Trading Chinese Mutual Fund: A Fund of Fund Strategy

By:

Charles Deng 12334214
Sosuke Kohda 12160736
Owen Howell 12104004
Holly Cao 12333838

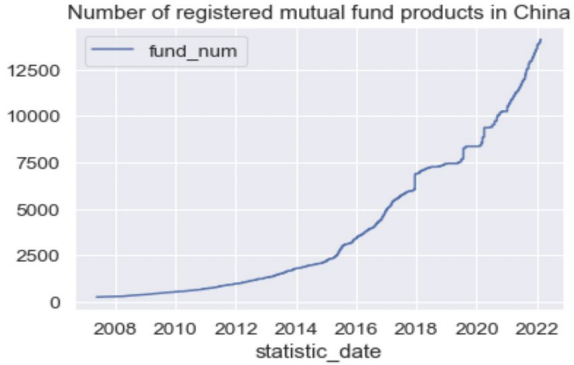
Final Project Draft

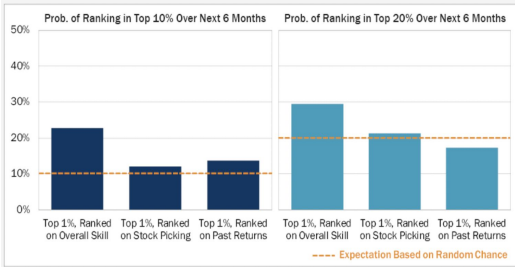
Regression Analysis and Quantitative Trading Strategies
with Brian Boonstra Winter 2022

THE UNIVERSITY OF CHICAGO | Department of Mathematics
Financial Mathematics Program

Background

- The Chinese mutual fund industry has seen robust growth over the past 5 years, with total asset management exceeding 3.04 trillion USD as of December 2020, as the total number of mutual fund products grow exponentially since 2001
- According to the annual report by the Shanghai Stock Exchange, retail investors account of 82% of the exchange trading volume, leading to potential market mispricing that grants institutional investors an edge.
- Chinese equity mutual fund managers has been able to deliver a risk-adjusted annual return of 7-8% , in contrast to near zero alpha after fees delivered by mutual funds in the US market according to Yu et al. (2020).





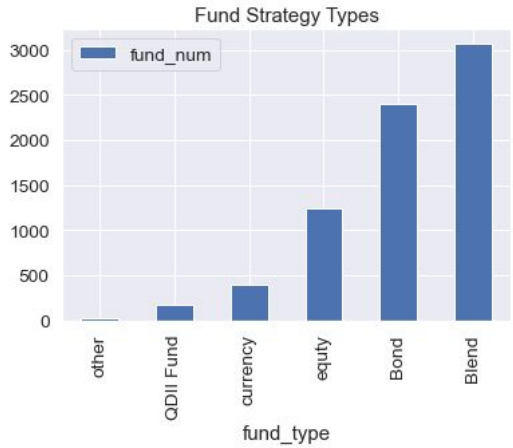
Source: Cornell, Hsu, Kiefer, and Wool (2020)

Strategy

- Prediction of cross-sectional mutual fund returns via performance factors such as sharpe ratio, max draw downs etc as well as ‘fundamental’ factors including abnormal cash holding, management fees, total assets etc.
- Long-only quantile trading strategy to buy the top funds based on factor ranking.
- Backtest with monthly rebalance.
- Expense and turnover consideration of individual funds

Data Overview

- The CHFDB database is a proprietary industry-grade data source that provides extensive access to information related to the Chinese mutual fund industry.
- It covers over >12000 mutual fund products with daily, weekly and monthly frequencies and provides up-to-date information relating to past fund performance, fund operations and structures, management details, money flows etc.
- The data is accessed through SQL



Investment Universe

- The investment universe contains mutual funds that are:
  - Equity Focused
  - At least 2-year old of NAV
  - Main Fund*
  - Use CSI 300 as benchmark
  - Open-End funds with AUM of at least ¥50 million (~8 million USD)
  - With no minimum investment requirement
- The Strategy takes:
  - Long only positions in mutual fund

*Master fund that could potentially have several subsidiary products with different expense structure

Competitive Edge

- Systematic approaches are not yet widely adopted in the Chinese for the portfolio management industry. Recent reports* shows that the Chinese fund of fund industry is still at its infancy, with bond funds accounts fo 51% of all FOF products.
- Access to unique industry-grade dataset that covers all perspectives of mutual fund performances and fundamentals, whereas most retail investors only look at historical returns.
- Extensive backtesting, with consideration of fees and turnovers

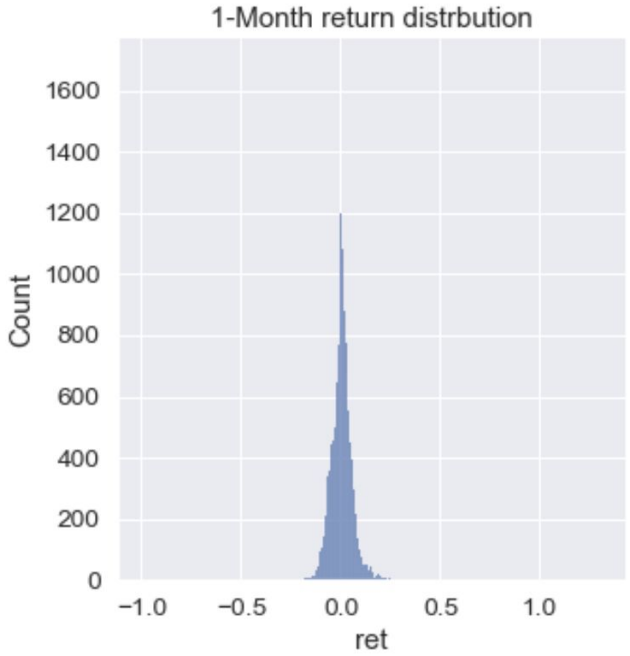
*http://www.21jingji.com/article/20211020/herald/16c7fa9b7e6e431b78aeb7df0d33d57e.html

Data Engineering

THE UNIVERSITY OF CHICAGO | Department of Mathematics
Financial Mathematics Program

Fund Filtering

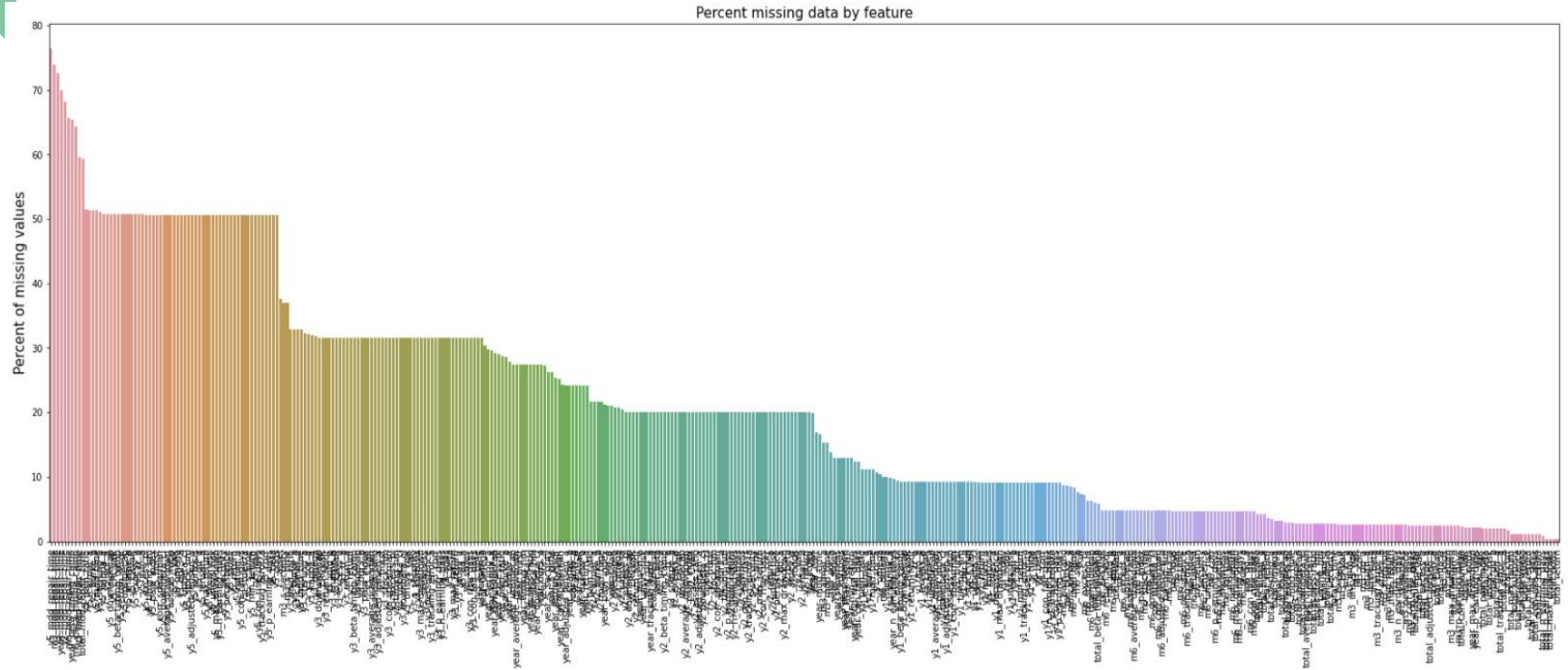
- The database contains 14117 mutual fund products in total, 608 funds are selected after criteria filter.
- One-month look ahead return is computed using net asset value that accounts for dividend with history starting from 2017-01-01



Feature Engineering: Performance Features

- Features that are based on performance measures of past returns.
- 431 pre-computed performance features including:
  - sharpe ratio
  - Max drawdown
  - Hurst ratio
  - Performance persistence
  - Win rate
  - etc
- 190 performance features after omitting features with >10% missing values
- Relative performance features are computed using CSI300 index as benchmark

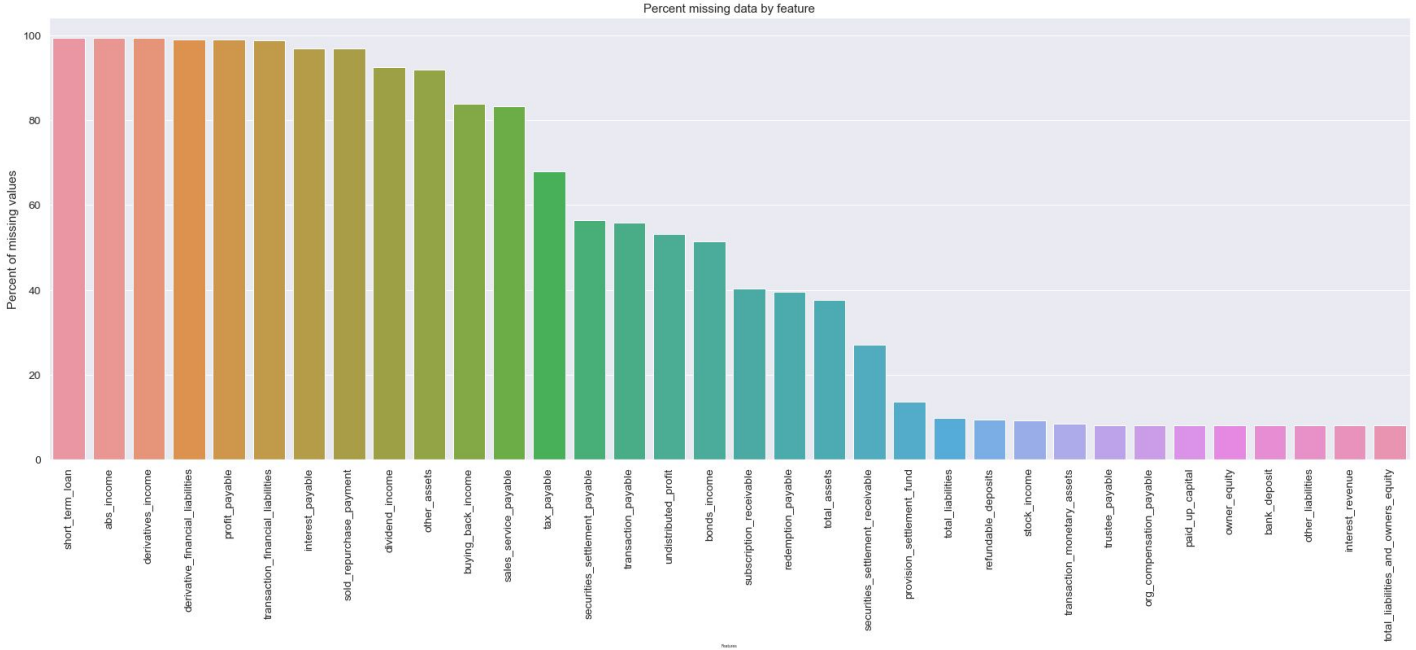
Feature Engineering: Performance Features



Feature Engineering: Fundamental Features

- Features that are based on information relating to fund operations, fee structures and manager skills.
- Convert such features to time series and front fill missing values
- Such features include:
  - Fee: Management and purchase fee charged to the product
  - Fund flow: Total purchase amount - total redemption amount
  - Fund asset: Total asset owned by individual product
  - Short term loan: Short term loans taken by mutual fund to fund its liquidity
- Start with 34 fundamental features and retain 19 of such features after filtering out features with >10% missing values

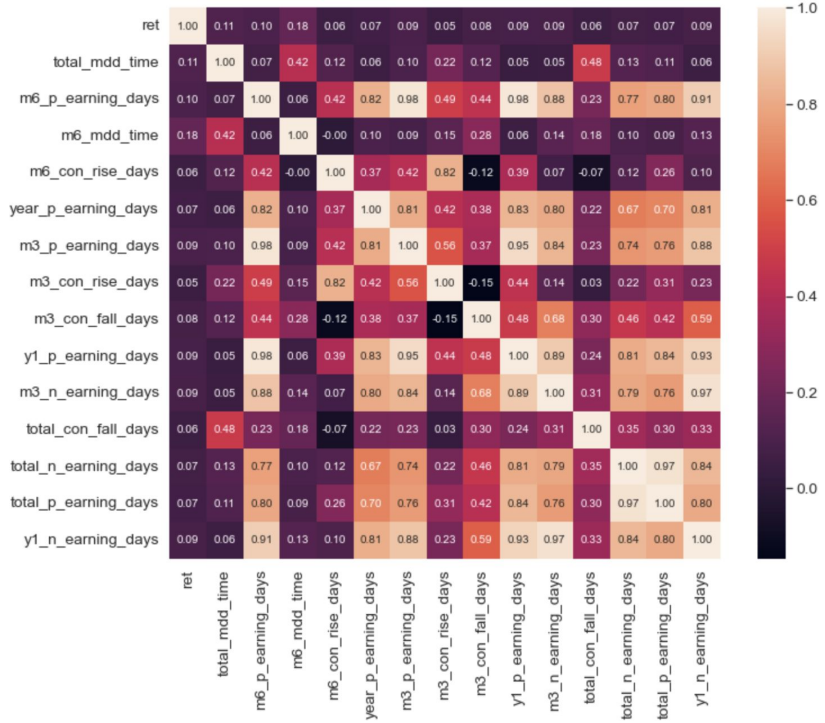
Feature Engineering: Fundamental Features



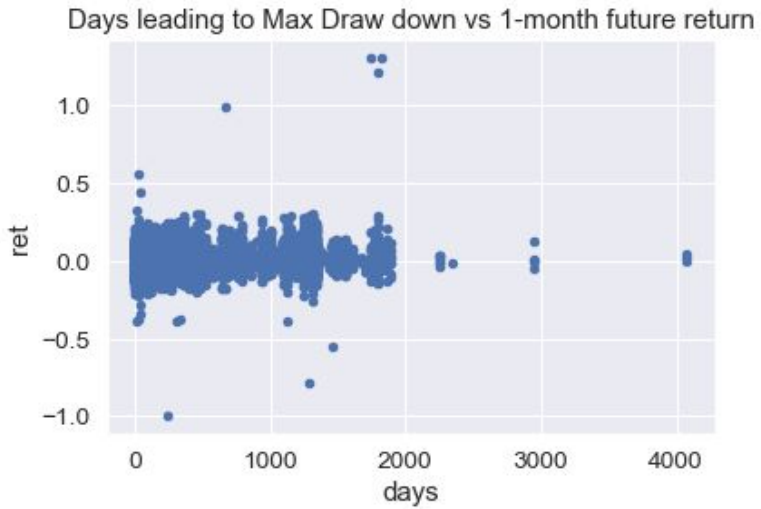
Feature Engineering: Summary Statistics for Mentioned Features

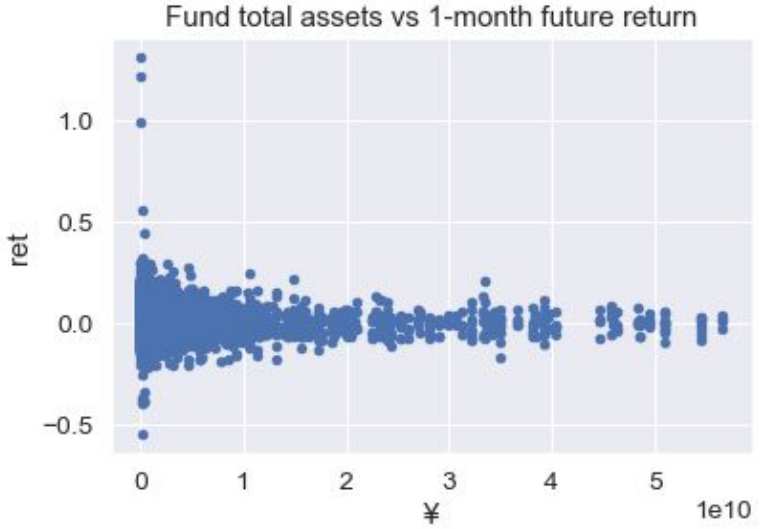
	ret	fee	flow	asset	total_dd_a	total_persistence	total_odds
count	13583.000000	13583.000000	13583.000000	1.358300e+04	13583.000000	13583.000000	13583.000000
mean	0.001959	0.007552	0.127822	1.564238e+09	0.154597	0.039835	0.510694
std	0.054855	0.003326	5.762593	5.003346e+09	0.047045	0.146345	0.094393
min	-0.550832	0.001500	-101.830000	4.648638e+06	0.002449	-0.665886	0.125000
25%	-0.029400	0.005000	-0.200000	8.802744e+07	0.132655	-0.045670	0.452381
50%	0.004132	0.006000	-0.020000	2.443052e+08	0.156296	0.024197	0.512195
75%	0.031467	0.010000	0.090000	8.387007e+08	0.180804	0.111658	0.565554
max	1.312039	0.015000	198.760000	5.661257e+10	0.370616	0.626889	0.916667

Feature Engineering: 15 most correlated features with 1- month lookahead return



Feature Engineering: Bivariate plot





Feature Engineering: Data Normalization

- Demean mutual fund features

- Use z-score to perform cross-sectional ranking for quantile trading.

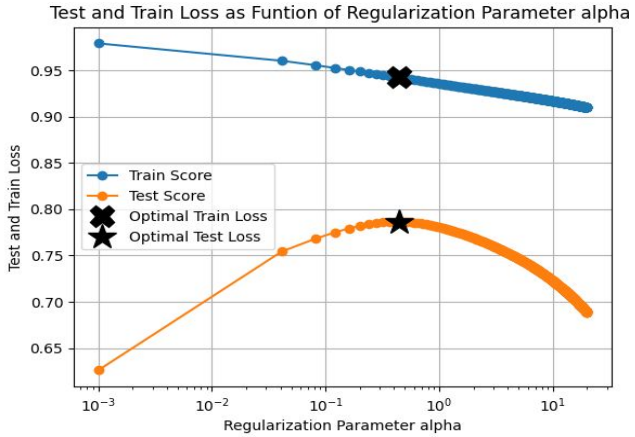
Hypothesis on features

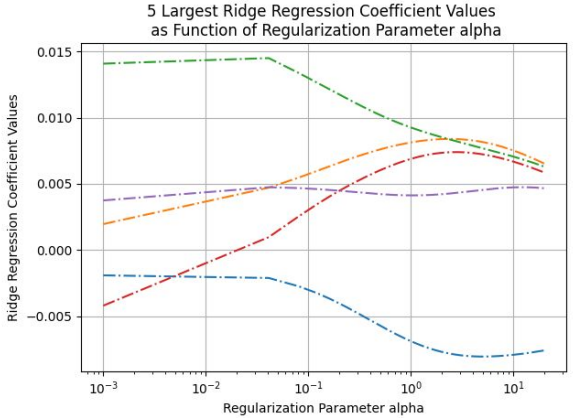
- Elton, Gruber, Das, Hlavka (1993) found that funds with higher fees and higher turnover resulted in poor performance.
- Gruber (1996) and Zheng (1999) found that higher fund flows are associated with higher returns.
- Hendricks, Patel, and Zeckhauser (1993) found that persistence exists in mutual funds - specifically, one-year past returns can predict future performance.
- Chen, Hong, Huang, Kubik (2004) found that size of a fund can negatively affect the performance.
- These are examined on mutual funds in the U.S. but can potentially be applied to funds in China.

Exploratory Data Analysis

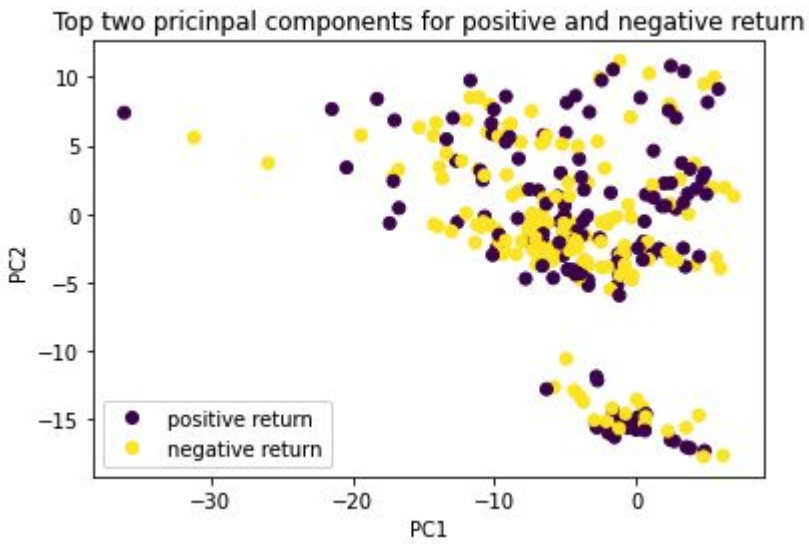
The most relevant features through a preliminary ridge regression analysis include:

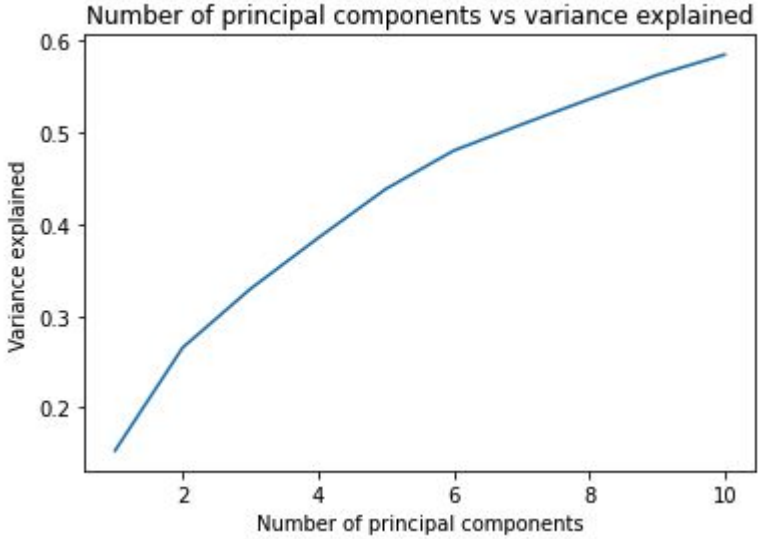
- month_sterling_a
- total_mdd_time
- y5_mdd_time
- m1_sterling_a
- m1_pain_ratio
- total_n_earning_days
- month_pain_ratio
- total_p_earning_days
- m3_sterling_a
- 1year_sterling_a





Exploratory Data Analysis - PCA on features





Backtesting: Business Constraint

- The portfolio value is ¥20 million, with individual fund allocation not exceeding ¥4 million (20 %).
- Rebalances can be completed in no less than 3 business days
- Rebalances need to happen at least once per month
- Expenses vary depending on fund structure, holding period and purchase amount
- No leverage or shorting is allowed.
- Only active fund is traded

Next Steps

- Feature selection with statistical methods
- Predictive analysis for one month look ahead returns
- Predictive analysis on weekly data
- Factor construction and analysis
- Quantile Trading Backtest Construction
- Performance analysis
- Paper trading
- Test run of allocated capital