Dynamic portfolio insurance strategy: a robust machine learning approach*

Siamak Dehghanpour and Akbar Esfahanipour

Department of Industrial Engineering and Management Systems, Amirkabir University of Technology, Tehran, Iran

ABSTRACT
In this paper, we propose a robust genetic programming (RGP) model for a dynamic strategy of stock portfolio insurance. With portfolio insurance strategy, we divide the money in a risky asset and a risk-free asset. Our applied strategy is based on a constant proportion portfolio insurance strategy. For determining the amount for investing in the risky asset, a critical parameter is a constant risk multiplier that is calculated in our proposed model using RGP to reflect market dynamics. Our model includes four main steps: (1) Selecting the best stocks for constructing a portfolio using a density-based clustering strategy. (2) Enhancing the robustness of our proposed model with an application of the Adaptive Neuro-Fuzzy Inference Systems (ANFIS) for forecasting the future prices of the selected stocks. The findings show that using ANFIS, instead of a regular multi-layer artificial neural network improves the prediction accuracy and our model’s robustness. (3) Implementing the RGP model for calculating the risk multiplier. Risk variables are used to generate equation trees for calculating the risk multiplier. (4) Determining the optimal portfolio weights of the assets using the well-known Markowitz portfolio optimization model. Experimental results show that our proposed strategy outperforms our previous model.

ARTICLE HISTORY
Received 30 November 2017
Accepted 19 January 2018

KEYWORDS
Robust genetic programming (RGP); portfolio insurance strategy; machine learning; constant proportion portfolio insurance (CPPI)

1. Introduction
In the last decades, countless studies have tried to present a model for investors that not only would generate the best payoff, but would also reduce the risk of the investment as low as possible. The theory of mean-variance model (Markowitz, 1959) was the base of the most of these studies. Markowitz was the first person who suggested that investors should consider both return and risk of the investing in their decision-making.

One of the most important techniques in constructing a portfolio is to divide it into risk-free assets and risky assets. This is the original idea of the popular constant proportion portfolio insurance (CPPI) strategy. The CPPI strategy uses the investor’s risk preference to calculate the amount that should be invested in the risky assets and clearly, the rest...
invested in risk-free assets (Black & Jones, 1987). Since the risk multiplier in the CPPI strategy is basically determined by investor’s personal view, it may not work well in some unusual market movements. For instance, consider the situation where there is a sudden downward movement in the stock’s price, then with the conventional CPPI strategy, the insurance model cannot prevent the loss from the investment in risky assets in such a short time. As a result, a dynamic model should be developed that can adopt itself to such market conditions. Chen and Chang (2005) proposed a dynamic model strategy called dynamic proportion portfolio insurance that uses genetic programming (GP) to calculate the risk multiplier according to market condition. Then, they compared the performance of the proposed strategy with the conventional CPPI strategy which showed their model outperformed the CPPI strategy.

Chen and Liao (2007) introduced a goal-directed strategy based on the popular CPPI. They combined this goal-directed model with regular CPPI and eventually proposed piecewise linear goal-directed CPPI strategy. Balder, Brandl, and Mahayani (2009) proposed a model for portfolio insurance that considered the discrete-time trading in the CPPI strategy.

In the case of comparison of the CPPI strategy with other portfolio insurance strategies, Jiang, Ma, and An (2009) proposed a VaR-based portfolio insurance strategy which is called VBPI strategy, they compared the effectiveness of this model with the other strategies such as CPPI and the buy and hold strategy (Dichtl, Drobetz, & Wambach, 2017).

In our last paper, we presented a dynamic portfolio insurance strategy with the use of GP and we combined our model with an optimization tool to get the best possible result (Dehghanpour & Esfahanipour, 2017). In the current study, we implement a density-based clustering strategy as a new clustering method for choosing the best stocks. We implement the Adaptive Neuro-Fuzzy Inference Systems (ANFIS) for the prediction of the future prices of the selected stocks and we present a dynamic model with the use of robust genetic programming (RGP) for calculating the risk multiplier in the conventional CPPI strategy and finally, we combine the model with an optimization tool to get the best possible outcome from the investment.

The rest of this paper is organized as follows. Section 2 presents the literature review and research background in the concept of portfolio insurance, the Markowitz model, robust GP, ANFIS and clustering. Section 3 describes our proposed model and its algorithms. In Section 4, the data gathering and experimental results will be discussed and finally, in Section 5 conclusions are presented.

This paper is an extended version of our INISTA 2017 conference paper (Dehghanpour & Esfahanipour, 2017). Hereunder, we describe the new features of our proposed model regarding in comparison of the previous work:

- In the last paper, we chose stocks from New York stock exchange arbitrarily. However, in this paper, we propose a density-based clustering strategy for choosing the best possible stocks for constructing the portfolio. Findings show that implementation of the proposed clustering strategy reduces the risk of the investment and makes the portfolio more profitable than our previous work.
- In the last paper, for enhancing the robustness of our model in different market conditions, we applied artificial neural network (ANN) for the prediction of stocks prices. However, in this paper, we applied the ANFISs method instead of the ANN for the
stock price prediction. This leads to increase the prediction accuracy of the current model.

- Experimental results on 30 stocks from the Dow Jones Industrial Average (DJIA), rather than 5 stocks in our previous work, show that this paper model outperforms our previous work for portfolio insurance performance.

2. Theoretical background-related work

In this section, we provide a theoretical background for our model and a literature review of the most important studies related to our paper.

2.1. Constant proportion portfolio insurance

CPPI is one of the most popular strategies for portfolio insurance. It was introduced by Black and Jones (1987). Investing with this strategy contains risk-free assets (usually treasury bills) and risky assets, such as stocks or bonds. The amount that should be allocated to risky assets is calculated with the following equation:

\[ K = M \times (A - F), \]  

in which, \( A \) is the current value of a portfolio and \( F \) stands for the floor, which is the lowest value that is acceptable for the portfolio. The difference between current value and floor is known as the cushion. \( M \) is the risk multiplier that in this strategy is defined by investor’s personal view and is commonly between 3 and 6. The most important element of this equation is \( M \) (i.e. risk multiplier), the more risk averse the investor is, the less he or she chooses the \( M \)'s value. Then, the amount of \( K \) should be invested in risky assets and the rest, in the risk-free assets. This strategy is easy to understand and implement for investors, which is why it is so popular around the world.

2.2. The Markowitz model

In the classic portfolio optimization problem, the objective is to reduce the risk of the investment and/or to increase the return of the investing. Variance is one of the most commonplace risk measures used in this kind of problems. Markowitz (1959) introduced variance as a risk measure for the first time, his proposed model is described as a formulation of (2)–(5):

\[
\text{Min } \sigma^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} W_i \times W_j \times \sigma_{ij}
\]  

s.t.

\[
\sum_{i=1}^{n} W_i = 1,
\]  

\[
\sum_{j=1}^{n} W_i f_j \geq R,
\]
\[ W_i \geq 0, \quad (5) \]

where \( W_i \) is the weight of the \( i \)th stock and \( \sigma_{ij} \) is the covariance between the \( i \)th and \( j \)th stock, \( r_i \) is the return of each stock, and \( R \) is the investor’s expected return from the portfolio. In this model, the objective, shown in Equation (2), is to choose proper weights for each of the stocks in the portfolio in a way that with a fixed expected rate of return from that portfolio, we minimize the risk of the investment. Inequality (3) states that total weights of the stocks should be equal to 1. Inequality (4) ensures that the portfolio rate of return is equal or greater than the expected rate of return. And, inequality (5) states that short selling is not allowed, which means each stock in the portfolio must have a positive weight.

### 2.3. Genetic programming

GP is an extension of classic Genetic Algorithm (GA) with an important difference that offers a different kind of solution candidates. In GP, solution candidates are shown in form of trees. GP was first introduced by Koza (1990) and was then developed through the study of Koza (1994). The structure of GP is similar to that of GA: after initialization of population, random changes can occur either by changing the functions or the subtrees.

For applying a GP framework, the required elements are as follows: nodes of trees, population initialization, selection method, genetic operators such as crossover and mutation, fitness function and termination condition (Chen & Chang, 2005). These elements are briefly described next.

**Nodes:** the nodes in the tree structure are defined in two categories: functions and terminals. Functions set usually consists of different types of mathematical or computer functions. Terminals set include variables and constants according to the requirements of the problem on hand.

**Initialization:** GP starts with a randomly produced initial population.

**Selection:** selection is a process in which the selection of a parent for reproduction is determined. Usually, better parents are expected to produce better offsprings. The roulette wheel selection method is the most popular selection method.

**Crossover:** crossover is a process that generates offsprings from chosen parents by substituting their subtrees.

**Mutation:** mutation occurs in a way that a limited number of parents changes randomly.

**Termination condition:** the most common termination conditions are: exact number of reproduction, fitness target and fitness convergence (Banzhaf, Nordin, Keller, & Francone, 1998).

### 2.4. Clustering

Clustering is a data mining technique that is used to divide a set of data into similar and related groups. The task is to place data into clusters where the elements in one cluster have the most possible similarity and the clusters have the most possible differences (Kanungo et al., 2002; Nanda, Mahanty, & Tiwari, 2010; Singh, Nigam, Pal, & Mehrotra, 2014).
In recent decades, many studies have used clustering methods for solving their problems. Also, there are a number of studies in the literature that compared the results of different clustering methods. Chiu, Chen, Kuo, and He (2009) used K-means (KM) for intelligent market classification. Kim and Ahn (2008) applied GA in their KM in an online shopping market problem. Rana, Jasola, and Kumar (2010) applied cluster density (KD)-tree KM clustering algorithm that determines the number of clusters based upon the percentage change of close prices. Some other clustering related works in the literature apply ANNs for clustering and self-organizing maps (SOM) is one of the most popular strategies. Nanda, Mahanty, and Tiwari (2010) considered the KM, Fuzzy C-means (FCMs) and SOM in their study of applying clustering in Indian stock market. The works of Budayan, Dikmen, and Birgonul (2009); Delibasis, Mouravliansky, Matsopoulos, Nikita, and Marsh (1999) and Mingoti and Lima (2006) show the comparison of different clustering methods results. One of the best algorithms for clustering problems especially for stock markets is the well-known KM algorithm (Nanda, Mahanty, & Tiwari, 2010). KM clustering starts with one cluster that is centred at the mean of the data set. Then the original cluster splits into two and the mean of each data set is trained as the new cluster centres. This process continues until the pre-determined number of clusters is obtained. The FCM algorithm was first introduced by Bezdek (1981). In (FCM) clustering, any data-point can be assigned to more than just one cluster. Hence, as an indicator, the degree of membership is defined and calculated.

2.5. Robustness

Robust optimization, first introduced by Mulvey, Vanderbei, and Zenios (1995) has been adopted as an effective tool for optimal design and management in uncertain environments. In general, this term refers to the potential of a programme to maintain its applicability in the presence of external or internal disturbance (Mousavi, Esfahanipour, & Zarandi, 2014).

2.6. Adaptive Neuro-Fuzzy Inference Systems

Fuzzy logic (FL) and Fuzzy Inference Systems (FISs) were first proposed by Zadeh (1968). Working in the context of FL includes data and elements that can belong to more than just one set. There are three different types of FISs: (1) the Mamdani system that was introduced by Mamdani (1976), where the output must be defuzzified, (2) the Sugeno system in which the output is a real number (Sugeno, 1985; Takagi & Sugeno, 1985) and (3) the system of Tsukamoto that uses functions (Jang, 1993).

ANFIS structure consists of both ANN and FL (Atsalakis & Valavanis, 2009; Baneshi, 2015; Cheng, Wei, & Chen, 2009; Degmar & Tresi, 2011). There are a number of if-then rules in ANFIS calculation, as well. In recent years, soft computing methods have been used to help investors in predicting the future prices of stocks. ANNs, GAs and fuzzy systems are some of the intelligent systems that researchers have been working on in recent years. Trinkle (2006) implements ANFIS for forecasting the annual returns of three companies and then the prediction ability of ANFIS is compared with an autoregressive moving average model (Boyaciglou & Avci, 2010). Abbasi and Abouec (2008) study the trend of the stock price of Iran Khodro Corporation at Tehran Stock Exchange by training an ANFIS, and their findings show that trends in stock price can be forecasted with the high prediction accuracy. Chang and Liu
(2008) developed a Takagi–Sugeno–Kang (TSK)-type fuzzy rule-based system for forecasting Taiwan Stock Exchange (TSE) price deviation. This model successfully forecasts stock price variation with an accuracy close to 97.6% in TSE index and 98.08% in MediaTek (Boyacioglu & Avci, 2010). Esfahanipour and Aghamiri (2010) utilize a Neuro-Fuzzy Inference System based on a TSK-type fuzzy rule-based system that was developed for the stock price prediction. The TSK fuzzy model applies the technical index as the input variables, and the consequent part is a linear combination of the input variables. FCM clustering was implemented for identifying the number of rules (Boyacioglu & Avci, 2010). Esfahanipour and Mardani (2011) implemented an ANFIS system based on subtractive clustering for predicting stock prices in Tehran Stock Exchange, and the findings show that their strategy outperforms the ANN method. Atsalakis and Valavanis (2009) provide a thorough review of ANFIS-based methods in the last decades in the literature.

3. The proposed model

In this paper, we present a new model that selects the best possible stocks with the use of a clustering method. Then, with the implementation of an ANFIS system, we obtain a powerful prediction method that forecasts the future prices and trends in prices for chosen stocks. Afterwards, we construct a portfolio and, with an application of robust GP, we insure the portfolio based on popular CPPI strategy. Simultaneously, application of the well-known Markowitz model ensures that the portfolio has the best possible outcome. In summary, Figure 1 shows the steps in our study according to the four main steps in our proposed model.

3.1. Stock selecting using the clustering strategy

Clustering is a data mining technique that has been used in many studies recently and its goal is to divide a set of data into different clusters. Now we describe the use of clustering in our model.

We base our clustering method on the well-known FCM strategy and we develop this strategy to obtain better results. In our model, we use cluster density as an inherited features of the data set and implement a density-based algorithm for the clustering. We name this strategy Fuzzy C-Means Cluster Density-Based (FCM-CDB).

We consider data set \( X = \{x_1, x_2, \ldots, x_n\} \).

For each element \( x_i \), the element density is calculated with the following equation:

\[
z_i = \frac{1}{\sum_{j=1, j \neq i}^{n} d_{ij}}, \quad d_{ij} \leq e, \quad 1 \leq i \leq n, \quad (6)
\]

where, \( e \) denotes the effective radius, \( d_{ij} \) is the distance between \( i \)th and \( j \)th data, and \( n \) is the number of elements in the data set.

But to calculate the cluster density as a weighted linear regression of density of elements we use the following equation:

\[
\hat{z}_i = \frac{\sum_{j=1}^{n} \alpha_{ij} w_{ij} z_j}{\sum_{j=1}^{n} \alpha_{ij} w_{ij}}, \quad 1 \leq i \leq c, \quad (7)
\]

where, \( \alpha_{ij} \) and \( w_{ij} \) are the data label and weight of \( x_i \), respectively.
The improved distance indicator $\hat{d}_{ij}^2$ in this strategy is calculated via the following equation:

$$\hat{d}_{ij}^2 = \frac{\|x_j - v_i\|}{\hat{z}_i}, \quad 1 \leq i \leq c, \quad 1 \leq j \leq n,$$

(8)

where $\|x_j - v_i\|$ denotes the Euclidean norm of the distance between $x_j$ and $v_i$. 

**Figure 1.** Flowchart of our proposed model.
And, finally with the use of the Lagrange multiplier method in the objective function, we obtain the following equations:

$$u_{ij} = \frac{\hat{d}_{ij}^{2/(m-1)}}{\sum_{k=1}^{c} \frac{1}{d_{kj}^{2/(m-1)}}}, \quad (9)$$

$$v_i = \frac{\sum_{j=1}^{n} u_{ij}^m x_j}{\sum_{j=1}^{n} u_{ij}^m}, \quad 1 \leq i \leq c. \quad (10)$$

In which, $v_i$ refers to cluster centres, $u_{ij}$ is the fuzzy partition matrix and $m$ is the fuzzy index of the algorithm.

We use rate of return and risk as the two features of our clustering method. To determine the optimal number of clusters, the algorithm is primarily initiated with two clusters. Then, at each step, the number of clusters is increased and the fitness value for each iteration is stored. And based on the well-known Levenberg–Marquardt algorithm, if the fitness after three consecutive iterations maintains its decreasing trend, then we set our optimal number of clusters as three steps before the current number of clusters in the algorithm. In this study, the optimal number of clusters was found to be five. There are a number of indices to measure the validity of the quality of the clustering and optimal number of clusters. We use Xie and Beni’s index, Partition index, Davies-Bouldin index and Silhouette index for our fuzzy-based clustering (Nanda, Mahanty, & Tiwari, 2010). Xie and Beni’s index is a validity measure of clustering with the goal of calculating the ratio of the total variation between clusters and the separation of cluster. The minimum value of this index indicates the optimal number of the clusters (Nanda, Mahanty, & Tiwari, 2010; Xie & Beni, 1991). Partition index is the ratio of the total compactness and separation of the clusters. The lower the value of the partition index, the better the partitioning (Bensaid et al., 1996). For the Davies-Bouldin index, the lower the value of the index, the better the cluster structures (Kasturi, Acharya, & Ramanathan, 2003). And, the higher the Silhouette index, the better the quality of the clustering (Chen et al., 2002).

### 3.2. Enhancing the robustness of the model with ANFIS implementation

In this study, we use robust GP to make our model’s fitness insensitive to daily changes in stock’s price. The purpose of our robust model is to make adjustments in stocks historical prices.

For instance, if in one period a big sudden upward change occurs in the price of a stock, we cannot interpret that the stock is the better choice than the others. Thus, the use of robustness helps us to not base our decision on the noise in the data and data that could damage the accuracy of our decision. In this paper, we have implemented an ANFIS algorithm for the prediction of daily stock price. For evaluating the prediction accuracy of our ANFIS, we have used Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) and coefficient of determination ($R^2$) (Armstrong, 2001; Atsalakis & Valavanis, 2009; Chen & Kwon, 2012; Dehghanpour & Esfahanipour, 2017).

ANFIS is a combination of soft computing techniques that was first introduced by Jang (1993). ANFIS is a proper tool to apply in the prediction of time-series in nonlinear environments such as stock markets. For the prediction, identifying the data behaviour is an
important factor for achieving promising results (Esfahanipour & Mardani, 2011; Karray & Silva, 2004).

The first step of the ANFIS is to generate an FIS. In this paper, we propose an ANFIS application using all the three clustering methods for structure identification including subtractive clustering, FCM and grid partitioning. We use the method that has the least error among these three strategies. Use of these clustering methods in the implementation of the ANFIS is critical because Fuzzy systems are rule-based systems and, the cluster centres can be considered for developing a rule and the rule’s parameters would be identified (Esfahanipour & Mardani, 2011; Jang, 1993; Lotfi, Darini, & Karimi, 2016).

For implementing ANFIS in our study, we consider first-order TSK FIS with an input set of \( x \) and two fuzzy rules. The mathematical definition of input set and fuzzy rules are described in the following equations:

\[
x = (x_1, x_2) \in U,
\]

\[
\text{if } x_1 \text{ is } L_1^i \text{ and } x_2 \text{ is } L_2^i \text{ then } y^l = \gamma_0 + \gamma_1 x_1 + \gamma_2 x_2 \quad l = 1, 2.
\]

In the above equation, which stands for a TSK system, \( L_i^l \) is the linguistic value, \( \gamma_i^l \) represents constant coefficients and \( l \) stands for the number of the rules in the system.

Jang, (1993) precisely described the different layers of the ANFIS structure in his work. Hereunder, we briefly present the different layers of our ANFIS.

**Layer one:** this layer contains four adaptive nodes with two inputs and two Membership Functions (MFs) amalgamated to each one. The MF of \( L_i^l \) is \( \mu_{L_i^l}(x) \) which is considered as the output of a node. The parameters of the MF could be any differentiable function between 0 and 1. In this study, we use the Gaussian function that Esfahanipour and Mardani (2011) used in their study due to the similarity of the nature of our works.

**Layer two:** there are two nodes that generate the weight of the \( i \)th rule using the following equation:

\[
w^l = \prod_{i=1,2} \mu_{L_i^l}(x_i) = \mu_{L_1^l}(x_1) \mu_{L_2^l}(x_2).
\]

**Layer three:** there are two nodes that generate the normalized weight of each rule using the following equation:

\[
w^{-l} = \frac{w^l}{(w^1 + w^2)}.
\]

**Layer four:** there are two adaptive nodes that calculate \( y^l \) as the rule outputs based on the following equation:

\[
y^l = \gamma_0^l + \gamma_1^l x_1 + \gamma_2^l x_2.
\]

**Layer five:** the input is \( w^{-l}y^l \), that is sent from each node in the fourth layer. And, the sum of the inputs is calculated in this layer, which is equivalent to the output of the first-order TSK FIS. The following equation shows how to calculate layer five’s output:

\[
f(x) = \sum_{l} w^{-l} y^l = w^{-1} y^1 + w^{-2} y^2.
\]
We control the step size, step size decrease rate, the increase rate and the number of training epochs in implementation of the ANFIS using MATLAB. The ANFIS in this study is trained to generate a one-step-ahead the prediction of stock prices. After our clustering method chooses the best possible stocks for constructing the portfolio, then we use the ANFIS to predict future prices of the selected stocks. Inputs for the model are the current and previous price changes (Atsalakis & Valavanis, 2009). The ANFIS controller is presented in the following equation:

\[
y(k + 1) = f(y(k), y(k - 1), y(k - 2), y(k - 3), y(k - 4)).
\]  

(17)

The training is based on an ANFIS that maps \([y(k), y(k - 1), y(k - 2), y(k - 3), y(k - 4)]\) to the predicted stock price change at \(y(k + 1)\) (Atsalakis & Valavanis, 2009; Baneshi, 2015; Wang & Chan, 2007).

### 3.3. Insuring the portfolio using RGP

As mentioned in the previous section, the risk multiplier in the CPPI strategy is determined by investor’s personal view and is constant during the insurance period (Branke, 1998). This method could lead to a serious loss for the investors in specific market conditions such as high volatility in short period of times. Hence, in their model, Chen and Chang (2005) proposed a dynamic model in which the risk multiplier is calculated using RGP to ensure that the multiplier would not be constant all the time. Thus, in any market conditions with risk variables being considered for RGP, an adapted risk multiplier will be calculated. Consider the following equation for this new strategy:

\[
K = M(\tau)(A - F).
\]  

(18)

In summary, our proposed model determines the risk multiplier \((M)\) by \(\tau\) which is a set of constants and risk variables. The value of \(M\) is obtained through an RGP model. Then the amount of money which should be invested in risky assets (i.e. \(K\) in Equation (18)) should be allocated to different stocks. The terminal set of our RGP is risk variables, including both market volatility factors and technical indicators. The beta in the capital asset pricing model is usually considered to be constant; thus, in this study, in addition to exchange rate, risk-free rate and market index are used instead of the beta as market volatility factors. We also use relative strength index (RSI), stochastic oscillator (STO) and Williams %R (%R) as our three technical indicators. Hence, for terminal set of our RGP, we consider market index, exchange rate (US dollar to Euro), Williams %R (%R), RSI, STO and T-bill (10-year). These risk variables and their abbreviations are listed in Table 1. Note that since the risk variables have different ranges of values, they are normalized to the same range before being used in RGP.

For the function set in this paper, we use the four commonly used arithmetic operators as follows (Berutich, López, Luna, & Quintana, 2016):

\[+ , - , / , * .\]
The fitness function for evaluating our model is calculated by the following equation:

\[ f = \frac{r + 1}{\sigma + 1}, \]  

(19)

In the above formula, \( r \) is the portfolio rate of return and \( \sigma \) is the portfolio standard deviation. To avoid encountering negative fitness values, the rate of return is added by one in the numerator. To prevent a zero denominator, the standard deviation is also added by one.

3.4. Determining portfolio weights with the application of the Markowitz model

In the previous section, we determined how much of the investment should go into risky assets or the stocks. Then, for optimizing the constructed portfolio, we use the Markowitz model. Thus, our proposed model is an RGP for portfolio insurance. Simultaneously, we run the Markowitz model to obtain the best possible portfolio weights that ensure the constructed portfolio has optimal results.

4. Experimental results

In our model, we have used historical data for the 30 stocks from the DJIA from 1 January 2012 to 30 December 2016. In total, we have 1258 days for each of the stocks. We run our clustering strategy based on the two most important features: return and risk of the stocks.

The algorithm has divided our data set into six clusters and after ranking the clusters according to maximum return and minimum risk, the best cluster consists of six stocks including CSCO, GE, INTC, KO, PFE and VZ. Note that we have formed our Markowitz model according to Equations (20–23). Also, we have considered the expected return \( (R) \) in this model to be 15%:

Min \( \sigma^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} W_i * W_j * \sigma_{ij} \)  

(20)

s.t.

\[ \sum_{i=1}^{n} W_i = 1, \]  

(21)

\[ \sum_{i=1}^{n} W_i r_i \geq R, \]  

(22)

| Variable                           | Abbreviation          |
|------------------------------------|-----------------------|
| Market index                       | NYSE (NYSE composite) |
| Exchange rate                      | US2Euro (US dollar to Euro) |
| Williams %R                        | (R %)                 |
| Relative strength index            | RSI                   |
| Risk-free rate                     | T-bill (10-year)      |
| Stochastic oscillator              | STO-D                 |
In Table 2, the name of the selected stocks and their symbols are presented. To evaluate our clustering method (FCM-CDB), we have used different indices. The result of these indices is presented in Table 3.

As can be seen in Table 3, almost all the indices indicate that the optimal number of clusters is 6. The only exception is the Silhouette index which obtains 7 as the optimal number of clusters. However, this is justifiable since the Silhouette index uses the hard partitioning method and may not be very countable for fuzzy-based clustering strategies.

For implementing the ANFIS, we have run it with three clustering methods including FCM, subtractive clustering and Grid partitioning. Then, the outcomes of the three clustering approaches are compared with each other to choose the best strategy. The data set was divided into the training set (75% of total data-points) and the test set (25% of total data). We have used MATLAB for running these three clustering methods. In this study, we use the Taguchi’s method for parameters setting for each FIS generation approach (Mast, 2004).

For comparing the accuracy of these three clustering approaches in implementing the ANFIS, we have applied three commonly used performance measures: RMSE, MAPE and coefficient of determination ($R^2$). The following equations describe the calculation of these measures:

$$ RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^{T} (a_t - o_t)^2}, $$

$$ MAPE = \frac{(100/T) \sum_{t=1}^{T} |(a_t - o_t)/a_t|,} $$

### Table 2. The selected stocks for forming the portfolio.

| Symbol | Company                        |
|--------|--------------------------------|
| CSCO   | Cisco Systems, Inc.            |
| INTC   | Intel Corporation              |
| GE     | General Electric Company       |
| KO     | The Coca-Cola Company          |
| PFE    | Pfizer Inc.                   |
| VZ     | Verizon Communications Inc.    |

### Table 3. Evaluation indices result for FCM-CDB clustering strategy.

| Indices           | 2       | 3       | 4       | 5       | 6       | 7       | 8       | 9       | 10      |
|-------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| Davies-Bouldin    | 1.538   | 1.449   | 1.653   | 1.347   | 1.297   | *       | 1.399   | 1.464   | 1.602   | 1.587   |
| Silhouette        | 0.186   | 0.143   | 0.113   | 0.187   | 0.260   | *       | 0.273   | 0.149   | 0.099   | 0.081   |
| Partition index   | 0.021   | 0.019   | 0.010   | 0.008   | 0.005   | *       | 0.006   | 0.011   | 0.019   | 0.019   |
| Xie and Beni’s    | 2.531   | 2.322   | 1.756   | 1.098   | 0.849   | *       | 0.973   | 0.988   | 1.017   | 1.220   |

*The optimal number of clusters based on each measurement index.*
\[ R^2 = 1 - \left[ \frac{\sum_{t=1}^{T} (a_t - o_t)^2}{\sum_{t=1}^{T} (a_t - \bar{a})^2} \right], \quad (26) \]

In the above equations, \( T \) is the number of testing patterns, \( a_t \) is the actual output and \( o_t \) is the predicted output of the model. Also, \( \bar{a} \) is the average of all actual outputs related to the \( T \) testing patterns (Esfahanipour & Mardani, 2011). In Table 4, we compare the effectiveness of the three clustering strategy on performance of the ANFIS model.

The implementation of the ANFIS using the FCM clustering strategy has the best result among the three clustering methods. Figure 2 shows the performance of the FCM clustering strategy on the ANFIS model.

The training and testing periods of our GP are described in Table 5. From 2012 to the end of 2016, similar to the work of Chen and Chang (2005), we divided these 5 years to 10 six-month periods and we assigned them for our training and testing periods as shown in Table 5. We divided our data set into training and testing periods in order to evaluate the fitness of our robust GP in different periods.

In Table 6, the parameters of our robust GP are described (Dehghanpour & Esfahanipour, 2017).

Now we present the fitness values for testing periods of our model and we compare them with fitness values obtained from our previous work (Dehghanpour & Esfahanipour, 2017). In order to avoid different results due to the random nature of the GP, we run the model 30 times. The reported results in this section are the average fitness values of these 30 runs. The result of this comparison shows that this new model’s fitness values are better than our previous work. This implies that the proposed model yields superior results in comparison with our previous model. While, in our previous work (GP1), we implemented a multi-layer neural network for improving the robustness of the decision model, in this work (GP2) however, ANFIS and clustering techniques are used for data processing and forecasting. The comparison of the fitness values is presented in Table 7. The difference between the two models is the application of ANFIS for enhancing the robustness of our model, instead of the multi-layer ANN that we used in our last paper (Dehghanpour & Esfahanipour, 2017). Since we had to run the GP1 model with the same stocks that the clustering strategy of the GP2 model had selected, this fitness comparison may not thoroughly show the improvement of our new model. Because, one of the most significant contributions of this paper is the implementation of the proposed clustering strategy for stock selecting. However, in comparing the fitness of the two models (GP1 and GP2), we could not emphasize on this important feature of our new work.

Figures 3 and 4 show the graphic comparison between the new model and the previous one.

| Clustering method          | RMSE  | MAPE  | \( R^2 \) |
|---------------------------|-------|-------|-----------|
| FCM                       | 4.9234\(^a\) | 0.4143\(^a\) | 0.9982\(^a\) |
| Grid partitioning         | 4.9287 | 0.5226 | 0.9851    |
| Subtractive clustering    | 4.9262 | 0.4158 | 0.9950    |

\(^a\)The clustering method with the best performance.
Note that GP2 is GP using ANFIS and GP1 is GP fitness using ANN that we presented in our previous work.

We implement $t$-test on fitness values of our optimized robust GP model and our previous model. The outcome of the $t$-test shows that our model has the better performance.
This means that combining optimization methods with the robust GP model with the use of ANFIS and implementation of our clustering is a better choice for investors. The $t$-test is calculated using MATLAB $t$-test function, with 95% confidence level, and the null hypothesis is that the mean fitness values of our proposed model are less than the mean fitness values of our previous model. The outcome indicates that the $t$-test does reject this null hypothesis.

Table 6. Parameters of the robust GP.

| Parameter            | Value |
|----------------------|-------|
| Population size      | 500   |
| Maximum tree depth   | 8     |
| Selection method     | Roulette wheel |
| Crossover method     | Subtree crossover |
| Mutation method      | Subtree replacement |
| Number of iterations | 10    |
| Crossover rate       | 0.7   |
| Mutation rate        | 0.3   |

Table 7. Comparison of fitness values of the two models. Each model has been run 30 times. The reported values are the average fitness values of 30 times.

| Testing period | GP fitness using ANN (GP1) | GP fitness using ANFIS (GP2) |
|----------------|-----------------------------|------------------------------|
| 2012.07-12     | 1.232                       | 1.373                        |
| 2013.01-06     | 1.107                       | 1.146                        |
| 2013.07-12     | 1.152                       | 1.213                        |
| 2014.01-06     | 1.173                       | 1.299                        |
| 2014.07-12     | 1.096                       | 1.148                        |
| 2015.01-06     | 1.089                       | 1.188                        |
| 2015.07-12     | 1.127                       | 1.235                        |
| 2016.01-06     | 1.091                       | 1.175                        |
| 2016.07-12     | 1.152                       | 1.275                        |

Figure 3. The comparison between our new strategy and the previous model.
5. Conclusion

In this paper, we propose an RGP model for portfolio insurance. The proposed model can determine the risk multiplier of the CPPI strategy using GP to consider different market situations. To be more practical, we use Markowitz’s portfolio optimization theory to determine the best portfolio weights for investors.

Our strategy is to generate the risk multiplier of the popular CPPI strategy with GP by using some risk variables in order to use the best risk multiplier according to different market conditions. Then with this dynamic strategy, we determine how much of the investor’s money should be invested in risky assets and how much of it in risk-free assets. We implement robust GP, in order to make sure that our model is more effective and reliable for the future stock price prediction. The next step that we used is to run the well-known Markowitz optimization model in order to determine the portfolio weights to be invested in risky assets. In our new model, we use the density-based clustering strategy according to risk and return features that are the most important traits in investment-related decision-making. Also, instead of the multi-layer ANN that we used in our previous work for predicting the future trends in stocks prices, we use the ANFIS strategy for the prediction of the future prices. The findings show that developing our previous model by adding the implementation of the FCM-CDB clustering strategy and ANFIS significantly improves the performance.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Siamak Dehghanpour received his B.Sc in Industrial Engineering and his M.Sc in Financial Engineering both from the Department of Industrial Engineering and Management Systems at Amirkabir University of Technology, Tehran, Iran. His research interests are in the field of investment analysis, portfolio insurance and corporate finance. His recent publication has appeared in the proceedings of IEEE INISTA 2017.
Akbar Esfahanipour received his B.Sc in industrial engineering from Amirkabir University of Technology, Tehran, Iran in 1995. His M.Sc and PhD degrees are in industrial engineering from Tarbiat Modares University, Tehran, Iran in 1998 and 2004, respectively. He is currently an Assistant Professor at Industrial Engineering Department, Amirkabir University of Technology. Prior to joining Amirkabir University, he was a Postdoctoral Fellow at DeGroote School of Business, McMaster University, Hamilton, ON, Canada. He has worked as a senior consultant for over 10 years in his specialized field of expertise in various industries. His research interests are in the areas of forecasting in financial markets, application of soft computing methods in financial decision-making, and analysis of financial risks. Dr. Esfahanipour has published his research articles on financial decision-making in journals such as European Journal of Operational Research, Journal of Management Information Systems, Expert Systems with Applications, Knowledge-Based Systems, Applied Soft Computing, and Resources Policy.

ORCID

Akbar Esfahanipour  http://orcid.org/0000-0003-3222-5186

References

Abbasi, E., & Abouec, A. (2008). Stock price forecast by using neuro-fuzzy inference system. Proceedings of World Academy of Science, Engineering and Technology, 36, 320–323.
Armstrong, J. S. (2001). Evaluating forecasting methods. In J. S. Armstrong (Ed.), Principles of forecasting: A handbook for researchers and practitioners (pp. 7923–7930). New York, NY: Kluwer.
Atsalakis, G. S., & Valavanis, K. P. (2009). Forecasting stock market short-term trends using a neuro-fuzzy based methodology. Expert Systems with Applications, 36(3), 10696–10707.
Balder, S., Brandl, M., & Mahayani, A. (2009). Effectiveness of CPPI strategies under discrete-time trading. Journal of Economics Dynamics & Control, 33, 204–220.
Baneshi, M. (2015). Using ANFIS and neural networks to predict the volume percentage of matrix and fluid. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 37, 2247–2258. Taylor & Francis Online.
Banzhaf, W., Nordin, P., Keller, R. E., & Francone, F. D. (1998). Genetic programming – an introduction: On the automatic evolution of computer programs and its applications. San Francisco: Morgan Kaufmann.
Bensaid, A. M., Hall, L. O., Bezdek, J. C., Clarke, L. P., Silbiger, M. L., Arrington, J. A., & Murtagh, R. F. (1996). Validity guided (re)clustering with applications to image segmentation. IEEE Transactions on Fuzzy Systems, 4, 112–123.
Berutich, J. M., López, F., Luna, F., & Quintana, D. (2016). Robust technical trading strategies using GP for algorithmic portfolio selection. Expert Systems With Applications, 46, 307–315.
Bezdek, J. C. (1981). Pattern recognition with fuzzy objective function algorithms. New York, NY: Plenum Press.
Black, F., & Jones, R. (1987). Simplifying portfolio insurance. Journal of Portfolio Management, 14(1), 48–51.
Boyacioglu, M. A., & Avci, D. (2010). An adaptive network based fuzzy inference system (ANFIS) for the prediction of stock market return: The case of Istanbul Stock Exchange. Journal of Expert Systems with Applications, 37(12), 7908–7912.
Branke, J. (1998). Creating robust solutions by means of evolutionary algorithms. In Proceedings of the 5th international conference on parallel problem solving from nature PPSN V (pp. 119–128). London: Springer-Verlag.
Budayan, C., Dikmen, I., & Birgonul, M. T. (2009). Comparing the performance of traditional cluster analysis, self-organizing maps and fuzzy C-means method for strategic grouping. Expert Systems with Applications, 36(9), 11772–11781.
Chang, P.-C., & Liu, C. H. (2008). A TSK type fuzzy rule based system for stock price prediction. Expert Systems with Applications, 34(1), 135–144.
Chen, J. S., & Chang, C. L. (2005). Dynamical proportion portfolio insurance with genetic programming. In L. Wang, K. Chen, & Y. S. Ong (Eds.), Advances in Natural Computation. ICNC 2005. Lecture Notes in Computer Science (Vol. 3611). Berlin: Springer.

Chen, G., Jaradat, S. A., Banerjee, N., Tanaka, T. S., Ko, M. S. H., & Zhang, M. Q. (2002). Evaluation and comparison of clustering algorithms in analyzing ES cell gene expression data. Statistica Sinica, 12, 241–262.

Chen, C., & Kwon, R. H. (2012). Robust portfolio selection for index tracking. Computers & Operations Research, 39, 829–837.

Chen, J. S., & Liao, B. P. (2007). Piecewise nonlinear goal-directed CPPI strategy. Journal of Expert Systems with Applications, 33, 857–869.

Cheng, C. H., Wei, L. Y., & Chen, Y. S. (2009). Fusion ANFIS models based on multi-stock volatility causality for TAIEX forecasting. Journal of Neurocomputing, 72(16–18), 3462–3468.

Chiu, C. Y., Chen, Y. F., Kuo, I. T., & He, C. K. (2009). An intelligent market segmentation system using k means and particle swarm optimization. Expert Systems with Applications, 36, 4558–4565.

Degmar, B., & Tresi, J. (2011). Financial forecasting using neural network China-USA. Business Review, 10(3), 169–175.

Dehghanpour, S., & Esfahanipour, A. (2017). A robust genetic programming model for a dynamic portfolio insurance strategy. IEEE international conference on INNovations in Intelligent SysTemS and Applications (INISTA), Gdynia, Poland.

Delibasis, K. K., Mouravliansky, N., Matsopoulos, G. K., Nikita, K. S., & Marsh, A. (1999). MR functional cardiac imaging: Segmentation, measurement and WWW based visualisation of 4D data. Future Generation Computer Systems, 15(2), 185–193.

Dichtl, H., Drobeta, W., & Wambach, M. (2017). A bootstrap-based comparison of portfolio insurance strategies. Taylor and Francis. Available at SSRN.

Jiang, C., Ma, Y., & An, Y. (2009). The effectiveness of the VaR-based portfolio insurance strategy: An empirical analysis. International Review of Financial Analysis, 18(4), 185–197.

Karray, F., & Silva, C. D. (2004). Soft computing and intelligent system design. Vancouver: Addison Wesley.

Kasturi, J., Acharya, R., & Ramanathan, M. (2003). An information theoretic approach for analyzing temporal patterns of gene expression. Bioinformatics, 19, 449–458.

Kim, K. j., & Ahn, H. (2008). A recommender system using GA K-means clustering in an online shopping market. Expert Systems with Applications, 34, 1200–1209.

Koza, J. R. (1990). Genetic programming: A paradigm for genetically breeding populations of computer programs to solve problems (Tech. Rep. STAN-CS-90-1314). Computer Science Department, Stanford University.

Koza, J. R. (1994). Introduction to genetic programming. In K. J. Kinnear (Ed.), Advances in genetic programming (pp. 21–42). Cambridge, MA: The MIT Press.

Lotfi, E., Darini, M., & Karimi-T, M. R. (2016). Cost estimation using ANFIS. The Engineering Economist, 61(2), 144–154.

Mamdani, E. H. (1976). Advances in the linguistic synthesis of fuzzy controllers. International Journal of Man-Machine Studies, 8, 669–678.

Markowitz, H. (1959). Portfolio selection. New York, NY: Wiley.
Mast, J. (2004). A methodology comparison of three strategies for quality improvement. *Journal of Quality Reliability, 21*(2), 198–213.

Mingoti, S. A., & Lima, J. O. (2006). Comparing SOM neural network with fuzzy cmeans, K-means and traditional hierarchical clustering algorithms. *European Journal of Operational Research, 174*, 1742–1759.

Mousavi, S., Esfahanipour, A., & Zarandi, M. H. F. (2014). A novel approach to dynamic portfolio trading system using multitree genetic programming. *Knowledge-based Systems, 66*, 68–81.

Mulvey, J. M., Vanderbei, R. J., & Zenios, S. A. (1995). Robust optimization of large-scale systems. *Operations Research, 43*, 264–281.

Nanda, S., Mahanty, B., & Tiwari, M. K. (2010). Clustering Indian stock market data for portfolio management. *Expert System with Applications, 37*, 8793–8798.

Rana, S., Jasola, S., & Kumar, R. (2010). A hybrid sequential approach for data clustering using K-means and particle swarm optimization algorithm. *International Journal of Engineering and Science Technology, 2*(6), 167–176.

Singh, K. K., Nigam, M. J., Pal, K., & Mehrotra, A. (2014). A fuzzy Kohonen local information C-means clustering for remote sensing imagery. *IETE Technical Review, 31*(1), 75–81.

Sugeno, M. (1985). *Industrial application of fuzzy control*. New York, NY: Elsevier Science. ISBN: 0444878297.

Takagi, T., & Sugeno, M. (1985). Fuzzy identification of systems and its application to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics, 15*, 116–132.

Trinkle, B. S. (2006). Forecasting annual excess stock returns via an adaptive network-based fuzzy inference system. *Intelligent Systems in Accounting, Finance and Management, 13*(3), 165–177.

Wang, J., & Chan, C. H. (2007). Stock market trading rule discovery using pattern recognition and technical analysis. *Expert Systems with Applications, 33*, 304–315.

Xie, X. L., & Beni, G. A. (1991). Validity measure for fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 13*(8), 841–847.

Zadeh, L. A. (1968). Fuzzy algorithm. *Information and Control, 12*, 94–102.