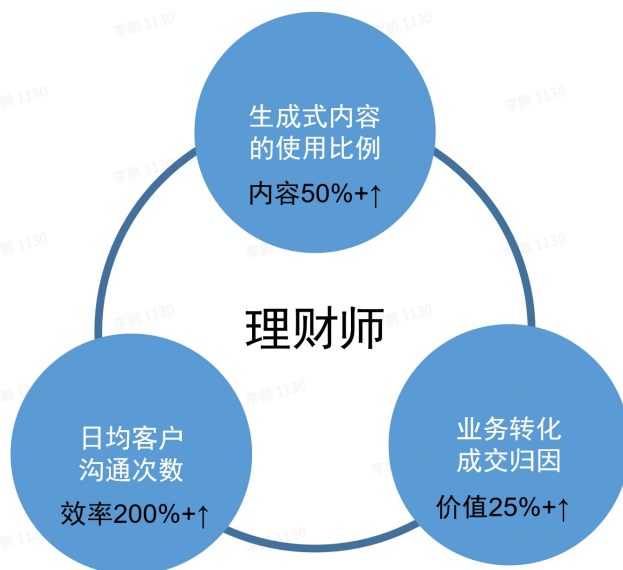
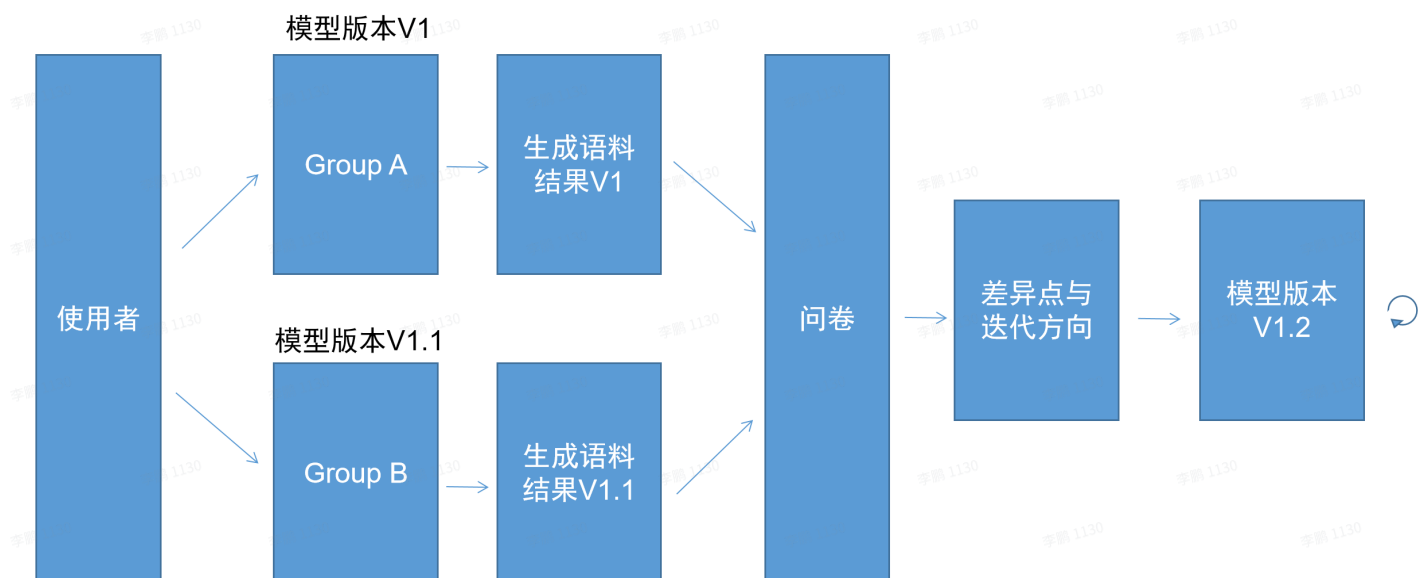


图4-1：三维度业务效果评测

测试指标及验收标准

相比其他具有标准答案的应用，很难量化地对生成式内容进行评测。为此，我们设计了生成式内容评价体系，引入AB-test实现对模型能力的比对和校验，向使用者收集问卷，针对生成内容的一致性、拟人性、逻辑性和创新性进行评估和打分，以便于对模型的有效迭代。



问卷示例

- **一致性**：检查生成文案是否与输入数据一致，评分标准为1-5分。
 - 例：生成的持仓信息是否准确反映客户数据。
- **拟人性**：评估生成内容是否具有自然流畅的人类语言特征，评分标准为1-5分。

- 例：生成的建议是否像理财师的真实话术。
- **逻辑性**：检查文案的逻辑结构和连贯性，评分标准为1-5分。
 - 例：基金对比是否清晰，推荐理由是否合理。
- **创新性**：评估生成内容的新颖度和创意性，评分标准为1-5分。
 - 例：是否提供了有价值的新视角或独特的建议。

通过这些测试指标，我们能全面评估大模型在基金对比软文创作中的表现，并根据评分结果进行模型优化和改进。

问卷内容截图：

AI基金比较验收问卷

1. AI输出的内容中，提到的基金名称数据是否与页面选中数据一致？

☐ 一致

☐ 50%一致

☐ 不一致

2. AI输出的内容中，提到的基金指标数据是否与页面选中数据一致？

☐ 一致

☐ 60%一致

☐ 30%一致

☐ 不一致

3. AI输出的段落内容是否与各段落标题内容匹配？

☐ 匹配

☐ 80%匹配

☐ 60%匹配

☐ 40%匹配

☐ 20%匹配

☐ 不匹配

4. AI输出的内容是否有前后矛盾或常识性的冲突？

☐ 没矛盾

☐ 未见明显矛盾

☐ 有矛盾但不影响

☐ 有个别矛盾

☐ 有一些矛盾

☐ 矛盾问题大

5. AI输出的内容专业程度你打几分？

☐ 5

☐ 4

☐ 3

☐ 2

☐ 1

☐ 0

6. AI输出的内容是否可以直接发给客户听？

☐ 可以

☐ 基本可以

☐ 调整一些口条即可

☐ 需做进一步修正

☐ 完全不可以

7. AI输出的内容是否有业务合规性和风险提示方面的信息？

☐ 有

☐ 没有

8. AI输出的内容有没有标题以外的内容呈现？

☐ 有

☐ 没有

9. AI输出的9个段落中，效果最差的是哪一段？

☐ 1

☐ 2

☐ 3

☐ 4

☐ 5

☐ 6

☐ 7

☐ 8

☐ 9

10. AI输出的9个段落中，效果最好的是哪一段？

☐ 1

☐ 2

☐ 3

☐ 4

☐ 5

☐ 6

☐ 7

☐ 8

☐ 9

提交