

Predicting Optimal Portfolio by Algorithm Analysis Systems

Mehdi Darvishan

Ph.D. Candidate, Department of Accounting, Shahrood Branch, Islamic Azad University, Shahrood, Iran. (Email: sepehr11384@gmail.com)

Mohammadreza Abdoli *

*Corresponding Author, Associate Professor, Department of Accounting, Shahrood Branch, Islamic Azad University, Shahrood, Iran. (Email: Mrab830@yahoo.com)

Mohammad Mehdi Hosseini

Assistant Professor, Department of Electrical Engineering, Shahrood Branch, Islamic Azad University, Shahrood, Iran. (Email: hosseini_mm@yahoo.com)

Esmail Alibeiki

Assistant Professor, Department of Electrical Engineering, Ali Abad Katoul Branch, Islamic Azad University, Ali Abad Katoul, Iran. (Email: esmail_alibeiki@aliabadiau.ac.ir)

Abstract

Choosing the proper investment mechanism is one of the main tasks of any investor that requires careful analysis and research on all available information. Since no investor exactly knows whether his or her expectations for a particular stock return will be met, they need to build their strategy in such a way as to eliminate as much damage as possible in the event of an adverse outcome. This study aims to predict the optimal portfolio using Algorithm Analysis Systems. In this regard, 98 firms listed on the Tehran Stock Exchange were examined in 2015-2019. Then, random portfolios were selected to test the research hypotheses by separating value stocks and growth stocks. For analysis, two algorithms of Support Vector Machines and an Adaptive Neuro-Fuzzy Inference System were used to select the most desirable portfolio. According to the support vector machine algorithm analysis, the results confirm the difference between the Sortino and Marquitz portfolios. To build their portfolios, decision-makers often rely on growth stocks which can boost their expected returns. Therefore, recognizing the analytical nature of portfolio formation in specialized areas can help improve investment analysis and pave the way for higher returns.

Keywords: Predicting Optimal Portfolio, Growth and value stocks, Optimization Algorithms.

Introduction

Predicting and examining the price behavior of securities to form an investment portfolio is a category that financial scientists and investors are always looking for. The main reason for investing in the stock market is to make a profit, which requires accurate information about the stock market, stock changes, and its future trend; therefore, the investor needs powerful and reliable tools to predict risk and return (Schellinger, 2020). Portfolio optimization as a strategic tool is considered a basis for forecasting that, if appropriate to market conditions and appropriate analytical methods, can help increase the accuracy of investors' forecasts (Kandahari et al., 2017). The issue of portfolio optimization has come a long way since Markowitz introduced mean-variance optimization. The most important achievement of the Markowitz model was the introduction of variance as a risk indicator and, in fact, the introduction of a quantitative criterion for it. Post-Markowitz research has shown that the use of variance as a risk factor has shortcomings that cause

the traditional tools in forecasting to be criticized (Raei et al., 2020). Studies such as Herberger and Reinle (2020); Senarathne (2019); Gupta et al. (2014), and Ince and Trafalias (2007) state that using traditional forecasting tools and methods has a high error rate and they have poorer performance compared to newer methods and nonlinear models. In this study, one of the artificial intelligence methods called support vector machine, along with one of the most widely used algorithms in this field, the adaptive neural fuzzy inference system, are examined to predict the most desirable stock portfolio because the goal is to choose the appropriate forecasting method in selecting the desired portfolio.

Recent advances in artificial intelligence have proposed new methods for prediction that are more accurate than traditional methods (Fattahi Nafchi et al., 2019). The most common of these methods is the artificial neural network algorithm. However, neural networks have shortcomings, such as the need for high control parameters, difficulty in achieving stable results, and so on. Due to such weaknesses, better methods have been designed to improve the neural network model. Support vector machine (SVM) is a supervised learning method in classification and regression. This method is one of the relatively new methods that, in recent years, has shown good performance compared to older classification methods, including neural networks (Faghihinejad & Minaei, 2018). The root of support vector is in statistical learning theory and many applications in regression and classification; they have clustering and approximation functions. This approach initially included only two-class classification but was later extended to multi-class classification using various blending techniques (Wang et al., 2017).

On the other hand, the adaptive neural fuzzy inference system is considered a neural-fuzzy analysis approach; the most significant advantage is that it can take advantage of the neural network learning capability and avoid the time-consuming regulation of an inference engine in a conventional fuzzy logic system. Practically, there are no restrictions on the node functions of adaptive networks; unless they need to be derivative. The only structural limitation of the shape of networks is the type of feed they are. Against this minor limitation, the use of adaptive networks has become increasingly widespread in the field of applications (Petkovic et al., 2014). The system works so that the neurons' output is typically calculated to the last layer as they move forward at each training session. And then, the result parameters are calculated by the least squares error method (Vakilifard et al., 2014). Therefore, despite the analytical interpretations of the two methods, this study seeks to choose the most desirable portfolio from the two portfolios of Sortino

and Markowitz to increase investment effectiveness and select the most effective hybrid analysis systems algorithm based on characteristics based on growth and value metrics. In fact, this research helps companies' managers improve their level of analytical knowledge and form an effective investment portfolio to achieve higher returns. Therefore, this study aims to predict the effectiveness of the difference between Sortino and Markowitz's portfolios based on the algorithm of hybrid analysis systems.

Literature Review

Portfolio Selection

The framework of Van Neumann-Morgan Stein's theory of utility function is always one of the most reliable theories for deciding whether to form a portfolio. Suppose the investor faces two portfolios, X and Y, as in this study (Sortino and Markowitz portfolios). In that case, selecting the appropriate portfolio requires using the expected utility function (Thompson, 2020). For example, if position Y is assumed to have a 50% probability of a 5% return and another 50 percent chance of 15 percent risk, a person without knowledge of the utility function will likely become more conservative and make decisions based on risk aversion. For this person, the utility function is as follows:

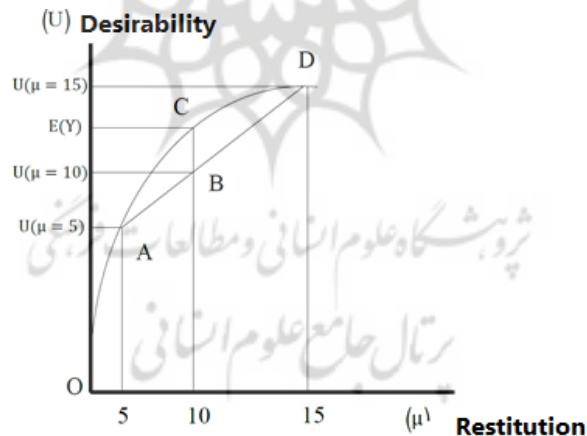


Figure 1. Portfolio selection utility function (Sources: Wang, 1999)

The utility function of the whole investor is written as follows:

$$U = F(\mu)$$

According to this function, position Y with a yield of 5% is equal to $U = U(Y)$, which is equal to the starting point of OA. In the case of a random position Y, the expected utility must be calculated.

$$E(u) = \sum W_i U(Y_i = a)$$

$$Eu(Y) = W_1 \times U(Y = 0/05) + W_2 \times U(Y = 0/15)$$

Accordingly, the expected utility in calculating the return on the Y portfolio is the midpoint of the AD line segment. By matching the return of portfolio X, we can estimate the expected desirability in portfolio selection (Falahpour et al., 2014). However, it is noteworthy that the investment is evaluated based on the utility function, which is a function of the risk and return of portfolios. Moreover, if one or more portfolios are random and their returns are not definitively determinable, the expected utility of the portfolios will be the basis for the decision:

$$E(U) = E[F(\mu, \sigma)]$$

The expected utility has a positive relationship with return and a negative relationship with risk. If a portfolio such as A is considered, its risk position and return compared to other portfolios are shown in the chart above. All points in space B indicate higher risk and lower returns, so it has a lower desirability level than point A. The points in space C represent projects with less risk and efficiency; therefore, their level of desirability is different from point A, which is based on an ambiguous situation. The points located in space D have the lowest risk and the lowest return in the desired portfolio, so they are a reference to point A (Firoozdehghan et al., 2019).

Growth and value stocks

The stock offering is considered one of the ways to finance companies in the capital market and stock exchanges of countries with specific frameworks, rules, and regulations. Potential investors consider important financial factors when trading stocks, including earnings per share, dividends per share, price and price growth, and stock returns. Meanwhile, the stocks of companies with higher returns than the market average are expected to trade at a higher price, but this may only be the case in some circumstances. Because the stock of a Firm that grows at a high average yield and trades at a high price may not be stable. On the other hand, a stable profit, even if it is average or below the market average, gives the investor confidence that in the future, at least, the price will not decrease. Therefore, the investor is faced with trading two types of stocks, called value and growth stocks, that need their own strategies to select them (Salehi & Salehi, 2016). Haugen (2001), in his definition of growth

stocks, stated that growth stocks are stocks whose price is higher than the average compared to cash flow, profit, dividend, and current book value. Therefore, growth stocks will have lower returns in the long run than value stocks, which have a lower-than-average price according to the mentioned criteria. In a comprehensive definition, growth stocks are the stocks of companies that will have a high positive profit. This profit is higher than the average rate of return commensurate with their risk because their stocks have a lower intrinsic value. Value stocks are stocks that, for some reason, other than the possibility of potential revenue growth, are priced at lower intrinsic value (Asadi & Eslami Bidgoli, 2014). Proponents of growth stocks believe that the main reason for investing in this type of stock is to invest in the future growth of the Firm's profits. Therefore, they put stock buying on the agenda, the profit of which is expected to grow almost rapidly. Growth investors are looking for stocks of companies that have grown faster than average throughout history; therefore, they have high growth potential. Growth is measured by factors such as the increase in profits or sales of the Firm. Growth stock managers tend to accumulate profits and refuse to pay dividends because they want to reinvest any available cash in the firm. Thus, growth investors mainly derive their return on investment from rising stock prices (Kistner, 1995).

In contrast, value-seeking investors do not rely on the Firm's future growth forecasts. They try to identify stocks trading at prices below value, which is reflected in the fundamental factors of the Firm. Analysts commonly recognize value stocks as low (P / E) to low (P/B) ratios. However, a clear definition of growth and value stocks must be defined.

Research Methodology

The present study is applied in terms of purpose and terms of data collection method is a quasi-experimental post-event research in the field of positive accounting research which is done using the proposed hybrid models. The statistical population studied in this study is companies listed on the Tehran Stock Exchange, which was studied from 2015 to 2019. Therefore, in order to select the research sample, the systematic screening method has been used, and companies that have the following set of conditions are selected as the sample:

1. Companies whose date of admission to the Stock Exchange Organization is before 2015 will be on the list of listed companies until the end of 2019.
2. Their fiscal year should end at the end of March, as it helps the research to compare the accounting data of different companies and avoid calendar effects in comparison between companies.

3. They have kept their activities and fiscal year the same during the mentioned years.
4. Not be part of investment and financial intermediation companies (investment companies were not included in the statistical community due to the difference in the nature of their activities with other companies).
5. The interruption of transactions in these companies during the mentioned period should be at most six months.

By applying the above restrictions, 98 firms were selected as the statistical sample of the study. The data required by the selected companies were extracted by referring to the financial statements and explanatory notes and the companies listed on the Tehran Stock Exchange available in the Codal system, the stock exchange website.

This study aims to test the various structures of a meta-heuristic algorithm to predict the effectiveness of selecting the most effective portfolio by considering the data of 98 stock exchange companies. The most important question is the efficiency of the above systems in predicting the effectiveness of choosing the right portfolio of companies from the two portfolios of Sortino and Markowitz. The second question examines which portfolio will yield better results in the companies surveyed in each industry. Moreover, it is possible to arrive at a single model to predict the effectiveness of selecting the desired portfolio. The third question also evaluates the possibility of extracting influential variables through research analysis.

Markowitz portfolio

To select the optimal portfolio, Markowitz proposed the "mean-variance" model, in which mean is a measure of return and variance is a measure of risk, and standard deviation and variance as a measure of risk assuming an average rate of return distribution (Dalagonol et al., 2009, p. 729). This model was introduced in 1952, and Markowitz stated that the investor, in addition to maximizing returns (as much as possible), also wants to ensure returns. Nevertheless, to justify his argument, he says, "If investors only wanted to maximize the expected return, they would only invest in the type of asset that has the highest expected return." at a glance, it can be seen that investors own "a collection of a portfolio of portfolio securities." In justifying this behavior, investors pay attention to the two phenomena of risk and return. Thus, investors seeking to maximize expected returns and minimize uncertainty (means risk) have these two opposing goals that must be balanced against each other. One of the exciting results of these two conflicting goals is that the

investor must diversify by buying several types of securities (Baptista, 2012). Markowitz's approach to investment can be better explained by a more precise definition of the concept of primary value and final value (Talebna & Fathi, 2010). In this model, a firm's stock is presented as a risky asset because the instability (randomness) of the total rate of return (weekly, monthly, and annual) is the cause. Because these rates change over time, a probability distribution function can be formed for them, and the criteria required by the Markowitz model, such as mean, standard deviation, covariance, etc., can be obtained from them. The Markowitz model is based on the following assumptions:

Investors are risk averse and have the expected incremental utility, and the ultimate utility curve of their wealth decreases. Investors select their portfolios based on the expected average return variance. Therefore, their indifference curves are a function of the expected rate of return and variance. Every investment option is infinitely divisible. Investors at a certain level of risk prefer higher returns and vice versa. Based on this, investors pay attention to two factors in their choice:

- ❖ A) High expected returns, which is a favorable factor.
- ❖ B) Uncertainty returns which is an undesirable factor.

To obtain a portfolio selection based on the Markowitz method, which is the minimum variance for a particular level of return. We have the following linear programming model:

$$\text{Min } z = \delta_p^2 \quad (1)$$

$$\text{St. } \bar{R}_p = \sum_{j=1}^n W_j \bar{R}_j \quad (2)$$

$$\sum_{j=1}^n W_j = 1 \quad (3)$$

$$W_j \geq 0$$

Which is in:

W_j weight related to i share in the stock portfolio; \bar{R}_p expected portfolio return; \bar{R}_j stock returns i ; δ_p^2 Stock portfolio return variance. The variance of stock portfolio returns is calculated according to equation (4) below:

$$\delta_p^2 = \sum_{i=1}^n \sum_{j=1}^n w_i w_j \text{cov}(\bar{r}_i, \bar{r}_j) \quad (4)$$

Sortino Portfolio

The Sortino index is obtained if an undesirable risk module is used in the

performance appraisal instead of SD (SD). In fact, if X is the portfolio return variable and $f(x)$ is a function of the probability challenge of this variable, μ is the mean, and r is the acceptable rate of return or MAR, the Sortino index knowledge can be represented as follows (Mamoghli & Daboussi, 2008).

$$SOR = \frac{(\mu-r)}{\sigma} \quad (5)$$

Which is in:

σ The half-standard deviation of returns is below the target rate and is:

$$\sigma^2 = \int_{-\infty}^r (r-x)^2 f(x) dx \quad (6)$$

In this research, equation (7) has been used to calculate the undesirable risk:

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (\text{Max}\{R_f - R_{ij}\})^2} \quad (7)$$

In the Sortino criterion, the mean return is adjusted for undesirable risk. This risk focuses on returns that have yet to increase returns without risk. The IMNEX index is designed to assess the performance of stock exchange companies and the fact that there is an indicator that indicates the performance of stocks in general. This index is designed by weight, and various adjustments are foreseen. The formula for calculating this index in equation (8):

$$IMNEX_t = \frac{\sum_{i=1}^n NAV_{it} \times NU_{it}}{C_t} \quad (8)$$

n : Number of Shares, NAV_{it} : The net value of each stock of company i at time t ; NU_{it} Number of the company i at time t ; C_t The base number is at time t , which is considered equal to 1000.

Different statistical methods were used to test the research hypotheses depending on the nature of the hypotheses. In the test of the first hypothesis, to better separate the two portfolios of Markowitz (X) and Sortino (Y) for selection and investment, stocks were classified into five categories: value stocks to growth stocks. (K1) to (K5) were used to classify stocks so that value stocks (K1) and growth stocks (K5) were considered. The method of stock separation was also calculated based on the defined ratios, Which are presented in Table (1), so those stocks with lower ratios are value stocks with higher returns than growth stocks.

Then, to distinguish between the two portfolios in this study, the hyperbola algorithm based on the criterion of random and equal weights was used separately. Equal weights According to the model of Fama and French (1992), after the stocks of the surveyed companies in each quintile, were selected, all

stocks with equal weight quintiles were placed in the portfolios of this study, namely Markowitz (X) and Sortino (Y). Thus, a growth portfolio and a value portfolio were formed each year. Also, random weights were first formed by a vector of random values based on the number of shares of the surveyed companies in each quintile. The values were standardized and, in other words, were examined, of the sum of all components became a vector equal to one that portfolios with random weights.

Due to the increase in the number of portfolios formed during a year and the research period, not all portfolios were necessarily formed with equal weight. In other words, ten growth and ten value portfolios with random weights were created each year, and over five years, 50 portfolios were created randomly for comparison.

This method has two advantages: First, it increases the number of samples; second, it shows whether the selection of a value or growth portfolio differs from the daily performance of stocks. It should be noted that since the stock weight in the portfolios is random, the return numbers and standard deviations may vary slightly with each program run; this difference is not so crucial in the analysis of the results. Based on this, the components of separating value and growth stocks should be determined to select the appropriate portfolio.

Separation of value stocks and growth stocks

This research is an adaptation and following various research such as Doering et al. (2019), Hubner and Lejeune (2017), and Carazo et al. (2010) of 7 book value criteria at share market price (B / P); Earnings to share price ratio (E / P); Cash flows to price ratio (CF / P); Total debt to total asset Ratio (TD / TA); Ratio of fixed assets to total assets (FA / TA); The ratio of sales to total assets (S / TA) and the ratio of net earning to total assets (NE / TA) are used. It is noteworthy that for the calculation of each ratio, the balance sheet data of year $t-1$ and the first price data after July 31 of year t are used. Accordingly, companies are divided into five categories for each ratio. The companies with the highest value ratios will be the value companies, and those with the lowest will be the growth companies.

Table 1. The basis for Measuring Criteria Growth and Value Stocks

Criteria's name	Abbreviation	Base	How to measure	Source
The ratio of book value to share market price	(B/P)	A1	Divide the book value of each share at the end of the fiscal year by the price market value.	Hubner and Lejeume (2017)
The ratio of profit to share price	(E/P)	A2	Divide earnings per share at the end of each year by the price market	Doering et al. (2019)
The ratio of cash flows to stock prices	(CF/P)	A3	Divide the operating cash flow per year by the price	Hubner and Lejeume (2017)
Total debt to total assets ratio	(TD/TA)	A4	Divide total debates by total assets	Carazo et al. (2010)
The ratio of fixed assets to total assets	(FA/TA)	A5	Divide total fixed assets by the sum of total assets	Carazo et al. (2010)
The ratio of sales to total assets	(S/TA)	A6	Divide sales by total assets	Carazo et al. (2010)
The ratio of net earnings to total assets	(NE/TA)	A7	Divide net earnings by the sum of total assets	Doering et al. (2019)

Preprocessing

Due to the unequal number of Sortino and Markowitz portfolios, it is necessary to apply a balancing (equalizer) structure to equalize the number of companies surveyed in terms of the above two portfolios. Suppose the data is entered without using the data balancing process. In that case, the group system analysis process will be detected accurately. The opposite group will be detected with lower accuracy due to the low number of input samples. Therefore, as explained, based on the separation of value stocks and growth stocks (companies with the highest value ratios, value companies, and companies with the lowest value ratios will be growth companies) has been balanced the Sortino and Marquitz portfolios.

The structure of analysis in a hybrid analysis algorithm

In this research, the general structure of the analysis process is in the form of Figure (2).

maximize its distance from the nearest data. Support vectors are the closest training points to the maximum margin of the super plate. Such a hyperplane, if it exists, is called a hyper-plane with a maximum margin. If the training points are $[x_i, y_i]$, the input vector is $x_i \in R_n$, and the value of the class is $i = 1, \dots, l$; is defined, $y_i \in \{-1, 1\}$, then, when the data are linearly separable, the decision rules are defined and will be concerning (1) with an optimal plane that separates the binary decision classes.

$$y = \text{Sign}(\sum_{i=1}^N y_i a_i (X \cdot X_i) + b) \quad (1)$$

Where: y is the output of the equation, y_i is the value of the educational sample class, and x_i indicates the internal coefficient. The vector $x = (x_1, x_2, \dots, x_n)$ represents input data, and the vectors $x_i: i = 1 \dots, N$ are backup vectors. In relation (1), parameters b and a_i determine the effectiveness. If the data are not linearly separable, equation (1) changes to equation (2).

$$Y = \text{Sign}(\sum_{i=1}^N y_i a_i K(X, X_i) + b) \quad (2)$$

The $K(X, X_i)$ function is a kernel function that generates internal beats to create machines with different types of nonlinear decision levels in the data space (Falahpour et al., 2013). Different kernels are used for the support vector machine regression model (first hypothesis): linear, quadratic, Gaussian, and polynomial. The Gaussian radial basis kernel function (RBF) usually has a radial function to predict better performance. The equation of this kernel function is equation (3):

$$K(X, X_i) = \exp\left(\frac{\|x_i - y_i\|^2}{2\sigma^2}\right) \quad (3)$$

After entering the corporate data into the SVM algorithm, this algorithm obtains the alpha coefficients of the model σ and then applies the data that the algorithm has not yet estimated to the model to measure the accuracy of the prediction (Salehi & Aminifard, 2012), and then based on three MSE criteria for grade point mean; MAE is the mean absolute value of the error and R2 is the coefficient of determination of the portfolio difference level.

Adaptive Neural-Fuzzy Inference System (ANFIS): The adaptive neural-fuzzy inference system is a hybrid system that combines the ability to make fuzzy logic with neural network computing capabilities and offers a complex and high level of modeling and estimation. This system has the advantages of both models, which means that it uses the educability of neural networks and the high decision-making power of fuzzy systems in conditions of uncertainty

and uncertainty. This model uses two error propagation algorithms and a combined method (combination of descending gradient method and least squares error method) for network training, which can reduce the complexity of the algorithm and, at the same time, improve network learning (Nasrollah Sarva Aghaji et al., 2016). Also, the fuzzy inference system used in it is the Sugino model, which is used to extract the fuzzy rules and the output of the system (Singh et al., 2012). Adaptive Neural-Fuzzy Inference System (ANFIS) is a fuzzy inference system with two x inputs, and one F output is assumed, as shown in Figure (3).

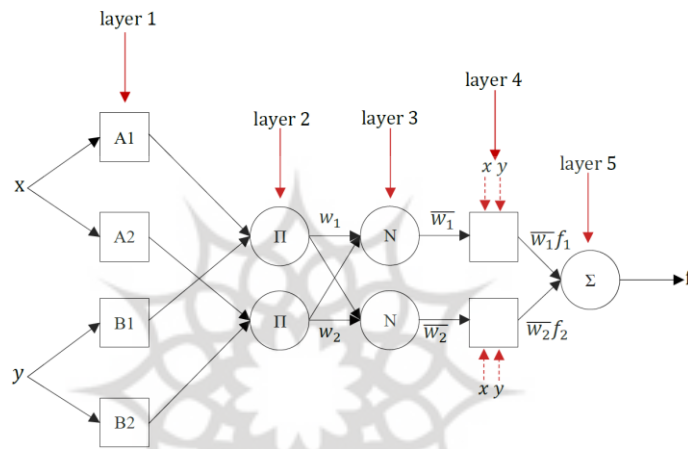


Figure 3. Adaptive Neural Fuzzy Inference System (ANFIS)

Based on the above figure, circle, and square shapes have been used to show how different nodes work. Each square node is an adaptive node with adjustable parameters, and each node of a circular shape is known as a fixed node. In the first layer, the values of the inputs are compared with the values of the corresponding membership functions, and the degree of compliance of each input with the corresponding membership function is selected as the output of the node. Accordingly, for the first time in the Sugeno fuzzy model, a rule consisting of two-phase If-Then sets is given in relation to Equation (4) and Equation (5):

$$\text{Rule 1: if } x \text{ is } A_1 \text{ and } y \text{ is then } f_1 = p_1x + q_1y + r_1 \tag{4}$$

$$\text{Rule 2: if } x \text{ is } A_2 \text{ and } y \text{ is then } f_2 = p_2x + q_2y + r_2 \tag{5}$$

The node in position i -th of the k -th layer is shown as O_k, i , and the node functions in the same layer of the same function family are as follows: layer 1 is the input layer, and each node in this layer is a square node with a

membership function, given in Equation (6) and Equation (7):

$$O_{1,i} = \mu A_i(x) \quad \text{for } i = 1,2 \quad (6)$$

$$O_{1,i} = \mu B_{i-2}(y) \quad \text{for } i = 3,4 \quad (7)$$

$O_{1,i}$ is a function of membership of A_i . The input Gaussian membership function of Equation (8) is a maximum of 1 and a minimum of zero. Experimental evidence and data analysis show that this distribution predicts the desirability and effectiveness of a relatively stable and reasonable portfolio.

$$\mu A_i(x) = e^{-\frac{(x-c)^2}{2\sigma^2}} \quad (8)$$

In the above relation, c is the mean, and σ is the variance of the membership function. Each node in layer 2 is a circular node labeled (norm function). The multiplication of the input signal is expressed by Equation (9):

$$O_{2,i} = w_i = \mu A_i(x) \times \mu B_i(y) \quad \text{for } i = 1,2 \quad (9)$$

Each node in layer three is labeled with a circle in Figure (4). The weights in this path are normally related to (10):

$$O_{3,i} = \bar{w}_i = \frac{w_i}{w_1 + w_2} \quad \text{for } i = 1,2 \quad (10)$$

Each node in layer 4 enters the membership function for the same node.

$$O_{4,i} = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \quad \text{for } i = 1 \quad (11)$$

Where p_1, q_1 , and r_1 are variables. This layer has a circular node labeled Sigma element whose final output equals the set of inputs (Equation 12).

$$O_{5,i} = \sum_{i=1}^2 \bar{w}_i f_i = \frac{\sum_{i=1}^2 w_i f_i}{\sum_{i=1}^2 w_i} \quad (12)$$

According to the given explanations, it should be stated that by using the possible structures of neural network and support vector (using vectorization), it is possible to provide the best and most effective portfolio selection in the above models, that depending on the type of data of the surveyed companies, its significant difference can be used as a basis for predicting the most effective portfolio tool.

Research Hypothesis

1. Based on the meta-heuristic algorithm, the Sortino (X) portfolio significantly differs from the Markowitz (Y) portfolio.
2. The accuracy of the Adaptive Neural Fuzzy Inference Analysis System (ANFIS) is higher than that of the Support Vector Machine Analysis

(SVM) system to select the most effective portfolio from the Sortino and Marquitz portfolio.

To design the main structure of the system, MATLAB software version R2018b has been used. The relationship between the two software is coded for standard reporting of Visual Basic for Applications (VBA) system in the Microsoft Excel 2013 environment. In addition to the above, parts of the Java language in the MATLAB environment have been used to improve system performance and efficiency. SQL Server 2014 is also used to communicate with the database and convert and transfer data.

Findings

In order to better explain the differences between the portfolios studied in this section, out of a total of 98 firms selected as the sample size, we compare 14 companies with a high level of investment in both Marquez and Sortino methods:

Table 2. Comparison of Markowitz and Sortinor portfolios

Row	Firm's Name	Symbol	Industry's name	Sortino portfolio	Markowitz portfolio
1	Pars Khazar Industrial	Lakhzer	Equipment and Machinery	0.082	0.243
2	Iran Khodro	Khodro	Car and parts manufacturing	0.111	0.382
3	Pars Khodro	Khpars	Car and parts manufacturing	0.114	0.313
4	Tractor forging	Khahan	Car and parts manufacturing	0.091	0.408
5	Iran Khodro Diesel Firm	East	Car and parts manufacturing	0.128	0.312
6	Saipa Glass	Casapa	Other non-metallic mineral products	0.141	0.173
7	Chini Iran	Kachina	Ceramic tiles	0.082	0.209
8	Jaam Darou	Fajam	Manufacture of metal products	0.052	0.155
9	Pars Darou	Depars	Pharmaceutical materials and products	0.039	0.181
10	Pars Industrial Soot	Shadows	Chemical products	0.071	0.125
11	Tractor casting	Khatrak	Car and parts manufacturing	0.063	0.190
12	Sarma Afarin	Lesarma	Equipment and Machinery	0.113	0.176
13	Mazandaran Cement	Samazan	Cement, lime, and gypsum	0.142	0.255
14	Pars Minoo	Ghapino	Food and beverage products	0.073	0.101

According to the information provided by 14 companies, the highest portfolio is related to the Markowitz portfolio. The highest share is related to the investment portfolio of Tractor Forging Firm with a portfolio of 0.408, given that the average number of years under review has been used; in practice, there may be fluctuations in the stock performance of the Firm during the years under review. Also, the lowest share of the surveyed companies in the investment portfolios relates to Jaam Darou Firm manufacturing metal products with a portfolio of 0.041.

Table 3. Percentage of Markowitz and Sortino portfolios

Row	Firm's name	Symbol	Industry's name	Percentage of Sortino portfolio	Percentage of Markowitz's portfolio
1	Pars Khazar Industrial	Lakhzer	Equipment and Machinery	35.9%	14.11%
2	Iran Khodro	Khodro	Car and parts	9.018%	49.12%
3	Pars Khodro	Khpars	Car and parts	73.8%	15.13%
4	Tractor forging	Khahan	Car and parts	40.4%	8.19%
5	Iran Khodro	East	Car and parts	6.12%	22.13%
6	Saipa Glass	Casapa	Other non-metallic mineral	95.3%	10.54%
7	Chini Iran	Kachina	Ceramic tiles	18.6%	13.62%
8	Jaam Darou	Fajam	Manufacture of metal products	17.2%	6.3%
9	Pars Darou	Depars	Pharmaceutical materials	65.2%	7.10%
10	Pars Industrial Soot	Shadows	Chemical products	34.7%	15.42%
11	Tractor casting	Khatrak	Car and parts	44.3%	10.12%
12	Sarma Afarin	Lesarma	Car and parts	37.10%	18.59%
13	Mazandaran Cement	Samazan	Cement, lime, and gypsum	61.11%	19.5%
14	Pars Minoos	Ghapino	Food and beverage products	16.4%	10.13%

Based on the comparison of portfolio percentages, it was determined that the highest percentage of the portfolio is related to Iran Khodro Industrial Firm, with 49.12%. This result shows that in the period of research in most of the investment portfolios of Tractor Forging Firm, it has a proportionate percentage of investment. 49.12% of Iran Khodro shares belong to the Markowitz portfolio, the highest percentage among the portfolios of the companies under review.

Inferential analysis

As explained in the third part of the study, in the first hypothesis test, stocks were classified into five to better separate the two portfolios of Markowitz (X) and Sortino (Y) for selection and investment categories: value stocks to growth stocks. (K1) to (K5) were used to classify stocks so that value stocks (K1) and growth stocks (K5) were considered. The method of stock separation was also calculated based on the defined ratios so that those stocks with lower ratios are value stocks with higher returns than growth stocks. Then, the meta-algorithm based on the random and equal weights criterion was used separately to distinguish between the two portfolios in this study. Equal weights According to the model of Fama and French (1992), after the stocks of the surveyed companies in each quintile, were selected, all stocks with equal quintuplets were placed in the portfolios of this study, namely Markowitz (X) and Sortino (Y). Thus, a growth portfolio and a value portfolio were formed each year. Also, random weights were first formed by a vector of random values based on the number of shares of the surveyed companies in the form of each quintile; the values were then standardized, that is, the sum of all the components became vectors equal to one, and the portfolios were examined at random weights. Due to the increase in the number of portfolios formed during a year and the research period, not all portfolios were necessarily formed with equal weight. In other words, ten growth and ten value portfolios with random weights were created each year, and over five years, 50 portfolios were created randomly for comparison. Therefore, according to the first hypothesis of the study, which examines the significant difference between the Sortino (X) portfolio and the Markowitz (Y) portfolio based on the meta-heuristic algorithm, Table (4) shows the results of this hypothesis based on the growth of growth stocks and value stocks.

Table 4. Inferential Analysis

Portfolio	Fold	Criteria	K1	K2	K3	K4	K5	MSE	MAE	R ²
Sortino Portfolio (X)	1	Equal weights	0.232	0.241	0.355	0.527	0.543	0.042	0.187	0.572
		Random weights	0.608	0.387	0.444	0.589	0.695			
	2	Equal weights	012.0	0.009	0.056	0.154	0.116	0.042	0.173	0.641
		Random weights	093.0	0.124	0.147	0.189	0.156			
	3	Equal weights	007.0	0.012	0.035	0.044	0.024	0.031	0.152	0.443
		Random weights	047.0	0.063	0.078	0.053	0.101			
	4	Equal weights	487.0	0.263	0.275	0.412	0.439	0.053	0.189	0.687
		Random weights	078.0	0.309	0.272	0.218	0.367			
	5	Equal weights	064.0	0.014	0.70	0.088	0.111	0.037	0.148	0.485

		Random weights	145.0	0.098	0.122	0.139	0.202			
	6	Equal weights	019.0	0.010	0.038	0.027	0.064	0.055	0.193	0.694
		Random weights	080.0	0.053	0.044	0.063	0.052			
	7	Equal weights	242.0	0.217	0.287	0.315	0.548	0.041	0.169	0.603
		Random weights	264.0	0.204	0.256	0.341	0.647			
	8	Equal weights	084.0	0.015	0.076	0.068	0.183	0.054	0.186	0.654
		Random weights	046.0	0.022	0.065	0.061	0.172			
	9	Equal weights	013.0	0.008	0.043	0.031	0.074	0.046	0.174	0.665
		Random weights	042.0	0.039	0.062	0.059	0.092			
	10	Equal weights	0.061	0.012	0.069	0.083	0.106	0.061	0.198	0.702
Random weights		0.139	0.094	0.128	0.135	0.211				
Mean								0.046	0.1786	0.621
Portfolio	Fold	Criteria	K1	K2	K3	K4	K5	MSE	MAE	R²
Markowitz Portfolio (X)	1	Equal weights	0.183	0.109	0.192	0.223	0.217	0.028	0.108	0.219
		Random weights	0.201	0.213	0.315	0.387	0.405			
	2	Equal weights	0.018	0.110	0.063	0.142	0.139	0/103	0/106	0/278
		Random weights	0.054	0.112	0.139	0.152	0.108			
	3	Equal weights	0.016	0.037	0.065	0.079	0.101	0/029	0/111	0/237
		Random weights	0.081	0.093	0/117	0.109	0.162			
	4	Equal weights	0.113	0.156	0.219	0.308	0.414	0/038	0/126	0/232
		Random weights	0.063	0.090	0.115	0.135	0.210			
	5	Equal weights	0.076	0.093	0.116	0.149	0.232	0.030	0.106	0.285
		Random weights	0.104	0.096	0.130	0.176	0.304			
	6	Equal weights	0.076	0.112	0.143	0.105	0.202	0.043	0.121	0.312
		Random weights	0.083	0.061	0.049	0.077	0.094			
	7	Equal weights	0.311	0.328	0.405	0.443	0.516	0.62	0.182	0.511
		Random weights	0.273	0.226	0.296	0.353	0.584			
	8	Equal weights	0.102	0.097	0.115	0.209	0.317	0.034	0.116	0.301
		Random weights	0.83	0.066	0.093	0.155	0.228			
	9	Equal weights	0.066	0.054	0.083	0.109	0.152	0.028	0.107	0.220
		Random weights	0.103	0.118	0.140	0.137	0.214			
	10	Equal weights	0.074	0.061	0.089	0.117	0.261	0.054	0.177	0.427
		Random weights	0.167	0.210	0.318	0.359	0.532			
Mean								0.038	0.127	0.322

Based on the statistical analysis according to the analysis of the backup vector machine regression algorithm, it was determined that the mean absolute value of the MAE error of the Sortino (X) portfolio is higher than the Markowitz (Y) portfolio. This means there is a significant difference between the two portfolios due to the use of growth and value data separation in terms of random weights. It is also clear from the random weights in both portfolios that the Sortino portfolio is based on growth stocks because it has higher ratios of weights, which has safer returns; Markowitz portfolios, on the other hand, are based on value stocks due to lower ratios, which pursue short-term returns on investment. It also shows the difference in the percentage of prediction R2 between the two portfolios; both portfolios can predict the probabilities and enter other influential data on investment returns. However, the R2 portfolio of the Sortino (X) portfolio is higher than the Markowitz (Y) portfolio, which means more reliance on investing through the Sortino (X) portfolio for investors who are looking for longer-term investments based on a focus on growth stocks. In order to test the second hypothesis of the research, which states that the accuracy of the adaptive neural fuzzy inference analysis (ANFIS) system is higher than the accuracy of the support vector machine analysis (SVM) system to select the most effective portfolio from Sortino and Markowitz portfolio, the effectiveness of the portfolio should be estimated based on the research components, based on which the accuracy or error calculation indices should be determined based on the mean squared error (MSE). The root means squared error (RMSE), which is presented in relations (13) and (14).

$$MSE = \frac{1}{N} \sum_{n=1}^N (\text{actual} - \text{prediction})^2 \quad (20)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^N (\text{actual} - \text{prediction})^2} \quad (21)$$

In the above actual relations is equal to the actual value; prediction is equal to the predicted value, and N is equal to the total number of data. Also, the target values and output of the algorithm for training data (companies) and epoch are presented for comparison. As Figure (4) shows, the adaptive neural fuzzy inference system (ANFIS) model adequately predicts portfolio effectiveness utility values close to the actual value.

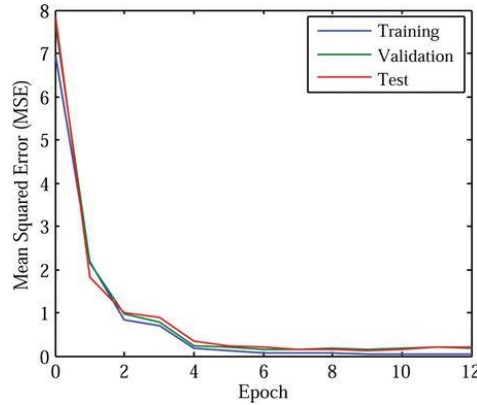


Figure 4. Error related to predicting educational and calibration data

As it turns out, the adaptive neural fuzzy inference system (ANFIS) based on principal component analysis, with a reduction in error, a very high correlation is observed between the actual data and the data predicted by the system, and it demonstrates the accuracy of the adaptive neural fuzzy inference system in predicting the effectiveness of a portfolio for investment. In order to differentiate the basis between the accuracy of Support Vector Machine Analysis (SVM) and the Adaptive Neural Fuzzy Inference System (ANFIS), the difference between the root mean square error (RMSE) and the mean absolute value of error (MAE) based on two portfolios is used. Table (5) shows a statistical comparison of optimization algorithms based on the Sortino portfolio (X) of the Markowitz portfolio (Y).

Table 5. Error Index Values Based on Optimization Algorithms

Error	Data type	NO.	ANFIS	SVM
Error index based on optimization algorithms	Education	42	0/14	0/06
	Epoch	12		
	Total	53		

This result shows that combining the optimization meta-heuristic algorithm with an adaptive neural fuzzy inference system based on research component analysis adjusts the optimal values for the parameters of these hybrid analysis systems, by reducing the error rate and accuracy of ANFIS adaptive neural fuzzy inference analysis compared to the accuracy of support vector machine (SVM) analysis, based on the difference between the root mean square error (RMSE) and the mean absolute value of the error (MAE), the ANFIS adaptive neural fuzzy inference is a better basis for selecting the effectiveness of the Sortino portfolio (X) from the Markowitz portfolio (Y).

Discussion and Conclusion

Relying on theoretical basics concerning the formation of a portfolio at the level of investment decisions of capital market companies, an attempt was made to examine the level of difference between the portfolios of Markowitz and Sortino in the form of meta-innovative analysis. In the second step, the basis of the adaptive neural fuzzy inference analysis (ANFIS) is to argue for the accuracy of the support vector machine analysis (SVM) system to select the most effective portfolio from the Sortino and Markowitz portfolios. Based on the results of the first hypothesis of the research, it was found that in order to better separate the two portfolios of Markowitz (X) and Sortino (Y) for selection and investment by using meta-heuristic algorithm analysis, the absolute mean error MAE of Sortino portfolio (X) from Markowitz portfolio (Y) is higher and this means that a significant difference was confirmed between Markowitz (X) and Sortino (Y). Therefore, relying on the analytical basis of the meta-innovative algorithm, which tries to select linear data from the data through linear classification of data, which has a more reliable margin, It should be noted that Firm managers tend to focus on growth stocks in order to choose a more investment portfolio because higher returns on random weights indicate potential returns on investment portfolio selection, they are more likely to use growth stocks. In this way, achieving longer-term than short-term returns plays a role in investment decisions. Referring to the theoretical basis of the portfolios considered in this study, it is clear that in the view of Markowitz's portfolio, the risk is considered a measure of possible volatility for future economic returns. Sortino portfolio, on the other hand, defines risk in terms of adverse deviations from the target rate of return, which means that higher returns are achieved through long-term risk control. Relying on the understanding of the concepts, it can be seen that Sortino's portfolio seeks the best return by estimating the desirability and disadvantage of returns and risk. Furthermore, since achieving higher returns requires time to invest, managers through Sortino's portfolio seek to control risk and increase returns in the long run. Simply put, the meta-heuristic algorithm analyzed the accuracy of estimating the returns on the Sortino portfolio to form a more important portfolio. Moreover, this shows the approach of focusing on growth stocks relative to value stocks. Based on the second result of the research hypothesis, it was found that the research hypothesis stated that the accuracy of the adaptive neural fuzzy inference analysis (ANFIS) system is higher than the accuracy of the support vector analysis (SVM) system in selecting the most effective portfolio from Sortino and Markowitz, based on the difference between the two roots of the mean square error (RMSE) which was equal to

(0.12) and the mean absolute value of the error (MAE) which was equal to (0.09), was confirmed. This result reflects the fact that to provide a reference for listed companies to evaluate the effective portfolio between the two Sortino (X) portfolios of the Markowitz (Y) portfolio, Adaptive Neural Fuzzy Inference Analysis (ANFIS) has a better ability to predict future investments. Studies have shown that the extracted fuzzy rules are logically correct and in accordance with reality, which is a sign of high power and good performance of the proposed algorithm. Using the analysis of the present result, Firm managers can make important decisions in optimal asset and debt management and change the Firm's performance by identifying the basis of investment portfolios to prevent a reduction in total return on investment and irreversible economic losses. This result suggests that using adaptive neural fuzzy inference analysis helps managers make better estimates of uncertainty and, therefore, achieve higher returns. On the other hand, Firm managers can perform better than backup vector analysis by relying on the computational intelligence of this analysis to classify their portfolio data to form portfolio returns. Because this system uses the ability to learn and optimize the neural network and linguistic expression of fuzzy inference simultaneously and by relying on a more accurate estimate of the mean squared error (RMSE), it can reduce the level of difference between the actual value and the predicted prediction value as much as possible. Finally, ANFIS adaptive neural fuzzy inference-based analysis provides the best predictive functions by checking all possible scenarios. Moreover, by mastering these fuzzy rules and realizing the knowledge gained by the Adaptive Neural Inference System (ANFIS), corporate executives can help the importance and relationships between the effective components in predicting the effectiveness of the right portfolio selection. At the same time, it is possible to identify investment opportunities with high accuracy according to the obtained results; it is suggested that the managers of companies, in choosing their investment portfolio while surrounding environmental changes, base their investment expectations on these changes. This way, they can choose the best strategy to form their investment portfolio and gain more returns and risk control. Because the lack of knowledge of environmental changes and alignment with the level of stakeholder expectations in investment leads to incorrect determination of investment risk coefficient and significant differences in optimal allocations. Also, the use of forecast analysis systems can sometimes strengthen the functions of forming a portfolio of investments, that having an analytical knowledge of the content and processes of those analyses, it is necessary to choose a balanced structure, such as adaptive neural inference analysis, to increase the accuracy of portfolio formation estimates.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest concerning the research, authorship and, or publication of this article.

Funding

The authors received no financial support for the research, authorship and, or publication of this article.

References

- Asadi, Gholam Hossein., Islami Bidgoli, Saeed (2014). Comparing the performance of value and growth strategies; Single Ratios and Combined Criteria, *Financial Research Quarterly*, Volume 16, Number 1, Spring and Summer, 1-24.
- Carazo, A. F., Gomez, T., Molina, J., Hernandez Diaz, A. G., Guerrero, F. M., Caballero, R. (2010). Solving a Comprehensive Model for Multi-objective Project Portfolio Selection", *Computers and Operations Research*, 37(4): 630-639. <https://doi.org/10.1016/j.cor.2009.06.012>
- Doering, J., Kizs, R., Juan, A. A., Fito, A., Polat, O. (2019). Metaheuristics for rich portfolio optimisation and risk management: Current state and future trends, *Operations Research Perspectives*, 6(4): 1-19. <https://doi.org/10.1016/j.orp.2019.100121>
- Falahpour, Saeed., Gol Arzi, Gholam Hossein., Fatorehchian, Nasser. (2013). Predicting stock price movement using support vector machine based on genetic algorithm in Tehran Stock Exchange, *Financial Research*, 15 (2): 269-288.
- Fallahpour, S., Safari, H., Omrani, N. (2014). Portfolio Selection Using Combination of Logarithmic Fuzzy Preference Programming Method and PROMETHEE. *Financial Management Strategy*, 2(2), 103–120.
- Faqhinejad, Mohammad Taghi., Minaei, Behrooz. (2018). Predicting stock market behavior based on artificial neural networks with intelligent collective learning approach, *Industrial Management*, 10 (2): 315-334.
- Fattahi nafchi, H., Arabsalehi, M., Esmaelian, M. (2019). Selection of optimal stock portfolios using accounting information, value-based information and balanced scorecard information. - Study: Companies listed in Tehran. , 11(2), 285–320.
- Firoozdehghan, Mohammad., Saeedi, Hadi., Mohammadi, Shaban., Elahi, Qhasem. (2020). High Frequency Data Portfolio Selection: Fixed Relative Avoidance and Liquidity Priorities, *Financial Engineering, and Securities Management*, 10 (38): 180-214.

- Gupta, P., Mehlawat M, K., Inuiguchi, M., Chandra S. (2014). Multi-criteria Portfolio Optimization Using Support Vector Machines and Genetic Algorithms. In: Fuzzy Portfolio Optimization. Studies in Fuzziness and Soft Computing, 3(16): 214–255. https://doi.org/10.1007/978-3-642-54652-5_10
- Haugen R (2010). *The New Finance: Overreaction, Complexity, and Uniqueness*, 5th Edition, Prentice Hall, New Jersey.
- Herberger, T, A., Reinle, F. (2020). A framework for screening and portfolio selection in corporate venture capital, *Studies in Economics and Finance*, 37(3): 475-495. <https://doi.org/10.1108/SEF-06-2019-0224>
- Ince, H., Trafalias, Th, B. (2007). Kernel principal component analysis and support vector machines for stock price prediction, *IIE Transactions*, 39(6): 629-637. <https://doi.org/10.1080/07408170600897486>
- Kandahari, Mahsa., Shamshiri, Azima., Fathi, Saeed. (2017). Stock portfolio optimization based on nonparametric estimation methods, *production and operations management*, 8 (1): 175-184.
- Kistner, W. G (1995). Growth Versus Value Investment Strategies. *Healthcare Financial Management*, 49(7).
- Nasrullah Sura Aghaji, Sajjad., Alimardani, Reza., Sharifi, Mohammad., Taghizadeh Yazdi, Mohammad Reza. (2016). Prediction of waste generated in Tehran using fuzzy-neural comparative inference system and artificial neural networks; *Iranian Biosystem Engineering*, 47 (1): 183-175.
- Petkovic, D., Issa, M., D. Pavlovic, N., Zentner, L., Nor Ridzuan Daud, M. and Shamshirband, S. (2014). Contact positions estimation of sensing structure using adaptive neuro-fuzzy inference system, *Kybernetes*, 43(5): 783-796. <https://doi.org/10.1108/K-03-2013-0041>
- Rai, Reza., Baskha, Hamed., Mahdikhah, Hossein. (2021). Stock portfolio optimization using Mean-CVaR method and symmetric and asymmetric conditional variance heterogeneity approach, *Financial Research*, 22 (2): 149-159.
- Salehi, Allah Karam., Salehi, Borzoo (2016). Comparison of Explanatory Power of Fama and French Three-Factor and Five-Factor Models in Explaining Value and Growth Stock Returns, *Investment Knowledge Quarterly*, Fifth Year, No. 19, Fall, 129-143.
- Salehi, M., Aminifard, Z. (2012). Examination of Variables Affecting Dividend Forecast Using Hybrid Models of PSO-LARS and PSO-SVR Algorithms. *Empirical Research in Accounting*, 2(4), 111-130.
- Schellinger, B. (2020). Optimization of special cryptocurrency portfolios, *Journal of Risk Finance*, 21(2) 127-157. <https://doi.org/10.1108/JRF-11-2019-0221>

- Senarathne, C. W. (2019). Do Fama–French common risk-factor portfolio investors herd on a daily basis? Implications for common risk-factor regressions, *Journal of Capital Markets Studies*, 3(2): 137-156. <https://doi.org/10.1108/JCMS-06-2019-0034>
- Thompson, T H. (2020). An instructional note to obtain an efficient portfolio with a short sale restriction, *Managerial Finance*, Vol. ahead-of-print No. ahead-of-print. <https://doi.org/10.1108/MF-07-2020-0400>
- Vakilifard, Hamidreza., Pilehvari, Nazanin., Zaidi, Seyedeh Samaneh. (2014). Presenting a Model for Predicting Bankruptcy of Companies Listed on the Tehran Stock Exchange Using the Adaptive Neural Fuzzy Inference System (ANFIS), *Financial Engineering and Securities Management*, 5 (18): 17-30.
- Wang, M Y. (1999). Multiple-Benchmark and Multiple-Portfolio Optimization, *Financial Analysts Journal*, 55(1): 63–72. <https://doi.org/10.2469/faj.v55.n1.2242>
- Wang, Y., Xie, Z., Lou, I., Ung, W.K. and Mok, K.M. (2017). Algal bloom prediction by support vector machine and relevance vector machine with genetic algorithm optimization in freshwater reservoirs", *Engineering Computations*, 34(2): 664-679. <https://doi.org/10.1108/EC-11-2015-0356>

Bibliographic information of this paper for citing:

Darvishan, Mehdi; Abdoli, Mohammadreza; Hosseini, Mohammad Mehdi & Alibeiki, Esmail (2023). Predicting Optimal Portfolio by Algorithm Analysis Systems. *Iranian Journal of Finance*, 7(2), 68-93.
