

# Estimating the Return on Investment Opportunities in Financial Markets and Establishing Optimized Portfolio by Artificial Intelligence

**Dr.Farзад Karimi**

Assistant Professor of Economics, Islamic Azad University, Mobarakeh Branch, Isfahan, Iran

**Dr.Alireza Zare'ie**

Assistant Professor of Accounting, Islamic Azad University, Mobarakeh Branch, Isfahan, Iran

**Mehdi SalemiNajafabadi<sup>1</sup>**

M.A. in Accounting, Islamic Azad university, Mobarakeh Branch, Isfahan, Iran

DOI: 10.6007/IJARBSS/v3-i7/45

URL: <http://dx.doi.org/10.6007/IJARBSS/v3-i7/45>

## Abstract

This project is looking for increasing return on investment, by presenting models based on artificial intelligence. Investment in financial markets could be considered in short-term (daily) and middle-term (monthly) basis/ hence the daily data in Tehran Stock Exchange and the rates of foreign exchange and gold coins have been extracted for the period Mar. 2010 to Sep. 2012 and recorded as the data into the neural networks and the genetic programming model. Also the monthly rate of return and risk of 20 active companies of the stock exchange, and the monthly risk values of foreign exchange and gold coin, as well as bank deposits were used as genetic algorithms in order to provide optimum investment portfolios for the investors.

The results obtained from executing the models indicates the efficiency of both methods of artificial neural network and also genetic programming in the short-term financial markets predictions, but artificial neural networks show a better efficiency. Also the efficiency of geneticalgorithm was approved in improving the rate of return and risks, via identifying the optimum investment portfolios.

**Keywords:** Financial markets, Return, Artificial Neural Network (ANN), GeneticProgramming (GP), GeneticAlgorithm (GA).

## Introduction

Investment is regarded as the economical basis of any country that if it is done perfectly, it could increase national production and economic growth. There are 4 main markets in any economy that include: Commodity market, Labor market monetary and capital market. Generally, monetary and capital markets are called the financial part of economy (financial markets). These two parts resemble the wheels of a bicycle that should move together in order

---

<sup>1</sup>Email: [saleminajafabadi@gmail.com](mailto:saleminajafabadi@gmail.com)

to cause economic growth and development (Sonmez, 2007). These markets affect each other and are effective on each other. The investment process starts by using its resource; i.e. savings. If we look at the places that savings are inclined to, we will see that these savings are mainly invested in the following fields:

- 1- Transaction of gold and gold coin
- 2- Depositing in banks
- 3- Transaction of foreign exchanges
- 4- Transactions of bonds and shares
- 5- Buying and selling lands and constructions
- 6- Investing for future transactions in Mercantile exchange
- 7- Other opportunities

Continuous changes of stock exchange index and extensive fluctuations of gold and foreign exchange and other financial markets confusion the investors in identifying investment risks and returning opportunities and recognizing the most profitable section. Moreover, determining which section could provide the highest rate of profit in long-term, medium-term or short-term periods are difficult. The confusion causes the drift of investments within financial markets and sometimes causes unwanted and sentimental fluctuations of the capital market. Also, in addition to the investors' losses, the market efficiency is not performed properly in allocation of the capital, and confronts the national production and economic growth with serious damages. According to the mentioned points, the main question in this research is that if a reliable model could be designed to predict the performance of different investable market for the investors to identify appropriate investment opportunities by them or not.

For the daily prediction of stock exchange index, the average of prices for the week ending the same day in other financial markets is used in this research. These variables have interaction relations, meaning that by increasing the prices in a market, the investments are inclined toward that market and the prices are affected in other markets. In this research, to analyze this interaction relations, the newest technologies for artificial intelligence, i.e. artificial neural networks and genetic programming are used. Also the genetic algorithm was used to prepare the optimum investment portfolio in the financial market. We are going to briefly introduce these techniques as follows:

#### **Artificial neural networks (ANN):**

Neural networks are calculating structures with learning and generalizing potentials. Neural networks are made of a series of layers including simple processing components called "neurons" that act in parallel to each other. Each entering layer is related to one or more medium layers and medium layers are connected to the output layer, where the network response appears in the role of system output. Artificial neurons are the basis of any neural networks, and its design and structure are inspired from biological neural networks. There are 2 simple similarities between a biological and an artificial network. Firstly, the shape and structure of both of them are simple and secondly, the connection between the neurons determines the network working techniques (Panahiyan, 2000). The structure of multi-layer perceptron neural network includes an input layer, one or more hidden layers and output layer. Each of these layers is formed by one or more nodes. The input layer has nodes with the same number of independent variables and similarly the output layer has the same number of

dependent variable, but defining the hidden layer is difficult (Hagan, 2002; Hoglund, 2012). Neural networks could be distinguished according to 3 different factors (Brabazon and O'Neill, 2006):

**A: Connection topology:**

Connection topology expresses the number of layers and the connection of neurons. Connection of neurons is divided into inter-layer connections and intra-layer connections. In intra-layer connections a neurons could be connected or may not be connected to other neurons in its layer. But in inter-layer connections each neuron could be connected to all neuron in its next layer or any neuron could be connected to a special neuron in the next layer. Also a neuron of layer could only be connected to the next layer., or be connected to some different layers. Moreover the connection between layers could be unilateral or bilateral and hence the two main neural networks could be distinguished.

- feed forward networks that have no feedback loops and the data is transferred only in one direction, from input to the output.
- Recurrent networks or feedback networks with feedback loops, i.e. the data is transferred from the input to the output, or vice versa.

**B: Training method:**

One important characteristic of artificial neural networks that approaches to human performance is the learning potentials. Learning may include two types of data processing: inductive and deductive. In inductive processing, the general paradigms and regulations are determined by pure and raw data and by using experiences, but in deductive processing, the general regulations are used to determine the realities although giving a precise definition of learning is difficult, but it can be said that the learning process in neural networks and the subject of updating the connection weights of it on such that a network to do a specific task efficiently.

**C: Learning algorithm:**

Learning algorithms determines the effectiveness of errors in the instruction trend and also the state of updating neural networks to reduce the errors. Learning algorithms is the approach towards reaching optimized coefficients in neural networks.

Apparently, White (1988) was the first person who used neural networks in predicting markets. He was curious about using neural networks in extracting nonlinear regulations of economic time series and finding the regulations that had not been discovered in the subject of assets values , previously, such as the fluctuations in ordinary share prices.

**Genetic algorithm (GA):**

Genetic algorithm is a searching algorithm in accordance to genetic structures of living animals. This algorithm combines selecting principle (or Survival of the Fittest) with a set of structure random information, creating a search algorithm with the characteristics of a natural evolution methods. The principle of Darwin (survival of the fittest) states that the generations of species survive from a population that has the best characteristics. The ones without such characteristics are destroyed gradually. From the combination of better chromosome, better generations are created on the nature. Sometimes mutations occur on them that many result in better generations . Genetic algorithm solves problems with inspiring from such an idea. The applied regulations in genetic algorithm that inspire from the rules of nature are as follows:

- A- Living creatures fight against natural powers to survive. The ones with more conformity with such powers (stronger, faster, larger) are more likely to survive.
- B- The ones with more compatibility survive and regenerate. In other words, more suitable species of the primary population survive for future regenerations.
- C- Children look like their parents but they are not identical (due to inheriting from both parents and genetic mutations); hence children may or may not be more appropriate.
- D- Children follow their parents and after some generations , the surviving children are the ones with more compatibility to their ancestors.

The working of genetic algorithm is also similar to the above, but the difference is that the genetic algorithm program should be defined with regards to the considered living creatures, reproduction style and their rate of compatibility. Before anything else, a strategy is defined for genetic algorithms to exhibit the solutions for the problems, as a chromosome. Then a set of chromosomes that are indeed expressing the solutions are considered as the primary population. After this stage, by using genetic operators, production of new chromosomes called “infants” is to be done, After generating a number of selected chromosomes with the same no.as the primary population, some members should be selected for the next reproduction . Selection process is based on the fitness of each strip. Most often, the fitness function is considered for each strip to be equal with the target function of optimization aspect. So far, one repetition of a generation or a generation from an algorithm is produced. After generating some generations , the algorithm is covered towards the optimum solution (Masih'abadi and Abdolahi, 2009; Monajemi et al, 2009).

#### **Genetic programming (GP):**

Like genetic algorithm, genetic programming is an algorithm inspired from the nature trying to improve computerizing systems. This method was first considered by Koza (1992). In this method, the solutions are created via simulating the evolution trends in creatures and the law of “the survival of the fittest”. The fundamental difference between genetic algorithm and genetic programming is in their solutions, or in other words in their outputs. Genetic algorithm gives a string of numbers that indicate an optimum solution, while genetic programming provides a calculating model that could make the output values and data by benefiting from the input data . This model is established by a set of connected nodes and these nodes could be in the form of “function sets” or “terminalsets”. The terminal sets for the independent variables and the function sets could be in the form of simple mathematical functions like summation, subtraction, multiplication and division, or they could be in the form of other mathematical functions like trigonometric, logarithmic, exponential functions, etc. (AlQuraishi, 2009).

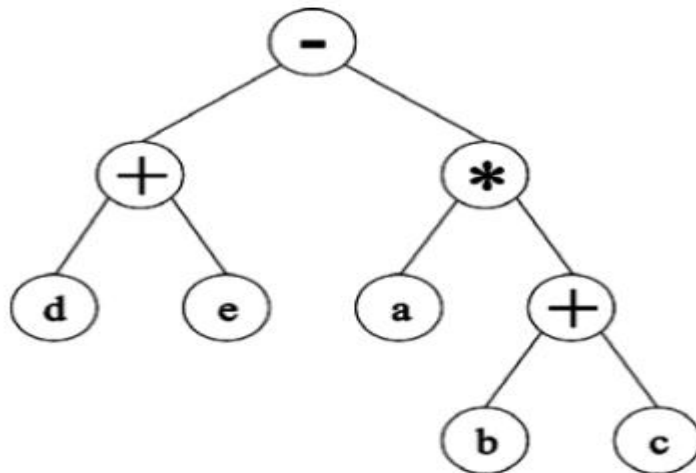


Fig.1. The presentation of a GP tree

The output of the genetic programming model could be shown as a tree like in fig.1. In these trees, the nodes with mathematical signs indicate function sets, while the nodes having alphabets show terminal sets (Tae et al. 2009). This algorithm with the function of genetic operators on the tree is looking for a solution or an equation with the highest rate of fitness. This algorithm is executed in different stages as follows:

- a: Creating a primary population of solutions- At this stage, the algorithm creates different solutions in random form.
- b: Measuring the rate of fitness of each tree-At this stage.
- c: Applying genetic operators on the trees.
- d: Steps (b) and (c) are replaced as much for the stopping conditions to be obtained. The stopping condition could be the definite number of repeating algorithm, the number of reviewing target function passing the definite time, reaching acceptable error levels or lack of improving solutions in some successive repetitions.

### Research hypotheses

- H.1: A reliable model could be established for predicting the behavior of capital market by using genetic programming
- H.2: A reliable model could be established for predicting the behavior of capital market by using artificial neural network
- H.3: Artificial neural network is more efficient than genetic programming model, for predicting the capital market.
- H.4: optimized investment portfolios could be identified by using genetic algorithm model.

### Research methodology

It was tried in the present research to determine the optimum investment model for the investors. We had our efforts to find the best model for short term prediction for the capital market by using genetic programming model and the artificial neural networks, such that by using the average of prices in a week ending the mentioned period, it was tried to predict the price for the next day. Then, the efficiency of genetic algorithm in finding the market efficiency boundary and providing optimum investment portfolios, for guiding for investment in medium

term basis was tested. The judging criterion about the efficiency of this algorithm is the Reward to Variability Ratio (sharp Index).

This research is an applied research. The research plan is semi-experimental by using post-occurring approach (via previous information). The independent variables are: USD market price, gold coin bubbling price, world price of an ounce of 24 carat gold and OPEC oil price. Also the dependent variable for these models is the price index in Tehran stock exchange.

We used 3-layer perception, feed-forward, neural networks in this research. Learning algorithm is Levenberg-Marquardt back propagation and connection structure is one way, with complete connection of each neuron to all the neurons in the next layer. The best network was selected by testing the no. of different neurons in the middle layer (from 1 to 20).

For genetic programming model a model with the population of 40,600 generations, mutation rate of 0.15 and crossover rate of 0.6 was used to predict the price index of the stock exchange. The optimum algorithm had 40 genes, determined by trial and error.

### Results

The best produced equation by the genetic programming model could predict the test data with the accuracy of 97.41%. the network mean squared error is 588.687 for the test data. Hence the first hypothesis was approved.

Regarding the 2<sup>nd</sup> hypothesis, after designing and executing the neural network with different no. of neurons in the middle layer and comparing them with each other, it was found that the neural network with 17 neurons in the middle layer could have minimum errors in predicting the stock exchange price index. Min. mean squared error was 99092, mean network error was 14.512 and the percentage of network error in test data was 1.06% of the stock exchange price index, and the network could predict the price index with 98.94% of accuracy. Therefore the 2<sup>nd</sup> hypothesis was approved. Testing the 3<sup>rd</sup> hypothesis was done by comparing the values of error percentage and the models root mean squared errors. The comparisons show that in both cases, artificial neural network is better than genetic programming. Therefore the 3<sup>rd</sup> hypothesis was approved

	Error percentage	RMSE
Artificial neural network	1.06%	314.788
Genetic programming	2.59%	588.687

In addition to quantitative criteria, the other superior points of neural networks as compared to genetic programming could be as follows:

- Much higher speed
- Less calculating complexities
- Easy applications and less essential adjustments to find the responses
- Existence of advanced and comprehensive software to execute the models.

The rate of return and risks of investment opportunities were extracted for the last month of the research period to test the 4<sup>th</sup> hypothesis. These opportunities include the shares of 20 more active companies of the stock exchange during the research, USD, gold coin and depositing in banks. The used rate of return in this model was the rate of return due to keeping the properties; i.e. the price of the end of the period minus the price of the beginning of the period plus the paid dividends. Risk is found by calculating the 2<sup>nd</sup> order moment around the

return rate resulted from keeping the properties. Capability of this model was calculated according to comparing the sharp index of optimum portfolios with sharp index of financial assets. Analysis shows that all the optimum portfolios have sharp index higher than all financial assets.

Thus the 4<sup>th</sup> hypothesis was confirmed. To follow, the matrix for the assets correlation coefficient and the specifications of the established optimum portfolios are presented.

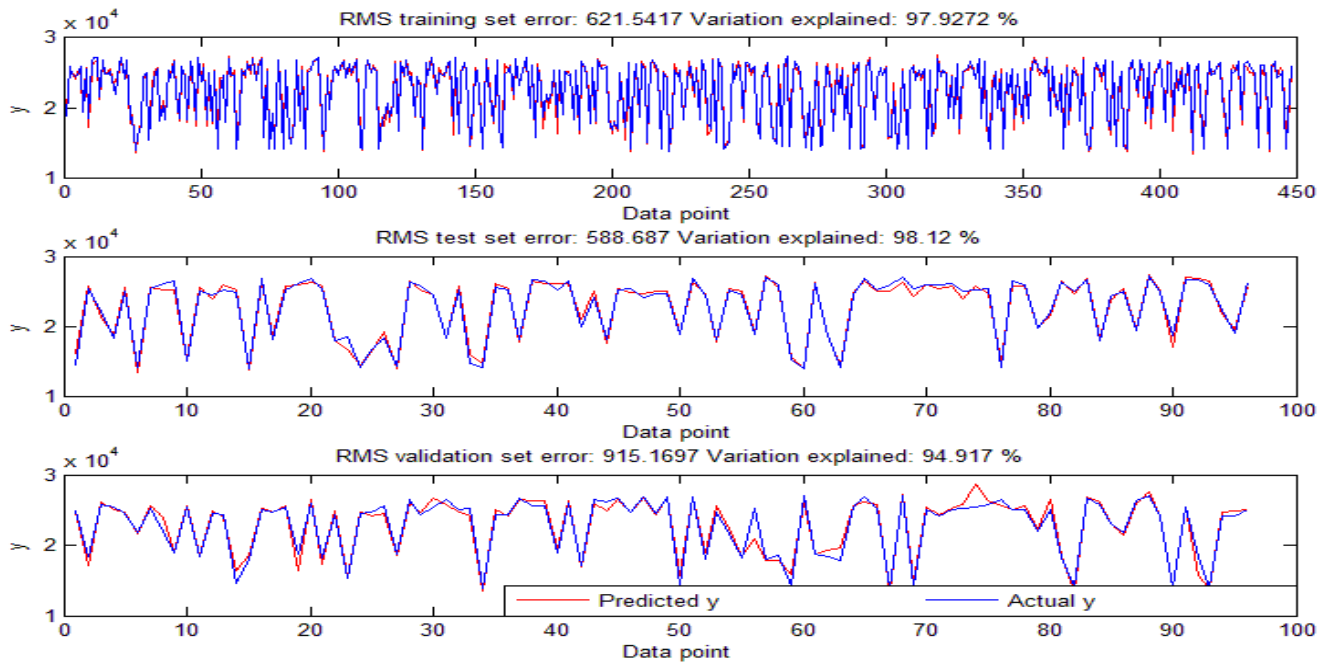


Fig.2. Best GP model prediction

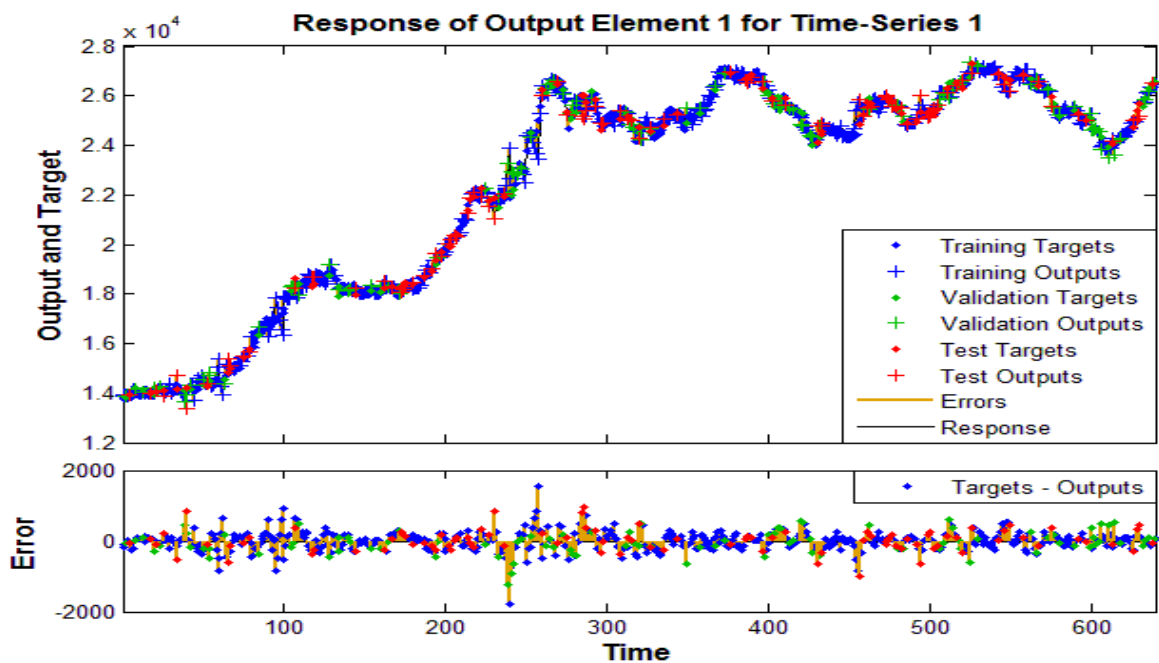


Fig.3. Best ANN prediction

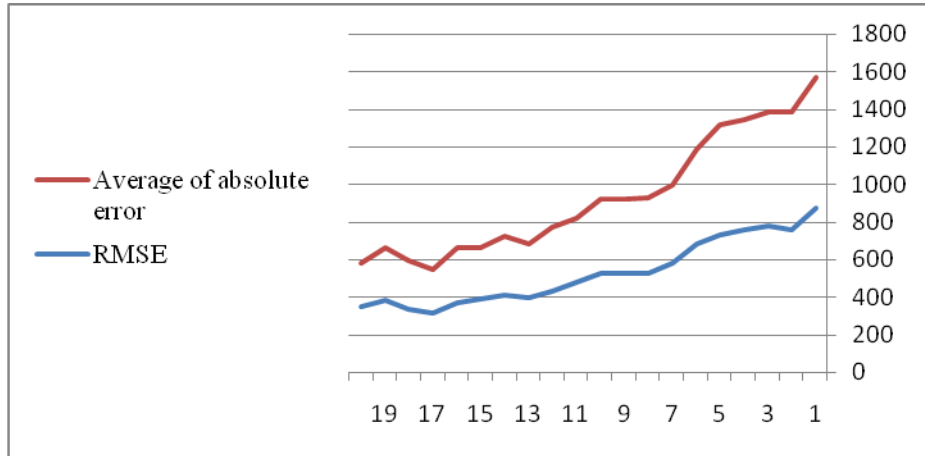


Fig.4. RMSE and error of ANNs by 1-20 neurons in middle layer

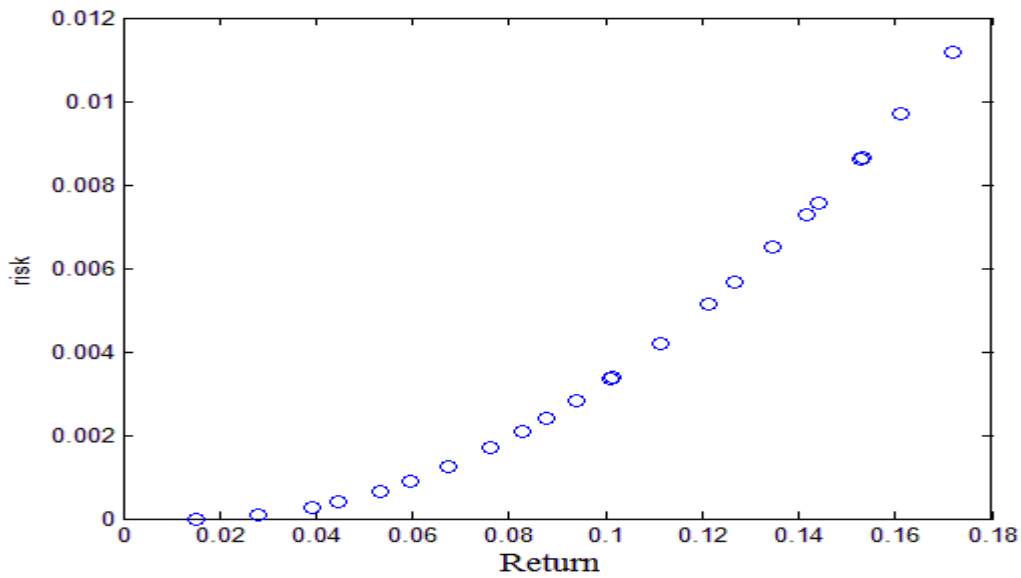


Fig.5. Optimized Portfolios Return and Risk



Fig.6. Optimized portfolios properties

Sharpe Index	Risk	Return	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	Gold	USD	Deposit in Bank
170.9833	7.5E-05	0.027909	0.001563	0.001563	0.001563	0.001563	0.001563	0.001429	0.001563	0.001563	0.001563	0.001563	0.001563	6.74E-09	0.001563	0.049453	0.001563	0.001563	0.001563	0.001563	0.001563	0.001563	0.001563	0.001563	0.1919425
91.3041	0.000265	0.039174	0.000297	0.000297	0.000297	0.000297	0.000297	0.002675	0.000297	0.000297	0.000297	0.000297	0.000297	1.09E-09	0.000297	0.092611	0.000297	0.000297	0.000297	0.000297	0.000297	0.000297	0.000297	0.000297	0.849411
74.12012	0.000402	0.044779	0.003605	0.003605	0.003605	0.003605	0.003605	0.003296	0.003605	0.003605	0.003605	0.003605	0.003605	5.1E-09	0.003605	0.114081	0.003605	0.003605	0.003605	0.003605	0.003605	0.003605	0.003605	0.003605	0.814127
57.81788	0.00066	0.053175	0.004622	0.004622	0.004622	0.004622	0.004622	0.004225	0.004622	0.004622	0.004622	0.004622	0.004622	9.66E-09	0.004622	0.146247	0.004622	0.004622	0.004622	0.004622	0.004622	0.004622	0.004622	0.004622	0.761718
49.69475	0.000894	0.059416	0.006377	0.006377	0.006377	0.006377	0.006377	0.004915	0.006377	0.006377	0.006377	0.006377	0.006377	1.87E-08	0.006377	0.170153	0.006377	0.006377	0.006377	0.006377	0.006377	0.006377	0.006377	0.006377	0.722767
41.97346	0.001253	0.067587	0.006366	0.006366	0.006366	0.006366	0.00582	0.006366	0.006366	0.006366	0.006366	0.006366	0.006366	3.39E-08	0.006366	0.201455	0.006366	0.006366	0.006366	0.006366	0.006366	0.006366	0.006366	0.006366	0.671763
35.99971	0.001703	0.076312	0.007422	0.007422	0.007422	0.007422	0.006785	0.007422	0.007422	0.007422	0.007422	0.007422	0.007422	1.64E-08	0.007422	0.254883	0.007422	0.007422	0.007422	0.007422	0.007422	0.007422	0.007422	0.007422	0.617303
32.53073	0.002086	0.08285	0.008214	0.008214	0.008214	0.008214	0.007509	0.008214	0.008214	0.008214	0.008214	0.008214	0.008214	3.37E-08	0.008214	0.25993	0.008214	0.008214	0.008214	0.008214	0.008214	0.008214	0.008214	0.008214	0.576491
30.35111	0.002396	0.087723	0.008804	0.008804	0.008804	0.008804	0.008048	0.008804	0.008804	0.008804	0.008804	0.008804	0.008804	3.6E-08	0.008804	0.278596	0.008804	0.008804	0.008804	0.008804	0.008804	0.008804	0.008804	0.008804	0.546078
28.00541	0.002814	0.095814	0.009541	0.009541	0.009541	0.009541	0.008722	0.009541	0.009541	0.009541	0.009541	0.009541	0.009541	8.72E-09	0.009541	0.301932	0.009541	0.009541	0.009541	0.009541	0.009541	0.009541	0.009541	0.009541	0.508062
25.67826	0.003347	0.100957	0.010406	0.010406	0.010406	0.010406	0.009512	0.010406	0.010406	0.010406	0.010406	0.010406	0.010406	8.68E-09	0.010406	0.329295	0.010406	0.010406	0.010406	0.010406	0.010406	0.010406	0.010406	0.010406	0.46548
25.5211	0.003389	0.101486	0.01047	0.01047	0.01047	0.01047	0.009571	0.01047	0.01047	0.01047	0.01047	0.01047	0.01047	3.11E-08	0.01047	0.331322	0.01047	0.01047	0.01047	0.01047	0.01047	0.01047	0.01047	0.01047	0.460172
22.90959	0.004205	0.111345	0.011663	0.011663	0.011663	0.011663	0.010662	0.011663	0.011663	0.011663	0.011663	0.011663	0.011663	5.82E-09	0.011663	0.369091	0.011663	0.011663	0.011663	0.011663	0.011663	0.011663	0.011663	0.011663	0.398641
20.72454	0.005139	0.121503	0.012893	0.012893	0.012893	0.012893	0.011786	0.012893	0.012893	0.012893	0.012893	0.012893	0.012893	0	0.012893	0.408006	0.012893	0.012893	0.012893	0.012893	0.012893	0.012893	0.012893	0.012893	0.335238
19.723	0.005674	0.126911	0.013548	0.013548	0.013548	0.013548	0.012385	0.013548	0.013548	0.013548	0.013548	0.013548	0.013548	7.13E-09	0.013548	0.428724	0.013548	0.013548	0.013548	0.013548	0.013548	0.013548	0.013548	0.013548	0.30148
18.42893	0.006499	0.134769	0.014499	0.014499	0.014499	0.014499	0.013254	0.014499	0.014499	0.014499	0.014499	0.014499	0.014499	5.44E-09	0.014499	0.458829	0.014499	0.014499	0.014499	0.014499	0.014499	0.014499	0.014499	0.014499	0.252431
17.40804	0.007284	0.141793	0.01535	0.01535	0.01535	0.01535	0.014032	0.01535	0.01535	0.01535	0.01535	0.01535	0.01535	1.07E-08	0.01535	0.485737	0.01535	0.01535	0.01535	0.01535	0.01535	0.01535	0.01535	0.01535	0.208589
17.07742	0.007568	0.144248	0.015647	0.015647	0.015647	0.015647	0.014303	0.015647	0.015647	0.015647	0.015647	0.015647	0.015647	8.4E-09	0.015647	0.495141	0.015647	0.015647	0.015647	0.015647	0.015647	0.015647	0.015647	0.015647	0.193268
16.00904	0.008612	0.152873	0.016691	0.016691	0.016691	0.016691	0.015258	0.016691	0.016691	0.016691	0.016691	0.016691	0.016691	6.24E-09	0.016691	0.528185	0.016691	0.016691	0.016691	0.016691	0.016691	0.016691	0.016691	0.016691	0.139431
15.98716	0.008636	0.153062	0.016714	0.016714	0.016714	0.016714	0.015279	0.016714	0.016714	0.016714	0.016714	0.016714	0.016714	6.5E-09	0.016714	0.528907	0.016714	0.016714	0.016714	0.016714	0.016714	0.016714	0.016714	0.016714	0.138253
15.08571	0.009694	0.161273	0.017708	0.017708	0.017708	0.017708	0.016187	0.017708	0.017708	0.017708	0.017708	0.017708	0.017708	4.7E-09	0.017708	0.560364	0.017708	0.017708	0.017708	0.017708	0.017708	0.017708	0.017708	0.017708	0.087001
14.05727	0.01117	0.172016	0.019008	0.019008	0.019008	0.019008	0.017376	0.019008	0.019008	0.019008	0.019008	0.019008	0.019008	0	0.019008	0.60152	0.019008	0.019008	0.019008	0.019008	0.019008	0.019008	0.019008	0.019008	0.019946
14.05727	0.01117	0.172016	0.019008	0.019008	0.019008	0.019008	0.017376	0.019008	0.019008	0.019008	0.019008	0.019008	0.019008	0	0.019008	0.60152	0.019008	0.019008	0.019008	0.019008	0.019008	0.019008	0.019008	0.019008	0.019946

## Reference

- AlQuraishi, A. (2009). Determination of Crude Oil Saturation Pressure Using Linear Genetic Programming, **Energy & Fuels**, (23), pp. 884–887.
- Brabazon, A & O'Neill, M. (2006). *Biologically Inspired Algorithms for Financial Modeling*. Berlin: Springer, 1st ed: 15-20.
- Hagan, M. T., Demuth, H. B. and M. Beal. (2002). *Neural Network Design*. Singapore: Thomson Asia Pte Ltd.
- Hoglund, H. (2012). Detecting Earnings Management with Neural Networks. **Expert Systems with Applications**, (39): 9564-9570.
- Koza, J. (1992). *Genetic programming: On the Programming of Computers by Means of Natural Selection*, England, London: MIT Press, Cambridge, Sixth ed.
- Masih'abadi, A., Abdolahi, A. (2009). The Concepts and applications of genetic algorithm in accounting, **Journal of Accounting knowledge and Research**, 17, pp. 24-29.
- Monajemi, S.A., Abzari, M. and Ra'yati, A. (2009). Forecasting the stock price in stock exchange market using the combined model of Neural Networks and Genetic Algorithm and comparing it with Artificial Neural Networks, **Quarterly Journal of Quantitative Economics**, 6 (3), pp. 1-26.
- Panahiyan, H. (2000). *The application of Neural Networks in forecasting the Stock Prices in Stock Exchange of Tehran*, Ph.D Thesis, [Islamic Azad University, Science and Research](#).
- Sonmez, F. saryal. (2007). Does inflation have an impact on conditional stock market volatility? Evidence from turkey and Canada. **international research journal at finance and economics**, (11): 122-133
- Tae-Mun Hwang, Hyunje Oh, Youn-KyooChoung, Sanghoun Oh, MoonguJeon, Joon Ha Kim, Sook Hyun Nam, Sangho Lee. (2009). Prediction of membrane fouling in the pilot-scale microfiltration system using genetic programming, **Desalination**(247), pp. 285–294.
- White, H. (1988). *Economic Prediction Using Neural Networks: The Case of IBM Daily Stock Return*, Proceeding of the second IEEE international conference on neural network.