

2023 科研营——量化交易在金融市场价格预测的应用

1 简介

"低买高卖"，这听起来很容易....在现实中，交易获利一直是一个难以解决的问题，在今天快速流动和复杂的金融市场中更是如此。电子交易允许在几分之一秒内发生数以千计的交易，从而产生了几乎无限的机会，有可能发现并利用实时的价格差异。在一个完全有效的市场中，买家和卖家将拥有做出理性交易决定所需的所有机构和信息。因此，产品将始终保持其 "公平价值"，永远不会被低估或定价过高。然而，金融市场在现实世界中并非完全有效。制定交易策略来识别和利用无效市场是具有挑战性的。即使一个策略现在是有利可图的，它在未来也不一定是有利可图的，而且市场的波动性使得我们无法确定地预测任何特定交易的盈利能力。因此，要区分好运气和做出好的交易决定可能很困难。

本项目旨在通过理论与实践相结合的方式，带你建立自己的量化交易模型，利用全球主要证券交易所的市场数据实现收益最大化。接下来，你将根据未来的市场回报测试你的模型的预测性，并在排行榜上获得反馈。你的挑战将是利用你所掌握的历史数据、数学工具和技术工具，创建一个尽可能接近确定性的模型。你将会看到一些潜在的交易机会，你的模型必须选择是接受还是拒绝。一般来说，如果一个人能够产生一个高预测性的模型，选择正确的交易来执行，他们也将发送市场信号方面发挥重要作用，推动价格接近 "公平 "的价值。也就是说，一个更好的模型将意味着市场在未来会更有效。然而，开发好的模型将具有挑战性，原因有很多，包括信噪比非常低，潜在的冗余性，强烈的特征相关性，以及难以提出适当的数学公式。

通过学习相关理论基础、增强实际动手编程能力、培养创新思维和团队协作能力，为金融数据分析的进一步学习和实践应用打下必要的基础。

此实习项目专门为计划申请计算机、软件工程、经济、金融、商科的学生所设计。学生将跟随导师共同工作，实际进行量化交易建模处理等工作，最后利用所学知识实现基于量化交易的金融产品价格预测系统。

2 内容

本实习项目为量化金融的初步应用。将会设计一个程序，提取当前的金融产品信息，实时预测下一时间点的价格，为金融监管提供支持。

在实习过程中，指导老师全程带领学生进行该科研项目的各个环节。通过参与该项目，学生可以了解如何系统地进行金融数据分析的科研工作。

3 师资背景

任职教师为中科院知名研究所副研究员，博士生导师。每年均成功推荐多位学生赴海外名校深造，有良好的科研短训经验。主要研究方向为自然语言处理、金融大数据，机器学习等，在 IEEE Transactions on Knowledge and Data Engineering ,Special Interest Group (SIG) on Knowledge Discovery and Data Mining , International Conference of Data Mining , IEEE International Conference on Communications 等国际著名期刊及会议上发表论文 50 余篇。现担任某 SCI 国际期刊的编委。

4 招生对象及要求

大一以上优秀本科生，计划申请计算机（包括数据挖掘、机器学习等）、软件工程、经济、金融、商科等相关专业，为了让学生可以更好地完成科研项目，项目组会以面试的形式

对学生进行筛选。

5 行程安排

时间	内容
第一次课	<div>1、金融大数据背景介绍</div> <div>2、初识量化交易</div> <div><div><input type="checkbox"/> 1：量化交易场景下Python基础知识的准备</div><div><input type="checkbox"/> 2：用NumPy来进行金融数据操作</div><div><input type="checkbox"/> 3：用Pandas来进行金融数据操作</div><div><input type="checkbox"/> 4：用Matplotlib对金融数据可视化</div></div>
第二次课	<div>1、Python环境的搭建</div> <div><div><input type="checkbox"/> 1：安装Python的发行版Anaconda</div><div><input type="checkbox"/> 2：安装Python IDE PyCharm</div></div>
第三次课	<div>1、金融数据的获取</div> <div><div><input type="checkbox"/> 1：通过Tushare库获取历史数据</div><div><input type="checkbox"/> 2：通过新浪财经 API 获取历史数据</div><div><input type="checkbox"/> 3：通过 Pandas_datareader 获取历史数据</div><div><input type="checkbox"/> 4：其他获取历史数据的方式</div></div>
第四次课	<div>1、量化交易的利器</div> <div><div><input type="checkbox"/> 1：Ta-Lib 库的介绍与安装</div><div><input type="checkbox"/> 2：市场技术指标的计算</div><div><input type="checkbox"/> 3：K 线组合的模式识别</div></div>

	<ul style="list-style-type: none"> □ 4：FFn 库的介绍与安装 □ 5：风险指标的计算 □ 6：两种经典策略的实现
第五次课	<p>1、时间序列分析</p> <ul style="list-style-type: none"> □ 1：Statsmodels库的介绍与安装 □ 2：时间序列的基本概念 □ 3：时间序列相关性分析 □ 4：时间序列平稳性分析 □ 5：时间序列协整性分析 □ 6：时间序列模型 □ 7：时间序列模型在股票市场中的应用
第六次课	<p>1、基于配对交易策略的回测框架的搭建</p> <ul style="list-style-type: none"> □ 1：配对交易介绍 □ 2：配对交易回测框架的实现
第七次课	<p>1、机器学习实战——利用支持向量机（SVM）进行趋势预测</p> <ul style="list-style-type: none"> □ 1：机器学习库Sklearn的介绍与安装 □ 2：机器学习基本知识介绍 □ 3：支持向量机介绍 □ 4：支持向量机预测模型的实现
第八次课	<p>1、深度学习实战——利用循环神经网络（RNN）进行价格预测</p> <ul style="list-style-type: none"> □ 1：深度学习库TensorFlow的介绍与安装 □ 2：TensorFlow的基本概念与结构

	<ul style="list-style-type: none"> □ 3：循环神经网络介绍 □ 4：循环神经网络预测模型的搭建
第九次课	<p>1、接触实盘——利用vn.py进行量化交易</p> <ul style="list-style-type: none"> □ 1：初识vn.py □ 2：vn.py运行环境的准备 □ 3：vn.py国内期货CTP的配置 □ 4：通过vn.py进行策略回测 □ 5：基于vn.py实现R-Breaker策略 □ 6：通过vn.py进行自动化交易
第十次课	<p>1、Kaggle比赛</p> <ul style="list-style-type: none"> □ Jane Street Market Prediction
第十一次课	<p>1、论文查阅方法介绍</p> <ul style="list-style-type: none"> □ 如何阅读英文论文 □ 阅读笔记工具NoteExpress介绍 □ Google学术搜索、微软Bing学术搜索
第十二次课	<p>1、论文撰写方法介绍</p> <ul style="list-style-type: none"> □ 如何撰写英文论文 □ 论文编辑Latex工具介绍 □ 公式编辑器MathType工具介绍
第十三次课	<p>1、论文实验方法介绍</p> <ul style="list-style-type: none"> □ 做图工具Matlab介绍
第十四次课	<p>1、论文修改方法介绍</p>

	<input type="checkbox"/> 语法修改 <input type="checkbox"/> 实验修改
第十五次课	1、论文投稿介绍 <input type="checkbox"/> 会议投稿 <input type="checkbox"/> 期刊投稿
第十六次课	1、项目回顾 <input type="checkbox"/> 代码回顾 <input type="checkbox"/> 论文回顾