



# Multi-verse metaheuristic and deep learning approach for portfolio selection with higher moments

Veena Jain<sup>1</sup> · Rishi Rajan Sahay<sup>2</sup> · Nupur<sup>3</sup>

Received: 21 June 2023 / Revised: 18 November 2023 / Accepted: 22 November 2023 / Published online: 27 December 2023  
© The Author(s) under exclusive licence to The Society for Reliability Engineering, Quality and Operations Management (SREQOM), India and The Division of Operation and Maintenance, Lulea University of Technology, Sweden 2023

**Abstract** The market has become very volatile these days in the presence of a war-like situation with a lot of political turmoil and the rapid occurrence of natural disasters the world over. It is difficult to predict the economic condition of the country and hence the company's financial position. This paper proposes a novel approach that integrates clustering techniques, deep learning, and a metaheuristic algorithm to enhance the process of asset selection and allocation. First, S&P BSE 500 index companies have been clustered into ten groups by using the Expectation Maximization (EM) clustering technique based on 11 fundamental characteristics of the companies. The Prowess financial database has been used to collect the required data. For diversification of the portfolio across clusters and sectors, the best-performing companies are chosen based on Sharpe Ratio. Advanced analytical tools like machine learning and deep learning have been employed to increase the accuracy and precision of estimating the returns on the stocks of the selected companies. The expected return on stocks of these selected companies has been estimated with the help of Neural Basis Expansion Analysis for Interpretable Time Series (N-BEATS), a deep learning neural network-based forecasting technique.

A portfolio multi-objective optimization model has been formulated by considering entropy and higher moments like skewness and kurtosis in the objective function. A metaheuristic algorithm named multi-verse is used to solve the optimization model, and hence the selection of the assets with their proportion of investment in the portfolio has been suggested under different scenarios.

**Keywords** Entropy · Higher-moments · Expectation maximization · N-BEATS · Multi-verse · Portfolio optimization

## 1 Introduction

The selection of risky assets for investment to be included in a portfolio that can optimize the return and risk of an investor is a herculean task in the current scenario of political instability, economic upheaval, increased terrorism, and supply chain disruption due to natural calamities and epidemics throughout the world. Due to the increasing uncertainty in economic conditions, while companies are facing recession and laying off staff to stay afloat in the market, the stock trading market has also become unpredictable, and hence the formation of an optimal portfolio of risky assets has become tough. In this context, the Modern Portfolio Theory developed by Harry Markowitz (1952) remains valuable. The theory provides a solid foundation for investors not only during normal times but also during times of uncertainty and crisis.

In the mean–variance portfolio framework of Markowitz, diversified portfolio development secures the highest possible expected return for a given degree of risk tolerance (Cheong et al. 2017). By clustering the assets based on certain characteristics and selecting assets from different clusters, one can ensure that the portfolio includes a

✉ Nupur  
cnupur91@gmail.com

Veena Jain  
veenaace@yahoo.com

Rishi Rajan Sahay  
rajansahay@sscsdsu.ac.in

- 1 Deen Dayal Upadhyaya College, University of Delhi, New Delhi, India
- 2 Shaheed Sukhdev College of Business Studies, University of Delhi, New Delhi, India
- 3 Department of Operational Research, University of Delhi, New Delhi, India

well-diversified mix of assets. Moreover, the experimental results have shown that the use of clustering algorithms can improve the reliability of the portfolio (Long et al. 2014). Chen and Huang (2009) and Nanda et al. (2010) applied clustering in their work to mitigate the complexity of diversification. Since then, various authors have employed clustering techniques to address the diversification of stocks in portfolio selection problems (Ashfaq et al. 2021; Long et al. 2014).

Forecasting the returns of individual stocks is a crucial step in the portfolio construction process. The theoretical studies have shown that mean–variance portfolio selection problems are very sensitive to small forecast errors in the means and covariances (Du 2022; Chopra et al. 1993; Goldfarb and Iyengar 2003). In recent times, researchers have shown growing interest in extending the mean–variance model by improving the accuracy of the expected returns using various forecasting methods (Ashrafzadeh et al. 2023; Du 2022; Wu et al. 2021; Gu et al. 2020). Conventional forecasting techniques are not suitable for time series data with non-linearity and non-stationarity. In such situations, deep learning machine algorithms outperform the conventional methods of forecasting (Du 2022).

Selecting the right forecasting technique is crucial, but improving model performance also depends on various other aspects such as determining an initial approximate solution, optimizing the model's meta-parameters, the training approach, etc. To achieve the best results, researchers have introduced several metaheuristic methods designed to search for the optimal set of hyperparameters. Metaheuristic techniques also enable the researchers to approximate optimal solutions to the portfolio optimization problem in an efficient manner (Erwin and Engelbrecht 2023). Some of these algorithms include the genetic algorithm (Gupta 2022; Cheong et al. 2017; Chang et al. 2009), the firefly algorithm (Wang and Liu 2019), and the particle swarm optimization (Wang and Liu 2019; Song et al. 2023). Unlike exact methods, which are suitable for solving simpler optimization problems under strict assumptions, metaheuristic methods are applicable to a broad range of more complex problems.

In this paper, we have proposed a novel approach that integrates clustering techniques, deep machine learning, and a metaheuristic algorithm to enhance the process of asset selection and allocation. First, data is extracted from the ProwessIQ database for S&P BSE 500 index companies. We apply the Expectation–Maximization (EM) clustering technique to categorize the S&P 500 companies into groups based on similar financial performance indicators. Subsequently, we concentrate on predicting the return of assets that have been chosen using the clustering technique. To do this, a deep neural-network based learning method called Neural Basis Expansion Analysis for Interpretable Time

Series (N-BEATS) is employed. The portfolio optimization problem considered in this paper takes numerous objectives into account, such as variance, skewness, kurtosis, and entropy. Since in real life, any individual investor or any corporation is having a limited budget for investment and having a goal of getting a minimum return on investment, we also incorporate constraints relating to mean return, capital allocation, and budget limits to improve the practicality of the problem. In the end, the multi-verse optimization (MVO) approach is utilized to solve the portfolio problem which can help the company/investor to decide about how to distribute his wealth optimally among different assets at minimum risk.

The contribution of the present work is significant for more than one reason. Firstly, to the best of our knowledge, no prior research has utilized an MVO approach for solving a Multi-objective portfolio optimization problem. Secondly, although EM clustering and the N-BEATS forecasting technique are valuable methods, not many authors have paid attention to these approaches. Furthermore, the combination of MVO, N-BEATS forecasting, and EM clustering has not been explored previously by any researcher. Consequently, the present study offers substantial value to researchers, practitioners, and investors by addressing these important research gaps.

The rest of the paper is organized as follows. Section 2 presents an in-depth analysis of the literature. Section 3 provides the theoretical background by explaining all the methods involved in this paper like the EM, N-BEATS, MVO. Section 4 explains the proposed problem formulation, notation, and assumptions. Section 5 describes the data used for the problem, the analysis of the data, and the obtained results. Section 6 concludes the paper with a brief discussion of the findings. Finally, Sects. 7 and 8 discuss about future research opportunities and limitations of the research works.

## 2 Literature review

Harry Markowitz (1952) introduced the formula for calculating the risk of the portfolio by including the covariation term between the returns of risky assets and emphasized diversification of the portfolio to reduce the overall risk by including least-correlated, zero-correlated, or even negatively correlated assets. Diversification is a portfolio allocation technique that seeks to reduce idiosyncratic risk. A perfect positive correlation between assets in a portfolio raises the portfolio's standard deviation, or risk. Portfolios can be diversified in numerous ways, like across industries, asset classes, and markets (i.e., countries).

The unsupervised machine learning technique called “clustering” also helps in diversifying the portfolio. Cluster analysis is a tool used for grouping objects that

have common features and is used by investors to create a subsystem trading strategy that assists them in building a diverse portfolio by picking stocks from different clusters. If implemented correctly, the individual clusters will have little association with one another. Investors obtain all the benefits of diversification under this setting: decreased downside losses, preservation of capital, and the opportunity to make riskier transactions without increasing overall risk. Diversification is a key tenet of investment, and clustering is simply one method for attaining it. In recent times, authors and researchers have shown growing interest in using clustering in their studies (Long et al. 2014; Cheong et al. 2017; Rezani et al. 2020; Sehgal and Jagadesh 2023; Sass and Thös, 2021; Wang and Aste 2023; Menvouta et al. 2023). K-means is one of the most widely used clustering techniques. It is easy to understand, computationally efficient, and works well when clusters are spherical and have similar sizes. Few of the recent studies on K-means clustering (Aithal et al. 2023; Navarro et al. 2023; Wu et al. 2022; Cheong et al. 2017; Nanda et al. 2010). However, it is sensitive to the initial placement of centroids, and it might converge to the suboptimal solutions in some cases. EM clustering, which is an extension of K-means, is, on the other hand, a more general framework that works well with data distributions that are not necessarily spherical or have equal sizes. It is often used in cases where the clusters have overlapping or complex shapes. EM clustering employs a probabilistic approach to clustering and is based on the expectation maximization algorithm. EM clustering is more flexible than K-means as it allows for more complex cluster shapes and sizes. It is also more robust to the choice of initial parameters due to its probabilistic nature. However, it can be computationally more intensive and might require careful initialization of parameters to converge on a good solution. A comparison of both the techniques is given by Jung et al. (2014) and Moghadaszadeh and Shokrzadeh (2018). Ng and Chin Khor (2014) built a plantation stock portfolio for Bursa Malaysia index using EM clustering technique.

To address the univariate time series forecasting problem using deep learning, Oreshkin et al. (2019) introduced a deep neural architecture, N-BEATS, which incorporates backward and forward residual connections as well as a very deep stack of completely linked layers. The M3, M4, and TOURISM competition datasets have shown cutting-edge performance for two N-BEATS configurations. A few of the studies available on N-BEATS are as follows: (Oreshkin et al. 2021; Sbrana and Lima De Castro 2023; Ma et al. 2023; Kaja et al. 2021). The only research available where N-BEATS forecasting is used in the field of finance is by Singhal et al. (2022), in which they describe a technique for improving stock market index forecasting that blends wavelet processing

with the deep learning architecture, N-BEATS. This leaves a research gap in the field.

Mirjalili et al. (2016) offered a nature-inspired algorithm called multi-verse optimizer (MVO). The main inspirations for this algorithm come from three cosmological concepts: white holes, black holes, and wormholes. These three concepts are designed mathematically to be executed: exploration, exploitation, and local search, respectively. The multi-objective multi-verse optimizer (MOMVO) is a multi-objective variation of the MVO suggested by Mirjalili et al. (2017). The competitive multiverse optimizer (CMVO), a unique population-based optimization approach, is introduced by Benmessahel et al. (2020). Although it uses a different framework, this unique approach is fundamentally based on MVO. Abualigah (2020) reviewed existing literature on MVO and presented a comprehensive survey of the work. None of the studies employed MVO to solve the portfolio optimization problem.

The mean–variance portfolio optimization model of Markowitz assumes that the return of assets follows a normal distribution, which might not be true, as proved by many researchers (Malek et al. 2009; Z. Zhu et al. 2020; Saranya and Prasanna 2014). In cases of violation of the normal distribution of return, only two moments, mean and variance, are inadequate to consider while finding an optimal portfolio. Higher-order moments like skewness and kurtosis also need proper attention in selecting an optimal portfolio of assets, as has been proven in many studies (Nguyen 2016; Abdelaziz and Chibane 2023; Sihem and Slaheddine 2014; Mirlohi et al. 2021). The introduction of higher-order moments can help in identifying assets with low correlation to the traditional risk factors, thus enhancing the diversification benefits of the portfolio (Barkhagen et al. 2023; Naqvi et al. 2017; Khan et al. 2020). To ensure portfolio diversification, entropy is another measure. It helps in diversifying the portfolio and hence in increasing its performance, and it has been used in many studies along with higher moments (Gupta et al. 2019; Gonçalves et al. 2022; Nabizadeh and Behzad 2018; Batra and Taneja 2022; Pourrafiee et al. 2020; Ji et al. 2017). Zhou et al. (2013) examined the concepts and principles of entropy as well as their applications in finance, particularly portfolio selection and asset pricing.

By taking moments of return like variance, skewness, kurtosis, and entropy into our objective function, the portfolio optimization model becomes a multi-objective portfolio selection model with conflicting objectives of maximizing skewness and entropy while minimizing variance and kurtosis. Many authors have used goal programming to solve related problems (Ashfaq et al. 2021; Siew et al. 2021; Aksaraylı and Pala 2018). Many have used metaheuristics like (Li et al. 2023; Chen and Zhou 2018). Milhomem and Dantas (2020) conducted a thorough examination of the exact and heuristic approaches, software and programming

languages, restrictions, and forms of analysis (technical and fundamental) employed in the solution of the portfolio optimization problem.

## 2.1 Motivation and contribution

An in-depth exploration of the literature review helped us identify the research gaps. Motivated by the research gap, the present study was undertaken. The main contribution of the present work lies in employing the MVO for the first time in the present context. Furthermore, the integration of the EM clustering technique and the N-BEATS forecasting method with the multi-verse portfolio optimization problem presents a novel approach. Elements such as variance, skewness, kurtosis, and Gini Simpson entropy are incorporated in the objective function for the analysis. The multi-objective problem encompasses a combination of objectives and constraints that have not been previously addressed, making it useful for complex portfolio allocation situations. Additionally, the utilization of the clustering method and forecasting techniques remains relatively unexplored by most researchers in this field. Our findings provide researchers and practitioners with valuable insights into how different combinations of objectives can impact portfolio performance (Table 1).

## 3 Methods description

### 3.1 Expectation maximization (EM) clustering method

EM is a general iterative optimization algorithm used to estimate the parameters of statistical models, particularly in situations involving missing or hidden data, as explained by Do and Batzoglou (2008). EM is a specific type of probabilistic clustering which uses the concept of Gaussian Mixture Models (GMMs).

The EM algorithm is often associated with Sir Ronald A. Fisher and developed further by other statisticians and researchers. However, its application to clustering and Gaussian mixture models can be attributed to many contributors, including Dempster et al. (1977). They introduced the algorithm and its application to statistical modelling in their paper titled "Maximum Likelihood from Incomplete Data via the EM Algorithm."

In the context of clustering, the EM algorithm for Gaussian mixture models iteratively updates the estimates of the mixture model's parameters by alternating between two steps:

1. **Expectation Step (E-step)** In this step, for each data point, the algorithm calculates the probabilities of belonging to each cluster based on the current estimates

of cluster parameters. These probabilities represent the "expectation" of the hidden or missing cluster assignments.

2. **Maximization Step (M-step)** In this step, the algorithm updates the parameters (means, variances, and mixing proportions) of the Gaussian distributions in a way that maximizes the likelihood of the observed data given the current cluster assignments.

By iteratively repeating these steps, the algorithm aims to find a set of parameters that maximize the likelihood of the observed data. This process helps in estimating the underlying cluster structure of the data.

In summary, while the EM algorithm itself is not attributed to a single individual, its application to clustering, particularly Gaussian mixture models, has been developed by a combination of researchers in the fields of statistics and machine learning.

### 3.2 N-BEATS (neural basis expansion analysis for interpretable time series)

Oreshkin et al. (2019) proposed N-BEATS, which is a deep neural network based on backward and forward residual links as well as a very deep stack of fully connected layers. It is a univariate model. The architecture (shown in Fig. 1) of the model is founded on a few fundamental ideas.

- The foundation framework should be simple, general, and descriptive (deep).
- The design should not rely on feature engineering or input scaling that is time series-specific (like trend and seasonality).
- For investigating interpretability, the architecture should be expandable so that its outputs can be easily interpreted by a human.

The data is fed into the model as a lookback period. The lookback period is the back horizon, which is used to make predictions on the forecast horizon. If the length of the forecast horizon is  $H$ , then the length of the backcast horizon should be  $2H-7H$ .

The model is divided into a collection of blocks and stacks.

**Block** A block is simply four fully connected (FC) layers that give rise to two forks. The first one attempts to recreate the back horizon input, whereas the second attempts to forecast the horizon. FC layers give rise to the  $\theta^b$  and  $\theta^f$  coefficients. They are expansion coefficients.  $g^b$  and  $g^f$  are the basis vectors. Then a linear combination of coefficients and basis vectors suffices to generate a prediction.

**Stack** A stack is made up of multiple basic blocks that are organized following the double residual stacking

**Table 1** Feature comparison of present study with existing studies in the literature

Feature	Aksarayli and Pala (2018)	Abolmakarem et al. (2023)	Pahade and Jha (2021)	Chen et al. (2020)	Ashfaq et al. (2021)	Siew et al. (2021)	Gupta et al. (2019)	Zhu et al. (2011)	Proposed approach
Problem type	Multi-objective, non-linear	Multi-objective, non-linear	Multi-objective, non-linear	Multi-objective, non-linear	Multi-objective, non-linear	Multi-objective, non-linear	Multi-objective, non-linear	Single-objective, non-linear	Multi-objective, non-linear
Mean as objective/constraint	Objective	Objective	Objective	Objective	Objective	Objective	Objective	Constraint	Constraint
Variance	✓	✓	✓	✓	✓	✓	✓	✓	✓
Skewness	✓	✓	✓	✓	✓	✓	✓	✗	✓
Kurtosis	✓	✓	✗	✓	✓	✓	✓	✗	✓
Entropy as objective/constraint	Objective	Constraint	✗	✗	✗	✗	✗	✗	Objective
Clustering technique	✗	✗	✗	✓	✗	✗	✗	✗	✓
Forecasting technique	✗	✓	✗	✓	✗	✗	✗	✗	✓
Bound constraint on assets	✗	✓	✓	✗	✗	✗	✓	✗	✓
Solution methodology	Goal programming	Goal programming	Goal programming	GA	Goal programming	Goal programming	Goal programming	PSO	MVO

GA genetic algorithm, MVO multi-verse optimizer, PSO particle swarm optimization

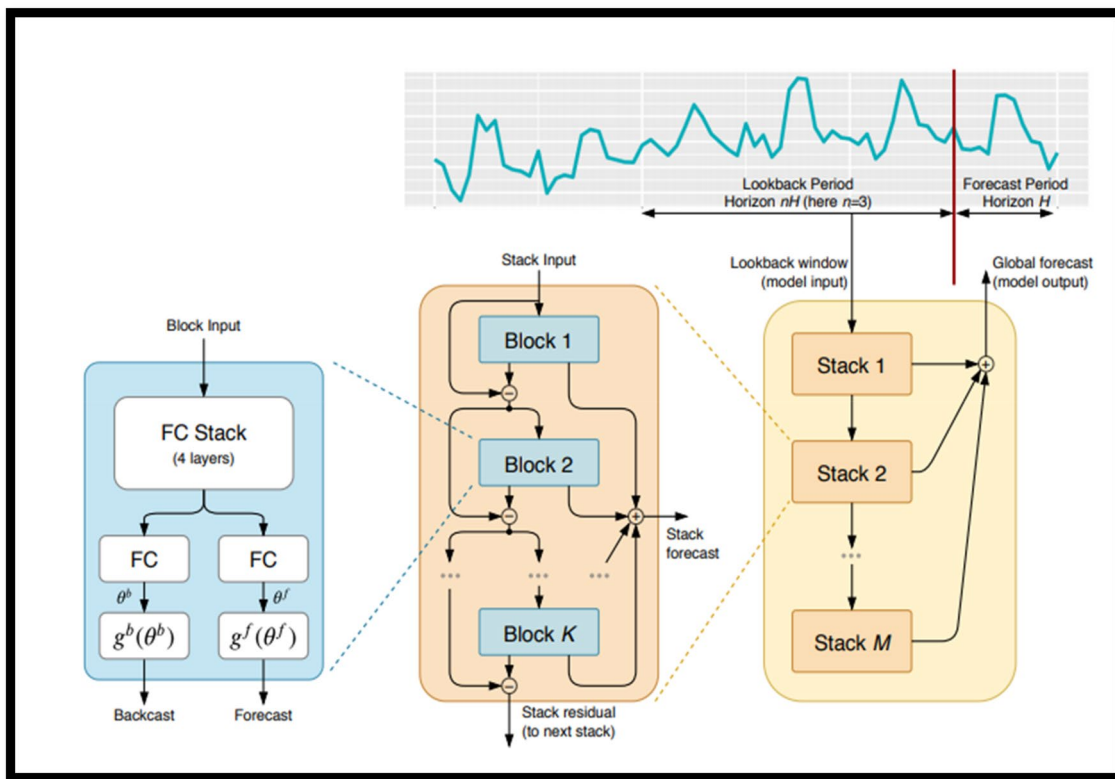


Fig.1 N-BEATS architecture (Adapted from: (Oreshkin et al. 2019))

concept. The output of the Basic Block undergoes two arithmetic operations (backcast and forecast), hence the phrase double residual stacking.

Multiple linked blocks yield better results; the subsequent blocks attempt to forecast the missing part of their predecessors, and the outputs are finally summed. These blocks form a stack, and the sum of numerous stacks yields the final output.

At this point, the model is in its generic form. The base expansion function, represented as  $g$  in the illustration, is trainable. To achieve the best outcomes, the neural network constructs a problem-specific function.

But to make the model interpretable, the author has incorporated trend and seasonality in the model in the form of polynomial and Fourier basis, respectively. As a result, in the interpretable version of the architecture, the model contains only two stacks: one for predicting a trend component and the other for forecasting a seasonal component. The predictions are then pooled to generate a final output.

The first block receives the actual input,  $n * H$ . The subsequent blocks receive the backcast of the prior block as their input.

For brevity, the mathematical aspect of the model has been explained for  $k^{th}$  block.

Suppose  $k^{th}$  block receives its input as  $x_k$  and then it gives out two outputs. One is  $\hat{x}_k$  and the other is  $\hat{y}_k$ , which are the backcast and forecast of the,  $k^{th}$  block. In the subsequent block,  $k + 1^{th}$  block receives the backcast from prior block  $k$  as its input, i.e.,  $x_{k+1} = x_k - \hat{x}_k$ .

The size of the input matrix is determined by batch size and back horizon. The input in each block passes through a pack of four fully connected layers and the ReLU (rectified linear unit) activation function, which produces the backward and forward expansion coefficients,  $\theta_k^b$  and  $\theta_k^f$ , respectively. Then these coefficients form a linear combination with vector basis,  $g_k^b$  and  $g_k^f$ , respectively and make a prediction (forecast and backcast). Only one block is enough for making a prediction, but subsequent blocks are added to improve the result.

The operation of first part of  $k^{th}$  block is described below:

$$\begin{aligned}
 h_{k,1} &= FC_{k,1}(x_k), & h_{k,2} &= FC_{k,2}(h_{k,1}), & h_{k,3} &= FC_{k,3}(h_{k,2}), \\
 h_{k,4} &= FC_{k,4}(h_{k,3}) \\
 \theta_k^b &= LINEAR_k^b(h_{k,4}), & \theta_k^f &= LINEAR_k^f(h_{k,4})
 \end{aligned}$$

The second part of the  $k^{th}$  block projects expansion coefficients  $\theta_k^b$  and  $\theta_k^f$  to vector basis and gives  $\hat{x}_k = g_k^b(\theta_k^b)$  and  $\hat{y}_k = g_k^f(\theta_k^f)$ . This operation can generally be described as follows:

$$\hat{x}_k = \sum_{i=1}^{dim(\theta_k^b)} \theta_{k,i}^b v_i^b, \hat{y}_k = \sum_{i=1}^{dim(\theta_k^f)} \theta_{k,i}^f v_i^f$$

where  $v_i^b$  and  $v_i^f$  are backcast and forecast vector basis.

As mentioned above, the N-BEATS model has two configurations: one is generic and the other is interpretable. Generic architecture does not depict time series-specific information. In it,  $g_k^b$  and  $g_k^f$  are linear projections of the output of the preceding layer. To make the model more interpretable, trend and seasonality components are introduced.

Stack-level indexing, which was excluded in the generic model, is considered an interpretable one. Like,  $\hat{y}_{s,k}$  denotes the partial forecast of  $k^{th}$  block in stack  $s$ .

Trend model: A common feature of a trend is that it is almost always a monotonic function, or at least a slowly changing one. To replicate this behavior, constrain  $g_{s,k}^b$  and  $g_{s,k}^f$  are polynomials of small degree  $p$ , a function that slowly varies over the prediction window:

$$\hat{y}_{s,k} = \sum_{i=1}^p \theta_{s,k,i}^f t^i$$

Here,  $t = [0, 1, 2, \dots, H - 2, H - 1]^T / H$  is time vector.

### 3.3 Multi-verse optimizer

MVO is a population-based, nature inspired metaheuristic algorithm. It is inspired by the multiverse theory. Mirjalili et al. (2016) proposed this method for solving numerical optimization problems. The MVO algorithm is based on principles of physics. In the multiverse theory, multiple worlds interact and may even collide. Each universe, according to MVO, would have its own set of physical principles. The three fundamental constituents of multiverse theory are white holes, black holes, and wormholes. The big bang could be regarded as a white hole and possibly the key component in the development of the universe. Black holes attract everything, including light beams, due to their immense gravitational attraction. Wormholes are holes in the cosmos that connect different portions of it. They serve as time and space travel tubes in the multiverse approach. These tunnels allow objects to travel between any two corners of a universe or even from one universe to another in a moment.

Key principles of the MVO optimization process:

- As the inflation rate (fitness value) rises, so does the likelihood of having a white hole, whereas the likelihood of having a black hole decrease.
- Objects are more likely to pass through white holes in universes with a higher inflation rate than through black holes with a lower inflation rate.
- Regardless of the inflation rate, objects in all universes may transfer at random to the best universe via wormholes.

Each solution is a universe, and each variable in that universe is an object. Furthermore, an inflation rate is applied to each solution that is proportional to the fitness function value associated with the solution.

To describe the mathematical model of the white and black hole tunnels as well as the transportation of items across universes, a roulette wheel mechanism is used. The roulette process is used to select one universe from among all possible universes for the white holes. At each iteration, we will rank the universe by fitness value and choose one using roulette.

According to multiverse theory, there are several universes:

$$U = \begin{bmatrix} y_1^1 & y_1^2 & \dots & y_1^m \\ y_2^1 & y_2^2 & \dots & y_2^m \\ \vdots & \vdots & \ddots & \vdots \\ y_n^1 & y_n^2 & \dots & y_n^m \end{bmatrix} \text{ where } m = \text{number of objects, } n = \text{number of universes.}$$

Mathematical model for the selection of universe using roulette wheel selection process:

Mathematical model for the selection of universe using roulette wheel selection process:

$$y_i^j = \begin{cases} y_k^j, & R_1 < NI(U_i) \\ y_i^j, & R_1 \geq NI(U_i) \end{cases}$$

where  $y_k^j = j^{th}$  parameter of  $k^{th}$  universe, selected by roulette selection process

$$y_i^j = j^{th} \text{ parameter of } i^{th} \text{ universe}$$

$$R_1 \in [0, 1], \text{ a random value}$$

$$U_i = i^{th} \text{ universe}$$

$$NI = \text{normalized rate of inflation}$$

Assume that each universe has wormholes to ensure the random interchange of objects via space. They shift objects at random without regard for their inflation rates. Assume that wormhole tunnels are always formed between a universe and the optimal universe (to provide local changes to each universe). Below is the mathematical formulation of this mechanism:

$$y_i^j = \begin{cases} \begin{cases} Y_j + TDR \times ((UB_j - LB_j) \times R_4 + LB_j), R_3 < 0.5 \\ Y_j + TDR \times ((UB_j - LB_j) \times R_4 + LB_j), R_3 < 0.5 \end{cases} \\ y_i^j, R_2 \geq WEP \end{cases}, R_2 < WEP$$

where  $Y_j = j^{th}$  parameter of the fitted universe

$y_i^j = j^{th}$  parameter of  $i^{th}$  universe

UB (Upper bound) = Maximum limit.

LB (Lower bound) = Minimum limit

$\{R_2, R_3, R_4\} \in [0, 1]$  and are random values

WEP (wormhole existence probability) and TDR (travelling distance rate) are coefficients. WEP is used to define the likelihood of the existence of wormholes in other universes. TDR helps in determining the distance rate (variation) at which an object can be transferred by a wormhole around the best universe obtained thus far. Unlike WEP, TDR is enhanced over iterations to allow for more precise exploitation and local search around the best-obtained universe (Fig. 2). They are formulated below:

**PSEUDO CODE: MULTI VERSE OPTIMIZER**

```

Create random universe (U)
Initialize WEP, TDR, and Best universe
SU = Sorted universes
NI = Normalize inflation rate (fitness) of the universes
while the end criterion is NOT satisfied do
    Evaluate the fitness of all universes
    for each universe indexed by i do
        Update WEP and TDR
        Black hole index = i;
        for each object indexed by j do
            R1 = random ([0,1]);
            if R1 < NI(Ui) then
                White hole index = Roulette Wheel Selection(-NI);
                U (Black hole index, j) = SU (White hole index, j);
            end
            R2 = random ([0,1]);
            if R2 < WEP then
                R3 = random ([0,1]);
                R4 = random ([0,1]);
                if R3 < 0.5 then
                    U (i, j) = Best Universe (j) + TDR
                    × ((UB(j) - LB(j)) × R4 + LB (j));
                else
                    U (i, j) = Best Universe (j) - TDR
                    × ((UB(j) - LB(j)) × R4 + LB (j));
                end
            end
        end
    end
    end
    end
Return the Best Universe
    
```

Fig.2 Pseudo Code for multi-verse optimizer (Adapted from: (Mirjalili et al. 2016))



$$WEP = \min + e \times \left( \frac{\max - \min}{E} \right),$$

$e = \text{current iteration}, E = \text{maximum iteration}$

$$TDR = 1 - \frac{e^p}{E^p}, p = \text{exploitation accuracy}$$

### 3.4 Higher moments

Portfolio optimization using higher moments refers to the process of constructing a portfolio by considering not just the expected return and risk (the first and second moments of the return distribution), but also higher-order moments such as skewness and kurtosis.

The inclusion of higher moments in portfolio optimization can lead to more diversified portfolios that recognize the dangers of asymmetric returns and fat-tail risk.

Let  $R_p$  be a random variable representing the portfolio return.  $R = (R_1, R_2, \dots, R_n)$  be the return vector of  $n$  assets where  $R_i$ 's are rate of return of  $i^{th}$  asset.

Further, let  $X = (x_1, x_2, \dots, x_n)$  be the weight vector where,  $x_i$ 's represents the proportion of investment in  $i^{th}$  asset.

Then the first four moments (Kemalbay et al. 2011) (Aksaraylı and Pala 2018) of portfolio return,  $R_p$  can be calculated as follows:

$$Mean = E(R_p) = E[X^T R] = \sum_{i=1}^n x_i \mu_i = X^T \mu = X^T M_1$$

$$Variance = V(R_p) = E[X^T R - E[X^T R]]^2$$

$$= \sum_{i=1}^n \sum_{j=1}^n x_i x_j \sigma_{ij} = X^T V X = X^T M_2 X$$

$$Skewness S(R_p) = E[X^T R - E[X^T R]]^3$$

$$= \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n x_i x_j x_k S_{ijk}$$

$$= X^T S(X \otimes X) = X^T M_3$$

$$Kurtosis K(R_p) = E[X^T R - E[X^T R]]^4$$

$$= \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n x_i x_j x_k x_l k_{ijkl}$$

$$= X^T K(X \otimes X \otimes X) = X^T M_4$$

Here,  $\mu = E[R] = (\mu_1, \mu_2, \dots, \mu_n)$  are the mean return of each asset vector,

$V = E[R - E[R]]^2$  is  $n \times n$  variance–covariance matrix consisting of values like  $\sigma_{ij}$ 's  $\forall (i, j) \in [1, \dots, n]$  and  $\sigma_{ij} = E[(R_i - E[R_i])(R_j - E[R_j])]$ ,

$S = E[R - E[R]]^3$  is  $n \times n^2$  skewness coskewness matrix consisting of values like  $s'_{ijk}$ 's  $\forall (i, j, k) \in [1, \dots, n]$  and  $s_{ijk} = E[(R_i - E[R_i])(R_j - E[R_j])(R_k - E[R_k])]$ ,

$K = E[R - E[R]]^4$  is  $n \times n^3$  kurtosis cokurtosis matrix consisting of values like  $k_{ijkl}$ 's  $\forall (i, j, k, l) \in [1, \dots, n]$  and  $k_{ijkl} = E[(R_i - E[R_i])(R_j - E[R_j])(R_k - E[R_k])(R_l - E[R_l])]$ .  $M_1, M_2, M_3$  and  $M_4$  denote these moments.  $\otimes$  denotes Kronecker product.

### 3.5 Gini-Simpson (GS) entropy

The following expression describes GS entropy as proposed by Aksaraylı and Pala (2018):

$$GSentropy : 1 - \sum_{i=1}^n x_i^2 = 1 - X^T X$$

## 4 Proposed (Variance–Skewness–Kurtosis–Entropy) VSKE optimization model

The optimization model utilized to calculate the best percentage of investments to be made in all assets in the portfolio problem, as well as the assumptions and concepts employed in the problem, are discussed in this section.

### 4.1 Assumptions and notations

It is assumed that the investor is risk averse and will be interested in investing in the efficient frontier portfolio of assets which minimizes his risk at a given level of return. Rate of return of asset is following a probability distribution and investor is interested in maximizing his utility of wealth. Further, no taxes, commission, or transaction fee is involved.

Following notations have been used in the subsequent analysis:

$w_i$ 's: weights assigned to different goals in the objective function,

$x_i$ 's : proportion of investment in each asset,

$k_{ijkl}$  and  $s_{ijk}$  : cokurtosis and coskewness matrix,  $i, j, k = 1, 2, \dots, n$ . ( $n$  is number of assets),

$MinRet$  : minimum value of return aspired by investor,

$LB$  &  $UB$  : lower and upper bounds on the investment proportion of assets.

Following assumptions have been made in the proposed optimization model:

- i. An investor allocates his/her wealth among  $n$  assets offering random rates of return.
- ii. The minimum target return for the investment is set at 5%.
- iii. Available capital should completely be invested.
- iv. The capital invested in each asset is assumed to be bounded between a lower and an upper bound.
- v. Predicted returns are not normally distributed; therefore, skewness and kurtosis are utilized in the analysis.

$$\sum_{i=1}^n x_i \mu_i \geq MinRet$$

$$\sum_{i=1}^n x_i = 1$$

$$LB \leq x_i \leq UB$$

### 4.2 Constraints of the model

4.2.1 *Constraint on return: no less than a certain amount of return allowed i.e.*

$$\sum_{i=1}^n x_i \mu_i \geq MinRet$$

4.2.2 *Capital budget constraint: capital should be completely invested i.e.*

$$\sum_{i=1}^n x_i = 1$$

4.2.3 *Bound constraint: bound on capital invested in each asset, i.e.*

$$LB \leq x_i \leq UB$$

where LB is the minimum proportion of investment and UB is the maximum value of investment in each asset.

### 4.3 Problem formulation

The multi-objective nonlinear optimization problem VSKE is formulated as follows:

$$Min w_1 \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^n x_i x_j x_k x_l k_{ijkl} - w_2 \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n x_i x_j x_k s_{ijk} + w_3 (1 - \sum_{i=1}^n x_i^2) + w_4 \sum_{i=1}^n \sum_{j=1}^n x_i x_j \sigma_{ij}$$

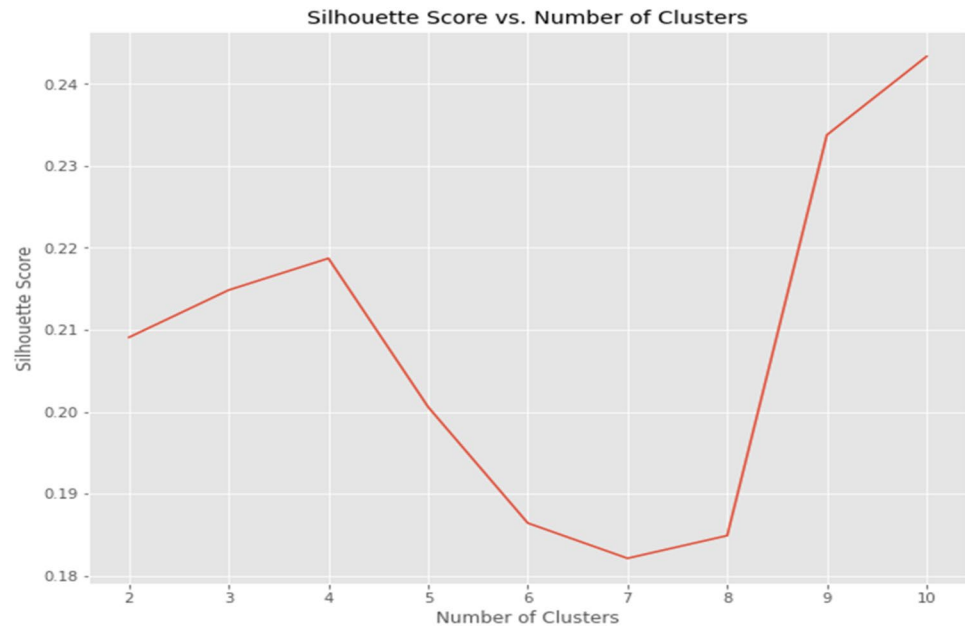
s.t.

The proposed optimization problem for optimal asset allocation is a multiple objective quadratic problem. Obtaining an optimal solution of the considered problem is a difficult task. A metaheuristic method, MVO, is utilized to solve the given problem. The reason for applying MVO is twofold. This method has not been applied previously in the present context. Furthermore, MVO can cater to a wide range of complex situations.

## 5 Data analysis and results

In this study, firstly, 10 years (January 2011–January 2022) of data on fundamental indicators like adjusted opening price, adjusted high price, adjusted low price, adjusted closing price, market capitalization, total returns, earnings per share (EPS), price-to-earnings (P/E) ratio, price-to-book (P/B), book value per share (BVPS), and turnover of S&P BSE 500 index companies have been collected from ProwesIQ’s financial database. These features cover a variety of financial indicators, providing a comprehensive understanding of the data’s characteristics. After normalising the data, we performed principal component analysis (PCA) to reduce the dimensionality of the data, which reduced the dimension of the data to 7 components, where the explained variance of these 7 components is 0.99678. This step is particularly beneficial for mitigating the problem of dimensionality and improving the stability of clustering algorithms. Then Expectation Maximization Technique, was applied to the data to make clusters where 500 companies were grouped according to the similarity and dissimilarity of the 7 components obtained through PCA. A total of 10 clusters were formed, where 70 companies belonged to the first cluster, 32 to the second, and subsequently 88, 30, 80, 33, 46, 35, 52, and 34 to the rest of the clusters. We assessed the quality of our clustering solution using the silhouette score. The silhouette score measures the separation between clusters and their compactness. A score of 0.243 for 10 clusters indicates a reasonable level of separation and compactness among the clusters. The graph for different values of the silhouette score is shown below (Fig. 3).

We have effectively managed the sensitivity to initialization in EM clustering by implementing controlled initialization strategies, employing dimensionality reduction through

**Fig.3** Silhouette Score

PCA and evaluating clustering quality using the silhouette score. This comprehensive methodology provides robustness and reliability to our clustering results, enhancing their practical applicability and interpretation.

The composite performance measure, the Sharpe Ratio, is used to select the best-performing asset from each cluster. To further increase the diversification, companies are selected from each cluster in such a way that all selected assets belong to different sectors such as consumer durables, infrastructure, telecom, finance, real estate, health care, power, oil and gas, industrial, auto, commodities, consumption, and so on (Table 2).

A portfolio of 10 companies' risky assets has been formed, and their five-year (October 2017–October 2022) daily return data was calculated using  $(NewACP - OldACP)/OldACP$ , for which the Adjusted Closing Price (ACP) data has been collected from the Yahoo Finance database. Now the N-Beats method was applied to

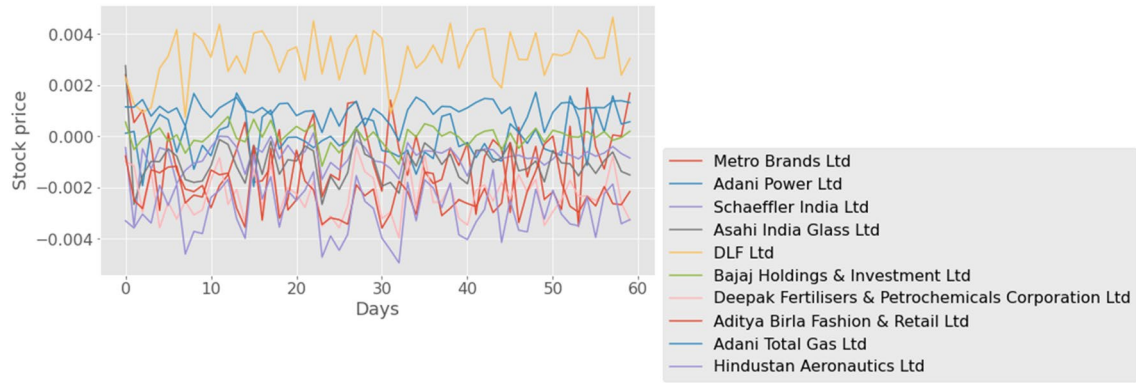
predict their return for the next 60 periods by using 7\*1 as the lookback period and 1 as the horizon, whereas the rest of the hyperparameters were taken as they are in Oreshkin et al. (2019). The mean absolute error and root mean square error values after testing the model are calculated as 0.02569 and 0.03119, respectively, which signifies good performance of the model (Fig. 4).

As asset returns do not necessarily follow normal distribution always, which was assumed in the Markowitz Mean–variance optimization model, the normality of predicted returns is checked by applying the Shapiro–Wilk test. The normal distribution assumption was found to be violated by the estimated return data. Higher order moments like skewness & Kurtosis other than mean and variance, along with Gini entropy were considered for optimization.

The following table comprises the list of descriptive statistics for the 10 stocks selected for analysis.

**Table 2** Sharpe Ratio and Sector of selected companies

S.No	Company name	Sharpe ratio	Sector
1	Metro Brands Limited	2.05	Consumer Durables
2	Adani Power Limited	2.55	Power
3	Schaeffler India Limited	2.24	Industrials
4	Asahi India Glass Limited	1.32	Auto
5	DLF Limited	2.55	Real Estate
6	Bajaj Holdings and Investment Limited	1.01	Finance
7	Deepak Fertilizers & Petrochemicals Corporation Limited	1.77	Commodity
8	Aditya Birla Fashion & Retail Limited	1.18	Consumption
9	Adani Total Gas Limited	2.01	Oil and Gas
10	Hindustan Aeronautics Limited	1.8	Infrastructure



**Fig. 4** Distribution of predicted return values

DLF Limited has the highest mean. Bajaj Holdings and Investment Limited has the lowest variance. Asahi India Glass Limited has the highest skewness, and Hindustan Aeronautics Limited has the lowest kurtosis. The Shapiro–Wilk test findings are also supplied in the last column. From Table 3, it can be seen that five of the companies’ returns are normally distributed and the other five are not (bold values of *p*), which justifies including skewness and kurtosis as one of the objectives to get better results.

Now the MVO is applied to solve multi-objective optimization problems by setting a minimum return of 5% and varying weights assigned to different objectives. Results have been tabulated in Table 4.

Table 4 displays the percentage of investment in ten selected stocks along with the risk and return of the portfolio. To begin with, we allocated equal weights to all objectives, yielding a 5.21 percent return with a 47.82 percent portfolio risk. Then we only analysed three targets at a time by keeping the weight assigned to one at zero, yielding returns of 5.06 percent, 5.14 percent, and 5.04 percent. It shows that when all objectives are taken together, the

portfolio is showing the best performance in terms of risk and return.

### 6 Conclusion

The present work is related to the portfolio allocation problem in Indian context. The findings of the study gave encouraging results. We applied EM clustering to create a well-diversified portfolio and N-BEATS to estimate future returns for further investigations. A multi-objective portfolio optimization problem involving variance skewness, kurtosis, and GS entropy as objectives and mean return with a minimum value of 5% as an additional constraint, was considered. This problem was solved using MVO metaheuristic technique. The maximum return was obtained when all objectives—kurtosis, skewness, entropy, and variance were considered and given equal weight. By neglecting kurtosis as an objective, the worst outcome was attained. The inclusion of higher moments improves the overall quality of the result. With the help of the present study, portfolio managers can fine tune their analysis for determining where & how

**Table 3** Mean, Variance, Skewness, Kurtosis, Shapiro–Wilk Statistics, and *p*-value of selected companies

S.No	Mean	Variance	Skewness	Kurtosis	Shapiro–Wilk Statistics	<i>p</i> -value
1	−0.00087	1.95E-06	0.263619	−0.67314	0.97706	<b>0.31715</b>
2	−4.8E-05	6.41E-07	−0.03927	−0.08722	0.98832	<b>0.83614</b>
3	−0.00079	2.95E-07	−2.12669	8.937137	0.84604	0.000002
4	−0.00105	6.29E-07	1.640721	7.823963	0.88067	0.000028
5	0.003102	9.03E-07	−0.62282	−0.06174	0.9516	0.01853
6	−3E-06	1.46E-07	−0.70644	1.149003	0.96293	<b>0.06551</b>
7	−0.00231	7.16E-07	0.353169	−0.6516	0.97222	<b>0.18717</b>
8	−0.00216	8.34E-07	0.58337	−0.41082	0.95273	0.02096
9	0.001033	1.50E-07	−1.05442	1.250317	0.92052	0.0008
10	−0.00293	9.44E-07	0.032406	−0.82199	0.97767	<b>0.33827</b>

**Table 4** Distribution of weights

Weights	$w_1 = w_2 = w_3 = w_4 = 0.25$	$w_1 = 0, w_2 = w_3 = w_4 = 0.33$	$w_2 = 0, w_1 = w_3 = w_4 = 0.33$	$w_4 = 0, w_2 = w_3 = w_4 = 0.33$
$x_1$	30.48%	1.04%	25.52%	23.63%
$x_2$	23.56%	4.57%	17.20%	17.39%
$x_3$	8.33%	37.75%	5.84%	1%
$x_4$	1.03%	1.61%	1.44%	1%
$x_5$	26.34%	19.19%	27.53%	32.32%
$x_6$	1.09%	17.61%	8.67%	1%
$x_7$	3.23%	1.76%	1.01%	1%
$x_8$	3.45%	2.26%	10.66%	20.56%
$x_9$	1.32%	13.22%	1.13%	1.09%
$x_{10}$	1.16%	1%	1.01%	1.01%
Iterations	100	100	100	100
Best fitness value	1,000,000.086	1,000,099.86	1,000,000.113	1,000,000.046
Portfolio return	5.21%	5.06%	5.14%	5.04%
Portfolio risk	47.28%	52.43%	42.93%	43.29%
Sum of weights	100%	100%	100%	100%

much proportion of their wealth should be invested and they can improve their decisions. It provides a perfect quantitative approach. A future study might concentrate on incorporating more constraints by relaxing some of the assumptions of the model into the suggested portfolio optimization model. It would also be fascinating to experiment with various metaheuristics and variants of MVO. The variability of the results of the forecasting technique is a significant limitation of the work. The results of the study rely primarily on the availability and accuracy of past data as well as parameter adjustment of applied techniques.

## 7 Discussion

The comprehensive study of a ten-year dataset containing crucial financial indicators is the basis of our research. PCA has been useful in enhancing the stability and interpretability of clustering methods by reducing data dimensionality. Our use of the EM approach yielded 10 well-defined clusters, laying the groundwork for asset selection and diversification techniques. The Sharpe Ratio is used to find the best-performing assets within each cluster, which improves the portfolio's risk-return profile. The N-BEATS approach for return prediction, which incorporates deep learning, has exhibited solid performance, providing useful insights for investment decisions. Furthermore, including higher-order moments into the multi-objective optimization model acknowledges the non-normal distribution of asset returns, resulting in a more thorough risk assessment. This study

provides portfolio managers and investors with a realistic and methodical strategy to managing the difficulties of current financial markets.

**Funding** This research received no grant from any funding agency.

### Declarations

**Conflict of interest** There is no conflict of interest.

**Human or animal rights** No human or animal participation was involved in this research.

**Informed consent** All authors have approved the manuscript and agree with its submission to the journal for publication.

## References

- Abdelaziz FB, Chibane M (2023) Portfolio optimization in the presence of tail correlation. *Econ Model* 122(May):106235. <https://doi.org/10.1016/j.econmod.2023.106235>
- Abolmakarem S, Abdi F, Khalili-Damghani K, Didekhani H (2023) Predictive Multi-period multi-objective portfolio optimization based on higher order moments: deep learning approach. *Comput Ind Eng*. <https://doi.org/10.1016/j.cie.2023.109450>
- Abualigah L (2020) Multi-verse optimizer algorithm: a comprehensive survey of its results, variants, and applications. *Neural Comput Appl* 32(16):12381–12401. <https://doi.org/10.1007/s00521-020-04839-1>
- Aithal PK, Geetha M, Dinesh U, Savitha B, Menon P (2023) Real-time portfolio management system utilizing machine learning techniques. *IEEE Access* 11:32595–32608. <https://doi.org/10.1109/ACCESS.2023.3263260>

- Aksaraylı M, Pala O (2018) A polynomial goal programming model for portfolio optimization based on entropy and higher moments. *Expert Syst Appl* 94(March):185–192. <https://doi.org/10.1016/j.eswa.2017.10.056>
- Ashfaq S, Ayub U, Mujtaba G, Raza N, Gulzar S (2021) Gainers and losers with higher order portfolio risk optimization. *Phys A Stat Mech Appl*. <https://doi.org/10.1016/j.physa.2020.125416>
- Ashrafzadeh M, Taheri HM, Gharehgozlou M, Zolfani SH (2023) Clustering-based return prediction model for stock pre-selection in portfolio optimization using PSO-CNN+MVF. *J King Saud Univ Comput Inf Sci* 35(9):101737. <https://doi.org/10.1016/j.jksuci.2023.101737>
- Barkhagen M, García S, Gondzio J, Kalcsics J, Kroeske J, Sabanis S, Staal A (2023) Optimising portfolio diversification and dimensionality. *J Glob Optim* 85(1):185–234. <https://doi.org/10.1007/s10898-022-01202-7>
- Batra L, Taneja HC (2022) Portfolio optimization based on generalized information theoretic measures. *Commun Stat Theory Methods* 51(18):6367–6384. <https://doi.org/10.1080/03610926.2020.1861294>
- Benmessahel I, Xie K, Chellal M (2020) A new competitive multi-verse optimization technique for solving single-objective and multiobjective problems. *Eng Rep*. <https://doi.org/10.1002/eng2.12124>
- Chang TJ, Yang SC, Chang KJ (2009) Portfolio optimization problems in different risk measures using genetic algorithm. *Expert Syst Appl* 36(7):10529–10537. <https://doi.org/10.1016/j.eswa.2009.02.062>
- Chen B, Zhong J, Chen Y (2020) A hybrid approach for portfolio selection with higher-order moments: empirical evidence from shanghai stock exchange. *Expert Syst Appl*. <https://doi.org/10.1016/j.eswa.2019.113104>
- Chen C, Zhou YS (2018) Robust multiobjective portfolio with higher moments. *Expert Syst Appl* 100(June):165–181. <https://doi.org/10.1016/j.eswa.2018.02.004>
- Chen LH, Huang L (2009) Portfolio optimization of equity mutual funds with fuzzy return rates and risks. *Expert Syst Appl* 36(2 PART 2):3720–27. <https://doi.org/10.1016/j.eswa.2008.02.027>
- Cheong D, Kim YM, Byun HW, Kyong Joo Oh, Kim TY (2017) Using genetic algorithm to support clustering-based portfolio optimization by investor information. *Appl Soft Comput* 61(December):593–602. <https://doi.org/10.1016/j.asoc.2017.08.042>
- Chopra VK, Hensel CR, Turner AL (1993) Massaging mean-variance inputs: returns from alternative global investment strategies in the 1980s. *Source Manag Sci* 39. <https://www.jstor.org/stable/2632423>
- Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. *J R Stat Soc Ser B (Methodol)* 39
- Do CB, Batzoglou S (2008) What is the expectation maximization algorithm? *Nat Biotechnol* 26(8):897–899. <https://doi.org/10.1038/nbt1406>
- Du J (2022) Mean–variance portfolio optimization with deep learning based-forecasts for cointegrated stocks. *Expert Syst Appl*. <https://doi.org/10.1016/j.eswa.2022.117005>
- Erwin K, Engelbrecht A (2023) Meta-heuristics for portfolio optimization. *Soft Comput*. <https://doi.org/10.1007/s00500-023-08177-x>
- Goldfarb D, Iyengar G (2003) Robust portfolio selection problems. *Math Oper Res* 28(1):1–38. <https://doi.org/10.1287/moor.28.1.1.14260>
- Gonçalves G, Wanke P, Tan Y (2022) A higher order portfolio optimization model incorporating information entropy. *Intell Syst Appl*. <https://doi.org/10.1016/j.iswa.2022.200101>
- Gu S, Kelly B, Xiu D (2020) Empirical asset pricing via machine learning. *Rev Financ Stud* 33(5):2223–2273. <https://doi.org/10.1093/rfs/hhaa009>
- Gupta P (2022) Portfolio optimization using elliptic entropy and semi-entropy of coherent fuzzy numbers. *Inf Sci* 614(October):240–262. <https://doi.org/10.1016/j.ins.2022.09.032>
- Gupta P, Mehlatat MK, Yadav S, Kumar A (2019) A polynomial goal programming approach for intuitionistic fuzzy portfolio optimization using entropy and higher moments. *Appl Soft Comput J*. <https://doi.org/10.1016/j.asoc.2019.105781>
- Harry Markowitz (1952) American Finance Association
- Ji R, Lejeune MA, Prasad SY (2017) Properties, formulations, and algorithms for portfolio optimization using mean-gini criteria. *Ann Oper Res* 248(1–2):305–343. <https://doi.org/10.1007/s10479-016-2230-4>
- Jung YG, Kang MS, Heo J (2014) Clustering performance comparison using k-means and expectation maximization algorithms. *Biotechnol Biotechnol Equip* 28:S44–48. <https://doi.org/10.1080/13102818.2014.949045>
- Kaja S, Shakshuki E, Yasar A (2021) N-beats for backup routing optimization in cloud acknowledgement scheme for a node network. In: *Procedia Computer Science*, vol 191. Elsevier B.V., pp 272–81. <https://doi.org/10.1016/j.procs.2021.07.034>
- Kemalbay G, Murat Özkut C, Franko C (2011) Portfolio selection with higher moments: A polynomial goal programming approach to ISE-30 index. *Ekonometri ve İstatistik Dergisi* (13):41
- Khan KI, Waqar SM, Naqvi A, Ghafoor MM, Akash RSI (2020) Sustainable portfolio optimization with higher-order moments of risk. *Sustainability (Switzerland)*. <https://doi.org/10.3390/su12052006>
- Li X, Li Bo, Jin T, Zheng P (2023) Uncertain random portfolio optimization with non-dominated sorting genetic algorithm-II and optimal solution criterion. *Artif Intell Rev*. <https://doi.org/10.1007/s10462-022-10388-x>
- Long NC, Wisitpongphan N, Meesad P, Unger H (2014) Clustering stock data for multi-objective portfolio optimization. *Int J Comput Intell Appl*. <https://doi.org/10.1142/S1469026814500114>
- Ma X, Yan B, Wang H, Liao H (2023) Decision-level machinery fault prognosis using N-BEATS-based degradation feature prediction and reconstruction. *Mech Syst Signal Process* 198(September):110435. <https://doi.org/10.1016/j.ymssp.2023.110435>
- Malek, Marc H, Dobrovolsky S, Clare A, Motson N (2009) The Journal of Alternative Investments Volatility Exposure of CTA Programs and Other Hedge Fund Strategies Locking in the Profits or Putting It All on Black? An Empirical Investigation into the Risk-Taking Behavior of Hedge Fund Managers.” [www.ijournals.com](http://www.ijournals.com).
- Moghadaszadeh M, Shokrzadeh H (2018) An overview of expectation maximization and K-means family clustering algorithms in data mining applications. In: Oct. 17–19, 2017 Dubai (UAE). Dignified Researchers Publication (DiRPUB). <https://doi.org/10.15242/DirPUB.DIR1017002>
- Menvouta EJ, Serneels S, Verdonck T (2023) Portfolio optimization using cellwise robust association measures and clustering methods with application to highly volatile markets. *J Financ Data Sci* 9(November):100097. <https://doi.org/10.1016/j.jfds.2023.100097>
- Milhomem DA, Dantas MJP (2020) Analysis of new approaches used in portfolio optimization: a systematic literature review. *Production* 30:1–16. <https://doi.org/10.1590/0103-6513.20190144>
- Mirjalili S, Jangir P, Mirjalili SZ, Saremi S, Trivedi IN (2017) Optimization of problems with multiple objectives using the multi-verse optimization algorithm. *Knowl-Based Syst* 134(October):50–71. <https://doi.org/10.1016/j.knsys.2017.07.018>
- Mirjalili S, Mirjalili SM, Hatamlou A (2016) Multi-verse optimizer: a nature-inspired algorithm for global optimization. *Neural Comput Appl* 27(2):495–513. <https://doi.org/10.1007/s00521-015-1870-7>

- Mirlohi M, Seyyed, Souri A, Bahman E (2021) Higher moments portfolio optimization with unequal weights based on generalized capital asset pricing model with independent and identically asymmetric power distribution. *Adv Math Financ Appl* 6(2):263–83. <https://doi.org/10.22034/amfa.2020.1909590.1484>
- Nabizadeh A, Behzad A (2018) Higher moments portfolio optimization considering entropy based on polynomial idealistic programming. *Financ Res J* 20(2):191–208. <https://doi.org/10.22059/frj.2018.255731.1006645>
- Nanda SR, Mahanty B, Tiwari MK (2010) Clustering Indian stock market data for portfolio management. *Expert Syst Appl* 37(12):8793–8798. <https://doi.org/10.1016/j.eswa.2010.06.026>
- Naqvi B, Mirza N, Naqvi WA, Rizvi SKA (2017) Portfolio optimisation with higher moments of risk at the Pakistan stock exchange. *Economic Research-Ekonomika Istrazivanja* 30(1):1594–1610. <https://doi.org/10.1080/1331677X.2017.1340182>
- Navarro MM, Young MN, Prasetyo YT, Taylar JV (2023) Stock market optimization amidst the COVID-19 Pandemic: technical analysis, K-means algorithm, and mean-variance model (TAKMV) approach. *Heliyon* 9(7):e17577. <https://doi.org/10.1016/j.heliyon.2023.e17577>
- Ng K-H, Khor KC (2014) Rapid profiling of plantation stocks in BursaMalaysia with expectation maximization clustering network intrusion detection view project. <https://www.researchgate.net/publication/328019059>
- Nguyen TT (2016) Portfolio selection under higher moments using fuzzy multi-objective linear programming. *J Intell Fuzzy Syst* 30(4):2139–2156. <https://doi.org/10.3233/IFS-151927>
- Oreshkin BN, Carpov D, Chapados N, Bengio Y (2019) N-BEATS: neural basis expansion analysis for interpretable time series forecasting. <http://arxiv.org/abs/1905.10437>
- Oreshkin BN, Dudek G, Peřka P, Turkina E (2021) N-BEATS neural network for mid-term electricity load forecasting. *Appl Energy*. <https://doi.org/10.1016/j.apenergy.2021.116918>
- Pahade JK, Jha M (2021) Credibilistic variance and skewness of trapezoidal fuzzy variable and mean–variance–skewness model for portfolio selection. *Results Appl Math*. <https://doi.org/10.1016/j.rinam.2021.100159>
- Pourrafiie M, Nafei AH, Banihashemi S, Pourmohammad Azizi S (2020) Comparing entropies in portfolio diversification with fuzzy value at risk and higher-order moment. *Fuzzy Inf Eng* 12(1):123–138. <https://doi.org/10.1080/16168658.2020.1811481>
- Rezani MA, Hertono GF, Handari BD (2020) Implementation of Iterative k -Means+ and ant colony optimization (ACO) in portfolio optimization problem. In: AIP conference proceedings, vol. 2242. American Institute of Physics Inc. <https://doi.org/10.1063/5.0008149>
- Saranya K, Krishna Prasanna P (2014) Portfolio selection and optimization with higher moments: evidence from the Indian stock market. *Asia-Pac Finan Mark* 21(2):133–149. <https://doi.org/10.1007/s10690-014-9180-0>
- Sass J, Thös A-K (2021) Risk reduction and portfolio optimization using clustering methods. *Econom Stat*. <https://doi.org/10.1016/j.ecosta.2021.11.010>
- Sbrana A, Lima De Castro PA (2023) N-BEATS perceiver: a novel approach for robust cryptocurrency portfolio forecasting. <https://doi.org/10.21203/rs.3.rs-2618277/v1>
- Sehgal R, Jagadesh P (2023) Data-driven robust portfolio optimization with semi mean absolute deviation via support vector clustering. *Expert Syst Appl* 224(August):120000. <https://doi.org/10.1016/j.eswa.2023.120000>
- Siew LW, Jaaman SH, Hoe LW (2021) Higher moment of portfolio optimization with polynomial goal programming approach. *J Phys Conf Ser*. <https://doi.org/10.1088/1742-6596/1988/1/012001>
- Sihem M, Slaheddine H (2014) The impact of higher order moments on market risk assessment. *Procedia Econ Financ* 13:143–153. [https://doi.org/10.1016/S2212-5671\(14\)00436-5](https://doi.org/10.1016/S2212-5671(14)00436-5)
- Singhal V, Mathew J, Agarwal M (2022) Fusion of wavelet decomposition and N-BEATS for improved stock market forecasting. <https://doi.org/10.21203/rs.3.rs-2003731/v1>
- Song Y, Liu Y, Chen H, Deng W (2023) A multi-strategy adaptive particle swarm optimization algorithm for solving optimization problem. *Electronics (Switzerland)*. <https://doi.org/10.3390/electronics12030491>
- Wang J, Liu G (2019) A novel firefly algorithm with self-adaptive step strategy. *Int J Innov Comput Appl* 10(1):18. <https://doi.org/10.1504/IJICA.2019.100525>
- Wang Y, Aste T (2023) Dynamic portfolio optimization with inverse covariance clustering. *Expert Syst Appl* 213(March):118739. <https://doi.org/10.1016/j.eswa.2022.118739>
- Wu D, Wang X, Shaocong Wu (2022) Construction of stock portfolios based on K-means clustering of continuous trend features. *Knowl-Based Syst* 252(September):109358. <https://doi.org/10.1016/j.knosys.2022.109358>
- Wu W, Chen J, Yang ZB, Tindall ML (2021) A cross-sectional machine learning approach for hedge fund return prediction and selection. *Manag Sci* 67(7):4577–4601. <https://doi.org/10.1287/mnsc.2020.3696>
- Zhou R, Cai R, Tong G (2013) Applications of entropy in finance: a review. *Entropy*. MDPI AG. <https://doi.org/10.3390/e15114909>
- Zhu H, Wang Yi, Wang K, Chen Y (2011) Particle Swarm Optimization (PSO) for the constrained portfolio optimization problem. *Expert Syst Appl* 38(8):10161–10169. <https://doi.org/10.1016/j.eswa.2011.02.075>
- Zhu Z, Thavaneswaran A, Paseka A, Frank J, Thulasiram R (2020) Portfolio optimization using a novel data-driven EWMA covariance model with big data. In: Proceedings—2020 IEEE 44th annual computers, software, and applications conference, COMPSAC 2020, 1308–13. Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/COMPSAC48688.2020.00-75>

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.