



2024 AI+ 研发数字峰会

AI+ Development Digital summit

AI驱动研发迈进数智化时代

中国·上海 05/17-18

大语言模型赋能的证券领域业务规则 规约与测试用例自动生成方法

陈小红 华东师范大学

科技生态圈峰会 + 深度研习



—1000+ 技术团队的选择



K+峰会

上海站

K+ 全球软件研发行业创新峰会

时间: 2024.06.21-22

K+峰会

敦煌站

K+ 思考周®研习社

时间: 2024.10.17-19

K+峰会

香港站

K+ 思考周®研习社

时间: 2024.11.10-12



K+峰会详情



AIDD峰会

上海站

AI+研发数字峰会

时间: 2024.05.17-18

AIDD峰会

北京站

AI+研发数字峰会

时间: 2024.08.16-17

AIDD峰会

深圳站

AI+研发数字峰会

时间: 2024.11.08-09



AIDD峰会详情



陈小红

软件需求工程践行者

华东师范大学软件工程学院，副教授，硕士生导师，主要研究方向为需求工程、形式化方法以及IoT最终用户编程。主持并参与了多项国家自然科学基金项目、重点研发、863、973项目以及省部级项目，主要工作包括需求建模与分析、需求形式化验证、基于投影的需求问题自动分解、基于因果关系的约简与基于用户意图的IoT最终用户编程、智能系统的需求确认等，在国内外知名期刊会议RE、TITS、TCAD、SCP、软件学报等发表论文50余篇，成果应用于轨道交通、航空航天、汽车、证券、智能家居等领域。

目录

CONTENTS

1. 背景
2. 规约与测试用例生成框架
3. 大语言模型微调
4. 生成方法细节
5. 方法评估与实践
6. 总结

PART 01

背景

▶ 背景

- 业务规则在证券领域扮演着关键角色，它们是一系列指导和控制证券市场运作的规定和指南。
- 各大证券交易所都有一套详细的业务规则文档来定义和约束交易。



首页 党建 披露 监管 数据 产品 服务 规则 关于 一网通办 | “三开门”专栏

规则 > 本所业务规则 > 债券 > 债券交易 >

债券交易通用

代码 / 简称 / 关键词



本所业务规则

- 规则总览
- 最新规则
- 章程
- 股票
- 债券
 - > 发行上市审核
 - > 发行承销
 - 上市 (挂牌)
 - 债券交易
 - > 债券交易通用
 - > 国债预发行
 - > 债券质押式三方回购
 - > 债券质押式协议回购
 - > 国债买断式回购交易
 - > 信用保护工具
 - > 上市公司可转债
- 基金
- 基础设施公募REITs
- 期权
- 交易通用

基本业务规则

关于发布施行《上海证券交易所债券交易规则》及有关事项的通知	2022-01-27
关于发布《银行间债券市场与交易所债券市场互联互通业务暂行办法》的通知	2022-01-20

业务细则

关于发布《上海证券交易所 中国证券登记结算有限责任公司债券借贷业务办法》及有关事项的通知	2023-11-24
关于发布《上海证券交易所债券市场投资者适当性管理办法（2023年修订）》的通知	2023-10-20
关于发布《上海证券交易所可转换公司债券交易实施细则》的通知	2022-07-29
关于发布《上海证券交易所 中国证券登记结算有限责任公司境外机构投资者债券交易及登记结算业务实施细则》的通知	2022-06-28

业务指引

关于发布施行《上海证券交易所债券交易规则适用指引第3号——债券做市业务》的通知	2022-01-27
关于发布施行《上海证券交易所债券交易规则适用指引第2号——债券通用质押式回购交易》的通知	2022-01-27
关于发布施行《上海证券交易所债券交易规则适用指引第1号——债券交易参与人管理》的通知	2022-01-27

▶ 背景

- 这些业务规则是证券交易系统需求和约束的来源
- 与软件需求相关的部分将被用作交易系统软件需求规约的一部分，以便进行设计、实现和验收测试
- 从自然语言描述的业务规则文档中发现需求，进一步挖掘这些需求之间的关系以形成软件需求规约，以及据此生成测试用例是交易系统开发中的关键步骤

深圳证券交易所融资融券交易实施细则 (2023年修订)

第一章 总则

1.1 为了规范融资融券交易行为，维护证券市场秩序，保护投资者的合法权益，根据《证券公司融资融券业务管理办法》《深圳证券交易所交易规则》及其他有关规定，制定本细则。

1.2 本细则所称融资融券交易，是指投资者向具有深圳证券交易所（以下简称本所）会员资格的证券公司（以下简称会员）提供担保物，借入资金买入证券或者借入证券并卖出的行为。

1.3 在本所进行的融资融券交易，适用本细则。本细则未作规定的，适用《深圳证券交易所交易规则》和本所其他有关规定。

上海证券交易所债券交易规则

第一章 总则

第一条 为规范债券交易行为，保护投资者合法权益，维护债券市场秩序，根据《中华人民共和国证券法》《证券交易所管理办法》等法律、行政法规、部门规章、规范性文件及《上海证券交易所章程》，制定本规则。

第二条 在上海证券交易所（以下简称本所）开展的债券、资产支持证券和其他具有固定收益特征的产品交易或者转让（以下简称债券交易）适用本规则。本所以对上市公司发行的可转换公司债券的交易另有规定的，从其规定。

在本所开展的债券交易类型包括债券现券交易、债券回购交易、债券预发行交易以及本所认可的其他交易类型。

本规则未作规定的，适用本所其他相关业务规则的规定。

问题

- 随着市场的快速发展和监管要求的不断变化，业务规则展现出易变性
- 如何提高业务规则规约和测试生成的效率成为一个关键问题

现有工作：人工

大语言模型时代，对证券交易所业务规则文档进行智能化分析，并自动化地生成测试用例？

附件 1

深圳证券交易所债券交易规则

第一章 总则

1.1 为规范债券交易行为，保护债券投资者合法权益，维护债券市场秩序，根据《中华人民共和国证券法》《证券交易所管理办法》等法律、行政法规、部门规章、规范性文件及《深圳证券交易所章程》，制定本规则。

1.2 在深圳证券交易所（以下简称本所）开展的债券、资产支持证券及其他具有固定收益特征产品的交易及转让（以下简称债券交易）适用本规则。上市公司发行的可转换公司债券的交易，本所另有规定的，从其规定。

在本所开展的债券交易类型包括债券现券交易、债券回购交易、债券预发行交易以及本所认可的其他交易类型。

本规则未作规定的，适用本所其他相关业务规则的规定。

1.3 债券交易应当遵循公开、公平、公正以及自愿、有偿、诚实信用等原则。

债券投资者（包括债券交易参与人和其他债券投资者）可以根据本所相关规定和自身交易需求，自主选择债券交易品种和交易结算方式。

1.4 本所为债券交易提供设施和相关服务，并依据本规则

- 1 -



规则id	3.1.5.3.1.1.1	用例id	3.1.5.3.1.1.1_1	关注点	时间
操作	状态	交易品种	交易市场	交易方向	交易方式
申报	未申报	债券	深圳证券交易所	买入	采用竞买成交
交易时间	债券品种	结果状态	应价方提交应价申报时间	卖方提交竞买发起申报时间	
09:00:00-11:30:00	债券现券	未成交	10:00:00-11:30:00	09:00:00-10:00:00	
结果					
成功					

规则id	3.1.5.3.1.1.1	用例id	3.1.5.3.1.1.1_2	关注点	时间
操作	状态	交易品种	交易市场	交易方向	交易方式
申报	未申报	债券	深圳证券交易所	买入	采用竞买成交
交易时间	债券品种	结果状态	应价方提交应价申报时间	卖方提交竞买发起申报时间	
00:00:00-09:00:00 或 11:30:00-24:00:00	债券现券	未申报	10:00:00-11:30:00	09:00:00-10:00:00	
结果					
不成功					

▶ LLMs用于规约与测试经验：焦点小组

基于大语言模型(LLMs)规约与测试的经验总结



8名参与者：

- 2名研究人员担任主持人
- 2名5年以上证券领域测试人员
- 2名1-2年经验证券领域测试人员
- 2名2-3年其他领域测试人员

这个主持人首先介绍了焦点小组的目的和形式会议，然后让参与者有时间分享他们自己的经历并讨论。

总结了不同领域的人在使用LLMs进行测试时遇到的三个常见问题，包括：

- (1) **对人的要求高。** 希望在测试等任务中有效使用LLMs的人应该掌握特定知识和提示词技能。
- (2) **不可控和难以处理的输出。** LLMs的输出是不可控和难以处理的，这使得生成的测试用例难以复制和有效追踪。
- (3) **有限的领域知识。** LLMs所具有的领域特定知识有限，受限于输入的数据。

具体示例

Model Input

Rule 1. If the matching trading method is adopted, the declared quantity of cash bonds shall be 100,000 yuan in face value or an integral multiple thereof. The portion less than 100,000 yuan in face value when sold shall be declared in one go.

Rule 2. The minimum change unit of the declared price using the matching trading method is 0.01 yuan, and if other trading methods are used, the minimum change unit of the declared price is 0.005 yuan.

Model Input

Simple Prompt

Please help me generate test cases for the above financial trading rules.

Introduction of scenario and task

Requirements on test data

Requirements on knowledge

Model Input

Carefully Designed Prompt

Suppose you are a testing expert in the financial field. Now there are some rules that require you to help generate test cases. It is required that the test cases should be as detailed as possible and use testing strategies such as equivalence classes and boundary values. The test cases should contain all the elements given in the rules, such as declaration methods, trading varieties, etc. In addition, test cases should also include other elements such as the trading direction, trading market, etc. Note that if there are abstract or referential words in the sentence, make them concrete. Please help me generate test cases for the above financial trading rules.

Model Output

Rule 1:

Test Case 1: Declare quantity of 300,000 yuan. *Valid.*
Test Case 2: Declare quantity of 250,000 yuan. *Invalid.*
Test Case 3: Declared quantity of 100,000 yuan. *Valid.*
Test Case 4: Declared quantity of 90,000 yuan. *Invalid.*
Test Case 5: Declared quantity of 0 yuan. *Invalid.*

Rule 2:

Test Case 1: Matching trading method with price change of 0.01 yuan. *Valid.*
Test Case 2: Matching trading method with price change of 0.007 yuan. *Invalid.*
Test Case 3: Other than matching trading method with price change of 0.005 yuan. *Valid.*
Test Case 4: Other than matching trading method with price change of 0.003 yuan. *Valid.*
Test Case 5: Any trading method with price change of 0 yuan. *Invalid.*

Model Output

Rule 1:

Equivalence Classes: Valid Class: Quantities of 100,000 yuan and its integral multiples. Invalid Class: Quantities less than 100,000 yuan and not multiples of 100,000 yuan.
Boundary Values: Exactly 100,000 yuan. Just above 100,000 yuan (e.g., 100,001 yuan). Just below 100,000 yuan (e.g., 99,999 yuan)

Test Case 1: *Method: Matching, Quantity: 100,000 yuan, Price Change: 0.01 yuan, Direction: Buy, Market: Domestic (Result? Valid or not?)*
Test Case 2: *Method: Other, Quantity: 200,000 yuan, Price Change: 0.005 yuan, Direction: Sell, Market: International*
Test Case 3: *Method: Matching, Quantity: 99,999 yuan (Invalid), Price Change: 0.01 yuan, Direction: Buy, Market: Domestic*
Test Case 4: *Method: Other, Quantity: 150,000 yuan, Price Change: 0.004 yuan (Invalid), Direction: Sell, Market: International*
Test Case 5: *Method: Matching, Quantity: 100,001 yuan, Price Change: 0.01 yuan, Direction: Buy, Market: Domestic*

▶ 证券领域业务规则规约与测试用例生成的难点

附件 1

深圳证券交易所债券交易规则

第一章 总则

1.1 为规范债券交易行为，保护债券投资者合法权益，维护债券市场秩序，根据《中华人民共和国证券法》《证券交易所管理办法》等法律、行政法规、部门规章、规范性文件及《深圳证券交易所章程》，制定本规则。

1.2 在深圳证券交易所（以下简称本所）开展的债券、资产支持证券及其他具有固定收益特征产品的交易及转让（以下统称债券交易）适用本规则。上市公司发行的可转换公司债券的交易，本所另有规定的，从其规定。

在本所开展的债券交易类型包括债券现券交易、债券回购交易、债券预发行交易以及本所认可的其他交易类型。

本规则未作规定的，适用本所其他相关业务规则的规定。

1.3 债券交易应当遵循公开、公平、公正以及自愿、有偿、诚实信用等原则。

债券投资者（包括债券交易参与人和其他债券投资者）可以根据本所相关规定和自身交易需求，自主选择债券交易品种和交易结算方式。

1.4 本所为债券交易提供设施和相关服务，并依据本规则

- 1 -

自然语言表示

- ? 不受限句式，有很多省略
- ? 领域术语多
- ? 抽象表示多

不仅仅是自然语言处理的问题，还是规约的问题，从抽象到具体可操作规则，需要领域知识的支持

规则间关系多

- ? 显式关系：字面表示
- ? 隐式关系：领域知识

不仅仅是自然语言处理的问题，还需要领域知识的支持

规约问题：自然语言处理+领域知识

▶ 我们的工作

为了解决上述问题并提高证券领域中交易规则文档规约和测试生成的效率
提出了一种结合**大语言模型（LLM）与领域知识库**的证券领域业务规则自动规约和测试用例
生成方法

- (1) 提出了第一个完全自动化的LLM驱动方法，无缝整合人工智能模型和算法，从自然语言表达的业务规则中生成覆盖面广的证券交易软件验收测试用例。
- (2) 实现了提出的方法原型工具LLM4Fin，用于真实世界的交易系统。据我们所知，这标志着第一个行业级别的完全自动化的LLM驱动的业务规则测试应用。
- (3) 我们进行了广泛的评估，展示了其在大约7秒内生成测试用例时，其在业务场景覆盖率上的重大进步，最高可达98.18%。

PART 02

方法框架

► 具体的问题

1. 自然语言处理的问题

(1) 筛选任务

交易所官网交易业务规则文档内除了具体交易业务规则描述，还**包含业务无关的内容**，因此必须要**准确识别并筛选出**哪些业务规则描述与交易系统业务需求、测试相关。

1.1 为规范债券交易行为，保护债券投资者合法权益，维护债券市场秩序，根据《中华人民共和国证券法》《证券交易所管理办法》等法律、行政法规、部门规章、规范性文件及《深圳证券交易所章程》，制定本规则。



软件需求无关

3.3.6 采用匹配成交方式的，债券现券的申报价格最小变动单位为 0.001 元，债券通用质押式回购的申报价格最小变动单位为 0.005 元；采用其他交易方式的，债券交易的申报价格最小变动单位为 0.0001 元。本所另有规定的除外。



软件需求相关

► 具体的难点

1. 自然语言处理的问题

(2) 抽取任务：不受限句式处理，补全，处理大量的领域术语

规则3.3.4 采用**匹配成交**方式的，**债券现券**的申报数量应当为10万元面额或者其整数倍，卖出时不足10万元面额部分，应当**一次性申报**；**债券通用质押式回购**的申报数量应当为1000元面额或者其整数倍。

2. 规则规约的问题

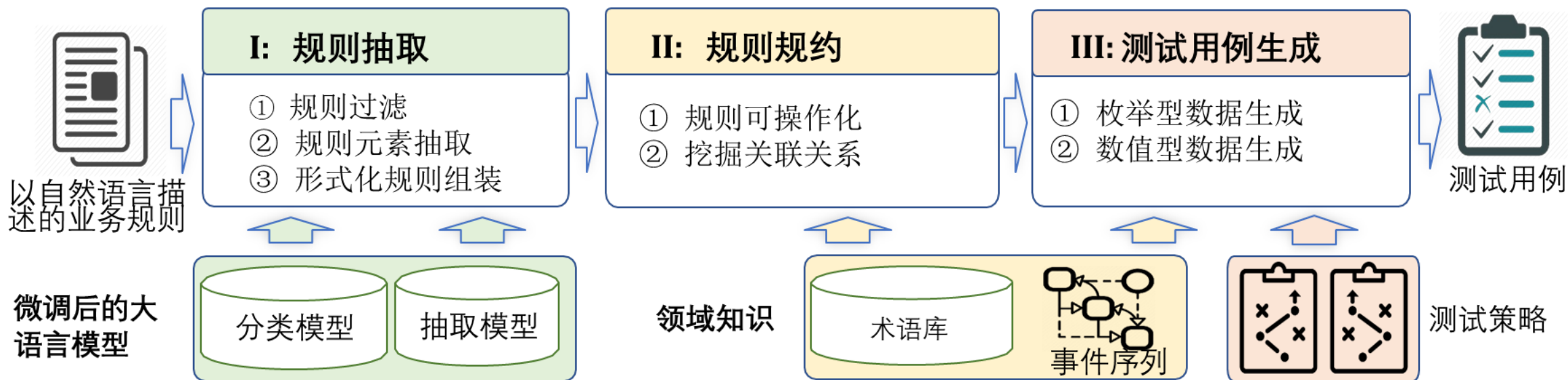
从抽象到具体规则：领域知识

规则间关系的深度理解：领域知识

规则3.3.12 **除本规则规定的不接受撤销申报的时间段外**，其他接受申报的时间内，未成交的申报可以撤销；部分达成成交的，未成交的部分可以撤销。

规则4.1.10 每个交易日 9:20 至 9:25 的开盘集合匹配阶段，本所交易系统**不接受匹配成交的撤销申报**。

方法框架



PART 03

大语言模型微调

▶ 分类模型

分类的标注准则：

- 领域知识：描述了证券领域的专业知识、术语和概念，用于帮助理解和应用业务规则。
- 需求相关规则：约束了证券系统行为和操作，涉及到时间、数量、价格、成交方式、定价方式等具体细节，可以通过观察、测量和操作来验证其准确性和符合性，是需要软件严格遵循和执行的规则。
- 需求无关的规则：不直接对软件的操作功能提出要求。

分类模型训练数据：

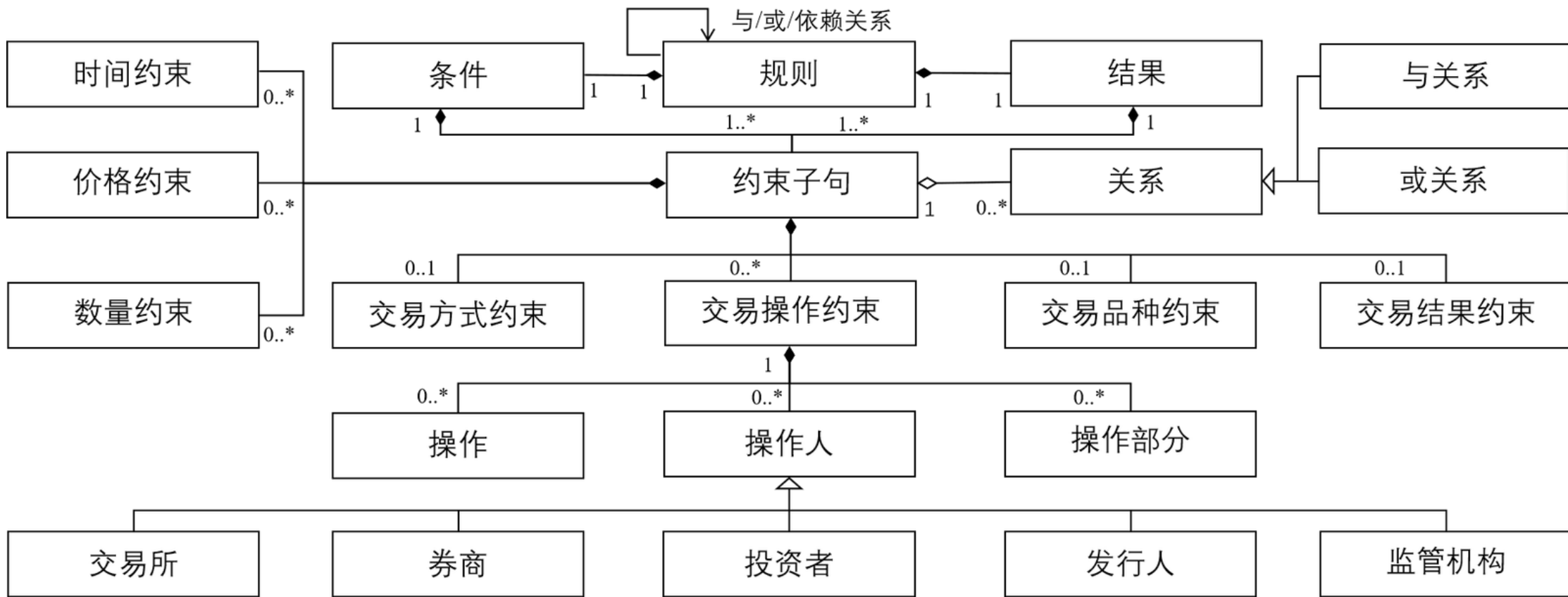
- 数据增强训练数据32945条，测试数据333条

分类模型的训练：

- 1.选择中文预训练模型**Mengzi 103M**作为基础模型框架
- 2.在手动标注的语料库上对Mengzi模型进行了微调，以实现噪声规则的过滤
- 3.在模型训练过程中采用了交叉熵作为损失函数，并使用AdamW优化器执行梯度下降等方法来不断优化模型的处理能力
- 4.最终规则分类模型的准确率达到**99.1%**

抽取模型

标注：构建证券领域元模型



证券领域元模型

抽取模型

基于上述交易规则元模型和对自然语言规则文档的结构分析制定用于信息提取的标签

15
↑
个
标
签

交易市场：深圳证券交易所、上海证券交易所等

交易品种：债券、基金、股票、债券现券、债券通用质押式回购等

交易方式：匹配成交、竞买成交等

操作：申报、撤销等交易动作

操作人：会员、经纪客户、申报人等交易操作人

操作部分：未成交的申报等，为操作的宾语

结果：成功、失败等

状态：部分成交的、临时停市的、不能一次全部成交时等交易进行的状态

事件：经纪客户委托会员申报并达成债券交易的等交易事件

结合规则：本规则第3.1.5条等

op：超过、小于等表达比较关系的词

时间：交易日9:00-10:00等表示交易时间的词

数量：10万元面额或者其整数倍等表示交易数量的词

价格：1.0元等表示交易价格的词

key：其他标签

value：key对应的值

▶ 抽取模型

1. 使用**Mengzi 103M**作为预训练模型来完成规则抽取任务
2. 按照上述15个标签将18篇文档中的所有规则进行标注
3. 使用数据增强技术对训练集进行扩充
 - 扩充后的训练集具有41215条标注好的数据,
 - 验证集有634条
4. 训练的设置和分类模型类似, 不同的是
 - 模型 $2e-5$ 的学习率下训练20个epoch,
 - 并额外添加0.002的权重衰退防止过拟合
5. 最终规则抽取的模型准确率达到了**87.0%**

PART 04

生成方法细节

▶ 第一步：规则抽取

① 规则过滤：使用构建好的分类模型对自然语言表示的业务规则进行过滤。

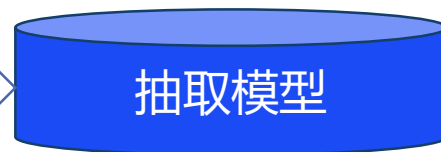
Rule 1: 采用点击成交方式的，债券交易的申报数量应当为 10 万元面额或者其整数倍。
Rule 2: 债券交易参与者应当加强对债券交易信用风险、市场风险以及流动性风险等风险的监测和管理。
.....



Rule 1: 可测试的规则
Rule 2: 领域知识
Rule 3: 无需测试的规则

② 规则元素抽取：

Rule 1: 采用点击成交方式的，债券交易的申报数量应当为 10 万元面额或者其整数倍。



交易方式：点击成交
交易品种：债券
申报数量：10万元面额或者其整数倍

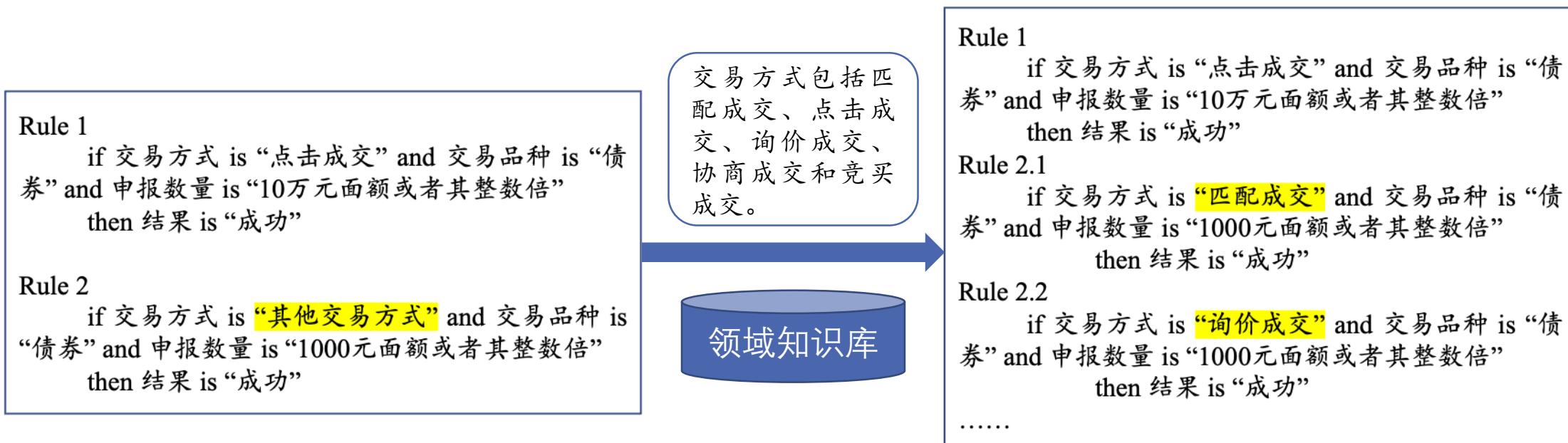
③ 形式化规则组装：

Rule 1
if 交易方式 is “点击成交” and 交易品种 is “债券” and 申报数量 is “10万元面额或者其整数倍”
then 结果 is “成功”

第二步：规则规约

① 规则可操作化：业务规则经常包含抽象表达或依赖于特定上下文的表示方式，需要将其具体化，从而使规则易于理解和测试。

这一步骤主要依赖于领域知识，将规则中的抽象表述替换成可操作的具体概念，确保生成可执行的规则。

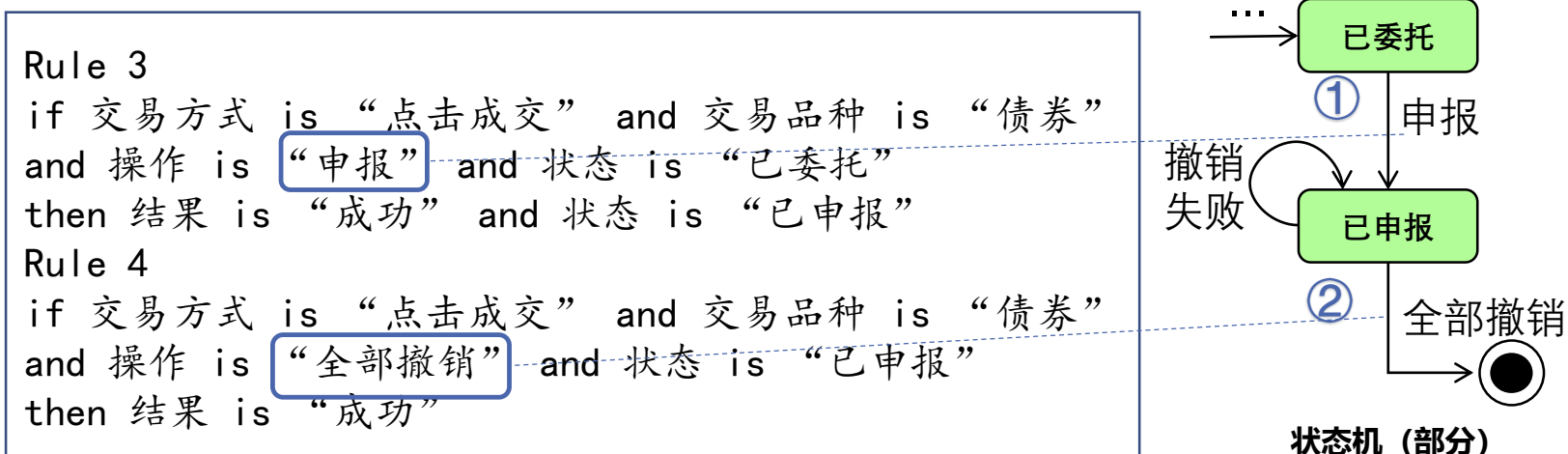


第二步：规则规约

② 挖掘关联关系：业务规则之间存在着大量显式与隐式的依赖关系。

生成完整的、可操作的测试用例需要全面考虑不同规则之间的关联。

这一步包括针对同一要素（例如申报数量）的多条规则的组合、挖掘规则的前置条件和后置条件并结合领域知识库中的状态机关联不同规则的先后顺序等。



Rule 3 先于 Rule 4

▶ 第三步：测试用例生成

➤ 测试场景+测试数据生成

- 枚举型数据生成：结合知识库遍历标签的所有值，每个值对应一条测试用例
- 数值型数据生成：将数值型约束使用数学表达式表示，并结合边界值、等价类等多种测试策略，使用SMT求解器生成满足与不满足约束的两类测试数据。

交易市场	交易品种	交易方式	交易方向	申报数量	结果
深交所	债券	点击成交	买入	100000、200000	成功
深交所	债券	点击成交	卖出	100000、1000000	成功
深交所	债券	点击成交	买入	99999、100001、1	不成功
深交所	债券	匹配成交	卖出	1000、2000	成功
.....					

PART 05

实验和评估

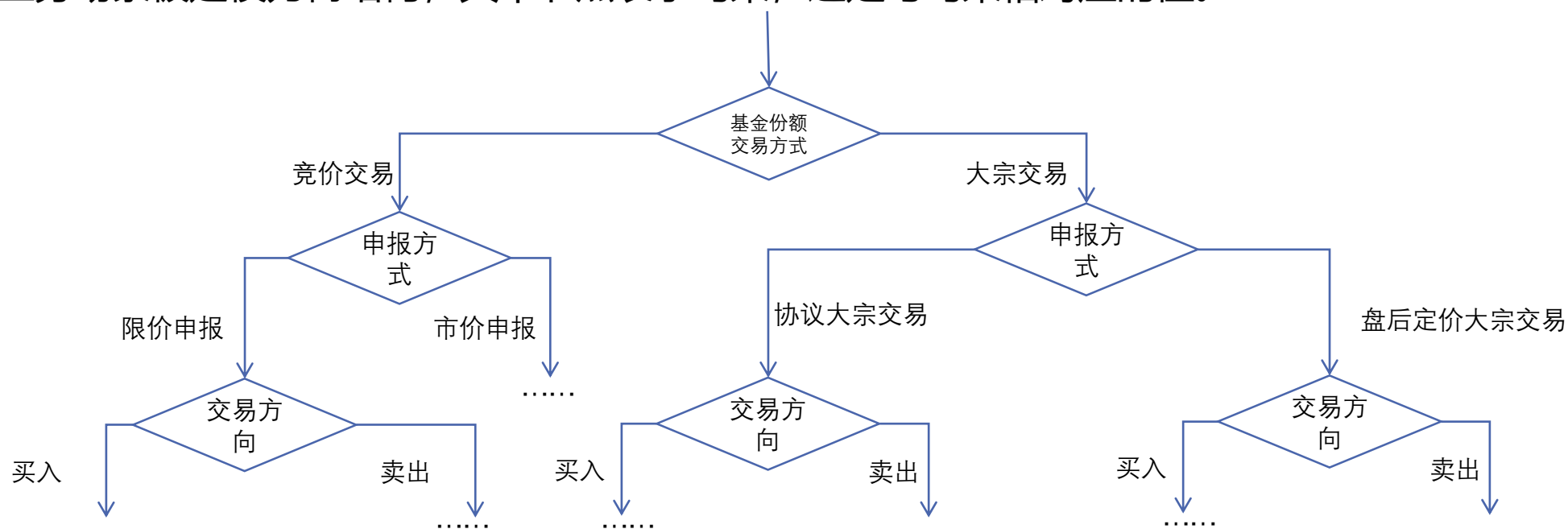
实验评估

➤ 评估问题：方法的有效性和效率、大语言模型对本方法的影响

➤ 评估指标

- 业务场景覆盖率 (BSC)

- ✓ 业务场景被建模为树结构，其中节点表示约束，边是与约束相对应的值。



- ✓ $BSC = p_t / p_a$, p_t 是触发场景的数量, p_a 是总数量

- ✓ BSC越高, 生成的测试用例覆盖范围越广

▶ 实验评估

➤ 评估指标

- **代码覆盖率 (CC)**

- ✓ 从源代码端度量生成的测试用例的覆盖率
- ✓ 使用广泛采用的指标，如语句块覆盖 (SBC) 和修改的条件/决策覆盖 (MC/DC) 来评估

- **时间**

- ✓ 使用生成或组合测试用例的时间成本来评估效率

▶ 实验评估

➤ 比较对象

- 目前从业务规则文档生成测试用例的工作主要由人工完成。
 - ✓ 参与实验的人类参与者包括3位领域专家和2位非专家。
- 大语言模型也能够自动生成测试用例
 - ✓ ChatGPT, ChatGLM

Table 1: Details of Competitors.

Competitor	Description
Expert testers	Three senior testing engineers with 5-7 years of experience in FinTech software centers.
Non-expert testers	Four graduate students with approximately one year of software testing knowledge acquired from courses.
ChatGPT [43]	GPT-4, including 1800 billion parameters.
ChatGLM [13]	A LLM fine-tuned in Chinese with 130 billion parameters, leading the c-eval [20] leaderboard.

实验评估

评估数据集

数据集构成

交易所：上海和深圳证券交易所

交易品种：债券、股票、基金和创业板等

交易方式：竞价交易、大宗交易和盘后定价交易

随机选择了具体业务

Table 2: Details of Evaluation Datasets. #Rules refers to the number of rules in the dataset and #BS refers to the number of business scenario.

Dataset	Sub-domain	# Rules	# BS
Dataset 1	<i>GEM after-hours pricing trading</i> : trading method conducted outside regular market hours.	11	12
Dataset 2	<i>Stock block trading</i> : buying or selling a large block of shares in a single transaction.	12	40
Dataset 3	<i>Fund trading</i> : transaction where investors contribute funds managed by fund managers.	13	37
Dataset 4	<i>Convertible bond trading</i> : trading variety with characteristics of both bonds and stocks.	8	24
Dataset 5	<i>Stock auction trading</i> : a trading method in which trades are matched by price and time priority.	13	400

▶ 实验结果一：业务场景覆盖率

本方法（LLM4Fin）与专家、非专家、大语言模型ChatGPT、ChatGLM 在五个数据集上的测试用例数目（#TC）、业务场景覆盖率（BSC）和时间（Time）的比较：

Datasets	Experts				Non-Experts				ChatGPT				ChatGLM				LLM4Fin		
	#TC	BSC (%)	Impr. (%)	Time	#TC	BSC (%)	Impr. (%)	Time	#TC	BSC (%)	Impr. (%)	Time	#TC	BSC (%)	Impr. (%)	Time	#TC	BSC (%)	Time
Dataset 1	24	83.43	15.1	33m	22	45.83	109.6	105m	24	44.61	115.3	30m	54	27.48	249.6	20m	218	96.06	7.01s
Dataset 2	50	83.80	2.5	38m	34	38.58	122.6	90m	42	50.48	70.1	25m	67	45.01	90.8	9m	672	85.36	6.85s
Dataset 3	168	71.15	38.0	30m	33	46.79	109.8	110m	32	61.81	58.8	17m	33	43.88	123.7	20m	270	98.18	6.82s
Dataset 4	30	77.34	25.4	35m	29	42.47	128.3	80m	34	25.00	287.8	10m	35	83.02	16.8	30m	88	96.96	5.57s
Dataset 5	67	60.10	38.0	35m	30	43.64	90.0	110m	41	66.35	25.0	15m	51	8.77	845.4	20m	880	82.91	8.72s
Average	68	75.16	22.3	34m	30	43.46	111.4	99m	35	49.65	85.1	19m	48	41.63	120.7	20m	426	91.89	6.99s

本方法在五个数据集上的业务场景覆盖率最高，同时时间消耗最少。

▶ 实验结果二：代码覆盖率

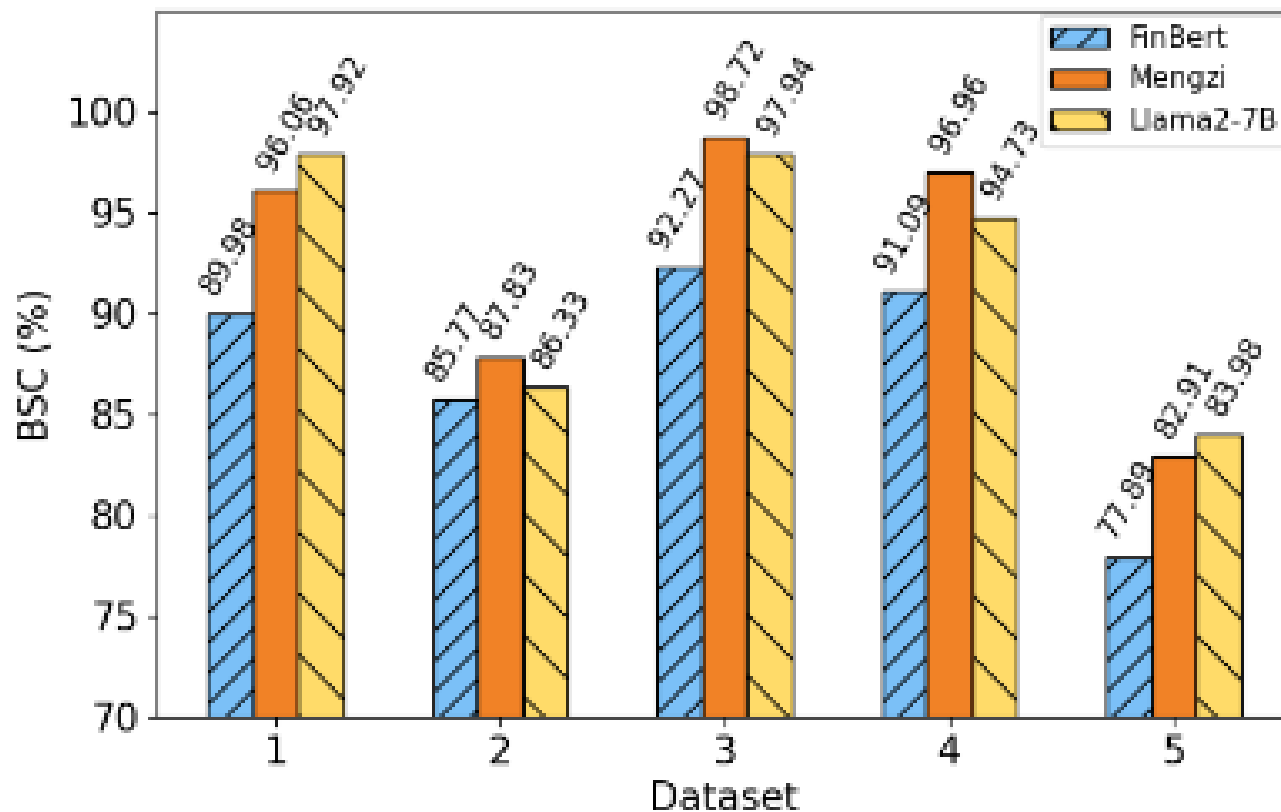
本方法 (LLM4Fin) 与专家、非专家、大语言模型ChatGPT 在数据集5上的语句块覆盖率 (SBC)、修订的条件/判定覆盖率 (MC/DC) 的比较：

	Experts	Non-Experts	ChatGPT	LLM4Fin
SBC (%)	86.90 (1466)	71.01 (1198)	87.61 (1478)	93.72 (1581)
MC/DC (%)	82.30 (1990)	64.06 (1549)	81.76 (1977)	90.86 (2197)

本方法在数据集上的语句块覆盖率最高，且修订的条件/判定覆盖率最高。

▶ 实验结果三：大语言模型的影响

- Mengzi
 - FinBert
 - Llama2 (7B)
- 只用在第一步规则抽取



无论使用何种LLM，BSC始终保持在80%以上。

高水平的LLM可以作为替代，确保我们的方法的可转移性。

PART 06

总结

▶ 总结

- 我们开发了一种全自动的方法，无缝集成LLMs和确定性算法
 - ✓ 在证券软件验收测试中生成高覆盖率的测试用例
 - ✓ 确保确定性和可追溯性
 - ✓ 这种结合在克服解释非结构化业务规则的固有挑战方面起着关键作用，使整个过程变得更有效和可管理
- 我们的方法得到了实际应用
 - ✓ 专门为证券领域的复杂环境而定制的原型
- 我们的评估突出显示了出色的性能
 - ✓ BSC的改进率高达98.18%，平均改进率为20% - 110%
 - ✓ 代码覆盖率达93.72%
 - ✓ 同时将时间成本从20分钟减少到仅7秒
- 更多探索
 - ✓ 未来考虑RAG(检索增强生成)，更通用的方法
 - ✓ 改进人类专家定义的规则测试生成、一致性检测等

科技生态圈峰会 + 深度研习



—1000+ 技术团队的选择



K+峰会

上海站

K+ 全球软件研发行业创新峰会

时间: 2024.06.21-22

K+峰会

敦煌站

K+ 思考周®研习社

时间: 2024.10.17-19

K+峰会

香港站

K+ 思考周®研习社

时间: 2024.11.10-12



K+峰会详情



AIDD峰会

上海站

AI+研发数字峰会

时间: 2024.05.17-18

AIDD峰会

北京站

AI+研发数字峰会

时间: 2024.08.16-17

AIDD峰会

深圳站

AI+研发数字峰会

时间: 2024.11.08-09



AIDD峰会详情

THANKS

