Clustering based portfolio optimization using investor information in Indian stock market

Patoliya Meetkumar Krushnadas (Roll No. 150101045)
Supervisors: Dr. N. Selvaraju and Dr. S. V. Rao

Department of Computer Science and Engineering, IIT Guwahati.

Introduction

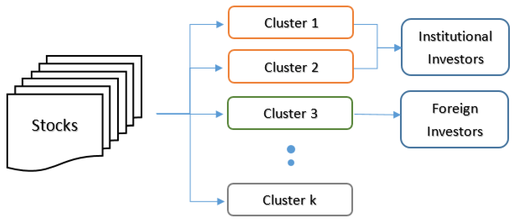
Various types of investors - individual, foreign and (domestic) institutional investors engage in constant competition. Studies have shown that performances of institutional and foreign investors are better than individual investors. Therefore, it appears evident that emulating foreign and institutional investor portfolio strategies can lead to better performance. Goal of an investor is to allocate capital to portfolio of stocks in a way that maximizes returns while minimizing their risk. Markowitz developed modern portfolio theory, but it is not easy to assign relative weights using that. This study presents scheme for portfolio optimization, in which a clustering-based stocks selection using investor information is followed by portfolio optimization using genetic algorithm.

Theory

- **Portfolio Management:** Portfolio management involves choosing stocks to incorporate into the portfolio and choosing how much capital to distribute to each stock according to risk tolerance level of investor. We will use active portfolio management technique.
- **Investor Information:** Studies has shown that foreign investors exhibit strong market-timing abilities and thus achieve better performance while institutional investors appear to have informational advantages relative to other investor types. It is seen that both of these types of investor outperforms individual investors.

Stocks Selection

In this step, highly traded stocks are selected based on investor information. These stocks are expected to give higher return. To select these stocks, k-means clustering is used.



Step 1: Stock Selection

The features used to do k-means clustering are- foreign, individual and institutional investor TVP (trading volume proportion), stock return, volume ratio. After k-means clustering, clusters with highest mean TVP for institutional and foreign investor are chosen. The portfolio is constructed and optimized from stocks of chosen clusters.

Portfolio Optimization

In this step, our aim would be to assign weights to stocks in portfolio to minimize the portfolio risk for given return or maximizing the portfolio return for given risk. Let w_j is proportion invested in stock j. We need to optimize weights w_js to maximize return such that they follow below constraints:

$$\sum_{j=1}^{n} w_j = 1, \quad w_j \geq 0 \quad \forall j = 1, 2, ..., n$$

Genetic algorithm is used to find solution in search space. Fitness functions used for evaluations are-

- Minimum variance $\sigma_p{}^2 = \min w_p C_p w_p{}^T$
- Maximum Sharpe ratio $S_p = \max \frac{R_p - R_f}{\sigma_p}$

Processes performed while evolution are-

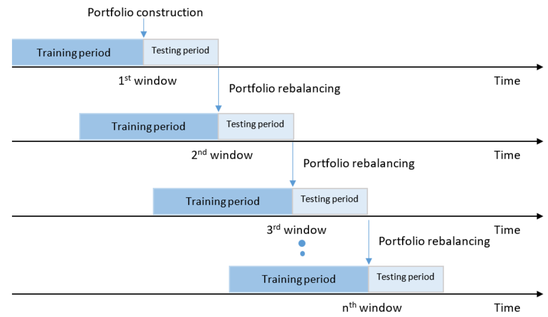
- Generation of initial population
- Selection
- Mutation
- Crossover

Other weight vectors considered are-

- Equal weight vector
- Market capitalization weight vector

Sliding Window Technique

Whole time-line is divided into windows. Each window is having training period and testing period.



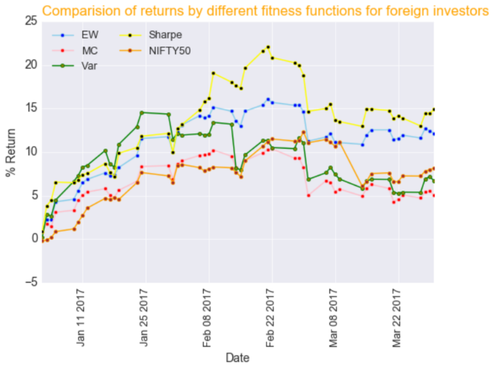
- **Stocks Selection:** k, investor type
- **Portfolio optimization** population size, mutation rate, crossover rate, number of generations, fitness function

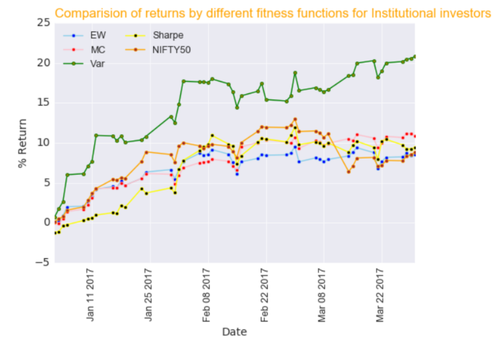
Data Collection

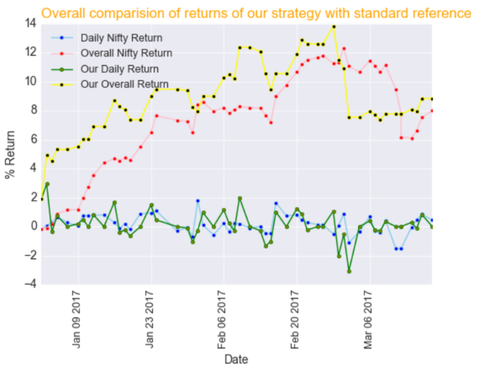
Proposed model is applied on Top 100 stocks of National Stock Exchange of India (NSE) in terms of market capitalization. From **CAPITALINE** database, daily share price, market capitalization, daily volume trade and quarterly shareholding pattern data of these shares are used to calculate TVPs of foreign, institutional and individual investors.

Empirical Results

We applied our algorithm in first quarter of year 2017. Population size was kept 50, crossover rate was kept 0.5 and mutation rate was kept 0.5/(length of chromosome). Total number of generations was taken as 500. Risk free rate for calculation of Sharpe ratio is taken to be 6.50% For foreign investors maximum Sharpe ratio fitness is best suitable while for institutional investors minimum variance fitness is best suitable.







Proposed model works well on long term (6-months) training on trading volume data of investors there by yielding high return portfolios.

Contact Information

- Email: patoliya@iitg.ac.in
- Phone: +91 9409076890