



# 量化实盘实践

## 量化实盘系统介绍



讲师：张泽亮

# CONTENTS

## ▶ PART 1

量化交易系统架构介绍

## ▶ PART 2

量化交易实盘管理要求

## ▶ PART 3

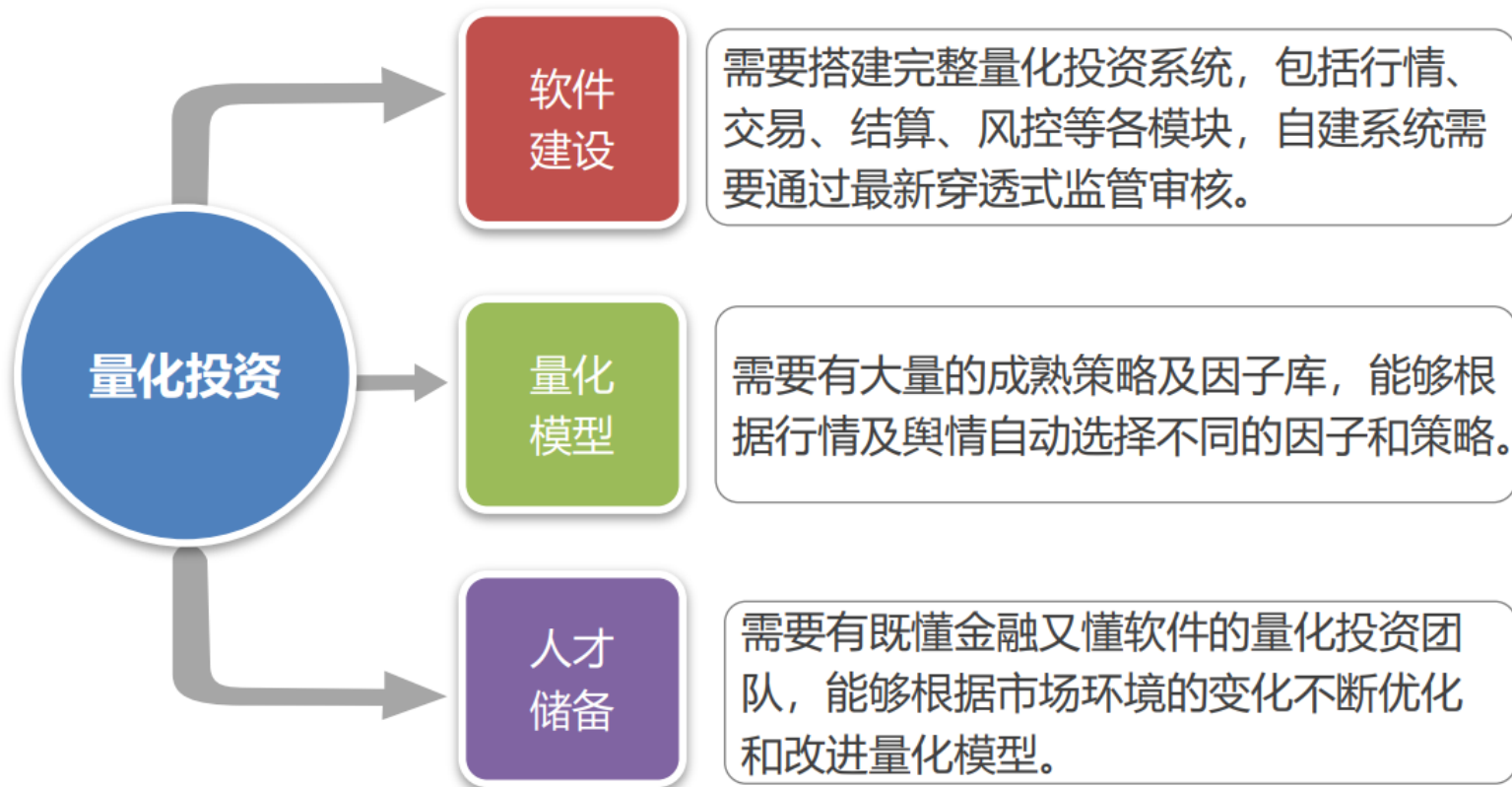
量化交易实盘系统介绍



- 量化交易系统构成：  
包括实时行情、历史行情、事件处理、风控、下单接口、回测统计等模块。
- 量化交易系统平台架构设计  
从输入和输出两端来说，任何交易系统都是通过两条路径和交易所发生交互：
  - 1、接收交易所公布的市场数据
  - 2、发送买卖订单并接收交易所应答

量化交易系统从交易所或者行情提供商获取最新的行情报价，包含信息有当前成交价格、成交量和委托订单队列。通常，交易员还需要参考历史行情数据、其它数据例如基本面信息等做出决定。所以交易系统一般有历史数据库来存储非实时市场数据，也方便金融工具使用数据库。







### 抄（照搬）

学习他人成熟经验  
量化投资并不是新生事情，有大量机构和学术研究成果及产品可供学习及借鉴。

### 改（适配）

融入自有成功经验  
大量雷同的策略往往导致收益降低，有特色的策略才能获取更高的收益。

### 造（探索）

探索前沿投资课题  
打破常规模式，探索前沿课题，利用最新金融科技及金融前沿理论实现弯道超车。

# CONTENTS

## ▶ PART 1

量化交易系统架构介绍

## ▶ PART 2

量化交易实盘管理要求

## ▶ PART 3

量化交易实盘系统介绍



- 量化交易是指利用计算机程序来进行大规模金融交易，由于量化交易可能会引发金融市场的波动，因此，量化交易监管措施的重要性日益凸显。
- 量化交易监管措施应该加强对量化交易活动的监管。通过实施强有力的监管措施，从而限制量化交易者的行为，避免市场的暴涨暴跌，保护投资者免受损失。
- 量化交易监管措施还应该加强量化交易者的责任追究制度。
- 量化交易监管措施还应该加强量化交易者的风险管理能力。
- 量化交易监管措施包括加强监管、实施程序性规则、加强责任追究制度以及加强量化交易者风险管理能力等多种措施，旨在确保量化交易的稳定性和安全性，维护金融市场的稳定运行。





2013年8月16日上午11点06分左右，上证指数瞬间飙升逾100点，最高冲至2198.85点。沪深300成分股中，总共71只股票瞬间触及涨停，且全部集中在上海证券交易所市场。其中22只金融股触及涨停。这一颠覆了以往常识的市场表现，让所有的市场参与者震惊。午后光大证券停牌，并发布公告称，公司策略投资部门自营业务在使用其独立的套利系统时出现问题。该事件也被称为“光大乌龙指”。

### 事件触发原因

策略投资部使用的套利策略系统出现了问题，套利策略系统在并不稳定的情况下就用于实盘交易。光大证券策略系统包含订单生成系统和订单执行系统两个部分。

在发出第三组买入180ETF成分股订单指令后，有24只股票未成交，交易员电脑上演示按下“重下”按钮后，系统显示补单买入24只股票被执行，但事实上程序实际执行的指令为“买入24组ETF一篮子股票”。由于订单生成系统存在的缺陷，导致2秒内，瞬间重复生成26082笔预期外的市价委托订单。

由于订单执行系统存在的缺陷，上述预期外的巨量市价委托订单被直接发送至交易所。光大策略投资部的策略是进行ETF套利，由于套利空间小，对交易速度的要求很高，所以没有使用业内较多采用的恒生交易系统，可能这种市价单处理方式导致验资过程中出了问题，或是相关人员为了提高交易速度而屏蔽掉了部分风控模块的功能。



此次乌龙指事件，从表面上来看是因为光大证券策略部使用的套利系统出现问题，但实际上该事件折射出光大证券风控的失效。虽然光大证券策略部的交易员都来自国内外知名券商，并且有丰富的从业经验，但其风控的失效是此次事件更为深层的因素。我们甚至怀疑该部门系统完全独立于公司其他系统，未置于公司风控系统监控下。

按照正常的交易流程，交易委托单在通过报盘机发送交易所之前要经过一系列风控审核，其中包括两个关键步骤：验资和额度审核。验资是要检查账户上是否有足够的资金来完成这笔交易，限价单的委托价格明确，验资相对容易，而市价单的验资则需要先估算一个成交价格，不同交易系统开发商的估算方式会有不同。额度审核则是要检查该笔交易的额度是否超过了之前风控设置的额度上限。策略投资部主要是追求风险对冲，策略投资部2亿限额的实盘，而实际成交72.7亿，可见交易系统的额度审核功能没有起到任何控制作用。

我们分别从以下几个角度来分析此次风险管理失控的原因：

交易员级：对于交易品种、开盘限额、止损限额三种风控，后两种都没发挥作用。

部门级：根据光大证券的公告，部门核定的交易员当日现货交易额度为8000万元，并在交易开始前由审核人员进行8000万元的额度设定。由此判断，如果订单执行系统风控模块不出问题的话，是可以卡住订单，不发送至交易所的。

公司级：公司监控系统没有发现234亿元巨额订单，同时，或者动用了公司其他部门的资金来补充所需头寸来完成订单生成和执行，或者根本没有头寸控制机制。

交易所：上交所对股市异常波动没有自动反应机制，对券商资金越过权限的使用没有风控，对个股的瞬间波动没有熔断机制。



- 期货量化交易没有资金门槛要求，但是要求交易系统必须进行期货公司测试审核，并填写相关资料，然后期货公司提供实盘量化交易接口，包括行情接口和交易接口两个独立通道。
- 2019年6月19日，期货CTP接口必须升级，以便满足穿透式监管的要求。
- 股票量化交易券商不提供接口，只提供实盘交易系统，其中自带策略编程功能，以Python为主要开发语言。
- 不同券商对量化交易的门槛各不相同，有的要求100万元以上资金才能开通量化交易权限，有的要求50万元即可，还有一些券商没有任何资金要求。

# CONTENTS

## ▶ PART 1

量化交易系统架构介绍

## ▶ PART 2

量化交易实盘管理要求

## ▶ PART 3

量化交易实盘系统介绍



- 传统交易系统提供多种指标供用户定义交易条件，而量化交易系统则使用编程方式实现判断逻辑。
- 传统交易系统强调用户界面友好性和操作便利性（有大量快捷键），而量化交易系统强调系统性能，淡化用户界面友好性和操作便利性。
- 传统交易系统不具有回测及模拟盘功能。
- 传统交易系统只能用户本机运行，而有些券商提供量化交易系统托管服务，直接把量化交易系统部署在交易所机房，交易（下单）更加迅速。



## 一、什么是极速交易系统

量化交易软件的极速交易系统是一种基于高速交易通道、专业软件和网络技术的交易系统，高速交易通道可以实现超快速的数据处理和运算，专业软件可以提供精确的市场数据和分析工具，网络技术可以实现高速稳定的交易通讯。这些技术的结合，可以实现超快速的下单和撤单，以及实时监控市场情况，进行自动化的交易。

## 二、极速交易系统的优势

- 1、高效率。极速交易系统利用计算机技术的高效性，对市场数据进行快速分析和处理，使交易者能够快速获取和分析市场信息，制定有效的交易策略。交易者不再需要花费大量时间和精力去分析市场数据，而是能够在短时间内获取准确的市场信息，从而更好地掌握市场动态，提高交易效率。
- 2、高稳定。极速交易系统通过计算机算法对股票进行自动交易，能够更准确地执行任务，不受人为因素的影响。极速交易系统能够持续不断地执行交易任务，同时，交易系统的算法设计和代码编写也经过了严格的测试和验证，确保系统的稳定性和正确性。
- 3、风险控制。极速交易系统可以帮助投资者更好地控制交易风险。投资者可以通过设置止损、止盈等参数快速规避风险，同时也能够及时调整交易策略，降低交易风险。
- 4、体验好。极速交易系统可以为投资者提供更快、更稳定、更高效的交易体验，满足投资者对于高品质交易服务的需求，能告别下单慢、卡单等情况。



### 君弘君融交易系统

4.8.9.1

提供篮子交易、算法交易、快捷交易、策略交易及君融策略等特色功能

下载



### 国信iQuant策略交易平台

策略交易、Python

下载

软件大小：4176M

软件版本：V3.4

更新时间：2023-06-20

语言：简体中文



- iQuant 学堂:

<https://www.guosen.com.cn/gs/iquant/classroom.html>

- iQuant Python API说明:

<https://www3.guosen.com.cn/guosen/newxwfiles/20190902/1567416998056.pdf>

- iQuant 跨品种套利说明:

[https://www.guosen.com.cn/gs/iquant/iQuantClassRoom\\_detail.html?channelid=000100010002000300020003&id=31485](https://www.guosen.com.cn/gs/iquant/iQuantClassRoom_detail.html?channelid=000100010002000300020003&id=31485)

- iQuant 系统操作演示



# Thank you!





# 量化实盘交易实践

## 量化实盘与智能投研



讲师：张泽亮

# CONTENTS

## ▶ PART 1

### 智能投研简介

## ▶ PART 2

### 知识图谱在投研中的应用

## ▶ PART 3

### 智能投研案例实战



- 智能投研是指基于人工智能、大数据、云计算等现代科技手段，获取、处理、分析海量相关数据，生成投资观点和报告。其中关键的人工智能技术包括自然语言处理（NLP，包括自然语言理解NLU、自然语言生成NLG等）、知识图谱等，人工智能提供算法和工具，大数据提供素材，云计算则提供算力。
- 与智能投顾的侧重点不同，智能投研的主要目的是深层次改造投资研究流程，提升资产管理能力和风险评估能力，创造AUM（资产管理规模）高质量增长。





- 传统投研业务流程需要投研人员通过各种渠道寻找数据和判别信息，依靠个人的知识储备和历史经验对数据进行加工整合，并通过逻辑分析和数理建模，最终以报告的形式输出观点并指导投资决策。整个过程对投研人员信息搜集、数据处理、逻辑分析和知识结构都提出了较高的要求，而且会花费大量时间，有些窗口时间较短的市场投资机会往往稍纵即逝。
- 即使部分金融数据服务商目前已经提供了相应的数据库产品，但数据的颗粒度和产业链关联性仍难以满足多元需求。同时依靠人工分析，最终的研究结果还容易受到个人情绪以及知识体系等方面的影响，稳定性和可控性不高；而且一旦较为资深的投研人员发生离职等工作变动，投研工作也会遭受较大影响。
- 相比之下，智能投研可以在整个业务流程的各个环节运用金融科技手段节省大量时间，提高投研工作效率。而且与智能投研所应用的技术手段相比，人类对非线性关系难以直观理解，而机器学习方法更擅长从复杂的历史数据中提炼非线性关系，达到投资决策胜率的目的。同时，机器的情绪控制明显好于人类，可以提高分析结果的稳定性和可控性。另外前期开发成本虽然可能较高，但系统的复制推广和运营成本极低，也可以降低少数个别人员离职变动带来的风险。



- 传统投资研究流程主要包括四个环节：数据收集、数据处理、分析研究、结果输出。而智能投研的业务流程从本质上看并没有显著的区别，最重要的一个区别是综合应用金融科技手段赋能前述四个环节，实现从数据输入到策略输出的过程，不仅可以大幅提高投研人员的工作效率和投资水平，而且可以降低决策成本，排除隐藏风险。
- 目前智能投研的主要应用和服务对象为证券公司和基金公司等专业投资机构，正在逐步向银行、保险、个人等拓展，为投资研究过程提供智能化辅助服务。
- 简单分类来看，智能投研应用的核心技术主要包括知识图谱和机器学习。其中知识图谱主要用来辅助搭建投研的语义网络，直观描述实体和概念间的关系属性；机器学习作为统计学和计算机科学概念和研究的混合领域，最终目标是探索投资策略的自开发和自学习。



- 知识图谱作为知识地图的高级表现形式，是用来描述各种实体和概念以及它们之间关联性的数据库，本质上是一种利用信息可视化技术构建的语义网络。知识图谱通过将各类实体之间的关联性以网络的方式进行直观呈现后，当某个实体或关系发生变化时，还可以通过分析得出可能的影响和后果，从而辅助投研人员更直观地掌握全面信息并加深研究层次。
- 金融领域的知识图谱节点包括企业、产品、证券资产等各类实体，节点之间的关系主要包括股权关系、负债关系、担保关系、上下游关系、竞争或合作关系、价格传导联动关系等。
- 智能投研领域知识图谱构建过程中的关键部分是识别主体和构建关系，这是整个知识图谱的“上层建筑”，考验投研人员对相关知识的积累和沉淀。其中识别主体需要从公司公告、研究报告、新闻报道、视频图像等海量的非结构化数据中批量化自动提取实体信息，这里可能需要用到情感分析、自然语言处理（NLP）等技术，在对原始数据经过知识提取或数据整合的方式进行转换后，加入数据模型，而构建关系则需要机器学习等技术手段加以实现。



- **智能投研智能水平有待提高。**从智能投研所应用的核心技术理论来看，通过知识图谱或者机器学习方法进行决策的前提都是基于历史数据，而对于新出现的事件，例如对“黑天鹅”政治经济事件的发生对资产价格影响的判断力仍有待考验。目前智能投研自主推理和思考能力仍有待提高，仍需依靠投研人员的分析判断。
- **系统无法判断市场极端情况。**底层数据的数量和质量也会影响智能投研的效果，同时构建模型时往往会在数据处理的过程中剔除极端数据，以追求时间序列的平稳性，但由非极端数据建立的模型可能在解释或应对极端情况下出现错误，而在现实的市场交易中如何应对极端情况往往是决定盈利或规避风险的关键时刻。
- **智能投研难克服同质化市场。**目前我国金融市场的可交易标的种类还不够丰富，资产交易规则的灵活性整体也不及海外成熟的金融市场，无法支持更多种类的算法模型交易，导致主要策略框架的数量相对有限，容易形成同质化策略。同质化策略给出相近的投资建议后，导致羊群效应，可能在短期内对市场产生较大冲击。



# CONTENTS

▶ **PART 1**

智能投研简介

▶ **PART 2**

知识图谱在投研中的应用

▶ **PART 3**

智能投研案例实战



- 知识图谱 (Knowledge Graph) , 在图书情报界称为知识域可视化或知识领域映射地图, 是显示知识发展进程与结构关系的一系列各种不同的图形, 用可视化技术描述知识资源及其载体, 挖掘、分析、构建、绘制和显示知识及它们之间的相互联系。
- 具体来说, 知识图谱是通过将应用数学、图形学、信息可视化技术、信息科学等学科的理论方法与计量学引文分析、共现分析等方法结合, 并利用可视化的图谱形象地展示学科的核心结构、发展历史、前沿领域以及整体知识架构达到多学科融合目的的现代理论。它把复杂的知识领域通过数据挖掘、信息处理、知识计量和图形绘制而显示出来, 揭示知识领域的动态发展规律, 为学科研究提供切实的、有价值的参考。
- 在金融领域, 面对大量的上市公司, 投研人员很难对所有上市公司的经营情况与业绩预期分析得较为透彻。如果投研系统可以将上市公司基本情况、产业或行业以及公司经营情况整合到知识图谱中, 同时将零碎的、多渠道实时舆情信息、券商研报等纳入知识图谱管理, 对于其中特定公司事件或者行业新政产生的影响进行测算, 将会得到一幅上市公司全息动态图。这对于投研人员来说会起到事半功倍的效果, 而且消息的实时性将得到很大的保障。



- **知识图谱**是由节点和关系所组成的图谱，为真实世界的各个场景直观地建模。通过不同知识的关联性形成一个网状的知识结构，看起来就如同一张图谱。人工智能的价值在于让机器具备认知能力，而构建知识图谱这个过程本质在于，让机器形成认知能力，从而去理解这个世界。在构建知识图谱的基础上，基于一定的关系规则进行推理，信息经由关系（即知识图谱中节点之间的边）进行传导，可以挖掘出知识图谱中隐藏的信息。
- **企业知识图谱**：企业知识图谱是将每一家上市公司的股权关系以及控股股东、企业主营业务及对外投资情况、公司专利技术、涉及司法诉讼情况、公司行业地位等情况按照知识图谱的数据结构表示出来，将控股股东的企业知识图谱分别刻画清楚。
- **行业知识图谱**：行业知识图谱的构建需要采集行业上下游的数据，这比较考验投研人员的专业能力以及数据获取与分析能力。行业知识图谱构建后需要及时更新、动态调整。
- **市场实时信息图谱**：是在前两个图谱的基础上，进一步拓展而形成的一种辅助决策方式。通过采集市场相关信息，运用自然语言处理技术进行加工处理后，经过多渠道相互印证，提取正确的市场信息，对其中不同类型的消息设置不同的权重。将这些市场信息输入到知识图谱中进行推导、计算，得到市场上所有的消息对相关企业的产生的影响，并将影响以及影响力传导的路径、大小以图形的形式展示出来。



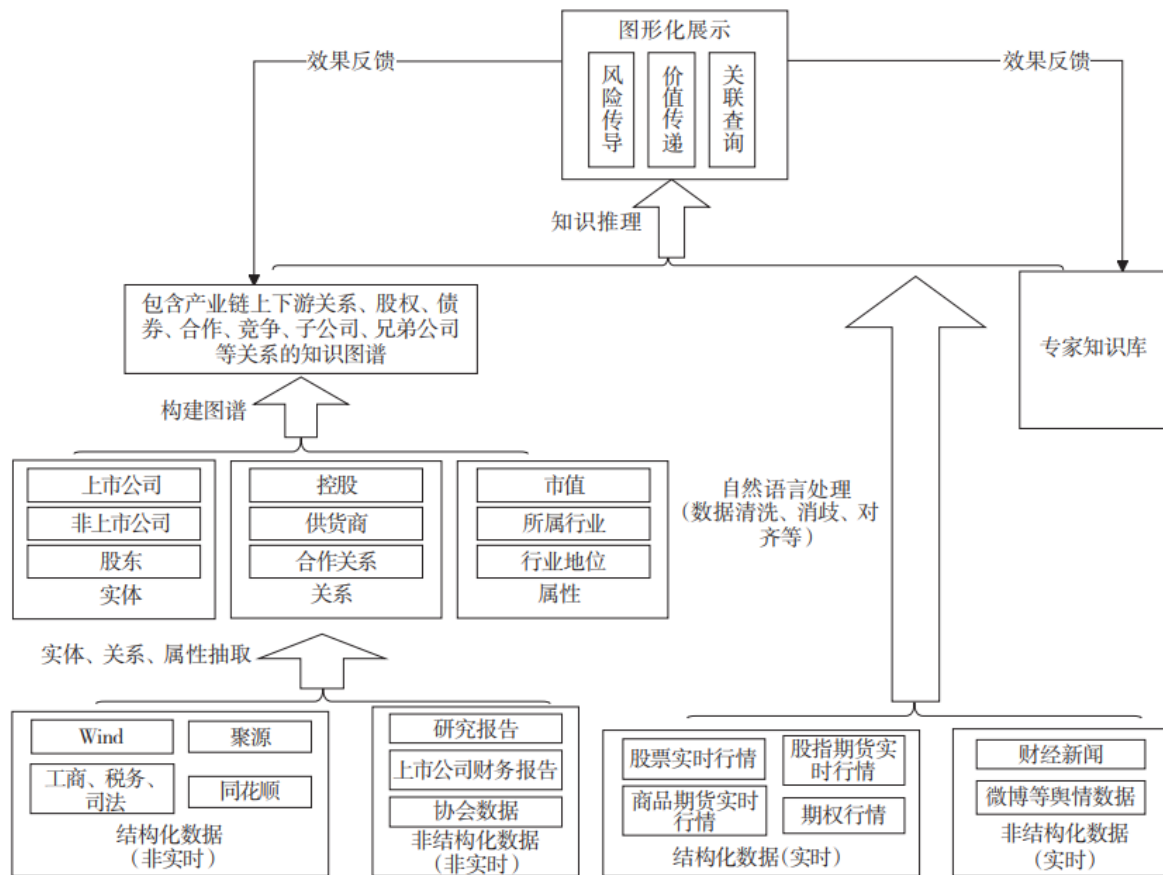
- **投资方向分析：**目前已经有一些成型算法能够处理行业、产业上下游等关联信息，根据知识图谱中相关信息来分析投资方向，选择投资标的，并且根据舆情信息，通过知识图谱等实时计算发现潜在风险标的并预警。
- **集团公司的识别与构建：**对于大型企业集团而言，存在集团公司、一、二、三级子公司等多层架构，且子公司内部可能也具有集团公司的多级架构，还可能分别处于不同的行业，每层级的子公司都有可能成为证券的发行人。单纯通过人工识别难以识别企业集团的全貌，低估下属企业的风险。以股权关系将企业集团户以一个包含多级数据源的知识图谱形式构建出来，则可以有效推进集团户信息归集，将其应用于投资、质押、授信等多个投资研究应用场景中。
- **业务关联关系识别与预警：**股权关联关系可以依靠工商信息等进行检索、处理，但业务关联关系则分散在多个数据源中，难以获取，需要依靠技术化的手段进行处理。结合上下游的供货、销售关系构建关联知识图谱，可以快速检索其自身及其关联企业的上下游关联企业，在节点企业出现突发事件时快速对该企业及上下游企业进行识别并采取相应措施。此外，知识图谱还可与信息一致性检验等方法相结合，就信息真实性进行核验。
- **市场走势归因：**利用程序化方法对历史上相似的市场走势进行归因，寻找共同的驱动因素，并形成逻辑推演的知识图谱。在驱动因素再次出现时，可以依次验证各个逻辑节点上的驱动因素是否符合规律，并进一步推断市场走势。



- 构造知识图谱的技术路线一般可以分为两种，一种是自顶向下，另一种是自下而上。其中自顶向下构建知识图谱的方案是从数据中找到本体和关系等信息，将结构化数据源添加到基础知识中。自下而上构建知识图谱的方法是基于公共资源模型收集数据，获得可靠性高的信息并添加到基础知识中。
- 本课程所构建的知识图谱是基于已有的非实时数据与抽取的实时数据自下而上构建的用于智能投研的知识图谱，主要分为三个步骤：知识抽取、知识融合以及知识加工。知识抽取指的是从各种类型的数据源中提取出实体、属性以及实体间的相互关系，并形成本体化的知识表达。由于知识的来源不同，同一个实体可能对应多种表达，或者同一种表达也可能对应不同的实体。
- 知识融合指的是对多个来源的知识进行整合，消除矛盾和歧义。知识加工指的是对于融合后的知识，经过质量评估环节后将符合条件的内容作为新增内容加入知识库中。



# 知识图谱建设路径



最左侧部分是基于公司已有的非实时数据所构建的基础架构知识图谱。该图谱构建完成后，投研人员就可以看到上市公司及非上市公司与投研相关的市值、所属行业、行业地位等信息以及各个公司之间的相互关系，每一种关系间都有表示相关性的权重值。

中间部分原始数据主要通过对股票、期货、商品等交易所数据或者爬虫收集的财经信息、社交平台舆情数据等实时获取。这部分数据对于跟踪市场情绪变化以及应对市场突发事件具有重要的意义。

最右侧部分是专家知识库，投研系统很多知识的补全需要依靠行业专家或者投研人员来完成，尤其对于宏观政策、行业趋势、企业运营等信息的分析，需要专业人员将结果汇集到专家库中。

将这三部分图谱或者数据整理完成后，就可以基于一定的知识推理算法来学习埋藏在数据里面的业务价值，并将最终生成的结果进行反馈，以便于后续的投资研究与知识推理。



- 知识图谱存储方式主要有两种形式：
- RDF 存储格式和图数据库。RDF 存储格式的优势在于数据易发布以及共享，图数据库的优势在于高效的图查询和搜索。RDF 存储格式以三元组的方式来存储数据而且不包含属性信息，但图数据库一般以属性图为基础表示形式，所以实体和关系还包含属性。其中属性包含内容相对比较广泛，例如节点所属行业、概念、业绩等。
- 在智能投研方面，使用图数据库相对比较合适，即采用实体、关系、属性的形式来存储。对于数据存储，有一些数据可能存放到关系型数据库更适合数据查询与分析，对于这样的数据就可以存放到关系型数据库里面。例如节点属性、宏观信息、行业信息与公司信息以及两者之间的对应关系、每个行业的龙头等信息都可以存放到关系型数据库中。对于其中的宏观信息、行业信息一般需要借助专家经验来构建，形成整体的框架，而且这些信息对于整个投研体系具有基础性、框架性影响。其中宏观信息主要包括GDP、PMI、通胀、货币与信贷等信息。公司与产品信息大多可以通过人工智能算法来采集，丰富投研信息。



- 从算法的角度来讲，有两种不同的传导过程：一种是基于规则的，另一种是基于概率的。
- 基于规则的传导计算是指基于某些事件或者政策，根据知识图谱中设置的权重以及相应的规则，从知识图谱中提取一些特征，而且这些特征一般基于深度的搜索。这些规则一般来说是知识图谱中提前设置好的要素，不仅设置节点数据，而且关系和属性数据都要进行设置。
- 基于概率的传导计算运用社区挖掘、标签传播、聚类等技术。这种方式更加开放与智能，是基于数据归纳出来的一个结果，推理方式需要人工辅助验证是否能够真正应用于实践。基于概率的传导计算不需要人为地去定义规则，但需要大量的数据进行训练学习。
- 鉴于目前人工智能技术现状，基于规则的方法论在应用中占据主导地位，但随着数据量的增加以及方法论的提升，基于概率的模型也会逐步带来更大的价值。





- **应用场景层面：**基于知识图谱的发展趋势，结合金融行业投资研究业务的需要，知识图谱在未来或将向更加集成化、智能化的方向发展。多个知识图谱的有机集合或将发挥更大作用。例如，基于股权关联关系的集团户知识图谱与基于业务关联关系的知识图谱可以有机结合，在资产质量分析、关联交易识别、财务造假辨别、现金流转记录等方面发挥更大作用。借助知识图谱可以明确应收账款的债务方是否为同一控制下的关联方，以及上下游客户是否有重合等，借助知识图谱的分析可以预防欺诈与造假。在走势归因的基础上，基于当前市场与走势判断，智能化地提出投资操作建议，并根据策略进行业绩回测，实现智能投研。
- **算法层面：**一是多信源信息推理算法。随着移动通信技术的快速发展，人们可以随时上传文字、语音、视频和图像到互联网上。如何有效利用这些丰富的信息成为一个重要且有挑战性的问题。目前，多信源信息推理算法已经开始显现出潜力，但还是在发展的初步阶段，未来可以设计更好的算法模型，充分利用多信源的信息。二是多语言知识图谱推理。很多知识图谱都有多语言的版本，从多语言的网站上抽取信息。多语言知识图谱推理在机器翻译、跨语言剽窃探测和信息抽取方面起到重要的作用。但目前只有少数图谱算法是在多语言上做推理，未来多语言推理将是知识图谱发展的一个重要方向。



- Neo4j是一个开源的、功能完备的高性能图形数据库，它结合了传统关系型数据库的高性能和高可用性，和现代图数据库的复杂关系功能。Neo4j具有高效的查询性能，并且具有简单的图遍历API，可以帮助您轻松使用图数据构建和管理复杂的利用程序。
- 熟悉Neo4j的数据模型对高效使用Neo4j是必要的。Neo4j的数据模型将数据存储存储在节点和关系之间，而不是像传统关系型数据库一样存储在表和列之间。Neo4j支持属性绑定，这意味着可以把属性（字段和值）附加到节点和关系，以实现更加准确的查询和快速访问数据。
- Neo4j也能够使用区别的语言来创建，保护和操作图形数据库，这使得Neo4j成为强大的基础设施，可为利用程序开发提供高度灵活的编程模型。这使得用户可以通过一系列API来快速创建和管理Neo4j实例。Neo4j 提供Java API、.Net API、Go API、Python API、REST API和Javascript API等多种API。



- 下载Neo4j: <https://neo4j.com/download-center/#community>
- 下载安装JDK: <https://www.oracle.com/java/technologies/downloads/#jdk20-windows>
- 环境及配置
- 启动: neo4j console
- 停止: neo4j stop



- MATCH(n) DETACH DELETE n     //删除所有图
- 创建节点:  

```
CREATE (n:Industry{name:'百货',code:'retailing'});  
CREATE (n:Stock{name:'上海九百',ts_code:'600838.SH'});  
CREATE (n:Stock{name:'杭州解百',ts_code:'600814.SH'});
```
- 创建关系:  

```
MATCH (a:Stock {name:'上海九百'}), (b:Industry {name:'百货'}) MERGE (a)-[:INDUSTRY]->(b)  
MATCH (a:Stock {name:'杭州解百'}), (b:Industry {name:'百货'}) MERGE (a)-[:INDUSTRY]->(b)
```
- 显示节点及关系:  

```
MATCH (a)--() RETURN a
```

# CONTENTS

▶ **PART 1**

智能投研简介

▶ **PART 2**

知识图谱在投研中的应用

▶ **PART 3**

智能投研案例实战



- 供需平衡

产量、月末开工产能、国内消费量、进口量、出口量、库存增量、供需关系评价对铝价的影响方向、供需关系影响幅度、供需关系影响力权重

- 交易心理

市场状态（抛售/正常/惜售）、大企业态度、交易心理评价、交易心理对铝价的影响方向、交易心理影响幅度、交易心理影响力权重

- 市场性质

竞争状态（充分竞争/垄断竞争/寡头垄断/完全垄断）、竞争状态对铝价的影响方向、竞争状态影响幅度、竞争状态影响力权重

- 成本

氧化铝涨跌幅、预焙阳极涨跌幅、电价涨跌幅、其他因子、行业加权平均成本增幅、行业加权平均成本对铝价的影响方向、行业加权平均成本影响幅度、行业加权平均成本影响力权重



- 环保政策

碳达峰与碳中和（碳政策对生产的影响、对铝价的影响方向、影响幅度、影响力权重），大气2+26城市（是否影响对生产的影响、对铝价的影响方向、影响幅度、影响力权重），固废（是否影响对生产的影响、对铝价的影响方向、影响幅度、影响力权重）

- 产能政策

指标置换（减停产/新立项或重启、对铝价的影响方向、影响幅度、影响力权重），各省地区的地方政策（减停产/新立项或重启、对铝价的影响方向、影响幅度、影响力权重）

- 收抛储

收抛储可能性、收抛储数量、收抛储对铝价的影响方向、收抛储影响幅度、收抛储影响力权重



- 产业宏观

房地产数据（新房屋开工面积、新房屋竣工面积、对铝价的影响方向、影响幅度、影响力权重）

汽车行业数据（产量、产量同比、产量环比、销量、销量同比、销量环比、对铝价的影响方向、影响幅度、影响力权重），

电力行业数据（新增220千伏及以上线路长度、同比、环比、影响幅度、影响力权重）

- 宏观经济

中国货币政策（加/降息预期、存款准备金率预期、对铝价的影响方向、影响幅度、影响力权重）

美联储货币政策（加/降息预期、对铝价的影响方向、影响幅度、影响力权重）

美元指数（美元指数对铝价的影响方向、影响幅度、影响力权重）





# 财经数据及处理

MySQL Workbench

47.97.205.89 - Warning - not...

File Edit View Query Database Server Tools Scripting Help



Navigator

SCHEMAS

Filter objects

Tables

- F表观消费量
- F产能产量
- F工厂库存
- F价格
- F进出口
- F市场库存
- F下游行业数据
- F消费结构
- F行业成本
- F行业利润
- 表观消费量
- 产能产量
- 工厂库存
- 价格

Administration Schemas

Information

Table: F产能产量

Columns:

年月	varchar(7) PK
产能	double
企业有效产能	double
产量	double
开工率	double
企业数	double

Object Info Session

Query Completed

RiskClass sys\_application

table\_mapping

api\_qixin

api\_qixin

api\_qixin

D价格分析

Limit to 1000 rows

1 • SELECT \* FROM 大宗商品价格预测.D价格分析;

Result Grid

Filter Rows:

Export:

Wrap Cell Content:

	年月	价格	产量	进口量	出口量	表观消费量	产能	企业有效产能	c产量	开工率	企业数	库存量	成本
▶	2019-01	13336.363636363636	2961400	19546.59	228.13	2980718.46	46786000	38691000	2961400	0.7596	130	139300	1464
	2019-02	13425.333333333334	2653700	3189.66	0	2656889.66	47030000	39184000	2653700	0.6771	130	131150	1466
	2019-03	13805.714285714286	2982600	1917.4	21214.76	2963302.64	47355000	39184000	2982600	0.7558	130	142500	1441
	2019-04	13955.238095238095	2929900	1551.61	20667.92	2910783.69	47365000	38994000	2929900	0.7423	130	113600	1427
	2019-05	14225	3003700	3680.19	4356.87	3003023.32	47515000	38859000	2992700	0.7558	130	177500	1423
	2019-06	13974.21052631579	2907300	2913.54	20.55	2910192.99	47495000	38839000	2907600	0.7346	130	193000	1430
	2019-07	13805.217391304348	3019500	12708.38	0	3032208.38	47640000	39234000	3019500	0.7606	130	211400	1392
	2019-08	14136.363636363636	2959600	6236.22	16.43	2965819.79	48190000	39434000	2959600	0.737	130	271300	1371
	2019-09	14324	2879300	5170.8	20.53	2884450.27	48210000	39454000	2879300	0.7167	130	219500	1344
	2019-10	13958.333333333334	2966000	4423.67	5116.21	2965307.46	48240000	39454000	2966000	0.7378	130	224800	1348
	2019-11	14020	2918400	5418.27	16255.22	2907563.05	48415000	39764000	2918400	0.7233	130	277250	1347
	2019-12	14260	3021700	8054.04	7401.1	3022352.94	47161000	40229000	3021659	0.7689	130	210750	1332

D价格分析 1 x

Read Only

Output





- 利用多元线性回归模型预测电解铝价格，影响因子有产量、进口量、出口量、表观消费量、产能、企业有效产能、开工率、企业数、库存量、成本、毛利、毛利润率、铝合金、铝材。
- 误差评估：误差预测指标判决系数 $R^2$ 得分为0.985，靠近1，精确度较高。
- F检验：F分布理论值为1.74，模型计算F统计值为640，拒绝原假设，多元线性回归是显著的。
- t检验：根据回归系数的t统计量值对应的概率值p，企业有效产能、开工率、企业数、库存量、成本、毛利六个影响因子的p值小于0.05，说明此六个影响因子通过t检验，是电解铝价格的重要影响因素。
- 正态性检验：偏度1.152，峰度5.591，通过定性图形法（如下QQ图法）对比正态分布的分位数和实际分布的分位数，散点近似在直线附近，价格变量近似服从正态分布。
- 线性相关性检验：计算皮尔逊相关系数，以及绘制散点图，可知产量、表观消费量、产能、企业有效产能、企业数、库存量、毛利、毛利润率与铝价存在较高的线性相关关系
- 由于进口量和出口量与表观消费量直接相关，产能与企业有效产能直接相关，会带来多重共线性问题，因此去掉这三个因子，判决系数 $R^2$ 得分仍为0.985。



```
import statsmodels.api as sm
import pymysql

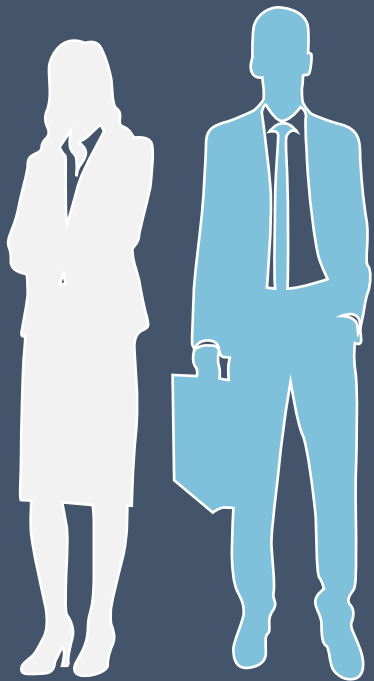
dbHost='47.97.205.89'
dbUser='test'
dbPassword='Aa123456'
dbName='大宗商品价格预测'
dbCharset='utf8'
con = pymysql.connect(host=dbHost, user=dbUser, password=dbPassword, db=dbName, charset=dbCharset) # 打

sql='select * from D价格分析 order by 年月'
profit=pd.read_sql(sql,con)
print('get data')

for a in ['2021-01','2021-02','2021-03','2021-04']:
    train=profit[profit['年月']<a]
    test=profit[profit['年月']==a]
    #根据train数据集建模
    model=sm.formula.ols('价格~产量+进口量+出口量+表观消费量+产能+企业有效产能+开工率+企业数+库存量+成本+毛利',
                        test_x=test.drop(labels='价格',axis=1))
    pred=model.predict(exog=test_x)
    diff=pd.DataFrame({'Prediction':pred,'Real':test['价格']})
    print(a,'diff:',diff)
```

# Thank you!





# 量化实盘交易实践

## 量化交易与智能风控



讲师：张泽亮

# CONTENTS

▶ **PART 1**

智能风控综述

▶ **PART 2**

KYC 尽职调查实战

▶ **PART 3**

财务舞弊识别

▶ **PART 4**

股价操纵与水军识别



智能风控是利用大数据、人工智能技术和科学决策方法，通过自动化预测、评级、决策等方式，替代风控中的人工操作，完成人工无法完成的任务，提高决策精度和效率，最终提高风控能力的综合体系。

智能风控是一个综合体系，并不是单一的某个算法或者技术。智能风控方法层面包含预测模型搭建方法、数据挖掘方法、风控策略制订方法等，通过一系列方法，可以构建智能风控的基本架构；

工具层面包括机器学习、深度学习和关系网络等智能算法，这些智能算法是方法体系中的核心部件，也是智能风控得以展现智能的关键；

实现方式是建立自动化体系，运用智能风控的方法论搭配智能算法构建智能决策的核心，结合工程技术实现自动化的风控决策和智能交互，自动进行模型迭代；

应用环节包括风控流程的全过程，都融入了智能风控的元素；从目标方面来讲，智能风控提高了风险控制精准度和效率，将风险控制在目标范围内，最终提高企业的盈利能力。

# CONTENTS

▶ **PART 1**

智能风控综述

▶ **PART 2**

KYC 尽职调查实战

▶ **PART 3**

财务舞弊识别

▶ **PART 4**

股价操纵与水军识别





## 什么是KYC尽职调查？

KYC的英文全名是Know Your Customer，中文直译过来的意思是：认识你的客户。所以通常会被简化为“客户尽调”。它的内在含义是指为了阻止、发现和报告洗钱活动的一系列程序。目前几乎所有国家地区的政府监管机构都规定在与金融和投资运作有关的工作中需要进行KYC。

投资领域的尽职调查又称谨慎性调查，是指投资人在与目标企业达成初步合作意向后，经协商一致，投资人对目标企业的历史数据和文档、管理人员的背景、市场风险、管理风险、技术风险和资金风险做一个全面深入的审核。



- 法务尽调：收集了解企业工商资料，掌握企业股权结构及对外投资等信息；
- 财务尽调：了解企业财务状况，分析企业经营风险；
- 合规尽调：掌握企业信用信息，评估企业信用风险。



- 法务尽调：收集了解企业工商资料，掌握企业股权结构及对外投资等信息；
- 财务尽调：了解企业财务状况，分析企业经营风险；
- 合规尽调：掌握企业信用信息，评估企业信用风险。



- 工商基本信息：工商信息、分支机构、工商变更；
- 对外投融资：对外投资、融资情况、股权质押；
- 核心人员：董、监、高人员名单及外部任职情况；
- 经营许可：资质及认证，行业许可证；
- 无形资产：商标、专利、软件著作权。



- 财务审计报告：财务三大报表、附注；
- 财务分析：盈利能力、偿债能力、营运能力和发展能力分析；
- 行业分析：行业对比分析。



- 企业风险：经营异常、纳税异常、涉法涉诉、失信被执行、限高；
- 企业社会责任：行政处罚、环保处罚、负面舆情；
- 国际制裁：联合国制裁、美国制裁、欧盟制裁。



- 数据来源：工商及信用信息-启信宝、财务信息-在线数据，国际制裁-百度；
- 实战案例：中国打印机制造商纳思达和新疆中泰化学；

**Ninestar**  
纳思达股份有限公司



**中泰化学**  
ZHONGTAI CHEMICAL

# CONTENTS

▶ **PART 1**

智能风控综述

▶ **PART 2**

KYC 尽职调查实战

▶ **PART 3**

财务舞弊识别

▶ **PART 4**

股价操纵与水军识别





- 财务舞弊是指有目的、有预谋、有针对性的财务造假和欺骗，从而导致会计报表产生不真实反映的故意行为。
- 财务舞弊特点
  - 1、财务舞弊是故意、有计划的欺骗行为；
  - 2、财务舞弊的实施者为管理层；
  - 3、财务舞弊对象则为会计数据。



## 动机 分析

### 发行上市

企业如果获得上市，那就意味着可以得到巨大的融资，及时缓解企业资金压力。由于上市有连续三年持续盈利的要求，因此企业最常用的舞弊手段是虚增利润。会计舞弊金额非常巨大，年均会计舞弊额为2.6亿万元/公司/年。

### 融资估值

对于要增发融资的公司，要求最近三个会计年度加权平均净资产收益率要高于6%的限定。因此，上市公司就开始根据企业经营状况略加调整，而那些根本无法满足条件的就冒险采取会计舞弊来达到目的。

### 避免退市

上市企业想要保证自己的良好声誉，为以后的融资活动创造条件，就会尽可能对公司的亏损情况进行隐瞒。那些被ST的公司，假如在第三年还不能通过正常渠道让公司的标志去除的话，就有可能选择会计作弊的形式来保证公司不会被暂停上市。



舞弊动机  
理论



财务舞弊手法  
和操纵方法



企业财务风  
险理论分类



五大能力指  
标体系

专家规则  
学院研究成果+行业实践经验

大数据机器学习  
统计方法、logit回归、增强型决策树

财务异常指标展示

异常指标业务解释

验证和算法迭代更新流程

系统软件工具

优化方向和迭代流程



### 1、远高于同行业的毛利率

一是毛利率明显高于或者低于同行业上市公司的水平，或者变动趋势与行业不一致；二是无视经济周期的影响，增长异常平稳；三是毛利率稳定或上升时，应收账款大增，存货周转变缓（营运比率下滑但是毛利率未表现出下滑趋势）

### 2、经营现金流

一是净现比：经营活动现金流量净额/净利润应该大于1；二是收现比，即经营活动现金流入/营业收入大于1，在1.16左右。波动过大也是异常。（但现金流异常的原因有，现金未收回来或者买入过多存货，具体分析）

### 3、存贷双高

货币资金和短期借款都很高，简称存贷双高。（大量资金却不用于还贷，资金利用率低，可能发生股东占款的现象，也可能是子公司过多存贷不均导致合并报表时出现双高问题）

### 4、隐瞒关联交易或严重依赖关联交易：通过关联交易拉升业绩

### 5、海外收入：国外营业收入以及毛利率过高

### 6、实控人股票质押比例过高：质押多为经营融资，过高易有庄股嫌疑，操控股价

### 7、分红少：净利润虚增造成分红少



上市公司财务问题的因素归纳为以下八个方面的特征指标:

- (1) 财务业绩指标: 资产周转率、应收账款周转率、销售收入增长比率、销售毛利比率、造假前1-2年是否亏损、可操纵性应计项、其他应收款、营业外收入、流动比率、资产负债比例;
- (2) 公司治理指标: 控制权是否是国企、机构投资者持股数、前五大股东股权集中度、独立董事比例、董事长和总经理是否合一、董事会规模、监事会规模;
- (3) 地区发展水平和行业特征: 公司所在地经济发展水平、市场化程度、是否处于造假高发行业;
- (4) 管理层特征及激励: CEO 年龄、董事长年龄、CEO 学历、董事长学历、CEO 任期、董事长任期、管理层持股数;
- (5) 市场信息: 经市场调整后的股票年收益率、股票月收益率波动率、经市场调整的股票换手率、股票月换手率波动率、是否再融资、股市周期、投资者情绪;
- (6) 审计信息: 审计意见、是否变更审计事务所、审计事务所规模、审计事务所是否是前十大;
- (7) ROE 指标
- (8) 特殊交易: 关联次数、关联交易总量与总收入之比、是否有经营租赁。



上市公司财务造假十大典型案例：

资源来源：<https://baijiahao.baidu.com/s?id=1760418602370048361&wfr=spider&for=pc>

- 一、5年虚增营收512亿，\*ST凯乐成2023年退市第一股
- 二、ST金正：虚构合同,空转资金，三年累计虚增收入230亿
- 三、证监会罕见定性，康美药业系统性财务造假
- 四、广州浪奇：自曝5.72亿元存货“不翼而飞”？两年虚增存货20亿，被罚1405万
- 五、康得新：四年虚增利润总额高达115亿元，实控人终身市场禁入
- 六、宜华生活：设“双系统”数据造假，新证券法下收大额罚单
- 七、雅百特：虚构跨国生意，跨境财务造假
- 八、豫金刚石：通过自有资金循环、虚假出售亏损子公司等方式虚增利润
- 九、龙力生物：删改、伪造大量会计凭证，造假方式“简单粗暴”
- 十、獐子岛：扇贝屡次“跑路”，一起“寅吃卯粮”、调节利润的恶性舞弊案件

# CONTENTS

## ▶ PART 1

智能风控综述

## ▶ PART 2

KYC 尽职调查实战

## ▶ PART 3

财务舞弊识别

## ▶ PART 4

股价操纵与水军识别



股价操纵是指某些股票投资者为了获得巨额盈利，通过控制其他投资者具有参考意义的股票投资信息，控制未来股票价格走势的行为。操纵股价不一定是机构，也可能是公司内部管理人员、大股东、隐形股东、个人等，只要集中资金优势、内部信息优势、管理优势等，其他投资不具备的特殊条件，达到使股票价格可持续的改变，为自己获得不正当的巨额利润的行为均属于操纵股价。

操纵股价者，机构较为常见。机构操纵股价一般会利用媒体舆论传递信息，或通过其他特别权威组织、个人、投资对象企业内部人员发布信息等炮制对其有利的信息控制股票价格走势。但是对于大多数的操纵股价者而言，做到信息面的控制显然不太现实，而且太过显眼容易被关注，失败风险高，不是所有操纵股价者能够做到的。





股价操纵的方法有很多种，其中比较常见的包括以下几种：

1. 假消息操纵法：通过发布虚假消息或者散布谣言来影响市场情绪，从而引起股价的波动。这种方法常常被用于短线操作，一旦消息被揭穿，股价就会迅速回落。
2. 集中买卖操纵法：通过大量买入或者卖出股票来影响市场供需关系，从而引起股价的波动。这种方法需要有足够的资金实力和市场影响力，一旦操作失误，可能会造成巨额亏损。
3. 操纵股价指数法：通过大量买入或者卖出股票指数成分股，来影响整个股票市场的走势，从而引起股价的波动。这种方法需要有足够的市场影响力和技术分析能力，一旦操作失误，可能会对整个市场造成影响。
4. 操纵交易量法：通过大量买入或者卖出股票来增加市场交易量，从而引起股价的波动。这种方法需要有足够的资金实力和市场影响力，一旦操作失误，可能会造成巨额亏损。



一只股票主力要做，股吧雇佣水军是最基本的要求。

- 1、水军是一个团队，有专职负责唱多的，有专职负责唱空的，目的就是一个：让散户低位交出筹码高位捂住筹码。
- 2、水军一般会比较统一的套路，还要有点基本的技术分析基本面分析，有的散户常常追随这些“大神”。
- 3、水军会尽可能发帖，让一些价值分析到位的帖子沉下去。
- 4、水军还会互相点击、互相品论，互相批评，都是为了影响散户的判断力。
- 5、出现水军说明主力在动作了，如果水军变少，说明主力开始撤退，散户可以作为一个操作参照。
- 6、水军长年累月在股吧发言，粉丝很少，关注很少或者没有。

# Thank you!





# 量化交易策略实操

## GPT对量化交易的影响



讲师：张泽亮

# CONTENTS

## ▶ PART 1

### ChatGPT 简介及发展历程

## ▶ PART 2

### ChatGPT 社会应用及争议

## ▶ PART 3

### GPT-4 对量化交易的影响

## ▶ PART 4

### GPT-API 开发实战



ChatGPT（全名：Chat Generative Pre-trained Transformer），美国OpenAI 研发的聊天机器人程序，于2022年11月30日发布。ChatGPT是人工智能技术驱动的自然语言处理工具，它能够通过理解和学习人类的语言来进行对话，还能根据聊天的上下文进行互动，真正像人类一样来聊天交流，甚至能完成撰写邮件、视频脚本、文案、翻译、代码，写论文等任务。





在OpenAI的官网上，ChatGPT被描述为优化对话的语言模型，是GPT-3.5架构的主力模型。ChatGPT具有同类产品具备的一些特性，例如对话能力，能够在同一个会话期间内回答上下文相关的后续问题。然而，其在短时间内引爆全球的原因在于，在网友们晒出的截图中，ChatGPT不仅能流畅地与用户对话，甚至能写诗、撰文、编码。

ChatGPT还采用了注重道德水平的训练方式，按照预先设计的道德准则，对不怀好意的提问和请求“说不”。一旦发现用户给出的文字提示里面含有恶意，包括但不限于暴力、歧视、犯罪等意图，都会拒绝提供有效答案。

2022年11月底，人工智能对话聊天机器人ChatGPT推出，迅速在社交媒体上走红，短短5天，注册用户数就超过100万。



- OpenAI公司在2018年推出具有1.17亿个参数的GPT-1模型；
- 2019年推出具有15亿个参数的GPT-2；
- 2020年推出有1750亿个参数的GPT-3，而ChatGPT是OpenAI公司对GPT-3模型微调后开发出来的智能对话机器人；
- 2022年11月推出GPT-3.5，参数量千亿级，预训练数据量百T级；
- 2023年1月末，ChatGPT的月活用户已突破1亿，成为史上增长最快的消费者应用。
- 2023年2月2日，OpenAI公司OpenAI发布ChatGPT试点订阅计划——ChatGPT Plus。订阅者可获得比免费版本更稳定、更快的服务，及尝试新功能和优化的优先权。
- 2023年2月2日，微软官方公告表示，旗下所有产品将全线整合ChatGPT，除此前宣布的搜索引擎必应、Office外，微软还将在云计算平台Azure中整合ChatGPT，Azure的OpenAI服务将允许开发者访问AI模型。
- 2023年3月发布了备受期待的大模型GPT-4，GPT-4可以理解图像内容。





## 指令模式

我们给模型提供一些特定信息，例如问题或关键词，模型需要生成与这些信息相关的文本。这种模式通常用于生成答案、解释或推荐等。特定信息可以是单个问题或多个关键词，具体取决于任务的要求。

**例：请帮我查一下ChatGPT如何使用**

## 代理模式

在这种模式下，可以充当了一个代理，代表某个实体（例如人、角色、机器人等）进行操作或交互。代理模式的核心思想是引入一个中介对象来控制对实际对象的访问，从而实现一定程度上的隔离和保护。

**例：如果你是一个售前工程师，帮我写一个能源类型的投标方案**

## 示例模式

我们给模型提供一些示例文本，模型需要生成与示例文本类似的文本。这种模式通常用于生成类似于给定示例的文本，例如自动生成电子邮件、产品描述、新闻报道等。

**例：红色代表温暖，黄色代表温暖，橙色代表什么**



ChatGPT受到关注的重要原因是引入新技术RLHF (Reinforcement Learning with Human Feedback, 即基于人类反馈的强化学习)。RLHF 解决了生成模型的一个核心问题, 即如何让人工智能模型的产出和人类的常识、认知、需求、价值观保持一致。ChatGPT是AIGC (AI-Generated Content, 人工智能生成内容) 技术进展的成果。该模型能够促进利用人工智能进行内容创作、提升内容生产效率与丰富度。





ChatGPT的使用上还有局限性，模型仍有优化空间。ChatGPT模型的能力上限是由奖励模型决定，该模型需要巨量的语料来拟合真实世界，对标注员的工作量以及综合素质要求较高。

ChatGPT可能会出现创造不存在的知识，或者主观猜测提问者的意图等问题，模型的优化将是一个持续的过程。若AI技术迭代不及预期，NLP模型优化受限，则相关产业发展进度会受到影响。

此外，ChatGPT盈利模式尚处于探索阶段，后续商业化落地进展有待观察。。

# CONTENTS

## ▶ PART 1

ChatGPT 简介及发展历程

## ▶ PART 2

**ChatGPT 社会应用及争议**

## ▶ PART 3

GPT-4 对量化交易的影响

## ▶ PART 4

GPT-API 开发实战



### 一. 内容生成。

GPT可以自动生成文章、小说段落、新闻稿、产品说明等内容。这大大提高了作者和记者的工作效率。

### 二. 图文翻译。

GPT不仅可以翻译文字,也可以翻译简单的图片和图表。这使得知识和信息的跨语言传播更加容易。

### 三. 代码补全。

对程序员来说,GPT可以智能补全代码,根据上下文推荐变量名、函数名以及其他编程元素。这可以显著提高编码效率。



### 四. 问答系统。

许多问答机器人和聊天机器人采用GPT可以自然地理解问题并给出问题解答。这让更多人可以方便地获知所需信息。

### 五. 自动邮件回复。

GPT也用于某些邮件自动回复系统中,可以理解邮件内容并生成恰当的回复,减少人工处理邮件的时间。

### 六. 机器翻译。

利用GPT机器翻译的质量可以得到较大提升,特别是可以实现更加流畅和适宜的翻译结果。这使得跨语言沟通更加方便。



### 七. 自动评论。

一些社交媒体上的自动评论功能也采用GPT,可以分析帖子内容并生成相关的评论,增加用户互动。

总之, GPT这个人工智能语言模型正被逐渐应用到更加广泛的领域,其强大的语言理解和生成能力可以显著提高效率,改善体验,并且还在继续中推陈出新。在不久的将来,它可能会是我们日常生活与工作中不可或缺的一部分。



多家学术期刊发表声明，完全禁止或严格限制使用ChatGPT等人工智能机器人撰写学术论文。

2023年3月31日，意大利个人数据保护局宣布，从即日起禁止使用聊天机器人ChatGPT，并限制开发这一平台的OpenAI公司处理意大利用户信息。同时个人数据保护局开始立案调查。一方面，个人数据保护局认为，3月20日ChatGPT平台出现了用户对话数据和付款服务支付信息丢失的情况。而该平台却并没有就收集处理用户信息进行告知，且缺乏大量收集和存储个人信息的法律依据。另一方面，该机构还指责，尽管根据OpenAI的条款声称该服务面对13岁以上的用户，但他们并没有使用任何过滤机制来验证用户的年龄。

当日晚些时候，OpenAI表示已在意大利将ChatGPT下线。





2023年4月4日，综合多家媒体报道，自3月11日韩国三星电子允许部分半导体业务部门员工使用ChatGPT开始，在20天内便爆出了三起机密资料外泄事件。三起机密资料外泄案件中，其中两起与半导体设备有关，另一起于内部会议有关。





2023年4月，据路透社报道，澳大利亚墨尔本西部赫本郡的市长指控OpenAI旗下的ChatGPT对其进行诽谤，或将对该公司提起诉讼，因为该聊天机器人在回答问题时错误地声称他在一桩贿赂丑闻中有罪。一旦正式提起，这将是全球首例针对生成式AI的诽谤诉讼。





2023年4月10日，中国支付清算协会表示，近期，ChatGPT等工具引起各方广泛关注，已有部分企业员工使用ChatGPT等工具开展工作。但是，此类智能化工具已暴露出跨境数据泄露等风险。为有效应对风险、保护客户隐私、维护数据安全，提升支付清算行业的数据安全管理水平，根据《中华人民共和国网络安全法》《中华人民共和国数据安全法》等法律规定，中国支付清算协会向行业发出倡议，倡议支付行业从业人员谨慎使用ChatGPT。

# CONTENTS

## ▶ PART 1

ChatGPT 简介及发展历程

## ▶ PART 2

ChatGPT 社会应用及争议

## ▶ PART 3

**GPT-4 对量化交易的影  
响**

## ▶ PART 4

GPT-API 开发实战



## 什么是GPT4?

2023年3月，美国人工智能研究实验室 OpenAI 为其爆红聊天机器人 ChatGPT 发布了最新 GPT-4 语言模型。相比ChatGPT（GPT-3.5），GPT-4的回答准确性不仅大幅提高，还具备更高水平的识图能力，且能够生成歌词、创意文本，实现风格变化。此外，GPT-4的文字输入限制也提升至2.5万字，且对于英语以外的语种支持有更多优化。



- 数据收集：首先需要收集大量的文本数据，这些数据将用于训练模型，数据可以是新闻、书籍等。
- 数据预处理：在将数据输入模型之前，需要对其进行预处理。这通常包括将文本转换为数值表示形式，并将其划分为训练、验证和测试集。
- 模型架构设计：GPT模型的架构通常是基于Transformer模型的。
- 模型训练：使用预处理后的数据训练模型。这通常需要使用大量的GPU资源，并且可能需要花费数天或数周的时间才能完成。
- 模型评估：训练模型后对其进行评估，以确定其在验证和测试数据集上的表现。
- 调整和改进：根据模型在评估过程中的表现调整和改进。
- 不断重复上述步骤得到更加智能的AI模型。



- 10秒钟做一个网站。
- 60秒钟做一个游戏。
- 可以开始识别图片，看懂图片内涵，处理更复杂的图片信息。
- GPT-4产生正确回应的可能性要比GPT-3.5高出40%。
- GPT-4在人类反馈的强化学习训练中加入了一个额外的安全奖励信号，通过训练模型来拒绝对此类内容的请求，从而减少有害产出。



GPT-4比上一代人工智能ChatGPT更加智能，但还是会有些缺点。下面是GPT-4的缺点：

- 1、没有完全解决饱受诟病的ChatGPT胡编乱造的问题，虽然已经减轻幻觉问题。比如可能会出现推理错误，会轻信用户明显的虚假陈述，在生成的代码中引入安全漏洞等。
- 2、虽然具备了基本的推理能力和逻辑能力，但还不能处理复杂问题。





- GPT的风刮向了金融领域。这项变革性的技术正在驱动金融行业使用AIGC，通过实际的应用产品去满足客户需求，进行金融服务创新。公开信息显示，摩根大通、摩根士丹利、美国支付巨头Stripe、高盛等都已纷纷入局AIGC，中国金融机构也开始跃跃欲试。
- 在与AI的互动中，金融机构都有相同的选择：都服务于公司内部，以间接的方式提升客户体验，并不直接对外开放。金融机构不约而同的选择并不奇怪，数字金融业态走向产业与场景的融合阶段下，AI提效最优解就是直接面向客户，金融机构也在围绕这个落点展开。



- GPT的风刮向了金融领域。这项变革性的技术正在驱动金融行业使用AIGC，通过实际的应用产品去满足客户需求，进行金融服务创新。公开信息显示，摩根大通、摩根士丹利、美国支付巨头Stripe、高盛等都已纷纷入局AIGC，中国金融机构也开始跃跃欲试。
- 在与AI的互动中，金融机构都有相同的选择：都服务于公司内部，以间接的方式提升客户体验，并不直接对外开放。金融机构不约而同的选择并不奇怪，数字金融业态走向产业与场景的融合阶段下，AI提效最优解就是直接面向客户，金融机构也在围绕这个落点展开。



- 来源文档: <https://zhuanlan.zhihu.com/p/626598836>
- 写一段股债动量轮动量化策略代码, 要求:从上证50指数基金, 创业板指数基金, 10年国债指数基金这3只基金中, 每天选出近20个交易日中涨幅最大的基金, 如果已经持有该基金则继续持仓, 如果未持有, 则清仓持有的基金买入该基金, 如果这3只基金都涨幅都为负就清仓。
- 请写一段代码回测上述策略的区间收益率、年化收益率、最大回撤, 并可以自定义滑点大小, 最后绘制净值曲线。



# GPT金融实训平台演示

ANF Analyst of Quantitative Finance

← → ↺ 不安全 | mygpt.bizanaly.cn/home.php/kb/kb\_qa?ref=addtabs



🏠 主页 🗑️ 清除缓存 👤 Admin ⚙️

## GPT金融实训平台



搜索菜单 🔍

财经知识库

知识问答

知识分类

财务报表分析

尽职调查报告

营销策划方案

投顾助手

投研助理

合规管理

风险管理

知识问答

🔄 + 添加 ✎ 编辑

编辑

<input type="checkbox"/>	ID	所属用户
<input type="checkbox"/>	7	1
<input checked="" type="checkbox"/>	6	1
<input type="checkbox"/>	5	1
<input type="checkbox"/>	4	1

显示第 1 到第 4 条记录, 总

问题:

我是一名券商机构的投资分析师, 要对商汤科技和海康威视进行全面的对比分析, 并给出充分的理由对以下问题进行说明: 如果投资者需要选一家公司进行投资, 应该选择哪一家公司?

GPT查询

答案:

作为一名券商机构的投资分析师, 我将对商汤科技和海康威视进行全面的对比分析, 并给出选择哪一家公司投资的建议。

1. 市值和财务情况

海康威视的市值约为1.2万亿元人民币, 而商汤科技的市值约为200亿美元。海康威视的财务情况相对较好, 截至2021年上半年, 海康威视的净利润为94.8亿元人民币, 同比增长约41.3%。而商汤科技的财务情况相对较弱, 截至2021年上半年, 商汤科技的净利润为负数。

2. 技术实力和研发投入

商汤科技是一家以人工智能技术为核心的公司, 其在人脸识别、图像识别等领域具有较强的技术实力和创新能力。商汤科技在研发方面的投入也相对较高, 2019年研发投入为14.6亿元人民币。

海康威视则是一家以安防技术为主的公司, 其在视频监控、智能交通、智能建筑等领域具有较强的技术实力和市场占有率。海康威视在研发方面的投入也相对较高, 2019年研发投入为78.8亿元人民币。

3. 业务拓展和应用场景

商汤科技的人工智能技术被广泛应用于安防、零售、金融等领域, 其提供的智

确定

重置

# CONTENTS

## ▶ PART 1

ChatGPT 简介及发展历程

## ▶ PART 2

ChatGPT 社会应用及争议

## ▶ PART 3

GPT-4 对量化交易的影响

## ▶ PART 4

**GPT-API 开发实战**



- 使用科学上网服务器，或者VPN访问 <https://chat.openai.com>;
- 使用公司邮箱及国外手机号码进行注册和验证;
- 使用虚拟信用卡充值并开通GPT-4服务。



- API调用前过滤（敏感词库）；
- API调用后过滤（敏感词库）；
- 标准提示词定义；
- 知识库建设。



- 服务器端直接调用;
- 封装成 web 服务供远程调用;



# Thank you!

