Optimal Portfolio Selection of Equity Mutual Funds using Genetic Algorithm: An Indian Perspective

Soumya Banerjee[1] (corresponding author), Sayan Gupta[2], Amlan Ghosh[3], Gautam Bandyopadhyay[4]

[1] Department of Management Studies, NIT Durgapur & Department of Statistics, Amity University Kolkata
[2,3,4] Department of Management Studies, NIT Durgapur

[1] Address: A-43 Satindra Pally, Garia, Kolkata – 700084, West Bengal, India, souma.ban@gmail.com

Article Info

Page Number: 4620 - 4631

Publication Issue:

Vol 71 No. 4 (2022)

Article History

***Article Received**: 25 March 2022*

***Revised**: 30 April 2022*

***Accepted**: 15 June 2022*

***Publication**: 19 August 2022*

Abstract

The key issue with financial investing selections is choosing the best portfolios. To maximise profit and reduce risk, investors must be able to choose the ideal mix and percentage of shares. This study aims to obtain an optimal portfolio selection of Indian Equity Mutual Funds by minimizing risk and maximizing return using the Genetic Algorithm. All the equity mutual funds with inception before January 2015 are considered and their monthly returns are calculated. The funds with positive average return and negative skewness of return are taken into consideration. Since a mutual fund's performance is evaluated in relation to the market, the portfolio is built by selecting those with a low standard deviation and a high beta value using an investor perception map. The market benchmark, the BSE 100, is used, and their monthly returns for the same time period are determined. Using the Genetic Algorithm, the funds in the portfolio have been given the proper weightage ensuring minimum risk and maximum return. This study will guide the investors to choose mutual funds wisely as well as instruct them on how to distribute proper investment weighting, which will aid them in making future investment decisions.

Keywords:-Portfolio, Return, Risk, Equity Mutual Funds, skewness, standard deviation, beta, Genetic Algorithm.

Introduction

Indian capital market provides numerous investment avenues to the investors, to assist them to take an investment position in different industries and to confirm the profitable outcome on return on investment. Among different financial avenues, a Mutual Fund (MF) ensures minimum risk with maximum return to the investors. The Indian MF industry provides an abundance of schemes and serves all types of investor needs. As of April 30, 2022, the size of the Indian MF market was approximately INR 38.04 trillion, and in the previous ten years, the corpus size has expanded up to five times (Association of Mutual Funds in India, 2020). It suggests a staggering 500 percent growth in just 10 years, which denotes people's interest in investing in MFs. The returns on broader schemes of these funds from 2013-14 are higher than the other financial instruments like post office monthly income schemes, National Savings Certificate VIII issue, etc. (CRISIL- AMFI Mutual Fund Performance Indices, (2018)) which have shown a significant downfall in the interest rates during these periods (RBI (2018)). Because of its higher return as compared to the other conventional financial instruments, the mutual fund has been a profitable investment avenue for investors and after demonetization with banks slashing the interest rates, made more people turned to mutual funds. After demonetization, equity mutual funds had significant net inflows of Rs. 1.23 trillion in just the first year (CMIE (2017). With this rapid increase of the mutual fund demand, close monitoring of mutual funds has become very essential, and choosing lucrative mutual funds for investment becomes a very critical issue.

One of the most difficult issues in the finance industry is portfolio optimization. Selecting the weights of assets to invest in a portfolio to meet the expectations about risk and return makes this problem more important. Despite there being many famous portfolio optimization models and techniques, the Modern Portfolio Theory (Markowitz (1952)) tends to remain the most used model because it analyses covariance, unlike other models. The mean-variance is one of the most important theories in finance studies where the Markowitz Efficient Frontier refers to the set of portfolios in which maximum expected returns are reached at a given level of risk or minimum risk is attained at a given level of return (Lee et al., (2010)). However, the model's application is constrained by cardinality and boundary restrictions (Fernandez and Gomez, 2007). Constrained optimization techniques are utilised to address the constraints imposed (Davidson, 2011). However, due to its linear foundation and preference for quadratic (deterministic) objective functions with a single aim, these restricted optimization techniques have some problems in portfolio optimization (Roudier, 2007). Sharpe (1971) and Stone (1973) made an effort to linearize the portfolio decision issue. The Markowitz model in its traditional version is still far from satisfying a professional investor, as demonstrated by Rudd and Rosenbeg (1979), and they proposed a realistic portfolio management. Konno (1990), Yamazaki and Konno (1991), Pang and Zenios (1993), and Speranza (1993) recommended measuring portfolio risk using the mean deviation. They have proposed a model that takes into account the asymmetric risk criterion, which addresses most of the problems with the optimization model. Several researchers suggesting several alternative means and methods to create an efficient portfolio capable of giving optimized returns with the least risk propensity prove the importance of the subject. Some of the suggested examples are the applicability of clustering techniques and their relative advantages in the making of an efficient stock portfolio in the Indian context (Nanda et. al., 2010), optimizing a stock portfolio with the help of linear-regressing ensembles (Nagy, et. al. 2015), formation of product portfolio using decision tree for design optimization (Tucker and Kim, 2009), ESP optimization applying artificial neural network modeling (Nazemi et. al. 2015), optimization for intelligent and timely decision-making using multi-criteria decision-making model (Ehrgott et. al., 2004). The different models presented in this literature act either at maximizing the return for a given risk or at minimizing the risk for a given return to determine the optimal choice of portfolio, but not both, at the same time. Here we apply the Genetic Algorithm (GA), based on artificial intelligence to overcome this problem.

Portfolio Optimization utilising GA produces better or superior outcomes when compared to other heuristic techniques (Elton et. al.2014; Woodside-Oriakhi et al. (2011). GA is more flexible than the other search techniques because they only require knowledge of the quality of the solution produced by each objective function values, as opposed to others that demand derivative information or even more, thorough knowledge of the problem structure (Bouktir et al, 2004). In order to enhance the optimal choice—that is, to have the highest return (portfolio value) and lowest risk—the genetic algorithm technique is used to find the best chromosome that interprets the weight or percentage of each stock. Thus, this study attempts to create an optimal portfolio of Equity Mutual Fund in the Indian Market ensuring minimum risk and maximizing return using GA.

Materials and Methods

Data

The sample for this study consists of Large Cap, Mid Cap, and Small Cap mutual funds with at least 5 years of operation in the Indian mutual fund market. The period of the study is January 2015 to December 2019. Therefore, we have included only those funds which are operational on or before January 2015. There are altogether 222 such funds and their monthly Net Asset Values (NAV) from January 2015 to December 2019 are collected (AMFI). The monthly returns are then calculated based on changes in their NAV values over time given as

$R_t = (NAV_t - NAV_{t-1}) / NAV_{t-1}$, t = 2, 3, …, 60 (Baliyan and Rathi, 2019)

where R_t: Return for month t and NAV_t and NAV_{t-1}: Net asset value for month t and t-1 respectively. (1)

Selection of Funds

The funds which have negative average returns are eliminated. The average monthly return of each fund is calculated as $\bar{x} = \sum_1^n x_i / n$; x_i = return of the i^{th} month of the fund. Here n = 60. Then we consider those funds which have negative skewness. Skewness is measured by Bowley's formula:

$$Sk = [(Q_3 - Q_2) - (Q_2 - Q_1)] / [(Q_3 - Q_2) + (Q_2 - Q_1)],$$

where Q_i: ith quartile of the distribution, i = 1,2,3. A distribution is said to be negatively skewed if more values are concentrated on the right side (tail) of the distribution graph while the left tail is longer. The distribution's negative skewness suggests that there won't be many significant losses for investors to anticipate. A lot of the trading tactics used by traders are based on distributions that are negatively skewed as they provide stable profits and higher returns with time (Corporate Finance Institute).

We then calculate the Standard deviation of the monthly returns and the Beta of each fund.

Standard deviation (SD) is a measure of an investment's total risk. It includes both systematic and unique risks. The deviation of the returns from their average return is expressed by standard deviation. It is defined as SD = $[\sum_1^n (x_i - \bar{x})^2 / 59\,]^{1/2}$. (2)

Beta is a measure of an investment's systematic risk compared to the market. It is calculated using the formula:

beta = Cov (x,y) / Var(y),

where Cov (x, y) = $\sum_{i,j=1}^n (x_i - \bar{x})(y_i - \bar{y})$ /59 and Var(y) = $\sum_{i=1}^n (y_i - \bar{y})^2$ / 59. (3)

Here x_i = return of the i^{th} month of the fund, y_i = market return of the i^{th} month, and $\bar{x}$ and $\bar{y}$ denote respectively their average returns.

In our study, the Bombay Stock Exchange (BSE) 100 is considered the market benchmark. BSE is one of the leading stock exchanges in the Indian market and the oldest stock exchange marketplace not just for India but in Asia as well, which offers high-speed trading to its customers.

Next, we select those funds whose SD is less than the combined SD and whose beta is greater than the average beta. This is done by drawing the Investor's Perception Map taking SD on X-axis and beta on Y-axis. Investor's Perception Map displays the position of beta and SD of a fund with respect to average beta and combined SD. Thus, the funds which fall in the second quadrant of the map are considered.

The combined SD is calculated by the formula:
combined SD = $(\sum n_i.s_i^2 + \sum n_i.(\bar{x}_i-\bar{x})^2 / \sum n_i)^{1/2}$ (Goon, Gupta and Dasgupta, 2008)
where s_i^2 = variance of monthly returns of the i^{th} fund, $\bar{x}_i$ = average monthly returns of the i^{th} fund, $\bar{x}$ = combined average, and n_i = number of monthly returns for each fund. (4)
The average beta is given by the average beta = sum of the beta values / n, where n = the number of funds remaining at this stage. (5) With these selected funds we then try to create an optimal portfolio using the Genetic Algorithm.

Genetic Algorithm

Introduced by John Holland (1960) and popularized by David Goldberg (1980), the Genetic Algorithm is a heuristic technique based on the concepts of genetics and the scientific selection process of Darwin's theory of evolution. A genetic algorithm manipulates a population with a fixed size to find the best answer iteratively. Chromosomes are candidate sites in this population. The chromosomes begin to compete with one another as a result of this process. Each chromosome contains a collection of components called genes that have a range of possible values and encodes a potential answer to the issue at hand. A brand-new population of the same size is produced at each iteration (generation). Better chromosomes "suited" to their environment, as shown by the selection function, make up this generation. The chromosomes will gradually gravitate in the direction of the optimal selective function. By utilising the genetic operators of selection, crossover, and mutation, a new population is created. Natural selection-based stochastic heuristic approaches known as GAs may solve nonlinear optimization problems with continuous and/or integer variables, non-smooth and even non-continuous objectives (Lin et al; 2005). The choice of GA parameters, such as the mutation and crossover procedures, might, nevertheless, have an impact on the performance of the GA (Bakhtyar et al., 2012).

The expected return of the individual fund is given by

$$E(w_i) = w_i . r_i \qquad (3)$$

where w_i denotes the weight of the ith fund and r_i denotes the expected return of fund i.

The expected return of portfolio P is then given by: $F = \sum_{i=1}^{n} E(w_i)$, where n is the number of funds, and the objective function of portfolio return to be maximized is written as:

Max F (4)

The Portfolio risk is given as

$$\sigma^2 = \sum_{i=1}^{n}(w_i^2\sigma^2(r_i)) + \sum_{i=1}^{n}\sum_{j=i+1}^{n} 2w_iw_j cov(r_i,r_j)$$

$where\ \sigma^2(r_i)$: Variance of ith fund

$cov(r_i,r_j)$: Covariance between fund *i* and fund j

The objective function of portfolio risk to be minimized is given as:

Min σ^2 (5)

Thus, the multi-objective function to be minimized is given as:

$$K = \sigma^2 - F \qquad (6)$$

Subject to the constraints:

$$w_i > 0,\ \sum_{i=1}^{n} w_i = 1 \qquad (7)$$

The fitness function is given as:

Fitness = K + 100. $[(\sum_{i=1}^{n} w_i - 1)^2 + \sum_{i=1}^{n}(Max\ (0, w_i - 1))^2 + \sum_{i=1}^{n}(Max\ (0, -w_i))^2]$

The penalty function is given as:

Penalty = $(\sum_{i=1}^{n} w_i - 1)^2 + \sum_{i=1}^{n}(Max\ (0, w_i - 1))^2 + \sum_{i=1}^{n}(Max\ (0, -w_i))^2$

To make sure that the constraints in the equation are met, the penalty function is applied. According to the constraints, short sales are prohibited, and the share weight must be positive. As a result, the portfolio with the lowest risk will be obtained, one with a minimal variance. The global minimal variance portfolio increases with the value of the penalty function. The penalty function is multiplied by 100 to speed up optimization to reach the global minimum.

Results and Discussions

The monthly returns of 222 funds are calculated using equation (1). After considering the funds with positive average return and negative skewness of returns, we have 68 funds (26 Large Cap, 20 Mid Cap, and 22 Small Cap). The average monthly returns, the standard deviation of monthly returns, and the beta values of these 68 funds are given in the Annexure. The combined SD and the average beta are respectively given by 0.045 and 0.622, using equations (2) and (3). The investor's perception map is prepared by taking SD on X-axis and beta on Y-axis with its origin at average beta (0.622) and combined SD (0.045).

Figure 1: Perception Map

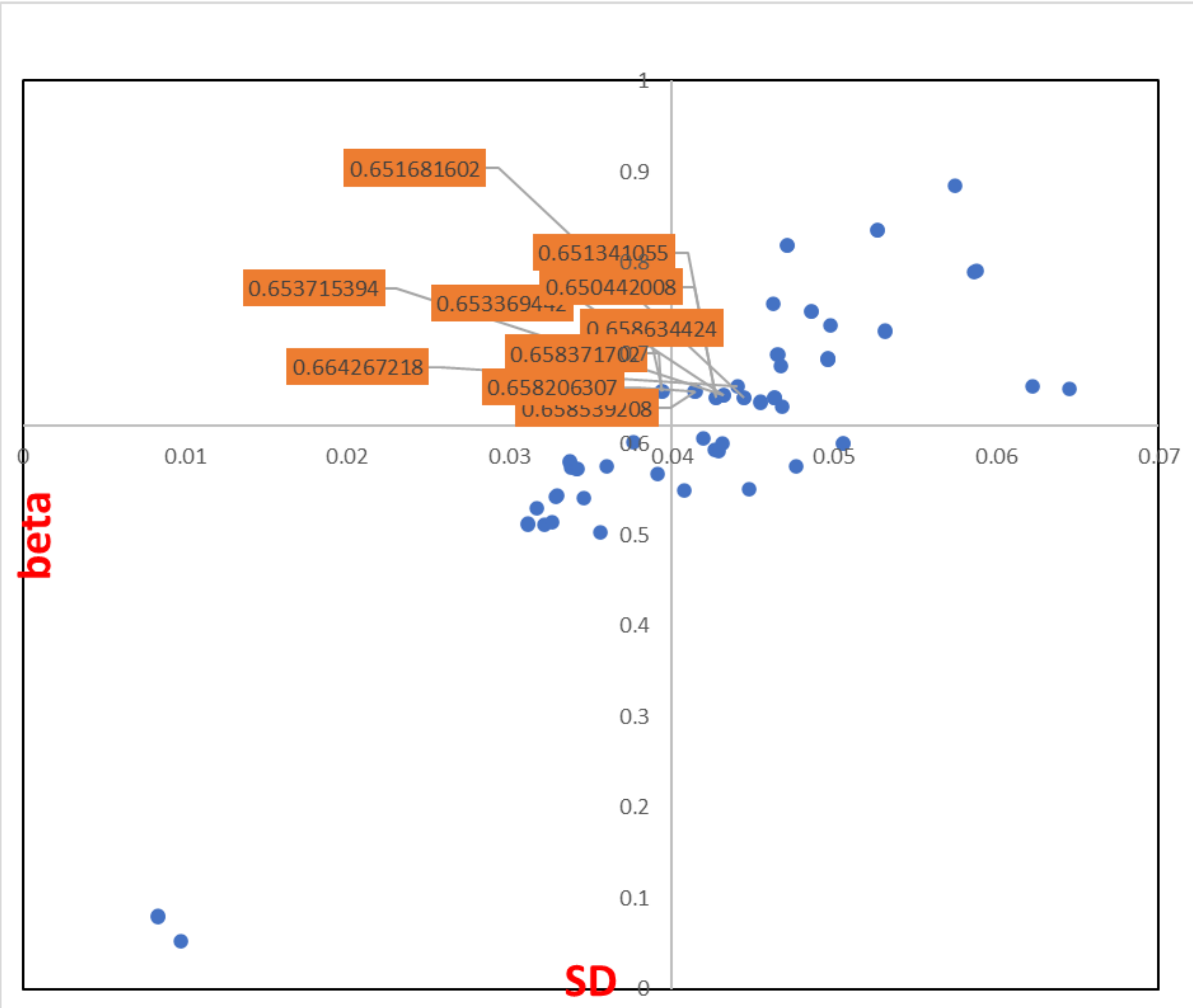

Table 1. List of funds belonging to the second quadrant along with their SD and beta

Category	Funds	Mean	SD	Beta
Large Cap Fund	ICICI Pru Nifty Next 50 Index - D (G)	0.014868	0.041443	0.658539
Large Cap Fund	ICICI Prudential Nifty Next 50 Index (G)	0.014506	0.041426	0.658206
Large Cap Fund	HDFC Top 100 Fund - Direct (G)	0.013080	0.039406	0.658634
Large Cap Fund	HDFC Top 100 Fund (G)	0.012496	0.039391	0.658372
Mid Cap Fund	L&T Midcap Fund -Direct (G)	0.020221	0.043211	0.653715
Mid Cap Fund	L&T Midcap Fund (G)	0.019500	0.043177	0.653369
Mid Cap Fund	ABSL Midcap Fund -Direct (G)	0.016938	0.042749	0.651341
Mid Cap Fund	Taurus Discovery (Midcap) - Direct (G)	0.016973	0.044086	0.664267
Mid Cap Fund	ABSL Midcap Fund (G)	0.016202	0.042706	0.650442
Mid Cap Fund	UTI Mid Cap (G)	0.016618	0.044421	0.651682

Source: authors calculation

None of the Small Cap funds falls in the second quadrant and thus are not considered in the portfolio.
The variance-covariance matrix of the monthly returns of the selected funds is then computed.

Table 2. Variance Covariance Matrix

0.001688									
0.001688	0.001687								
0.001469	0.001468	0.001527							
0.001469	0.001468	0.001526	0.001525						
0.001638	0.001638	0.00148	0.00148	0.001836					
0.001637	0.001636	0.001479	0.001479	0.001834	0.001833				
0.001631	0.001631	0.001455	0.001455	0.001773	0.001772	0.001797			
0.0017	0.0017	0.0015	0.001531	0.0018	0.0018	0.0017	0.0019		

09	09	32		08	07	84	11		
0.0016 29	0.0016 29	0.0014 54	0.001454	0.0017 71	0.0017 69	0.0017 95	0.0017 81	0.0017 93	
0.0016 46	0.0016 45	0.0014 68	0.001468	0.0018 09	0.0018 08	0.0017 68	0.0018 13	0.0017 66	0.0019 4

Source: author's calculation

The evolutionary algorithm method used to optimise portfolios yields the best fitness value of 7322.19 with 1000 generations, 50 population units, 0.8 crossover probabilities, 0.1 mutation probabilities, and elitism 2. Solutions that tend to be stable, where the fitness value does not vary in the following generation, have been discovered around the 334th generation.

Table 3: weight of shares in the formation of optimal portfolios with the genetic algorithm.

Category	Funds	Weights
Large Cap Fund	ICICI Pru Nifty Next 50 Index - D (G)	0.0994
Large Cap Fund	ICICI Prudential Nifty Next 50 Index (G)	0.1046
Large Cap Fund	HDFC Top 100 Fund - Direct (G)	0.0999
Large Cap Fund	HDFC Top 100 Fund (G)	0.0988
Mid Cap Fund	L&T Midcap Fund -Direct (G)	0.0987
Mid Cap Fund	L&T Midcap Fund (G)	0.0996
Mid Cap Fund	ABSL Midcap Fund -Direct (G)	0.0993
Mid Cap Fund	Taurus Discovery (Midcap) - Direct (G)	0.1004
Mid Cap Fund	ABSL Midcap Fund (G)	0.1007
Mid Cap Fund	UTI Mid Cap (G)	0.0987

Source: Author's calculation

The portfolio return is 0.016, which is quite high considering monthly data and the portfolio risk is 0.0016, which is fairly low. As can be seen, the optimal portfolio's proportions or stock weight obtained using the genetic algorithm accounts for stock risk as well as the average return on shares.

Table 4: Correlation matrix

1	1.688E-3	1.688E-3	1.469E-3	1.469E-3	1.64E-03	1.64E-03	1.63E-03	1.71E-03	1.63E-03
1.688E-3	1	0.99998	-0.48344	-0.48194	0.81708	0.8164	0.83517	0.90901	0.8343
1.688E-3	0.99998	1	-0.48352	-0.48203	0.81822	0.81754	0.8368	0.91015	0.83595
1.469E-3	-0.48344	-0.48352	1	0.99998	-0.54609	-0.5447	-0.62521	-0.40392	-0.62651
1.469E-3	-0.48194	-0.48203	0.99998	1	-0.54415	-0.54283	-0.62414	-0.40239	-0.62545
1.64E-03	0.81708	0.81822	-0.54609	-0.5441	1	0.999994	0.97576	0.95363	0.9755
1.64E-03	0.8164	0.81754	-0.54477	-0.54283	0.99999	1	0.97561	0.95372	0.97535
1.63E-03	0.83517	0.8368	-0.62521	-0.62414	0.9757	0.9756	1	0.94915	0.999995
1.71E-03	0.90901	0.91015	-0.40392	-0.40239	0.95363	0.95372	0.94915	1	0.94848
1.63E-03	0.8343	0.83595	-0.62651	-0.62545	0.9755	0.97535	0.999995	0.94848	1

Source: Author's calculation

The value of the determinant is 8.71E-53, which is almost zero.
Thus, it is evident that the portfolio is diversified, and thus unsystematic risk is eliminated. (Markowitz, (1952); Gupta et al., (2019)).

Conclusion

Mutual funds have become the most popular financial investment avenues for the diversity of investment since they can disperse investment risks to the smallest degree The key to gaining good profit from mutual fund investment is to understand where and in what proportion to invest in. In this study, an efficient portfolio of Indian Equity Mutual funds has been determined through dynamic allocation of the weights to the funds to minimize the portfolio risk.

All the Equity funds consisting of Large Cap, Mid Cap, and Small Cap funds with at least 5 years of operation in the Indian Mutual Fund market have been considered.10 funds have been selected applying two-stage criteria i.e., i) funds having positive average return and negative skewness of return and ii) funds that fall in the second quadrant of the Investor's Perception Map. Then, the portfolio is constructed by allocating weights to these selected funds ensuring minimum risk and maximum return using the Genetic Algorithm. The results are intriguing and support the effectiveness of the algorithm due to its quick convergence to a better solution and interesting processing time, which makes it capable of providing a better potential for finding an ideal portfolio with high returns and low risks. To increase the performance of the created portfolios, additional study is required to compare the performance of the genetic algorithm with alternative portfolio optimization models. Moreover, future research is required regarding forecasting the performance of these efficient funds with time by building appropriate statistical models.

Declaration

Funding: The authors did not receive support from any organization for the submitted work

Conflict of Interest: The authors have no conflicts of interest to declare that are relevant to the content of this article

Data Availability Statement: www.amfiindia.com, www.cmie.com, www.rbi.org.in, www.groww.in

Acknowledgment: The author extends his appreciation to the anonymous reviewers for their valuable suggestions.

Reference

1. Aranha, C., & Iba, H. (2009). The memetic tree-based genetic algorithm and its application to portfolio optimization. *Memetic Computing*, *1*(2), 139-151. [2] O.B. Augusto, S. Rabeau, Ph. Dépincé and F. Bennis, Multi-objective
2. Bakhtyar, R., Meraji, S. H., Barry, D. A., Yeganeh-Bakhtiary, A., & Li, L. (2012). An application of evolutionary optimization algorithms for determining concentration and velocity profiles in sheet flows and overlying layers. *Journal of offshore mechanics and Arctic engineering*, *134*(2).
3. Bouktir, T., Slimani, L., & Belkacemi, M. (2004). A genetic algorithm for solving the optimal power flow problem. *Leonardo Journal of Sciences*, *4*, 44-58.
4. Chang TJ, Yang SC, Chang KJ. (2009). Portfolio optimization problems in different risk measures using genetic algorithm. Expert Syst Appl 36(7):10529–10537.
5. Davidson, M., (2011). Portfolio Optimization and Linear Programming, *Journal of Money, Investment and Banking*, 129-137.
6. Ehrgott M, Klamroth K, Schwehm C. (2004) An MCDM approach to portfolio optimization. Eur J Oper Res 155(3):752–770.
7. Fernández, A., & Gómez, S. (2007). Portfolio selection using neural networks. *Computers & operations research*, *34*(4), 1177-1191.

8. Garkaz, M. (2011). The selection and optimization of stock portfolio using genetic algorithm based on mean-semi variance model. In *International Conference on Portfolio Selection Using Genetic Economics and Finance Research, IPEDR, LACSIT Press, Singapore* (Vol. 4, pp. 379-381). *(2011), 379-381*.

9. Goldberg, D., (2012). *Genetic Algorithms in Search, Optimization, and Machine Learning*.

10. Gupta, S., Bandyopadhyay, G., Bhattacharjee, M., Biswas, S. (2019). International Journal of Innovative Technology and Exploring Engineering, 8(10), 4078–4087.

11. Haupt, R. L., & Haupt, S. E. (2004). *Practical genetic algorithms*. John Wiley & Sons.

12. Jh, H. (1975). Adaptation in natural and artificial systems. Ann Arbor.

13. Lin, C. M., & Gen, M. (2007). An effective decision-based genetic algorithm approach to multiobjective portfolio optimization problem. *Applied mathematical sciences*, *1*(5), 201-210.

14. Lin, D., Li, X., & Li, M. (2005, August). A genetic algorithm for solving portfolio optimization problems with transaction costs and minimum transaction lots. In *International Conference on Natural Computation* (pp. 808-811). Springer, Berlin, Heidelberg.

15. Mahfoud, S., & Mani, G. (1996). Financial forecasting using genetic algorithms. *Applied artificial intelligence*, *10*(6), 543-566.

16. Markowitz, H. M. (1952). Portfolio Selection, the journal of finance. 7 (1).

17. Markowitz, H. M. (1991). Foundations of portfolio theory. *The journal of finance*, *46*(2), 469-477.

18. Nagy G, Barta G, Henk T. (2015). Portfolio optimization using local linear regression ensembles 251 in RapidMiner. arXiv:1506.08690.

19. Nanda SR, Mahanty B, Tiwari M.K. (2010). Clustering Indian stock market data for portfolio 255 management. Expert Syst Appl 37(12):8793–8798.

20. Nazemi A, Abbasi B, Omidi F. (2015). Solving portfolio selection models with uncertain returns using an artificial neural network scheme. Appl Intel 42:609–621. https://doi.org/10.1007/ 243 s10489-014-0616-z.

21. Pereira, R. (2000). Genetic algorithm optimisation for finance and investments.

22. Petridis, V., Kazarlis, S., & Bakirtzis, A. (1998). Varying fitness functions in genetic algorithm constrained optimization: the cutting stock and unit commitment problems. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *28*(5), 629-640.

23. Sharpe, W. F. (1970). *Portfolio theory and capital markets*. McGraw-Hill College.

24. Sefiane, S., & Benbouziane, M. (2012). Portfolio selection using genetic algorithm.
25. Yang, X. (2006). Improving portfolio efficiency: A genetic algorithm approach. *Computational Economics*, *28*(1), 1-14.