JOURNAL OF APPLIED INTELLIGENT SYSTEMS & INFORMATION SCIENCES
VOL. 4, ISSUE 1, PP. 32-42, FEBRUARY 2023.
Available at: **www.journal.research.fanap.com**
DOI: https://doi.org/10.22034/JAISIS.2023.377702.1059

A MULTI-OBJECTIVE MODEL FOR PORTFOLIO SELECTION WITH BUDGET CONSTRAINTS AND A DEEP LEARNING-BASED STOCK PRICE PREDICTION

Serveh Kakaei [1,2,*]

[1] *Department of Industrial Engineering, University of Kurdistan, Sanandaj, IRAN,*
[2] *Research Center, FANAP Co., Tehran, IRAN.*

ABSTRACT

The stock market is a key pivot in every growing and thriving economy, and every investment in the market is aimed at maximizing profit and minimizing associated risk. For this purpose, a multi-objective optimization model with the goals of reducing risk, and increasing profit are provided considering the budget constraint. The presented model helps investment companies to make more beneficial stock portfolios. In addition, in this paper, a deep learning model Long Short-Term Memory (LSTM) is used to predict all days of the next year's stock price according to historical data. The dataset consists of the TSE index traded in Tehran Stock Exchange financial market. First, among the 50 active industries in the TSE stock market, 10 highly profitable industries are selected. The stock price of the companies in these industries is predicted according to the data of the last 15 years. The results obtained from 170 stocks and 10 industries show that the automobile, investment, and pharmaceutical materials and products industries were the best industries for investment in 2023 in the Tehran stock exchange.

***KEYWORDS*:** Multi-objective optimization, Long short-term memory neural network, Stock market prediction, Portfolio selection

1. INTRODUCTION

The stock market is an essential component of the nation's economy, where most of the capital is exchanged around the world. Therefore, the stock market's performance has a significant influence on the national economy. It plays a crucial role in attracting and directing distributed liquidity and savings into optimal paths. Investors and speculators in the stock market aim to make better profits from the analysis of market information. The mathematical framework for assembling a portfolio of stocks in a way that maximizes the expected return for a given level of risk is called mean-variance analysis or modern portfolio theory (Markowitz, 1952). The Markowitz model is thus a theoretical framework for the analysis of risk and return and their inter-relationships. He used statistical analysis for the measurement of risk and mathematical programming for the selection of stocks in a portfolio in an efficient manner. His framework led to the concept of efficient portfolios. An efficient portfolio is expected to yield the highest return for a given level of risk or the lowest risk for a given level of return (Markowitz, 1952). Risk and profit are two aspects of investment considered by investors. The expected return may vary depending on the assumptions. The risk index is measured by the variance of the distribution around the mean, which is in statistical terms called variance and covariance. The qualification of risk and the need for optimization of return with the lowest risk are the contributions of Markowitz. Investment portfolio management is one of the most basic conditions for success in the stock market. In the stock market, investors

* *Corresponding Author, Email: s.kakaei@uok.ac.ir*

RECEIVED: 21 DECEMBER, 2022; **ACCEPTED**: 19 FEBRUARY, 2023; **PUBLISHED ONLINE**: 01 MARCH, 2023

want to invest in stocks that will bring them the most profit and the least risk. Since the goal is to choose the best stocks for the next year, the price of all the stocks should be predicted. Stock price prediction has been done with different methods until now. Time series data as stock markets data is a sequence of data points in chronological order to analyze past data and make future predictions. Time series prediction problems are a difficult type of predictive modeling problem. A powerful type of neural network designed to handle sequence dependence is called a recurrent neural network. The Long Short-Term Memory network or LSTM network is a type of recurrent neural network used in deep learning because very large architectures can be successfully trained. For this purpose, daily historical stock data for the past 15 years will be used, and stock price prediction will be done through the LSTM network. The LSTM network inputs will be four features open, high, low, and closing prices to predict the stock price for each stock.
The main contributions of this research are as follows:

- A multi-objective mathematical model will be presented to select the best industries in the Tehran Stock Exchange to bring the most profit and the least risk with the budget constraint.
- A special kind of recurrent neural network LSTM would be used to predict the stock price.

The proposed paper is structured in the following sections: Section 2, the literature review gives an overview of several research works in stock price prediction and mathematical models for portfolios. Section 3 describes the methodology used in the paper to select stocks and predict the stock price. The proposed mathematical portfolio model and the LSTM model for predicting the price of stocks as well as the results obtained are provided in Section 4. The result of a case study is shown in Section 5. Finally, some conclusions and outlines of the future scope of work are explained in Section 6.

2. LITERATURE REVIEW

Markowitz's modern portfolio theory has made a new paradigm of portfolio selection for investors to form a portfolio with the highest expected return at a given level of risk tolerance (Markowitz, 1952). Lots of efforts have been made by experts to solve and expand Markowitz's model. These attempts regarding the limitations of a factual market have been made to make his model more practical. In 2001, Konno and Wijayanayake proposed an algorithm for his portfolio optimization problems regarding transaction costs and minimum transaction lots (Konno and Wijayanayake, 2001). In the first years of the 20th century, some studies threw more light on the capabilities of heuristic algorithms in portfolio selection problems. Lin et al. (2001) considered the multi-objective genetic algorithm for the portfolio selection problem. Carama and Schyns (2003) utilized a Simulated Annealing to solve Markowitz's model with cardinality and Turn Over constraints. Soleimani et al. (2009) presented a portfolio selection model based on Markowitz's mean-variance portfolio selection model. The proposed model covered minimum transaction lots, cardinality constraints, and regards market capitalization. In order to solve this mixed-integer nonlinear programming model a genetic algorithm is used. Lim et al. Lim et al. (2014) developed a way of using DEA cross-efficiency evaluation in portfolio selection under the mean-variance framework. They applied the proposed approach to stock portfolio selection in the Korean stock market for a 9-year sample period from 2002 to 2011.
On the other hand, in the financial world, the forecasting of stock price gains significant attraction. For the growth of shareholders in a company's stock, stock price prediction has a great consideration to increase the interest of speculators in investing money in the company. The successful prediction of a future cost could return noteworthy benefits. Many methods have been proposed to predict stock prices, among which linear models like autoregressive (AR), moving average (MA), autoregressive and moving average (ARMA), and autoregressive integrated moving average (ARIMA) and Non-linear models like autoregressive conditionally heteroscedastic (ARCH), Generalized autoregressive conditionally heteroscedastic (GARCH), and threshold autoregressive (TAR) methods can be mentioned. These methods are the analysis of time series data. Time series data can be defined as a chronological sequence of observations for a selected variable (Selvin et al., 2017). In this paper, the variable is the stock price. Shi et al. (2012) proposed a hybrid method of ARMA and back propagation neural network (BPNN) and Markov model to forecast the stock price. ARMA and BPNN solve the linear and nonlinear components of the stock price series respectively. Lee et al. (2007) compared the forecasting performance of a neural network model and a time-series (SARIMA) model in the Korean stock exchange. They investigate whether the back-propagation neural network model outperforms the seasonal

autoregressive integrated moving average model in forecasting Korea Composite Stock Price Index and its return. Wijaya et al. (2010) compared the stock forecasting result of ANTM (PT Aneka Tam bang) using artificial neural networks and ARIMA in Indonesia stocks.

Nowadays, the most significant challenge in the stock market is to predict stock prices. The stock price data represents financial time series data which becomes more difficult to predict due to its characteristics and dynamic nature. Recently, with the expansion of the applications of artificial intelligence and machine learning algorithms, they can also be used in predicting the selection of stock portfolios and stock prices. Support Vector Machines (SVM) and Artificial Neural Networks (ANN) are widely used for the prediction of stock prices and their movements. Every algorithm has its way of learning patterns and predicting. Artificial Neural Network (ANN) is a popular method that also incorporates technical analysis for making predictions in financial markets. Pahwa and Agarwal (2019) used linear regression, the supervised learning approach to predict stock prices. The proposed research work outlines the entire process of using a given dataset to forecast the closing value, by studying the GOOGL stock and extracting approximately 14 years of data. In Persio and Honchar's (2016) work different artificial neural network approaches, namely Multi-layer Perceptron (MLP), Convolutional Neural Networks (CNN), and Recurrent Neural Networks (RNN), have been applied to the forecasting of stock market price movements for the S&P500 index. Kim and Han (2000) proposed a genetic algorithms approach to feature discretization and the determination of connection weights for artificial neural networks to predict the stock price index. Hiransha et al. (2018) applied four types of neural networks named MLP, RNN, LSTM, and CNN to test the performance of predicting two stock market prices. The dataset is taken from the NSE market including three different sectors of Automobile, Banking, and IT. The NYSE market contains the Finance and Petroleum sectors. ARIMA model was used in this paper as the comparison between linear and nonlinear models. The price in the stock market has already reflected all known information which means that the financial market is informationally efficient, and the price is influenced by news or events. Ding et al. (2015) made use of the work in Natural Language Processing and applied this to stock market prediction. In this article, the team proposed a novel deep neural network for event-driven stock market prediction. The experimental results show that the proposed model could achieve almost 6% improvements compared with state-of-the-art baseline methods.

Mehtab and Sen (2022) recently proposed another approach to stock price and movement prediction using convolutional neural networks on a multivariate time series. The predictive model proposed by the authors exploits the learning ability of a CNN with a walk-forward validation ability to realize a high level of accuracy in forecasting the future NIFTY index values, and their movement patterns. Three different architectures of CNN are proposed by the authors that differ in the number of variables used in forecasting the number of sub-models used in the overall system, and the size of the input data for training the models. The experimental results indicated that the CNN-based multivariate forecasting model was highly accurate in predicting the movement of NIFTY index values with a weekly forecast horizon. The use of LSTM networks in stock price prediction has also been proposed (Sen and Mehtab, 2021). Vargas et al. (2017) proposed a sophisticated deep learning model for the detection and analysis of complex patterns and interactions among stock price data. The deep learning framework consists of a CNN and an RNN model. The results show that CNN is more effective in catching the semantic meaning from the text inputs to the model, while RNN is superior in understanding the context information and modeling the temporal characteristics for the stock price prediction. Sunny et al. (2020) proposed two models including LSTM and Bi-Directional LSTM (BI-LSTM) to predict the stock price. Also, Ghosh et al. (2019) proposed a framework using the LSTM model to predict stock prices in Indian stock markets. Many researchers have tried using historical stock prices as the basis for time series analysis to forecast future stock prices. Recently different neural network models, and evolutionary algorithms were being applied for stock prediction with success.

Table 1- Research on portfolio selection and stock price prediction

Researchers	Year	Stock Price Prediction	Mathematical Model
Kim and Han	2000	ANN	
Lin et al.	2001		Mean-Variance
Konno and Wijayanayake	2001		Mean-Variance
Carama and Schyns	2003		Mean-Variance
Lee et al.	2007	ANN	
Soleimani et al.	2009		Mean-Variance

Wijaya et al.	2010	ANN, ARIMA	
Shi et al.	2012	ARMA, BPNN	
Lim et al.	2014		DEA- Mean-Variance
Ding et al.	2015	NLP	
Persio and Honchar	2016	MLP, CNN, RNN	
Selvin et al.	2017	AR, MA, ARIMA, ARCH, GARCH	
Vargas et al.	2017	CNN, RNN	
Hiransha et al.	2018	MLP, RNN, LSTM, ARIMA	
Pahwa and Agarwal	2019	Linear Regression	
Gosh et al.	2019	LSTM	
Sunny et al.	2020	LSTM	
Mehtab and Sen	2021	LSTM	
Mehtab and Sen	2022	CNN	

Table 1 shows the stock price forecasting methods used by researchers from 2000 to 2022. In these years, some researchers have used mathematical models to select stocks, but none of them have simultaneously used new methods to predict stock prices and choose the best stocks portfolio in the next years.

For this reason, we use LSTM to predict the stock price for the next year. The reason for using LSTM is that it is one of the best methods for predicting time series. Then based on the predicted prices, we select the best stock portfolio for the next year.

3. Methodology

3.1. Markowitz Model

The modern theory of the portfolio mean-variance model is derived from the pioneering research works of Harry Markowitz (Markowitz, 1952). After the publication of Markowitz's paper on portfolio selection, a revolution was created in the field of portfolio optimization. In 1959, Markowitz published his and many researchers' achievements in the field of portfolio selection and model solution methods in a book (Boonjing and Boongasame, 2017). The published model is as follows:

$$\text{Min} \quad \sigma_{R_p}^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} x_i\, x_j Cov(\overline{r_i}, \overline{r_j}) \tag{1}$$

S.t:

$$\overline{r_p} = E(r_p) = \sum_{i=1}^{N} \overline{r_i} x_i = R \tag{2}$$

$$\sum_{i=1}^{N} x_i = 1 \tag{3}$$

$$x_i \geq 0 \;\; , \;\; \forall i \in \{1,\ldots,\mathrm{N}\} \tag{4}$$

This model is quadratic. As seen in the model, the objective function is the correlations of the component stocks, for all stock pairs (r*i*, r*j*). The expected return for a set of stocks is equal to the weighted average of the expected return of all the stocks in the portfolio (equation (2)). The variable x_i represents the weight assigned to the stock *i* in the portfolio. The parameter N shows the total number of stocks in the portfolio. In equation (5), $E(R_P)$ is the expected rate of the portfolio, $E(R_i)$ is the expected rate of stock *i*, and W_i is the ratio of the value of the portfolio that is invested in each stock.

$$E(R_p) = \sum_{i=1}^{N} E(R_i)\, W_i \tag{5}$$

Estimating the combined risk of a portfolio is not as easy as calculating its return. Risk is the possibility of future fluctuations in the rate of return. Various metrics are used to determine fluctuations. The most important of these measures is the average absolute value of deviations, variance, semi-variance, beta index, and value at risk. It should be noted that in this research, the variance measure is used to calculate risk.

3.2. Long-Short-Term Memory Network

Long-Short Term Memory was first proposed by Hochreiter and Schmidhuber (1997) to solve the problem of vanishing and exploding gradients by proposing a different architectural approach to RNNs. This is achieved by protecting its hidden activation using gates between each of its transaction points with the rest of its layer. The hidden activation that is protected is called the cell state. The protection of the cell state is undertaken by the three LSTM gates, namely, the forget, input, and output gates. During the forward pass, the first gate that acts upon the cell is the forget gate. It determines which of the cell's activations are forgotten and by how much. It achieves this by multiplying all the cell's elements with a vector $f_t \epsilon\ (0,1)^{m_h}$. If the forget gate emits a value close to zero, the corresponding element in the cell state will be wiped and set to zero, whereas if it outputs a value close to one, then the cell will fully retain the value of that element. The next gate to act upon the cell is the input gate, which determines what portion of new information will be added to the protected state. This happens in combination with the calculation of a new candidate cell state c_t'. Similar to the forget gate, the input gate $i_t \epsilon\ (0,1)^{m_h}$ is multiplied by the candidate state c_t' and added to the cell state. This prevents unnecessary additions to the cell state.

Finally, there is the output gate, which is important for the backward pass. It determines which parts of the cell state need to be propagated forward and be included in the output of the network. The forward flow of information across the described gates is shown in Fig.1 The following equations describe the behavior of the LSTM model:

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \tag{6}$$

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \tag{7}$$

$$c_t' = tanh(W_{hc}h_{t-1} + W_{xc}x_t + b_c) \tag{8}$$

$$c_t = f_t c_{t-1} + i_t c_t' \tag{9}$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_i) \tag{10}$$

$$h_t = o_t\, tanh(c_t) \tag{11}$$

where f_t, i_t, and o_t are the activations of the input, forget, and output gates at timestep t, which control how much of the input and the previous state will be considered, and how much of the cell state will be included in the hidden activation of the network. The protected cell activation at time step t is denoted by c_t, whereas h_t is the activation that will be given to the next layer. The notations of the weight matrices $W_{(..)}$ are explicitly denoted with subscripts of their transform. $b_{(.)}$ denotes the bias vector of each gate. W_{xf} and W_{hf} are the matrices that transform the input x_t and hidden state h_{t-1}, respectively, to the forget gate dimension, W_{xi}, and W_{hi} transform the input and hidden state to the input gate dimension, and so on. There are simplified notations combining the input and hidden state into a single matrix, thus having a single matrix per gate W_f, W_i, W_c, W_o (Tsantekidi et al., 2022).

Due to its capability of storing past information, LSTM is very useful in predicting stock prices. This is because the prediction of a future stock price is dependent on the previous prices. Also, the results of past studies show that the LSTM model is more accurate than other neural networks in predicting stock prices.

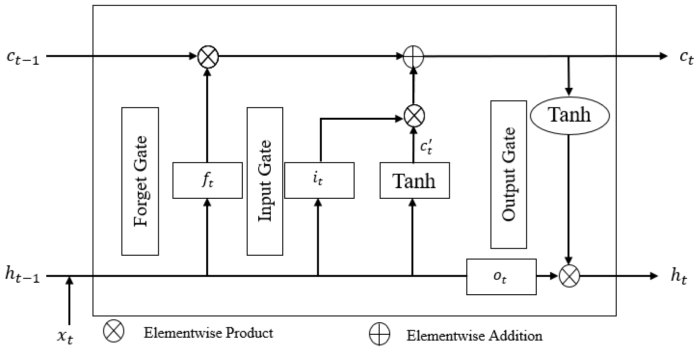

Fig. 1. LSTM computational graph

4. THE PROPOSED MODEL

Markowitz's model assumes that the main objective of the investor is to minimize risk at a certain level of return. Normally, the decision maker considers the rate of return to be constant and then minimizes the risk of the stock portfolio with return limits. But in the multi-objective model presented in this paper, risk minimization and return maximization are simultaneously performed in the objective function. Portfolio risk includes not only the risk of each stock but also the covariance between both stocks. The importance of the covariance term may be equal to the importance of combining the risk of each stock. Therefore, when one stock is added to the portfolio, the importance of the average covariance between the new stock and other stocks in the portfolio is more than the risk of the new stock. Considering that this multi-objective model has been validated in Kakaei and Janahgoshai's (2016) thesis and has good results, a similar model has been used in this paper. Therefore, the proposed model can be formulated as:

$$Min \quad \sum_{i=1}^{N_1}\sum_{j=1}^{N_1}\sum_{k=1}^{N_2} x_{ik}.x_{jk}.Cov(\overline{r_{ik}},\overline{r_{jk}}) \tag{12}$$

$$Max \quad \sum_{i=1}^{N_1}\sum_{k=1}^{N_2} r_{ik}.x_{ik} \tag{13}$$

S.t:

$$\sum_{i=1}^{N_1}\sum_{k=1}^{N_2} c_{ik}.x_{ik}.p_{ik} \le B \tag{14}$$

$$0 \le x_{ik} \le 1, \quad \lambda_k \in \{0,1\} \text{ binary}$$

The model indices are shown as i, j, and k are stocks and industry, respectively. The value of x_{ik} shows the weighting of stock i in industry k, $Cov(\overline{r_{ik}},\overline{r_{jk}})$ is the covariance of the returns on the two stock i, and j in industry k. Parameter r_{ik} is the return on stock i in industry k. Equation (14) limits the maximum amount of budget that investors can invest in the stock portfolio. The parameter B shows the amount of budget available for investment and parameter p_{ik} and c_{ik} shows the number of stocks offered in the stock market for stock i in industry k and the price of stock i in industry k, respectively.

In order to solve the proposed model, it is necessary to convert the multi-objective function into a single-objective function. One of these methods for solving multi-objective problems is the LP-metric method, which can be used to convert a multi-objective function into a single-objective. In this method, a point is considered an ideal solution, then we tried to identify the closest point of the solution space to this ideal point. For this reason, the solution obtained from this approach is called a compromise solution. To obtain a compromise solution, different compromise functions are used to measure the distance. In this paper, the compromise function used is as follows:

$$Min \quad FP = W_1 * \frac{F_1 - F_{1_{Optimal}}}{F_{1_{Optimal}}} - W_2 * \frac{F_2 - F_{2_{Optimal}}}{F_{2_{Optimal}}} \tag{15}$$

S.t:

$$W_1 + W_2 = 1 \qquad 0 \leq W_1 \leq 1, \qquad 0 \leq W_2 \leq 1 \tag{16}$$

The parameters w_1 and w_2 are the weights assigned to the objective function 1 and 2, respectively. Also, F_1 represents the first objective function, and the optimal value of the objective function is shown with $F_{1Optimal}$. Therefore, the proposed model by converting the multi-objective function to a single objective one will be as follows:

$$Min \quad FP = W_1 * \frac{\sum_{i=1}^{N_1}\sum_{j=1}^{N_1}\sum_{k=1}^{N_2} x_{ik}.x_{jk}.Cov(\overline{r_{ik}},\overline{r_{jk}}) - F_{1_{Optimal}}}{F_{1_{Optimal}}} - W_2 * \frac{\sum_{i=1}^{N_1}\sum_{k=1}^{N_2} r_{ik}.x_{ik} - F_{2_{Optimal}}}{F_{2_{Optimal}}} \tag{17}$$

S.t:

$$\sum_{i=1}^{N_1}\sum_{k=1}^{N_2} c_{ik}.x_{ik}.p_{ik} \leq B$$

$0 \leq x_{ik} \leq 1, \ \lambda_k \in \{0,1\}$ binary

$$W_1 + W_2 = 1 \qquad 0 \leq W_1 \leq 1, \qquad 0 \leq W_2 \leq 1$$

4.1. Experimental Setup

For the LSTM model, the four features open, high, low, and closing prices are used to predict stock prices. Ten years time period has been collected in these datasets starting from March 2008 to November 2022. The data is divided into two parts: The training part which is used to train the model and update the model parameters, and the other part which would be used for the testing part so that we would use the data to optimize the model for data forecasting.

In the LSTM model, the unit of the hidden layer is designed to be 140. One fully connected layer is chosen as the output layer. The sigmoid function and tanh function are used in the LSTM cell. For all methods, the time step is set to twenty, and the learning rate is 0.01. We choose Mean Square Error (MSE) loss as our loss function and use the Adam algorithm to optimize the loss function. The batch size is set to 64. All the input data will be shuffled before training. The experiments of all methods were implemented on Keras. In order to choose the stock portfolio with the presented model for investment in 2023, the average stock price predicted by LSTM in 2023 is considered for each stock. Since the presented model is a mixed integer nonlinear program, the model is solved in GAMS software with a DICOPT solver.

5. Results

In this study, the proposed model is implemented based on the extracted data from 10 industries and 170 existing stocks, and the results are analyzed. The rate of return on investment in different industries is shown in Table 2. Parameter B, which indicates the maximum number of budgets that the investor can invest, is set at 100000 billion rials in Iran. According to the assumptions, the results obtained from the GAMS and solver DICOPT. It shows that the automobile industry is the best industry for investment. The choice of this industry is not far from expected due to the political and economic conditions of Iran. The automobile industry has always been a profitable industry for investing. The results obtained for selecting only one industry among all industries (k=1) are shown in Table 3. The weight of each stock and the amount of each stock for the purchase of each stock is shown in Table 3.

Table 2- Stock return rate based on different industries

Industry	Rate of return			
	Min	Max	Mean	Variance
Automobile	0.01	0.3	0.2	0.0073
Basic metals	0.06	0.2	0.11	0.007
Rubber and plastic	0.05	0.3	0.14	0.010
Sugar	0.06	0.15	0.1	0.0058
Investments	0.11	0.16	0.14	0.0053
Ceramic tile	0.14	0.27	0.2	0.0074
Chemical products	0.05	0.28	0.14	0.009
Cement, lime, and plaster	0.05	0.23	0.16	0.0072
Aggregation of real estate	0.07	0.2	0.12	0.006
Pharmaceutical materials and products	0.1	0.34	0.18	0.0067

Table 3- The results of the proposed model with the assumption of (k=1).

Stocks	Weight	Price of prediction	Amount of stock	Budget of stock
Khesapa	0.134	2500	20234000000	50585000000000
Khazin	0.038	3100	152000000	471200000000
Khodro	0.032	3095	9632000000	29811040000000
Khekomak	0.091	22080	29120000	642969600000
Khefanar	0.173	15230	294100000	4479143000000
Khenasir	0.044	13020	74800000	973896000000
Khetoor	0.152	6090	304000000	1851360000000
khepooyesh	0.336	19050	1492146000	2172512140000

It should be noted that the following 35 companies are a subcategory of the automobile industry. Among these 35 active companies in this industry, 8 companies are known as the best companies in terms of objective functions, and investing in these stocks will bring the highest profit and the lowest risk for the investors. The high cost of importing cars and increasing sanctions in Iran has led to an increase in the price of cars in Iran. Car buyers are more willing to buy domestic cars, even at a high price. Therefore, investing in these companies has brought good profits for the investors.

In the second case, it is assumed that investors tend to invest in two different industries that have the highest profit and the lowest cost, and they want to invest 8 billion in both industries. Hence, we set the parameter k equal to 2 in the presented model (k=2). The results show that the automobile industry and investment industries were selected as the best industries for investment. The results obtained from solving the proposed model with the assumption of k = 2 are presented in Tables (4) and (5). Table (4) shows the results of the proposed model for the automobile industry and Table (5) shows the results related to the investment industry.

Table 4- The results of the proposed model with the assumption of (k=2).

Stocks	Weight	Price of prediction	Amount of stock	Budget of stock
Khesapa	0.064	2500	9664000000	24160000000000
Khazin	0.2	3100	800000000	2480000000000
Khodro	0.033	3095	9933000000	30742635000000
Khekomak	0.113	22080	36160000	798412800000

The automobile industry has 35 subsidiary companies, and the investment industry has 15 subsidiary companies. Among these 50 companies, 6 companies have been selected for investment. As stated earlier, the reason for choosing the automobile industry is explained in detail in the previous assumption. On the other hand, the reason for choosing the investment industry is the economic conditions and the price of coins, gold, and currency market in Iran and the world market, which has caused investment companies to invest in the gold, coin, and currency market. As a result, it is clear that the mentioned conditions led to investment in such markets and made a huge profit for the shareholders and investors of this particular industry.

Table 5- The results of the proposed model with the assumption of (k=2).

Stocks	Weight	Price of prediction	Amount of stock	Budget of stock
Vesapa	0.27	6200	4320000000	26784000000000
Vebahman	0.32	3128	4480000000	14013440000000

If the investors want to invest in three industries (k=3) in this case, the three industries of automobile, investment, and pharmaceutical materials and products are selected for investment. The pharmaceutical industry in Iran is one of the least risky industries in the stock market. Due to the coronavirus, this industry has been one of the most productive industries during this period. Among 95 companies, 8 companies are selected for investment. The results obtained from solving the proposed model with the assumption of k = 3 are presented in Tables (6), (7), and (8).

Table 6- The results of the proposed model with the assumption of (k=3).

Stocks	Weight	Price of prediction	Amount of stock	Budget of stock
Khesapa	0.074	2500	11174000000	27935000000000
Khodro	0.034	3095	10234000000	31674230000000
Khekomak	0.09	22080	28800000	635904000000

Table 7- The results of the proposed model with the assumption of (k=3).

Stocks	Weight	Price of prediction	Amount of stock	Budget of stock
Vesapa	0.17	6200	2720000000	16864000000000
Vebahman	0.062	3128	868000000	2715104000000

Table 8- The results of the proposed model with the assumption of (k=3).

Stocks	Weight	Price of prediction	Amount of stock	Budget of stock
Depars	0.32	30684	304000000	9327936000000
Dasveh	0.25	37150	187500000	6965625000000

6. Conclusions

Managing of stock portfolio and optimizing it is very important to success in the stock market. In this paper, a single-period multi-objective mathematical model for the stock portfolio selection has been presented based on the Markowitz model. The proposed model, considering the budget constraint, tries to enable investors to invest in stocks with the highest return and the lowest risk. In order to predict the stock price for 2023, a deep learning LSTM model has been used. The results obtained from the proposed model and its analysis with the data of the Tehran Stock Exchange can state that despite the severe sanctions against Iran, investment in the automobile industry due to the increase in the purchase price of domestic and foreign cars and according to the purchase demand automobile in Iran has been one of the best industries for investment. For future studies, one of the suggestions of this research is to examine the models of stock portfolio selection with the assumption of multi-period.

Reference

Boonjing, V., & Boongasame, L. (2017). Combinatorial portfolio selection with the ELECTRE III method: case study of the stock exchange of Thailand. *Afro-Asian Journal of Finance and Accounting*, *7*(4), 351-362.

Crama, Y., & Schyns, M. (2003). Simulated annealing for complex portfolio selection problems. *European Journal of operational research*, *150*(3), 546-571.

Di Persio, L., & Honchar, O. (2016). Artificial neural networks architectures for stock price prediction: Comparisons and applications. *International Journal of circuits, systems, and signal processing*, *10*(2016), 403-413.

Ding, X., Zhang, Y., Liu, T., & Duan, J. (2015, June). Deep learning for event-driven stock prediction. In the *Twenty-fourth international joint conference on artificial intelligence*.

Ghosh, A., Bose, S., Maji, G., Debnath, N., & Sen, S. (2019, September). Stock price prediction using LSTM on Indian Share Market. In *Proceedings of 32nd international conference on* (Vol. 63, pp. 101-110).

Hiransha, M., Gopalakrishnan, E. A., Menon, V. K., & Soman, K. P. (2018). NSE stock market prediction using deep-learning models. *Procedia computer science*, *132*, 1351-1362.

Kakaei, S. Jahangoshai, M (2016). *A multi-objective mathematical portfolio selection model for maximizing cover of companies in stock exchange industry groups (case study: Iran stock exchange)* [Unpublished master thesis]. The Urmia University of Technology.

Kim, K. J., & Han, I. (2000). Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. *Expert systems with Applications*, *19*(2), 125-132.

Konno, H., & Wijayanayake, A. (2001). Portfolio optimization problem under concave transaction costs and minimal transaction unit constraints. *Mathematical Programming*, *89*(2), 233-250.

Lee, C. K., Sehwan, Y., & Jongdae, J. (2007). Neural network model versus SARIMA model in forecasting Korean stock price index (KOSPI). *Issues in Information System*, *8*(2), 372-378.

Lim, S., Oh, K. W., & Zhu, J. (2014). Use of DEA cross-efficiency evaluation in portfolio selection: An application to Korean stock market. *European journal of operational research*, *236*(1), 361-368.

Lin, D., Wang, S., & Yan, H. (2001). A multi objective genetic algorithm for portfolio selection problem.

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, *7*(1), 77–91.

Mehtab, S., & Sen, J. (2022). Analysis and forecasting of financial time series using CNN and LSTM-based deep learning models. In *Advances in Distributed Computing and Machine Learning* (pp. 405-423). Springer, Singapore.

Pahwa, K., & Agarwal, N. (2019, February). Stock market analysis using supervised machine learning. In *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)* (pp. 197-200). IEEE.

Selvin, S., Vinayakumar, R., Gopalakrishnan, E. A., Menon, V. K., & Soman, K. P. (2017, September). Stock price prediction using LSTM, RNN and CNN-sliding window model. In *2017 international conference on advances in computing, communications, and informatics (icacci)* (pp. 1643-1647). IEEE.

Sen, J., & Mehtab, S. (2021, June). Accurate stock price forecasting using robust and optimized deep learning models. In *2021 International Conference on Intelligent Technologies (CONIT)* (pp. 1-9). IEEE.

Shi, S., Liu, W., & Jin, M. (2012, November). Stock price forecasting using a hybrid ARMA and BP neural network and Markov model. In *2012 IEEE 14th International Conference on Communication Technology* (pp. 981-985). IEEE.

Soleimani, H., Golmakani, H. R., & Salimi, M. H. (2009). Markowitz-based portfolio selection with minimum transaction lots, cardinality constraints and regarding sector capitalization using genetic algorithm. *Expert Systems with Applications*, *36*(3), 5058-5063.

Sunny, M. A. I., Maswood, M. M. S., & Alharbi, A. G. (2020, October). Deep learning-based stock price prediction using LSTM and bi-directional LSTM model. In *2020 2nd Novel Intelligent and Leading Emerging Sciences Conference (NILES)* (pp. 87-92). IEEE.

Tsantekidi, A., Passalis, N., & Tefas, A. (2022) Recurrent neural Networks in Alexandros Iosifidis (editor), Anastasios Tefas (editor) Deep Learning for Robot Perception and Cognition. *London: Academic Press.* (pp.108-109).

Vargas, M. R., De Lima, B. S., & Evsukoff, A. G. (2017, June). Deep learning for stock market prediction from financial news articles. In *2017 IEEE international conference on computational intelligence and virtual environments for measurement systems and applications (CIVEMSA)* (pp. 60-65). IEEE.

Wijaya, Y. B., Kom, S., & Napitupulu, T. A. (2010, December). Stock price prediction: comparison of Arima and artificial neural network methods-An Indonesia Stock's Case. In *2010 Second International Conference on Advances in Computing, Control, and Telecommunication Technologies* (pp. 176-179). IEEE.