



番禺彭于晏

# Contents

## 目录 ▶

01

团队介绍

02

方案综述

03

解决方案

04

方案成效

05

技术创新

# 01 团队介绍

赛题二：金融保险场景下多源文件长上下文一致性校验

队名：番禺彭于晏

## 队名：番禺彭于晏



陈键

IT语音

---

从事数字人一体机RAG、Agent方面的工作。



# 02 方案综述



## 问题理解

### 核心需求

#### 同一 material\_id 内 长文本一致性核验

确保同一标识下不同保险文档  
素材信息保持一致。



#### 动态规则适配

根据动态的校验规则，检测素  
材的合规性。

# 方案综述

01

## 保险文档复杂多样

保险文档包含条款、说明、宣传等多种类型，语义层次丰富，增加了判定的难度。

02

## 冲突类型多元

存在数值、条件、范围等多种冲突类型，需精准识别实质性矛盾。

03

## 成本与精度平衡

要在保证高可解释性和证据回溯的前提下，实现成本与精度的良好平衡。



## 技术路线

### 实现路径



#### 结构化知识抽取

对保险文档进行结构化处理，提取关键要素，为后续判定奠定基础。



#### Agent编排

利用Agent进行工具调用与证据回溯，实现流程的自动化和智能化。



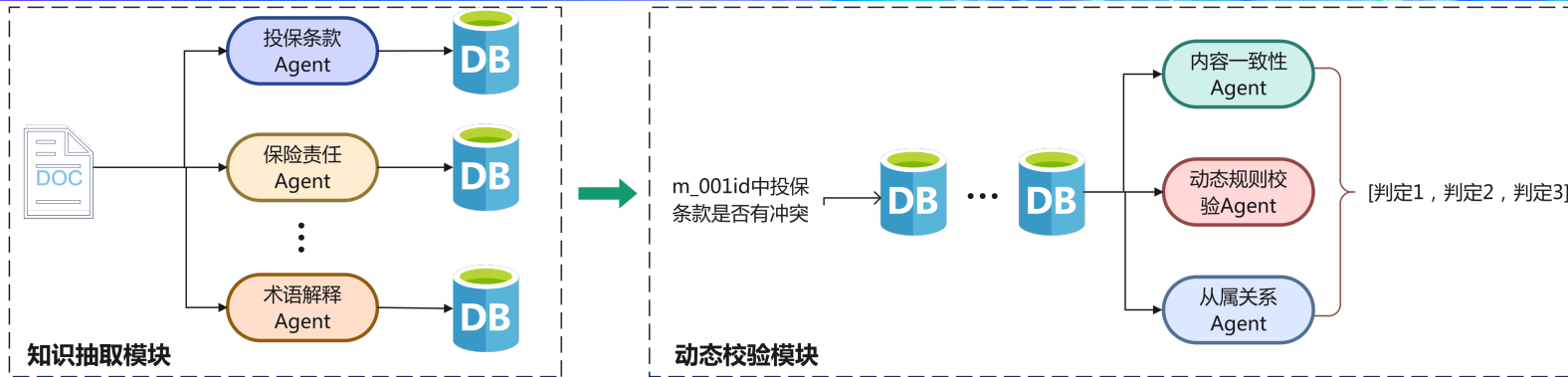
#### RAG双路校验

通过双路校验机制，增强对低置信度结果的验证，提高判定准确性。



# 03 解决方案



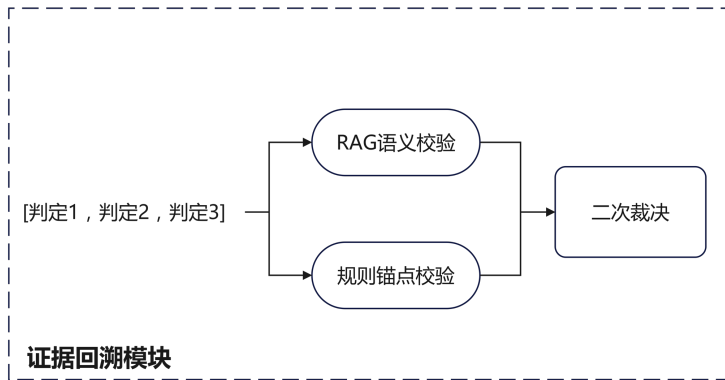


## 知识抽取模块：

11类专用Prompt：保障/时间/赔付/免责/续退保/术语等领域专门化；  
分块与召回：段落/句子/字符多层回退 + 重叠窗口，最大化覆盖处理；  
文档类型感知：营销语境与法律条文区别处理

## 动态校验模块：

策略：低置信分支二审、多视角重判，按 doc\_type 区分路由；会话缓存；  
语义锚点驱动：关键概念词表检索，等待期/免赔额等核心要素精准定位



## 证据回溯模块：

核心思想：对低置信结果进行二次验证，精准减少误判。

双路架构：语义与规则路径并行检索，实现优势互补。

校验流程：双路检索证据，综合比对后进行二次裁决。



# 04 方案成效



## v1 (端到端长文直判)

F1=0.81, 不经过抽取模块, 直接用原始文本进行两两判定。

## v2 (抽取+Agent判定)

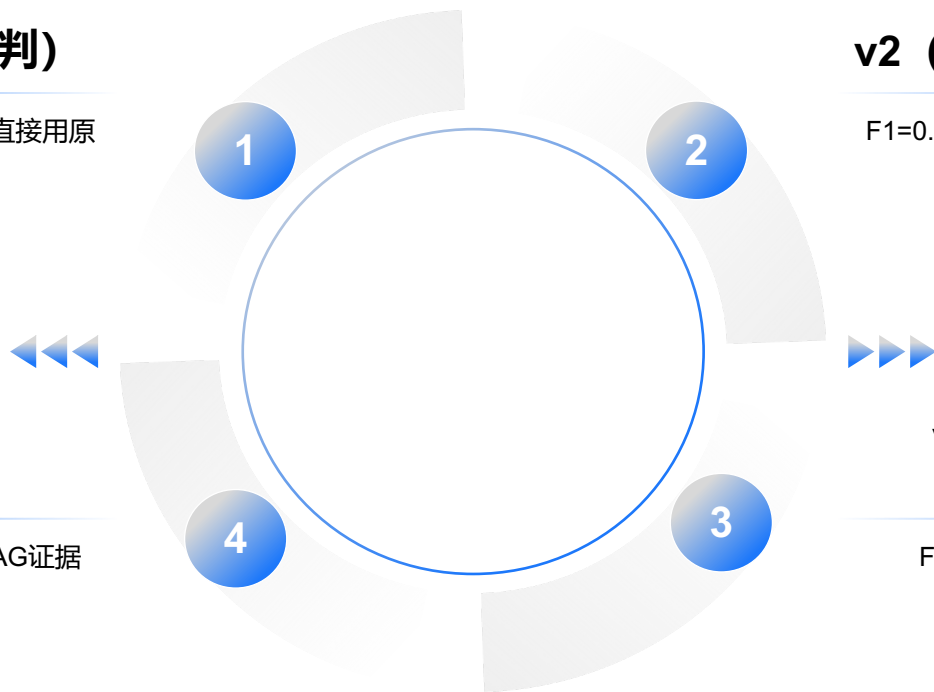
F1=0.822, 一个抽取Agent+一个判定Agent

## v4 (RAG回溯)

F1=0.8434, V3基础上加入RAG证据回溯

## v3 (结构化抽取+多Agent判定)

F1=0.838, 多抽取Agent+多判定Agent



## 失败尝试（复刻cursor）

核心：构建工具。不需要阅读全部文本。根据动态规则阅读对应部分。

线上分数：0.815

好的，让我分析 m\_00003s 材料的责任免除冲突。首先查看这个材料的目录结构：

Listed m\_00003s >

现在查看责任免除目录：

Listed LIABILITY\_EXCLUSION

Read 0.md

现在查看条款目录：

Listed CLAUSE

Grepped 责任免除|免责 [data/样例集/materials/m\_00003s/CLAUSE/0.md]

Grepped 第八条|第九条|第十条 [data/样例集/materials/m\_00003s/CLAUSE/0.md]

让我直接查看条款文件的内容，寻找责任免除部分：

Read 0.md L1-100

让我继续查看条款文件中的责任免除部分：

Read 0.md L100-199

现在我找到了冲突！让我对比一下两个文件中的责任免除条款：

LIABILITY\_EXCLUSION/0.md 中的责任免除条款：

≡ text

三、 被保险人故意自伤、或自本合同成立或者本合同效力恢复之日起3年内自杀，但被保险人自杀时为无民事行为能力人

CLAUSE/0.md 中的责任免除条款：

≡ text

三、 被保险人故意自伤、或自本合同成立或者本合同效力恢复之日起 2 年内自杀，但被保险人自杀时为无民事行为能力人的除外；

### m\_00003s 材料的责任免除冲突分析


冲突的具体表现：

1. 自杀免责期限不一致：

- LIABILITY\_EXCLUSION/0.md：规定为"3年内自杀"
- CLAUSE/0.md：规定为"2年内自杀"

2. 冲突的影响：

- 如果被保险人在保险合同成立或效力恢复后第2-3年期间自杀



# 05 技术创新

一句话总结



## 创新点

### Typed Claim 标准体系

将保险条款原子化，实现跨文档的精准事实对齐。

### Agent 驱动的判定流水线

编排多阶段判定与校验，实现动态、可回溯的智能核验。

### 语义+规则双路校验

结合向量检索与关键词锚点，精准验证实质性矛盾。