

200127

上海市浦东新区杨高南路288号19-21层 上海金融期货信息技术有限公司 李悦萌 女士

关于: 申请号为202210112284.8的专利申请案

申 请 人: 上海金融期货信息技术有限公司

发明名称:金融产品价格影响因素传导分析方法及系统

我方编号: CNJRQH-0151.21A579

李悦萌 女士

您好!

关于本案的第1次审查意见通知书,已按贵方指令于期限内答复国知局,附件是答复文本的副本,请查收。

如有问题请随时联系我们。谢谢! 顺颂 业琪

专利代理师: 施浩 2025年4月15日

Morta

Encls.

hdc.doc

地址: 上海市桂平路435号 电话: 86-21-34183200 邮箱: info@sptl.com.cn 邮政编码: 200233 传真: 86-21-64828651/2 网址: www.sptl.com.cn



分支机构: 北京 / 临港 / 海南 / 南京

意见陈述书

1	申请号或专利号 2022101122848	
专专		
利		
或	 申请人或专利权人(第一署名人)上海金融期货信息技术有限公司	
申	中国八头专利仪八(另一省石八)上/博並融州页目/总仪个有限公司 	
请 利	 关于费用的意见陈述请使用意见陈述书(关于费用)	
	: 大丁贡加的总元协业用使加总元协业 [1大丁贡加] [只能选择一项	
_		
★ 計对国家知识产权局于 2025 年 03 月 22 日发出的 第 1 次审查意见通知书 (发文序号 2025032200024860)陈述意见。		
□针对国家知识产权局于年月日发出的(发文序号)补充陈述意见。		
□针对国家知识产权局于年月日发出的药品专利权期限补偿审查意见通知书(发文序		
号) 陈述意见。		
□主动提出修改(根据专利法实施细则第 57 条第 1 款、第 2 款的规定)		
│ □ 公布公告事项 □ いいよう		
其他事宜		
③关于补交实验数据的情况		
□补交了实验数据		
④陈述的意见:		
意见陈述请见附件。		
⑤ 附件清单		
【附件名称】权利要求书		

100012

意见陈述书

【附件名称】修改对照页
【附件名称】其他证明文件
已备案的证明文件备案编号:
上海专利商标事务所有限公司

100012 2

意见陈述书正文

尊敬的审查员, 您好!

针对本案的审查意见,申请人的意见陈述如下。

一、针对权利要求1-12不具备创造性的审查意见

权利要求的修改:

将"其中,步骤2进一步包括:

步骤2.1:基于预处理后的数据,抽取出其中的因果触发词及所在句,作为原因、结果子句集合:

步骤2.2:针对原因、结果子句集合,使用基于规则模板的正则匹配进行事件抽取,抽取出显式因果语句中的原因子句和结果子句;

步骤2.3:针对显式因果语句中的原因子句和结果子句,基于依存句法分析确定事件表达,并构建出事理图谱;

其中,在步骤2.2中,基于规则模板的正则匹配方法抽取显式因果语句中的原因子句和结果子句,和因果触发词共同组成事件因果三元组,为后续构建金融衍生品价格影响因素事理图谱做好准备,其中,规则模板的格式为〈Pattern, Constraint, Priority〉,Pattern 是包含因果触发词的正则表达式,Constraint包含了句子中的句法约束,Priority 是当多个模板都匹配上时匹配模板的优先级;

其中, 步骤3进一步包括:

步骤3.1:对有标签的文本进行有监督学习;

步骤3.2: 使用无监督学习框架对无标签的文本进行一致性训练;

步骤3.3:针对无标签数据和有标签数据使用半监督学习学习框架进行训练,得到损失函数,作为模型的目标优化函数用于训练,以使模型学习到无标签数据和有标签数据中的信息;

其中,在步骤3的有监督学习中加入数据增强方式"的技术特征加入到原独 立权利要求1中; 将"其中,事理图谱构建模块进一步配置为:

抽取单元,基于预处理后的数据,抽取出其中的因果触发词及所在句,作为原因、结果子句集合:

匹配单元,针对原因、结果子句集合,使用基于规则模板的正则匹配进行 事件抽取,抽取出显式因果语句中的原因子句和结果子句;

构建单元,针对显式因果语句中的原因子句和结果子句,基于依存句法分析确定事件表达,并构建出事理图谱:

其中,在匹配单元中,基于规则模板的正则匹配方法抽取显式因果语句中的原因子句和结果子句,和因果触发词共同组成事件因果三元组,为后续构建金融衍生品价格影响因素事理图谱做好准备,其中,规则模板的格式为〈Pattern, Constraint, Priority〉,Pattern 是包含因果触发词的正则表达式,Constraint包含了句子中的句法约束,Priority 是当多个模板都匹配上时匹配模板的优先级:

其中, 传导分析模块进一步配置为:

有监督学习单元,对有标签的文本进行有监督学习;

无监督学习单元,使用无监督学习框架对无标签的文本进行一致性训练;

损失函数训练单元,针对无标签数据和有标签数据使用半监督学习学习框架进行训练,得到损失函数,作为模型的目标优化函数用于训练,以使模型学习到无标签数据和有标签数据中的信息:

其中,在传导分析模块的有监督学习单元中加入数据增强方式"的技术特征加入到原独立权利要求7中。

上述修改来自于申请文本原始记载的范围,符合专利法第33条的规定。

意见陈述理由:

申请人认为,修改后的独立权利要求1是具备创造性的,理由如下。

本案所要解决的技术问题是:金融衍生品包括股指期货、国债期货等在优 化资源配置、降低国家总体的金融风险、吸纳社会闲置资金、有效地增加市场的 流动性、提高交易效率等方面起着很大作用。近些年随着国内外经济环境日趋复杂,金融衍生品价格受到货币政策、财政政策、经济基本面、市场流动性,股价涨跌等多方面长短不同周期的因素影响,影响因素间错综复杂的关系,给业务人员分析金融衍生品关键影响因素带来了巨大的困扰,为机构投资者做好风险管理提出了新的挑战。因此人们迫切需要一种自动从海量新闻中构建金融衍生品价格影响因素的方法,帮助人们从海量金融新闻中找到影响金融衍生品价格的核心影响因素并总结其演化规律。我们需要的是一个自动从新闻中提取知识、组织并构建金融事件对金融衍生品价格因果传导因素的工具或系统。一般来说,要解决这个问题,需要人工阅读大量金融相关研报、新闻并根据专家经验对专家对宏观经济传导关系进行归纳总结,构建影响金融衍生品价格传导链路。然而由于金融领域研报和新闻获取成本高、效率低,无法满足对金融衍生品价格波动实时性分析的业务场景。如何做到可以在不依赖人工参与的情况下来解决金融衍生品影响因素传导分析,是目前业界亟待解决的问题。然而目前市场上没有一款产品,能够自动化解决金融衍生品价格波动成因分析。

为解决上述技术问题,本案的关键技术特征是:引入了两方面的技术处理: 一是结合句法模板和依存句法分析技术从研报中抽取事件因果关系构成事理图谱,二是基于语言预训练模型为金融新闻打上相应的金融分类标签便于业务人员检索和归因分析。这两方面的结合可以大大提升业务人员的新闻检索效率,在不依赖人工参与的前提下自动构建影响金融产品价格的事件传导关系,帮助业务人员从各方面分析金融产品价格波动分析成因提供参考。

详细而言,本发明的创新点如下:

创新点1:基于事理图谱的整个架构的设计。目前,市场上没有一款产品针对金融产品市场构建价格影响因素事理图谱,本发明第一次基于依存句法分析、语言预训练模型等自然语言处理技术解决了以往依赖人工总结金融产品价格传导因素成本高,耗时长的痛点,构建了自动从新闻中提取知识并构建事件因果传导因素的工具。

创新点2:本方法使用句法模板抽取事件因果关系对,且句法模板具有可扩展性。

创新点3:本发明针对金融领域标签数量少,获取难度高的特点,采用了基

于语言预训练模型的半监督学习框架,通过使用中文金融新闻的数据增强方法对 样本中的无监督数据进行数据增强,提高了无标签数据的利用率,有效提升新闻 标签的准确率。

对比文件1的技术摘要为: 传统的知识图谱大多关注实体的属性与关系知识, 而忽视了事件间的演化规律知识,为了弥补这一不足,研究者们提出了事理图谱 的概念。事理图谱中的节点是高度泛化的事件,边是事件间的演化关系,例如因 果关系与顺承关系。自然语言处理技术的发展使得信息抽取的准确率得到提升, 也使得从文本中自动挖掘事理知识、构建事理图谱成为可能。本文从金融领域入 手,对事理图谱自动构建的关键技术展开研究,具体内容包括端到端的事件因果 关系抽取、常识信息增强的事件表示学习以及数据驱动的因果关系强度计算。事 理知识的获取是构建事理图谱的基础,本文针对事件间因果关系的获取进行了探 索。本文将事件因果关系抽取建模为序列标注任务,提出了基于预训练模型的因 果抽取方法,以端到端的方式同时进行因果关系的识别与相关事件的抽取。为缓 解有标注数据不足的问题,本文讲一步提出基于噪声模型的半监督学习方法,利 用大量无标注数据提升因果抽取的效果。中英文两个因果抽取数据集上的实验结 果证明了该方法的有效性。事件是事理图谱的核心元素,为了更好地建模事件语 义,本文提出了常识信息增强的事件表示学习方法,使学习到的事件表示中融入 意图、情感、实体关系等常识信息, 以更好地帮助事理图谱的构建以及在其他任 务上的应用。事件相似度、脚本事件预测、股市预测等多个任务上的实验结果表 明我们的方法可以更准确地建模事件语义,并提升下游任务上的效果。为了更好 地建模事件间因果关系的强度, 本文探索了基于统计与基于预训练模型的因果 强度计算方法,从大量因果事件对中自动学习因果强度信息。因果推理数据集上 的实验结果表明,预训练模型可以有效地从大量因果事件对中学习因果知识,并 准确地建模因果关系强度。最后,本文基于上述研究成果设计并实现了金融领域 事理图谱构建系统,并在大规模金融语料上构建了包含数百万事件与因果关系的 事理图谱,验证了本文提出的事理图谱构建方法的可行性。

对比文件2的技术摘要为:一种面向金融场景的端到端自然语言处理训练框架与方法,涉及自然语言处理领域,所述训练框架包括:Google原生BERT模块、

FinBERT预训练模块、基于类似self-training思想从外部相关数据中挖掘新数据模块、下游任务语料上进行预训练模块、用半监督学习的框架来充分利用无标签语料模块、知识蒸馏模块、线上部署模块。所述训练方法包括以:步骤1、FinBERT预训练;步骤2、基于类似self-training思想从外部相关数据中挖掘新数据;步骤3、下游任务语料上进行预训练,对所述FinBERT进行再一次预训练,得到的模型称之为TASK FinBERT;步骤4、用半监督学习的框架来充分利用无标签语料,在TASK FinBERT的基础上,训练得到的模型称之为UDA FinBERT;步骤5、蒸馏学习,将学习到的知识和特征蒸馏到轻量级模型上。

将本发明的修改后的独立权利要求1的技术方案与对比文件1-2的技术方案相比,区别如下。

一、本发明中,步骤2:基于预处理后的数据,抽取事件因果对,基于事件因果对构建事理图谱。

其细化处理过程为:

步骤2进一步包括:

步骤2.1:基于预处理后的数据,抽取出其中的因果触发词及所在句,作为原因、结果子句集合:

步骤2.2:针对原因、结果子句集合,使用基于规则模板的正则匹配进行事件抽取,抽取出显式因果语句中的原因子句和结果子句;

步骤2.3:针对显式因果语句中的原因子句和结果子句,基于依存句法分析确定事件表达,并构建出事理图谱。

上述的细化处理过程,并没有被对比文件1-2所全部揭示。

且,该区别技术特征并非本领域的惯用技术手段。若审查意见认为是惯用手段,应不难通过检索提出对比文件,仅在未提出任何书面证据的情况下就简单断言区别技术特征是惯用手段,并不符合审查指南第二部份第八章第4.10.2.2节的规定:「审查员在审查意见通知书中引用的本领域的公知常识应当是确凿的,如果申请人对审查员引用的公知常识提出异议,审查员应当能够说明理由或提供相应的证据予以证明」。

二、本发明中,步骤3中是"针对无标签数据和有标签数据使用半监督学习学习框架进行训练,得到损失函数,作为模型的目标优化函数用于训练,以使模型学习到无标签数据和有标签数据中的信息",而且在步骤3的有监督学习中加入数据增强方式。

使用半监督学习框架且在有监督学习中加入数据增强方式。这在对比文件 1-2中没有全部揭示。

且此区别技术特征并非本领域的惯用技术手段。若审查意见认为是惯用手段,应不难通过检索提出对比文件,仅在未提出任何书面证据的情况下就简单断言区别技术特征是惯用手段,并不符合审查指南第二部份第八章第4.10.2.2节的规定:「审查员在审查意见通知书中引用的本领域的公知常识应当是确凿的,如果申请人对审查员引用的公知常识提出异议,审查员应当能够说明理由或提供相应的证据予以证明」。

且此区别技术特征可以带来如下的技术效果:针对金融领域标签数量少, 获取难度高的特点,采用了基于语言预训练模型的半监督学习框架,通过使用中 文金融新闻的数据增强方法对样本中的无监督数据进行数据增强,提高了无标签 数据的利用率,有效提升新闻标签的准确率。

三、在本发明中:基于规则模板的正则匹配方法抽取显式因果语句中的原因子句和结果子句,和因果触发词共同组成事件因果三元组,为后续构建金融衍生品价格影响因素事理图谱做好准备。其中,规则模板的格式为〈Pattern, Constraint, Priority〉。Pattern 是包含因果触发词的正则表达式; Constraint包含了句子中的句法约束, Priority 是当多个模板都匹配上时,匹配模板的优先级。

这一实现在对比文件1-2中没有揭示。

且此区别技术特征并非本领域的惯用技术手段。若审查意见认为是惯用手段,应不难通过检索提出对比文件,仅在未提出任何书面证据的情况下就简单断言区别技术特征是惯用手段,并不符合审查指南第二部份第八章第4.10.2.2节的

规定:「审查员在审查意见通知书中引用的本领域的公知常识应当是确凿的,如果申请人对审查员引用的公知常识提出异议,审查员应当能够说明理由或提供相应的证据予以证明」。

综上三点,修改后的独立权利要求1具有突出的实质性特点和显著的进步, 具备创造性。依附其的全部从属权利要求均依法具备创造性。

同理,修改后的独立权利要求4具有突出的实质性特点和显著的进步,具备创造性。依附其的全部从属权利要求均依法具备创造性。

以上陈述,如有不妥,请指正并给与再次修改的机会,也可通过 021-34183200-2409 和代理人联系。申请人愿意积极配合以使本申请能够早日授 权。谢谢!

权 利 要 求 书

- 1、一种金融产品价格影响因素传导分析方法,其特征在于,方法包括:
- 步骤 1: 对自动获取的原始数据进行预处理;
- 5 步骤 2: 基于预处理后的数据,抽取事件因果对,基于事件因果对构建事理图谱;
 - 步骤 3: 基于语言预训练模型的构建分类算法对步骤 2 所构建的事理图谱中的因果事件对自动赋予标签,基于语言预训练模型自动赋予的标签用于价格影响因素的传导分析;
- 10 其中,步骤2进一步包括:
 - 步骤 2.1: 基于预处理后的数据,抽取出其中的因果触发词及所在句,作为原因、结果子句集合;
 - 步骤 2.2: 针对原因、结果子句集合,使用基于规则模板的正则匹配进行事件抽取,抽取出显式因果语句中的原因子句和结果子句;
- 5 步骤 2.3: 针对显式因果语句中的原因子句和结果子句,基于依存句法分析确定事件表达,并构建出事理图谱;

其中,在步骤 2.2 中,基于规则模板的正则匹配方法抽取显式因果语句中的原因子句和结果子句,和因果触发词共同组成事件因果三元组,为后续构建金融衍生品价格影响因素事理图谱做好准备,其中,规则模板的格式为<Pattern, Constraint,

20 Priority>, Pattern 是包含因果触发词的正则表达式, Constraint 包含了句子中的句法约束, Priority 是当多个模板都匹配上时匹配模板的优先级;

其中, 步骤 3 进一步包括:

- 步骤 3.1: 对有标签的文本进行有监督学习:
- 步骤 3.2: 使用无监督学习框架对无标签的文本进行一致性训练;
- 25 步骤 3.3: 针对无标签数据和有标签数据使用半监督学习学习框架进行训练,得到损失函数,作为模型的目标优化函数用于训练,以使模型学习到无标签数据和有标签数据中的信息:

其中,在步骤3的有监督学习中加入数据增强方式。

2、根据权利要求 1 所述的金融产品价格影响因素传导分析方法,其特征在于,步骤 1 中的原始数据包括新闻、研报在内的舆情数据,是采用分布式爬虫系统自动从互联网上抓取;预处理包括对原始数据进行去重、分词、和词性标注。

5

- 3、根据权利要求 1 或 2 所述的金融产品价格影响因素传导分析方法,其特征在于,传导分析的结果以事件为节点,以事件之间的关系为边,存入图数据库中以实现可视化效果。
- 10 4、一种金融产品价格影响因素传导分析系统, 其特征在于, 系统包括: 预处理模块, 对自动获取的原始数据进行预处理:

事理图谱构建模块,基于预处理后的数据,抽取事件因果对,基于事件因果 对构建事理图谱;

传导分析模块,基于语言预训练模型的构建分类算法对事理图谱构建模块所 15 构建的事理图谱中的因果事件对自动赋予标签,基于语言预训练模型自动赋予的 标签用于价格影响因素的传导分析:

其中, 事理图谱构建模块进一步配置为:

抽取单元,基于预处理后的数据,抽取出其中的因果触发词及所在句,作为原因、结果子句集合;

20 匹配单元,针对原因、结果子句集合,使用基于规则模板的正则匹配进行事件抽取,抽取出显式因果语句中的原因子句和结果子句:

构建单元,针对显式因果语句中的原因子句和结果子句,基于依存句法分析确定事件表达,并构建出事理图谱;

其中,在匹配单元中,基于规则模板的正则匹配方法抽取显式因果语句中的 原因子句和结果子句,和因果触发词共同组成事件因果三元组,为后续构建金融衍生品价格影响因素事理图谱做好准备,其中,规则模板的格式为<Pattern, Constraint, Priority>, Pattern 是包含因果触发词的正则表达式, Constraint 包含了句子中的句法约束, Priority 是当多个模板都匹配上时匹配模板的优先级;

其中, 传导分析模块进一步配置为:

有监督学习单元,对有标签的文本进行有监督学习;

无监督学习单元,使用无监督学习框架对无标签的文本进行一致性训练;

损失函数训练单元,针对无标签数据和有标签数据使用半监督学习学习框架 进行训练,得到损失函数,作为模型的目标优化函数用于训练,以使模型学习到 无标签数据和有标签数据中的信息;

其中,在传导分析模块的有监督学习单元中加入数据增强方式。

- 5、根据权利要求 4 所述的金融产品价格影响因素传导分析系统, 其特征在于, 预处理模块中的原始数据包括新闻、研报在内的舆情数据, 是采用分布式爬虫系 10 统自动从互联网上抓取; 预处理包括对原始数据进行去重、分词、和词性标注。
 - 6、根据权利要求 4 或 5 所述的金融产品价格影响因素传导分析系统,其特征在于,系统还包括图形化模块,配置为将传导分析的结果以事件为节点,以事件之间的关系为边,存入图数据库中,以实现可视化效果。

15

5

权 利 要 求 书

- 1、一种金融产品价格影响因素传导分析方法,其特征在于,方法包括:
- 步骤 1: 对自动获取的原始数据进行预处理;
- 5 步骤 2: 基于预处理后的数据,抽取事件因果对,基于事件因果对构建事理图谱:
 - 步骤 3: 基于语言预训练模型的构建分类算法对步骤 2 所构建的事理图谱中的因果事件对自动赋予标签,基于语言预训练模型自动赋予的标签用于价格影响因素的传导分析;
- 10 其中,步骤 2 进一步包括:
 - 步骤 2.1: 基于预处理后的数据,抽取出其中的因果触发词及所在句,作为原因、结果子句集合;
 - 步骤 2.2: 针对原因、结果子句集合,使用基于规则模板的正则匹配进行事件 抽取,抽取出显式因果语句中的原因子句和结果子句;
- 5 步骤 2.3: 针对显式因果语句中的原因子句和结果子句,基于依存句法分析确定事件表达,并构建出事理图谱;

其中,在步骤 2.2 中,基于规则模板的正则匹配方法抽取显式因果语句中的原因子句和结果子句,和因果触发词共同组成事件因果三元组,为后续构建金融衍生品价格影响因素事理图谱做好准备,其中,规则模板的格式为<Pattern, Constraint,

20 <u>Priority>, Pattern</u> 是包含因果触发词的正则表达式, Constraint 包含了句子中的句法约束, Priority 是当多个模板都匹配上时匹配模板的优先级;

其中,步骤3进一步包括:

步骤 3.1: 对有标签的文本进行有监督学习;

步骤 3.2: 使用无监督学习框架对无标签的文本进行一致性训练;

25 步骤 3.3: 针对无标签数据和有标签数据使用半监督学习学习框架进行训练, 得到损失函数,作为模型的目标优化函数用于训练,以使模型学习到无标签数据 和有标签数据中的信息;

其中, 在步骤 3 的有监督学习中加入数据增强方式。

2、根据权利要求 1 所述的金融产品价格影响因素传导分析方法,其特征在于,步骤 1 中的原始数据包括新闻、研报在内的舆情数据,是采用分布式爬虫系统自动从互联网上抓取;预处理包括对原始数据进行去重、分词、和词性标注。

5

3、根据权利要求 1 所述的金融产品价格影响因素传导分析方法, 其特征在于, 步骤 2 进一步包括:

步骤 2.1: 基于预处理后的数据,抽取出其中的因果触发词及所在句,作为原因、结果子句集合:

10 步骤 2.2: 针对原因、结果子句集合,使用基于规则模板的正则匹配进行事件 抽取, 抽取出显式因果语句中的原因子句和结果子句:

步骤 2.3: 针对显式因果语句中的原因子句和结果子句,基于依存句法分析确定事件表达,并构建出事理图谱。

15 4、根据权利要求 1 所述的金融产品价格影响因素传导分析方法, 其特征在于, 步骤 3 进一步包括:

步骤 3.1:对有标签的文本进行有监督学习:

步骤 3.2: 使用无监督学习框架对无标签的文本进行一致性训练:

步骤 3.3: 针对无标签数据和有标签数据使用半监督学习学习框架进行训练, 20 得到损失函数,作为模型的目标优化函数用于训练,以使模型学习到无标签数据 和有标签数据中的信息。

5、根据权利有安全 4 所述的金融产品价格影响因素传导分析方法, 其特征在 于, 在步骤 3 的有监督学习中加入数据增强方式。

25

63、根据权利要求 1 至 5 中任一项或 2 所述的金融产品价格影响因素传导分析方法,其特征在于,传导分析的结果以事件为节点,以事件之间的关系为边,存入图数据库中以实现可视化效果。

74、一种金融产品价格影响因素传导分析系统,其特征在于,系统包括: 预处理模块,对自动获取的原始数据进行预处理;

事理图谱构建模块,基于预处理后的数据,抽取事件因果对,基于事件因果 对构建事理图谱:

5 传导分析模块,基于语言预训练模型的构建分类算法对事理图谱构建模块所构建的事理图谱中的因果事件对自动赋予标签,基于语言预训练模型自动赋予的标签用于价格影响因素的传导分析;

其中, 事理图谱构建模块进一步配置为:

抽取单元,基于预处理后的数据,抽取出其中的因果触发词及所在句,作为 10 原因、结果子句集合;

匹配单元,针对原因、结果子句集合,使用基于规则模板的正则匹配进行事件抽取,抽取出显式因果语句中的原因子句和结果子句;

构建单元,针对显式因果语句中的原因子句和结果子句,基于依存句法分析确定事件表达,并构建出事理图谱;

- 其中,在匹配单元中,基于规则模板的正则匹配方法抽取显式因果语句中的原因子句和结果子句,和因果触发词共同组成事件因果三元组,为后续构建金融衍生品价格影响因素事理图谱做好准备,其中,规则模板的格式为<Pattern, Constraint, Priority>, Pattern 是包含因果触发词的正则表达式, Constraint 包含了句子中的句法约束, Priority 是当多个模板都匹配上时匹配模板的优先级;
- 20 其中,传导分析模块进一步配置为:

25

有监督学习单元,对有标签的文本进行有监督学习;

无监督学习单元,使用无监督学习框架对无标签的文本进行一致性训练;

损失函数训练单元,针对无标签数据和有标签数据使用半监督学习学习框架 进行训练,得到损失函数,作为模型的目标优化函数用于训练,以使模型学习到 无标签数据和有标签数据中的信息:

其中,在传导分析模块的有监督学习单元中加入数据增强方式。

85、根据权利要求 **74** 所述的金融产品价格影响因素传导分析系统,其特征在于,预处理模块中的原始数据包括新闻、研报在内的舆情数据,是采用分布式爬

虫系统自动从互联网上抓取; 预处理包括对原始数据进行去重、分词、和词性标注。

9、根据权利要求 7 所述的金融产品价格影响因素传导分析系统, 其特征在于, 事理图谱构建模块进一步配置为:

抽取单元,基于预处理后的数据,抽取出其中的因果触发词及所在句,作为原因、结果子句集合:

匹配单元,针对原因、结果子句集合,使用基于规则模板的正则匹配进行事件抽取,抽取出显式因果语句中的原因子句和结果子句:

10 构建单元,针对显式因果语句中的原因子句和结果子句,基于依存句法分析确定事件表达,并构建出事理图谱。

10、根据权利要求 7 所述的金融产品价格影响因素传导分析系统,其特征在于, 传导分析模块进一步配置为:

15 有监督学习单元,对有标签的文本进行有监督学习;

无监督学习单元,使用无监督学习框架对无标签的文本进行一致性训练;

损失函数训练单元,针对无标签数据和有标签数据使用半监督学习学习框架 进行训练,得到损失函数,作为模型的目标优化函数用于训练,以使模型学习到 无标签数据和有标签数据中的信息。

20

5

11、根据权利要求 10 所述的金融产品价格影响因素传导分析系统,其特征在 于,在传导分析模块的有监督学习单元中加入数据增强方式。

126、根据权利要求 7至11中任一项4或5 所述的金融产品价格影响因素传 25 导分析系统,其特征在于,系统还包括图形化模块,配置为将传导分析的结果以 事件为节点,以事件之间的关系为边,存入图数据库中,以实现可视化效果。