Portfolio Optimization Using Sharpe Ratio with Genetic Algorithm

A Thesis submitted in partial fulfillment for the

Degree of **Master of Computer Application** of

Jadavpur University

By

Rohan Kumar Pandey

Registration No.-149869 of 2019-20

Roll No. - 001910503006
Exam Roll No. - MCA226006

Under the Guidance of

Prof. Susmita Ghosh

Department of Computer Science and Engineering

Jadavpur University

Kolkata-700032

To Whom It May Concern

I hereby recommend that the thesis entitled "**Portfolio Optimization Using Sharpe Ratio with Genetic Algorithm**" has been carried out by Rohan Kumar Pandey (University Registration No.: 149869 of 2019-20, Class Roll No: 001910503006) under my guidance and supervision and be accepted in partial fulfillment of the requirement for the Degree of Master of Computer Application in the Department of Computer Science and Engineering, Jadavpur University.

Prof. Susmita Ghosh

Department of Computer Science and Engineering

Jadavpur University, Kolkata-32

Prof. Anupam Sinha

Head of the Department

Department of Computer Science and Engineering

Jadavpur University, Kolkata-32

Prof. Chandan Mazumdar

Dean

Faculty of Engineering and Technology

Jadavpur University, Kolkata-32

FACULTY OF ENGINEERING AND TECHNOLOGY

JADAVPUR UNIVERSITY

<u>Certificate of Approval</u>

This is to certify that the thesis entitled "**Portfolio Optimization Using Sharpe Ratio with Genetic Algorithm"** is a bona-fide record of work carried out by Rohan Kumar Pandey in partial fulfillment of the requirements for the award of the degree of Master of Computer Application in the Department of Computer Science and Engineering, Jadavpur University. It is understood that by this approval the undersigned do not necessarily endorse or approve any statement made, opinion expressed or conclusion drawn therein but approve the thesis only for the purpose for which it has been submitted.

Signature of Examiner 1

Date:

--- -

Signature of Examiner 2

Date:

FACULTY OF ENGINEERING AND TECHNOLOGY

JADAVPUR UNIVERSITY

<u>Declaration of Originality and Compliance of Academic Ethics</u>

I hereby declare that this thesis entitled "**Portfolio Optimization Using Sharpe Ratio with Genetic Algorithm"** contains literature survey and original research work by the undersigned candidate, as part of his Master of Computer Application studies. All information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by thesis rules and conduct, I have fully cited and referenced all materials and results that are not original to this work.

Name- Rohan Kumar Pandey
Registration No. –149869 of 2019-20

Class Roll No. –001910503006

Exam Roll No.: MCA226006

Thesis Title – "Portfolio Optimization Using Sharpe Ratio with Genetic Algorithm"

Signature with Date

Acknowledgement

The writing of the thesis as well as the related work has been a long journey with input from many individuals, right from the first day till the development of the final project. I would like to express my deepest gratitude to my supervisor, Prof. Susmita Ghosh, Professor, Department of Computer Science and Engineering, Jadavpur University for her admirable guidance, care, patience and for providing me with an excellent atmosphere for doing research. Our numerous scientific discussions and her many constructive comments have greatly improved this work. I feel deeply honored that I got the opportunity to work under her. I would like to thank all the faculty members of the Department of Computer Science and Engineering of Jadavpur University for their continuous support.

This thesis would not have been completed without the inspiration and support of a number of wonderful individuals including my batch mates of Master of Computer Application in Jadavpur University — my thanks and appreciation to all of them for being part of this journey and making this thesis possible.

_ _

Name:-Rohan Kumar Pandey
Registration No.: 149869 of 2019-20
Class Roll No.: 001910503006
Exam Roll No.: MCA226006
Department of Computer Science & Engineering, Jadavpur University

CONTENTS

Chapter 1

1. Introduction

Portfolio is an appropriate collection of investments held by an individual or a financial institution. These investments or financial assets constitute shares of a company (often referred as equities), government bonds, fixed income securities, commodities (such as Gold, Silver, etc.), derivatives (incl. options, futures and forwards), mutual funds, and, various mathematically complex and business driven financial instruments [1, 2].

Theory of Modern Portfolio was introduced in 1952 by Harry Markowitz. According to Markowitz an optimal set of weights is one in which the portfolio achieves an appropriate expected rate of returns with minimum volatility [1].

Markowitz proposed the Mean-Variance (M-V) model in his Portfolio Theory, which defines variance as a measure of economic risk [1]. Alternative risk measures such as Value-at-Risk (VaR), Conditional Value-at-Risk (CVaR), Mean-Absolute Deviation (MAD), Minimax (MM), and Variance with Skewness (VwS) portfolio optimization models have been suggested in the portfolio theory literature [2, 3].

Due to the lack of constraints such as boundary constraints [4], cardinality constraints [5], Transaction cost and transaction lots create the limitation for an investor and weaken the direct applicability of basic Portfolio optimization models. Unconstrained portfolio optimizations can be solved by linear and quadratic programming.

1.1 Brief of Portfolio Optimization

Portfolio optimization selects assets that show promise on their own, then looking at how the assets are connected to each other to create a portfolio that will meet financial needs [6].
All investment professionals are dependent on portfolio optimization techniques grounded in Modern Portfolio Theory (MPT) to structure investment portfolios for individual investors [6]. Using statistical techniques and computer-assisted modeling, investment advisers are able to combine different types of assets such as stocks, bonds, real estate, and hedge funds to create portfolios that claim to offer the best possible return for specified level of risk, or to minimize the amount of risk for an investor and assume to achieve a specified amount of return.

1.2 Brief of Genetic Algorithm

Portfolio optimization using genetic algorithms gives better or superior results compared to other heuristic [16].
Genetic Algorithm has been developed by John Holland, his students and his colleagues at the Michigan University in 1975. Book “Adaptation in Natural and Artificial Systems” was the first by them that discusses Genetic Algorithm. Genetic Algorithm is search algorithm based on mechanics of natural selection and natural genetics. Firstly, they combined survival of the fittest among string structures with a structured yet randomized information exchange to make a search algorithm with some of the innovative flair of human touch. In every new generation (iteration) a new set of strings is created using bits and pieces of fittest of the old. Randomized, genetic algorithm is not simple random walk. They efficiently exploit historical information to make guesses on new search points with expected improved performance [7]. The Central theme of research on genetic algorithm has been robustness the balance between efficiency and efficacy necessary for survival in many different environments. Genetic Algorithms are theoretically & empirically proven to provide robust search in complex spaces. This algorithm are computationally simple and yet powerful in their search for improvement.

1.3 Literature of survey

Portfolio optimization is a process where an investor receives the right guidance with respect to selection of assets from the range of other options. Modern portfolio theory was introduced by Harry Markowitz [1]. In 1990, Harry Markowitz shared Nobel Prize in Economics for his work on modern portfolio theory. After Markowitz many researchers worked on this domain and came across different model for improvement. For the completion of my work “Portfolio Optimization using Genetic Algorithm”, I took help from many thesis and their work. In every model there is some room for improvement and researchers improves those work later. Like in mean variance model there is some drawback and those drawbacks were removed by the Konno and Yamazaki in their Mean absolute deviation model [13]. In different model different parameter is used for risk measurement. In mean-variance model, variance is measured as risk. In mean absolute deviation model, mean of the absolute deviations of the portfolio return in all periods as risk measure. In min max model proposed by Young, risk is measured by the minimum return over all periods.

1.4 Scope of the Thesis

Investing in any stocks, fund, assets is very tedious and risky matter. Investor should have a good knowledge of the market and the previous year return of a particular stock. So for a normal investor it is a quite challenging for choosing the correct stocks to invest. In this work, we aim to maximize the sharpe-ratio for a certain number of given stocks. Thus this work can simplify the investment plan for an investor.

Chapter 2

2. Sharpe Ratio

Sharpe ratio is the measure of risk-adjusted return of a financial portfolio. A portfolio with a higher Sharpe ratio is considered superior relative to its peers [8]. The measure was named after William F Sharpe, a Nobel laureate and professor of finance, emeritus at Stanford University.

Sharpe ratio is a measure of excess portfolio return over the risk-free rate relative to its standard deviation. Normally, the 90-day Treasury bill rate is taken as the proxy for risk-free rate.

The formula for calculating the Sharpe ratio is {R (p) – R (f)} /s (p)

Where

R (p): Portfolio return

R (f): Risk free rate of return

s (p): Standard deviation of the portfolio

Chapter 3

3. Genetic Algorithm

A **genetic algorithm** is a search heuristic that is inspired by Charles Darwin's theory of natural evolution. This algorithm reflects the process of natural selection where the fittest individuals are selected for reproduction in order to produce offspring of the next generation [7].

3.1 Chromosome representation

In Genetic Algorithm, a chromosome is a set of parameters that represent a solution to a given problem. Each such parameter is called gene and it is the basic unit of information within the evolutionary process [9]. For the optimization problems considered in this project report, the chromosome represents the weights of the assets in the portfolio. If the portfolio is made up of N risky assets, then the chromosome is represented by an array of N positive real numbers, whose sum is 1, each of which represents a percentage of wealth. Let **w** be a chromosome, then

$$\mathbf{w} = \{w_1, w_2, \cdots, w_{N-1}, w_N\}$$

Each gene, w_i, of the chromosome represents the wealth invested in an asset, i, and it must be positive [11].

3.2 Fitness function

For each round of population generation, the fitness value is evaluated for each chromosome in the population. In other words, the fitness function evaluates the evolved portfolio of weights. For example, if the task is to find the minimum variance portfolio, then the fitness value is the standard deviation of the portfolio and this would be minimized for the best solution [11]. If, on the other hand, the task is to find the largest Sharpe ratio, then the fitness value is the Sharpe ratio for the portfolio calculated as if the asset's allocation is as described by the chromosome [10].

3.3 Genetic operators

Crossover, elitism, gene mutation and new random genes have been used in this project to generate the new population. Crossover combines the genes of two parents to generate one or more offspring. Crossover causes the exchange of genetic materials between the two parents with the possibility that the new genes generated this way are fitter (e.g. have higher fitness function value for the maximization problem) than their parents. It can significantly speed up the evolutionary process in terms of number of generations required to find the optimal solution. Elitism ensures that the fittest individuals of one generation are carried on to the next. Mutation operator involves changing the genes of the chromosome [10, 12].

3.3.1 Mutation operator

To generate one or more new chromosomes, the mutation operator modifies one or more genes from the parent's chromosome. All other genes, however, are normalized [10] because the sum of the weights in the chromosome must be one. In this project report, two different mutation operators were used. Gene mutation (GM-) and Bump mutation (BM-) are two types of mutations. These mutation operators are applied to the chromosomes of the fittest parents [9].

In GM-μ, for each parent, two offspring are generated in the following way:

- The largest gene (weight) is increased and the smallest gene is decreased by a factor μ;
- The largest gene (weight) is decreased and the smallest gene is increased by a factor μ.

In BM-μ, for each parent, two offspring are generated by modifying a percentage of the parent's chromosome. A percentage, x, of the parent's chromosome is randomly selected, then half of those genes are increased and half are decreased by a factor μ.

The factor μ of the GM-μ and BM-μ can be selected by the user, and its value is between 0 and 1 [9].

3.3.2 Crossover operator

The crossover operators used in this project is the single point crossover, the flat crossover [10] and the blend crossover [10]. The single point crossover randomly selects a crossover point, then the parents are cut at the crossover point (randomly selected) and subsequently recombined with a piece of each other to generate two descendants

The flat crossover, introduced in [12], generates offspring from parents by selecting a random number from the range and using a linear combination of parents' genes, such that child's gene, i, is computed as follows:

$$w_{i,child} = \lambda_i \cdot w_{i,parent_A} + (1 - \lambda_i) \cdot w_{i,parent_B}$$

For the standard flat crossover λ is a uniform random number between 0 and 1. In this project's experiments the value of flat crossover lambda set to 0.7 has produced the best results.

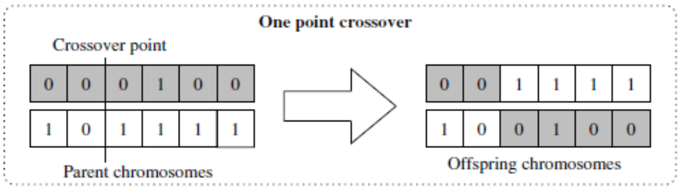
The blend crossover (BLX-α) was introduced in [11] as a generalization of the flat crossover. The blend crossover generates offspring by uniformly picking values within an interval that contains the two parents. The lower and upper bound of the interval are calculated from the parent's genes as follows:

$$\text{UpperBound}_i = \max(w_{i,parent_A}, w_{i,parent_B}) + \alpha \cdot d_i$$
$$\text{LowerBound}_i = \min(w_{i,parent_A}, w_{i,parent_B}) - \alpha \cdot d_i$$
$$d_i = | w_{i,parent_A} - w_{i,parent_B} |$$

where the coefficient, α, is selected within the interval [0, 1]. Several simulations with different portfolios and different time horizons have been run in order to find the best value for the coefficient [10].

Single Point Crossover

In a single-point crossover, a crossover point is generated at random, determining the point at which parents exchange information to form children.

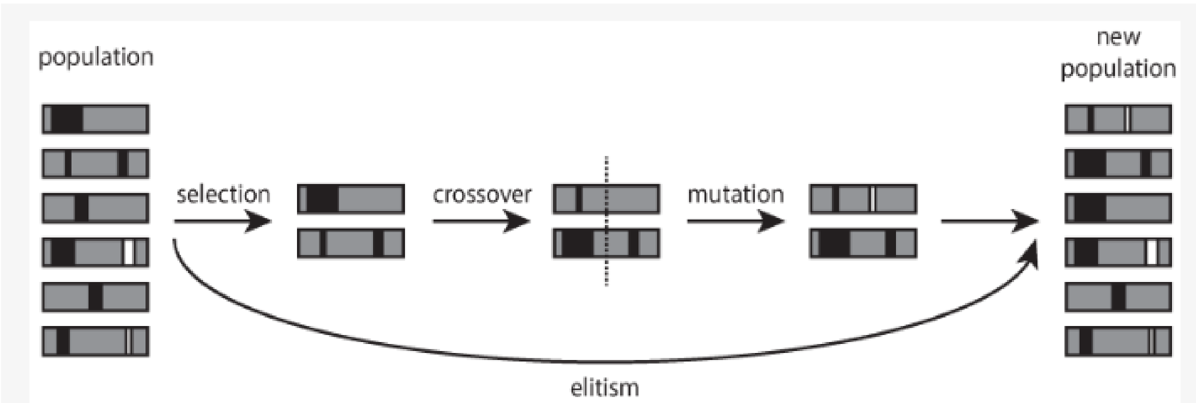


Source: https://medium.com

3.3.3 Elitism operator

Elitism employs the fittest members of the current generation as members of the following generation. The number of individuals copied to the next generation in the (μ + λ) evolutionary strategy is, and unless otherwise stated, = 20% is used in the numerical experiments in this project [9].

<u>Flowchart</u>:



Source: https://www.mdpi.com/2078-2489/11/12/587/htm

Chapter 4

4. Whale Optimization Algorithm:

I am comparing GA best result with WOA best result.

The Whale Optimization Algorithm (WOA) is a nature-inspired meta-heuristic optimization algorithm that mimics humpback whale hunting behavior. The bubble-net hunting strategy inspired the algorithm [13].

Bubble-net feeding is a term used to describe how humpback whales forage. Humpback whales prefer to hunt krill or small fish near the surface of the water. This foraging has been observed to be done by forming distinct bubbles along a circle or '9'-shaped path [14].

The 'upward-spirals' and 'double-loops' are two bubble net feeding maneuvers.

In the 'upward-spirals' maneuver, humpback whales dive to a depth of 12 meters, then begin to create a spiral of bubbles around their prey and swim up to the surface.

There are three stages to the 'double-loops' maneuvers: coral loop, lob tail, and capture loop.

Bubble-net feeding is a one-of-a-kind behavior seen only in humpback whales. The spiral bubble-net feeding maneuvers is mathematically modelled in the whale optimization algorithm (WOA) in order to perform optimization [14].

To chase the prey, WOA used a random or the best search agent to simulate hunting behavior.
WOA uses a spiral to simulate a humpback whale's bubble-net attack mechanism.
Surrounding the Prey

The current best candidate solution is assumed to be close to the target prey, while other solutions adjust their positions in relation to the best agent.

Encircling Prey:
The current best candidate solution is assumed to be close to the target prey, while other solutions adjust their positions in relation to the best agent [13].

Chapter 5

5. Optimizing Sharpe Ratio Using Genetic Algorithm:

I have used Genetic Algorithm to find the weights of stocks such that we maximize the returns and at the same time minimize the risk. Sharpe Ratio will be used to evaluate the fitness.

i. Read the data and combine them into one data-frame stocks from https://finance.yahoo.com using pandas data-reader I combine them and formed a data-frame

ii. Calculate the historical returns for 252 days for each of the stocks.

iii. Define Gene (Scalar): A fraction of the total capital assigned to a stock.

iv. Define Chromosome (1D Array): Set of genes i.e. fractions of total capital assigned to each stock. Sum of each chromosome should be equal to 1.

v. Generate Initial Population (2D Array): A set of randomly generated

chromosomes.

vi. Fitness function (Define a Function): The Sharpe ratio, S, is a measure for quantifying the performance (Fitness) of the portfolio which works on " Maximization of return (mean) and Minimization of risk (Variance) simultaneously " and is computed as follows: S = {R (p) – R (f)} /s (p)

Here R (p) is the return of the portfolio over a specified period or Mean portfolio return, R (f) is the risk-free rate over the same period and s (p) is the standard deviation of the returns over the specified period or Standard deviation of portfolio return.

I calculated the mean of the historical return, standard deviation of the historical return and the covariance of the historical return using the given data.

vii. Select Elite Population (Define a Function): It filters the elite chromosomes which have highest returns, which was calculated in fitness function. Here elite rate is 0.25.

viii. Crossover: When two chromosomes from the same generation (referred to as "parent chromosomes") interact, certain genes from one chromosome are swapped for genes from the corresponding positions in the other. Two new chromosomes are produced by this process.. The crossover probability is 0.4.

ix. Mutation: A function that will perform mutation in a chromosome. I used mutation probability as 0.01. We shall choose 2 numbers between 0, 5 and those elements we shall swap.

Iterate the process: Iterate the whole process till there is no change in maximum returns.

Chapter 6

6. Result and Analysis

I have executed both Genetic Algorithm and Whale Optimization Algorithm 10 times with 5 different population sizes 100,200,300,400 and 500 iterations and also the average of 10 times is calculated.
The best results of both the Genetic Algorithm and Whale Optimization Algorithm are listed in the below tables:

Table of results of Sharpe ratio optimization using Genetic Algorithm:

No. of Chromosomes	Sharpe Ratio It 1	Sharpe Ratio It 2	Sharpe Ratio It 3	Sharpe Ratio It 4	Sharpe Ratio It 5	Sharpe Ratio It 6	Sharpe Ratio It 7	Sharpe Ratio It 8	Sharpe Ratio It 9	Sharpe Ratio It 10	AVG
100	1.5451	1.5457	1.5454	1.546	1.5458	1.5456	1.5457	1.5452	1.5451	1.5458	1.5455
200	1.5485	1.5491	1.5488	1.5494	1.5492	1.549	1.5491	1.5486	1.5485	1.5492	1.5489
300	1.5489	1.5495	1.5492	1.5495	1.5496	1.5494	1.5498	1.549	1.5489	1.5496	1.5493
400	1.5434	1.544	1.5437	1.5443	1.5441	1.5439	1.544	1.5435	1.5434	1.5441	1.5438
500	1.5468	1.5474	1.5471	1.5477	1.5475	1.5473	1.5474	1.5469	1.5468	1.5475	1.5472

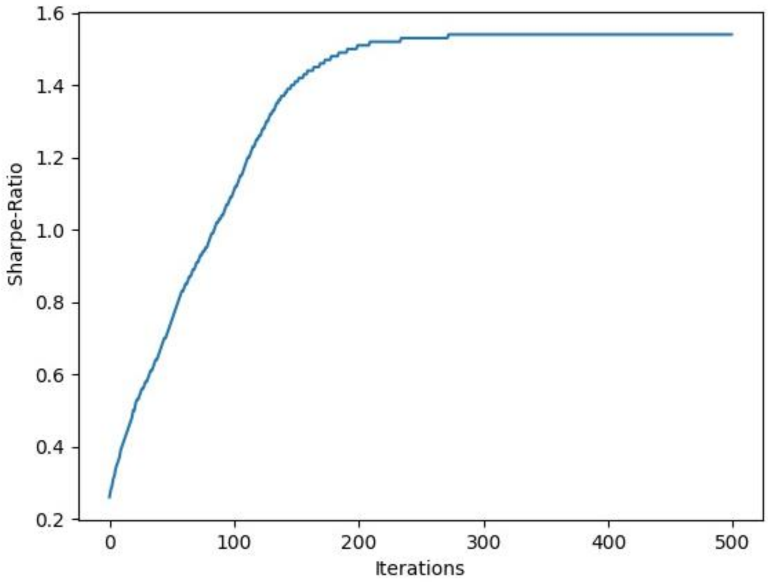
Table of Whale Optimization Algorithm:

No. of Populations	Sharpe Ratio It 1	Sharpe Ratio It 2	Sharpe Ratio It 3	Sharpe Ratio It 4	Sharpe Ratio It 5	Sharpe Ratio It 6	Sharpe Ratio It 7	Sharpe Ratio It 8	Sharpe Ratio It 9	Sharpe Ratio It 10	AVG
100	0.9856	0.9754	0.9778	0.9987	1.0032	0.9898	0.9878	0.9768	0.9845	0.9687	0.98483
200	0.9789	0.9997	0.9968	0.9779	0.9812	0.9661	0.9732	0.9665	0.9928	0.9888	0.98219
300	0.9989	1.076	0.9891	0.9845	0.9934	0.9712	0.9768	0.9718	0.9923	0.9823	0.99363
400	0.9768	0.9718	0.9923	0.9665	0.9928	0.9888	1.076	0.9891	0.9823	0.9887	0.99251
500	0.9991	0.9789	0.9997	0.9768	0.9845	0.9687	0.9779	0.9812	0.9661	0.9661	0.9799

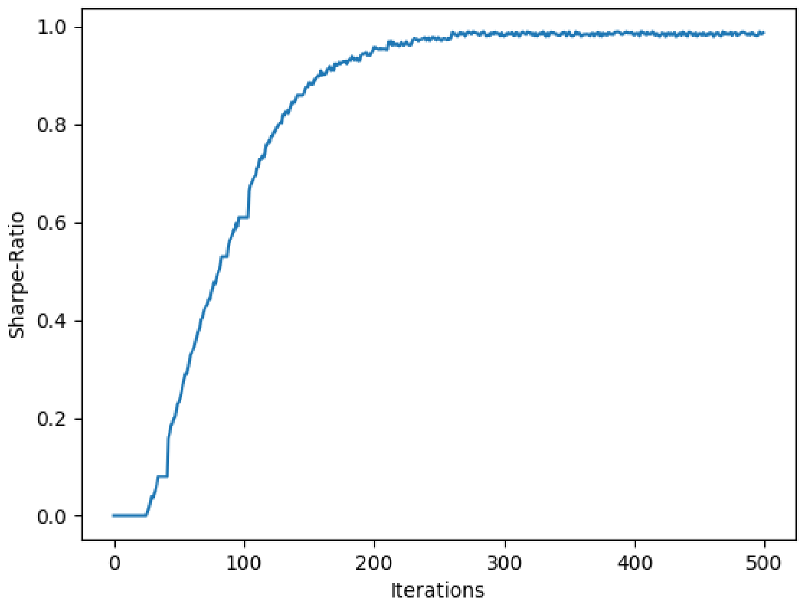
By looking at the above results generated by Genetic Algorithm and Whale Optimization Algorithm for 5 different population sizes executed 10 times each we observed that:
The best result was given by the Genetic Algorithm in 300 chromosome size and 7th Iteration 1.5498.
The best result of Whale Optimization Algorithm was in 300 population size and 1st iteration 0.9989

However considering all the iterations of each population and calculating
Average of the best result was given by GA is 1.5493 and the best average in Whale Optimization Algorithm is 0.99363.
GA was more consistent in results in comparison to Whale Optimization Algorithm.

The graphs of best individual result of both PSO and GA are given in the page below:



Graph of Best Result of Genetic Algorithm



Graph of Best Result of Whale Optimization Algorithm

Chapter 7

7. Conclusion:

This project report presents genetic algorithm and investigates how it can be utilized to solve optimization problems for portfolios. It has been shown that evolutionary algorithm can be used to maximize the sharpe-ratio in a relative short time frame.

I have compared GA with WOA, The best result in the individual iteration was given by GA.

Chapter 8

8. References:

1. H.M Markowitz , Portfolio selection,J-finance,1952.
2. H M Markowitz, Efficient Diversification of Investment, Yale university press,1959.
3. P. Jorion, Value at Risk: the new Benchmark for controlling Market Risk, Irwin Professional Pub,1997.
4. D. Blenstock, Computational study of a family of mixed-integer quadratic programming problems, Math. Program (1996) 121-140.
5. M.J. Magill, G.M. Constantinides, Portfolio Selection with transaction costs,J. Econ. Theor.(1976) 245-263.
6. Financial Portfolio Optimization: Computationally guided agents to investigate, analyse and invest!? By Ankit Dangi CENTRE FOR MODELING AND SIMULATION UNIVERSITY OF PUNE
7. Article SAIPO-TAIPO and Genetic Algorithms for Investment Portfolios Juan Frausto Solis * , José L. Purata Aldaz, Manuel González del Angel, Javier González Barbosa and Guadalupe Castilla Valdez
8. https://economictimes.indiatimes.com
9. Genetic Algorithm Application to Portfolio Optimisation Emanuele Stomeo
10. N. J. Radcliffe, "Genetic neural networks on MIMD computers", Ph.D. Thesis (Edinburgh University), 1990.
11. L. J. Eshelman and J. D. Schaffer, “Real-coded genetic algorithms and interval schemata,” Foundations of Genetic Algorithms, 1993, 2, 185–202
12. N. J. Radcliffe, “Equivalence class analysis of Genetic Algorithms,” Complex Systems, 1991, 5-2, 183-205
13. Mean-Absolute Deviatzon Portfolio Optimization Model Under Transaction Costs Hiroshi Konno Annista Wljayanayake Tokyo Institute of technology
14. https://www.sciencedirect.com/science/article/pii/S0965997816300163
15. Whale Optimization Algorithm for Multiconstraint Second-Order Stochastic Dominance Portfolio Optimization Q. H. Zhai,1 T. Ye,2 M. X. Huang,3,4 S. L. Feng , 4 and H. Li4
16. Woodside-Oriakhi M, Lucas C and Beasley JE ,Heuristic Algorithms for the Cardinality