



Sharif University of Technology
Scientia Iranica
Transactions E: Industrial Engineering
<http://scientiairanica.sharif.edu>



Mathematical modeling for a new portfolio selection problem in a bubble condition using a new risk measure

A. Ghahtarani^a, M. Sheikhmohammady^{a,*}, and A.A. Najafi^b

a. *Faculty of Industrial and Systems Engineering, Tarbiat Modares University, Tehran, Iran.*

b. *Faculty of Industrial Engineering, K.N. Toosi University of Technology, Tehran, Iran.*

Received 8 August 2018; received in revised form 30 October 2019; accepted 18 November 2019

KEYWORDS

Decision tree;
 Financial bubble;
 Fundamental value;
 Gene expression
 programming;
 Portfolio selection
 problem;
 Risk measure.

Abstract. A portfolio selection model is developed in this study using a new risk measure. The proposed risk measure is based on the fundamental value of stocks. For this purpose, a mathematical model is developed and transformed into an integer linear programming. In order to analyze its efficiency, the actual data of the Tehran Stock Exchange market are used in 12 scenarios to solve the proposed model. In order to evaluate the scenarios, data mining approaches are employed. Data mining methods used in this paper include Adaptive Neuro-Fuzzy Inference System (ANFIS), decision tree, random forest, Fisher Discriminant Analysis (FDA), and Gene Expression Programming (GEP). The best method for scenario evaluation is GEP based on numerical results. Hence, the market values are evaluated by this algorithm. Software packages like MATLAB, GEP xpero tools, and LINGO are used to solve the model. Different trends of market value and fundamental value volatility in the optimum stock portfolio are determined. It is possible to examine the optimum portfolio profitability in different scenarios. By using real-world data, trends are extracted and analyzed. Results show that the developed model can be effectively applied in bubble situations.

© 2021 Sharif University of Technology. All rights reserved.

1. Introduction

Portfolio selection problem is one of the most important issues in finance in which investors try to maximize return and minimize risk. Markowitz was the first researcher who developed a mathematical definition of risk measures [1]. However, his risk measure was not linear and was subject to some calculation issues. After developing the mentioned measure, many researchers have tried to develop new risk measures. Konno and Yamazaki [2] presented Mean Absolute Deviation

(MAD) as a risk measure. Their risk measure was based on the stocks' rate of return. It was linear, but MAD considered both positive and negative deviations from mean as a risk which was not practically acceptable. The next huge step toward a proper risk measure was taken by Rockafellar and Uryasev [3] who developed Conditional Value at Risk (CVaR). This new risk measure is convex and has a linear programming model. CVaR is a downside risk measure and uses historical data with the assumption that historical trends will happen in the future. Moreover, this risk measure uses the rate of return for its calculation. They also developed Conditional Drawdown at Risk (CDaR) which used the maximum price of a stock in the investment period as a threshold. However, this risk measure is too conservative. Bernardi et al. [4] presented multiple risk measures for multivariate dynamic heavy-tailed models. They studied the evolution of risk interdependence

*. *Corresponding author. Tel.: +98 21 82884394*
E-mail addresses: alighahtaran@gmail.com (A. Ghahtarani); msheikhm@modares.ac.ir (M. Sheikhmohammady); aanajafi@kntu.ac.ir (A.A. Najafi)

and proposed a new risk measure to capture tail co-movements. This risk measure uses the stock rate of return. Kuzubaş et al. [5] proposed network centrality measures and systemic risk. They used data from Turkish Interbank market in the financial crisis and did not consider fundamental values in their analysis. Fu et al. [6] developed a convex risk measure based on the generalized lower deviation. This risk measure indicates the risk aversion of investors in non-linear equations. They used stocks' rate of return in their risk measure. Fu et al. [6] illustrated deviations and asymptotic behavior of convex and coherent entropic risk measures. Gong and Zhuang [7] measured the financial risk and portfolio reversion with time changed and employed stochastic volatility through tempered stable Lévy processes to construct time changed tempered stable Lévy processes. Sorwar and Dowd [8] estimated financial risk measures for options. They applied simulation-lattice procedure estimation and did not consider the fundamental factors in their analysis. Many researchers have worked on developing financial risk measures [9–15]. However, these researches have a gap in considering fundamental factors and bubble conditions.

Many of these risk measures are based on the difference between the expected rate of return at the beginning and that in the last period of investment. One of the most important criticisms of these risk measures is about the nature of the stock market value which facilitates calculating risk measures. In fact, this assumption is based on the fact that the stock market value is an indication of all events and phenomena around corporates including good or bad news about them. However, some part of stock value volatility is the result of investors' emotions about events. These volatilities are mostly short-lived. In fact, after a while, stock price reverts to its former price channel. Nevertheless, traditional risk measures have no procedure so as to confront these volatilities. The stock market crash under economic crisis has shown that traditional risk measure cannot evaluate risk properly, especially in emotional and crisis situations. This circumstance creates a bubble phenomenon.

Some stocks' market prices are greater than the fundamental values, which cause a bubble. In recent years, some researchers have focused on bubble analysis in the financial market. Domeij and Ellingsen [16] proposed a quantitative analysis of rational bubbles and public debt policy. They used bubble analysis for securities in public sectors. Barberis et al. [17] wrote an article paper entitled “Extrapolation and Bubbles”. They developed a model for extrapolation of bubbles. Lee and Philips [18] proposed performing asset pricing considering financial bubble risk. They provided an empirical investigation into risk factors in bubble pricing. Bosi et al. [19] proposed an article

paper entitled “financial bubbles and capital accumulation in altruistic economies”. They studied the global dynamics of capital stocks and assets' values. Maynard and Ren [20] assessed the power of long-term predictive tests in a model that contains financial asset bubbles. Michaelides et al. [21] developed non-linearities in financial bubbles. They used theory and Bayesian evidence from S&P500 and attempted to find and date non-linear bubble episodes, which are captured using a neural network. Miao and Wang [22] surveyed banking bubbles and financial crises. Wigniolle [23] considered the role of optimism and pessimism concerning financial bubbles. They demonstrated that it was possible to extend the scope of rational bubbles when uncertainty was introduced in association with rank-dependent expected utility. Harvey et al. [24] developed tests for explosive financial bubbles in the presence of non-stationary volatility. Kunieda and Shibata [25] presented a tractable model in which asset bubbles could exist in spite of infinitely lived agents and found that the policy of purchasing an asset should involve predicting the possibility of financial crises and avoiding them.

Many researchers have focused on bubble analysis [26–30]. The above-mentioned researches are among the newest of researches in this context. However, there are still some important gaps in the literature. Researchers who have developed risk measures did not consider a fundamental factor and bubble analysis in their models. On the other hand, researchers who have worked on bubble analysis have not developed risk measures. Consequently, there is a gap in this research area. Developing a new risk measure which takes into account bubble conditions can fill this gap.

To overcome the weakness of traditional risk measures, we need the development of a new paradigm in risk measures. Many people in financial market invest in financial assets based on the fundamental value of stocks. The attitude of these investors is that if the market value of a stock is less than the fundamental value, then the probability of increasing market value to fundamental value is high. In this situation, investors attempt to buy stocks. On the other hand, if the market value of a stock is higher than the fundamental value, then the probability of decreasing market value to fundamental value is high. In this situation, investors often sell stocks. Accordingly, in this strategy, investors have profit and loss zones, as illustrated in Figure 1.

The aim of this paper is to present a formulation for portfolio selection problem based on a new definition of risk measure. Risk is measured in this paper based on the difference between fundamental value and the market value of stocks. The structure of this study is given in the following. In the second part, the innovation and research method are presented.

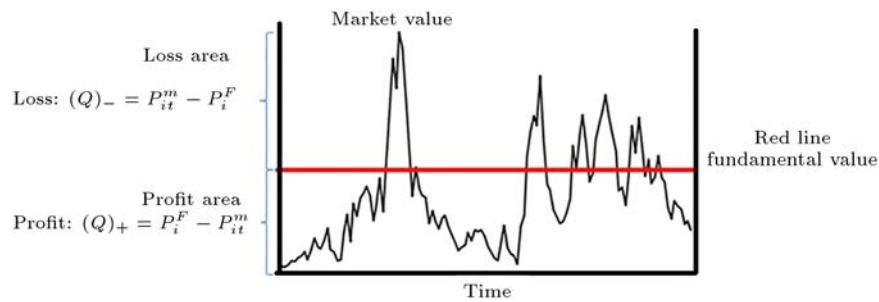


Figure 1. Definition of losing and profit areas (a hypothetical image).

The third section presents problem modeling and the method of calculating the stock fundamental value. The fourth section depicts market value prediction based on different scenarios. The fifth part contains the numerical results. Finally, the conclusion of the study is presented in the last part.

2. Innovation and research method

The most important innovation in this research is the conceptual and mathematical development of the portfolio selection model. In fact, in this research, a new concept is introduced as a risk and then, this new concept is transformed into a mathematical model using mathematical foundations. The concept which is developed in this research is generally different from financial risk measures in the literature of the financial field. On the other hand, the proposed risk measure is more consistent with the existing strategies of investment in the stock markets.

To develop a risk measure and develop a new portfolio selection model, this paper uses the concept of market and fundamental values of stocks. In the real world, many investors use the strategy of estimating the fundamental value and comparing it with the actual stock value to choose their optimum portfolio. In fact, investors first calculate the fundamental value of a stock and then, make a decision to buy or sell the stock by comparing that fundamental value with the market value. The key to this trading strategy is that based on diverse fundamental value calculations, the amount of fundamental value in an interval is roughly the same unless a particular event occurs for the company, e.g., an increase in capital, or for the outcome of a new development contract that improves the company's future cash flow. Furthermore, adverse events such as company losses or unexpected events or even macroeconomic problems in a country will reduce the value of the company's capital or its future cash flows. In this case, the fundamental value decreases, but these events are not issues that occur every day. Therefore, it is expected that the fundamental value of a stock will remain stable in a long-term interval. Generally, it is expected that the stock market value

will move towards its fundamental value. In reality, when market value is higher than the fundamental value, it is likely to be a transitory excitement in the market. However, prices will decrease over time. Even the demand pressure may create bubble in the market, which pushes the market value farther away from the fundamental value. If the market value collapses, investors will face a significant loss. On the other hand, if the market value is less than the fundamental value for various reasons, e.g., rumors or investors' inaccurate estimates, then the market value is expected to increase and reach its fundamental value after a while.

This brief explanation is a basis for introducing the concept of risk in this research. Investors should expect a significant risk when a stock's market value is greater than the fundamental value, thus forming a bubble. Furthermore, they should avoid investing in these conditions. On the other hand, when the market value is lower than the fundamental value of a stock, it is an investment opportunity and a chance for securing some profit. Therefore, two areas of profit and loss are formed in relation to the comparison between fundamental value and market value, which can be considered as the concept of risk.

The innovation of this research lies in the use of this trading strategy as the concept of risk and profitability as well as its mathematical modeling to provide the best possible optimum combination of stocks. The important point, considered in the modeling of this research, is the idea of developing a model that considers the market value in approaching different scenarios. Owing to the dynamics of the stock market and the volatility of market value, the proposed model has some advantages over the existing models to be counted below:

1. Traditional portfolio selection models focus only on the mathematical concepts of calculating the difference between the expected returns or expected level of confidence, but the proposed model in this research is more consistent with a practical strategy in the market. Hence, it is more understandable for investors,
2. The traditional approach relies heavily on investors'

expected returns. However, in a market with bubble conditions, even when the expected returns are reasonably relative to market conditions, if the bubble collapses, the reality will reveal itself and clearly, the expected returns of investors with real values and actual values of the market will be very different. For instance, assume that a particular asset in the market bubbles up and its prices grow at an average of 25%. If an investor expects a 25% profit, it seems reasonable and the investment seems to be successful. However, if the market bubble subsides and prices fall, the portfolio selection models that are designed and decided based on the expected return of 25% for the mentioned investor will be not only non-profitable but also highly detrimental, thus resulting in a huge loss of capital. On the contrary, the model and the concept developed in this study are not characterized by this drawback,

3. Traditional portfolio selection models have failed to consider the fundamental value of a stock. As a result, no attention has been paid to emotional issues and bubbles in the market, while the concept of risk as well as the proposed model in this study focus on this issue.

The actual data of 15 shares in the Tehran Stock Exchange under 12 scenarios are used in this study to solve the proposed model within the time duration of 5 years.

3. Mathematical formulation

The first step in the model formulation is the definition of parameters and decision variables. The present study used the fundamental value as a key factor to calculate a new risk measure. There are many different methods for estimating the fundamental value of a stock. One of the most commonly used ones is the free cash flow method. The free cash flow of a company is the cash flow obtained from company activities apart from the required capital expenditure for executing current operations with the increased production capacity. Free cash flow is calculated from the total net income after taxes minus depreciation costs, working capital costs, and capital expenditure. The stock pricing in this way is achieved in the form of the total current free cash flow value and cost of equity. The required indices of the model are listed in Table 1.

To define the new risk measure based on the difference between the fundamental and market values of the stock, it is required to perform some calculations for the model. The associated variables are given in Table 2.

Calculating the fundamental value of a share is of significance in this section. The operational free cash flow must be calculated in the first step:

Table 1. Indices definition of the model.

Name	Indices
Stock index	i
The maximum number of shares	N
Scenario index	t
The maximum number of scenario	s

$$OFCF = EBIT \times (1 - Ta) + DEPRECIATION - CAPEX - WORKING CAPITAL. \quad (1)$$

To calculate the fundamental value, the growth rate of the company should be calculated which is itself necessary for determining the discount rate. The growth rate is calculated through the following formula:

$$g = RR * ROIC. \quad (2)$$

In the above formula, there are two components that include the ratio of dividends to total and return on the invested capital. These two components are as follows:

$$RR = (1 - PAYOUT \text{ RATIO}), \quad (3)$$

$$PAYOUT \text{ RATIO} =$$

$$\frac{DIVIDENDS_PER_SHARE}{EARNING_PER_SHARE}, \quad (4)$$

$$ROIC = \frac{EBIT(1 - Ta)}{TOTAL_CAPITAL}. \quad (5)$$

To calculate the discount rate, two components are required. The next component is the weighted average cost of capital which is calculated as follows:

$$WACC = \frac{E}{V} * Re + \frac{D}{V} * Rd * (1 - T). \quad (6)$$

The required variables for calculating the weighted average cost of capital are presented in Table 2. The fundamental value is calculated as follows:

$$FIRM_VALUE = \frac{OFCF}{(1 + WACC - g)^t}. \quad (7)$$

Eq. (7) is a fundamental price based on a constant growth rate. This study defined two different zones based on the fundamental value. In case the market value is less than the fundamental value, the market value will get closer to the fundamental value level. If an investor buys stocks in this situation, some profit can be expected. On the other hand, if the market value is higher than the fundamental value, the stock value will decrease. This area is called the loss zone.

The amounts of loss and profit are calculated in the following:

$$Profit : (Q)_+ = P_i^F - P_{it}^m, \quad (8)$$

Table 2. Model variables.

Sign	Variable
EBIT	Profit before tax
Ta	Tax rate
CAPEX	Capital expenditure
DEPRECIATION	Depreciation costs
WORKING CAPITAL	Working capital
OFCF	Operating Free Cash Flow
G	Growth rate
RR	Retention Rate
ROIC	Return On the Invested Capital
PAYOUT RATIO	Payout ratio
WACC	Weighted Average Cost of Capital
Re	Cost of equity
Rd	Cost of debt
E	The market value of equity
D	The market value of debt
V	Sum of the market value of equity and market value of debts.
P_i^F	Fundamental value of stock i
P_{i0}^m	Market value of stock i at the beginning of the investment period
P_{it}^m	Market value of stock i based on scenario t
PP^F	Portfolio fundamental value
PP_t^m	Portfolio market value based on scenario t
x_i	The number of purchased shares from i share
B	Available budget
Lb	The lower limit of the authorized number of shares that can be bought from a share
Ub	The upper limit of the authorized number of shares that can be bought from a share
α	% of the expected profit
$p_t = \frac{1}{s}$	The probability of occurrence of the scenario t which we consider discrete uniform distribution function for it. Then, the probability is equal to $\frac{1}{s}$

$$Loss : (Q)_- = P_{it}^m - P_i^F. \quad (9)$$

Formulation needs two variables used for calculating both profit and loss. Here, Q_+ is a variable used in calculating the profit and must take the value of zero in case of loss. However, Q_- is a variable used for explaining loss while it must take zero in the profit zone. These two variables are as follows:

$$Q_+ = \max\{Q, 0\}, \quad (10)$$

$$Q_- = \min\{-Q, 0\}. \quad (11)$$

By using Eqs. (8) and (9) and substituting them in Eqs. (10) and (11), formulations are done as follows:

$$Q_+ = \max\{P_i^F - P_{it}^m, 0\}, \quad (12)$$

$$Q_- = \min\{P_{it}^m - P_i^F, 0\}. \quad (13)$$

Investors aim to maximize the profit and minimize the loss. This can be illustrated in a mathematical form through a fraction in which the numerator of the fraction is profit and the denominator is loss. Through maximization of this fraction, the numerator (profit) increases while the denominator (loss) decreases. This fraction is given below:

$$\max r = \frac{\max\{P_i^F - P_{it}^m, 0\}}{\min\{P_{it}^m - P_i^F, 0\}}. \quad (14)$$

Instead of one stock, the fraction can be used by the portfolio of stocks, as shown in the following:

$$\max r = \frac{\max\{PP^F - PP^m, 0\}}{\min\{PP^m - PP^F, 0\}}, \quad (15)$$

$$PP^F = \sum_{i=1}^n x_i P_i^F, \quad (16)$$

$$PP^m = \frac{1}{s} \sum_{i=1}^n \sum_{t=1}^s x_i P_{it}^m, \quad (17)$$

$$PP_t^m = \sum_{i=1}^n x_i P_{it}^m, \quad \forall t. \quad (18)$$

The fundamental value is calculated and shown in the following:

$$P_i^F = \frac{\text{Firm_value}}{\text{Total_number_of_stocks}}. \quad (19)$$

The portfolio selection problem based on a new concept of risk measure is:

$$\max r = \frac{PP^F - PP^m}{z}, \quad (20)$$

$$\frac{1}{s} \sum_{i=1}^n \sum_{t=1}^s x_i P_{it}^m - \sum_{i=1}^n x_i P_{i0}^m \geq \alpha B, \quad (21)$$

$$\sum_{i=1}^n x_i P_{i0}^m \leq B, \quad (22)$$

$$z = \sum_{t=1}^s d_t p_t \quad p_t = \frac{1}{s}, \quad (23)$$

$$d_t \geq PP_t^m - PP^F, \quad (24)$$

$$d_t \geq 0, \quad (25)$$

$$x_i \geq 0 \quad \text{and} \quad \text{integer}. \quad (26)$$

In this model, Eq. (20) is the objective function of the model that represents the profit-to-loss ratio. This type of ratio is very common in financial models. Among the most famous profit-to-loss ratios in financial problems are sharp and Omega ratios. This objective function aims to maximize profits and minimize losses. The numerator of the fraction is the calculation of the profit area, aiming at finding a combination of stocks where the market value is as low as possible compared to the fundamental value. The denominator is the value of z which is the area of loss, calculated through Eqs. (23)–(25). These limitations resurface exactly when measuring the area of losses. According to Guastaroba et al. (2016) and Mansini et al. (2003), these types of ratios, i.e., profit-to-loss, can be negative [31,32]. The negative value of this objective function in the proposed model points to the formation of bubble in the market. In case the market value is greater than the fundamental value, the objective function can get negative value which is a bubble condition. The proposed model enjoys the ability to clearly pinpoint the bubble formation by a negative value in the objective function.

Eq. (21) evaluates the yield of an investment. In this constraint, the average of market values of the portfolio based on different scenarios minus the market price of the portfolio at the beginning of the investment period shows the investment yield. Here, s is the number of scenarios based on different economic conditions, n the number of candidate stocks, P_{it}^m the market value of stock i in scenario t , and α the expected rate of return. In this formulation, the yield of an investment is a fraction of the budget.

Eq. (22) is related to the budget constraint for each scenario. In this research model, p_{i0}^m is the market pricing of stock i at the beginning of the investment period.

Constraints (23), (24), and (25) are used for measuring the area of loss, and Eq. (26) is also employed to determine whether the decision variable is greater than or equal to zero. Decision variables should also be an integer number.

In Objective Function (20), z is in the denominator the value of which should not be zero because; otherwise, the fraction is not defined and the model will be unacceptable; therefore, the value of z must be nonzero. As a result, the value of z is assumed to be larger than a nonzero ε value [31–33].

$$z \geq \varepsilon \quad \varepsilon \neq 0, \quad (27)$$

$$z > \frac{1}{M}, \quad (28)$$

$$\varepsilon = \frac{1}{M} \quad M \rightarrow \infty. \quad (29)$$

The model that considers new variables is given below:

$$\max r = \frac{PP^F - PP^m}{z}, \quad (30)$$

$$\sum_{i=1}^n x_i P_{i0}^m \leq B, \quad (31)$$

$$z = \sum_{t=1}^s d_t P_t \quad P_t = \frac{1}{s}, \quad (32)$$

$$d_t \geq PP_t^m - PP^F \quad \forall t, \quad (33)$$

$$d_t \geq 0, \quad (34)$$

$$x_i \geq 0. \quad (35)$$

The linearization of the problem is presented as follows. Initially, the fraction of Objective Function (30) is divided into two fractions shown below:

$$\max r = \frac{PP^F}{z} - \frac{PP^m}{z}. \quad (36)$$

There is a need for a variable change in this stage. This variable can change the objective function fraction to a linear function. This variable is:

$$\frac{PP^F}{z} = \nu. \quad (37)$$

On the contrary, there should be no factor in the form of a fraction; therefore, the change in the second variable is shown in the following:

$$\frac{1}{z} = \nu_0. \quad (38)$$

With the consideration of the two above-mentioned variables, the objective function can be rewritten as follows:

$$\max \quad \nu - PP^m \nu_0. \quad (39)$$

This variation of the introduced variables should be applied to all parts of the research model since the research model must be integrated on the basis of a series of decision variables. Therefore, the two sides of the research limit are divided by z . However, this change requires new variables for the model to change. These variations are as follows:

$$\tilde{d}_t = \frac{d_t}{z}, \quad (40)$$

$$\tilde{B} = \frac{B}{z} = B\nu_0, \quad (41)$$

$$\tilde{x}_i = \frac{x_i}{z}. \quad (42)$$

The three above-mentioned variables are used for situations in which the two sides of the inequalities of constraints are divided by the value of z . As already stated, this study simulates the variables and parameters of the model. An example of this operation is:

$$\sum_{i=1}^n x_i P_{i0}^m \leq \frac{B}{z} \xrightarrow{\div z}, \quad (43)$$

$$\sum_{i=1}^n \frac{x_i}{z} P_{i0}^m \leq \frac{B}{z}. \quad (44)$$

Accordingly, we can rewrite the model as follows:

$$\max \sum_{i=1}^n \tilde{x}_i P_i^F - \frac{1}{s} \sum_{i=1}^n \sum_{t=1}^s \tilde{x}_i P_{it}^m = \max \quad \nu - PP^m \nu_0, \quad (45)$$

$$\sum_{i=1}^n \tilde{x}_i P_{i0}^m \leq B\nu_0, \quad (46)$$

$$\frac{1}{s} \sum_{t=1}^s \tilde{d}_t = 1, \quad (47)$$

$$\tilde{d}_t \geq \frac{1}{s} \sum_{i=1}^n \sum_{t=1}^s \tilde{x}_i P_{it}^m - \sum_{i=1}^n \tilde{x}_i P_i^F \quad \forall t, \quad (48)$$

$$\frac{1}{s} \sum_{i=1}^n \sum_{t=1}^s \tilde{x}_i P_{it}^m - \sum_{i=1}^n \tilde{x}_i P_{i0}^m \geq \alpha B\nu_0, \quad (49)$$

$$\tilde{d}_t \geq 0, \quad (50)$$

$$\tilde{x}_i \geq 0, \quad (51)$$

$$\nu_0 \leq M. \quad (52)$$

Eqs. (45)–(51) represent Eqs. (20)–(26) of the non-linear programming models which have turned into linear programming models upon changing the introduced variables. Based on the concept of new risk introduced in this study, the model as well as a relaxed linear programming model offer a new approach to measuring risk. At the next step, an attempt is made to examine the analysis of the model presented with the real data.

4. Input data calculation (scenario production and data evaluation)

The proposed model is very sensitive to the input data. Therefore, market and fundamental values are required to measure the bubble. In addition, the corresponding risk measure needs these two values for each stock so that the bubble risk can be evaluated.

In this section, five data mining approaches with some similarities are presented. Decision tree, random forest, and Fisher Discriminant Analysis (FDA) are based on linear regression. Adaptive Neuro-Fuzzy Inference System (ANFIS), FDA, decision tree, and random forest are based on the inference engine. Gene Expression Programming (GEP), decision tree, and random forest are based on tree structure. In the case of using only one method, one could ask the reason for using that specific method. However, this comparison shows that GEP is the best data mining approach in this concept based on numerical results. In recent years, researchers in data science have compared different methods so that the best method can be selected for their own data. In this study, these studies have been referred to as the references to this approach [34–39]. In addition, this study employed this method and compared five data mining approaches to select the best one in order to calculate our input data.

This section is of high significance in this study. If we had used historical data as input, we would have used an inaccurate assumption (historical trends tend to be repeated in the future while it is not true in most cases).

Since there are numerous tables and graphs for each data mining method, the results of data mining approaches were presented for one stock only.

The data used in this study were divided into two categories of inputs and target. The input data were employed to assess the target that contain:

- Max price of the stock on a day;
- Min price of the stock on a day;
- The volume of stocks which have been bargained;
- Exchange rate;
- Oil price;
- Tehran stock index.

The output or the target is the market value of stock. The data used in this study comprised 15 stocks in the time period of five years from 2013 to 2017. The following approaches are employed in this study.

4.1. Decision tree

A decision tree is a tool for classification and prediction. It is a structure that includes a root node, branches, and leaf nodes. Each internal node denotes a test on an attribute, each branch denotes the outcome of a test, and each leaf node holds a class label. The topmost node in the tree is the root node. The decision tree is predicted based on rules, derived from the input data. The rules for prediction are established based on the relationship among the input data. Figure 2 shows a decision tree used for assessing the market value of stocks.

A decision tree is a machine learning method that maps the observable facts of inputs to target values through inference engine. In the decision tree, each internal node is a variable (in this case, input variables), and each leaf is the target value (in this case, the market value of stocks). A decision tree is

generated when each decision node in the tree contains a test on the values of some input variables. The terminal nodes of the tree contain the predicted output variable values. In this specific decision tree, regression is conducted based on the fractures of subtree. Initially, all records in the training set (pre-classified records that are used to determine the structure of the tree) are grouped into the same partition. The algorithm then begins allocating the data to the first two partitions or branches using every possible binary split in every field. The algorithm selects the split that minimizes the sum of the squared deviations from the mean in the two separate partitions. This splitting rule is then applied to each of the new branches. This process continues until each node reaches a user-specified minimum node size and becomes a terminal node [40]. Following steps shows the decision tree process:

Step 1: The standard deviation of the target is calculated;

Step 2: The dataset is then split into different attributes. The standard deviation for each branch is calculated. The resulting standard deviation is subtracted from the standard deviation before the split. The result is the standard deviation reduction;

Step 3: The attribute with the largest standard deviation reduction is chosen for the decision node;

Step 4: The dataset is divided based on the values of the selected attribute. This process is run recursively on the non-leaf branches until all data are processed.

After preprocessing the data and eliminating the lost data, both input data and target are applied to the

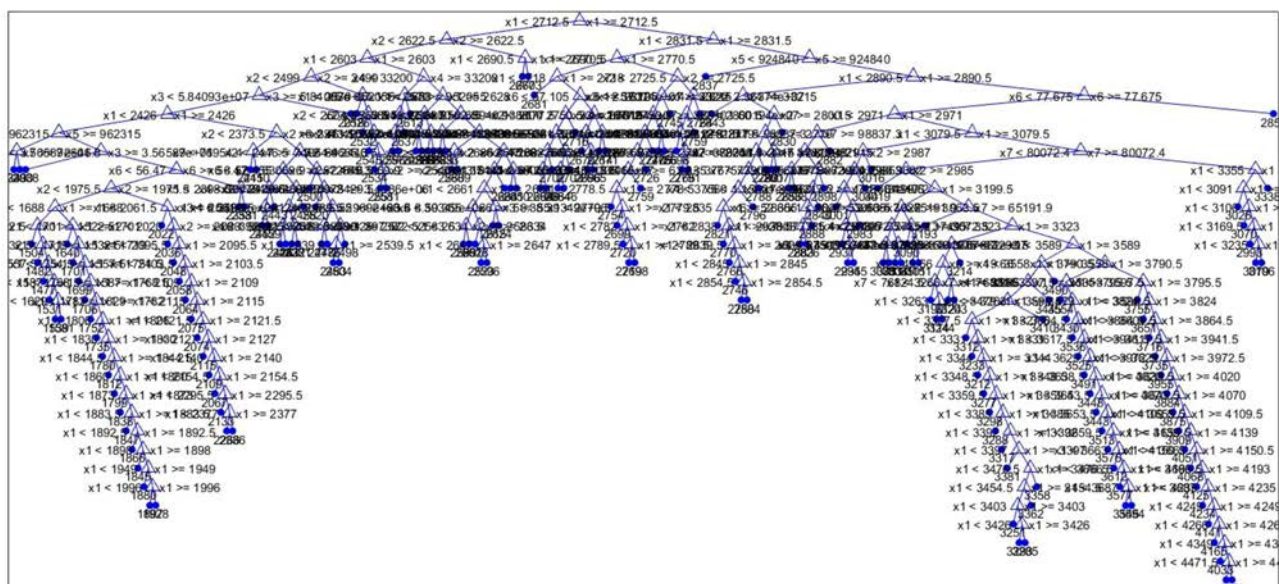


Figure 2. Decision tree for prediction of market value of one stock.

decision tree. Input variables are the independent variables, and target is the dependent variable. The training method is cross-validation. Each branch of the decision tree is based on rules, thus resulting in the creation of leaf, and each leaf of the tree is a target value. Figure 2 shows the decision tree used in this study. K-fold loss in this tree is 0.98. This algorithm was then applied to all 15 stocks in different scenarios.

4.2. Random forest

Random forest is a learning method based on classification and regression. The construction of a random forest is a multitude of decision trees. This algorithm is a group algorithm which is a set of decision trees. In a random forest, a set of trees is made. The classification accuracy for the random forest comes from voting among trees. The random forest produces enormous numbers of trees. Each tree gives a classification. Hence, each tree votes for a class. Trees with the largest number of votes are selected as the final forest. Random forest is built through a process known as binary recursive partitioning which is an iterative process that splits the data into partitions or branches; then, it continues splitting each partition into smaller groups as the method moves up each branch [41]. However, random decision forests correct the habit of overfitting to their training set for each decision tree. Figure 3 shows the random forest. This figure is like the decision tree. This figure also depicts the rules and equations built based on the relationship among the input variables.

One of the most prominent problems of random forest is overfitting. In order to solve this problem, the algorithm prunes the trees, i.e., every branch and

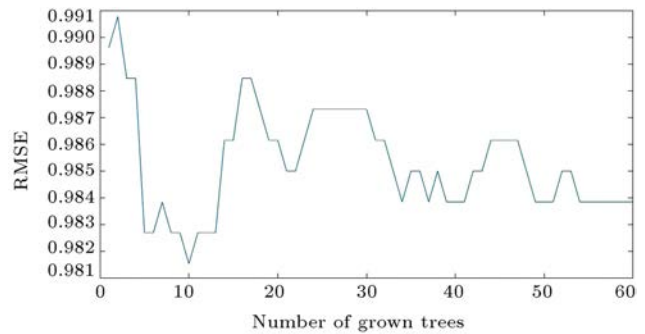


Figure 4. Pruning of random forest.

leaf are transferred to the preceding nodes. Figure 4 illustrates the pruning process.

Figure 4 shows that after pruning 60 times, the Root-Mean-Square Error (RMSE) declines to 0.98, indicating that there is no overfitting on the model.

4.3. Fisher Discriminant Analysis (FDA)

FDA is a statistical method employed to find a linear combination of traits in the best form. This method uses linear discriminant analysis to determine the pattern. Figure 5 illustrates a linear discriminant analysis that shows the priority of variables. Variables must be arranged in decision tree based on linear discriminant analysis.

Discriminant analysis works by creating one or more linear combinations of predictors, thus creating a new latent variable for each function. These functions are called discriminant functions. The number of functions possible is either $N_g - 1$, where N_g is the number of groups, or p (the number of predictors), whichever is smaller. The first created function maximizes the differences between groups on that function. The sec-

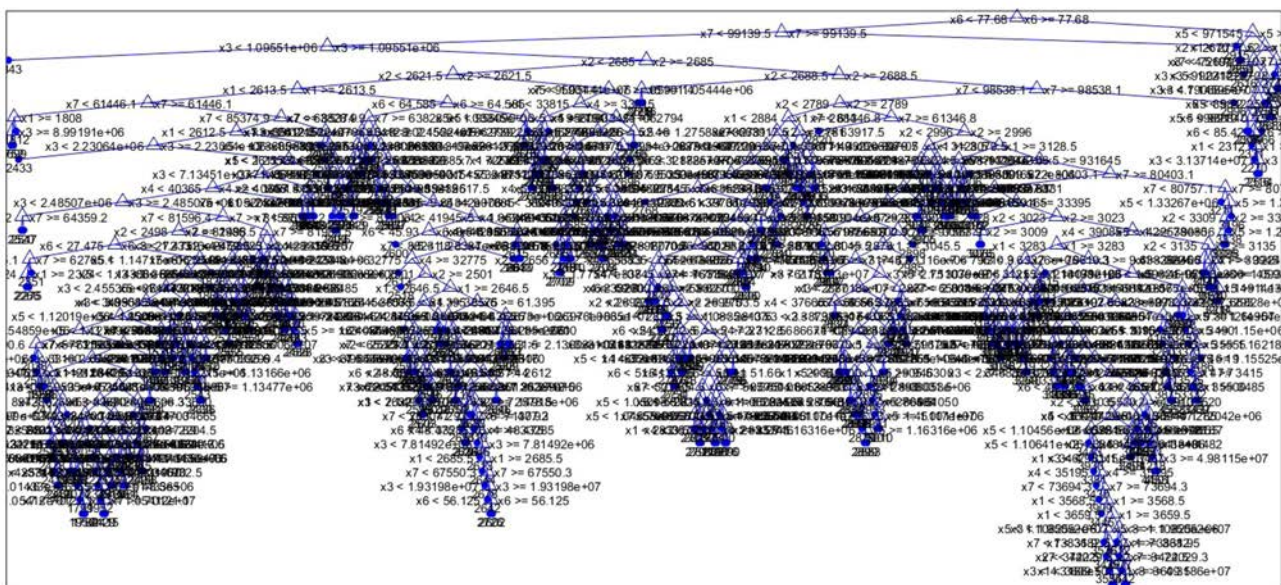


Figure 3. Random forest network for prediction.

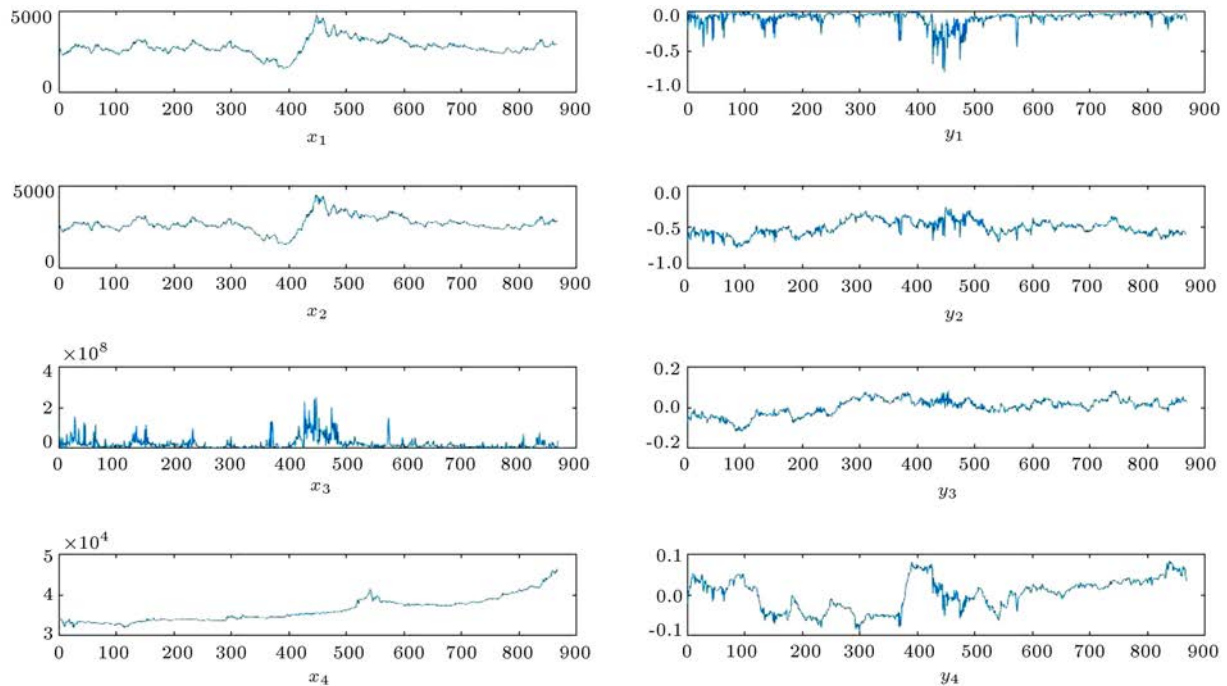


Figure 5. Linear discriminant analysis.

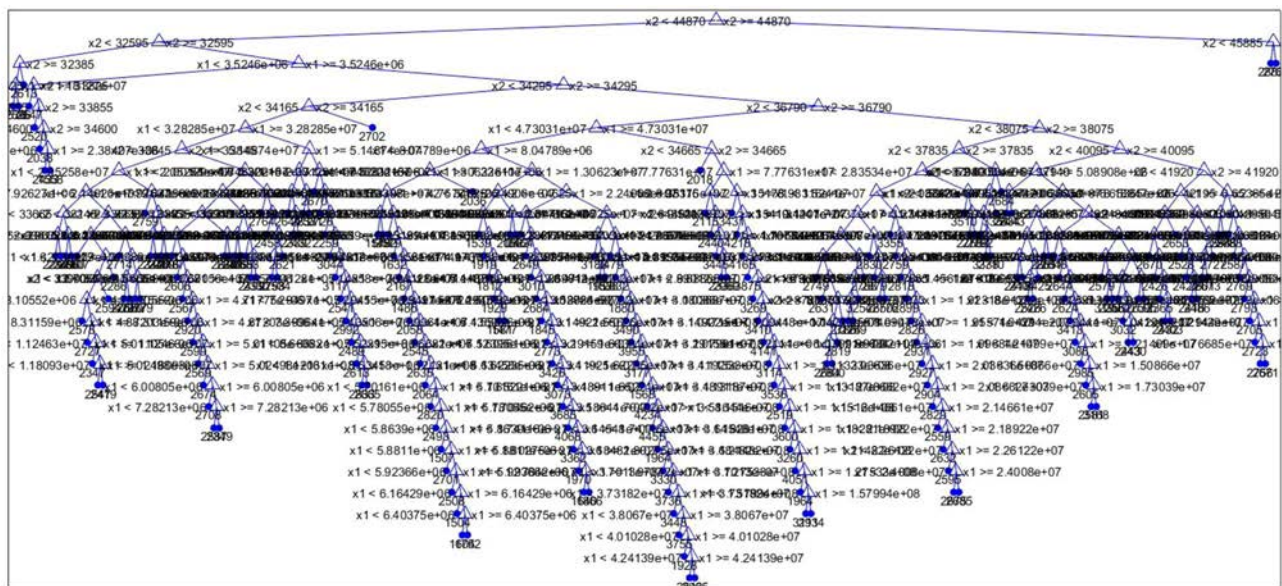


Figure 6. Decision tree based on Fisher Discriminant Analysis (FDA).

and function maximizes the differences between groups on that function, but must not be correlated with the previous function. The same holds for the subsequent functions provided that the new ones would not be correlated with any of the previous functions [38].

Figure 6 shows a decision tree based on FDA analysis. Upon determining the priority of variables, Figure 6 shows the decision tree based on the priority of the variables. This tree is better than the decision tree and random forest shown in this research given that FDA analysis shows the best combination of variables.

4.4. Adaptive Neuro-Fuzzy Inference System (ANFIS)

The ANFIS is a framework that operates based on if-then. This system has 3 parts:

1. Fuzzy rules;
2. Membership function;
3. Inference process.

In this research, we use TSK (Takagi-Sugeno-Kang) rules. The architecture is composed of five

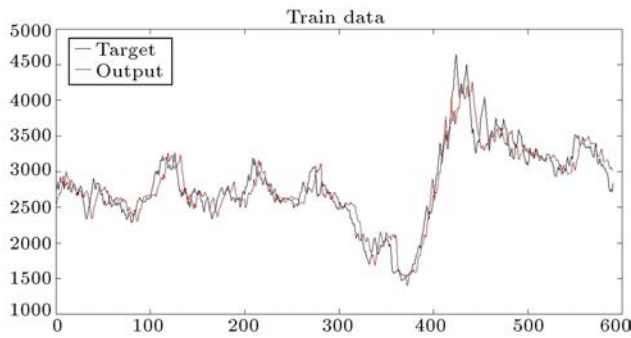


Figure 7. Fuzzy inference network (compatibility between real and predicted data).

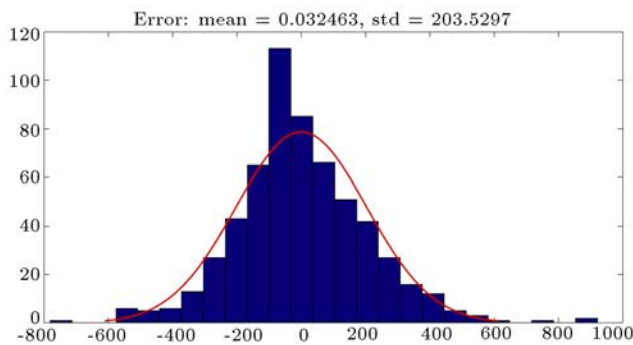


Figure 8. Root-Mean-Square Error (RMSE) index through training process.

layers. The first layer takes the input values and determines the membership functions belonging to them. It is commonly called fuzzification layer. The membership degrees of each function are computed by using the premise parameter set, namely $\{a, b, c\}$. The second layer is responsible for generating the firing strengths of the rules. In this task, the second layer represents “rule layer”. The role of the third layer is to normalize the computed firing strengths by dividing each value by the total firing strength. The fourth layer takes input as the normalized values and the resulting parameter set $\{p, q, r\}$. The values returned to this layer are the defuzzified ones and they are passed to the last layer to return the final output [42]. Eighty percent of inputs are training data and twenty percent belong to the trail. Figure 7 is the fuzzy inference network applied to the training data. The black curve is real data, while the red curve is predicted data. The graph shows that there exists a good fitness. Real and predicted data share acceptable compatibility.

Figure 8 shows the RMSE, which is the error index. RMSE in this stock is 0.0322. Figure 9 shows that the conformity of real data and prediction is 93 percent.

4.5. Gene Expression Programming (GEP)

GEP is a developed genetic algorithm. This algorithm operates based on population and each generation is selected based on fitness. Solution production was done through some genetic operators. The difference

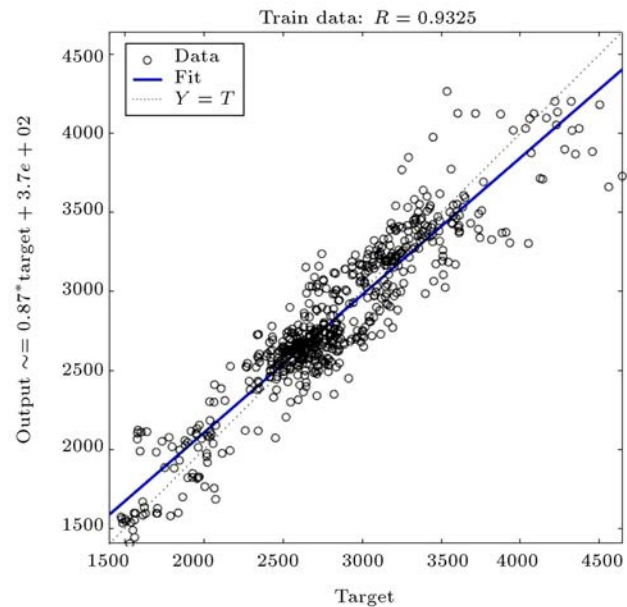


Figure 9. Conformity between real data and prediction.

between GEP and genetic algorithm is the essence of individuals. In GEP, chromosomes are the string with a constant length and they are expressed in a nonlinear string with different lengths. In this algorithm, 80 percent of data are the training data and 20 percent of data belong to the trail. Figure 10 shows the objective function of GEP, which is the mean square error. Basic steps of the GEP are as follows:

1. Select function set;
2. Select terminal set;
3. Load dataset for fitness evaluation;
4. Create chromosomes of initial population randomly;
5. For each program in population:
 - a) Express chromosome;
 - b) Execute program;
 - c) Evaluate fitness;
6. Verify stop condition;
7. Select programs;
8. Replicate selected programs to form the next population;
9. Modify chromosomes using genetic operators; and
10. Go to Step 5.

In Figure 10, the yellow curve is the target and the green curve is real data. The consistency between the real data and prediction is shown well in Figure 10. Some fitness indices used in this algorithm are shown in Table 3.

Coefficient of determination is 99 percent, which indicates the best data mining method among the used

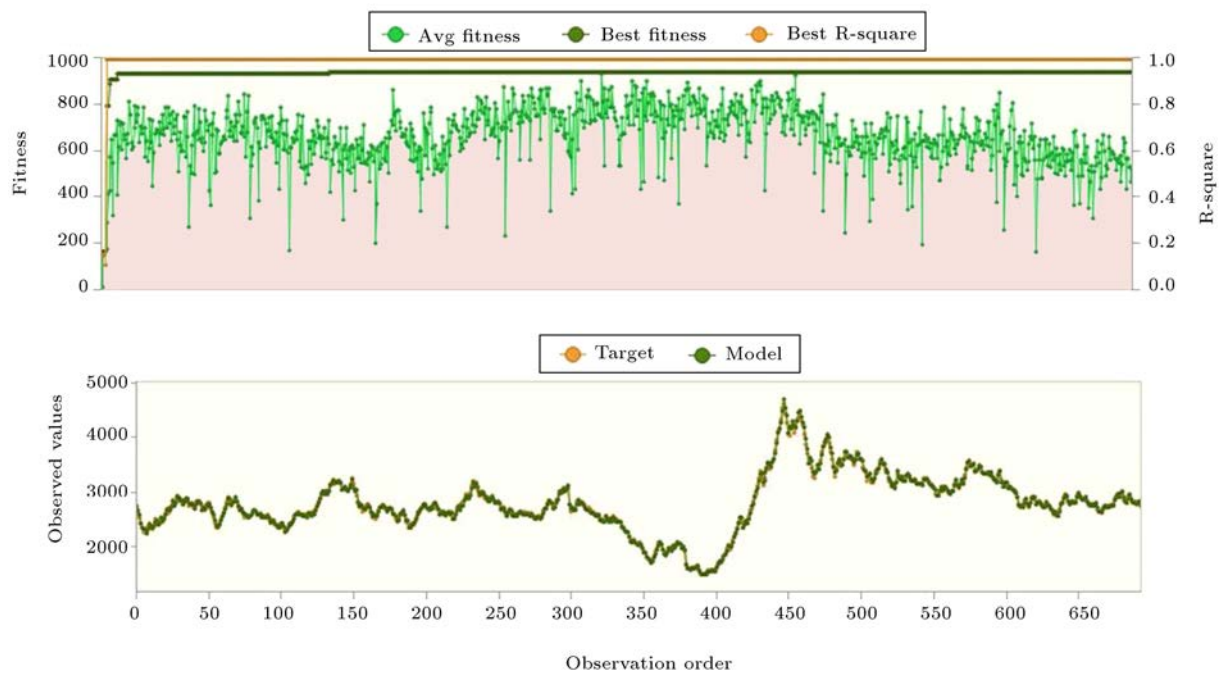


Figure 10. Gene Expression Programming (GEP) results.

Table 3. Fitness indexes.

Performance measures	Best of run-training	Best of run-validation
Fitness	939.114296655572	891.659462224849
R-square	0.9962	0.9883
Max fitness	1000	1000

Table 4. Best solution of Gene Expression Programming (GEP).

Generation	1000
Best fitness	939.114296655572
R-square	0.9962
Correl coefficient	0.9981

methods to assess the market value of stocks. The best solution is shown in Table 4.

GEP presents some tree graphs that demonstrate the prediction model. The prediction occurs based on these trees, as depicted in Figure 11.

The above-mentioned methods were used to predict the market value of stocks. The assessment was done for all 15 stocks. However, only the results of one stock are shown in this research. Based on the evaluation of different methods, the best predictor is GEP, which has a higher coefficient of determination and the lowest Mean Square Error (MSE). Coefficient of determination is the fitness index and MSE is

the error-index. Table 5 shows the accuracy of the prediction of each method for each stock.

The results of accuracy index show that GEP is the best method for assessing the market value under different scenarios. For all 15 stocks, GEP is the best data mining method. Therefore, 12 scenarios were employed to predict the market value of stocks in different economic situations. Table 6 shows the calculated data.

5. Numerical results

In this part, the proposed model was analyzed using the data of scenarios. The data used in this section include the shares of 15 companies in 12 scenarios. For each share, its fundamental value is also calculated. The scenarios are evaluated through GEP algorithm, which is proven to be the best data mining method for assessing the scenarios. Then, 12 different conditions are used for the oil price, exchange rate, and gold price. The proposed model is a linear programming that can be solved easily. This is another advantage of our model which can be solved using regular software packages.

In the first stage, the results of the model for a budget of one-million units and the expected growth estimated to be 10 to 30 percent in wealth are achieved. The results are given in Table 7.

The first column in Table 7 shows the expected profit margin. Figure 12 illustrates this margin. In fact, it is the percentage expected to be derived from the stock market. The second column shows the difference between the fundamental value and market value.

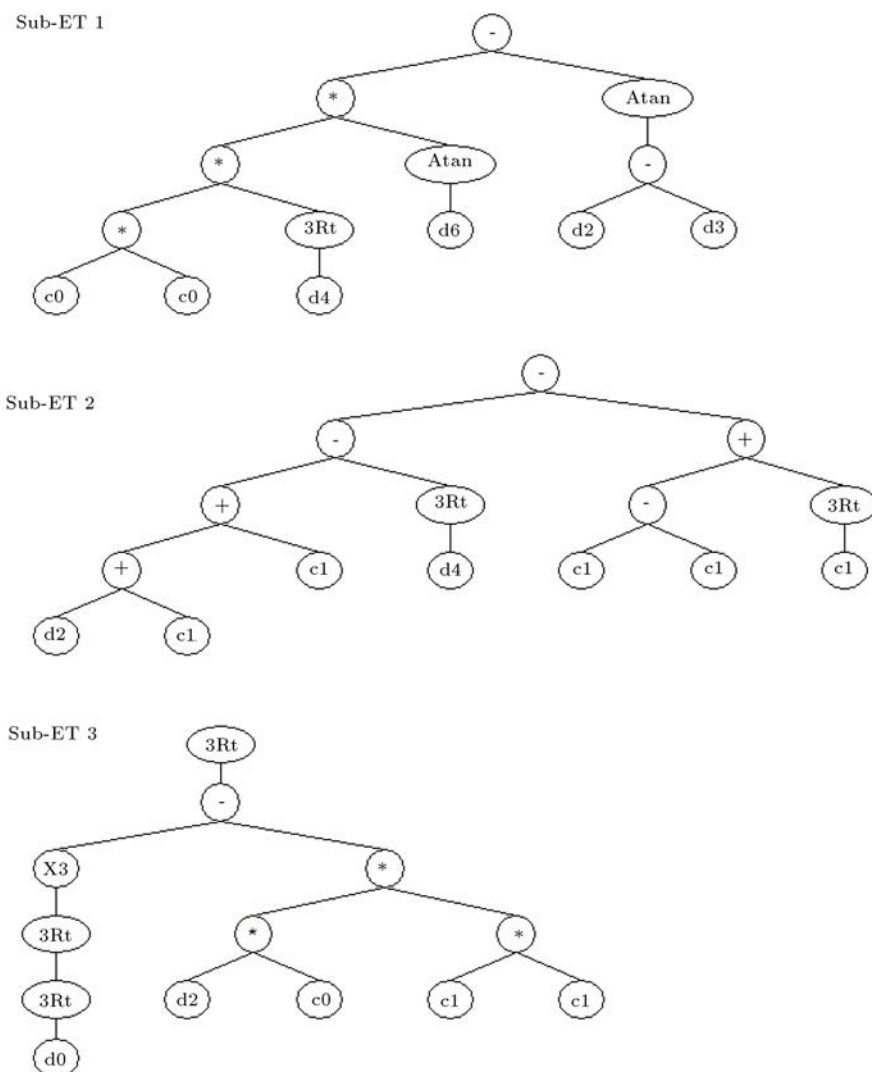


Figure 11. Tree graph of Gene Expression Programming (GEP) for prediction.

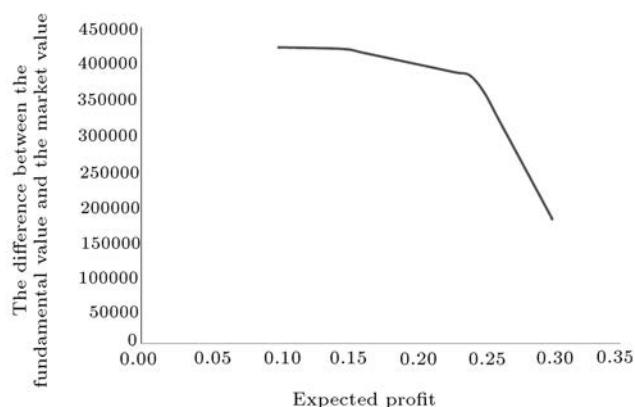


Figure 12. The difference between fundamental and market values.

The third column represents the fundamental value of the optimum stock portfolio, and finally the last column is the market value of the optimum portfolio.

In Section 2, it can be claimed that the proposed

model considers the difference between fundamental and market values in order to select the best stocks in the portfolio. Results show that upon increasing the expected profit level, the difference between the fundamental value of the portfolio and the market value of the optimum portfolio decreases. The optimum portfolio is based on the composition of the output that has a higher market value. In fact, if the expected profit is placed at its lowest, a combination of shares is chosen which has the greatest difference from the fundamental value. In fact, they have the highest gap with respect to their fundamental value. In this case, the risk is low. The optimum combination of stocks in a portfolio is the situation in which the shares are at their lowest price and the likelihood of a fall in prices is very low. In other words, the risk is very low. In such circumstances, the market value of stocks is expected to increase and reach its fundamental value. Column 3 of Table 7 indicates the status of the fundamental value of the optimum portfolio.

Table 5. Accuracy index of different methods.

Stock number	ANFIS	Decision tree	FDA	Random forest	GEP
1	0.93	0.29	0.3	0.48	0.99
2	0.72	0.31	0.27	0.46	0.999
3	0.64	0.38	0.31	0.49	0.965
4	0.96	0.4	0.33	0.463	0.976
5	0.88	0.33	0.291	0.51	0.982
6	0.45	0.33	0.31	0.42	0.96
7	0.87	0.34	0.28	0.36	0.96
8	0.72	0.32	0.33	0.48	0.98
9	0.57	0.31	0.36	0.441	0.95
10	0.74	0.42	0.39	0.53	0.98
11	0.52	0.37	0.33	0.47	0.97
12	0.8	0.35	0.36	0.476	0.97
13	0.88	0.39	0.38	0.48	0.99
14	0.87	0.36	0.33	0.34	0.99
15	0.81	0.37	0.41	0.39	0.98

Table 6. Calculated fundamental and market values.

Stocks no.	Market value (Sce. ^a 1)	Market value (Sce. 2)	Market value (Sce. 3)	Market value (Sce. 4)	Market value (Sce. 5)	Market value (Sce. 6)	Market value (Sce. 7)	Market value (Sce. 8)	Market value (Sce. 9)	Market value (Sce. 10)	Market value (Sce. 11)	Market value (Sce. 12)	Fun. ^b values
1	1164	1017	3326	1634	1936	1714	2222	2828	3746	1164	3287	1745	2443
2	2229	3874	3284	3555	2659	2753	3503	1487	3637	3874	2912	3379	1293
3	1032	1832	3486	2009	3484	1688	2100	1482	3619	3610	2972	2633	1064
4	2015	1862	2554	1621	1741	2585	1806	1863	2686	3231	3448	1196	1941
5	3775	2781	2033	3434	2933	3061	1266	2548	1589	3356	2143	2659	1360
6	2914	3914	2673	3063	3990	1143	2304	3018	1074	1575	1149	3195	3953
7	3222	2033	1916	2345	2686	3350	3811	1188	1225	2371	2607	1643	1563
8	3343	2163	1611	1094	1583	1319	3850	1870	3050	1551	2868	1174	1166
9	1685	1556	1552	2438	3892	1543	1274	3016	3814	2933	1900	2135	1759
10	1651	2723	3511	3855	2492	2040	1850	2072	1006	1556	1788	1414	1874
11	2961	2918	1234	3056	1580	3826	3746	1130	1265	1746	3654	2708	1092
12	3003	1204	1200	2557	1164	1110	1078	2194	1487	2919	1125	2038	3095
13	3566	1635	3039	3980	2176	3202	1088	3334	3033	3419	2982	2943	3137
14	1415	2447	3369	1887	2588	3759	2630	2662	1450	3302	1190	3128	1517
15	2781	2090	1501	3354	3005	1690	3913	1496	3337	3335	3514	2690	1791

^aSce.: Scenario; ^bFun.: Fundamental.

In Section 2, we claimed that the proposed model uses fundamental value as an important factor to make the optimum portfolio. Figure 13 points to this claim. According to Figure 13, this diagram exhibits a trend which is indicative of an increase in its value with a relatively gentle gradient and a decrease with a sudden steep slope. The reason behind this occurrence is the fact that it is feasible at a level of 27% in the range of 0.27 to 0.3; however, in the range of 27% to 30%, the model is unrecoverable; in fact, there is no optimal combination of stocks that yields profits of 28% and 29%. At these levels, the optimal level is 30% response. Of course, this trend may have different shapes for other data and other stocks. Generally, when the expected profit increases, it is expected that shares that are closer to their market value be selected. On the

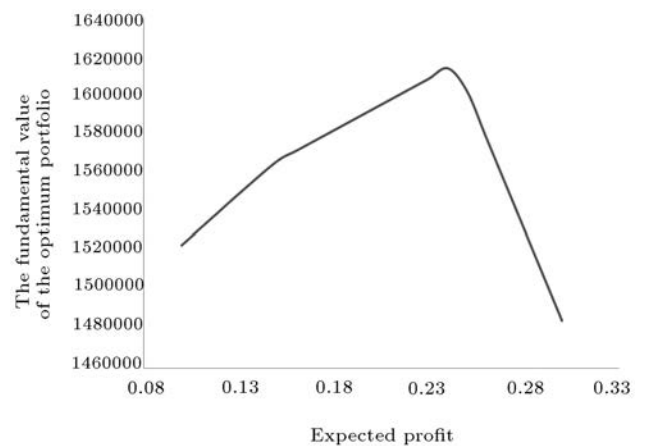
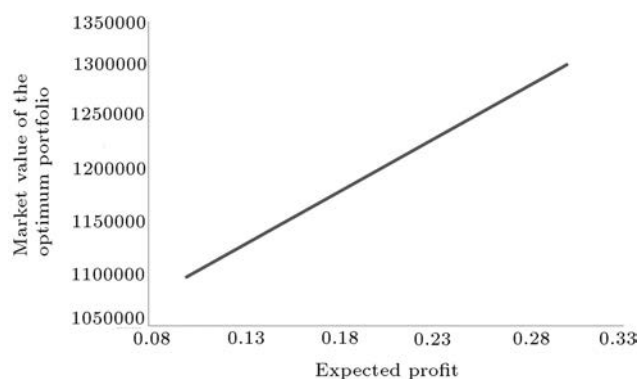
**Figure 13.** The fundamental value of the optimum portfolio.

Table 7. Summary of model results.

Expected profit percentage	The difference between the fundamental value and the market value	The inherent value of the optimal portfolio	Market value of the optimal portfolio
0.1	426263.5	1526263	1100000
0.11	425887.7	1535888	1110000
0.12	425511.9	1545512	1120000
0.13	425136.1	1555136	1130000
0.15	423579.4	1573579	1150000
0.16	419440.8	1579441	1160000
0.17	415302.2	1585302	1170000
0.18	411163.6	1591164	1180000
0.19	407025	1597025	1190000
0.2	402886.4	1602886	1200000
0.21	398747.8	1608748	1210000
0.22	394609.2	1614609	1220000
0.23	390470.7	1620471	1230000
0.24	386332.1	1626332	1240000
0.25	363431	1613431	1250000
0.26	327453	1587453	1260000
0.27	291475	1561475	1270000
0.3	183541	1483541	1300000

**Figure 14.** The market value of the optimum portfolio.

other hand, further increase in the distance between market value and fundamental value is expected when the market value is lower than the fundamental value. Hence, the model chooses a combination of stocks that are generally low in price and the gap between fundamental value and market value is high.

Figure 14 shows one of the most important advantages of the proposed model which is the raised expected profit. Hence, the market value of the optimum portfolio increases. The point is that the market value of the optimum portfolio at each expected rate of return is greater than the budget. Consequently, our model selects stocks based on both the fundamental point of view and market profitability. Column 4 of

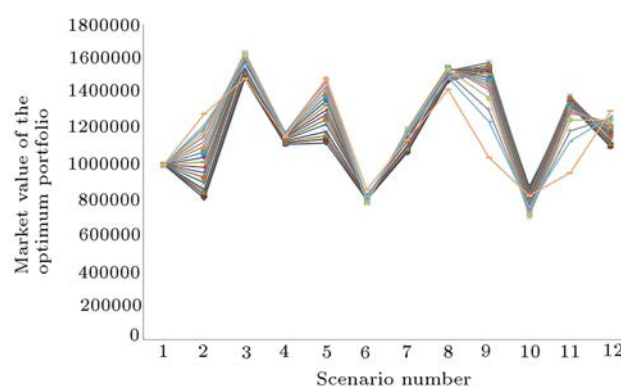
**Figure 15.** Market value trends of the optimum portfolio based on each scenario.

Table 7 shows the results of solving the model and market value over different scenarios. As mentioned before, 12 scenarios are used for analysis in this model. Table 8 shows the whole market values of the optimum portfolio based on different scenarios and different expected profits.

Figure 15 illustrates the trends of portfolio market values in different scenarios. This figure shows that based on each economic scenario, there exists the same pattern of portfolio market value. Each economic condition results in different market values for each stock. Consequently, the proposed model selects shares that are in the profit zone. Moreover, the proposed model

Table 8. Market values of optimum portfolios.

Scenarios	1	2	3	4	5	6	7	8	9	10	11	12
Expected rate 0.1	1000000	813993.2	1483648	1115305	1123296	805175.6	1068148	1467144	1526263	878191.1	1334465	1100000
Expected rate 0.11	1000000	830680	1502176	1120766	1144015	810224.2	1081471	1479130	1535888	871589.9	1345201	1110000
Expected rate 0.12	1000000	847366.7	1520705	1126227	1164734	815272.9	1094794	1491116	1545512	864988.7	1355938	1120000
Expected rate 0.13	1000000	864053.5	1539233	1131688	1185453	820321.6	1108118	1503102	1555136	858387.6	1366674	1130000
Expected rate 0.15	1000000	900018.1	1573579	1142309	1228118	828428.3	1133443	1525005	1568796	843538.8	1383613	1150000
Expected rate 0.16	1000000	928815	1579441	1146362	1254571	824173.1	1140595	1527324	1552299	829242.2	1373161	1160000
Expected rate 0.17	1000000	957611.8	1585302	1150415	1281023	819917.9	1147746	1529643	1535802	814945.6	1362708	1170000
Expected rate 0.18	1000000	986408.7	1591164	1154468	1307476	815662.7	1154897	1531962	1519305	800649	1352255	1180000
Expected rate 0.19	1000000	1015206	1597025	1158521	1333928	811407.5	1162048	1534281	1502808	786352.5	1341802	1190000
Expected rate 0.20	1000000	1044002	1602886	1162574	1360381	807152.3	1169200	1536600	1486311	772055.9	1331349	1200000
Expected rate 0.21	1000000	1072799	1608748	1166627	1386833	802897.1	1176351	1538919	1469815	757759.3	1320897	1210000
Expected rate 0.22	1000000	1101596	1614609	1170680	1413286	798641.9	1183502	1541238	1453318	743462.7	1310444	1220000
Expected rate 0.23	1000000	1130393	1620471	1174733	1439739	794386.7	1190653	1543557	1436821	729166.1	1299991	1230000
Expected rate 0.24	1000000	1159190	1626332	1178786	1466191	790131.4	1197805	1545876	1420324	714869.6	1289538	1240000
Expected rate 0.25	1000000	1183008	1613431	1177851	1477399	796082.4	1194429	1533524	1374345	721561.9	1250558	1250000
Expected rate 0.26	1000000	1203356	1587453	1173441	1477982	809146.7	1183717	1510948	1307817	742882.9	1191696	1260000
Expected rate 0.27	1000000	1223704	1561475	1169030	1478566	822211	1173005	1488372	1241290	764203.9	1132833	1270000
Expected rate 0.3	1000000	1284748	1483541	1155799	1480315	861404	1140868	1420644	1041707	828167	956245.8	1300000

selects stocks that are out of losing zone. Differences between optimum portfolios and fundamental factors have the same pattern in each scenario. In other words, regardless of economic conditions, the proposed model selects the best possible portfolio with the least bubble.

The results show that the increase and reduction in the market value of the optimum portfolio in scenarios are based on the difference between fundamental and market values. In fact, the model developed in this research attempts to minimize the average of the difference between fundamental and market values in the losing area under various scenarios.

6. Conclusion

In this research, a specific type of portfolio optimization model was examined and analyzed. The proposed model was first introduced with the objective of minimizing any deviation of market price and fundamental value. It was presented as a descriptive definition of risk. To develop a risk measure and a new portfolio selection model, this paper used the concept of market value and fundamental value of stocks. In the real world, many investors use the strategy of estimating the fundamental value and comparing it with the actual stock value to choose their optimum portfolio. In fact, investors first calculate the fundamental value of a share and then, by comparing that fundamental value with the market value, they make their decision whether to buy or sell the stock. Risk in this research was calculated as the difference between fundamental and market values. The proposed model uses the ratio of maximizing profits and minimizing losses. The model was transformed into linear programming by using mathematical

modeling techniques. The model was analyzed using real data from the Tehran Stock Exchange.

Subsequently, the proposed model was solved by using the data from the Tehran Stock Exchange for different conditions of expected profit. Finally, the results of the model were compared to the optimum portfolio profitability, the fundamental value of the portfolio as well as the optimum portfolio risk. The results represent specific trends in the risk and fundamental value of the portfolio for expected volatility in the profitability level.

Backed by the results of this research, investors might suggest making investments based on fundamental value under poor economic conditions and emotional price hikes in the stock market. Indeed, when the market value is considered to be emotionally high in the models of this research, the level of risk will greatly increase and the profitability of the portfolio as well as the fundamental value of the portfolio may fall. These results are not desirable for investors. When the market value is close to the fundamental value or even lower, the proposed model offers a portfolio that derives from market value fluctuations. On the other hand, the risk of reduction in prices will not cause losses for shareholders.

References

1. Markowitz, H. "Portfolio selection", *The Journal of Finance*, **7**, pp. 77–91 (1952).
2. Konno, H. and Yamazaki, H. "Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market", *Management Science*, **37**, pp. 519–531 (1991).
3. Rockafellar, R.T. and Uryasev, S. "Optimization of

- conditional value-at-risk", *Journal of Risk*, **2**, pp. 21–42 (2000).
4. Bernardi, M., Maruotti, A., and Petrella, L. "Multiple risk measures for multivariate dynamic heavy-tailed models", *Journal of Empirical Finance*, **43**, pp. 1–32 (2017).
 5. Kuzubaş, T.U., Ömercikoğlu, I., and Saltoğlu, B. "Network centrality measures and systemic risk: An application to the Turkish financial crisis", *Physica A: Statistical Mechanics and Its Applications*, **405**, pp. 203–215 (2014).
 6. Fu, T., Zhuang, X., Hui, Y., and Liu, J. "Convex risk measures based on generalized lower deviation and their applications", *International Review of Financial Analysis*, **52**, pp. 27–37 (2017).
 7. Gong, X. and Zhuang, X. "Measuring financial risk and portfolio reversion with time changed tempered stable Lévy processes", *The North American Journal of Economics and Finance*, **40**, pp. 148–159 (2017).
 8. Sorwar, G. and Dowd, K. "Estimating financial risk measures for options", *Journal of Banking & Finance*, **34**, pp. 1982–1992 (2010).
 9. Gao, J., Xiong, Y., and Li, D. "Dynamic mean-risk portfolio selection with multiple risk measures in continuous-time", *European Journal of Operational Research*, **249**, pp. 647–656 (2016).
 10. Beauchamp, J.P., Cesarini, D., and Johannesson, M. "The psychometric and empirical properties of measures of risk preferences", *Journal of Risk and Uncertainty*, **54**, pp. 203–237 (2017).
 11. Righi, M.B. and Borenstein, D. "A simulation comparison of risk measures for portfolio optimization", *Finance Research Letters*, **24**, pp. 105–112 (2018).
 12. Sawik, B. "Triple-objective models for portfolio optimisation with symmetric and percentile risk measures", *International Journal of Logistics Systems and Management*, **25**, pp. 96–107 (2016).
 13. Bernard, C., Denuit, M., and Vanduffel, S. "Measuring portfolio risk under partial dependence information", *Journal of Risk and Insurance*, **85**, pp. 843–863 (2018).
 14. Grechuk, B. and Zabaranin, M. "Inverse portfolio problem with coherent risk measures", *European Journal of Operational Research*, **249**, pp. 740–750 (2016).
 15. Ghahtarani, A. and Najafi, A.A. "Robust optimization in portfolio selection by m-MAD model approach", *Economic Computation & Economic Cybernetics Studies & Research*, **52**, pp. 279–291 (2018).
 16. Domeij, D. and Ellingsen, T. "Rational bubbles and public debt policy: A quantitative analysis", *Journal of Monetary Economics*, **96**, pp. 109–123 (2018).
 17. Barberis, N., Greenwood, R., Jin, L., and Shleifer, A. "Extrapolation and Bubbles (No. w21944)", In, *National Bureau of Economic Research* (2016).
 18. Lee, J.H. and Phillips, P.C. "Asset pricing with financial bubble risk", *Journal of Empirical Finance*, **38**, pp. 590–622 (2016).
 19. Bosi, S., Ha-Huy, T., Le Van, C., Pham, C.-T., and Pham, N.-S. "Financial bubbles and capital accumulation in altruistic economies", *Journal of Mathematical Economics*, **75**, pp. 125–139 (2018).
 20. Maynard, A. and Ren, D. "The finite sample power of long-horizon predictive tests in models with financial bubbles", *International Review of Financial Analysis*, **63**, pp. 418–430 (2019).
 21. Michaelides, P.G., Tsionas, E.G., and Konstantakis, K.N. "Non-linearities in financial bubbles: Theory and Bayesian evidence from S&P500", *Journal of Financial Stability*, **24**, pp. 61–70 (2016).
 22. Miao, J. and Wang, P. "Banking bubbles and financial crises", *Journal of Economic Theory*, **157**, pp. 763–792 (2015).
 23. Wigniolle, B. "Optimism, pessimism and financial bubbles", *Journal of Economic Dynamics and Control*, **41**, pp. 188–208 (2014).
 24. Harvey, D.I., Leybourne, S.J., Sollis, R., and Taylor, A.R. "Tests for explosive financial bubbles in the presence of non-stationary volatility", *Journal of Empirical Finance*, **38**, pp. 548–574 (2016).
 25. Kunieda, T. and Shibata, A. "Asset bubbles, economic growth, and a self-fulfilling financial crisis", *Journal of Monetary Economics*, **82**, pp. 70–84 (2016).
 26. Gong, B., Pan, D., and Shi, D. "New investors and bubbles: an analysis of the baosteel call warrant bubble", *Management Science*, **63**, pp. 2493–2508 (2016).
 27. Obayashi, Y., Protter, P., and Yang, S. "The lifetime of a financial bubble", *Mathematics and Financial Economics*, **11**, pp. 45–62 (2017).
 28. Fry, J. and Cheah, E.-T. "Negative bubbles and shocks in cryptocurrency markets", *International Review of Financial Analysis*, **47**, pp. 343–352 (2016).
 29. Sornette, D. and Cauwels, P., *Financial Bubbles: Mechanisms and Diagnostics*, Swiss Finance Institute Research Paper (2014).
 30. Hirota, S. and Sunder, S. "Price bubbles sans dividend anchors: Evidence from laboratory stock markets", *Behavioral Interactions, Markets, and Economic Dynamics*, Springer, pp. 357–395 (2016).
 31. Guastaroba, G., Mansini, R., Ogryczak, W., and Speranza, M.G. "Linear programming models based on Omega ratio for the enhanced index tracking problem", *European Journal of Operational Research*, **251**, pp. 938–956 (2016).
 32. Mansini, R., Ogryczak, W., and Speranza, M.G. "On LP solvable models for portfolio selection", *Informatica*, **14**, pp. 37–62 (2003).
 33. Ghahtarani, A., Sheikhmohammady, M., and Najafi, A.A. "Development of robust random variable for portfolio selection problem", *Industrial Engineering & Management Systems*, **17**, pp. 632–641 (2018).
 34. Barak, S., Arjmand, A., and Ortobelli, S. "Fusion of multiple diverse predictors in stock market", *Informa-*

tion Fusion, **36**, pp. 90–102 (2017).

35. Chong, E., Han, C., and Park, F.C. “Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies”, *Expert Systems with Applications*, **83**, pp. 187–205 (2017).
36. de Oliveira, F.A., Nobre, C.N., and Zarate, L.E. “Applying artificial neural networks to prediction of stock price and improvement of the directional prediction index-case study of PETR4, Petrobras, Brazil”, *Expert Systems with Applications*, **40**, pp. 7596–7606 (2013).
37. Krauss, C., Do, X.A., and Huck, N. “Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500”, *European Journal of Operational Research*, **259**, pp. 689–702 (2017).
38. Maroco, J., Silva, D., Rodrigues, A., Guerreiro, M., Santana, I., and Mendonça, A. de “Data mining methods in the prediction of Dementia: A real-data comparison of the accuracy, sensitivity and specificity of linear discriminant analysis, logistic regression, neural networks, support vector machines, classification trees and random forests”, *BMC Research Notes*, **4**, p. 299 (2011).
39. Weng, B., Ahmed, M.A., and Megahed, F.M. “Stock market one-day ahead movement prediction using disparate data sources”, *Expert Systems with Applications*, **79**, pp. 153–163 (2017).
40. Kohavi, R. and Quinlan, J.R. “Data mining tasks and methods: Classification: decision-tree discovery”, in *Handbook of Data Mining and Knowledge Discovery*, Oxford University Press, Inc., pp. 267–276 (2002).
41. Verikas, A., Gelzinis, A., and Bacauskiene, M. “Mining data with random forests: A survey and results of new tests”, *Pattern Recognition*, **44**, pp. 330–349 (2011).
42. Lin, C.-T., Lee, C.G., Lin, C.-T., and Lin, C., *Neural Fuzzy Systems: a Neuro-Fuzzy Synergism to Intelligent Systems*, Prentice Hall PTR Upper Saddle River NJ (1996).

Biographies

Alireza Ghahtarani studied Industrial Engineering at the Department of Industrial and Systems Engineering at Tarbiat Modares University, Iran. He is currently a Researcher and PhD student in Canada at

Dalhousie University. His main research area includes mathematical modeling and optimization techniques for portfolio selection problem.

Majid Sheikhmohammady received his PhD in Systems Design Engineering from University of Waterloo, Canada in January 2009. He was also a Post-Doctoral Fellow at the Department of Systems Design Engineering, University of Waterloo in 2009. He is currently a faculty member at the Department of Industrial and Systems Engineering, Tarbiat Modares University. Since January 2019, he has been an active member of Conflict Analysis Group at the Department of Systems Design Engineering, University of Waterloo, Canada. Dr. Sheikhmohammady’s primary research interests lie in the analysis of decisions, as the intersection of mathematics, engineering, and social sciences. His applications of game theory and related formal techniques include problems in conflict resolution, coalition formation, group decision and negotiation, fair division, and win-win solutions. He has pioneered the development of a methodology to model and analyze multilateral negotiations from a systems engineering perspective. More specifically, the new methodology called EMMN predicts the outcome of multilateral negotiations. Generally, Dr. Sheikhmohammady is interested in systems methodologies for innovatively solving real-world complex interdisciplinary problems that lie at the interface of society, technology, and the environment. In addition to his background in Industrial Engineering, he has carried out several research projects in other fields of study including natural resources, international relations, environmental studies, and defense science.

Amir Abbas Najafi received his BS degree in Industrial Engineering from Isfahan University of Technology in 1996, and his MS and PhD degrees in Industrial Engineering from Sharif University of Technology in 1998 and 2005, respectively. He is currently an Associate Professor of Industrial Engineering at K. N. Toosi University of Technology. His research interests include applied operations research, project scheduling and management, and portfolio selection models.