

大宗商品价格风险的智能识别与预警

——基于动态因果知识图谱

刘伟龙¹ 易娜²

1. 中山大学岭南学院 广东广州 510275; 2. 广州期货股份有限公司 广东广州 511458

摘要: 文章报告了一种基于动态因果知识图谱的大宗商品价格风险识别与预警方法。该方法以大连商品交易所交易品种为对象,以产业链上下游企业为“节点”,以17种关联关系(如产业链、参股、资金流等)为“边”,整合覆盖产业链关系、分支机构、参股等企业关系类型,并详细记录了每家企业的经营范围、地理位置、注册资本等30个属性,形成包含因果逻辑的知识图谱与事理图谱。基于结构因果模型(SCM)、反事实分析等因果推断技术,提炼出三大核心风险指标:产业区域分布(揭示地理集中度与经济联动性)、核心企业分析(识别关键节点及风险传导能力)、价格风险传染路径(量化“地缘冲突—供应链中断—价格波动”等因果链条),通过正负向指标综合评估企业风险等级,实现风险实时预警。研究突破传统统计关联分析局限,解决高维数据伪相关问题,提升风险预警精度。在全球供应链不确定性加剧背景下,为国家战略资源安全保障与企业风险管理提供数据驱动的智能化解方案,契合“十四五”规划强化经济安全风险预警的政策导向。

关键词: 大宗商品; 价格风险; 贸易摩擦; 风险预警; 粮食安全; 因果知识图谱

中图分类号: F224.7; F832.5 **文献标志码:** A **文章编号:** 1672-8505(2025)04-0082-11

doi: 10.12189/j.issn.1672-8505.2025.04.008

Intelligent Identification and Early Warning of Commodity Price Risks

—Based on Dynamic Causal Knowledge Graph

LIU Wei-long¹ YI Na²

1. Lingnan College, Sun Yat-sen University, Guangzhou, Guangdong, 510275, China;

2. Guangzhou Futures Exchange, Guangzhou, Guangdong, 511458, China

Abstract: This study introduces a novel method for identifying and forecasting commodity price risks through the construction of dynamic causal knowledge graphs. Focusing on products traded on the Dalian Commodity Exchange, the proposed method models upstream and downstream enterprises along the industrial chain as "nodes", with 17 types of inter-enterprise relationships (e.g., industrial linkage, equity investment, capital flow) represented as "edges". The graph integrates relationship types including industrial chain connections, subsidiary affiliations, and equity holdings, and incorporates 30 attributes for each enterprise—such as business scope, geographic location, and registered capital—to build a causally enriched knowledge and event graph. Through causal identification, includ-

收稿日期: 2025-02-24

基金项目: 大商所2023年度“百校万才”工程研究项目(项目编号: DCEYJ202301)。

第一作者: 刘伟龙, 男, 科研博士后, 主要从事金融科技、金融工程与量化投资等研究。

引用格式: 刘伟龙, 易娜. 大宗商品价格风险的智能识别与预警——基于动态因果知识图谱[J]. 西华大学学报(哲学社会科学版), 2025, 44(4): 82-92.

ing Structural Causal Models (SCMs) and counterfactual analysis, the study derives three key risk indicators: (1) the industrial and regional distribution of relevant commodities (for revealing the spatial clustering and inter-regional economic dependencies of key enterprises); (2) the status of core enterprises (for identifying central nodes and their capacities for risk transmission within the network); and (3) the contagion paths of price risk (for quantifying such causal chains as "geopolitical conflict-supply chain disruption-price volatility", etc.). Enterprise risk levels are comprehensively evaluated using both positive and negative indicators, enabling real-time risk alerts. This approach addresses the limitations of traditional correlation-based statistical models, mitigates spurious associations in high-dimensional data, and significantly enhances early-warning accuracy. In the context of increasing global supply chain uncertainties, the method offers a data-driven, intelligent solution for safeguarding national strategic resources and improving enterprise risk management, in alignment with the 14th Five-Year Plan's emphasis on strengthening economic security and early warning systems.

Key words: commodities; price risk; trade friction; risk early warning; food security; causal knowledge graph

一、研究背景与文献综述

大宗商品价格风险研究具有重要的背景和意义,这与中国经济的发展、资源依赖、国际贸易以及市场波动等因素密切相关。《中华人民共和国国民经济和社会发展第十四个五年规划和2035年远景目标纲要》强调“强化经济安全风险预警、防控机制和能力建设,实现重要产业、基础设施、战略资源、重大科技等关键领域安全可控,着力提升粮食、能源、金融等领域安全发展能力。”这一政策指引为我国的农业、能源和金融领域制定了发展方向,也明确了风险监测和预警系统在应对国际形势的不确定性方面的重要作用。通过完善风险监测和预警系统、强化经济安全措施,我国能够更有效地抵御俄乌战争以及全球范围内的贸易摩擦等突发事件所造成的粮食和能源安全问题对国民经济带来的冲击。中国是全球最大的商品进口国之一,其国际贸易对全球大宗商品市场产生着重要影响。因此,大宗商品价格的波动会直接影响中国的进口成本、通胀率等经济指标。另外,中国对于许多关键大宗商品(如原油、铁矿石、大豆等)的依赖性较高,保障这些资源的稳定供应对国家的经济安全和发展至关重要。

大宗商品在中国经济中占有举足轻重的地位,对经济的支撑作用至关重要,其价格反映了中国乃至整个世界经济的发展状况^[1]。作为基础原材料,大宗商品在能源、制造、农业等领域发挥着关键作用,确保了中国产业的稳定运行和经济的持续增长。以农产品价格为例,农产品价格不稳定很大程度上会引发严重的社会问题^[2]。中国作为全球最大的大宗商品进口国之一,其经济发展与国际大宗商品供应链的稳定密切相关,大宗商品贸易也推动了人民币国际化的进程,提升了中国在全球贸易中的地位和影响力。数据作为数字经济的基础构建单元,在大宗商品市场中具有巨大推动力。数据如今已经成为数字经济深化发展的核心动力,被视为在土地、劳动和资本之后的又一重要的经济社会生产要素^[3]。首先,数据的应用可以极大地提升市场的透明度^[4]。通过实时收集、分析和展示大宗商品交易相关的数据,市场在数字经济时代,大宗商品交易面临着诸多挑战,且大宗商品具有稀缺性、替代性较低、价格波动性大以及全球供需关系紧密等特征^[5]。

对大宗商品价格风险的探究可以追溯到20世纪初,早期的研究主要集中在对大宗商品价格的基本驱动因素的探索,包括供需动态、市场投机,国际贸易政策等。随着20世纪的发展,全球事件进一步强化了大宗商品价格风险,促进了风险管理策略的演变^[6],强调了经济和政治危机相互关联的重要性,以及在缓解这些危机对公司影响方面操作和财务灵活性的重要性^[7]。进入20世纪末和21世纪初,研究重点转向大宗商品价格波动对全球经济稳定性的影响。这种转变部分是由诸如亚洲金融危机和2008年国际金融危机等事件引起的,这些事件凸显了全球大宗商品市场与金融系统的相互联系。这一时期的研究开始探索由大宗商品价格波动带来的系统性风险以及可能导致的全球经济传染。例如, Frank^[8]的研究揭示了2008年金融危机对亚洲大米出口市场的影响,为理解金融动荡期间大宗商品市场的复杂动态提供了

见解。此外,Zheng等^[9]的研究显示地缘政治风险冲击对各类大宗商品期货价格波动的短期影响较大且持续时间较长,突出了经济和金融危机对大宗商品市场环境的更广泛影响。

国外关于大宗商品价格风险的研究起步较早,20世纪90年代就已经有了较系统的研究(如Pindyck和Rotemberg^[10],Deaton和Laroque^[11]等),世界银行1997年提出并设计了大宗商品价格预测模型,美国商品期货交易委员会还提出了气候对于大宗商品风险的影响^[12]。从研究视角看,既包括宏观经济角度,也有金融市场角度。从研究方法看,主要涉及机器学习、计量经济学和运筹学方法。目前已经积累了丰富的研究成果并形成了系统的理论体系,研究水平整体处于先进水平。

我国关于大宗商品价格风险的研究起步于2000年后(如褚珉海^[13],佟德庆^[14]等)。其研究涵盖了经济、金融、农业、能源等多个领域。从研究内容看,主要包括大宗商品市场的供求变化、价格波动的影响因素、市场预测和风险管理等方面。研究方法也基本能够紧跟西方发达国家的发展。经过20多年的发展,也积累了大量的研究成果,研究水平处于国际先进水平。但我国在某些领域面临数据可用性和准确性的挑战,这可能影响到研究的深度和广度,西方国家在一些方面有一定优势。另外,中国经济体制与西方国家存在差异,这会影响到价格风险研究中考虑的因素和政策应对。

随着近年来地缘政治冲突和贸易摩擦带来的全球不确定性的增加,近期,美国总统特朗普宣布对中国进口商品加征125%的关税,导致中国对美出口商品面临更高的贸易壁垒。作为回应,中国对美国进口商品加征84%的关税,进一步升级了贸易紧张局势。这种关税的相互加征直接影响了大宗商品的价格和贸易流向。例如,铜作为重要的工业金属,其价格在关税政策变化后出现波动。在特朗普宣布对中国商品加征关税后,铜价一度下跌,但在随后宣布90天的关税暂停期后,铜价又有所回升。这也证明了,地缘政治风险对大宗商品期货价格具有显著影响,尤其在冲突和不稳定时期,价格波动更加剧烈。

与此同时,如何保障粮食、能源和金融等领域的安全性已成为各国普遍关注的重大问题,大宗商品风险监测和预警系统完善也成为各国亟待解决的问题之一。目前大宗商品价格风险的研究热点主要包括宏观经济数据与商品价格相关性研究、高频数据在监测中的应用、机器学习和深度学习在模型构建中的运用、区块链技术在溯源和预警中的创新应用等。

而在这些创新应用中,因果知识图谱(Causal Knowledge Graph)在大宗商品价格风险研究中的应用,依托结构因果模型(SCM)、反事实推理等理论框架^[15],为解析复杂市场机制提供了突破性工具。以Judea Pearl^[15]的因果推断理论为基础,CKG通过有向无环图(DAG)构建变量间的因果依赖网络,例如量化“地缘冲突—供应链中断—价格波动”这一知识图谱的传导路径,突破仅仅依赖传统计量模型所捕捉统计相关性,而是展示出真正的因果关联。其动态知识图谱特性支持多源异构数据(如库存、政策、供应链节点)的实时融合,结合时间序列分析捕捉非线性演变规律。在高频数据应用中,CKG通过因果检验等,解决了高维数据下的伪相关问题,显著提升现实问题中的预警精度。面对中国场景的数据治理与体制差异挑战,可通过图神经网络(GNN)识别有效节点(如地方储备粮数据)同时利用因果干预,如do算子等^[16],对比分析行政干预与市场调节的因果效应。

从近年来大宗商品价格风险研究的发展看,未来发展趋势是加强模型的智能化和信息化建设,构建起多源异构数据驱动的智能化工大宗商品风险监测和预警系统。未来,预计大宗商品价格风险监测和预警系统将更加智能化、自动化和高效化,能够实时采集和处理大量的市场数据,利用先进的模型和算法进行数据分析和预测,提供更准确、全面的风险评估和预警服务。

二、知识图谱和事理图谱构建框架

(一) 大宗商品知识图谱和事理图谱实体要素及要素关系提取

现有的大宗商品大数据中包含了海量的能体现金融实体之间关联关系的信息,传统的金融风险链特征提取与指标体系,很难应用于此种关联型数据。本研究以品种为对象,以企业上下游和个人及行业实

体等金融风险对象作为“节点”,以实体之间的关联关系作为“边”,形成一个动态的大宗商品知识图谱网络。建立的知识图谱包括:①企业网络图——建立共同股东、共同高管和资金流水等相关联的企业关系网络图;②贸易关系网络图——构建企业间贸易网络图;③投资网络图——构建个人与企业的股权关系与信贷关系网络图;④构建企业高管的社交网络图——小微企业高管个人的社会资源的社交网络。并且同时通过因果语义重构实体与关系,将品种、企业、行业、政策及个人等节点赋予因果角色(如因变量、中介变量),并基于结构因果模型(SCM)^[16]升级关联关系为因果依赖。例如,通过因果发现算法基于已经建立的知识图谱,识别“共同高管决策→资金流水异常→供应链违约”的因果传导链;贸易网络则利用反事实分析验证企业断供事件对下游生产成本的因果效应;投资网络通过量化股权质押与信贷风险的传导权重,揭示个人行为对企业融资成本的深层影响;社交网络进一步挖掘高管社会资源与政策信息优势之间的隐性因果路径。通过因果假设检验(如 do-calculus)剔除伪相关链接,最终构建具有经济学解释力的因果网络。

(二) 大宗商品知识图谱和事理图谱网络特征分析

在构建了因果知识图谱后,如何对网络图中的结构性特征进行表示,以更好地利用数据挖掘方法研究,实现因果网络的可计算化。还需要解决的问题,主要包括以下内容。①知识图谱的向量化表示。知识图谱是一种非结构化的网络图,大多数传统的数据挖掘方法无法直接在网络图中应用,因此本部分拟将网络图中的结构化特征嵌入到数值向量空间,以进行下一步学习。②金融异构网络的特征融合。金融实体之间的交互具有多样性,以此形成的金融知识图谱中的网络连接往往具有不同的属性,因此是一种典型的异构网络。本文将研究如何将这种金融异构网络中的不同性质边连接序列形成的向量特征融合在同一个向量空间,对金融知识图谱的总体特征进行表示。③知识图谱特征的解耦学习。金融数据易受众多的经济性、社会性因素影响,提取的数据特征之间存在着广泛的线性与非线性相关关系。本部分研究还将对上一步形成的特征向量空间中各维度的相互依赖关系进行分析,并变换为具有线性独立,分布简单等优良特性的新的知识图谱特征表示。同时,采用因果图神经网络(Causal GNN)将节点表征分解为因果影响因子(如政策冲击、供应链中断)与非因果噪声(如市场情绪波动),并引入时间维度捕捉因果效应的动态演化(如地缘冲突初期价格飙升与长期供应链重构的差异)。针对金融异构网络的复杂性,设计多模态因果注意力机制,区分贸易关系、股权关联等不同边类型对价格波动的直接或间接因果贡献度,并且,通过后门调整分离混杂效应(如企业规模对贸易量与抗风险能力的双重干扰),结合因果不变性学习提取跨周期、跨市场的稳定因果特征(如“美联储加息→美元定价商品下跌”的普适规律),最终实现高纯度因果特征表示。

(三) 大宗商品图谱智能表征学习及价格风险监测

在构建因果知识图谱的特征表示基础上,进一步提炼大宗商品风险指标体系,构建融合可解释性与干预导向的风险量化模型,以形成金融风险度量与金融管理决策的依据。

本文提出或使用相关挖掘算法,从正负两个方面分别评价金融实体的风险状况,然后设计新的综合评分体系,整合正负面的评价指标,最终实现对企业的风险度量。具体来讲,课题拟首先基于金融知识图谱的特征空间,挖掘金融实体的影响力、路径关键性、社区分布等特征,以对知识图谱中的金融实体“影响力”进行评估,形成“正向”评分指标,而正向评分体系通过因果中心性度量节点全局影响力(如关键矿产企业断供的产业链级联效应),并识别敏感因果链(如“极端天气→减产→通胀预期”)的脆弱性与干预优先级;基于对知识图谱网络中结构性异常、风险子图匹配、长程风险依赖等方面问题的研究,对企业的风险进行评估,形成“负向”评价指标。与此同时,负向风险识别利用反事实风险值(CRV)模拟事件未发生的潜在差异(如无政策干预时的铜价涨幅风险);最后,提出不依赖于风险标签的无监督式的综合风险度量指标体系,及依赖少量风险标签信息的半监督式的最优化综合风险度量指标体系,将传统风险指

标的“描述性关联”升级为“机制性解析”,为精准调控关键因果节点提供靶向决策支持。

三、基于动态因果知识图谱的图谱构建

(一) 基于知识图谱的因果特征生成

通过多模态生成模型(如 GPT-4V)生成输入图像,并将其划分为若干等尺寸模块,每个模块对应图像的一个局部特征。同时,通过知识图谱引导的因果矩阵预测,将模块化特征输入因果预测模型,生成因果矩阵以区分因果与非因果区域。在此阶段,因果知识图谱(CKG)中的先验关系(例如“白色长牙→大象”等因果边)被编码为约束条件,辅助模型识别与目标变量(“大象”)直接相关的因果特征。

1. 注意力驱动的因果选择与知识关联

将因果矩阵与原始图像输入选择模块,利用注意力机制聚焦因果区域,并通过关联排名强化因果元素的关联性,进行因果和非因果的分离。非因果部分(如背景噪声)基于 CKG 中的无关实体关系被机制动态屏蔽,确保后续推理仅依赖因果相关特征。

2. 对比学习与知识嵌入对齐

利用编码器将筛选后的因果特征转换为低维表征向量,并基于此构建知识图谱驱动的对标损失。通过对标学习模块,将嵌入向量与 CKG 中对应的因果实体的语义嵌入对齐,最大化因果正样本对的相似性,最小化非因果负样本对的相似性,从而优化模型对因果结构的捕捉能力。

3. 反事实验证与知识图谱干预

反事实样本构建:从随机类别中提取反事实图像的因果元素,并将其与非因果背景(如固定场景)结合,生成干预样本。同时,对反事实样本进行分类,验证模型输出的因果特征是否在干预后保持一致性。若结果与 CKG 中定义的因果机制(如“阿司匹林是头痛缓解的必要条件”)一致,则证明模型具备因果鲁棒性。

4. 知识图谱的动态更新与迭代优化

反事实验证结果被反馈至 CKG,新增或修正因果边,形成闭环迭代。通过图神经网络(GNN)将 CKG 的结构化知识注入模型训练,确保因果预测与知识图谱的语义逻辑一致。

(二) 正样本对

正样本对指的是共享相同因果机制的图像数据对,它们在因果推理过程中起到核心作用。这些正样本对不仅仅是相似图像,而是具备相同因果特征的实例。在本节中,文章将详细介绍正样本对的生成方法、因果特征的提取与编码、对比学习的优化策略,并结合知识图谱(Knowledge Graph, KG)进一步增强因果推理能力。

1. 生成器

为了构造符合因果学习目标的数据集,首先通过文本到图像模型(GPT-4V)生成高质量图像数据,并进行结构化处理,使得模型能够学习稳定的因果关系。

(1) 文本到图像转换(Text-to-Image Generation)

在数据构造阶段,利用 GPT-4V 生成基于文本描述的图像。具体而言,给定一个文本描述 T ,生成的图像表示为: $I = G(T)$ 。

其中, $G(\cdot)$ 为文本到图像的转换函数。例如,文本“一只在森林里行走的大象”将生成带有大象和森林背景的图像。

(2) 图像结构化与网格划分(Image Partitioning)

为了增强因果建模的空间感知能力,可将图像划分为 $M \times N$ 的网格,每个网格单元表示一个特征区域:

$$X = \{x_{ij} | i \in [1, M], j \in [1, N]\}$$

这样处理能够精细化提取因果特征,同时提供更多结构化信息,便于后续计算因果矩阵。

(3) 知识图谱增强(Knowledge Graph Augmentation)

为了确保图像生成过程中具备合理的因果关系,引入知识图谱(KG),利用其中的实体—关系信息来优化数据构造。知识图谱的基本结构由三元组表示:

$$KG = \{(e_i, r, e_j)\}$$

其中, e_i 和 e_j 分别表示两个实体, r 代表它们之间的关系。例如,在“大象”这一概念中,知识图谱可以提供以下因果信息:(大象,生活在,森林/草原),(大象,具有,象牙),(森林,包含,树木)。通过KG,能够确保生成的正样本对共享一致的因果关系,而不仅仅是视觉上的相似性。

2. 因果预测模型

在获取结构化的图像数据后,使用因果预测模型 $f_{\mathcal{O}}(\cdot)$ 计算因果矩阵(Causal Matrix) \mathbf{C} ,以区分因果相关部分与非因果部分。

(1) 因果矩阵的定义(causal matrix definition)

因果矩阵 \mathbf{C} 是一个二值矩阵,其计算方式如下:

$$\mathbf{C} = f_{\mathcal{O}}(X)$$

其中, $\mathbf{C} = \{0, 1\}^{M \times N}$,若 $C_{ij} = 1$,则特征块 x_{ij} 因果相关(rationale),即对目标决策具有直接贡献。若 $C_{ij} = 0$,则特征块 x_{ij} 非因果(non-causal),可能是背景或噪声信息。

(2) 结合知识图谱的因果推理

利用图神经网络(graph neural network, GNN)对知识图谱中的实体进行嵌入,以增强因果推理能力。并采用图注意力网络(graph attention network, GAT)计算因果特征之间的关联:

$$h_i^{(l+1)} = \sigma \left(\sum_{j \in N(i)} a_{ij} W h_j^{(l)} \right)$$

其中:

$h_i^{(l)}$ 表示第 l 层的节点嵌入表示;

$$a_{ij} \text{是注意力系数: } a_{ij} = \frac{\exp(\text{LeakyRelu}(\mathbf{a}^T [W h_i | W h_j]))}{\sum_{k \in N(i)} \exp(\text{LeakyRelu}(\mathbf{a}^T [W h_i | W h_k]))}$$

3. 特征编码器

特征编码器的核心目标是从图像的局部区域中提取关键因果特征,并将其转换为高维嵌入表示,以便进行下游的对比学习和分类任务。相比于传统的图像特征提取方法,这里的编码器不仅利用深度卷积神经网络(CNN)对视觉信息进行建模,还结合了知识图谱和注意力机制以增强因果特征的代表能力。

(1) 因果特征的提取(Feature Selection)

在因果预测模型的作用下,获得了因果矩阵 \mathbf{C} ,该矩阵指示了输入图像中的哪些区域属于因果相关部分(rationale),哪些属于非因果部分(non-causal regions)。因此,可以基于 \mathbf{C} 进行特征筛选:

$$X_c = \{x_{ij} | C_{ij} = 1\}$$

其中, X_c 代表从图像中筛选出的因果特征子集。这一操作的本质是从原始图像的所有网格单元中,仅保留那些与目标任务因果相关的区域,在分类任务中,仅保留动物的主要身体关键部分,而过滤掉背景和无关

关元素。

相比于传统特征提取方法,这里的因果特征提取过程能够有效去除冗余背景信息,避免混杂因子的影响,提高模型的可解释性。

(2)多模态融合(Multimodal Fusion)

单独依赖图像特征进行因果建模可能存在信息不足的问题,因此进一步结合知识图谱(KG)和文本信息,构建一个更加全面的因果表示。

$$Z = f_{\phi}(X_c, KG, T)$$

其中:

X_c 由因果矩阵 C 提取出的因果相关图像特征。

KG : 知识图谱嵌入,提供关于图像中实体的背景知识,例如“大象”与“草原”或“森林”的关联性。

T : GPT-4V生成的文本描述,例如“这是一只成年非洲象”。

在多模态融合过程中采用Transformer结构,以建模跨模态信息的长程依赖关系:

$$Z = \text{Transformer}(X_c + \text{Emb}(KG) + \text{Emb}(T))$$

其中:

$\text{Emb}(\cdot)$ 表示嵌入层,将知识图谱和文本转换为可学习的高维向量。

Transformer采用自注意力机制(self-attention mechanism)计算不同模态信息之间的关联性,使得模型能够在不同来源的信息之间进行交互学习。

例如,在识别一只在森林里的大象时:视觉特征 X_c 提供关于大象形状和颜色的信息。知识图谱 KG 提供关于“大象通常生活在森林或草原”的先验知识。文本信息 T 提供GPT-4V生成的环境描述,使得模型能够更好地理解上下文。通过这种多模态融合,该特征编码器不仅仅依赖图像本身,还能综合利用外部知识,增强因果推理能力,提高因果特征的稳定性和泛化能力,构建一个更加全面的因果表示。

4. 对比学习

对比学习(contrastive learning)是一种无监督学习方法,主要用于学习数据的区分性表示。在本研究中采用对比学习来优化因果特征提取过程,使得模型能够更加有效地区分因果相关和非因果特征,提高对因果关系的学习能力。

在因果推理任务中,本研究希望模型能够学习到图像中哪些部分真正影响目标变量,并忽略背景或其他无关因素。对比学习的核心思想是:拉近正样本对的表示(positive pairs),即共享相同因果机制的样本应该具有相似的嵌入表示。推远负样本对的表示(negative pairs),即无因果关联的样本应该被区分开来。这一过程能够帮助模型学会区分因果特征和非因果特征,从而提高因果推理能力。

为了实现上述目标,采用对比损失进行优化:

$$\text{sim}(Z_i, Z_j) = \frac{Z_i Z_j}{\|Z_i\| \|Z_j\|}$$

其中 Z_i 和 Z_j 分别表示两个样本的嵌入表示,分子为它们的点积,分母为它们的范数,确保相似度值在 $[-1,1]$ 之间。基于这个相似度,使用对比损失函数来优化嵌入空间:

$$L_{\text{contrast}} = -\log \frac{\exp(\text{sim}(z_i, z_j | \tau))}{\sum_{k \neq i} \exp(\text{sim}(z_i, z_k | \tau))}$$

这个损失函数的作用是最大化正样本对之间的相似度,同时最小化负样本对之间的相似度,从而学

习到更加区分因果和非因果特征的表示空间。在传统的对比学习框架下,样本之间的相似度是基于数据本身的特征来计算的。然而,在因果推理任务中,知识图谱可以提供额外的语义信息,帮助模型更准确地理解样本之间的因果关系。因此,进一步引入知识图谱信息来指导对比学习。其中,利用知识图谱计算样本之间的语义相似度 w_{ij} :

$$w_{ij} = \frac{1}{1 + d(e_i, e_j)}$$

其中, e_i 和 e_j 是知识图谱中的两个实体(对应样本的概念)。 $d(e_i, e_j)$ 是知识图谱中的最短路径距离,表示两个样本在因果关系上的接近程度。

结合以上:

$$L_{KG} = \sum_{(e_i, r, e_j) \in KG} w_{ij} L_{\text{contrast}}$$

相比于传统的对比学习方法,这种方式能够更有效地利用因果知识,提高模型的推理能力。

(三) 负样本对

在因果推理和表示学习任务中,负样本对的作用至关重要。与正样本对不同,负样本对是指具有不同因果机制或不共享关键语义信息的图像对。负样本的构建和训练可以帮助模型区分因果与非因果特征,避免模型仅依赖背景或其他非因果因素进行预测,从而提高模型的稳健性和泛化能力。

1. 负样本构造

负样本对由以下三种方式构造。

(1) 不同类别样本对

随机选择两幅来自不同类别的图像 X_i 和 X_j 作为负样本对。例如, X_i 是一只“大象”的图像,而 X_j 是一只“狗”的图像,则它们构成一个负样本对。这种方法保证了负样本对在高层语义上具有显著差异。

(2) 同类别但不同背景样本对

为了确保模型不是仅依赖背景信息进行分类,从相同类别的图像中选择负样本对,但修改其背景。例如,原始图像可能是“在草原上的大象”,而负样本则是“在城市街道上的大象”。

这种构造方式能够有效测试模型是否真正关注了因果特征,而非仅依赖于背景线索进行识别。

(3) 反事实样本对

通过修改原始样本的因果特征,生成反事实负样本。将一张“大象”图像的关键因果特征(如象牙、耳朵)替换为其他动物(如狗的耳朵和鼻子),从而构造一个挑战性的负样本。

$$X' = g(X, \delta)$$

其中, $g(\cdot)$ 表示反事实变换函数, δ 代表修改的特征。

2. 非因果矩阵计算

为了确保模型在区分因果与非因果特征方面具备更强的能力,构建非因果矩阵 S ,用于屏蔽关键因果信息,使模型仅能依赖无关特征进行推理。非因果矩阵 S 由因果矩阵 C 反向计算得到:

$$S = 1 - C$$

其中: $S_{ij} = 1$ 表示该区域为非因果区域,即该部分的特征不应影响分类决策。 $S_{ij} = 0$ 表示该区域为因果相关区域,在负样本处理中被屏蔽。

为了进一步评估模型的鲁棒性,从原始图像网格 X 中选择仅包含非因果区域 X_s ,将这些非因果特征输入到分类器 $f_\delta(\cdot)$ 中,观察其预测结果,如果模型在仅使用非因果特征的情况下仍然能够正确预测类别,

则说明它可能过度依赖背景信息,而非关键因果特征,这需要进一步优化因果预测模型 $f_{\delta}(\cdot)$ 。

3. 反事实验证

为了确保模型真正学到了因果特征,基于以上所构建的反事实样本对,优先选择知识图谱中相邻类别的特征,例如,将“大象的象牙”替换为“犀牛的角”,而非完全随机地改变特征,观察模型的预测变化。

整体上,该框架不仅显著提升了因果特征的识别与解释能力,还通过闭环反事实验证和动态知识更新,有效改善了模型在分布外数据下的泛化性能,为多模态因果推理提供了一种系统、严谨且高效的解决方案。

四、知识图谱构建的案例研究

本研究通过构建一个分布式智能检索引擎,收集并整合了大宗商品企业的工商注册信息及其关联数据,创建了一个涵盖玉米、豆粕、铁矿石、焦煤、聚乙烯、聚氯乙烯等商品的庞大企业级知识图谱。该图谱包括超过10万家企业和15万条关系数据,覆盖产业链关系、分支机构、参股等17种企业关系类型,并详细记录了每家企业的经营范围、地理位置、注册资本等30个属性。利用这些详尽的数据,研究采用数据挖掘技术识别潜在风险模式,如企业的财务状况、信用记录和历史业绩,以便发现可能影响供应链稳定性或引起价格波动的因素。基于此,本研究开发了算法模型,评估市场需求变化、供应链稳定性和政策变动等多种风险因素对大宗商品价格和供应的影响。这些模型经过历史数据训练,能够预测未来风险趋势。此外,研究还建立了一个实时监控与预警系统。该系统运用机器学习技术跟踪关键企业和商品的最新动态,实时更新风险评估,并及时发出预警信号,帮助决策者迅速响应,有效管理和减轻潜在风险。

金融风险传播对金融稳定性构成重大威胁,迫切需要有效的检测与缓解策略。传统的时空图学习模型面临多重挑战,包括低效的推理能力、稀缺的标记数据、类别不平衡、有限的因果解释能力等。为应对这些问题,我们推出了一种创新的时空图学习框架——SIGNAL,专门用于金融风险的识别和因果分析。SIGNAL框架包括三个创新组件:①一种时空图掩码自编码器,有效降低计算复杂度并增强信息捕获;②一种半监督风险检测器,通过聚类技术在极少量标签数据的支持下实现高效风险识别;③一种图向量自回归模型,用于精确量化和验证风险传播的路径与强度。在两个开源数据集上的测试结果显示,SIGNAL在风险识别和因果解释方面均优于现有模型。这项研究不仅开创了使用时空图学习进行金融风险识别和因果分析的先例,还在计算效率和模型解释性方面取得了显著进展,为构建更为坚实的金融风险管理策略奠定了基础。

五、基于知识图谱的风险指标发现

(一) 产业区域分布

基于知识图谱的大宗商品区域集中度分析,可以更直观地理解大宗商品在不同地区的分布情况及其影响因素。知识图谱通过节点和边的关系,将大宗商品的生、消费、贸易等各个环节紧密地联系在一起。在大宗商品区域集中度的分析中,可以将各个地区作为节点,将地区间的大宗商品流动作为边,从而形成一个复杂而精细的网络。通过观察这个网络,可以发现某些地区在大宗商品的某些品种上具有较高的集中度。这种集中度可能源于该地区的资源禀赋、产业链优势、市场需求等多种因素。同时,知识图谱还能揭示出不同地区之间大宗商品流动的路径和规模,帮助我们更好地理解区域间的经济联系和相互影响。此外,知识图谱还可以用于预测大宗商品区域集中度的变化趋势。通过分析历史数据和当前市场动态,可以构建预测模型,预测未来大宗商品在不同地区的分布情况,为政府和企业制定相关政策和战略提供科学依据。

(二) 核心企业分析

基于知识图谱的大宗商品核心企业分析,可以清晰地揭示出在大宗商品领域中具有重要地位和影响

力的企业及其相互关系。

知识图谱通过构建节点和边的关系,将大宗商品核心企业及其业务往来紧密地联系在一起。在这个网络中,核心企业作为关键节点,其影响力可以通过边的关系传递到整个网络中。通过分析知识图谱,可以发现哪些企业在大宗商品领域具有更强的实力和竞争力。这些企业通常拥有完善的产业链布局、先进的生产技术和强大的市场拓展能力。同时,知识图谱还能揭示出核心企业之间的合作关系和竞争格局,帮助人们更好地理解大宗商品市场的运作机制和行业发展趋势。此外,基于知识图谱的分析还可以为投资者提供有价值的参考信息。通过关注核心企业的发展动态和市场表现,投资者可以更加准确地把握大宗商品市场的脉搏,从而作出更为明智的投资决策。

(三) 传染路径分析与企业风险预警

基于知识图谱的大宗商品传染路径分析,是理解和预测大宗商品市场动态的重要手段,表1展示了基于知识图谱的企业属性及风险评级样例。知识图谱能够将大宗商品、相关企业、地区及市场变动等复杂关系以结构化的方式展现,使得传染路径的追踪成为可能。在大宗商品市场中,一种商品的价格波动往往会对其他商品乃至整个市场产生影响,这种影响通过复杂的供应链和市场关联传递。通过分析知识图谱中的节点(代表大宗商品、企业、地区等)和边(代表它们之间的关系),由此可以追踪到某一商品的价格变动是如何沿着供应链或市场关联路径影响到其他商品的。这有助于揭示潜在的市场风险,并为政策制定者和投资者提供早期预警。此外,基于知识图谱的传染路径分析还可以用于评估不同大宗商品之间的关联性,以及这种关联性对市场稳定性的影响,为大宗商品市场的风险管理提供有力支持。

表1 基于知识图谱的企业属性及风险评级样例

名称	交易方向	企业评分	行业属性
***集团有限公司	供应商	83	交通运输、仓储和邮政业
**中海粮油工业有限公司	客户	83	批发业
****粮油工业张家港有限公司	客户	83	制造业
**生态食业有限公司	客户	83	制造业
**中国贸易有限责任公司	客户	83	批发业
***港股份有限公司	供应商	82	水上运输业
****钦州有限公司	客户	81	交通运输、仓储和邮政业
****粮油有限公司	客户	81	制造业
****营口有限公司	客户	81	农副食品加工业
****物流有限公司	供应商	80	租赁和商务服务业
****粮油工业有限公司	客户	80	农副食品加工业
****物流有限公司	客户	80	租赁和商务服务业
***南通有限公司	客户	79	农副食品加工业
***农业发展有限公司	客户	58	农副食品加工业
****饲料有限公司	客户	68	制造业

六、局限与展望

基于知识图谱的大宗商品风险分析,尽管为市场参与者提供了更为全面和深入的风险洞察,但仍面临一些局限,并需要在未来进行进一步的探索与展望。

首先,知识图谱的构建本身就是一个复杂的过程,需要收集、清洗和整合大量的数据。在大宗商品

市场中,数据的质量和完整性往往受到多种因素的影响,如信息披露的不充分、数据标准的不统一等,这可能导致知识图谱的准确性和完整性受到限制。

其次,大宗商品市场的风险具有多样性和复杂性,包括价格波动风险、供应链风险、政策风险等。知识图谱虽然能够揭示出风险之间的关联性和传导路径,但在具体风险的量化评估和预测方面,仍需要结合其他分析工具和方法,如统计模型、机器学习算法等。

随着大数据、人工智能等技术的不断发展,知识图谱在大宗商品风险分析中的应用前景将更为广阔。一方面,数据质量的提升和数据标准的统一将有助于提高知识图谱的准确性和完整性;另一方面,结合其他先进的分析工具和方法,知识图谱将能够更全面地揭示出大宗商品市场的风险特征和变化趋势,为市场参与者提供更加精准的风险预警和决策支持。同时,也需要关注知识图谱在隐私保护和合规性方面的挑战,确保在利用知识图谱进行风险分析的同时,能够充分保护个人隐私和遵守相关法律法规。

在当前国际局势日益复杂的背景下,尤其是受到中美贸易冲突等全球性事件的影响,大宗商品市场面临的风险格局在不断演化,知识图谱的构建与应用也面临新的挑战与机遇。未来,企业之间的跨境交易、供应链协同和金融传导路径将变得更加隐蔽和动态,传统的静态风险识别模式难以适应这种高频变动和政策敏感的市场环境。与此同时,在全球贸易不确定性持续增强的趋势下,知识图谱将在大宗商品风险防控中扮演愈发核心的角色,但也必须不断拓展其数据边界、分析维度与应用深度,方能更好服务于国家战略与企业决策。

参考文献:

- [1] 邱中行,徐小芳.地缘政治风险对大宗商品期货价格的影响研究[J].*财富涌现与流转*,2023,13(3):15-25.
- [2] 廖杉杉,鲁钊阳.农产品价格风险的成因及规避机制研究[J].*农村经济*,2013(3):27-30.
- [3] 黄莉群,官心果,钟宇.数字经济时代的数据安全研究——以金融行业为例[J].*商业经济*,2024(2):174-179.
- [4] ATHIQUE A. Integrated commodities in the digital economy[J]. *Media, culture & society*, 2020, 42(4): 554-570.
- [5] 郭金涛.大宗商品贸易融资模式创新与风险防控[EB/OL]. [2025-01-15]. <https://ojs.s-p.sg/index.php/fm/article/download/13734/12837>.
- [6] BEASLEY B Alex. Overview: The oil shocks of the 1970s[EB/OL]. [2025-01-15]. <https://energyhistory.yale.edu/the-oil-shocks-of-the-1970s/>.
- [7] SCHOLES M S. Crisis and risk management[J]. *American economic review*, 2000, 90(2): 17-21.
- [8] FRANK C. The financial crisis of 2008 and its impact on the Asian rice export market[J]. *Journal of advances in economics and finance*, 2017, 2(2):97-116.
- [9] ZHENG D Y, ZHAO C G, HU J Y. Impact of geopolitical risk on the volatility of natural resource commodity futures prices in China[J]. *Resources policy*, 2023(83): 103568.
- [10] PINDYCK R S, ROTEMBERG J J. The excess co-movement of commodity prices[J]. *The economic journal*, 1990, 100(403): 1173-1189.
- [11] DEATON A, LAROQUE G. On the behaviour of commodity prices[J]. *The review of economic studies*, 1992, 59(1): 1-23.
- [12] MONASTEROLO I. Climate change and the financial system[J]. *Annual review of resource economics*, 2020(12): 299-320.
- [13] 褚珺海.中国期货市场风险研究[M].北京:中国财政经济出版社,2001.
- [14] 佟德庆.期货市场风险及其监管研究[D].西安:西北大学,2005.
- [15] PEARL J. Causal diagrams for empirical research[J]. *Biometrika*, 1995, 82(4): 669-688.
- [16] GOLDSZMIDT M, PEARL J. Rank-based systems: A simple approach to belief revision, belief update, and reasoning about evidence and actions[C]. KR'92: Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning, 1992: 661-672.

[责任编辑 刘书亮]