刘莉. 黄金价格动态预测和影响因素研究[D]. 山东财经大学, 2016.

黄金是一种具有金融属性的产品,是商品市场上较为重要的现货、期货交易产品之一。在2008年的世界经济危机中,各种金融产品都表现出疲软的现状,而只有黄金的市场具有突出的表现,近期在中国的股市大幅度波动中,黄金价格变化也不是很大。基于大数据分析方法,探究黄金价格变化规律,无论对投资者、生产者还是政府管理者,都有积极的意义。本篇文章结构分为五个部分,分别从三个方面对黄金价格进行了研究。在对行文背景、研究思路、内容以及方法分析的基础上,首先利用协整突跃点检测模型、箱线图等方法对黄金价格收益率数据进行了检测,查出了所选数据区间内较异常点所对应的日期;其次分别利用去除异常点和不去异常点的数据进行ARIMA-GARCH模型的建立,研究表明去掉异常点之后所建模型的显著性有所提高,接着利用该模型对去异常点之后的黄金价格进行了波动性分析与预测,又利用数据挖掘技术中的支持向量机模型对相同的黄金价格数据进行预测,并将预测结果进行对比,发现支持向量机的预测效果比时间序列模型的预测效果更精确;再次,在对黄金价格进行波动性分析的基础上,研究了黄金价格的一些影响因素,并且引入了两个新的变量,在后续的分析中,我们利用不同的解释变量建立了不同的VAR模型。研究发现,就模型的整体效果来说,引入新变量后效果变好了。另外,我们还分别利用月度数据和日数据建立了VAR模型并提出了对比结果。最后根据一系列的分析给出了本文的总体结论与建议。经过理论与实证的分析,结合各种类型的对比分析,对黄金价格的各方面进行了较全面的分析,通过研究文献发现对黄金价格的研究已经从多方面进行了开展,比如波动性分析、影响因素分析等等,但是本文是对中国的黄金价格进行的分析,所以由于中国国情的特殊性,本论文的分析与结论都还需要进一步完善。

姚衡. 基于贝叶斯网络的大数据因果关系挖掘[D]. 云南财经大学, 2016.

贝叶斯网络是描述随机变量之间因果关系图的模型，是概率理论、因果推理与图形理论的结合，也是传统的基于数据的统计方法和强调知识的人工智能方法的统一，其重要应用之一是随机变量之间的因果知识表示和推理。贝叶斯网络由结构和参数两部分构成，分别用于定性与定量描述变量之间的因果关系，它具有多功能性、有效性和开放性等特征，能够有效的将数据转化成知识，然后利用这些转化后的知识进行推理，来解决现实世界中的不确定性方面的问题，其有效性已在金融风险分析、信息安全、DNA分析、软件智能化、医疗诊断、系统分析和控制等许多领域得到验证。 　　目前，对于非时序的常规数据，通常采用贝叶斯网络来挖掘其中的因果关系；而时序的常规的单时间序列一般采用格兰杰方法来挖掘其中特定的因果关系，但是这种方法存在诸多问题。随着大数据时代的来临，大数据技术为我们分析问题和解决问题提供了新的思路和方法。与常规数据集相比，在大数据环境下进行数据挖掘将得到更多更全面的信息。未来从大数据中发现因果关系以及在常规数据中挖掘一般因果关系将是一种趋势。 　　为了改善传统格兰杰模型在时间序列因果关系挖掘中出现的弊端，改进并完善因果关系挖掘模型，本文提出了在大数据环境下使用二阶贝叶斯网络模型进行因果关系挖掘。该模型采用最小描述长度（MDL）原理来进行打分。通过对期货样本数据分析，并对原始时间序列进行离散化、属性约简、重构等处理后进行二阶贝叶斯网络模型训练，不仅可以挖掘节点与节点之间的因果关系而且可以发现因果关系之间的联系。 　　本文的主要工作和主要研究成果如下: 　　1.分析对比现有因果关系挖掘模型和贝叶斯网络结构学习方法，选择基于MDL打分原理的贝叶斯网络模型作为本文的研究方法； 　　2.提出了一种新型的贝叶斯网络模型：二阶贝叶斯网络模型。设计出了新型模型构建的方法，并实现了相关算法。 　　3.利用二阶贝叶斯网络推理模型对期货时间序列进行仿真实验，实验不仅得到了单个期货时间序列内部节点之间的因果关系，而且得到了多个时间序列边与边之间的因果关系。 收起

王晓琳. 金融投资者行为及其对股票与期货市场的影响研究[D]. 哈尔滨工业大学, 2016.

关于投资者行为如何影响金融市场，传统金融学并没有给出答案，但在行为金融学框架下，投资者行为是影响证券价格和市场运行的重要因素。大数据的飞速增长正在影响人类日常生活的各个方面，大数据的出现为我们解决人类赖以生存复杂世界的基本问题提供了一个重要的新机会，金融市场是此类定量调查的主要目标。 　　在互联网迅速发展之前，理论界通常利用股票交易量、报纸头条新闻以及广告费用等作为考察投资者的代理变量，谷歌搜索趋势这一用户创造内容平台的出现为投资者提供了便捷的信息获取方式，互联网环境推动了投资者行为与金融市场之间关系的深入... 展开 关于投资者行为如何影响金融市场，传统金融学并没有给出答案，但在行为金融学框架下，投资者行为是影响证券价格和市场运行的重要因素。大数据的飞速增长正在影响人类日常生活的各个方面，大数据的出现为我们解决人类赖以生存复杂世界的基本问题提供了一个重要的新机会，金融市场是此类定量调查的主要目标。 　　在互联网迅速发展之前，理论界通常利用股票交易量、报纸头条新闻以及广告费用等作为考察投资者的代理变量，谷歌搜索趋势这一用户创造内容平台的出现为投资者提供了便捷的信息获取方式，互联网环境推动了投资者行为与金融市场之间关系的深入研究。系统地研究投资者在金融市场上的行为及其产生的影响，能够使投资者更清楚自己在预测股票价格时出现了怎样的偏差，并可能有助于投资者依据这些偏差获取超额收益；有助于更好的理解收益、波动性和交易量之间的互动机理；有利于为监管机构更好地掌握投资者的心理特征，从而为其更有针对性地调控市场提供理论基础。 　　本研究主要内容如下： 　　首先，本文运用标准普尔500股票的谷歌搜索趋势记录这一用户创造内容的平台作为投资者关注的代理变量，探讨投资者关注对分析师盈余预测行为的影响，以及分析师预测行为对谷歌搜索趋势的影响。实证结果发现搜索量与分析师预测修正、分析师总人数和分析师预测分歧相关。与之前的研究结论一致，我们发现分析师盈余预测分歧较大的股票其预期收益将显著下降，即分析师盈余预测分歧效应。然而，当投资者关注度较高时，这种分析师盈余预测分歧效应被显著削弱，研究结果为传统金融学无法阐明的金融异象提供了理论依据。 　　其次，本文首次发现我国股指期货市场流动性日内效应与周内效应，得出了期货市场流动性的日内模式在上午呈现反“J”形变化模式，下午呈现倒“U”型变化模式，期货流动性的周内模式呈现倒“V”形变化模式。实证结果表明中国股票指数期货市场流动性的周内模式和日内模式存在着显著的周期性。本文丰富了对股票指数期货市场流动性的订单驱动型市场的研究，所获得的实证结果有助于股票指数期货投资者更好地了解市场微观结构的投资者行为。 　　再次，采用股指期货高频数据对中国股指期货运行初期期货与现货市场的领先滞后关系进行研究，分别在投资者情绪高涨、投资者情绪低落与投资者情绪平稳三个区间进行实证分析，探讨股指期货价格发现功能在投资者情绪表现不同时是否均能得以实现。本文研究结果修正了之前关于股指期货不具有价格发现功能的结论，股指期货价格发现功能的确立，有助于我国金融市场的完善。 　　第四，基于指令失衡这一关键指标作为投资者交易行为的代理变量，分析投资者交易行为与价格行为之间的关系，具体包括投资者交易行为的影响因素，投资者交易行为与市场流动性变化关系，投资者交易行为对市场波动性与成交量关系的影响。研究结果发现在中国农产品期货市场指令失衡具有收益预测的能力。过度指令失衡引起的价格行为并不符合传统库存理论的解释，中国期货市场的价格行为主要由投机者的羊群效应导致。 　　本文运用多种统计学与计量经济学中常用的分析方法，包括：时间序列回归分析、协整检验、VECM模型估计、格兰杰因果检验与Fama-MacBeth截面回归模型等，具有前沿性、系统性、市场性和监管导向性。

孥磊磊, 吴友琦. 聚类算法在期货价格预测中的应用[J]. 知识经济, 2014(7):73-73.

王娜. 基于大数据的碳价预测[J]. 统计研究, 2016, 33(11):56-62.

为了研究大数据是否有助于预测碳排放权价格,本文讨论了结构化数据和非结构化信息对预测碳价所起的作用.结构化数据选取国际碳现货价格、碳期货价格和汇率,非结构化信息选择百度搜索指数和媒体指数.考虑到当解释变量很多时,平等对待每一个解释变量是不合理的,本文提出了网络结构自回归分布滞后(ADL)模型,在参数估计和变量选择的同时兼顾了解释变量之间的网络关系.实证分析表明,网络结构ADL模型明显优于其他模型,可以获得较高的预测准确性,更适合基于大数据的预测.