

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(10) 国际公布号

WO 2023/065545 A1

(43) 国际公布日
2023 年 4 月 27 日 (27.04.2023)

- (51) 国际专利分类号:
G06Q 10/04 (2012.01)
- (21) 国际申请号: PCT/CN2022/071241
- (22) 国际申请日: 2022 年 1 月 11 日 (11.01.2022)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
202111216347.6 2021 年 10 月 19 日 (19.10.2021) CN
- (71) 申请人: 平安科技(深圳)有限公司(PING AN TECHNOLOGY (SHENZHEN) CO.,LTD.) [CN/CN];
中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(72) 发明人: 肖京(XIAO, Jing); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。 李娜(LI, Na); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。 王磊(WANG, Lei); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。 赵盟盟(ZHAO, MengMeng); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。 王媛(WANG, Yuan); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。 谭韬(TANG, Tao); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(54) Title: RISK PREDICTION METHOD AND APPARATUS, AND DEVICE AND STORAGE MEDIUM

(54) 发明名称: 风险预测方法、装置、设备及存储介质

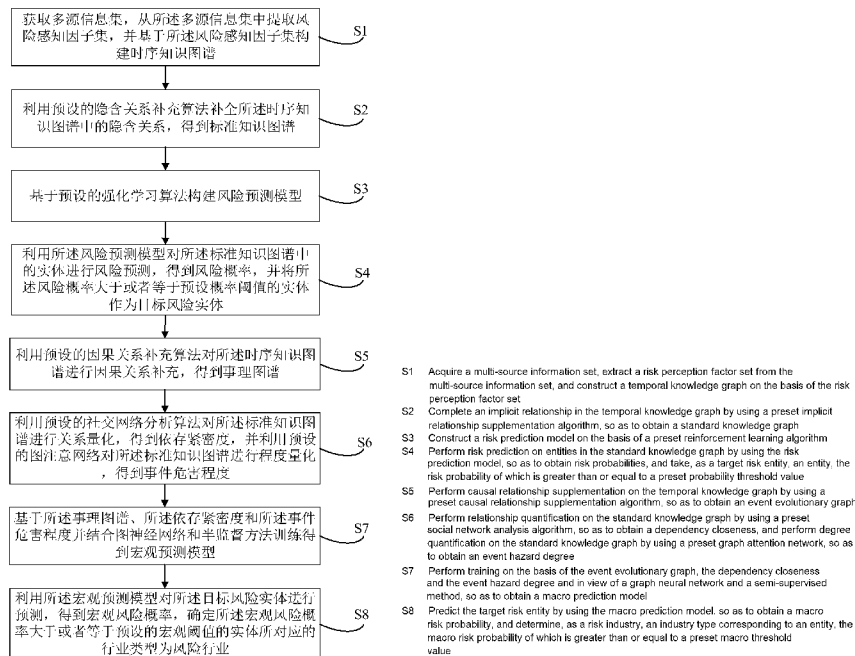


图 1

(57) Abstract: The present application relates to artificial intelligence technology. Disclosed is a risk prediction method, comprising: constructing a temporal knowledge graph on the basis of a risk perception factor set that is extracted from a multi-source information set, and performing implicit relationship supplementation and causal relationship supplementation on the temporal knowledge graph, so as to obtain a standard knowledge graph and an event evolutionary graph; performing prediction by using a risk prediction model that is constructed on the basis of a reinforcement learning algorithm, so as to obtain a target risk entity; and performing relationship

(CN)。 陈又新(CHEN, YouXin); 中国广东省深圳市福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳市沃德知识产权代理事务所(普通合伙)(SHENZHEN WORLD INTELLECTUAL PROPERTY AGENCY (GENERAL PARTNERSHIP)); 中国广东省深圳市福田区梅林街道中康路128号卓越城1期2栋301B于志光, Guangdong 518000 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

quantification and degree quantification on the standard knowledge graph, so as to obtain a dependency closeness and an event hazard degree, performing training on the basis of the event evolutionary graph, the dependency closeness and the event hazard degree and in view of a graph neural network and a semi-supervised method, so as to obtain a macro prediction model, and performing prediction by using the macro prediction model, so as to obtain a risk industry corresponding to the target risk entity. In addition, the present application further relates to blockchain technology. The event evolutionary graph can be stored in a node of a blockchain. Also provided in the present application are a risk prediction apparatus, an electronic device and a storage medium. By means of the present application, the accuracy of performing risk prediction on an industry can be improved.

(57) 摘要: 本申请涉及人工智能技术, 揭露一种风险预测方法, 包括: 基于从多源信息集中提取的风险感知因子集构建时序知识图谱, 对时序知识图谱进行隐含关系补充和因果关系补充, 得到标准知识图谱和事理图谱; 利用强化学习算法构建的风险预测模型预测得到目标风险实体; 对标准知识图谱进行关系量化和程度量化, 得到依存紧密度和事件危害程度, 基于事理图谱、依存紧密度和事件危害程度并结合图神经网络和半监督方法训练得到宏观预测模型, 利用宏观预测模型预测出目标风险实体对应的风险行业。此外, 本申请还涉及区块链技术, 事理图谱可存储于区块链的节点。本申请还提出一种风险预测装置、电子设备以及存储介质。本申请可以提高对行业进行风险预测的准确度。

风险预测方法、装置、设备及存储介质

本申请要求于 2021 年 10 月 19 日提交中国专利局、申请号为 CN202111216347.6、名称为“风险预测方法、装置、设备及存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及人工智能技术领域，尤其涉及一种风险预测方法、装置、电子设备及计算机可读存储介质。

背景技术

随着科技的发展和社会的进步，为了保证经济的稳定发展，提前进行关于行业的金融风险的预测是十分有必要的。行业的金融风险可能危及整个金融体系稳定，对实体经济造成严重负面效应的风险。

现有的风险预测方法主要基于较为简单的企业之间的网络结构，网络结构的生成要依赖于企业关系，如信贷关系。发明人意识到，企业间的关系复杂多样，例如合作研发关系、竞争关系、供应链关系等难以观测，因此会导致网络结构中缺失多种无法观测的隐含关系，进而导致对行业进行风险预测的准确度较低。

发明内容

本申请提供一种风险预测方法，包括：

获取多源信息集，从所述多源信息集中提取风险感知因子集，并基于所述风险感知因子集构建时序知识图谱；

利用预设的隐含关系补充算法补全所述时序知识图谱中的隐含关系，得到标准知识图谱；

基于预设的强化学习算法构建风险预测模型；

利用所述风险预测模型对所述标准知识图谱中的实体进行风险预测，得到风险概率，并将所述风险概率大于或者等于预设概率阈值的实体作为目标风险实体；

利用预设的因果关系补充算法对所述时序知识图谱进行因果关系补充，得到事理图谱；

利用预设的社交网络分析算法对所述标准知识图谱进行关系量化，得到依存紧密度，并利用预设的图注意网络对所述标准知识图谱进行程度量化，得到事件危害程度；

基于所述事理图谱、所述依存紧密度和所述事件危害程度并结合图神经网络和半监督方法训练得到宏观预测模型；

利用所述宏观预测模型对所述目标风险实体进行预测，得到宏观风险概率，确定所述宏观风险概率大于或者等于预设的宏观阈值的实体所对应的行业类型为风险行业。

本申请还提供一种风险预测装置，所述装置包括：

知识图谱构建模块，用于获取多源信息集，从所述多源信息集中提取风险感知因子集，并基于所述风险感知因子集构建时序知识图谱，利用预设的隐含关系补充算法补全所述时序知识图谱中的隐含关系，得到标准知识图谱；

目标风险实体预测模块，用于基于预设的强化学习算法构建风险预测模型，利用所述风险预测模型对所述标准知识图谱中的实体进行风险预测，得到风险概率，并将所述风险概率大于或者等于预设概率阈值的实体作为目标风险实体；

事理图谱生成模块，用于利用预设的因果关系补充算法对所述时序知识图谱进行因果关系补充，得到事理图谱；

图谱量化模块，用于利用预设的社交网络分析算法对所述标准知识图谱进行关系量化，得到依存紧密度，并利用预设的图注意网络对所述标准知识图谱进行程度量化，得到事件

危害程度；

宏观预测模块，用于基于所述事理图谱、所述依存紧密度和所述事件危害程度并结合图神经网络和半监督方法训练得到宏观预测模型，利用所述宏观预测模型对所述目标风险实体进行预测，得到宏观风险概率，确定所述宏观风险概率大于或者等于预设的宏观阈值的实体所对应的行业类型为风险行业。

本申请还提供一种电子设备，所述电子设备包括：

至少一个处理器；以及，

与所述至少一个处理器通信连接的存储器；其中，

所述存储器存储有可被所述至少一个处理器执行的计算机程序，所述计算机程序被所述至少一个处理器执行，以使所述至少一个处理器能够执行如下步骤：

获取多源信息集，从所述多源信息集中提取风险感知因子集，并基于所述风险感知因子集构建时序知识图谱；

利用预设的隐含关系补充算法补全所述时序知识图谱中的隐含关系，得到标准知识图谱；

基于预设的强化学习算法构建风险预测模型；

利用所述风险预测模型对所述标准知识图谱中的实体进行风险预测，得到风险概率，并将所述风险概率大于或者等于预设概率阈值的实体作为目标风险实体；

利用预设的因果关系补充算法对所述时序知识图谱进行因果关系补充，得到事理图谱；

利用预设的社交网络分析算法对所述标准知识图谱进行关系量化，得到依存紧密度，并利用预设的图注意网络对所述标准知识图谱进行程度量化，得到事件危害程度；

基于所述事理图谱、所述依存紧密度和所述事件危害程度并结合图神经网络和半监督方法训练得到宏观预测模型；

利用所述宏观预测模型对所述目标风险实体进行预测，得到宏观风险概率，确定所述宏观风险概率大于或者等于预设的宏观阈值的实体所对应的行业类型为风险行业。

本申请还提供一种计算机可读存储介质，所述计算机可读存储介质中存储有至少一个计算机程序，所述至少一个计算机程序被电子设备中的处理器执行以实现如下步骤：

获取多源信息集，从所述多源信息集中提取风险感知因子集，并基于所述风险感知因子集构建时序知识图谱；

利用预设的隐含关系补充算法补全所述时序知识图谱中的隐含关系，得到标准知识图谱；

基于预设的强化学习算法构建风险预测模型；

利用所述风险预测模型对所述标准知识图谱中的实体进行风险预测，得到风险概率，并将所述风险概率大于或者等于预设概率阈值的实体作为目标风险实体；

利用预设的因果关系补充算法对所述时序知识图谱进行因果关系补充，得到事理图谱；

利用预设的社交网络分析算法对所述标准知识图谱进行关系量化，得到依存紧密度，并利用预设的图注意网络对所述标准知识图谱进行程度量化，得到事件危害程度；

基于所述事理图谱、所述依存紧密度和所述事件危害程度并结合图神经网络和半监督方法训练得到宏观预测模型；

利用所述宏观预测模型对所述目标风险实体进行预测，得到宏观风险概率，确定所述宏观风险概率大于或者等于预设的宏观阈值的实体所对应的行业类型为风险行业。

附图说明

图1为本申请一实施例提供的风险预测方法的流程示意图；

图2为本申请一实施例提供的风险预测装置的功能模块图；

图3为本申请一实施例提供的实现所述风险预测方法的电子设备的结构示意图。

本申请目的的实现、功能特点及优点将结合实施例，参照附图做进一步说明。

具体实施方式

应当理解，此处所描述的具体实施例仅仅用以解释本申请，并不用于限定本申请。

本申请实施例提供一种风险预测方法。所述风险预测方法的执行主体包括但不限于服务端、终端等能够被配置为执行本申请实施例提供的该方法的电子设备中的至少一种。换言之，所述风险预测方法可以由安装在终端设备或服务端设备的软件或硬件来执行，所述软件可以是区块链平台。所述服务端包括但不限于：单台服务器、服务器集群、云端服务器或云端服务器集群等。所述服务器可以是独立的服务器，也可以是提供云服务、云数据库、云计算、云函数、云存储、网络服务、云通信、中间件服务、域名服务、安全服务、内容分发网络(Content Delivery Network, CDN)、以及大数据和人工智能平台等基础云计算服务的云服务器。

参照图 1 所示，为本申请一实施例提供的风险预测方法的流程示意图。在本实施例中，所述风险预测方法包括：

S1、获取多源信息集，从所述多源信息集中提取风险感知因子集，并基于所述风险感知因子集构建时序知识图谱。

本申请实施例中，可以从多源异构信息源中获取多源信息集，其中，所述多源异构信息源来自于互联网公开数据、业务数据和政府数据等多个渠道，包括但不限于视频、电子图片、遥感影像、文本等多种类型。

具体地，所述从所述多源信息集中提取风险感知因子集，包括：

识别所述多源信息集中的文本信息和图像信息；

利用预设的自然语言处理技术对所述文本信息进行因子提取，得到文本感知因子集；

利用预设的图像识别技术对所述图像信息进行因子提取，得到图像感知因子集；

将所述文本感知因子集和所述图像感知因子集进行汇总，得到风险感知因子集。

详细地，所述多源信息集中的信息包含不同的类型，筛选出所述多源信息集中的文本信息和图像信息，利用预设的自然语言处理技术对所述文本信息进行因子提取，其中，所述自然语言处理技术可以为利用 Word2Vec 算法进行词嵌入，利用 FastText 模型或者 ELMO (Embedding from language models, 双向语言模型) 模型进行文本特征编码，利用 BERT (Bidirectional Encoder Representations from Transformer, 双向编码器表征) 模型进行文本特征提取。利用预设的图像识别技术对所述图像信息进行因子提取，其中，所述图像识别技术可以为 YOLO V2~V4 或者 SSD 等目标检测算法、AlexNet 或者 ResNet 等目标识别算法、弱监督的语义分割方法等语义分割算法。

在本方案中，对于图片信息而言，例如可以利用卫星图像识别技术、光学字符识别 (OCR) 并结合 NLP 技术等技术完成信息提取。例如可以从超高分辨率卫星图像中识别农作物、航运货物、海陆运输等目标，进而对经济生产重要环节走势变化做出预警，可以使用 OCR 技术从财务票据、交易票据等非标准信息中提取用于风险审核的重要信息，而夜间灯光遥感数据则可用于动态预测人口密度、城市扩张速度。对于文本信息内容，可以利用自然语言处理 (NLP) 结合机器学习等技术完成信息提取。如可以从新闻、舆情、论坛资讯类文本数据中实时识别金融实体、发现金融事件的关联关系，提取刻画经济不确定性等的相关因子；从上市公司年报、IPO 招股说明书和公司前瞻性陈述类文本数据，可挖掘企业收入、业务发展规模、公司发展战略倾向等信息；也可从社交媒体类文本信息中，包括推特、微博、微信公众号和论坛帖子等，提取事件倾向评分、关注度指数、风险波动率等因子。

进一步地，所述基于所述风险感知因子集构建时序知识图谱，包括：

抽取所述风险感知因子集中的实体和实体关系；

基于所述实体和所述实体关系进行图谱构建，得到时序知识图谱。

其中，利用 Word2Vec 或者 LSTM+CRF 进行实体抽取，进行实体关系抽取的算法包括但不限于以下：监督学习(SVM,NN,朴素贝叶斯)、半监督学习(远程监督, Bootstrapping)、深度学习(Pipeline 如 Att-CNN&Att-BLSTM, Joint Model 如 LSTM-RNNs)。利用基于规则的推理 SWRL 或者基于图的推理 Path Ranking 进行图谱构建，得到时序知识图谱。

例如，以 PPP (Public-Private Partnership, 政府和社会资本合作) 合作模式为诱因引发的“东方园林”2018 年发债事件为例，所述实体可以为东方园林、10 亿公司债等，所述实体关系可以为宣布、发行等。

S2、利用预设的隐含关系补充算法补全所述时序知识图谱中的隐含关系，得到标准知识图谱。

本申请实施例中，通过补全所述时序知识图谱中的隐含关系可以丰富和校准所述时序知识图谱中各个实体之间的关系，扩宽所述时序知识图谱的应用范围。

其中，所述隐含关系是指实体之间难以直接获得的实体关系，所述时序知识图谱中包含多个实体和实体之间的关系，但时序知识图谱中的实体之间的关系为较为明显的观测得到的关系，因此需要对各个实体进行深度挖掘，得到隐含关系，并将所述隐含关系补充至所述时序知识图谱中，得到标准知识图谱。

具体地，所述利用预设的隐含关系补充算法补全所述时序知识图谱中的隐含关系，得到标准知识图谱，包括：

基于预设的稀疏图卷积网络对所述时序知识图谱进行图稀疏化处理，得到稀疏知识图谱；

利用训练好的关系图卷积网络对所述稀疏知识图谱进行关系预测，得到隐含关系；

将所述隐含关系补全在所述时序知识图谱中，得到标准知识图谱。

详细地，本方案采用图卷积网络模型对所述时序知识图谱进行隐含关系的补全，主要在于先对所述时序知识图谱进行图稀疏化处理，并将经过图稀疏化后得到的稀疏知识图谱作为关系图卷积网络的输入进行关系预测。这种方法可以避免在观测的实体关系样本不足、实体关系类型多样的情况下出现的由于特征稀疏而造成模型预测精度较低的问题，同时可以解决随着图形的大小增长，大型图上的节点分类和关系预测等功能也可能占用大量的时间和空间的问题。

其中，图卷积网络模型会将输入的知识图谱转换为稀疏知识图谱并进行下一步的关系预测。

进一步地，所述基于预设的稀疏图卷积网络对所述时序知识图谱进行图稀疏化处理，得到稀疏知识图谱，包括：

确定所述时序知识图谱对应的邻接矩阵及特征矩阵，并获取所述稀疏图卷积网络的预设权重矩阵；

基于所述邻接矩阵、所述特征矩阵和所述预设权重矩阵构建得到稀疏输出函数；

利用基于乘数的交替方向算法对所述稀疏输出函数进行优化处理，将优化后的所述稀疏输出函数中的变量对所述时序知识图谱进行变量更新，得到稀疏知识图谱。

详细地，若所述时序知识图谱为 $G = (V, E)$ ，其中， $V = \{v_1, \dots, v_n\}$ 为所述时序知识图谱中实体的集合， $E = \{e_1, \dots, e_m\}$ 为所述时序知识图谱中实体与实体之间关系的集合，而所述时序知识图谱对应的邻接矩阵即为存放实体与实体之间关系的二维数组。所述特征矩阵为每个实体对应的节点的特征，表示为 $X(v) = [x_1, \dots, x_k]$ ，当所述稀疏图卷积网络中包含两层子网络时，所述稀疏图卷积网络的预设权重矩阵可以为 $W^{(0)}$ 和 $W^{(1)}$ 。

其中，对于所述时序知识图谱为 $G = (V, E)$ 来说，其对应的邻接矩阵为：

$$A_{ij} = \begin{cases} 1, & (i, j) \in E, i, j \in V \\ 0, & otherwise. \end{cases}$$

具体地，所述基于所述邻接矩阵、所述特征矩阵和所述预设权重矩阵构建得到稀疏输出函数，包括：

$$\begin{aligned} Z(\hat{A}, W) &= \text{softmax}(\hat{A} \text{ReLU}(\hat{A} X W^{(0)}) W^{(1)}) \\ \hat{A} &= \tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} \\ \tilde{D} &= \text{diag} \left(\sum_j \tilde{A}_{ij} \right) \\ \tilde{A} &= A + I_N \end{aligned}$$

其中， $Z(\hat{A}, W)$ 为所述稀疏输出函数， W 为所述预设权重矩阵， $\hat{A}$ 为更新邻接矩阵， A 为所述邻接矩阵， I_N 为固定参数， diag 为对角矩阵， $W^{(0)}$ 和 $W^{(1)}$ 为所述稀疏图卷积网络的预设权重矩阵， ReLU 为线性整流函数， X 为所述特征矩阵。

进一步地，所述利用基于乘数的交替方向算法对所述稀疏输出函数进行优化处理，包括：

对所述稀疏输出函数进行简化处理，得到简化输出函数；

利用自适应矩估计优化器对所述简化输出函数中的变量的梯度进行更新，得到更新后的变量的梯度。

详细地，对所述稀疏输出函数进行简化处理是指所述稀疏图卷积网络的输出依赖于 $\hat{A}$ 和 W ，而但是 $\hat{A}$ 可以被表示为关于 A 的函数，则输出可以被表示为关于 A 和 W 的函数，即 $Z(A, W)$ ，因为 W 保持不变， $Z(A, W)$ 可被简化为 $Z(A)$ 。

其中，所述简化输出函数为：

$$Z(A) = \text{diag} \left(\sum_j (A + I)_{ij} \right)^{-\frac{1}{2}} (A + I) \text{diag} \left(\sum_j (A + I)_{ij} \right)^{-\frac{1}{2}} XW$$

具体地，利用自适应矩估计优化器对所述简化输出函数中的变量的梯度进行更新，将优化后的所述稀疏输出函数中的变量对所述时序知识图谱进行变量更新，得到稀疏知识图谱。

详细地，利用基于乘数的交替方向算法（ADMM）进行优化处理可以保留所述时序知识图谱中的结构和层次信息的网络主干，并在保持节点分类预测的性能的同时又保留的时序知识图谱的边缘信息。

进一步地，所述利用训练好的关系图卷积网络对所述稀疏知识图谱进行关系预测，得到隐含关系之前，所述方法还包括：

利用预设的关系图卷积网络中的实体编码器对所述稀疏知识图谱进行特征预测，得到所述稀疏知识图谱中的实体对应的潜在特征；

基于所述关系图卷积网络中的解码器对所述实体对应的潜在特征进行评分，将所述评分大于或者等于预设的评分阈值的对应潜在特征作为目标潜在特征；

根据所述目标潜在特征和预设的交叉熵损失函数计算得到交叉熵损失值；

当所述交叉熵损失值小于或者等于预设的损失阈值时，将所述关系图卷积网络输出为训练好的关系图卷积网络。

其中，所述关系图卷积网络中包含实体编码器和解码器，所述实体编码器用于产生实体的潜在特征表示，所述解码器用于通过一个评分函数对所述潜在特征表示进行评分处理。

具体地，利用预设的关系图卷积网络中的实体编码器对所述稀疏知识图谱进行特征预测即利用 R-GCN 作为编码器生成每个实体的实向量表示 e_i 。其中，R-GCN 模型按照预设的方式堆叠 L 层，R-GCN 模型中上一层的输出会作为下一层的输入，所述实体编码器使用 R-GCN 的输出作为每个实体的向量表示，即 $e_i = h_i^{(L)} \in \mathbb{R}^{d^{(L)}}$ 。 $h_i^{(l)} \in \mathbb{R}^{d^{(l)}}$ 是节点 v_i 在第 l 层

神经网络中的隐向量 (hidden state), $d^{(l)}$ 是该层向量表示的维度。

进一步地, 所述基于所述关系图卷积网络中的解码器对所述实体对应的潜在特征进行评分, 包括:

利用所述解码器中的 DistMult 因子分解模型对所述实体对应的潜在特征进行评分。

详细地, DistMult 因子分解模型是语义匹配模型的一种, 语义匹配模型利用基于相似性的评分函数, 通过匹配实体和关系在隐空间的潜在语义衡量此三元组存在或成立的可能性。

具体地, 利用所述解码器中的 DistMult 因子分解模型对所述实体对应的潜在特征进行评分, 包括:

$$f(s, r, o) = e_s^T R_r e_o,$$

其中, $e_s = h_s^{(L)} \in \mathbb{R}^{d^{(L)}}$ 是头实体 s 的隐性向量表示, $e_o = h_o^{(L)} \in \mathbb{R}^{d^{(L)}}$ 是尾实体 o 的隐向量表示, $R_r \in \mathbb{R}^{d \times d}$ 是对于关系类型 r 的邻接矩阵, d 是实体向量的维度。

进一步地, 所述根据所述目标潜在特征和预设的交叉熵损失函数计算得到交叉熵损失值, 包括:

所述交叉熵损失函数为:

$$\mathcal{L} = -\frac{1}{(1 + \omega)|\hat{\mathcal{E}}|} \sum_{(s, r, o, y) \in \mathcal{T}} y \log l(f(s, r, o)) + (1 - y) \log (1 - l(f(s, r, o)))$$

其中, $\mathcal{L}$ 为所述交叉熵损失值, $\mathcal{T}$ 是所有正负三元组样本的集合, 对于集合中的每个元素 (s, r, o, y) , $s, o \in \mathcal{V}$ 分别为头实体和尾实体, $r \in \mathcal{R}$ 为关系类型, y 则是一个指示器, $y = 1$ 表示正样本, $y = 0$ 表示负样本。

具体地, 利用训练好的关系图卷积网络对所述稀疏知识图谱进行关系预测, 得到隐含关系, 并将所述隐含关系补全在所述时序知识图谱中, 得到标准知识图谱。

例如, 对所述稀疏知识图谱进行关系预测, 得到隐含关系为“用 PPP 模式引发东方园林事件”, 则将所述隐含关系“用 PPP 模式引发东方园林事件”补充在所述时序知识图谱中, 得到标准知识图谱。

S3、基于预设的强化学习算法构建风险预测模型。

本申请实施例, 预设的强化学习算法是一个可以应用于顺序决策和控制任务的框架, 其中, 强化学习算法中的代理 (Agent) 通过与环境 (Environment) 交互来优化其行为。

具体地, 所述基于预设的强化学习算法构建风险预测模型, 包括:

获取原始风险状态数据, 对所述原始风险状态数据进行抽样处理, 得到预训练数据;

利用预设的深度神经网络对所述预训练数据进行拟合处理, 得到所述预训练数据对应的状态动作;

获取执行所述状态动作下的初始风险状态数据, 并计算所述初始风险状态数据和所述原始风险状态数据之间的时间差分;

以所述时间差分作为目标函数, 以所述强化学习算法为框架训练得到风险预测模型。

其中, 所述原始风险状态数据是指现在数据所属的风险情况的标识数据。可以抽取符合预设的抽样标准的数据得到预训练数据, 例如, 抽取风险情况满足高风险和中风险条件的数据作为预训练数据。

详细地, 所述强化学习算法为 Actor-Critic 算法, 在基于策略梯度的 Actor-Critic 方法中, 代理的角色被分为了参与者 (Actor) 和批评者 (Critic), 本质上 Actor 和 Critic 分别代表策略 (Policy) 和值函数 (Value function)。给定当前状态 x , 参与者仅负责生成动作 u 。批评者负责处理收到的奖励 r , 即通过调整值函数来评估当前策略的质量。在批评者执行了多个策略评估步骤后, 通过使用来自批评者的信息以更新参与者。

S4、利用所述风险预测模型对所述标准知识图谱中的实体进行风险预测, 得到风险概率, 并将所述风险概率大于或者等于预设概率阈值的实体作为目标风险实体。

本申请实施例中，可以利用所述风险预测模型对所述标准知识图谱中的多个不同的实体进行风险预测，将所述实体输入至所述风险预测模型中，得到所述实体对应的风险概率，并将所述风险概率大于或者等于预设概率阈值的实体作为目标风险实体。

其中，所述标准知识图谱中的实体包含不同类型的企业。

例如，在本方案中，即根据所述风险预测模型预测出所述标准知识图谱中风险概率大于或者等于所述预设概率阈值的企业作为目标风险企业。

S5、利用预设的因果关系补充算法对所述时序知识图谱进行因果关系补充，得到事理图谱。

本申请实施例中，所述因果关系补充是指将所述时序知识图谱中各个实体之间的因果顺序等关系进行补充。

具体地，所述利用预设的因果关系补充算法对所述时序知识图谱进行因果关系补充，得到事理图谱，包括：

获取训练文本集，对所述训练文本集进行事件抽取及因果关系归纳，得到多个因果三元组；

保留所述因果三元组中符合预设筛选标准的多个因果事件作为标准三元组；

对多个所述标准三元组进行事件融合，得到融合事件，并将所述融合事件补充至所述时序知识图谱中，得到事理图谱。

详细地，事件是指发生在某个特定时间点或者时间段、某个特定的地域范围，由一个或者多个角色参与的一个或者多个动作组成的事情或者状态的改变，因果关系是指两个事件之间一个事件导致或引起或造成了另一个事件，所述因果关系包含正向、负向、显性以及隐性关系，并且包含了其他例如排除转折、并列等可以帮助事件融合和推理的各种关系。所述因果关系归纳即将从所述训练文本集中抽取的具有因果关系的两个事件形成“原因事件-关系-结果事件”的三元组的数据形式的因果三元组。

具体地，可以通过预训练模型完成事件抽取及因果关系归纳，其中，所述预训练模型中的基础模型采用 BERT+CRF 的结构思想，实际上以 BERT 改进版本 RoBERTa 作为预训练模型。输入为词嵌入向量，通过段落嵌入、位置嵌入和标点嵌入后，再经过多个 transformer 模块，输出多个隐向量。然后通过 Seq2Seq 模型完成序列标签生成任务。此外，因为输入的每个字都会得到相应的标签（例如事件开头字、事件中间字、事件结尾字、其他字等），此模型同时能完成事件的标准化表达以及除去副词、助词等任务。该联合模型充分运用了预训练模型的语义信息，在事件抽取和归纳上都取得了较好的效果。

详细地，所述预设筛选标准为提前构建的规则或模板过滤不符合表达习惯或者表述不完整的事件节点，在得到标准化的事件表示和因果关系后，使用提前构建的规则或模板过滤不符合表达习惯或者表述不完整的事件节点。筛选后的不同事件节点可能因为其表义相同，指代的是现实世界中的同一实体，为此需要进行事件融合，得到融合事件。

进一步地，所述对多个所述标准三元组进行事件融合，得到融合事件，包括：

对多个所述标准三元组进行向量化，得到多个三元向量组；

计算多个所述三元向量组中事件之间的相似度；

若所述相似度大于预设的第一阈值时，将所述相似度对应的标准三元组划分至第一簇事件中；

若所述相似度小于预设的第二阈值时，将所述相似度对应的标准三元组划分至第二簇事件中；

若所述相似度小于所述第一阈值且大于所述第二阈值时，将所述相似度对应的标准三元组划分至缓冲簇事件中。

详细地，增量聚类算法可以得到实时的聚类结果，所以对部分事件样本进行聚类以后，可从聚类结果中抽取样本扩充训练集，对模型进行再训练，使模型能学习新的事件和文本

特征，增强聚类效果。最后，在上述步骤结束后，将聚类结果中所含样本较少的所有簇与 Buffer 结合作为不确定样本，利用多次训练后的模型对这些样本进行聚类，得到最终的聚类结果，从而完成此部分事件融合的过程。

S6、利用预设的社交网络分析算法对所述标准知识图谱进行关系量化，得到依存紧密度，并利用预设的图注意力网络对所述标准知识图谱进行程度量化，得到事件危害程度。

本申请实施例中，所述利用预设的社交网络分析算法对所述标准知识图谱进行关系量化，得到依存紧密度，包括：

根据度中心性和模块化分类度量的定义确定所述标准知识图谱的度中心性及模块化分类度量；

利用预设的接近中心性计算公式计算所述标准知识图谱的接近中心性，并将所述度中心性、所述模块化分类度量和所述接近中心性代入至预设的依存紧密度计算公式中，得到依存紧密度。

详细地，所述度中心性 (Degree centrality) 是指一个节点相连的边的数量，用于表示节点的连接程度，所述模块化分类度量 (Modularity class) 用于社区检测，用来衡量社区划分质量或稳定性，模块化分类度量等于组内的边数减去随机设置边的等效网络中的预期边数。所述接近中心性 (Betweenness centrality) 衡量了一个节点到达其他节点的难易程度。

具体地，所述利用预设的接近中心性计算公式计算所述标准知识图谱的接近中心性，包括：

$$C_B(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

其中， $C_B(v)$ 代表节点 v 的接近中心性值， $\sigma_{st}(v)$ 代表从节点 s 到节点 t 之间经过 v 的最短路径数之和， σ_{st} 代表从节点 s 到节点 t 之间的所有最短路径数之和， $v, s, t \in V$ 。

进一步地，所述预设的依存紧密度计算公式，包括：

$$T(v_i, v_j) = w_D \frac{C_D(v_i) + C_D(v_j)}{\sum_{k=1}^n C_D(v_k)} + w_B \frac{C_B(v_i) + C_B(v_j)}{\sum_{k=1}^n C_B(v_k)} + w_M \frac{C_M(v_k)}{\sum_{k=1}^n C_M(v_k)}$$

其中， $T(v_i, v_j)$ 是 v_i 和 v_j 之间的依存紧密度， $w_D, w_B, w_M \in (0, 1)$ 为各子指标的权重且 $w_D + w_B + w_M = 1$ 。

具体地，利用预设的图注意力网络对所述标准知识图谱进行程度量化，得到事件危害程度，所述图注意力网络 (GAT) 学习节点所有邻居上的注意力系数以进行特征聚合，并可以提高许多图学习任务的性能。利用图注意力网络对于风险危害程度进行打分，评分函数依赖于网络的注意力系数和相关的实体特征向量，然后以最小化均方差损失为目标训练模型，最后以矩阵形式输出每一对实体之间的风险危害程度分数。由此针对不同的知识图谱和实体特征，可以量化评估针对金融领域的实体关系，例如企业信贷关系、供应链关系、行业间投入产出等。

S7、基于所述事理图谱、所述依存紧密度和所述事件危害程度并结合图神经网络和半监督方法训练得到宏观预测模型。

本申请实施例中，所述 S7，包括：

将所述事理图谱、所述依存紧密度和所述事件危害程度汇总为标注数据，并以所述标注数据为风险标签构建有监督模型；

获取无标注数据，以所述无标注数据和所述标注数据为基础构建无监督时序模型；

利用预设的半监督贝叶斯算法将所述有监督模型和所述无监督时序模型结合为宏观预测模型。

S8、利用所述宏观预测模型对所述目标风险实体进行预测，得到宏观风险概率，确定所述宏观风险概率大于或者等于预设的宏观阈值的实体所对应的行业类型为风险行业。

本申请实施例中，所述宏观预测模型用于进行行业的风险预测，利用所述宏观预测

模型对所述目标风险实体进行预测，得到宏观风险概率，将所述宏观风险概率大于或者等于预设的宏观阈值的实体所对应的行业类型为风险行业。

例如，所述目标风险实体为建筑环保板块，利用所述宏观预测模型对所述目标风险实体进行预测，得到宏观风险概率为 0.6，若预设的宏观阈值为 0.5，则所述宏观风险概率大于所述宏观阈值，则将目标风险实体为建筑环保板块对应的建筑行业作为风险行业。

本申请实施例从预获取的多源信息集中提取风险感知因子集，并基于所述风险感知因子集构建时序知识图谱，所述时序知识图谱作为后续风险预测的数据基础，利用预设的隐含关系补充算法补全所述时序知识图谱中的隐含关系，得到标准知识图谱。基于预设的强化学习算法构建风险预测模型，利用所述强化学习算法进行模型的构建可以保证模型的稳定性，并利用所述风险预测模型对所述标准知识图谱中的实体进行风险预测，得到目标风险实体。对所述标准知识图谱分别进行关系量化和程度量化，得到依存紧密度和事件危害程度，结合进行因果补充得到的事理图谱训练得到宏观预测模型，并利用宏观预测模型对所述目标风险实体进行预测，得到对应的风险预测行业。所述风险预测模型可以实现从实体角度的风险预测，通过宏观预测模型对已经经过风险预测的目标风险实体进行行业角度的预测，提高了对行业进行风险预测的准确度。因此本申请提出的风险预测方法可以实现解决对行业进行风险预测的准确度不够高的问题。

如图 2 所示，是本申请一实施例提供的风险预测装置的功能模块图。

本申请所述风险预测装置 100 可以安装于电子设备中。根据实现的功能，所述风险预测装置 100 可以包括知识图谱构建模块 101、目标风险实体预测模块 102、事理图谱生成模块 103、图谱量化模块 104 及宏观预测模块 105。本申请所述模块也可以称之为单元，是指一种能够被电子设备处理器所执行，并且能够完成固定功能的一系列计算机程序段，其存储在电子设备的存储器中。

在本实施例中，关于各模块/单元的功能如下：

所述知识图谱构建模块 101，用于获取多源信息集，从所述多源信息集中提取风险感知因子集，并基于所述风险感知因子集构建时序知识图谱，利用预设的隐含关系补充算法补全所述时序知识图谱中的隐含关系，得到标准知识图谱；

所述目标风险实体预测模块 102，用于基于预设的强化学习算法构建风险预测模型，利用所述风险预测模型对所述标准知识图谱中的实体进行风险预测，得到风险概率，并将所述风险概率大于或者等于预设概率阈值的实体作为目标风险实体；

所述事理图谱生成模块 103，用于利用预设的因果关系补充算法对所述时序知识图谱进行因果关系补充，得到事理图谱；

所述图谱量化模块 104，用于利用预设的社交网络分析算法对所述标准知识图谱进行关系量化，得到依存紧密度，并利用预设的图注意网络对所述标准知识图谱进行程度量化，得到事件危害程度；

所述宏观预测模块 105，用于基于所述事理图谱、所述依存紧密度和所述事件危害程度并结合图神经网络和半监督方法训练得到宏观预测模型，利用所述宏观预测模型对所述目标风险实体进行预测，得到宏观风险概率，确定所述宏观风险概率大于或者等于预设的宏观阈值的实体所对应的行业类型为风险行业。

详细地，本申请实施例中所风险预测装置 100 中所述的各模块在使用时采用与上述图 1 所述的风险预测方法一样的技术手段，并能够产生相同的技术效果，这里不再赘述。

如图 3 所示，是本申请一实施例提供的实现风险预测方法的电子设备的结构示意图。

所述电子设备 1 可以包括处理器 10、存储器 11、通信总线 12 以及通信接口 13，还可以包括存储在所述存储器 11 中并可在所述处理器 10 上运行的计算机程序，如风险预测程序。

其中，所述处理器 10 在一些实施例中可以由集成电路组成，例如可以由单个封装的

集成电路所组成，也可以是由多个相同功能或不同功能封装的集成电路所组成，包括一个或者多个中央处理器（Central Processing unit，CPU）、微处理器、数字处理芯片、图形处理器及各种控制芯片的组合等。所述处理器 10 是所述电子设备的控制核心（Control Unit），利用各种接口和线路连接整个电子设备的各个部件，通过运行或执行存储在所述存储器 11 内的程序或者模块（例如执行风险预测程序等），以及调用存储在所述存储器 11 内的数据，以执行电子设备的各种功能和处理数据。

所述存储器 11 至少包括一种类型的可读存储介质，所述可读存储介质包括闪存、移动硬盘、多媒体卡、卡型存储器（例如：SD 或 DX 存储器等）、磁性存储器、磁盘、光盘等。所述存储器 11 在一些实施例中可以是电子设备的内部存储单元，例如该电子设备的移动硬盘。所述存储器 11 在另一些实施例中也可以是电子设备的外部存储设备，例如电子设备上配备的插接式移动硬盘、智能存储卡（Smart Media Card，SMC）、安全数字（Secure Digital，SD）卡、闪存卡（Flash Card）等。进一步地，所述存储器 11 还可以既包括电子设备的内部存储单元也包括外部存储设备。所述存储器 11 不仅可以用于存储安装于电子设备的应用软件及各类数据，例如风险预测程序的代码等，还可以用于暂时地存储已经输出或者将要输出的数据。

所述通信总线 12 可以是外设部件互连标准（peripheral component interconnect，简称 PCI）总线或扩展工业标准结构（extended industry standard architecture，简称 EISA）总线等。该总线可以分为地址总线、数据总线、控制总线等。所述总线被设置为实现所述存储器 11 以及至少一个处理器 10 等之间的连接通信。

所述通信接口 13 用于上述电子设备与其他设备之间的通信，包括网络接口和用户接口。可选地，所述网络接口可以包括有线接口和/或无线接口（如 Wi-Fi 接口、蓝牙接口等），通常用于在该电子设备与其他电子设备之间建立通信连接。所述用户接口可以是显示器（Display）、输入单元（比如键盘（Keyboard）），可选地，用户接口还可以是标准的有线接口、无线接口。可选地，在一些实施例中，显示器可以是 LED 显示器、液晶显示器、触控式液晶显示器以及 OLED（Organic Light-Emitting Diode，有机发光二极管）触摸器等。其中，显示器也可以适当的称为显示屏或显示单元，用于显示在电子设备中处理的信息以及用于显示可视化的用户界面。

图 3 仅示出了具有部件的电子设备，本领域技术人员可以理解的是，图 3 示出的结构并不构成对所述电子设备 1 的限定，可以包括比图示更少或者更多的部件，或者组合某些部件，或者不同的部件布置。

例如，尽管未示出，所述电子设备还可以包括给各个部件供电的电源（比如电池），优选地，电源可以通过电源管理装置与所述至少一个处理器 10 逻辑相连，从而通过电源管理装置实现充电管理、放电管理、以及功耗管理等功能。电源还可以包括一个或一个以上的直流或交流电源、再充电装置、电源故障检测电路、电源转换器或者逆变器、电源状态指示器等任意组件。所述电子设备还可以包括多种传感器、蓝牙模块、Wi-Fi 模块等，在此不再赘述。

应该了解，所述实施例仅为说明之用，在专利申请范围上并不受此结构的限制。

所述电子设备 1 中的所述存储器 11 存储的风险预测程序是多个指令的组合，在所述处理器 10 中运行时，可以实现：

获取多源信息集，从所述多源信息集中提取风险感知因子集，并基于所述风险感知因子集构建时序知识图谱；

利用预设的隐含关系补充算法补全所述时序知识图谱中的隐含关系，得到标准知识图谱；

基于预设的强化学习算法构建风险预测模型；

利用所述风险预测模型对所述标准知识图谱中的实体进行风险预测，得到风险概率，

并将所述风险概率大于或者等于预设概率阈值的实体作为目标风险实体；

利用预设的因果关系补充算法对所述时序知识图谱进行因果关系补充，得到事理图谱；

利用预设的社交网络分析算法对所述标准知识图谱进行关系量化，得到依存紧密度，并利用预设的图注意网络对所述标准知识图谱进行程度量化，得到事件危害程度；

基于所述事理图谱、所述依存紧密度和所述事件危害程度并结合图神经网络和半监督方法训练得到宏观预测模型；

利用所述宏观预测模型对所述目标风险实体进行预测，得到宏观风险概率，确定所述宏观风险概率大于或者等于预设的宏观阈值的实体所对应的行业类型为风险行业。

具体地，所述处理器 10 对上述指令的具体实现方法可参考附图对应实施例中相关步骤的描述，在此不赘述。

进一步地，所述电子设备 1 集成的模块/单元如果以软件功能单元的形式实现并作为独立的产品销售或使用，可以存储在一个计算机可读存储介质中。所述计算机可读存储介质可以是易失性的，也可以是非易失性的。例如，所述计算机可读存储介质可以包括：能够携带所述计算机程序代码的任何实体或装置、记录介质、U 盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器（ROM，Read-Only Memory）。

本申请还提供一种计算机可读存储介质，所述计算机可读存储介质存储有计算机程序，所述计算机程序在被电子设备的处理器所执行时，可以实现：

获取多源信息集，从所述多源信息集中提取风险感知因子集，并基于所述风险感知因子集构建时序知识图谱；

利用预设的隐含关系补充算法补全所述时序知识图谱中的隐含关系，得到标准知识图谱；

基于预设的强化学习算法构建风险预测模型；

利用所述风险预测模型对所述标准知识图谱中的实体进行风险预测，得到风险概率，并将所述风险概率大于或者等于预设概率阈值的实体作为目标风险实体；

利用预设的因果关系补充算法对所述时序知识图谱进行因果关系补充，得到事理图谱；

利用预设的社交网络分析算法对所述标准知识图谱进行关系量化，得到依存紧密度，并利用预设的图注意网络对所述标准知识图谱进行程度量化，得到事件危害程度；

基于所述事理图谱、所述依存紧密度和所述事件危害程度并结合图神经网络和半监督方法训练得到宏观预测模型；

利用所述宏观预测模型对所述目标风险实体进行预测，得到宏观风险概率，确定所述宏观风险概率大于或者等于预设的宏观阈值的实体所对应的行业类型为风险行业。

在本申请所提供的几个实施例中，应该理解到，所揭露的设备、装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述模块的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式。

所述作为分离部件说明的模块可以是或者也可以不是物理上分开的，作为模块显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。

另外，在本申请各个实施例中的各功能模块可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用硬件加软件功能模块的形式实现。

对于本领域技术人员而言，显然本申请不限于上述示范性实施例的细节，而且在不背离本申请的精神或基本特征的情况下，能够以其他的具体形式实现本申请。

因此，无论从哪一点来看，均应将实施例看作是示范性的，而且是非限制性的，本申请的范围由所附权利要求而不是上述说明限定，因此旨在将落在权利要求的等同要件的含义和范围内的所有变化涵括在本申请内。不应将权利要求中的任何附关联图标记视为限制

所涉及的权利要求。

本申请所指区块链是分布式数据存储、点对点传输、共识机制、加密算法等计算机技术的新型应用模式。区块链(Blockchain)，本质上是一个去中心化的数据库，是一串使用密码学方法相关联产生的数据块，每一个数据块中包含了一批次网络交易的信息，用于验证其信息的有效性(防伪)和生成下一个区块。区块链可以包括区块链底层平台、平台产品服务层以及应用服务层等。

本申请实施例可以基于人工智能技术对相关的数据进行获取和处理。其中，人工智能(Artificial Intelligence, AI)是利用数字计算机或者数字计算机控制的机器模拟、延伸和扩展人的智能，感知环境、获取知识并使用知识获得最佳结果的理论、方法、技术及应用系统。

此外，显然“包括”一词不排除其他单元或步骤，单数不排除复数。系统权利要求中陈述的多个单元或装置也可以由一个单元或装置通过软件或者硬件来实现。第一、第二等词语用来表示名称，而并不表示任何特定的顺序。

最后应说明的是，以上实施例仅用以说明本申请的技术方案而非限制，尽管参照较佳实施例对本申请进行了详细说明，本领域的普通技术人员应当理解，可以对本申请的技术方案进行修改或等同替换，而不脱离本申请技术方案的精神和范围。

权利要求书

- 1、一种风险预测方法，其中，所述方法包括：
获取多源信息集，从所述多源信息集中提取风险感知因子集，并基于所述风险感知因子集构建时序知识图谱；
利用预设的隐含关系补充算法补全所述时序知识图谱中的隐含关系，得到标准知识图谱；
基于预设的强化学习算法构建风险预测模型；
利用所述风险预测模型对所述标准知识图谱中的实体进行风险预测，得到风险概率，并将所述风险概率大于或者等于预设概率阈值的实体作为目标风险实体；
利用预设的因果关系补充算法对所述时序知识图谱进行因果关系补充，得到事理图谱；
利用预设的社交网络分析算法对所述标准知识图谱进行关系量化，得到依存紧密度，并利用预设的图注意网络对所述标准知识图谱进行程度量化，得到事件危害程度；
基于所述事理图谱、所述依存紧密度和所述事件危害程度，并结合图神经网络和半监督方法训练得到宏观预测模型；
利用所述宏观预测模型对所述目标风险实体进行预测，得到宏观风险概率，确定所述宏观风险概率大于或者等于预设的宏观阈值的实体所对应的行业类型为风险行业。
- 2、如权利要求 1 所述的风险预测方法，其中，所述利用预设的隐含关系补充算法补全所述时序知识图谱中的隐含关系，得到标准知识图谱，包括：
基于预设的稀疏图卷积网络对所述时序知识图谱进行图稀疏化处理，得到稀疏知识图谱；
利用训练好的关系图卷积网络对所述稀疏知识图谱进行关系预测，得到隐含关系；
将所述隐含关系补全在所述时序知识图谱中，得到标准知识图谱。
- 3、如权利要求 2 所述的风险预测方法，其中，所述基于预设的稀疏图卷积网络对所述时序知识图谱进行图稀疏化处理，得到稀疏知识图谱，包括：
确定所述时序知识图谱对应的邻接矩阵及特征矩阵，并获取所述稀疏图卷积网络的预设权重矩阵；
基于所述邻接矩阵、所述特征矩阵和所述预设权重矩阵构建得到稀疏输出函数；
利用基于乘数的交替方向算法对所述稀疏输出函数进行优化处理，将优化后的所述稀疏输出函数中的变量对所述时序知识图谱进行变量更新，得到稀疏知识图谱。
- 4、如权利要求 2 所述的风险预测方法，其中，所述利用训练好的关系图卷积网络对所述稀疏知识图谱进行关系预测，得到隐含关系之前，所述方法还包括：
利用预设的关系图卷积网络中的实体编码器对所述稀疏知识图谱进行特征预测，得到所述稀疏知识图谱中的实体对应的潜在特征；
基于所述关系图卷积网络中的解码器对所述实体对应的潜在特征进行评分，将所述评分大于或者等于预设的评分阈值的对应潜在特征作为目标潜在特征；
根据所述目标潜在特征和预设的交叉熵损失函数计算得到交叉熵损失值；
当所述交叉熵损失值小于或者等于预设的损失阈值时，将所述关系图卷积网络输出为训练好的关系图卷积网络。
- 5、如权利要求 1 所述的风险预测方法，其中，所述基于所述风险感知因子集构建时序知识图谱，包括：
抽取所述风险感知因子集中的实体和实体关系；
基于所述实体和所述实体关系进行图谱构建，得到时序知识图谱。
- 6、如权利要求 1 至 5 中任一项所述的风险预测方法，其中，所述利用预设的因果关系补充算法对所述时序知识图谱进行因果关系补充，得到事理图谱，包括：

获取训练文本集，对所述训练文本集进行事件抽取及因果关系归纳，得到多个因果三元组；

保留所述因果三元组中符合预设筛选标准的多个因果事件作为标准三元组；

对多个所述标准三元组进行事件融合，得到融合事件，并将所述融合事件补充至所述时序知识图谱中，得到事理图谱。

7、如权利要求 1 至 5 中任一项所述的风险预测方法，其中，所述从所述多源信息集中提取风险感知因子集，包括：

识别所述多源信息集中的文本信息和图像信息；

利用预设的自然语言处理技术对所述文本信息进行因子提取，得到文本感知因子集；

利用预设的图像识别技术对所述图像信息进行因子提取，得到图像感知因子集；

将所述文本感知因子集和所述图像感知因子集进行汇总，得到风险感知因子集。

8、一种风险预测装置，其中，所述装置包括：

知识图谱构建模块，用于获取多源信息集，从所述多源信息集中提取风险感知因子集，并基于所述风险感知因子集构建时序知识图谱，利用预设的隐含关系补充算法补全所述时序知识图谱中的隐含关系，得到标准知识图谱；

目标风险实体预测模块，用于基于预设的强化学习算法构建风险预测模型，利用所述风险预测模型对所述标准知识图谱中的实体进行风险预测，得到风险概率，并将所述风险概率大于或者等于预设概率阈值的实体作为目标风险实体；

事理图谱生成模块，用于利用预设的因果关系补充算法对所述时序知识图谱进行因果关系补充，得到事理图谱；

图谱量化模块，用于利用预设的社交网络分析算法对所述标准知识图谱进行关系量化，得到依存紧密度，并利用预设的图注意网络对所述标准知识图谱进行程度量化，得到事件危害程度；

宏观预测模块，用于基于所述事理图谱、所述依存紧密度和所述事件危害程度并结合图神经网络和半监督方法训练得到宏观预测模型，利用所述宏观预测模型对所述目标风险实体进行预测，得到宏观风险概率，确定所述宏观风险概率大于或者等于预设的宏观阈值的实体所对应的行业类型为风险行业。

9、一种电子设备，其中，所述电子设备包括：

至少一个处理器；以及，

与所述至少一个处理器通信连接的存储器；其中，

所述存储器存储有可被所述至少一个处理器执行的计算机程序，所述计算机程序被所述至少一个处理器执行，以使所述至少一个处理器能够执行如下步骤：

获取多源信息集，从所述多源信息集中提取风险感知因子集，并基于所述风险感知因子集构建时序知识图谱；

利用预设的隐含关系补充算法补全所述时序知识图谱中的隐含关系，得到标准知识图谱；

基于预设的强化学习算法构建风险预测模型；

利用所述风险预测模型对所述标准知识图谱中的实体进行风险预测，得到风险概率，并将所述风险概率大于或者等于预设概率阈值的实体作为目标风险实体；

利用预设的因果关系补充算法对所述时序知识图谱进行因果关系补充，得到事理图谱；

利用预设的社交网络分析算法对所述标准知识图谱进行关系量化，得到依存紧密度，并利用预设的图注意网络对所述标准知识图谱进行程度量化，得到事件危害程度；

基于所述事理图谱、所述依存紧密度和所述事件危害程度，并结合图神经网络和半监督方法训练得到宏观预测模型；

利用所述宏观预测模型对所述目标风险实体进行预测，得到宏观风险概率，确定所述

宏观风险概率大于或者等于预设的宏观阈值的实体所对应的行业类型为风险行业。

10、如权利要求 9 所述的电子设备，其中，所述利用预设的隐含关系补充算法补充所述时序知识图谱中的隐含关系，得到标准知识图谱，包括：

基于预设的稀疏图卷积网络对所述时序知识图谱进行图稀疏化处理，得到稀疏知识图谱；

利用训练好的关系图卷积网络对所述稀疏知识图谱进行关系预测，得到隐含关系；

将所述隐含关系补充在所述时序知识图谱中，得到标准知识图谱。

11、如权利要求 10 所述的电子设备，其中，所述基于预设的稀疏图卷积网络对所述时序知识图谱进行图稀疏化处理，得到稀疏知识图谱，包括：

确定所述时序知识图谱对应的邻接矩阵及特征矩阵，并获取所述稀疏图卷积网络的预设权重矩阵；

基于所述邻接矩阵、所述特征矩阵和所述预设权重矩阵构建得到稀疏输出函数；

利用基于乘数的交替方向算法对所述稀疏输出函数进行优化处理，将优化后的所述稀疏输出函数中的变量对所述时序知识图谱进行变量更新，得到稀疏知识图谱。

12、如权利要求 10 所述的电子设备，其中，所述利用训练好的关系图卷积网络对所述稀疏知识图谱进行关系预测，得到隐含关系之前，所述计算机程序被所述至少一个处理器执行时还实现如下步骤：

利用预设的关系图卷积网络中的实体编码器对所述稀疏知识图谱进行特征预测，得到所述稀疏知识图谱中的实体对应的潜在特征；

基于所述关系图卷积网络中的解码器对所述实体对应的潜在特征进行评分，将所述评分大于或者等于预设的评分阈值的对应潜在特征作为目标潜在特征；

根据所述目标潜在特征和预设的交叉熵损失函数计算得到交叉熵损失值；

当所述交叉熵损失值小于或者等于预设的损失阈值时，将所述关系图卷积网络输出为训练好的关系图卷积网络。

13、如权利要求 9 所述的电子设备，其中，所述基于所述风险感知因子集构建时序知识图谱，包括：

抽取所述风险感知因子集中的实体和实体关系；

基于所述实体和所述实体关系进行图谱构建，得到时序知识图谱。

14、如权利要求 9 至 13 中任一项所述的电子设备，其中，所述利用预设的因果关系补充算法对所述时序知识图谱进行因果关系补充，得到事理图谱，包括：

获取训练文本集，对所述训练文本集进行事件抽取及因果关系归纳，得到多个因果三元组；

保留所述因果三元组中符合预设筛选标准的多个因果事件作为标准三元组；

对多个所述标准三元组进行事件融合，得到融合事件，并将所述融合事件补充至所述时序知识图谱中，得到事理图谱。

15、如权利要求 9 至 14 中任一项所述的电子设备，其中，所述从所述多源信息集中提取风险感知因子集，包括：

识别所述多源信息集中的文本信息和图像信息；

利用预设的自然语言处理技术对所述文本信息进行因子提取，得到文本感知因子集；

利用预设的图像识别技术对所述图像信息进行因子提取，得到图像感知因子集；

将所述文本感知因子集和所述图像感知因子集进行汇总，得到风险感知因子集。

16、一种计算机可读存储介质，存储有计算机程序，其中，所述计算机程序被处理器执行时实现如下步骤：

获取多源信息集，从所述多源信息集中提取风险感知因子集，并基于所述风险感知因子集构建时序知识图谱；

利用预设的隐含关系补充算法补全所述时序知识图谱中的隐含关系，得到标准知识图谱；

基于预设的强化学习算法构建风险预测模型；

利用所述风险预测模型对所述标准知识图谱中的实体进行风险预测，得到风险概率，并将所述风险概率大于或者等于预设概率阈值的实体作为目标风险实体；

利用预设的因果关系补充算法对所述时序知识图谱进行因果关系补充，得到事理图谱；

利用预设的社交网络分析算法对所述标准知识图谱进行关系量化，得到依存紧密度，并利用预设的图注意网络对所述标准知识图谱进行程度量化，得到事件危害程度；

基于所述事理图谱、所述依存紧密度和所述事件危害程度，并结合图神经网络和半监督方法训练得到宏观预测模型；

利用所述宏观预测模型对所述目标风险实体进行预测，得到宏观风险概率，确定所述宏观风险概率大于或者等于预设的宏观阈值的实体所对应的行业类型为风险行业。

17、如权利要求 16 所述的计算机可读存储介质，其中，所述利用预设的隐含关系补充算法补全所述时序知识图谱中的隐含关系，得到标准知识图谱，包括：

基于预设的稀疏图卷积网络对所述时序知识图谱进行图稀疏化处理，得到稀疏知识图谱；

利用训练好的关系图卷积网络对所述稀疏知识图谱进行关系预测，得到隐含关系；

将所述隐含关系补全在所述时序知识图谱中，得到标准知识图谱。

18、如权利要求 17 所述的计算机可读存储介质，其中，所述基于预设的稀疏图卷积网络对所述时序知识图谱进行图稀疏化处理，得到稀疏知识图谱，包括：

确定所述时序知识图谱对应的邻接矩阵及特征矩阵，并获取所述稀疏图卷积网络的预设权重矩阵；

基于所述邻接矩阵、所述特征矩阵和所述预设权重矩阵构建得到稀疏输出函数；

利用基于乘数的交替方向算法对所述稀疏输出函数进行优化处理，将优化后的所述稀疏输出函数中的变量对所述时序知识图谱进行变量更新，得到稀疏知识图谱。

19、如权利要求 17 所述的计算机可读存储介质，其中，所述利用训练好的关系图卷积网络对所述稀疏知识图谱进行关系预测，得到隐含关系之前，所述计算机程序被处理器执行时还实现如下步骤：

利用预设的关系图卷积网络中的实体编码器对所述稀疏知识图谱进行特征预测，得到所述稀疏知识图谱中的实体对应的潜在特征；

基于所述关系图卷积网络中的解码器对所述实体对应的潜在特征进行评分，将所述评分大于或者等于预设的评分阈值的对应潜在特征作为目标潜在特征；

根据所述目标潜在特征和预设的交叉熵损失函数计算得到交叉熵损失值；

当所述交叉熵损失值小于或者等于预设的损失阈值时，将所述关系图卷积网络输出为训练好的关系图卷积网络。

20、如权利要求 16 所述的计算机可读存储介质，其中，所述基于所述风险感知因子集构建时序知识图谱，包括：

抽取所述风险感知因子集中的实体和实体关系；

基于所述实体和所述实体关系进行图谱构建，得到时序知识图谱。

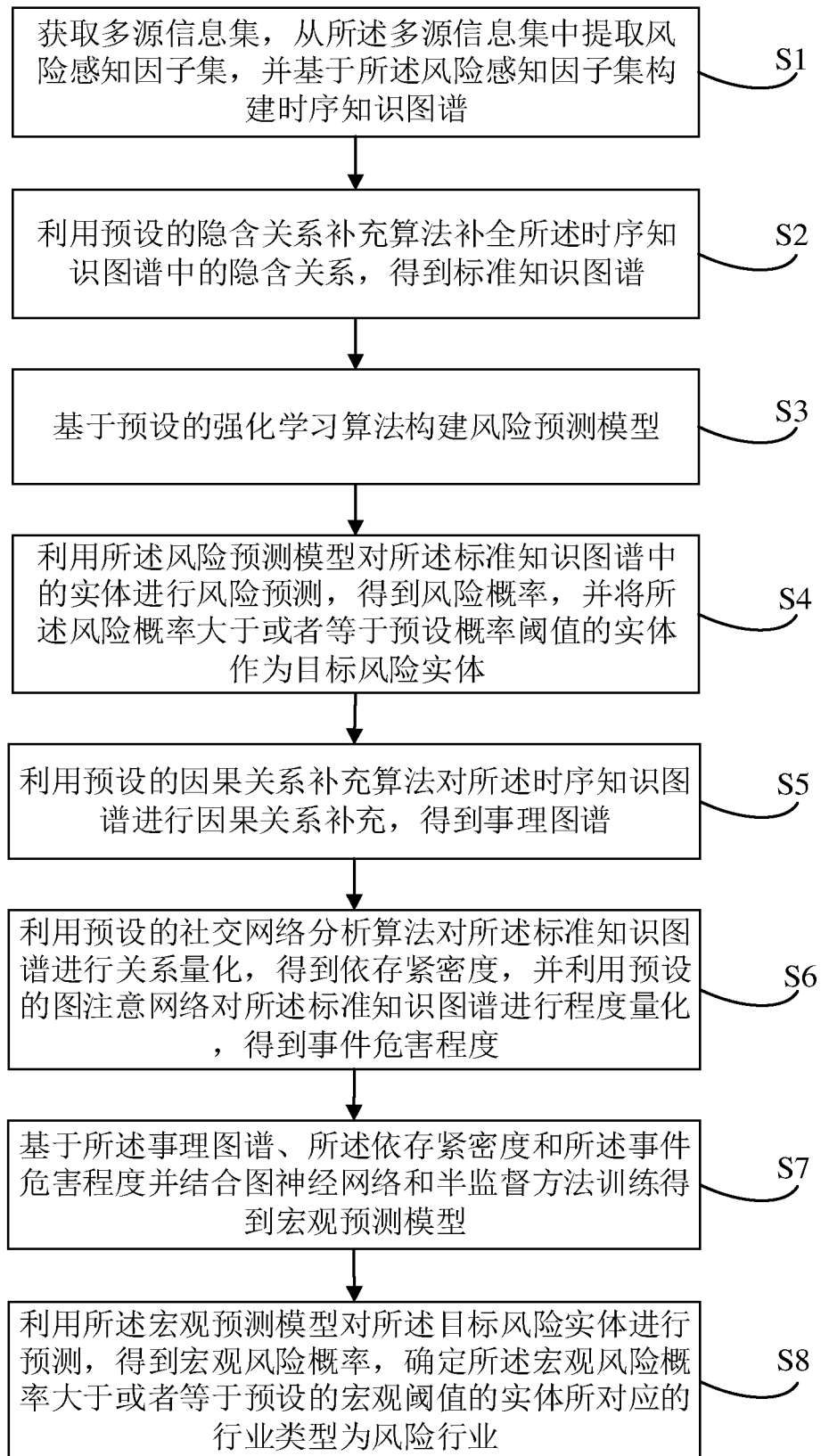


图 1

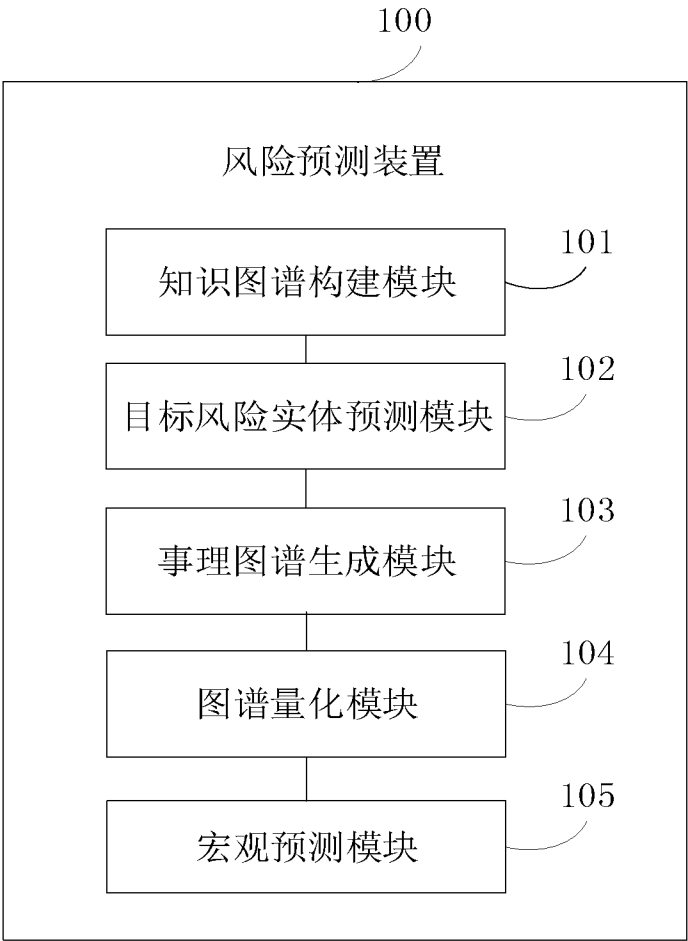


图 2

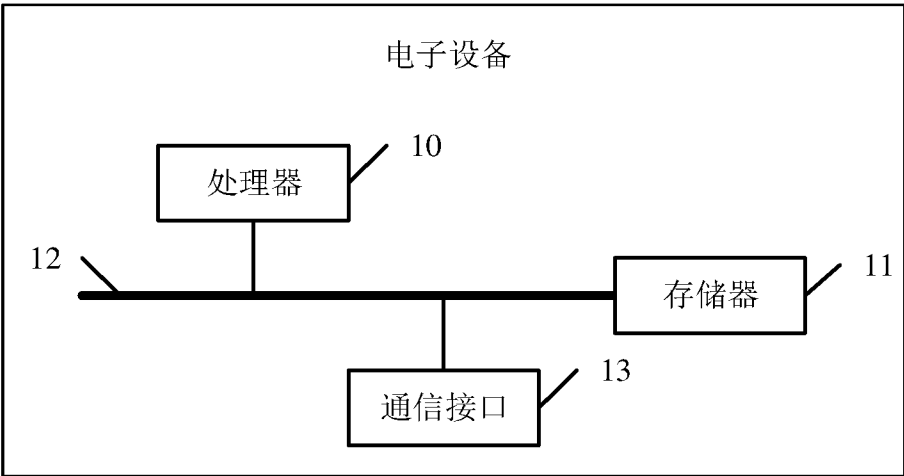


图 3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2022/071241

A. CLASSIFICATION OF SUBJECT MATTER G06Q 10/04(2012.01)i According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G06Q Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) CNPAT, CNKI, DWPI, EPODOC, IEEE: 风险, 危害, 风控, 控制, 预测, 概率, 知识图谱, 时序, 感知, 社交网络, artificial w intelligence, AI, risk w industr+, risk w business, financial w risk, risk w profession+, predict+ d risk?, analy+ d risk?, risk w percept+, knowledge w graph, time w sequence		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 113822494 A (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) 21 December 2021 (2021-12-21) entire document	1-20
A	CN 110795568 A (BEIJING QIYU INFORMATION TECHNOLOGY CO., LTD.) 14 February 2020 (2020-02-14) description, paragraphs 0050-0075	1-20
A	CN 111026874 A (HISENSE GROUP CO., LTD.) 17 April 2020 (2020-04-17) description, paragraphs 0138-0140	1-20
A	CN 112395423 A (BEIJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS) 23 February 2021 (2021-02-23) entire document	1-20
A	US 2021019674 A1 (QOMPLX, INC.) 21 January 2021 (2021-01-21) entire document	1-20
☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.		
* Special categories of cited documents: “A” document defining the general state of the art which is not considered to be of particular relevance “E” earlier application or patent but published on or after the international filing date “L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) “O” document referring to an oral disclosure, use, exhibition or other means “P” document published prior to the international filing date but later than the priority date claimed “T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention “X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone “Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art “&” document member of the same patent family		
Date of the actual completion of the international search 29 June 2022		Date of mailing of the international search report 20 July 2022
Name and mailing address of the ISA/CN China National Intellectual Property Administration (ISA/CN) No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088, China Facsimile No. (86-10)62019451		Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2022/071241

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	113822494	A	21 December 2021	None			
CN	110795568	A	08 June 2021	None			
CN	111026874	A	07 June 2019	None			
CN	112395423	A	23 February 2021	WO	2022052374	A1	17 March 2022
US	2021019674	A1	21 January 2021	None			

国际检索报告

国际申请号

PCT/CN2022/071241

A. 主题的分类 G06Q 10/04(2012.01)i 按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类				
B. 检索领域 检索的最低限度文献(标明分类系统和分类号) G06Q 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库(数据库的名称,和使用的检索词(如使用)) CNPAT,CNKI,DWPI,EPDOC,IEEE: 风险,危害,风控,控制,预测,概率,知识图谱,时序,感知,社交网络, artificial w intelligence, AI, risk w industr+, risk w business, financial w risk, risk w profession+, predict+ d risk?, analy+ d risk?, risk w percept+, knowledge w graph, time w sequence				
C. 相关文件				
类 型*	引用文件,必要时,指明相关段落	相关的权利要求		
PX	CN 113822494 A (平安科技深圳有限公司) 2021年12月21日 (2021 - 12 - 21) 全文	1-20		
A	CN 110795568 A (北京淇瑞信息科技有限公司) 2020年2月14日 (2020 - 02 - 14) 说明书第0050-0075段	1-20		
A	CN 111026874 A (海信集团有限公司) 2020年4月17日 (2020 - 04 - 17) 说明书第0138-0140段	1-20		
A	CN 112395423 A (北京邮电大学) 2021年2月23日 (2021 - 02 - 23) 全文	1-20		
A	US 2021019674 A1 (QOMPLX, INC.) 2021年1月21日 (2021 - 01 - 21) 全文	1-20		
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。				
<table><tr><td>* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件,或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件</td><td>“T” 在申请日或优先权日之后公布,与申请不相抵触,但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件,单独考虑该文件,认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件,当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时,要求保护的发明不具有创造性 “&” 同族专利的文件</td></tr></table>			* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件,或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件	“T” 在申请日或优先权日之后公布,与申请不相抵触,但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件,单独考虑该文件,认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件,当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时,要求保护的发明不具有创造性 “&” 同族专利的文件
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件,或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件	“T” 在申请日或优先权日之后公布,与申请不相抵触,但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件,单独考虑该文件,认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件,当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时,要求保护的发明不具有创造性 “&” 同族专利的文件			
国际检索实际完成的日期 2022年6月29日		国际检索报告邮寄日期 2022年7月20日		
ISA/CN的名称和邮寄地址 中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		授权官员 霍玉明 电话号码 010-53961327		

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2022/071241

检索报告引用的专利文件			公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN	113822494	A	2021年12月21日	无	
CN	110795568	A	2021年6月8日	无	
CN	111026874	A	2019年6月7日	无	
CN	112395423	A	2021年2月23日	WO 2022052374 A1	2022年3月17日
US	2021019674	A1	2021年1月21日	无	