Machine Learning Applications in Asset Management

12/15/20 5:07 PM

Speaker: Tugce Karatas (Columbia University) and Satyan Malhortra (Flexstone Partners)

A. Overview

Machine Learning's base pillars

- Data
  - Needs to be carefully processed to be useful
- Analytical Tools
  - Signals
    - How does price / volume move
    - Which sector / asset is going to outperform
    - What is product's success rate
  - Sentiment
    - Up or down trend
    - What is closing probability
  - Scrapping
    - What is real time impact of news
  - Clustering
    - What should allocation be
    - What are cohorts and replications
    - What are key questions
- Visualization
  - What is best representation for the application

B. Liquid products - Mutual Funds

1. Manager Selection

Which managers may outperform given objective function?
(Which funds to be included in our portfolio?)

- Features: alternative data, holding data
  - Alternative data: information about company but published from external sources
  - Holding data: anything related to the holding of fund (e.g. turning ratio)
- Optimization: LSTM

2. Portfolio Construction / Asset Allocation

How to allocate assets given object functions?

- Traditional: from history return data, optimize expected return for given risk level
- Sample nested clustered optimization
- LSTM for initiate weights, then Reinforcement Learning to decide final weights of portfolio
- Asset allocation visualization

3. Stress testing / Real time impacts

Test the market conditions and portfolios against historic scenarios and news to access impacts

- We can do stress testing in order to see resilience of models under events like 2008 crisis or COVID-19
  - Take our current allocation to go back to the scenarios
  - To see if we have these events again, how our performance will be
- Real time impacts:
  - NLP and sentiment to access potential impacts on sectors, portfolios, and other

4. Market indicators

Identify the relative potential and attractiveness of sectors and sub-sectors, given the economic condition

- Framework
  - Data collection
    - Macroeconomic indicators and performance benchmark data by sector
  - Feature engineering
    - Determine the most important features for each sector
  - Prediction by sectors
    - ML and DL to predict performance benchmark data for each sector
  - Ranking
    - Rank sectors based on predicted values
- One example: Echo State Networks Application (Jaeger, Herbert)
- Extensions: Jump prediction, sentiment analysis

C. Illiquid Products - Private Equity

1. Manager Selection
2. Investment Selection
3. Market Indicator

*These three are essentially same with liquid products but with unique datasets

4. Cash Flow Forecasting

- Cash flow questions:
  - What are the contribution and distributions profiles?
  - What impact do unplanned events have on profile?
  - How close is the profile to the generic sectors or sub-sectors profiles?
  - What is the tracking error and reinforcement method?
- Data challenges
  - Sparse and difficult to access
  - Not standard
  - Infrequent updates
- Machine Learning applications: allow building of assumption-free models
  - Key challenge: Inconsistent and insufficient data
    - Solution: Brownian interpolation
  - Contributions and distributions are predicted independently