



SPM 2023

ChatGLM 金融大模型挑战赛

——基于生成式AI与判别式AI相融合 的上市公司年报金融问答模型方案

龙盈战队：闫括、刘锦霖、王彦博（队长）

华夏银行&龙盈智达

项目答辩说明

赛道	模型开发赛道
赛题	SMP 2023 ChatGLM金融大模型挑战赛
项目进展程度	进展中项目

目录 CONTENTS

1

团队介绍

2

赛题分析

3

技术方案

4

业务成效

5

应用价值

目录 CONTENTS

1

团队介绍

2

赛题分析

3

技术方案

4

业务成效

5

应用价值

赛题分析

1.比赛背景:

本次ChatGLM 金融大模型挑战赛要求参赛选手**通过解析给定的企业年报PDF数据，并以ChatGLM模型为中心构建一个问答模型，回答用户的金融相关的问题。**本次比赛评估模型能力按照涉及问题的复杂程度大体分为初级、中级、高级三种类型。

- a.初级：数据基本查询，包含财务报告基本数据查询、公司基础信息查询等问题。
- b.中级：数据统计分析查询，包含财务报告基本数据的计算等问题。
- c.高级：开放性问题，包含基于某特定公司的开放域问题和完全开放域问题。

2.赛题分析:

首先使用ChatGLM原始模型结合粗粒度的数据，初步测试模型在QA问答上存在的问题。

问题1

存在问题

6b模型对于较长文本的**抽取能力较差**



解决方案

有必要压缩上下文信息，提供给大模型关键信息

问题2

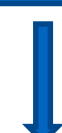
大模型**容易产生幻觉**



需要采取设计合理的Prompt、调节模型Temperature、微调、后处理等手段

问题3

信息**定位不准确**



不同类问题出现在年报不同位置：

- 计算型、数值查询型问题对应财务报告中相关报表。
- 公司信息类问题对应公司简介类报表。
- 基于公司开放域问题涉及非结构化数据。

问题4

测试问题样例**存在噪声**，影响开放域问题相关性文本检索



问题中的关键短语提取对于提升 开放域问题检索准确率较为重要。

- "Question": "根据2020年****有限公司的年报，给我简要介绍一下报告期内公司的社会责任工作情况？"
- "Keywords": **社会责任工作情况**

赛题分析

3.评估指标

$\begin{cases} \max_{\text{similar}}(\text{sentence1}, \text{sentence2}, \text{sentence3}), \\ 0.25 + 0.25 + \max_{\text{similar}}(\text{sentence1}, \text{sentence2}, \text{sentence3}) * 0.5, \\ 0.25 + 0 + \max_{\text{similar}}(\text{sentence1}, \text{sentence2}, \text{sentence3}) * 0.5, \\ 0, \end{cases}$	无基础信息及关键词 基础信息正确，关键词正确 基础信息正确，关键词错误 基础信息错误
---	---

比赛评估指标理解：

- 1.在图中，Prom_answer是基础信息，基础信息错误不给分。
- 2.若基础信息正确，需要结合基础信息、关键词、文本相似度进行综合评分，因此文本生成的格式也要遵从一定的规范。

```
{ "Question" : "2019年**银行财务费用是多少元?" ,  
  
  "Prompt": {"财务费用": "12345678.9元", "Key_word": "财务费用、2019", "Prom_answer": "12345678.9元"},  
  "Answer": "2019年**银行财务费用是12345678.9元。"} 
```

通过以上分析可知，基础信息占分比重较高，并且6b模型对于表格的识别精准度不高，因此采取以下策略：

- 01 根据表格的表头和行数据，将**报表转化为文本描述**。
- 02 使用Few-shot Learning，**定义【回答示例】**，让模型严格按照格式作答。

目录 CONTENTS

1

团队介绍

2

赛题分析

3

技术方案

4

业务成效

5

应用价值

技术方案

➤ 整体效果

复赛C榜

排名	参与者	组织	score	type1Score	type2Score	type3-1Score	type3-2Score
1	🐼馒头科技	Manteia/京发科技/龙程电子	88.62	91.64	91.90	77.06	89.54
2	🐼安硕硕眼探企	上海安硕信息技术股份有限公司	88.33	92.97	93.09	71.31	89.44
3	🐼流宝真人	杭州大道一以科技有限公司	87.82	91.35	90.89	75.89	88.79
4	🐼chatglm反卷总局	腾讯	87.23	87.33	92.55	75.22	89.64
5	🐼吃辣子	饺子研究院	86.32	83.73	92.93	75.15	89.97
6	🐼nsddd	中国科学院	85.89	85.14	91.30	74.24	89.83
7	🐼龙盈战队	龙盈智达（北京）科技有限公司	85.85	86.36	88.67	77.61	89.50
8	🐼Tablsabaopilong	上海交通大学	85.04	86.02	88.69	73.77	89.97

复赛B榜

排名	参与者	组织	score	type1Score	type2Score	type3-1Score	type3-2Score
1	🐼馒头科技	Manteia/京发科技/龙程电子	88.58	92.70	89.57	80.09	89.20
2	🐼安硕硕眼探企	上海安硕信息技术股份有限公司	88.52	93.19	92.02	74.14	89.31
3	🐼流宝真人	杭州大道一以科技有限公司	85.93	87.36	89.03	76.24	88.68
4	🐼chatglm反卷总局	腾讯	85.84	86.79	89.45	75.64	88.99
5	🐼南哪都队	南京大学	84.29	88.35	85.98	72.61	88.68
6	🐼吃辣子	饺子研究院	82.20	77.09	87.82	74.65	90.13
7	🐼IntelliSeekers		81.76	82.17	87.43	65.90	89.62
8	🐼龙盈战队	龙盈智达（北京）科技有限公司	81.71	83.37	81.54	75.94	88.94

Type1Score	Type2Score	Type3-1Score
86.36	88.67	78.23

上图为私有榜分数各个Type的最高分，从分数可以看出该方案在Type3开放域问题上分数为前八名**最高分**。

技术方案

➤ 创新性

01. 创新提出3P框架

将判别式AI与生成式AI相融合，通过将判别式AI应用在大模型推理的**数据预处理**（Pre-processing）、**生成**（Processing）、**和后处理**（Post-processing），保证了大模型的输入流和输出流的稳定性和可靠性。

02. 发挥判别式AI辅助作用

Pre-processing阶段引入了问题分类的判别式AI，针对不同类问题采用不同优化路线。Processing阶段，为了提升开放域问题对话质量，引入了用于检索的判别式AI与大模型融合，我们采用**对话关键词抽取+建立检索文本二级索引**的方式构建向量库。对话关键词抽取去除了用户问题的噪声，保证了检索准确率。本方案也将年报切割，并建立了二级索引，一级索引为章、节小标题，二级索引为原始文本，并将二者一一映射。

03. 设计Few-shot机制

使用Few-shot 方法给与大模型一些提示，并对Prompt做了一些优化，使大模型理解了文本回答的格式，并在一定程度上减少了幻觉问题

数据预处理Pre-processing

✓ 数据探查—判别式AI对问题进行分类

鉴于不同问题的答案出现在年报中不同的位置，因此在Pre-processing阶段对问题进行精准分类将有助于捕捉和问题相匹配的文本片段，给予大模型更好的提示。

● 基于预训练模型的判别式AI

考虑到用户问题的口语化、多样化，使用Bert等预训练模型训练一个判别式AI，找到一个边界或决策面，以将不同类别的数据点正确地分开是最初考虑方案。在训练数据较少的情况下，原则上可以通过回译、关键词替换等方法来进行**数据增强**，并且通过GPT4扩充样本。由于比赛测试数据的Question本身的特征并非很难区分，且由于时间限制，**本方案最终采用关键词识别的方法对问题分类**。

● 关键词识别

通过关键词识别，总结出下述七类问题。

1. “2021年哪七个上市公司的资产总金额最高？”--**统计类问题**
2. “提供鞍*七*化*股份有限公司2019年其他非流动资产，单位为元。”--**财务报表类基本数据查询问题**
3. “请提供拉*拉*付股份有限公司2020年的无形资产增长率并保留2位小数”--**财务报表类计算问题**
4. “2021年天*渤*化*股份有限公司法定代表人与2020年相比是否都是相同的？”--**公司基本信息类问题**
5. “唐*港*团股份有限公司2019年的销售人员人数为？”--**公司员工信息类问题**
6. “概述一下关键审计事项的情况，针对珠*光*科*股份有限公司2019年的年报。”--**针对某公司的开放域问题**
7. “长期股权投资的投资回报和价值变动如何影响公司的资本利用和盈利能力？”--**完全开放域问题**

数据预处理Pre-processing

✓ PDF文本提取及数据库报表字段构建

- 根据PDF文件的目录和章节信息，通过正则去除噪音，并对内容进行精确的切块处理。
- 定义并构建了资产负债表、现金流量表、利润表、公司信息表、研发人员表等多个金融数据库报表。

合并资产负债表

['项目', 2019, 2018]
['货币资金', '601,869,143.23', '500,593,539.53']
['应收账款', '3,035,935.04', '4,300,249.47']
['预付款项', '72,535,574.60', '66,193,514.95']
['其他应收款', '293,543,891.28', '154,152,832.20']
['应收利息', '178,314.19', '']
['存货', '1,473,938,091.38', '1,306,508,492.06']
['其他流动资产', '113,141,532.84', '140,469,917.13']
['流动资产合计', '2,558,064,168.37', '2,172,218,545.34']
['可供出售金融资产', '', '122,819,734.46']
['其他权益工具投资', '121,054,358.40', '']
['投资性房地产', '138,880,251.56', '312,293,015.77']
['固定资产', '1,272,996,256.10', '1,382,129,786.23']
['在建工程', '47,648,213.79', '65,808,392.88']
['无形资产', '91,693,766.23', '91,229,209.05']
['长期待摊费用', '220,534,122.33', '233,706,714.40']

合并利润表

['项目', 2019, 2018]
['营业总收入', '3,383,409,896.94', '3,468,780,535.04']
['营业收入', '3,383,409,896.94', '3,468,780,535.04']
['营业总成本', '3,289,153,247.18', '3,473,923,150.39']
['营业成本', '2,624,476,902.41', '2,758,563,285.71']
['税金及附加', '20,447,732.43', '24,342,548.57']
['销售费用', '197,326,184.46', '200,753,245.52']
['管理费用', '138,255,984.85', '142,535,828.85']
['研发费用', '313,000,219.30', '362,544,741.21']
['财务费用', '-4,353,776.27', '-14,816,499.47']
['利息费用', '602,519.00', '144,275.00']
['利息收入', '4,622,929.20', '3,766,056.63']
['其他收益', '35,942,024.47', '28,268,880.40']
['投资收益', '22,864,140.70', '31,825,734.72']

公司基本情况表

['公司注册地址', '余姚市姚北工业新区']
['公司注册地址的历史变更情况', '无']
['公司办公地址', '浙江省余姚市朗霞街道朗马路188号']
['公司办公地址的邮政编码', '315480']
['公司网址', 'www.langdi.com']
['电子信箱', 'ldzd@langdi.com']

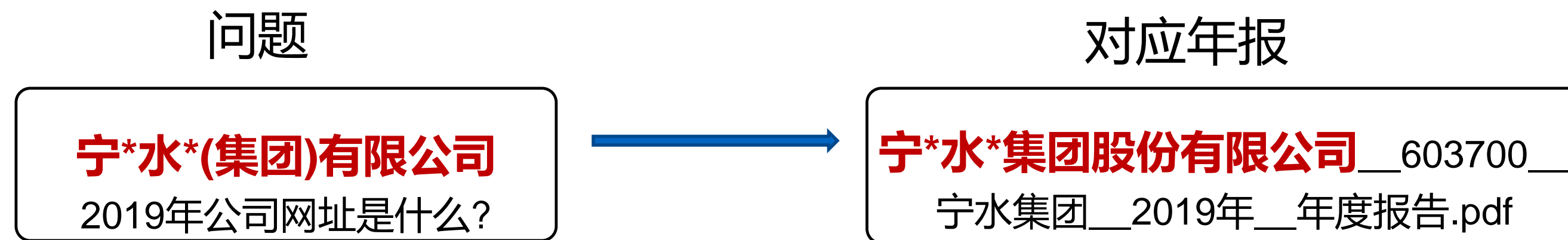
员工情况表

['在职员工的数量合计', '2708']
['生产', '1695']
['销售', '198']
['技术', '463']
['财务', '115']
['行政', '206']
['其他', '31']
['职工总数', '2708']
['专科及以上', '280']
['高中及以下', '2428']

数据预处理Pre-processing

✓ 基于文本相似度搜索召回的判别式AI

问题当中存在某些特殊的公司名称，其通过关键词无法定位具体年报（年报文件定位失败），这导致模型回答错误是必然的。因而可通过相似性搜索的方式将问题与对应年报建立映射。



使用m3e-base预训练模型构建包含所有公司名称和年份的语义向量库，并通过相似性计算搜索到和问题相关性最高的年报文件。

生成Processing

✓ 构建语义向量数据库

开放性问题中表格数据和非表格数据均有涉及，如果提供给大模型的信息太过于冗余庞杂，则会掺杂太多噪音，影响答案的质量。我们发现测试问题当中的**关键词和年报小标题**具有高度相关性。如下图示例。

(8) 主要销售客户和主要供应商情况

公司主要销售客户情况

前五名客户合计销售金额（元）	196,097,521.79
前五名客户合计销售金额占年度销售总额比例	48.96%
前五名客户销售额中关联方销售额占年度销售总额比例	0.00%

公司前 5 大客户资料

序号	客户名称	销售额（元）	占年度销售总额比例
1	客户 1	71,327,036.73	17.81%
2	客户 2	63,017,021.94	15.73%
3	客户 3	32,655,866.88	8.15%
4	客户 4	15,158,832.83	3.78%
5	客户 5	13,938,763.41	3.48%
合计	--	196,097,521.79	48.96%

主要客户其他情况说明

√ 适用 □ 不适用

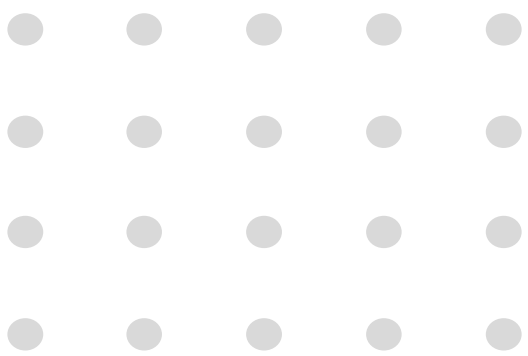
公司前五大客户与公司不存在关联关系，公司董事、监事、高级管理人员、核心技术人员、持股5%以上股东、实际控制人和其他关联方在主要客户中未直接或间接拥有权益。

问题示例

- “Question”: "根据四*中*防*科*股份有限公司2019年的年报数据，能否简要介绍公司报告期内**主要供应商的详情**。"
- 问题中的关键信息是 **主要供应商的详情**，与之对应的是小标题**【主要销售客户和主要供应商情况】**，因此我们可构建包含小标题的语义向量用于文本检索任务。

方案

因此将年报的章节和子目录精确切割。并使用**m3e-base**预训练模型和相关**Text2Vec**技术，从文本中提取出包含标题信息的语义向量，并筛选和问题最相关的**Top 4**，送入大模型。



生成Processing

✓ Prompt的分类

- 1 Zero-shot: 直接用大模型做任务而不给其参考示例, 需要模型具有很强的泛化能力。
- 2 **Few-shot**: 与zero-shot相对, 在与大模型交互时, 在prompt中给出少量示例。
- 3 Role prompt: 让大模型想象自己是某方面专家、从而获得更好的任务效果。
- 4 Chain-of-thought: 把一个多步推理问题分解出多个中间步骤, 并且让 LLM 更加可解释。
- 5 Multimodal prompt: 多模态prompt。输入包含了众多模态的信息如同时输入文本和图像与多模态大模型进行交互。
- 6 **Prompt 微调**: 包括AutoPrompt (通过梯度优化, 自动从一系列候选词中生成对目标任务最佳prompt) 和 Prefix-tuning、P-tuning (其在Transformer的输入层中加入可学习的prompt embedding, 进行微调) 等方法。

生成Processing

✓ Few-shot的使用

由于Few shot方法存在**多标签偏见**的问题，即Few-shot例子中所提供的标签如果分布不平衡，则会严重影响到测试任务。**因此可对于一阶段得到的不同分类问题设计不同的Prompt，保证Few-shot Example与Test样本相似性距离更近。并通过调整Temperature变量为0.1，减少模型输出结果的随机性。**

例如：对于法定代表人是否相同问题的判定。

上下文内容中给出了从2019-2021年公司的法定代表人名字，请根据上下文内容回答问题，请严格按照上下文内容来回答问题，不要胡编乱造，下文已经给出几个回答示例。

【回答示例1】：
上下文内容：
2019年法定代表人名字是：王颖。
2020年法定代表人名字是：王刚。
2021年法定代表人名字是：王颖。
问题：浦*金*2020-2021年的法定代表人是否都相同？
答案：不相同。浦*金*2020年法定代表人是王刚。
2021年法定代表人是王颖。

【回答示例2】：
上下文内容：
2019年法定代表人名字是：李刚。
2020年法定代表人名字是：李刚。
2021年法定代表人名字是：李刚。
问题：工*银*2019-2021年的法定代表人是否都相同？
答案：相同。工*银*2019-2021年法定代表人都是李刚。

===
上下文内容：2019年法定代表人名字是：张**。
2020年法定代表人名字是：张**。
2021年法定代表人名字是：杨**。
问题：2021年晋*车*股份有限公司法定代表人与2020年相比是否都是相同的？
答案：

加入Few-shot 模型结果

不相同。晋*车*股份有限公司2021年法定代表人是杨**，2020年法定代表人是张**。



无 Few-shot 模型结果

不相同。晋*车*股份有限公司2021年和2020年的法定代表人都是张**。



如果所示，在回答多年来法定代表人是否相同的问题时，须遵循

- 【1.提供**相同/不相同**字眼】
 - 【2.提供法定代表人**具体名字**】
- 的格式来作答。

我们使用few shot提供两种不同的示例，让模型减少了一些幻觉问题，并严格根据规范作答。

但是大模型在某些问题上还是存在部分幻觉问题，我们在这些问题上采用后处理规则的方式做了小幅度的修正。

后处理Post-processing

✓ 生成文本监测与过滤

上下文

上下文内容：
广*能*年报数据：
证券简称为广汇能源。
证券代码为600256。



Question

请告诉我广*能*股份有限公司2020年的证券代码的具体情况



模型预测Answer ✗

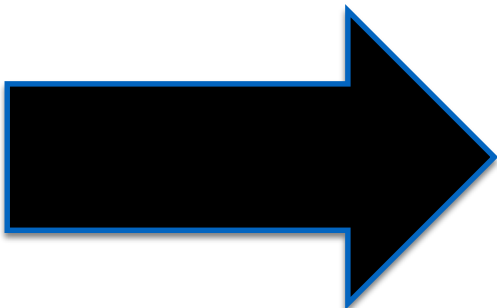
抱歉，我无法回答这个问题，因为我的知识截止于2023年，而且我无法从上下文中获取关于2020年广*能*股份有限公司的证券代码信息。

正确答案Answer ✓

广*能*股份有限公司2020年的证券代码为600256。

如上图所示，基于上下文中给出正确答案，模型也会出现无法回答的情况。实际业务中可能会涉及敏感信息过滤等复杂问题。

未来

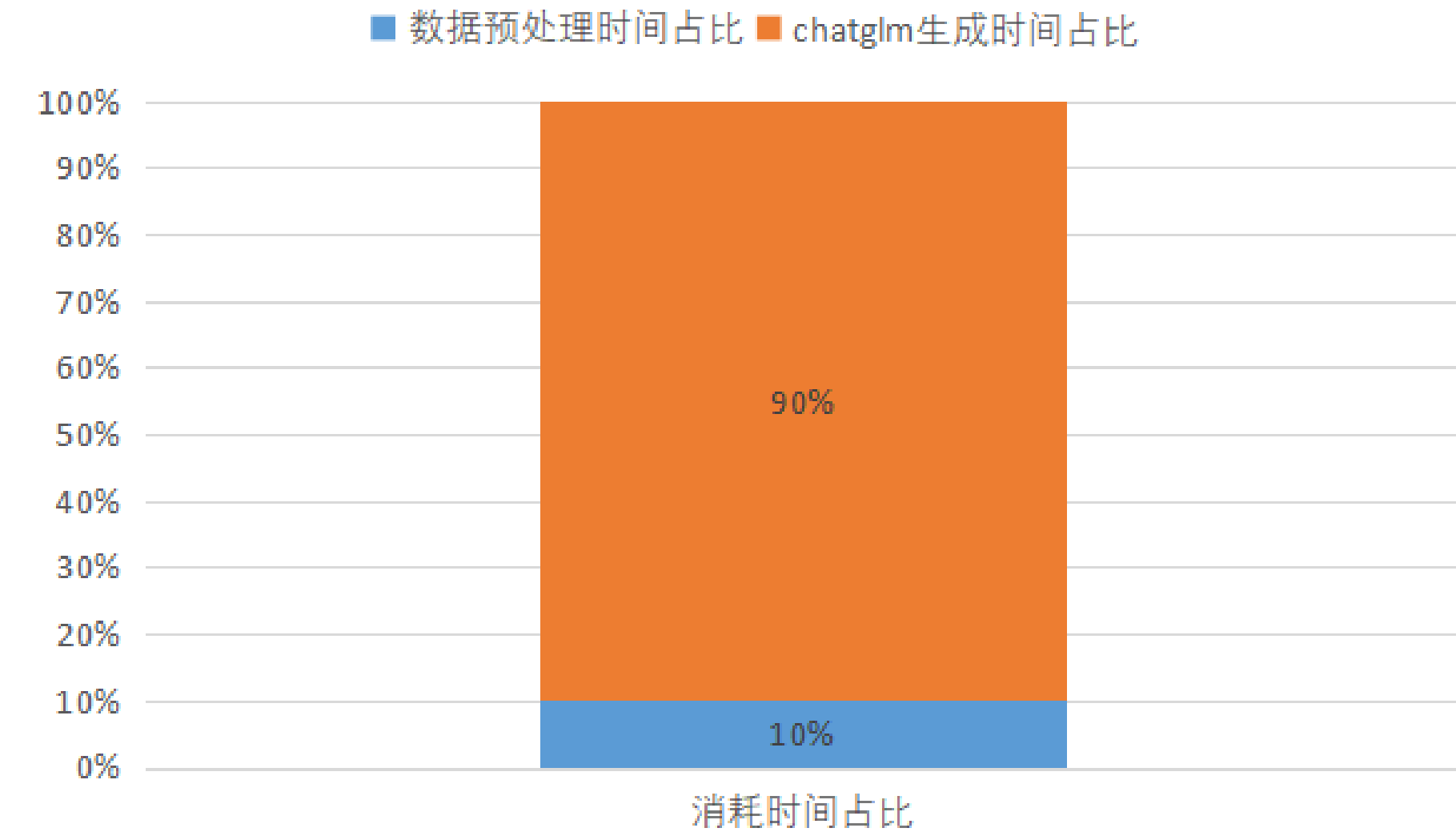
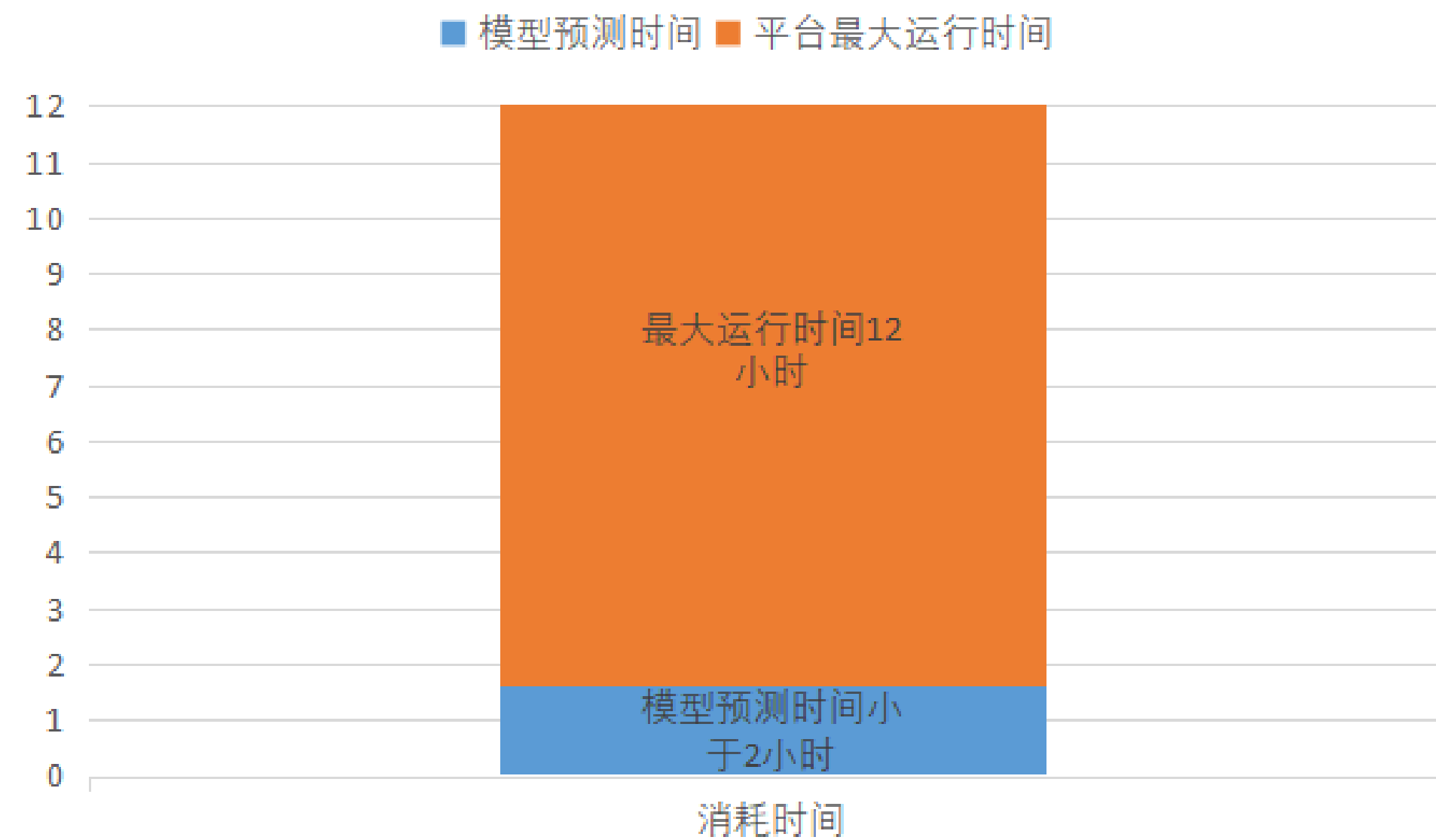


基于判别式规则：通过规则将此类问题筛选出来，然后使用模型二次生成、后处理修正等手段解决此类问题。

基于判别式AI：通过小模型对生成文本进行监测过滤，找到不符合业务逻辑的信息。

技术方案

➤ 复杂度



- 本方案模型对于2000条测试问题的预测时间为**1小时40分钟左右**，平均每条样本预测时间为2-3秒，包括数据处理和大模型生成，**远远小于天池平台提供的最大运行时间12小时**，即仅占运行时限的14%，**复杂度低**。
- 本方案仅仅使用了官方提供的由PDF转化的txt文本做数据预处理，**数据预处理耗时仅10分钟左右**，占整个预测时常的**10%**，**复杂度低**。

技术方案

➤ 可靠性

Type1Score	Type2Score	Type3-1Score	Type3-2Score	Final_Score
85.74	88.33	78.19	89.22	85.62
86.36	88.67	77.61	89.5	85.85
84.69	88.34	78.23	89.41	85.33

- 上图为本方案模型的三次提交结果，可以看出模型在各个Type分数和综合分数上均没有太大波动，**分数稳定、可靠。**
- 鉴于大模型结果的随机性，本方案为了保证模型输出结果的可靠，一方面降低模型温度（Temperature）**减少随机性**，另一方面对于质量较为不可靠的生成结果，**进行多次预测**，并加入了**Post-processing控制模型输出。**

技术方案

➤ 泛化性

- Pre-processing阶段使用判别式AI对用户问题进行了精准有效的分类。
- Processing阶段本方案对于不同类别问题设计不同的处理策略。
- Post-processing阶段使用判别式规则将低质量答案筛选出来，并进行了二次处理，从而进一步提升模型的泛化能力。

分数	名次
B榜81.76	7
C榜85.85	5

在分数和名次上，C榜相比B榜均有提升，表明本方案模型具有较好的泛化能力

目录 CONTENTS

1

团队介绍

2

赛题分析

3

技术方案

4

业务成效

5

应用价值

业务成效

➤ 精准情况

年报对应片段

十七、社会责任情况

1、履行社会责任情况

具体内容详见公司同日刊载于中国证监会指定创业板信息披露网站巨潮资讯网（www.cninfo.com.cn）上的《2019年度社会责任报告》。

2、履行精准扶贫社会责任情况

公司报告年度暂未开展精准扶贫工作，也暂无后续精准扶贫计划。

3、环境保护相关的情况

上市公司及其子公司是否属于环境保护部门公布的重点排污单位是

公司或子公司名称	主要污染物及特征污染物的名称	排放方式	排放口数量	排放口分布情况	排放浓度	执行的污染物排放标准	排放总量	核定的排放总量	超标排放情况
深圳迈瑞生物医疗电子股份有限公司	-	-	-	-	-	-	-	-	-
深圳迈瑞生物医疗电子股份有限公司	化学需氧量 五日生化需氧量 氨氮 总磷 悬浮物	纳管排放	1	园区东侧工业废水总排口	52mg/L 11.4mg/L 1.45mg/L 0.42mg/L 94mg/L	300mg/L 150mg/L 40mg/L 4.5mg/L 200mg/L	工业废水排放总量：48,230.96	工业废水排放核定总量：	未超标

模型生成结果

2019年,迈瑞医疗在社会责任方面表现良好。公司发布了《2019年度社会责任报告》,详细介绍了公司在社会责任方面的实践和成果。在环境和社会责任方面,公司取得了一定的进展。

在**环境保护方面**,迈瑞医疗采取了多种措施来保护环境。公司对生产废水进行了合规处置,将实验室产生的废液收集并交给有资质的危废厂商进行处理。公司还建有危险废物间,定期将危险废物交给有资质的危废厂商进行处置。此外,公司建立了废水在线监测设备,确保工业废水排放浓度远低于排放限值。

在环境保护行政许可方面,迈瑞医疗也在不断完善。公司建设项目均已完成竣工环境保护验收,并取得审批,最新批复为“深南环水评许[2016]282号”。

在**精准扶贫方面**,迈瑞医疗暂未开展精准扶贫工作,也没有后续精准扶贫计划。

综上所述,迈瑞医疗在社会责任方面表现良好,为推进可持续发展做出了积极的贡献。

针对开放域问题中涉及的多类型细节信息，本方案模型可以综合全面且精准地予以回答。

例如：某公司社会责任情况涉及“环境保护”和“精准扶贫”两方面内容，本方案模型可以全面综合提取两个方面的信息，并予以回答，精准性很高。

业务成效

理解能力

合并资产负债表

['项目', 2019, 2018]
['货币资金', '601,869,143.23', '500,593,539.53']
['应收账款', '3,035,935.04', '4,300,249.47']
['预付款项', '72,535,574.60', '66,193,514.95']
['其他应收款', '293,543,891.28', '154,152,832.20']
['应收利息', '178,314.19', '']
['存货', '1,473,938,091.38', '1,306,508,492.06']
['其他流动资产', '113,141,532.84', '140,469,917.13']
['流动资产合计', '2,558,064,168.37', '2,172,218,545.3']
['可供出售金融资产', '', '122,819,734.46']
['其他权益工具投资', '121,054,358.40', '']
['投资性房地产', '138,880,251.56', '312,293,015.77']
['固定资产', '1,272,996,256.10', '1,382,129,786.23']

将表格数据转化为文本数据：
2019年固定资1272996256.10元，2018年其他流动资产为1382129786.23元。

合并利润表

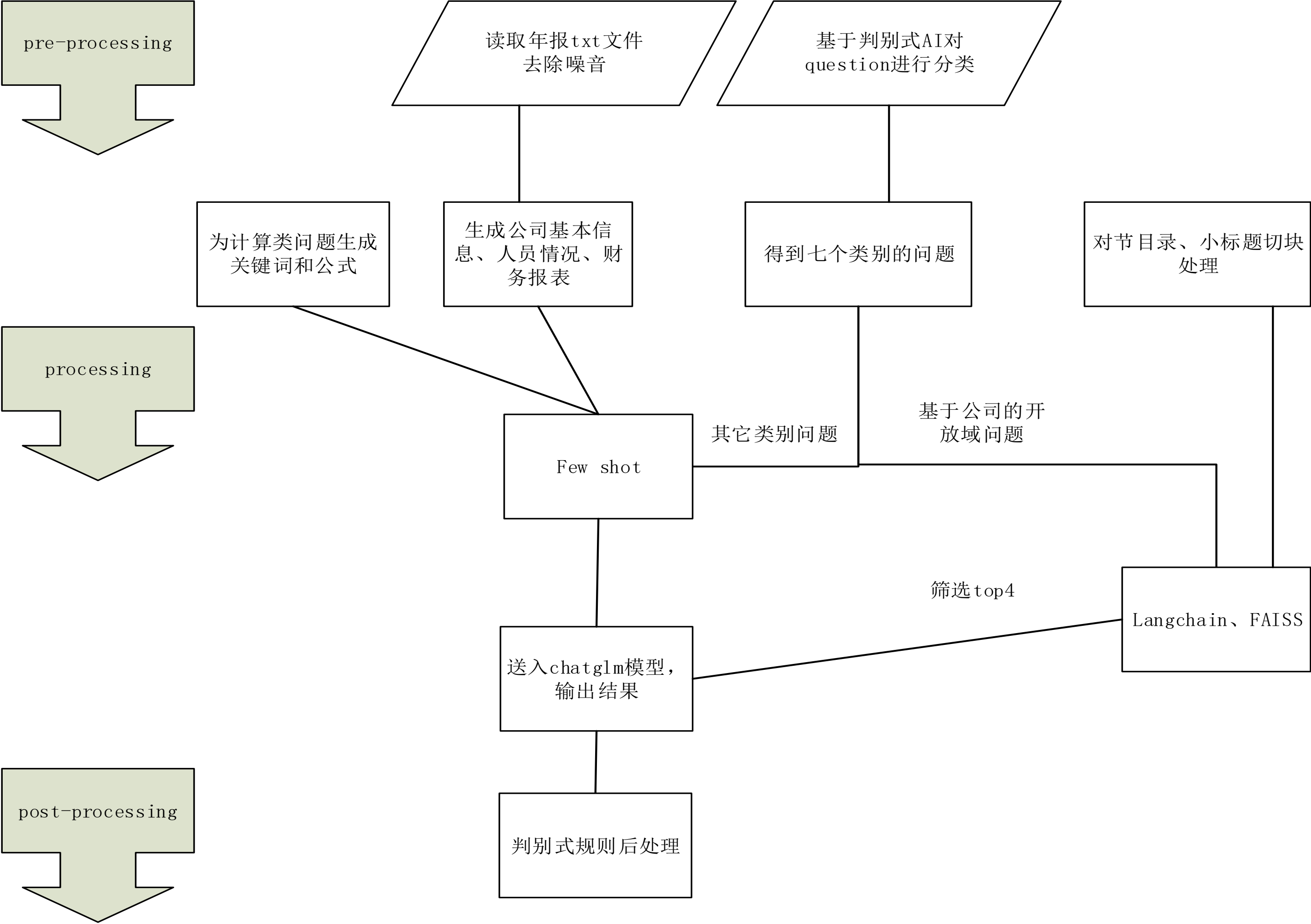
['项目', 2019, 2018]
['营业总收入', '3,383,409,896.94', '3,468,780,535.04']
['营业收入', '3,383,409,896.94', '3,468,780,535.04']
['营业总成本', '3,289,153,247.18', '3,473,923,150.39']
['营业成本', '2,624,476,902.41', '2,758,563,285.71']
['税金及附加', '20,447,732.43', '24,342,548.57']
['销售费用', '197,326,184.46', '200,753,245.52']
['管理费用', '138,255,984.85', '142,535,828.85']
['研发费用', '313,000,219.30', '362,544,741.21']
['财务费用', '-4,353,776.27', '-14,816,499.47']
['利息费用', '602,519.00', '144,275.00']
['利息收入', '4,622,929.20', '3,766,056.63']
['其他收益', '35,942,024.47', '28,268,880.40']
['投资收益', '22,864,140.70', '31,825,734.72']

将表格数据转化为文本数据：
2019年利息收入4622929.20元，2018年其他流动资产为3766056.63元。

前期测试发现6b模型对于表格的提取易出现混乱，对于表格的理解能力较差。本方案将表格数据转化为大模型更容易理解的文本描述，保证统计数据抽取的正确性。

业务成效

➤ 流程情况



- 设计了从Pre-processing到Post-processing多模块处理 workflow。
- 从输入端到输出端设计不同的判别式AI模型把控生成结果的质量。
- 多模块设计保证了算法的灵活性，每个模块可以单独维护。

业务成效

➤ 专业度

- 团队**熟悉银行业务场景**，长期深耕大模型在金融科技领域的应用探索，并在**投研投顾**、**金融制度**问答等领域均形成了一系列成果。
- 团队多次斩获国际相关领域奖项。例如：在国际会议（**NAACL 2022**） Workshop SUKI Task 研讨会发布的**FinQA(金融表格数据QA问答竞赛)**获得**全球三等奖**。

Abstract

This research mainly focuses on the financial numerical reasoning dataset - FinQA. In the shared task, the objective is to generate the reasoning program and the final answer according to the given financial report containing text and tables. We use the method based on **DeBERTa pre-trained language model**, with additional optimization methods including **multi-model fusion, training set combination on this basis**. We finally obtain an execution accuracy of **68.99%** and a program accuracy of **64.53%**.

Example

In the task, you are given a financial report containing both text and table, the goal is to answer a question requiring numerical reasoning. An example is shown below:

Page 31 from the annual reports of GERM (German Ltd.)

The fair value for these options was estimated at the date of grant using a Black-Scholes option pricing model with the following weighted average assumptions for 2006, 2005 and 2004:

	2006	2005	2004
Weighted average fair value of options granted	\$20.01	\$9.48	\$7.26
Expected volatility	0.3534	0.3224	0.3577
Dividend yield	1.95%	0.98%	1.30%
Expected life of options in years	6.3	6.3	6.3
Risk free interest rate	5%	4%	4%

The total fair value of shares vested during 2006, 2005, and 2004 was \$8,413, \$6,249, and \$6,418 respectively. The aggregate intrinsic value of options outstanding and exercisable at December 31, 2006 was \$204.1 million and \$100.2 million, respectively. (... additional 10 sentences ...)

Question: Considering the weighted average fair value of options, what was the change of shares vested from 2005 to 2006?

Answer: - 400

Calculations:

$$\left(\frac{9413}{20.01} \right) - \left(\frac{6249}{9.48} \right) = -400$$

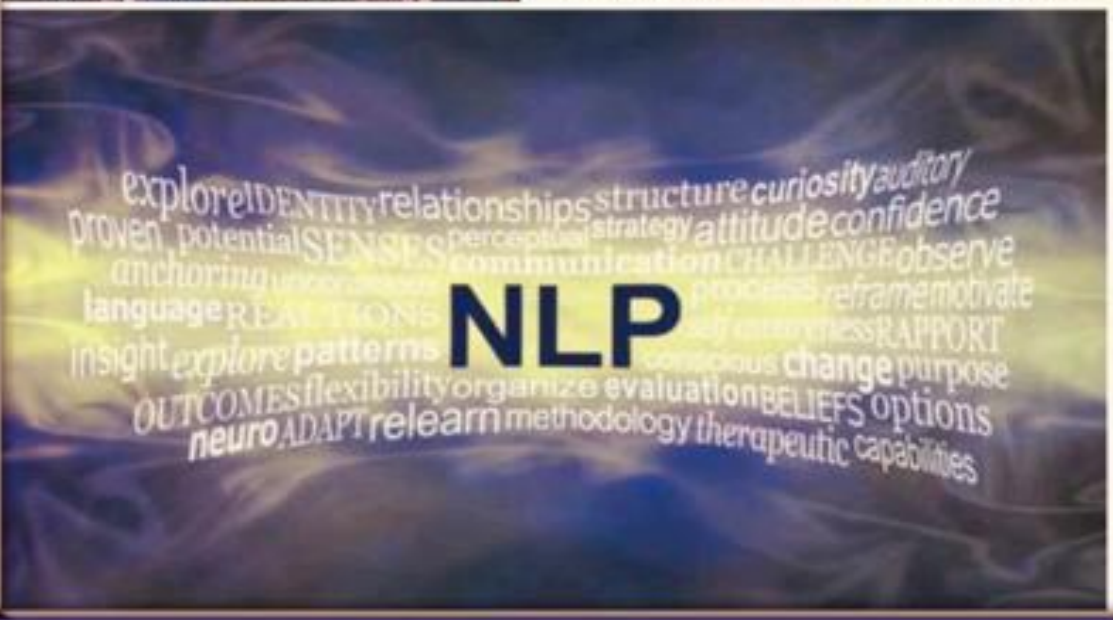
Program:

$$\text{divide} (9413, 20.01) \quad \text{divide} (6249, 9.48)$$
$$\text{subtract} (90, 91)$$

2022 Annual Conference of the North American Chapter of the Association for Computational Linguistics

SUKI: Structured and Unstructured Knowledge Integration

Workshop at NAACL 2022 at Seattle, WA



Program Schedule

Workshop Location: 800 - Cowley Index
Poster Session Location: Registry A & B and Galileo Room

Thursday, July 14, 2022 Pacific Daylight Time

- 08:00-09:00 - Opening Remarks
- 09:00-09:45 - Invited Talk: Hongdi Sun, Google
- 09:45-10:30 - Invited Talk: Perry Liang, Stanford
- 10:30-11:15 - Invited Talk: Jonathan Berant, The Allen Institute for AI
- 11:15-12:00 - Invited Talk: Hanna Hajjari, Google
- 12:00-12:30 - Lunch Break
- 12:30-01:00 - Poster Session Location: Registry A & B and Galileo Room
- 01:00-02:15 - Invited Talk: William Cohen, Google
- 02:15-03:00 - Invited Talk: Julian Jern, Google
- 03:00-04:00 - Shared Task Session
 - Chairman: Team ARI (AI Risk)
 - Team ARI
 - Runner-up: Team (MASC) from Reading University
 - Honorable Mention: Team TATTS from Google Research
 - Honorable Mention: Team U2D2-Prithvi
- 04:00-04:45 - Invited Talk: Lina Ding, Google
- 04:45-05:30 - Closing Remarks

- 团队具备扎实的**Prompt工程**经验，同时对于6b模型生成的异常结果已开展一系列探索。

业务成效

➤ 可解释性

请根据上下文内容回答问题，请仔细阅读上下文，正确答案均在上下文内容中。下文已经给出回答示例，请按照示例的风格来作答。

【回答示例】：
上下文内容：
工商银行年报数据：在职员工的数量合计人员数量为13989。生产人员数量为10600。销售人员数量为35。技术人员数量为1363。
问题：2021年中国工商银行销售人员数量是什么？
答案：2021年中国工商银行销售人员数量是35人。

=====
上下文内容：
亿*锂*年报数据：在职员工的数量合计（人）为9669。
当期领取薪酬员工总人数（人）为9669。
生产人员数量为6390。
销售人员数量为370。
技术人员数量为1751。
财务人员数量为86。
问题：请提供2020年惠州亿*锂*股份有限公司职工总人数的详细数据。
答案：

基于公司员工信息的示例

上下文内容中给出了从2019-2021年公司的法定代表人名字，请根据上下文内容回答问题，请严格按照上下文内容来回答问题，不要胡编乱造，下文已经给出几个回答示例。

【回答示例1】：
上下文内容：
2019年法定代表人名字是：王颖。
2020年法定代表人名字是：王刚。
2021年法定代表人名字是：王颖。
问题：浦*金*2020-2021年的法定代表人是否都相同？
答案：不相同。浦*金*2020年法定代表人是王刚。
2021年法定代表人是王颖。

【回答示例2】：
上下文内容：
2019年法定代表人名字是：李刚。
2020年法定代表人名字是：李刚。
2021年法定代表人名字是：李刚。
问题：工*银*2019-2021年的法定代表人是否都相同？
答案：相同。工*银*2019-2021年法定代表人都是李刚。

===
上下文内容：2019年法定代表人名字是：张**。
2020年法定代表人名字是：张**。
2021年法定代表人名字是：杨**。
问题：2021年晋*车*股份有限公司法定代表人与2020年相比是否都是相同的？
答案：

基于公司法定代表人变更情况的示例

- 用户可以**通过调整输入样式来引导模型的行为，从而提供一定程度的可解释性。**
- 对于**不同种类的问题设计了不同的回答示例，使得大模型在面对不同问题时能够有不同的回答风格，保证了模型每次输出格式与所给示例的协调一致性。**

目录 CONTENTS

1

团队介绍

2

赛题分析

3

技术方案

4

业务成效

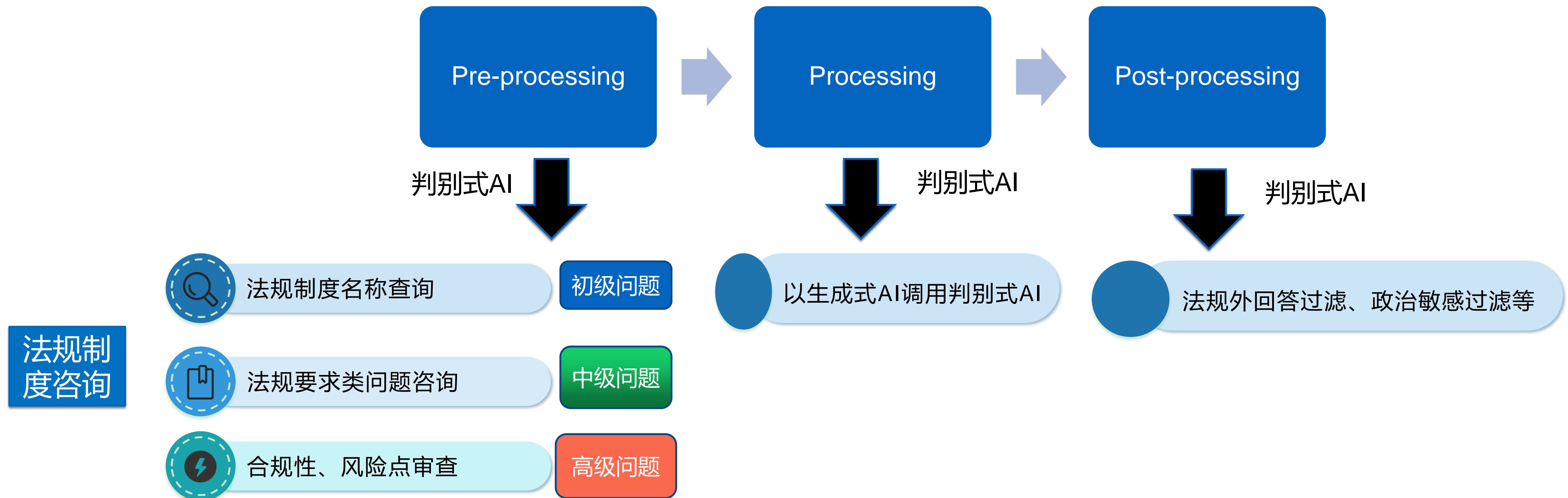
5

应用价值

应用价值

➤ 构建标准体系

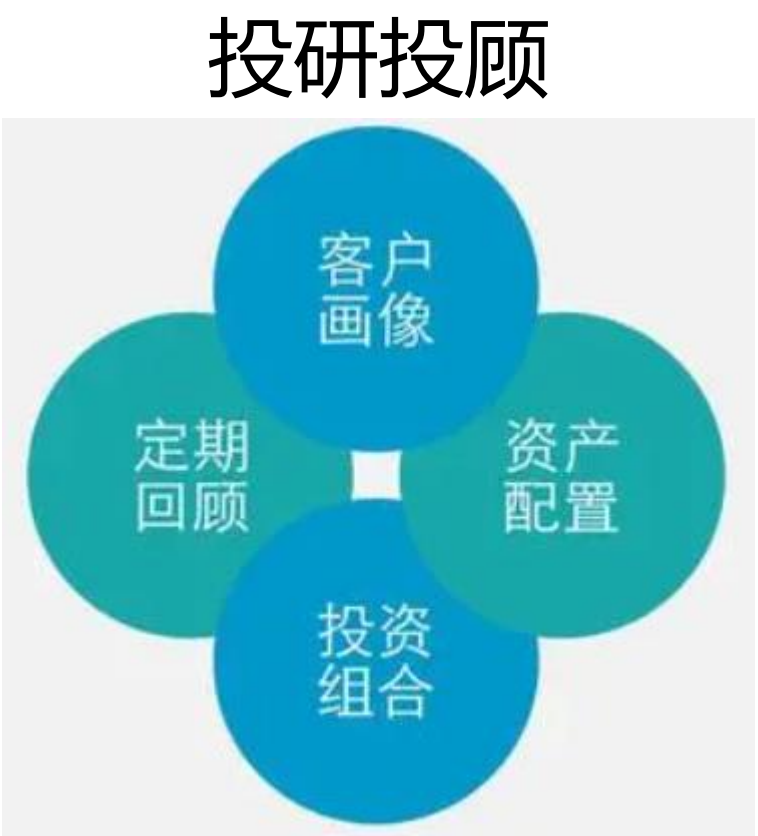
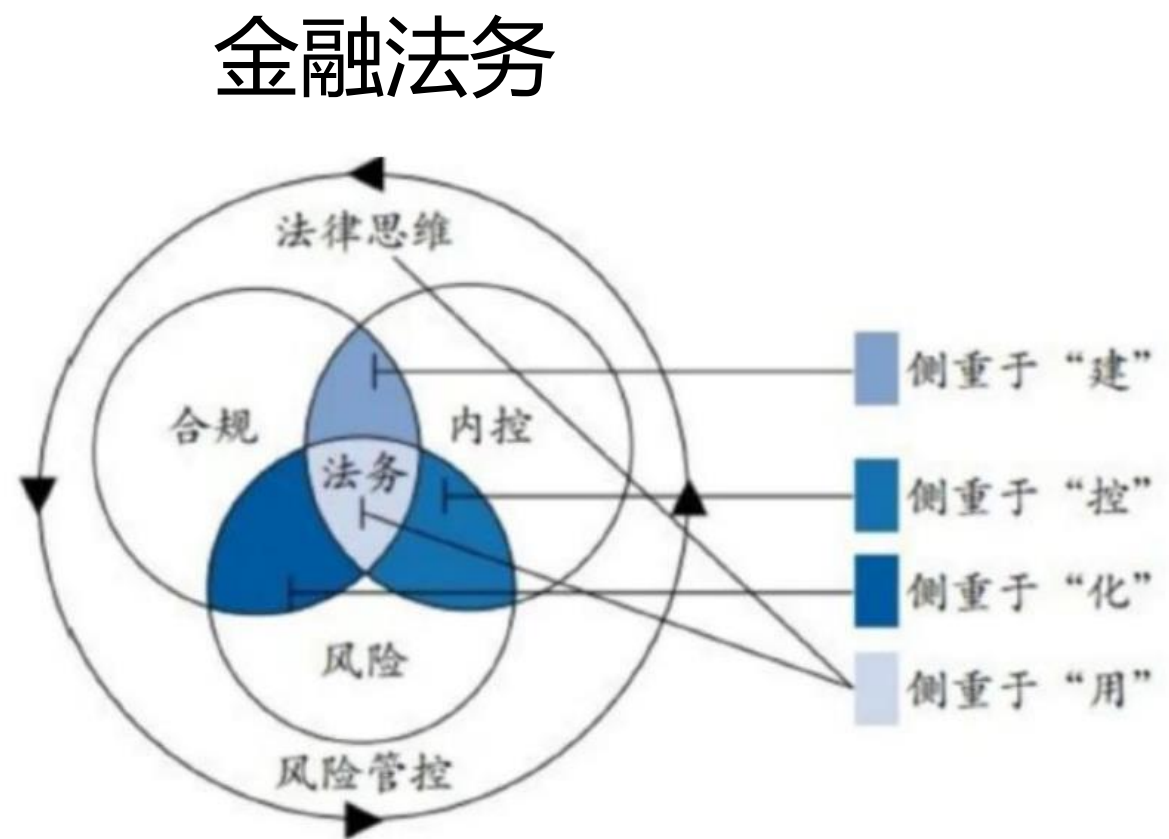
- 在问答系统中，用户问题是多样化的，大模型在处理不同种类问题上有不同的准确率，并且模型的输出质量从输入端到输出端受多种因素影响。因此构建Pre-processing、Processing到Post-processing模块化管理标准体系很有必要。



应用价值

➤ 应用拓展

可将模块化处理方式**拓展到金融行业其它细分领域**，例如面向金融客户的**投研投顾服务**、面向金融客户的**财富管理服务**、面向金融从业者的**会议纪要**以及**金融法务服务**等。



会议纪要

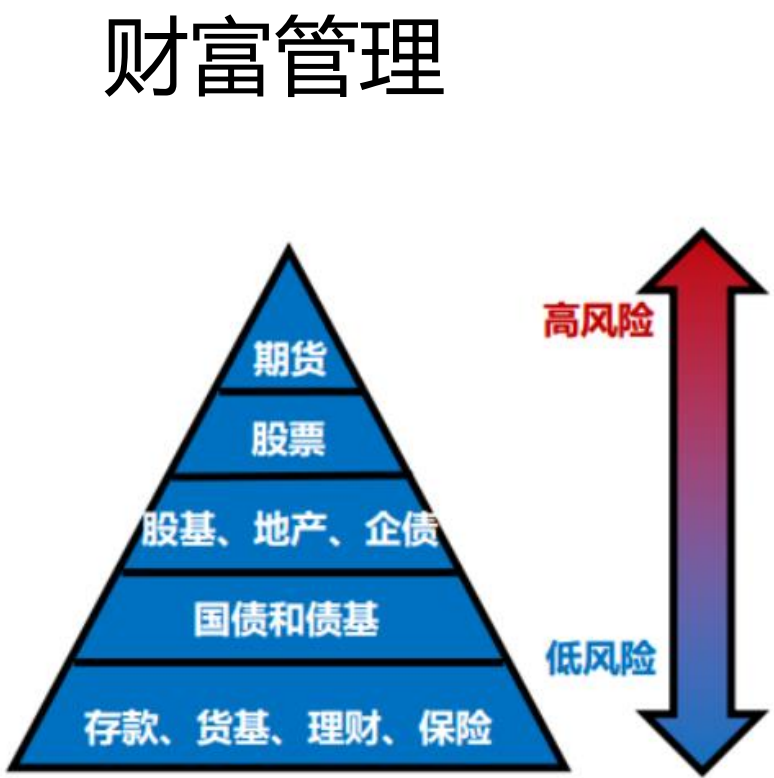
会议纪要
MEETING MINUTES

➤ 工程部、营销部、组织部。

🕒 会议主题：	
👤 会议时间：	
📍 会议地点：	
🗣️ 主持人：	
📝 记录人：	
👥 参会人员：	

📖 会议内容：

请在这里写入会议内容文字；请在这里写入会议内容文字；请在这里写入会议内容文字；请在这里写入会议内容文字；请在这里写入会议内容文字；请在这里写入会议内容文字；请在这里写入会议内容文字；请在这里写入会议内容文字；请在这里写入会议内容文字；请在这里写入会议内容文字。





智谱·AI



marsoft

安硕信息



北京交通大学
BEIJING JIAOTONG UNIVERSITY



ModelScope
魔搭社区

阿里云

敬请指正!