

Class Note

2018年8月7日 18:17

Introduction for Data Science

1. Course Outline

- a. 金融统计进阶
 - i. 概率分布
 - ii. 线性回归
 - iii. 时间序列模型
 - iv. 主成分分析(PCA)
 - v. 因子分析(Factor Analysis)
 - vi. 风险模型
- b. 机器学习,分类聚类分析
 - i. Package: Scikit-learn的介绍
 - ii. Package: Keras的介绍
 - iii. 决策树, KNN, 朴素贝叶斯, K-Means, SVM
 - iv. 人工神经网络在金融中的应用
- c. 网络金融数据的获取, API以及爬虫
 - i. Package: BeautifulSoup的介绍
 - ii. Html解析以及网络爬虫
 - iii. API的使用和沟通方法
 - iv. 常用金融数据的收集与获取
 - v. 基础自然语言处理 (NPL)

2. 金融行业分工及数据科学

- 卖方
 - 投资银行, 商业银行, 证券公司 ...
- 买方
 - 个人, 公司, 保险公司, 基金, 券商资管 ...
- 服务提供商
 - 评级机构, 金融研究, 数据, 系统 ...
- 监管机构, 政府机构
 - 央行, 银保监会, 证监会, 政府职能部门



3. 服务提供商

- 评级机构
 - 标普, 惠誉, 穆迪, Equifax, Transunion
- 投研机构
 - 券商, 投行, 专业投研机构
- 数据
 - 市场数据 (Market Data) : Bloomberg B-pipe, Thomson Reuters, ACTIV Financial, Interactive Data (ICE) ...
 - 基础数据 (Fundamental Data) : FactSet, MSCI Barra, IBES (Thomson Reuters), RavenPack, 万得, 东方财富, 朝阳永续 ...
 - 另类数据 (Alternative Data) : Quandl ...
- 系统
 - 综合系统: Bloomberg Terminal, Thomson Reuters Eikon ...
 - 交易系统: Citi Velocity, FlexTrade, 通达信, 恒生 ...
 - 风险管理系统:
 - 服务器: 数据中心 (Data Center) , AWS, 阿里云 ...

4. 量化交易 (quantitative trading)

量化交易是指以数学模型替代人的主观判断, 利用计算机技术从庞大的历史数据中海选能带来超额收益的多种“大概率”事件以制定策略, 极大地减少了投资者情绪波动的影响, 避免在市场极度狂热或悲观的情况下作出非理性的投资决策。

a. 量化交易策略种类

股票 (Equity) :

- 价值投资
- 多因子选股 (multi-factor model)
- 机器学习

CTA (Commodity Trading Advisor) :

- 趋势策略 (trending strategy)
- 回归策略 (mean reversion strategy)
- 统计套利 (statistical arbitrage, pairs trading)
- 波动性策略 (volatility strategy)

高频 (High Frequency Trading)

- 做市 (market making)
- 套利 (arbitrage)
- 短期区趋势 (directional trading)

b. 量化交易的Architect

