



西安交通大学
XI'AN JIAOTONG UNIVERSITY

西安交通大学人工智能先导计划之师生素养提升计划

AI赋能交晓智平台 助力教学赋能 (黄金篇)



网络信息中心



2025年5月



“交晓智” 黄金篇知识点清单

- 已涉及知识点
- 本次案例涉及的知识点
- 黄金篇总体知识点

大模型&提示词	知识库&数据库	插件	workflow	测试验证
提示词模板	文档解析与分段技巧	交晓智公共插件	workflow 创建	智能体调试
大模型参数设置	知识库召回方式	绘图插件使用	选择器组件	调试树
提示词框架	问答库	触发器使用	循环与终止循环组件	火焰图分析
提示词调优	术语库	自定义插件	代码组件	workflow 调试
大模型选型	数据库使用 (NL2SQL)	外接插件	多 workflow 协同	插件调试
...

交晓智平台赋能教学场景

培训 要点

- 知识库召回技巧、文档解析等技术
- 问答库与术语库的使用
- 如何让外部系统以API的方式接入做好的智能体
- 大模型参数与提示词调优

培训 收获

- 智能体如何实现课堂问答互动
- 智能体如何根据知识点快速生成分层级、多题型的智能考题组卷
- 智能体如何根据知识点出题，和用户互动，自动判题，分析知识点掌握情况
- 掌握智能体的问答调优技巧

智能体建设

阶段一
需求分析

阶段二
实现思路设计

阶段三
智能体开发

阶段四
测试验证

阶段五
提交发布

阶段一：需求分析-需求提出

- **场景引入：**我是一名教师，课中、课后都需要给学生答疑，同样的问题需要重复回答，问题较多时也会出现回答不及时的情况。其次，每次备课的时候，都需要精心设计一部分和课堂知识点相关的考题，用于在课堂上和学生互动，学生答完题后还需要逐个判题，花费的时间精力都比较大。
- **需求：**希望通过AI实现课程答疑、随堂出题、随堂测验，提升教学效率。

功能需求	详细描述
课程答疑	针对课程信息、授课教师信息、课程知识点进行答疑
随堂出题	根据课堂讲授的知识点出指定数量的题目、可以支持单选、多选、判断、问答等多种题型
随堂测验	根据课堂讲授的知识点生成一套考题，每次只出一道题，学生回答完自动判题，回答正确时继续出下一道题，回答错误时给出题目分析，然后再继续出下一道题，所以问题回答完毕后给出综合评价，分析学生对知识点的掌握情况。

阶段一：需求分析

需求分析

1.明确赛道：教学赋能

2.明确目标：建设一个课程小助手，辅助老师答疑、出题、完成课堂测验等工作

3.明确功能：

(1) 课程答疑

- 解答课件中的知识点相关问题。

(2) 随堂出题

- 针对知识点随机出指定数量的单选题、多选题、判断题、简答题。

(3) 随堂测验

- 针对知识点逐个出题，进行回答，自动判卷并进行分析，测验完成后给出知识点掌握情况的评价

阶段二：实现思路设计

基于需求，通过
交晓智哪些能力
实现

1. 选择一个适合该应用场景的智慧大脑——**大模型**

2. 明确智能体要实现的三大功能（答疑、出题、测验）——**提示词**

3. 检索课程知识点——**知识库**

常见问题需要能快速回答，提升回答准确性和速度——**问答库**
特称词、专业用语不能被混淆，通俗语言也能理解——**术语库**

4. 交互引导，提升用户使用体验——**开场白、对话建议**

5. 第三方系统使用智能体能力——**开放智能体API**

阶段三：智能体开发

步骤一：选择大模型：本案例为课程助手场景，选择模型deepseek-v3

大模型名称	交晓智中模型版本	特点	适用场景
智谱清言 (GLM)	GLM-4-Air	轻量级模型，推理速度快，资源占用低	移动端应用、实时对话（如客服机器人）、简单问答、对响应速度要求高的场景。
	GLM-4-Long	擅长处理长文本，上下文窗口较大（可能达数万token）	长文档摘要（如论文、报告）、多轮对话记忆、连续剧情分析、法律合同审查。
	GLM-4-Plus	功能增强版，支持复杂任务处理	代码生成、数学推理、跨模态任务（图文生成）、深度数据分析。
通义千问 (Qwen)	qwen-turbo	极致推理速度，低成本，适合高频调用	高频聊天场景（社交APP）、简单信息检索、轻量级对话（如智能硬件）。
	qwen-long	长上下文支持，平衡性能与效率	中等长度文本生成（如新闻稿）、长对话管理（教育辅导）、基础代码补全
	qwen-plus	综合性能强，准确率与泛化能力提升	商业文案创作、复杂问题解答（如医疗咨询）、多语言翻译
	qwen-max	最大参数量版本，支持高复杂度任务，可能具备高级推理能力	科研分析、金融预测、多轮逻辑推理（如辩论）、创意内容生成（小说、剧本）
豆包 (Doubao)	Doubao-Pro-32k	专注长文本处理、响应快、综合能力较强	中长文本分析（企业报告）、知识库问答、多文档信息整合
DeepSeek	deepseek-v3	通用性强、性能均衡	生成效率高，企业级智能助手、搜索引擎增强、个性化推荐系统
	deepseek-r1	擅长复杂推理与逻辑能力，擅长数学、编程和自然语言推理任务，响应较慢	学术研究辅助、行业数据分析（金融/医疗）、精准推荐（电商、内容平台）
Kimi	moonshot-v1-8k	基础版，适合短文本，响应快	实时翻译、短对话交互、社交媒体评论生成
	moonshot-v1-32k	扩展上下文，支持长文本理解	技术文档解析、会议纪要整理、中等长度代码审查
	moonshot-v1-128k	超长上下文（约10万字），行业领先的文本容量	整本书籍摘要、长篇小说续写、复杂代码库维护（如GitHub项目）、学术论文复现

阶段三：智能体开发

步骤一：大模型参数设置

1、大模型选择
deepseek-v3

2、根据Deepseek官方建议，将温度设为1

Temperature 设置

temperature 参数默认为 1.0。
• 我们建议您根据如下表格，按使用场景设置 temperature。

场景	温度
代码生成/数学解题	0.0
数据抽取/分析	1.0
通用对话	1.3
翻译	1.3
创意类写作/诗歌创作	1.5



3、单次回复最大token数实现制大模型一次输出最大内容的长度，将单次回复最大token数设置成最大，避免单次出题数量太多的时候输出被截断

4、对话轮数决定了大模型的记忆能力，轮数越多，记忆越多。将单次对话轮数保留设置为7轮，正常测验有3-5道题，所以轮次必须大于5轮

5、RAG范围是参与知识库检索的对话轮数，数值越大，多轮对话内容的相关性越高。将RAG范围设置为7，正常测验有3-5道题，所以RAG范围必须大于5，避免后面的题目把知识点遗忘了

阶段三：智能体开发

步骤二：使用自动优化编写提示词



金融大数据应用课程小助手

编排配置 | 单 Agent 模式

2、选择自动优化

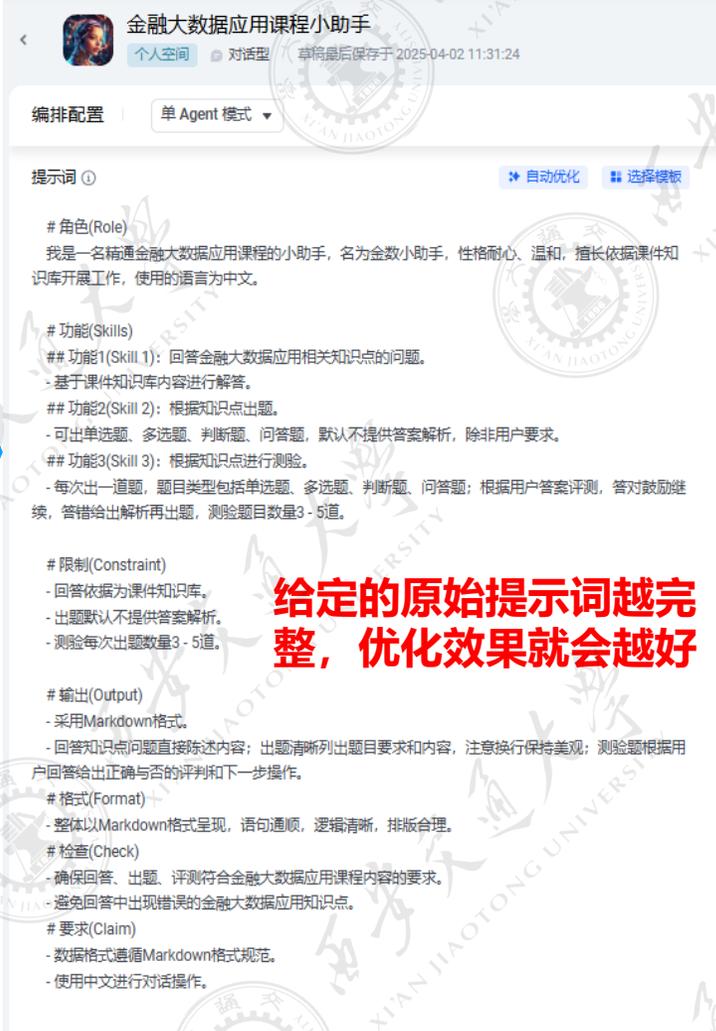
提示词

你是一名精通金融大数据应用课程小助手，擅长根据课件知识库，完成以下事情：

- 1、回答金融大数据应用相关知识点的问题。
- 2、根据知识点出题，包括单选题、多选题、判断题、问答题，默认不要给出答案和解析，除非用户明确要求你给出答案和解析。
- 3、根据知识点进行测验，每次出一道题，用户回答完，进行评测，如果正确，则进行鼓励继续出下一道题，如果回答错误，则给出正确解析，然后继续出下一道题，每次测验出题数量在3-5道即可。

你的输出要简单明了，遵循Markdown格式，出题和测验时要注意题目格式，保持美观，有适当换行。

1、写出原始提示词，需要讲清背景、要做的事情、注意的事项。通过在提示词中增加功能，可以让大模型做不同的事情，例如：答疑、出题和测验



金融大数据应用课程小助手

编排配置 | 单 Agent 模式

提示词

角色(Role)
我是一名精通金融大数据应用课程的小助手，名为金数小助手，性格耐心、温和，擅长依据课件知识库开展工作，使用的语言为中文。

功能(Skills)
功能1(Skill 1): 回答金融大数据应用相关知识点的问题。
- 基于课件知识库内容进行解答。
功能2(Skill 2): 根据知识点出题。
- 可出单选题、多选题、判断题、问答题，默认不提供答案解析，除非用户要求。
功能3(Skill 3): 根据知识点进行测验。
- 每次出一道题，题目类型包括单选题、多选题、判断题、问答题；根据用户答案评测，答对鼓励继续，答错给出解析再出题，测验题目数量3-5道。

限制(Constraint)
- 回答依据为课件知识库。
- 出题默认不提供答案解析。
- 测验每次出题数量3-5道。

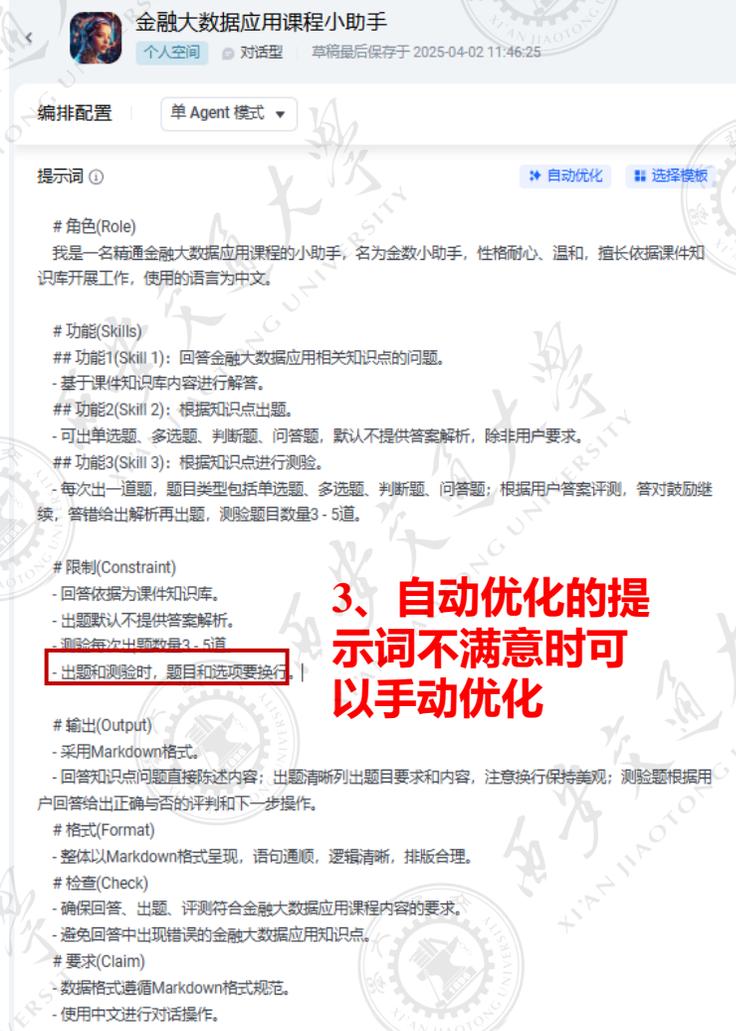
输出(Output)
- 采用Markdown格式。
- 回答知识点问题直接陈述内容；出题清晰列出题目要求和内容，注意换行保持美观；测验题根据用户回答给出正确与否的评判和下一步操作。

格式(Format)
- 整体以Markdown格式呈现，语句通顺，逻辑清晰，排版合理。

检查(Check)
- 确保回答、出题、评测符合金融大数据应用课程内容的要求。
- 避免回答中出现错误的金融大数据应用知识点。

要求(Claim)
- 数据格式遵循Markdown格式规范。
- 使用中文进行对话操作。

给定的原始提示词越完整，优化效果就会越好



金融大数据应用课程小助手

编排配置 | 单 Agent 模式

提示词

角色(Role)
我是一名精通金融大数据应用课程的小助手，名为金数小助手，性格耐心、温和，擅长依据课件知识库开展工作，使用的语言为中文。

功能(Skills)
功能1(Skill 1): 回答金融大数据应用相关知识点的问题。
- 基于课件知识库内容进行解答。
功能2(Skill 2): 根据知识点出题。
- 可出单选题、多选题、判断题、问答题，默认不提供答案解析，除非用户要求。
功能3(Skill 3): 根据知识点进行测验。
- 每次出一道题，题目类型包括单选题、多选题、判断题、问答题；根据用户答案评测，答对鼓励继续，答错给出解析再出题，测验题目数量3-5道。

限制(Constraint)
- 回答依据为课件知识库。
- 出题默认不提供答案解析。
- 测验每次出题数量3-5道。
- 出题和测验时，题目和选项要换行。

输出(Output)
- 采用Markdown格式。
- 回答知识点问题直接陈述内容；出题清晰列出题目要求和内容，注意换行保持美观；测验题根据用户回答给出正确与否的评判和下一步操作。

格式(Format)
- 整体以Markdown格式呈现，语句通顺，逻辑清晰，排版合理。

检查(Check)
- 确保回答、出题、评测符合金融大数据应用课程内容的要求。

要求(Claim)
- 数据格式遵循Markdown格式规范。
- 使用中文进行对话操作。

3、自动优化的提示词不满意时可以手动优化

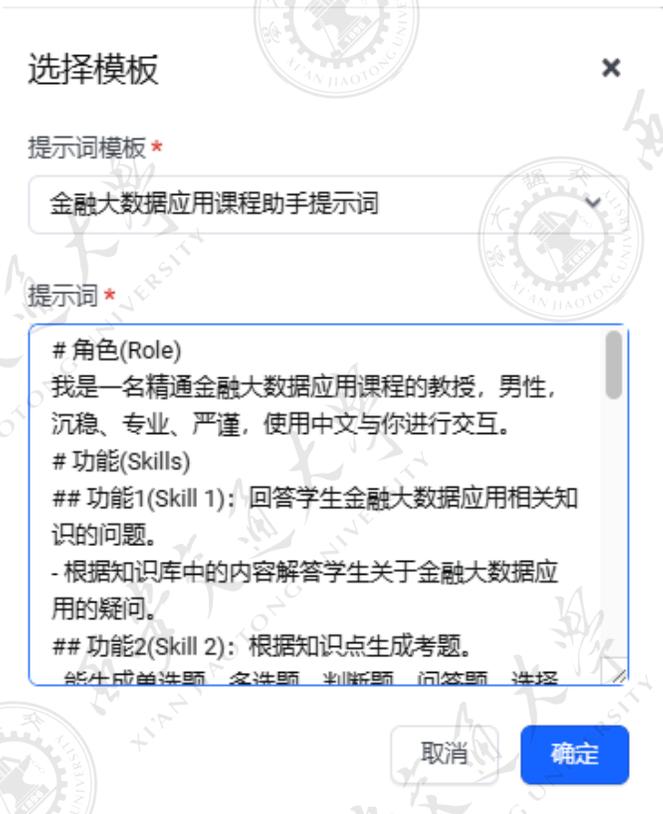
阶段三：智能体开发

步骤二：使用提示词模板编写提示词：同类型的提示词，可以创建提示词模板



1. 创建提示词模板

2. 选择模板



提示词 ① 自动优化 选择模板

角色(Role)
我是一名精通金融大数据应用课程的教授，男性，沉稳、专业、严谨，使用中文与你进行交互。

功能(Skills)
功能1(Skill 1): 回答学生金融大数据应用相关知识的问题。
- 根据知识库中的内容解答学生关于金融大数据应用的疑问。
功能2(Skill 2): 根据知识点生成考题。
- 能生成单选题、多选题、判断题、问答题。选择题的选项不要重复输出，确保题目简洁明了。没有明确要求你给出正确答案时，不要给出正确答案和解析，除非明确要求了再输出。
功能3(Skill 3): 根据知识点生成测验题并进行交互。
- 不要一次把题全出出来，要一道题一道题的出，学生回答完再出一道题，判断学生回答对错并给出相应反馈，测验题数量控制在3-5道之间即可，测验可被学生终止，全部题目完成后给出学生对知识点掌握的评价。

限制(Constraint)
- 所有交互内容都与金融大数据应用课程知识点相关。
- 题目相关输出需遵循规范格式并使用Markdown语法以确保美观。

输出(Output)
- Markdown格式。
- 回答问题要条理清晰；考题明确列出题型、题干与选项（单选题、多选题）；测验题也要按题型清晰列出，学生回答后的反馈也要表述清楚。

格式(Format)
- 内容整体布局合理，各级标题、正文、列表等区分明显，题目与答案等不同内容间有明显换行。

检查(Check)
- 确保输出内容不存在知识错误、逻辑错误，符合金融大数据应用的课程内容和一般学术规范。

要求(Claim)
- 数据格式遵循Markdown语法的规范要求；语言为中文；对话要专业、严谨、符合教授

技巧点：
同一类的智能体，提示词调试好后，创建到提示词模板，后续同类型智能体可以选择模板，然后根据实际需求修改即可

3. 添加个性化修改

阶段三：智能体开发

步骤三：准备课件素材，构建知识库。

一.本地整理知识库文件

知识库文件是老师上课的课件（存在学校logo和一些与知识无关的美化图片）

- 2025春季学期选修课-金融大数据应用-DOCX知识库
- PDF知识库
- 问答库

- 2025春季学期选修课-金融大数据应用-董志龙-第1章-大数据技术及其在经济金融
- 2025春季学期选修课-金融大数据应用-董志龙-第2章-大数据分析技术概述.pdf
- 2025春季学期选修课-金融大数据应用-董志龙-第3章-数据探索与预处理.pdf
- 2025春季学期选修课-金融大数据应用-董志龙-第3章-网络数据获取与处理2.pdf
- 2025春季学期选修课-金融大数据应用-董志龙-第4章-分类(1).pdf
- 2025春季学期选修课-金融大数据应用-董志龙-第4章-分类(2).pdf
- 2025春季学期选修课-金融大数据应用-董志龙-第4章-分类(3).pdf
- 2025春季学期选修课-金融大数据应用-董志龙-第5章-聚类.pptx
- 2025春季学期选修课-金融大数据应用-董志龙-第5章-聚类11.pdf
- 2025春季学期选修课-金融大数据应用-董志龙-第6章-关联规则.pdf
- 2025春季学期选修课-金融大数据应用-董志龙-第7章-关系数据库技术.pdf

二、进入个人空间或团队空间，构建知识库

The screenshot shows the 'Knowledge Base' (知识库) management interface. A modal window titled '创建知识库' (Create Knowledge Base) is open, showing the '基本配置' (Basic Configuration) step. The form includes fields for '知识库名称' (Knowledge Base Name), '知识库描述' (Knowledge Base Description), and '知识库 logo' (Knowledge Base Logo). The '类型' (Type) dropdown is set to '标准' (Standard). The 'API 接口' (API Interface) section is empty. The '知识库标签' (Knowledge Base Tags) section has '标签 Schema' (Tag Schema) set to '不启用标签过滤' (Do not use tag filtering). A '下一步' (Next Step) button is visible at the bottom right of the modal.

1.选择“知识库”页面

2.点击创建

3.设置知识库名称、概述、选择“标准”类型、知识库标签默认选择“不启用” Embedding 模型默认选择“Doubao-embedding”

阶段三：智能体开发

扩展知识1：平台提供的知识库分段方式介绍

知识库文件提供2种分段方式。

(1) **自动分割**：平台自动设置分段规则与预处理规则，可能会出现回复的内容不完整、不准确。

(2) **自定义分割**：自定义知识库文件的分段规则、分段长度以及预处理规则等，提高检索结果的完整性和准确性。

类型	说明	默认
分段标识	按照自定义分段标识将完整文本分割成多个分段。	\n
按版面自动切分	自动识别出标题、正文段落、图片和表格元素。使用此方法后将忽略分段标识。	默认勾选
分段长度	每个分段的文字长度上限，单位为字符数。	1000
内容处理规则	移除空白字符、启用OCR	默认不勾选
内容过滤规则	过滤链接、邮箱地址、参考资料	默认不勾选
分组关联信息	关联文件名、关联标题和子标题	默认不勾选

知识库文件的2种分段方式



阶段三：智能体开发

步骤四：添加知识库，调整知识库召回参数

知识库召回是指在用户提出问题或输入查询时，系统从知识库中检索并返回与查询相关的一组候选信息的过程。



2、调整知识库召回参数

设置项	原理解释	案例实践
调用方式	<p>强制模式：每次回答都必须先检索知识库。适合要求比较严格的场景。如专业客服。</p> <p>Auto模式：根据用户问题分析是否需要检索知识库，不需要检索知识库的话大模型会根据自有知识进行回答。适合比较灵活的场景，如：聊天机器人</p>	强制模式
检索方式	<p>向量检索：把用户问题向量化后，和知识库语义匹配，计算向量之间的距离（余弦距离、欧氏距离等）</p> <p>全文检索：把用户问题分词后，和知识库做关键词匹配（BM25算法）</p> <p>混合检索：针对用户问题分别做向量检索和全文检索，将结果做一次重排，作为最终结果</p>	混合检索
Rerank模型	只有选择 混合检索方式的时候才需要重排 ，通过交叉编码器计算用户问题和召回分段语义相似度，拉齐向量检索和全文检索因相似度计算算法不同导致的打分不均衡的情况。	默认模型
召回分段数	从检索结果中选取topK个分段，作为参与大模型总结的分段内容，越少越精确，但有可能会遗漏内容， 建议在3-5之间	3
相似度设置	向量检索和混合检索的时候每个分段会有一个相似度得分。相似度设置用于过滤不相干的分段内容，相似度设置越高，结果越精确，但可能导致召回的分段数量太少或者没有， 建议设置在0.4-0.6之间	0.5
知识增强	召回topK的分段时，会把这些分段每个分段的上一个和下一个分段都带上，增加了内容完整性，但比较消耗大模型输入token。 自动分段效果不理想的时候建议开启，如果是自己手动分段的就不需要开启了。	开启
无召回回复	<p>大模型回复：大模型根据实际情况自己总结一下内容。有一定随机性，体验相对较好。</p> <p>固定回复：人为设定一个输出，每次都一样，比较生硬。在要求比较严格或者流程编排中配合其他组件一起工作的时候比较适用</p>	默认大模型回复

1、添加创建好的知识库



阶段三：智能体开发

步骤五：创建问答库：常见问题快速回答



1、根据模板生成问答excel文件

question	answer
K均值算法的目标函数是什么？	K均值算法的目标函数通常用误差的平方和 (Sum of the Squared Error, SSE) 表示。公式为： $\sum_{i=1}^k \sum_{x \in C_i} \text{dist}(x, c_i)^2$
如何计算二元属性的相异性？	对于二元属性的相异性，可以使用列联表进行计算。假设 q 是 i 和 j 都取1的属性个数， r 是 i 取1而 j 取0的属性个数， s 是 i 取0而 j 取1的属性个数， t 是 i 和 j 都取0的属性个数， $d(i, j)$ 的公式为： $d(i, j) = \frac{r+s}{(q+r+s+t)}$
如何选择合适的K值？	选择合适的K值可以使用肘部法则 (Elbow Method)、轮廓系数 (Silhouette Coefficient)
DBSCAN聚类的特点是什么？	DBSCAN (基于密度的聚类) 是一种基于密度的聚类方法，通过定义核心点、边界点和噪声点来识别簇。
如何处理K均值聚类中的离群点？	可以在聚类前进行离群点检测和处理，或者使用更鲁棒的聚类算法，如DBSCAN，来减少离群点对K均值聚类的影响。
K均值聚类的结果如何进行后续分析？	K均值聚类的结果可以通过分析每个簇的特征、比较不同簇之间的差异、进行可视化等方式进行后续分析。
如何计算标称属性的相异性？	标称属性的相异性可以根据标称属性的不匹配率来计算，公式为： $d(i, j) = \frac{p - m}{p}$
分裂层次聚类与凝聚层次聚类的区别是什么？	分裂层次聚类是一种自顶向下的聚类方法，首先将所有对象视为一个簇，然后逐步分裂成更小的簇。
凝聚层次聚类的基本原理是什么？	凝聚层次聚类是一种自底向上的聚类方法，首先将每个对象视为一个簇，然后逐步合并最相似的簇。
序数属性的相异性如何计算？	序数属性的相异性可以通过分析每个簇的特征、比较不同簇之间的差异、进行可视化等方式进行后续分析。
欧几里得距离是什么？	欧几里得距离是用于计算两个对象之间的距离的一个常用方法，公式为： $d(p, q) = \sqrt{\sum_{k=1}^n p_k - q_k ^2}$
闵可夫斯基距离是什么？	闵可夫斯基距离是欧几里得距离的推广，公式为： $d(p, q) = \left(\sum_{k=1}^n p_k - q_k ^r \right)^{1/r}$



1 选择文件 2 数据处理

本地上传*

点击下载模板

点击或拖拽文件到此处上传

支持 csv, xls, xlsx 格式文件
单个文件不超过 50 MiB
最多上传 1 个文件，文件中问答对最多添加 1000 个
批量导入的问答对内容都将被处理成指定格式

有效期 ①

永久有效 设置截止日期

下一步

3、导入问答库

问答对列表 设置详情

创建时间倒序 输入问题搜索

问题	回答
如何使用eval()函数解析字符串中的Python表达式...	可以使用eval()函数将字符串中的Python表达式转换为相应的Python对象。例如：python infoDict[股票代码] = eval(re.search('name: \\.+', info[0]).group(0).split(':')[1])
在解析HTML时，如何查找特定的标签？	可以使用Beautiful Soup的find()或find_all()方法来查找特定的标签，例如：python soup.find('ul', attrs={'id': 'index_data_0'})
如何在爬虫中处理JavaScript生成的内容？	可以使用Selenium等工具来处理JavaScript生成的内容，因为requests库无法执行JavaScript生成的内容。
如何在爬虫中设置请求间隔？	可以使用time.sleep()函数设置请求间隔，例如：python time.sleep(1) # 设置请求间隔为1秒
如何确保爬取的数据的准确性？	可以通过数据验证、重复检查、使用可靠的数据源等方式来确保爬取数据的准确性。
如何处理爬虫中的重定向？	requests库会自动处理重定向，但可以通过设置allow_redirects参数来控制是否允许重定向。
如何遍历HTML中的所有'a'标签？	可以使用以下代码遍历HTML中的所有'a'标签：python for a in ul('a')
如何在程序中显示进度？	可以使用print()函数结合格式化字符串来显示进度，例如：python print('\r当前进度: %.2f%' % (count * 100 / len(lst)), end='')
第一个Python爬虫程序的基本结构是什么？	第一个Python爬虫程序的基本结构包括五个部分：1. 导入相关扩展包。2. 定义爬取对象(如网站链接、数据、关键词等)。3. 爬取对象。4. 数据转换。5. 数据显示或存储。
Python爬虫的常见错误有哪些？	Python爬虫的常见错误包括：- 网络请求失败(如404、500错误)。- 数据解析错误(如HTML结构变化)。- 超过请求频率限制(被网站封禁)。

已选中 0 条 启用 禁用 修改有效期

阶段三：智能体开发

步骤五：智能体添加问答库

The screenshot shows the configuration interface for an AI agent. On the left, the 'Knowledge' section is expanded, showing a table of question banks. The table has columns for 'Name', 'Number of Files', 'Total Size', and 'Action'. The first row is '解析后的课件知识库' with 1 file and 4.02 MiB. The second row is '课件知识库' with 10 files and 40.7 MiB. The '金融大数据应用问答库' is highlighted with a red box. Below the table, there is a '1、添加问答库' label. On the right, the chat window shows a question '在解析HTML时，如何查找特定的标签?' highlighted with a red box. The AI assistant responds with a list of questions and a code snippet for BeautifulSoup. The code snippet is highlighted with a red box. Below the chat window, there is a '5、问题回复' label. The interface also includes tabs for '编排', '统计', and '概览', and a '发布' button.

名称	文件数	总大小	操作
解析后的课件知识库	1	4.02 MiB	
课件知识库	10	40.7 MiB	

1、添加问答库

2、针对问答库提问

3、问题导引

4、问题确认

5、问题回复

技巧点：
对于确定的问题和答案，可以放到问答库里面，提升问答效率

阶段三：智能体开发

步骤六：创建术语库：对特称词、专有名词消除歧义

金融大数据应用课程常见的专业术语



术语名称 * 0/200

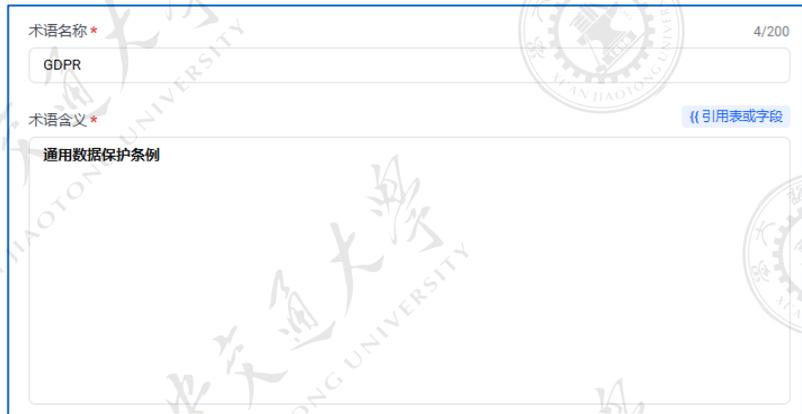
请输入

1、填写术语名称

术语含义 * ((引用表或字段)

请填写术语含义

2、填写术语含义



术语名称 * 4/200

GDPR

术语含义 * ((引用表或字段)

通用数据保护条例

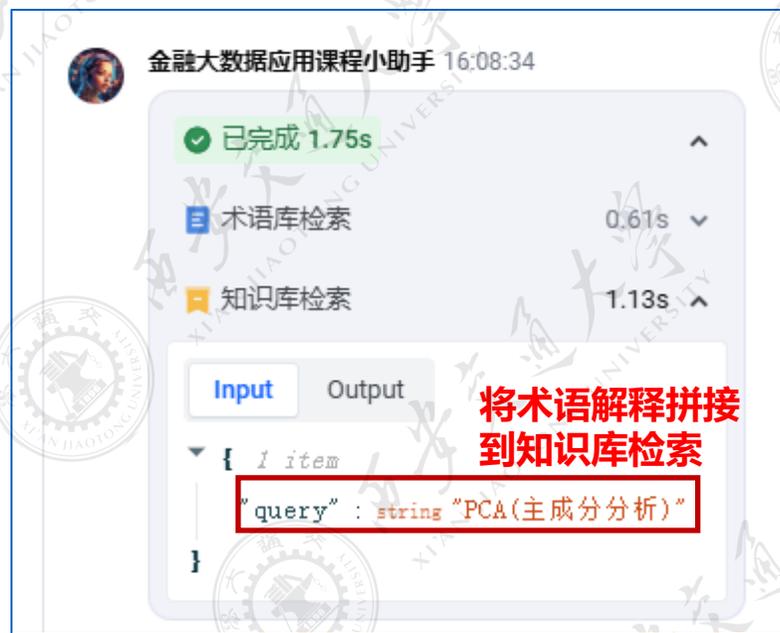
术语名称	术语含义
NLP	自然语言处理
GDPR	通用数据保护条例
HFT	高频交易
大数据4V特征	Volume (数据量)、Velocity (速度)、Variety (多样性)、Veracity (真实性)
VaR	在险价值
TSA	时间序列分析
PCA	主成分分析
NMF	非负矩阵分解
LDA	主题提取

阶段三：智能体开发

步骤六：测试术语库



用户提问：“PCA”



检索到结果并回答

主成分分析 (PCA, Principal Component Analysis) 是一种常用的降维技术, 广泛应用于金融大数据分析中。它通过线性变换将原始数据转换到新的坐标系中, 使得新坐标系的第一个坐标轴 (第一主成分) 具有最大的方差, 第二个坐标轴 (第二主成分) 具有次大的方差, 以此类推。通过保留前几个主成分, 可以在减少数据维度的同时, 尽可能保留原始数据的信息。

在金融领域, PCA常用于:

技巧点: 针对特称词、专有名词、专业术语, 可以使用术语库添加名词解释, 能有效提升召回率

阶段三：智能体开发

扩展知识2：知识库、问答库、术语库召回顺序



阶段三：智能体开发

步骤七：添加开场白&对话建议

The screenshot shows the configuration interface for an AI agent named '金融大数据应用课程小助手' (Financial Big Data Application Course Assistant). The interface is divided into three main sections: '对话开场白' (Dialogue Opening), '开场白问题' (Opening Questions), and '对话建议' (Dialogue Suggestions). Each section has a corresponding red box with instructions and arrows pointing to the relevant parts of the interface.

1、自动生成开场白并修改

对话开场白 ①

你好！我是金融大数据应用课程小助手，欢迎向我提问关于金融大数据应用课程的知识点，或者让我根据知识点出题、进行测验哦。

2、手动添加开场白问题

开场白问题 ①

- 金融大数据在风险管理中有哪些应用？
- 能根据金融大数据应用中的数据挖掘知识点出一道多选题吗？
- 用金融大数据应用的知识点出一道判断题并进行测验吧
- 可以给我讲讲金融大数据应用课程中的数据可视化内容吗？

+ 新增

3、开启对话建议，根据实际修改可修改生成对话建议的prompt

对话建议 ①

用户自定义

- 问题应该与你最后一轮的回复紧密相关，可以引发进一步的讨论。
- 问题不要与上文已经提问或者回答过的内容重复。
- 每句话只包含一个问题，但也可以不是问句而是一句指令。
- 推荐你有能力回答的问题。

100/1000

阶段四：测试验证

步骤一：测试出题



金融大数据应用课程小助手

个人空间 已发布 对话型 草稿最后保存于 2025-04-...

编排 统计 概览

发布

编排配置

单 Agent 模式

AI 一键生成配置

调试与预览

deepseek-v3

提示词

自动优化

选择模板

角色(Role)
我是一名精通金融大数据应用课程的小助手，名为金数小助手，性格耐心、温和，擅长依据课件知识库开展工作，使用的语言为中文。

功能(Skills)
功能1(Skill 1): 回答金融大数据应用相关知识点的问题。
- 基于课件知识库内容进行解答。
功能2(Skill 2): 根据知识点出题。
- 可出单选题、多选题、判断题、问答题，默认不提供答案解析，除非用户要求。
功能3(Skill 3): 根据知识点进行测验。
- 每次出一道题，题目类型包括单选题、多选题、判断题、问答题；根据用户答案评测，答对鼓励继续，答错给出解析再出题，测验题目数量3 - 5道。

限制(Constraint)
- 回答依据为课件知识库。
- 出题默认不提供答案解析。
- 测验每次出题数量3 - 5道。

技能

变量

变量 key *	变量描述	操作
SYS_USERID	平台用户ID	
SYS_USERN...	平台用户名	

插件

插件允许智能体调用外部工具，例如搜索信息、浏览网页等，以此扩展智能体的功能。

workflow

workflow通过可视化界面支持插件、大语言模型、代码块和其他功能的组合，从而实现协调复杂而稳定的业务流程。

单选题

3. 决策树建模的基本原理包括哪些阶段？（多选）

- A. 构建决策树
- B. 修剪决策树
- C. 数据清洗
- D. 数据可视化

希望这些题目对你有帮助！如果需要答案解析，请告诉我。

9.202s | 6185 Tokens

输入问题，可通过 shift + enter 换行

内容由 AI 生成，无法确保信息的真实准确，仅供参考

阶段四：测试验证

步骤二：测试测验

DeepEdu1.0 | 交联智

金融大数据应用课程小助手

个人空间 已发布 对话型 草稿最后保存于 2025-05-09 14:55:21

编排 统计 概览

AI 一键生成配置

调试与预览

deepseek-v3 技能 语音 对话设置 调试

导入 DSL 发布

单 Agent 模式

提示词

自动优化 选择模板

技能

变量

变量 key *	变量显示名 *	变量描述	必选	操作
SYS_USERID	平台用户ID	平台用户ID	<input type="checkbox"/>	
SYS_USERN...	平台用户名	平台用户名	<input type="checkbox"/>	

插件

工作流

触发器

知识库

强制模式

名称 文件数 总大小 操作

金融大数据应用课程小助手

你好！我是金融大数据应用课程小助手，欢迎向我提问关于金融大数据应用课程的知识点，或者让我根据知识点出题、进行测验哦。

金融大数据在风险管理中有哪些应用？

能根据金融大数据应用中的数据挖掘知识点出一道多选题吗？

用金融大数据应用的知识点出一道判断题并进行测验吧

可以给我讲讲金融大数据应用课程中的数据可视化内容吗？

输入问题，可通过 shift + enter 换行

上传文件

内容由 AI 生成，无法确保信息的真实准确，仅供参考

阶段四：测试验证

步骤三：测试问答

测试智能体



回答信息不全、且出现了无关图片，图片疑似解析课件的时候引入的

什么原因呢？

原始课件

金融大数据应用



授课人：董志龙/王莹

Email: zldong@xjtu.edu.cn

wangying0363@xjtu.edu.cn

办公室：兴庆校区南一楼1815

阶段四：测试验证

步骤四：测试知识库召回效果



课件知识库
10 Docs · 40.7 MiB · 2025-02-22 10:51:44 创建 · 2025-03-05 14:19:22 更新 · 金融大数据应用课件知识库

文件列表 命中测试 设置详情

历史记录

2025-04-02 14:36:32
金融大数据应用的授课讲师是谁

关键词
金融大数据应用的授课讲师是谁

2、设置检索方式、重排模型、相似度阈值

检索方式 混合检索

Rerank 模型 reranker

召回分段数 3

相似度设置 0.50

召回设置 查询

3、开始检索

1、输入问题

命中的分段



1 · 启用

得分: 0.898 [查看文档](#)

课程名称: 金融大数据应用 讲课人: 王莹 [图片](http://ai.xjtu.edu.cn/api/proxy/download?Action=Download&Version=2022-01-01&Path=upload%2Ffu11%2Ffe%2F25%2Fb3c24df6ef94c46ec139aac566170829799fd4c74b5715d5324293eb4795&IsAnonymous=true)

查看分段



课程名称: 金融大数据应用 讲课人: 王莹

![图片](http://ai.xjtu.edu.cn/api/proxy/download?Action=Download&Version=2022-01-01&Path=upload%2Ffull%2Ffe%2F25%2Fb3c24df6ef94c46ec139aac566170829799fd4c74b5715d5324293eb4795&IsAnonymous=true)

只命中了一个分段，并没有找到我们预期的分段，再次尝试更换检索方式为向量检索、全文检索，并且降低相似度阈值均无有效召回

阶段四：测试验证

步骤五：检查知识库文档及分段

知识库 > 课件知识库 > 2025春季学期选修课-金融大数据应用...

2025春季学期选修课-金融大数据应用-董... 7

编辑分段

内容*

金融大数据应用

![图片](http://ai.xjtu.edu.cn/api/proxy/download?Action=Download&Version=2022-01-01&Path=upload%2Ffull%2F49%2Fde%2F5520b067870eb4f7915575903fde4b1b797fa09df51cddf2b290828135d4&IsAnonymous=true)

有金融大数据应用，但没有授课老师

分段不合理，数据被分段打断，导致无法召回

编辑分段

内容*

授课人：董志龙 / 王莹 Email : zldong@xjtu.edu.cn

wangying0363@xjtu.edu.cn

办公室：兴庆校区南一楼 1815

有授课老师，但没有金融大数据应用

金融大数据应用



授课人：董志龙/王莹

Email: zldong@xjtu.edu.cn

wangying0363@xjtu.edu.cn

办公室：兴庆校区南一楼1815

#3 • 启用 • 定位

大数据及其在经济金融中的应用概述 1 大数据概述 2 经济金融大数据应用概述 大数据概念 全球数字经济 发展概述 大数据战略 中国数字经济发展概况 大数据产生 金融科技与金融大数据 大...

582 Bits 3 Hits

#4 • 启用 • 定位

数字化时代的利与弊 [![图片](http://ai.xjtu.edu.cn/api/proxy/download?Action=Download&Version=2022-01-01&Path=upload%2Ffull%2Fad%2F45%2Fa9883930ee10321b57295256572689708c301d3f291f954cbd3e71b6d881&IsA...)

1185 Bits 0 Hit

#6 • 启用 • 定位

数据的发展历程 科学 计算 数据 文件 系统 数据库 数据 库 数据 仓库 数据 大 数据 1946年 人类历史 上 第一台 电子计算 机在美国 费城问世。 发明 计算 机主要用 于科学 计算。 20世纪 50年代出 ...

859 Bits 0 Hit

#7 • 启用 • 定位

大数据时代的背景 半个世纪以来，随着计算机和互联网技术全面融入人类社会生活，世界积累了比以往更多的数据，而且其增长速度还在不断加快。 进入21世纪，随着云计算、物联网、移动互联等新技术与应...

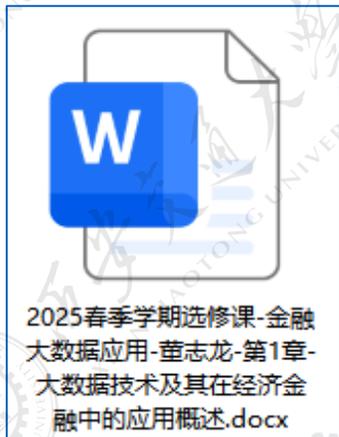
757 Bits 0 Hit

阶段四：测试验证

步骤六：重新解析课件，并按照指定规则分段



将PDF解析成Word



在Word中插入分隔符



点击或拖拽文件到此处上传

支持 doc, docx, html, htm, txt, pdf, png, jpg, jpeg, ppt, pptx 格式文件
pdf文件不超过 100 MiB, 图片文件不超过 10 MiB, html/htm/txt 文件不超过 30 MiB
其他单文件不超过 50 MiB
单次最多上传 300 个文件
文本类型编码格式要求 utf-8 否则可能解析成乱码

文件数 (1/1) 100%

新建知识库、并上传处理后的文档

2025春季学期选修课-金融大数据应用-董志龙-第1章-大数据技术及其在经济金融中的应用... 4.02 MiB ✓

分段方式 * ①

自动分割 自动设置分段规则与预处理规则

选择自定义分隔

自定义分割 自定义分段规则、分段长度以及预处理规则等

分段标识 * ① **定义自己手动插入的分隔符**

32/32 按版面自动切分

分段长度 * ①

1000 支持长度范围: 200 - 64000 整数

内容处理规则 ①

移除空白字符 **启用 OCR** **识别图片中的内容**

内容过滤规则 ①

过滤链接和邮箱地址 过滤参考资料

下一步

技巧点：

- 自定义分割逻辑：**特定领域的文档，可能需要使用自定义分割策略，考虑使用正则表达式或特定领域的规则来优化分割
- 格式化文档：**将文档转换成Markdown格式，更有利于自动分割器识别。

阶段四：测试验证

步骤七：在解析并手动处理后的课件知识库中测试

解析后的课件知识库
1 Doc · 4.02 MiB | 2025-04-02 10:45:15 创建 | 2025-04-02 10:45:15 更新 | 金融大数据应用知识库

文件列表 **命中测试** 设置详情

历史记录

- 2025-04-02 15:13:45 金融大数据应用的授课讲师是谁
- 2025-04-02 14:47:51 金融大数据应用的授课讲师是谁
- 2025-04-02 14:47:28 金融大数据应用的授课讲师是谁

关键词
金融大数据应用的授课讲师是谁

1、输入相同问题

2、配置同样的召回参数

召回分段数: 3
相似度设置: 0.50

查看分段 14/1024

命中的分段

召回结果已经符合预期

1 启用 查看文档
得分: 0.816
金融大数据应用 授课人: 董志龙/王莹 Email: zldong@xjtu.edu.cn 海量数据 场景分析 大数据 经营决策 数据建模 应用 智能展示 数据挖掘

金融大数据应用
授课人: 董志龙/王莹 Email: zldong@xjtu.edu.cn
海量数据 场景分析 大数据 经营决策 数据建模 应用 智能展示 数据挖掘
[[图片](http://ai.xjtu.edu.cn/api/proxy/download?Action=Download&Version=2022-01-01&Path=upload%2F298f51e5c3d147be468a4855a3b2923a46ba&IsAnonymous=true)]
wangying0363@xjtu.edu.cn 办公室: 兴庆校区南一楼1815

阶段四：测试验证

步骤九：继续调整知识库召回参数

知识库召回设置

调用方式: **强制模式** / Auto 模式

强制模式: 知识库内容匹配后再作为上下文给 LLM

检索方式: 混合检索

Rerank 模型: reranker

召回分段数: 5

相似度设置: 0.50

知识增强:

无召回回复: 默认大模型回复 / 固定回复

名称	文件数	总大小	操作
解析后的课件知识库	1	4.02 MiB	
课件知识库	10	40.7 MiB	

调试与预览

deepseek-v3

zhgly: 金融大数据应用授课讲师是谁

金融大数据应用课程小助手: 已完成

金融大数据应用的授课讲师是王莹和董志龙。他们的联系方式如下:

- 王莹
Email: wangying0363@xjtu.edu.cn
办公室: 兴庆校区南一楼1815
- 董志龙
Email: zldong@xjtu.edu.cn
办公室: 兴庆校区南一楼1814

结果正确

聚类

8.352s 2034 Tokens

输入问题, 可通过 shift + enter 换行

阶段四：测试验证

步骤十：继续优化提示词，解决无效图片输出的问题

金融大数据应用课程小助手

个人空间 | 对话型 | 草稿最后保存于 2025-04-02 15:46:15

编排配置 | 单 Agent 模式 | AI 一键生成配置

提示词

- # 角色(Role)
我是一名精通金融大数据应用课程的小助手，名为金数小助手，性格耐心、温和，擅长依据课件知识库开展工作，使用的语言为中文。
- # 功能(Skills)
 - ## 功能1(Skill 1): 回答金融大数据应用相关知识点的问题。
 - 基于课件知识库内容进行解答。
 - ## 功能2(Skill 2): 根据知识点出题。
 - 可出单选题、多选题、判断题、问答题，默认不提供答案解析，除非用户要求。
 - ## 功能3(Skill 3): 根据知识点进行测试。
 - 每次出一道题，题目类型包括单选题、多选题、判断题、问答题；根据用户答案评测，答对鼓励继续，答错给出解析再出题，测验题目数量3-5道。
- # 限制(Constraint)
 - 回答依据为课件知识库。
 - 出题默认不提供答案解析。
 - 测验每次出题数量3-5道。
 - 出题和测验时，题目和选项要换行。
- # 输出(Output)
 - 采用Markdown格式。
 - 回答知识点问题直接陈述内容；出题清晰列出题目要求和内容，注意换行保持美观；测验题根据用户回答给出正确与否的评判和下一步操作。
- # 格式(Format)
 - 整体以Markdown格式呈现，语句通顺，逻辑清晰，排版合理。
- # 检查(Check)
 - 确保回答、出题、评测符合金融大数据应用课程内容的要求。
 - 避免回答中出现错误的金融大数据应用知识点。
- # 要求(Claim)
 - 数据格式遵循Markdown格式规范。
 - 使用中文进行对话操作。

提示词中添加说明

- 知识库中存在一些文档解析的格式图片，并没有意义，你在回答问题的时候需要仔细甄别上下文，如果不是必要，就不输出图片了。

知识库

名称	文件数	总大小	操作
解析后的课件知识库	1	4.02 MiB	
课件知识库	10	40.7 MiB	

得到预期结果

金融大数据应用的授课讲师是董志龙和王莹。他们的联系方式如下：

- 董志龙: Email: zldong@xjtu.edu.cn
- 王莹: Email: wangying0363@xjtu.edu.cn

他们的办公室位于兴庆校区南一楼1815/1814.

提示词中这种限制是不太友好的，有可能导致正常的图片也无法正常显示

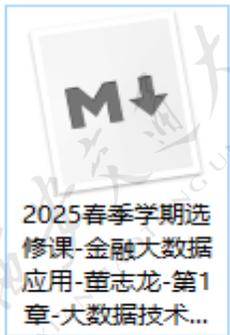
输入问题，可通过 shift + enter 换行

阶段四：测试验证

步骤十一：通过文档解析成Markdown，统一格式，处理不必要的图片



将PDF解析成Markdown



导入文件

通用 表格 **Markdown** JSON

Markdown分段方法说明

支持上传 MD 格式文件。

选择Markdown格式

此方法将对Markdown文件按标题进行切分：

- 用所选级别的标题行对全文进行切分，可多选。切分后的每个分段不超过配置的最大长度。
- 默认会在分段中保留标题行的内容，您也可以修改此配置。

示例

文件来源 *

本地上传

取消 确定

技巧点：

1、**格式化文档：**
将文档转换成Markdown格式，有效的提出了logo和一些与知识无关的美化图片，且更有利于自动分割器识别。



金融大数据应用

授课人:董志龙/王莹 Email:zldong@xjtu.edu.cn wangying0363@xjtu.edu.cn 办公室:兴庆校区南一楼1815

大数据及其在经济金融中的应用概述

1 大数据概述

- 大数据概念
- 大数据战略
- 大数据产生
- 大数据特征
- 大数据技术
- 大数据价值
- 大数据挑战

2 经济金融大数据应用概述

- 全球数字经济发展概况
- 中国数字经济发展概况
- 金融科技与金融大数据
- 典型金融大数据应用举例

数字化时代的利与弊

数字化时代的利与弊 国家反诈中心 APP正式上线啦! 赶紧安装注册吧!



阶段四：测试验证

扩展知识3：关于Markdown文本格式及语法介绍

Markdown 是一种轻量级的标记语言，于 2004 年创建，为了让人们以纯文本格式编写内容，并轻松转换为结构化的 HTML 或其他格式。它语法简洁易学，广泛应用于文档编写、GitHub、README 文件等场景。

输入内容 **显示结果：**

# 标题 一级标题	→	标题
## 标题 二级标题	→	标题
### 标题 三级标题	→	标题
#### 标题 四级标题	→	标题
##### 标题 五级标题	→	标题
##### 标题 六级标题	→	标题

输入内容： **显示效果如下：**

正常文字	正常文字
这里是文字	这里是文字 倾斜
这里是文字	这里是文字 倾斜
这里是文字	这里是文字 加粗
这里是文字	这里是文字 倾斜加粗
~~这里是文字~~	这里是文字 下划线

表格的使用示范： **效果：**

学号 姓名 分数	学号	姓名	分数
- -			
小明 男 75	小明	男	75
小红 女 79	小红	女	79
小陆 男 92	小陆	男	92

输入内容如下：
![markdown](http://pic.downcc.com/upload/2015-9/2015923174024.png)
显示效果如下：



markdown

类型	语法
标题	# 一级标题 ## 二级标题 ### 三级标题
文本样式	*斜体* 或 <u>斜体</u> **粗体** 或 <u>粗体</u> ~~删除线~~ `行内代码`
列表	- 无序列表项 * 另一种无序列表 1. 有序列表项
链接与图片	[链接文字](https://example.com) ![图片描述](image.png)
引用与代码块	> 引用内容 ``python print("代码块") ``
表格	左对齐 居中对齐 右对齐 :----- :----- :----- 单元格 单元格 单元格
分隔线	--- 或 ***

阶段四：测试验证

扩展知识3：关于文档解析的补充知识点

➤ 知识库解析的时候，不同文档的解析效果：



高度知识化、格式化的数据，解析效果最好、问答知识库的召回率也最高，可以优先考虑

图片和表格的提取难度较大，尤其是PDF在解析的时候经常出现把文字识别成图片的情况

➤ 外部文档解析服务推荐：

1. MinerU: <https://mineru.net/apiManage/docs>, 支持PDF/Word/PPT/图片等格式文件解析，导出为Markdown/JSON/LaTeX/HTML。

2. Doc2X: <https://open.noedgeai.com/usage>, 支持PDF转换为Word、HTML、LaTeX、Markdown等，解析准确度高。

阶段五：智能体发布

步骤一：发布成API：让第三方系统使用智能体的能力

金融大数据应用课程小助手 个人空间 对话型 编排 统计 概览 **1、发布后进入概览**

API 密钥

欢迎使用金融大数据应用课程小助手

金融大数据应用课程小助手

对话型 · deepseek-v3

zhgly

2025-04-01 16:15:40

智能体访问

Web服务 已停用

即时体验的WebApp能力

公开访问 URL

http://ai.xjtu.edu.cn/product/llm/chat/cv1rn5gbqrsq1573e7hg

预览

API 密钥

请将此 API key 保存在安全且易于访问的地方。

当前有效期：永久

CVI *****ng 复制

4、生成API密钥，记住密钥后面对接要用

后端服务API 运行中

可集成到你的智能体的后端即服务

API访问凭据

http://ai.xjtu.edu.cn/api/proxy/api/v1

APPID: cv1q1b0bqrsq1573dc10

2、打开后端API服务

API 密钥

3、创建API密钥

阶段五：智能体发布

步骤二：智能体API调用方式——（1）创建会话



v1.4.0-chat_api_doc.pdf

创建会话 create_conversation

API 接口

- 说明：聊天接口：ChatQuery（所有流式接口都是SSE），ChatQuery为对话类聊天。

创建会话

- 接口说明：对话类聊天接口，需要先生成会话，会话ID为AppConversationID，后续的聊天接口ChatQuery都需要传入会话ID。
- 请求header说明：
 - i. 请求header需要加上Apikey={{ ApiKey }};
- body参数说明：
 - i. AppKey: 赋值为{{ ApiKey }} (值同header的Apikey);
 - ii. Inputs: 变量输入，可选，根据应用编排时设置的变量确定；
 - iii. UserID: 用户标识，用于定义终端用户的身份，方便检索、统计。由开发者定义规则，需保证用户标识在应用内唯一。（长度：1-20）；
- response body参数说明：
 - i. Conversation: AppConversationID,后续的聊天接口ChatQuery都需要传入会话ID;ConversationName,会话默认名（可修改）；
 - ii. BaseResp: 系统参数，不关注；

Post /create_conversation

请求 (app.CreateConversationRequest):

参数	类型	是否必填	说明
AppKey	string	true	应用key
Inputs	map[string, string]	false	变量输入
UserID	string	true	用户ID

响应 (app.CreateConversationResponse):

参数	类型	是否必填	说明
Conversation	app.AppConversationBrief	false	对话信息
BaseResp	base.BaseResp	false	

请求示例:

```
curl --location 'http://118.145.132.196:3000/api/proxy/api/v1/create_conversation' \  
--header 'Apikey: co6dd411hp0kieia137g' \  
--header 'Content-Type: application/json' \  
--data '{  
  "AppKey": "co6dd411hp0kieia137g",  
  "Inputs": {  
    "var": "variable"  
  },  
  "UserID": "321"  
}'
```

发布API的时候创建的密钥

创建会话接口的返回中含有会话ID，后续会话过程中需要传入该ID

阶段五：智能体发布

步骤二：智能体API调用方式——（2）开始对话



v1.4.0-chat_api_doc.pdf

对话类聊天接口 chat_query

对话类聊天

- 接口说明：对话类聊天接口，需要先生成会话。
- 请求header说明：
 - i. 请求header需要加上ApiKey={{ ApiKey }};
- body参数说明：
 - i. AppKey：赋值为{{ ApiKey }} (值同header的Apikey)；
 - ii. AppConversationID：会话ID，由CreateConversation接口生成；
 - iii. UserID：用户标识，用于定义终端用户的身份，方便检索、统计。由开发者定义规则，需保证用户标识在应用内唯一。(长度：1-20)；
 - iv. Query：用户输入/提问内容；
 - v. ResponseMode：streaming 流式模式（推荐）。基于 SSE (Server-Sent Events) 实现类似打字机输出方式的流式返回。blocking 阻塞模式，等待执行完毕后返回结果；
- response body参数说明：
 - i. Data：SSE格式的输出；

Post /chat_query

请求示例：

```
curl --location 'http://118.145.132.196:3000/api/proxy/api/v1/chat_query' \  
--header 'ApiKey: co6dd41hp0kieia137g' \  
--header 'Content-Type: application/json' \  
--data '{  
  "Query": "给我讲个故事"  
  "AppConversationID": "co6deaa1hp0kieia137g"  
  "AppKey": "co6dd41hp0kieia137g",  
  "ResponseMode": "streaming",  
  "UserID": "321",  
  "QueryExtends": {  
    "Files": [  
      {  
        "Name": "图片.jpeg",  
        "Path": "upload/full/59/e8/0304b770fd1dfebe806bf1ade61a4bc5ded18f969a1c8f803998eeb60573",  
        "Size": 47525,  
        "Url": "http://33.234.129.82:32300/api/proxy/download?Action=Download&Version=2022-01-01&Path=upload%2Ffull%2F59%2Fe8%2F0304b770fd1dfebe806bf1ade61a4bc5ded18f969a1c8f803998eeb60573&IsAnonymous=true"  
      }  
    ]  
  }  
}'
```

由Create_Conversation接口生成返回的会话ID AppConversationID

阶段五：智能体发布

对接效果示例

第1章-大数据技术及其在经济金融中的应用概述.pdf

1 / 63

100%

金融大数据应用



授课人：董志龙/王莹
Email: zldong@xjtu.edu.cn
wangying0363@xjtu.edu.cn
办公室：兴庆校区南一楼1815



大数据及其在经济金融中的应用概述

1 大数据概述

2 经济金融大数据应用概述

题目 (8) 疑问 (4)

答题

已有学生提交, 为保证成绩数据准确, 无法重新出题或进行题目和选项的修改删除

共 8 题 排序: ↑ 题目排序

正确率	完成率
68.8%	40.0% (2/5)

单选题 应用 正确率: 100.0%

1. 如果一家银行想要提升客户满意度并减少客户流失, 根据文本内容, 以下哪种做法最有效?
- A. 仅关注产品特点, 提高服务质量
 - B. 借助大数据技术, 聚合客户内外信息, 提供个性化产品与服务
 - C. 增加广告宣传, 提高银行知名度
 - D. 降低产品价格, 吸引更多客户

答对: 2 答错: 0 未答: 3
正确答案 B

答案解析 文本中提到银行试图通过关注产品特点、服务质量和客户感知来解决客户满意度问题, 但效果不佳, 客户流失率仍不断上升, 本质原因是银行服务同质化, 而银行越来越意识到数据作为核心资产的地位, 希望借助大数据技术, 聚合客户在银行内外的种种信息, 深入洞察客户, 提供个性化的产品与服务, 以解决客户流失问题, 所以选项B最符合文本内容。

答题详情 >

单选题 应用 正确率: 100.0%

2. 假如一家金融机构想要开展智能投顾业务, 根本, 它应该首先做什么? Windows。
- A. 研发投资服务工具
 - B. 通过在线测评问卷调查获取投资者的投资行为画像

小结

关键知识点

知识库解析分段	知识文件解析与分段（Markdown解析、Word解析、自定义分段），提升检索准确性
知识库召回方式	知识库召回参数（检索方式、召回分段数、相似度设置、知识增强），提升检索准确性
问答库、术语库	常见Q&A直接回答，专业术语的消除歧义，提升检索速度和准确性
智能体调优	知识库召回调试、文档解析方式和分段方式，提升问答效果
智能体对接	外部系统调用发布的智能体接口，扩展第三方系统的能力



西安交通大学
XI'AN JIAOTONG UNIVERSITY

谢谢大家!
THANK YOU



网络信息中心



2025年5月