Effective components on the forecast of companies' dividends using hybrid neural network and binary algorithm model

Mahdi Salehi1*, Behad Kardan2 and Zohresh Aminifard3
1Department of Accounting, Ferdowsi University of Mashhad, Iran
2Department of Accounting, Ferdowsi University of Mashhad, Iran
3Department of Accounting, Science and Research of Hormozgan Branch, Islamic Azad University, Hormozgan, Iran

*mahdi_salehi54@yahoo.com

Abstract
The issue of accounting profit has been noticed from long time by investors, managers, financial analysts and creditors. Due to the importance of dividend per share is disclosed by companies and the role of dividend in decisions and because the most important source of information for investors and managers and other users in the stock, is the forecasted dividend by companies, this study follows to recognize the affecting factors on 23 chemical companies in the Tehran Stock Exchange dividend using genetic algorithms combined with artificial neural network. Finally, the variables affecting the output are used to predict dividends in the model that is by neural network designed. The error is calculated and be the basis of comparison with other methods. The study included chemical companies accepted in Tehran Stock Exchange during 2006-2010. The independent variables in this study are accounting ratios and stock cash dividend is dependent variable.

Keywords: Prediction, Dividends, Neural network, binary algorithm

Introduction
Cash return and cash stock dividends, due to the objectivity and tangible, has a special consideration among some of stakeholders. In fact, actual and potential users of financial information have been eager to inform about the ability to create cash and sometimes its distribution among the stakeholders of the company. Because this information not only offers a clear picture of the current situation of company, but it also can provide the estimates of the future status that certainly is important in their decision making process. This issue has a serious importance for corporate managers, to use information due to the use of information obtained in the management process of company and market evaluations of their performance. Therefore part of the attention of corporate managers paid to the subject that is titled "dividend policy". But more important than the dividend policy, is the root of reasons to adopt a specified dividend policy by the company. This may help in the important economic decisions for different groups of stakeholders, particularly the investors. The reasons and factors obtained from this study, not only help to explain the behavior of companies in the past, but rather provides a tool to predict the movement and its future direction (Jahankhani & Ghorbani, 2006). In general, the company's attractiveness to investors, in addition to the current situation, depends on its potential of creating the future income.

The company which is currently profitable and its profitability continuing in the future are more attractive than the company whose potential benefits will be reduced soon. It can be claimed that the most important criteria for investment decisions, is the current and future profitability of companies, so that investors in different investment strategies, mainly decide due to profit making of companies and one of the most important decisions of financial managers, is decisions related to stock cash dividends (Nouraini, 2006; Hghihat et al., 2011). Due to the importance of decisions related to stock cash dividends there are always a lot of models for predicting these factors that it can be pointed to one of the strongest of them, the neural network. In the proposed model in this study, using a combination of two methods (binary genetic algorithms and neural networks), have been tried to identify the factors that mostly influence on the chemical company's stock cash dividends.

Gounopoulos (2003) evaluates the forecast accuracy and the factors affecting the predicted accuracy of companies that their shares were public offered for the first time in the Athens Stock Exchange during 1994-2001. The results show existence of an inverse relationship between firm size and the prediction error. While the more the time period, it causes the prediction error to decrease. Finally, this research did not show a positive relationship between the level of financial leverage, shareholding structure and macro-economic conditions and the prediction error.

Kato et al., (2009) investigated the optional forecast of dividends by management in listed companies in Japan during 1997-2006. The results showed that most of the dividend forecast is higher than actual dividend, but forecasts adjusted and reduced during the year. While the predictions have average information, the anticipated dividend of companies with poor performance and with managers that had poor and optimistic predictions, have little and low reliability information.

Savov and Weber (2006) study sample consisted of German companies during 1982 -2003. Results show that the market situation, leads the decision to increase dividend. In Iran the following study had been carried out. Mashayehk and Shahrokhi (2007) compared the accuracy of manager's prediction to future earnings per share, with predictions based on the random walk model. For this purpose prediction to the 279 companies
including 639 observed were analyzed during 2002-2004 using mean difference. The findings of this research show, there are significant differences between the manager's prediction error and the error of prediction based on random walk. Furthermore, the comparison of mean difference of two models, indicating a higher prediction accuracy of managers than prediction based on random walk. The other results of study assumptions show that predictions of managers have optimistic deviations and prediction accurately is different due to the size of the company, profitable or detrimental and type of the industry.

Sarebanha and Ashtab (2008) look at factors affecting the prediction of error of 107 listed companies in Tehran Stock exchange during 1999-2006 and concluded, the correlation between company size, company life, time horizon forecasting, profitability ratios, leverage ratio and auditor's valid only profitability ratios factor affecting dividend forecast errors, that are inversely associated with the prediction error.

Azad (2004) studied 58 companies listed on the Stock Exchange. The results showed that the dividend anticipated by the companies and the real profits have also a significant relationship. Also, the relation between the benefits anticipated by the company and stock returns are significant.

Samadzadeh (1993) evaluates the dividend strategies and their impact on companies stock value. The study results show the dividend policies in the stock market, are not known for corporate managers and shareholders do not notice to cash dividends as an adjustment that have message.

Research methodology

The choice of affecting factors on dividend forecasts using neural networks and binary genetic algorithms hybrid model can correctly reduce the dividend predict error of chemical companies’ active in the stock exchange. This research is applicable and is based on field research, based on information collected from the Tehran Stock Exchange, research hypotheses and then test results can be generalized to the whole population. Model, is a combination of, binary genetic algorithms, and neural network that genetic algorithm is known mostly as a way to optimize the function. Implementation of genetic algorithm begins with an initial population of chromosomes. The initial solution is then evaluated according to their level of competence to be given the opportunity to reproduce. To be made the best generation at the end.

In the first stage the variables that are likely to affect the dividend will be entered into the binary genetic algorithms. The fitness function of genetic algorithm is a neural network, which its inverse error criterion is used to fit chromosomes in binary genetic algorithms. After the genetic algorithm binary, binary, chose the best combination of input variables, the variables entered into artificial neural network that designed to suit and predict selected variables, after that data variables being elected the previous section (genetic algorithm), trained neural network, the network is tested with evaluation data and thus the prediction error will be measurable.

To achieve the desired objectives of research and to reduce the effect of other factors such as type of industry and circumstances prevailing in the industry, from various industries active in the Tehran Stock Exchange, on dividends. The chemical industry which was one of most active companies in the stock and its data are available for 2006 years to 2010 is chosen as the statistical community, as a result, 23 companies are selected from 30 chemical company active in the stock market.

Regarding to the specifications of neural networks that higher the number of test data, results a better response from the network (Azar & Karimi, 2009), all companies in the survey were selected as sample and sampling is not done.

Artificial neural networks

This area is one of the most dynamic areas of contemporary research that has attracted many researchers from various scientific fields. Using neural networks and genetic algorithms has increased more and more these days to solve complex applied problems. The importance of this model is that it can describe and examine processes that depend on several parameters and with different degrees of importance and then provide a satisfactory answer. Way of dealing with computational methods of neural networks is capturing the fundamental strategic principles of brain processes and their application in computer systems. Artificial neural networks are intelligent dynamic free-model systems that with the processing of experimental data transmit the knowledge, or the law behind the data, to the network structure. These systems based on computational intelligence, are trying to model the human neurosynaptic brain structure. Major components of computational intelligence or soft computing are neural networks (neuronal computations), fuzzy logic (approximate calculation) and genetic algorithms (genetic analysis) that each one have a kind of brain modeling. Neural networks, inspired by the extensive distribution and parallel computing in the human brain, the biological and learning neural networks in these systems, model the synaptic connections and neuronal structure. A neural network has been composed of a large number of nodes and some directional lines that are connected to nodes, the nodes in the input layer are sensory nodes and the ones in output layer nodes are called respondents, between input and output neurons, hidden neurons are located. Information enters from input nodes to the network and then connects to hidden layers from connections and finally the network output of output layer nodes is achieved (Bill & Jackson, 2007; Nikbakht & Sharifi, 2010) (Fig.1).
Fig. 1. An example of a neural network structure

![Neural Network Diagram](image)

**Binary genetic algorithm**

Genetic Algorithm is an optimization method inspired by the living nature that can be classified as a numerical method, random and direct search. This algorithm is based on the repetition that its basic principles adapted of genetics and has been invented duplication of some of the processes observed in natural evolution, effectively uses the old knowledge of a population, to provide new and improved solutions.

**Fig. 2. Conceptual framework of the system**

![Conceptual Framework](image)

In the beginning, the initial population that shows the answer will be randomly selected. Each of the members of these population are called chromosomes is one of the answers of problem. Each of these chromosomes is selected of equal-length string of numbers, which each of these numbers is called a gene. Genetic algorithm based on repeated acts. The population in each stage is called generation. Each of the members of this generation is evaluated based on the value function. In these algorithms, the new generation is trying to allocate more value from the value function and with this performance to be closer to target function. At each step of the replication, each of the chromosomes, with certain probability, cross with others, or married so that the outcome one or more new chromosome is called a child. In these children may have genetic mutations at certain probability, as this will change the amount of one or more genes of chromosome (Akhbari, 2008). In the final stage, the children evaluated based on their value function and new generation will be produced according to their value and the worth of their parents, the first generation. These processes are repeated until the present generation will converge to the optimal solution or one of optimal sub-solutions. There are different methods for cross and mutation operators that due to the complexity of the problem one of them is selected.

**Dividend forecasts using neural network and binary algorithm**

Fig. 2 can be considered as a conceptual framework in which, X is the vector of input variables (here, variables influencing the prediction of a stock dividend) and Y is the vector of output variables (in this case, the stock dividend) (Makvandi et al., 2008). The main purpose of this paper is to find the optimum combination of input variables of the system and in this case, binary genetic algorithm (collective motion of particles algorithm) is used. In fact, we find out what ratios and variables must be extracted from financial statements and entered their values into the prediction model to predict the stock dividends.

In this section, the design and implementation of the proposed hybrid approach, is expressed. Scheme 1 shows the complete process of the proposed method. There are 5 processes (level) in the proposed method that in order are: the data selection, data cleaning and preparation and, factors affecting choice of cash forecasting profits of stock, multi-layer perceptron neural network training process based on selected components and finally evaluation of models trained with test data that has not been observed so far by the algorithm.

The MATLAB software version 7.6.0 is used for algorithms implementation. This software package provides many ready tools to work with that are makes it very comfortable.

**Table 1. The independent variables of the study**

<table>
<thead>
<tr>
<th>No</th>
<th>Name of independent variable</th>
<th>No</th>
<th>Name of independent variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Capital</td>
<td>13</td>
<td>Quick Ratio</td>
</tr>
<tr>
<td>2</td>
<td>Current Ratio</td>
<td>14</td>
<td>Return on capital</td>
</tr>
<tr>
<td>3</td>
<td>Dividend yield</td>
<td>15</td>
<td>Return on equity%</td>
</tr>
<tr>
<td>4</td>
<td>Earnings per share</td>
<td>16</td>
<td>Return on sales%</td>
</tr>
<tr>
<td>5</td>
<td>Fixed asset turnover</td>
<td>17</td>
<td>Sales to end capital</td>
</tr>
<tr>
<td>6</td>
<td>gross income to sales%</td>
<td>18</td>
<td>Sales to market value</td>
</tr>
<tr>
<td>7</td>
<td>Inventory to assets Ratio%</td>
<td>19</td>
<td>Stock price before the annual general meeting</td>
</tr>
<tr>
<td>8</td>
<td>Inventory turnover</td>
<td>20</td>
<td>Total asset turnover</td>
</tr>
<tr>
<td>9</td>
<td>Net sales</td>
<td>21</td>
<td>Total debt</td>
</tr>
<tr>
<td>10</td>
<td>Operating income</td>
<td>22</td>
<td>total debt to total assets ratio%</td>
</tr>
<tr>
<td>11</td>
<td>Price to earnings Ratio</td>
<td>23</td>
<td>year</td>
</tr>
<tr>
<td>12</td>
<td>Profit after tax</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Data selection
Financial data for 23 different chemical companies in Tehran Stock Exchange with 23 independent variables and a dependent variable (stock cash dividends) of 2006 to 2010 years were collected from various sources. A total of 108 samples were collected. Independent variables are shown in alphabetical order in Table 1.

Data cleaning and preparation
The second stage is data cleaning. At this stage, the data that their independent variables information were incomplete or were not calculated are removed. One of the criteria used to evaluate a classification, is the error rate, which are of various types. In general we can't properly judge about abilities of algorithm by comparing the errors that is calculated on the learning data. The error rate on learning data is usually lower than the error rate on that have not been seen in the learning process. With this reasoning, learning error can't be used to compare the two algorithms. This is because for more complex models, classifications that usually have more parameters, the boundaries are more complicated. This reduces the error on the learning data compared with simpler models (Abbas Kia, 2009). So in addition to learning data sets, a set of data for evaluation is required. Learning data is used for the model training and evaluation data is used for calculate the algorithm error rate on data that has not seen. In this paper, a random 70% data as training data and the remaining 30% was considered as evaluation data. In order to prepare data for training and evaluation with estimates, initially each of the variables is normalized using the following equation to reduce the impact of large numbers:

\[ \tilde{S}_i = \frac{(S_i - S_{\text{min}})}{S_{\text{max}} - S_{\text{min}}}, \quad i = 1, \ldots, 290 \]  

(1)

That \( S_{\text{min}} \) and \( S_{\text{max}} \) shows the minimum and maximum value of the variable and \( S_i \) shows the normalized value of \( S_i \).

Choice of influencing factors on predict stock cash dividends
This section introduces the approach to choice of effective components of prediction of stock cash dividends, using a combination of multi-layer perceptron neural network with binary genetic algorithm. Due to the neural network algorithm being used, at first it is discussed how to implement a neural network and then the method to predict the factors affecting stock cash dividends using binary genetic algorithm, will be discussed. Multi-layer perceptron neural network

How to code variables: the main purpose of this review is to provide a model to identify the variables that have more influence on decision-making, through the elimination of these low-impact and non-impact variables on the output, to improve the forecasting process. Variables should be coded so that they are considered individually and combined (Makvandi et al., 2008). The number of input variables in the present analysis is 23. These variables are coded in binary method into a string. To achieve the desired result, we perform the following way: An n-bit string (n is the number of variables) can be prepared. In this string, each bit represents a variable, the presence of these variable on the final result means bit=1 and if it is not present then bit=0.

Type of network and the learning rule: an artificial perceptron multilayer neural network used in the present model with error back propagation learning rule. The number of neurons and layers: the network consists of three layers with one neuron in the output layer and the middle layer, has respectively 10 and 5 neurons. Number of network layers and middle layer neurons is determined using the method of trial and error (Fig.3).

Stimulus functions
Since it is necessary to calculate the moment derivative we need to use stimulus functions that can be derived. In the present context hyperbolic tangent function is the best function. This function is as follows (Makvandi et al., 2008):

![Fig. 3. The number of neurons and layers](image)
Error function: error functions in neural networks, in terms of meaning, do not differ much from each other. Hence, for this case, mean square error is considered (Makvandi et al., 2008).

\[
MSE = \frac{\sum (A_i - F_i)^2}{n}
\]

That \( A_i \) is the actual value and \( F_i \) is the predicted value. Neural network test: to test network performance and determine fitness value, percentage of data that previously did not participate in the training process, acted as test data, and measures the prediction and generalization power of neural network (Makvandi et al., 2008).

Choice of influencing factors on stock cash dividends predict using binary genetic algorithm

As previously described, the aim of algorithm of the present model is to identify the combination of variables that mostly influences the prediction of the output variable (dividend shares). In the process of implementing this algorithm, the fitness value of individual members of each generation is calculated, according to fitness values, the next generations are produced by application of 3 operator, selection, transplantation and mutation. This random search process, continues until the termination criterion is achieved. The general algorithm of this process presented in Fig. 4.

**Genetic Algorithm Implementation details**

String presentation: In this study, the variables are binary encoded with fixed length strings. Due to the discontinuousness nature of the variables used, each bit corresponding to each generation of chromosomes, represents one of the used variables (Abbas Kia, 2009).

Fitness calculation: For evaluating the appropriate answers and a natural choice, the need for an index of the appropriate answers from improper. This index is an objective function, which can be selected based on a mathematical model or a computer simulation or qualitative criteria, which better answer can be recognized by it. The fitness function can be the mean square error (MSE). Whatever the MSE be less for a chromosome or an answer, the answer is more appropriate. In this study, the fitness function has been equaled to reverse the errors of the neural network obtained from the training neural network in any string (Akhbari, 2008).

Population size: population size is often one of the affecting parameters the efficiency of the algorithm. For example, if the population size is considered small, may lead to premature convergence. And if it is considered large, the algorithm execution time may be too much. In the used algorithm in this study, the population size of 50 has been considered (Akhbari, 2008).

2- Operators of genetic algorithm

Selection: parent chromosomes are selected randomly of initial population, so that the answers have higher fitness values, more selection are possible. In other words, the better answer is preferred to the worse answer (Abbas Kia, 2009).

Crossover: The most important operator in genetic algorithm is crossover operator. Crossover is process which the old chromosomes are combined and mixed together to create a new generation of chromosomes. The pairs that were considered as parents in selection part exchange their genes and create new members.
Crossover causes the fragmentation or loss of genetic variation is eliminated. Because it allows the appropriate genes find each other. Here, the single point crossover is used. In single point crossover at first the pair of parent chromosome is cut at appropriate point along the string then parts of the cut-off point change together. This creates two new chromosomes that any point of it, inherit genes from parent chromosomes (Abbas Kia, 2009) (Fig.5).

Mutation: mutation is the third operator in genetic algorithms. In genetic algorithm, after the creation of a new member each of its gene may be mutated. In mutation, a gene may be eliminated from the population, or to be added. The mutation operator behavior is so that for every person in the set mutation probability, which is usually less than one percent, is examined. If the mutation should be done a bit of chromosome randomly selected and its value is converted from zero to one, or vice versa (Abbas Kia, 2009) (Fig.6).

Convergence: In every iteration algorithm, the obtained population has to evaluated, in terms of convergence. For this purpose, an index can be defined as the ratio of minimum value of answer to the average value of answer in the population and was compared with predefined value, and increase mutation probability, if the population id converged (Akhbari, 2008).

**Research findings**

As previously described, 23 independent variables from 23 companies during years 2006 to 2010 were chosen in total, 108 samples were selected. In order to prepare the data for training and evaluation, at first each of the variables are normalized using equation (1) to reduce the impact of large numbers. To evaluate the regression models, two criteria’s have been used, mean square error (MSE) and root mean square error (RMSE), that is calculated using the following relations:

\[ RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - d_i)^2} \]

That \( d_i \) and \( y_i \) are respectively real dividends and predicted dividend of samples by each algorithm and \( n \) is the number of samples. The much MSE and RMSE values are closer to zero; the prediction algorithm is closer to reality.

**Influencing factors on predict stock cash dividends and results**

<table>
<thead>
<tr>
<th>Table 2. Parameters of genetic algorithm applied on model</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type/amount</strong></td>
</tr>
<tr>
<td>-------------------------</td>
</tr>
<tr>
<td>50</td>
</tr>
<tr>
<td>Single point</td>
</tr>
<tr>
<td>0.80</td>
</tr>
<tr>
<td>0.010</td>
</tr>
<tr>
<td>Roulette wheel</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3. The results of training data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number</strong></td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4. Independent variables influencing the prediction of stocks cash dividends with genetic algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Inventory turnover</strong></td>
</tr>
<tr>
<td>-------------------------</td>
</tr>
<tr>
<td>Price to earnings Ratio</td>
</tr>
</tbody>
</table>

RMSE is the root mean square error and is calculated using the following relation:

\[ RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - d_i)^2} \]

That \( d_i \) and \( y_i \) are respectively real dividends and predicted dividend of samples by each algorithm and \( n \) is the number of samples. The much MSE and RMSE values are closer to zero; the prediction algorithm is closer to reality.
neural network with these variables. The neural network that is designed for this experiment can be like as Fig. 7. Fig.8 presents MLP neural network with 23 variables selected for prediction.

Now neural network with training data (14 independent variables in Table 3) is taught and then evaluated data affected to the trained neural network. Table 5 and 6 show the results of training and assessment data for this model. In the second stage, all 23 independent variables listed in Table 1 are used to predict stock cash dividends. Now neural network with training data is taught and then evaluated data affected to the trained neural network. Table 7 and 8 show the results of training and assessment data for this model.

According to the results, the performance of model 1 is better than model 2. This suggests that a combination of genetic algorithms and neural network model can lead to a substantial increase in explanatory power of the model. Thus the research hypothesis is accepted.

Conclusion
The overall aim of this study, is selecting the affecting components on forecast cash dividends of stock, using neural networks and genetic algorithms. The results show that, combining neural networks and genetic algorithms, in order to select the optimal parameters, significantly increases the predictive power compared to the only use of neural networks.

It should be noted that artificial neural networks, are known as the black box test, meaning that despite of the strength of these models to recognize relationships between variables, they do not show the user the form of this relationship (Reber et al., 2005).

Since only fourteen variables of twenty-three independent variables in this study were selected by genetic algorithm for training neural networks, it shows that the not elected variables do not have significant effects on cash dividends in studied sample.

Research suggestions
We suggest the following topics for future studies:
1. Other methods of artificial intelligence such as artificial-fuzzy networks, other genetic algorithms or other hybrid models can be used to forecast dividends.
2. The research has done in an active industry in stock exchange, so it can be done in a larger level, or in all accepted companies in stock exchange.

References