A Hybrid Method Based on Neural Networks and a Meta-Heuristic Bat Algorithm for Stock Price Prediction

Marjan Golmaryami\textsuperscript{1*}, Marziyeh Behzadi\textsuperscript{1} and Marziyeh Ahmadzadeh\textsuperscript{1}
\textsuperscript{1}Department of Computer Engineering and Information Technology, Shiraz University of Technology, Shiraz, Iran

\textsuperscript{*}Corresponding Author’s E-mail: m.golmaryami@gmail.com

Abstract

Recent developments in the stock market have created an urgent need for efficient methods to help stockholders take appropriate decisions about their stocks. Since large fluctuations occur in the stock market over time and there are many parameters which influence this, it seems difficult to make good decisions that are also well-timed. The purpose of this study is to apply artificial neural networks (ANNs), which can deal with time series data and nonlinear parameters, to predict the next day’s stock price. This research has trained the proposed ANN with a meta-heuristic bat algorithm which has a fast and powerful convergence. The recommended method has been applied to stock price forecasting for the first time. This work has used a seven-year dataset of a private bank stocks in order to prove the performance of the suggested method. After data pre-processing, three types of ANNs (back propagation-ANN, particle swarm optimization-ANN and bat-ANN) were employed to predict the stocks’ closing price. Afterwards, MATLAB was used to evaluate the performance of these three methods by scoring the target of the mean absolute percentage error (MAPE). This paper indicates that the bat algorithm adjusts the weight matrix of ANN more precisely than the two other algorithms. The results may be adapted to other companies’ stocks.

Keywords: Artificial Neural Networks, Back Propagation Algorithm, Particle Swarm Optimization Algorithm, Bat Algorithm, Stock Market, Forecasting.

I. Introduction

Since the main goal of investment is to achieve expected profits and capital growth over time, the most important thing is to buy stocks at a low price and sell them at a higher price, which involves stock price prediction. Stockholders should take many factors into consideration in stock price analysis, which are classified into two categories: 1) fundamental analysis and 2) technical analysis. Fundamental analysis uses real data and considers factors such as inflation, unemployment, investment, the economic growth of a company, etc., while technical analysis uses historical data and takes mathematical indexes into consideration [1, 2]. There are two general methods for predicting stock prices: 1) statistical methods and 2) artificial intelligence (AI) methods. Statistical methods include ARIMA\textsuperscript{3}, STAR\textsuperscript{4}, etc. which assume linearity and normality, while in stock markets these assumptions may not exist. Thus, it is better to use AI methods which have no restrictive assumptions. Artificial neural networks, evolutionary population-based algorithms and fuzzy systems are some examples of AI methods [3]. In general, it is difficult to find an algorithm which performs well in all circumstances and all applications at any time [4]. ANNs are one of the best methods for forecasting stock prices because they learn by experience and they can also learn any variations in hidden rules in time series data, and use these to forecast the future. They follow the structure and
function of the brain, and simulate many of the brain’s abilities such as pattern recognition, generalization based on observations, etc. In recent decades, many researchers have considered using ANNs in order to solve optimization problems and have studied ANNs training methods, structure designs and their applications [5]. Various evolutionary population-based algorithms such as genetic algorithm [6], ant colony algorithm [7], honey bee algorithm [8], cuckoo search algorithm [9], teaching learning-based algorithm [10], etc. are employed in the ANNs training phase. This study has looked at ANNs training techniques and their application to the stock market. The main purpose has been to improve the neural network training algorithm based on the meta-heuristic bat algorithm (BA), the superiority of which over the back propagation (BP) algorithm and the particle swarm optimization (PSO) algorithm has been demonstrated.

BA can modify the weight matrix of ANNs better than the two previous algorithms and in discovering optimum solutions it converged faster than those two algorithms.

Since stock market data types are time series and many factors affect stock market behaviour, scientists have made great efforts to find new methods which perform well in stock price prediction. Time delay neural networks and adaptive time delay neural networks which have been optimized by genetic algorithm (GA) [11], have a better performance than the standard mode of these two networks. Although these networks work with sequential data and consider the dynamic nature of input data [12], their functioning is weak compared to recurrent neural networks [13]. On the other hand, a neural network model which is trained by GA based on reinforcement learning, has forecasted the next day’s stock price more precisely than recurrent neural networks [6], however, the problem is the low rate of convergence of the GA algorithm.

Yang and Chen utilize a back propagation artificial neural network (BP-ANN) to predict stock prices in [2]. They separate the variables that have an effect on stock prices from the others and put them into the network as input variables, and assume a two-layer structure with three hidden nodes as the neural network structure. But, if they had used a relation between the number of input neurons and hidden neurons like the case in [14], their model would have had a better performance. However, the deficiency of these two models is evident when confronting many changes in the stock market. This is due to the BP algorithm’s characteristics which do not have a proper efficiency when there are many variations in the stock market.

An ANN which is trained by a PSO algorithm (PSO-ANN) surpasses the BP-ANN in stock exchange forecasting [15]; however, the PSO’s defects are its high dependency on initial parameters and its trapping in local optima.

To tackle the problems mentioned above, this study has proposed a method which is based on ANNs and a meta-heuristic bat algorithm. This method solves those issues and increases the degree of certainty of the stock price prediction. BA’s abilities are its low dependency on initial parameters and fast and powerful convergence [5]. In fact, BA engages the features of PSO and simulated annealing algorithms; in addition it makes some modifications to the searching method in order to find optimum solutions and this is BA’s superiority compared to the two previously explained algorithms [16].

The rest of this paper is organized as follows: in the second section, the existing research gap is described. Section three presents the suggested method. The experimental design, the analysis of results and discussion are presented in section four, and finally the fifth section presents the research conclusions and future works.

II. Research gap

Due to the variability of stock market parameters over time and the nonlinear relationship between them, it is hard to predict stock prices. Thus, the necessity of utilizing AI techniques is evident. Among the AI techniques, the ANNs technique has been studied which has many benefits
such as its learning capability, its generalization and its adaptability; furthermore, it is powerful in solving hard problems [17, 5].

In BP-ANN the network is trained as follows: at first random values are assigned to weight and bias matrices, then the sum of the weighted inputs is entered into an activation function, afterwards the calculated outputs are compared with the expected values (values in the data set) and the error rate is computed which is used in the modification of the weight matrix [18, 19]. This study, based on experiments, assumes a structure for ANN as shown in Fig. 1. It is a multi-layer perceptron (MLP) with two layers and contains four input neurons, one hidden layer with five neurons, and one output neuron.

Despite the advantages of ANN when it is trained by BP algorithm, it may be trapped in local optima in a complex search space; moreover it is not able to forecast the price well when there are many changes in the stock market.

One of the AI tools which produces promising results is swarm intelligence. Swarm intelligence consists of a population of simple agents interacting locally with one another and with their environment, and all together trying to achieve a single goal. One of the swarm intelligence algorithms is PSO algorithm, which is population-based. It is inspired by the swarm behaviour of birds and fish. Each member of this group has access to only a limited amount of local information such as position, movement direction and the speed of its neighbours. The PSO algorithm is initialized with random particles and seeks the optimum solution by updating the particles iteratively in the problem space. In every iteration, each particle is updated with regard to these cases: the local best solution that the particle has achieved until that point with its fitness and the other case is the global best solution with its fitness [9, 20]. In an ANN, training means determining the optimum values of the weight and bias matrices which are acquired by means of the particles’ position in the PSO algorithm. One of the known issues in PSO algorithm is converging the whole population to a local optimum in a complex search space and as a consequence a premature convergence occurs [21]. To overcome the problems mentioned and to increase the accuracy of prediction in the stock exchange, this study proposes a solution that is discussed below.

III. Proposed method

This section describes the original BA and then explains the suggested method. BA is a new meta-heuristic population-based algorithm that is inspired by the behaviour of bats when they find their prey. In this algorithm, possible solutions are presented by the position of the bats.

A more appropriate performance of BA is demonstrated compared to the other evolutionary algorithms, such as GA and PSO algorithm [5], shuffled frog leaping algorithm [22], honey bee algorithm [23] and ant colony algorithm [24]. BA operates based on the following assumptions [16]:

1) bats use echolocation to sense distance, and recognize the difference between prey and food in a number of ways; 2) each bat with position \( x_i \) flies randomly with velocity \( v_i \) and produces pulses with frequency \( f_i \) and loudness \( A_i \) and a rate of pulse emission \( r_i \).

They can adjust the frequency of their pulses automatically and also adjust the rate of pulse emission based on the proximity of their prey; 3) the loudness can change in different ways; we suppose it alters from a large positive value \( A_0 \) to a minimum fixed value \( A_{min} \). In the beginning, BA starts with a random population of bats, subsequently at each step, in order
to update the position of each bat, the following formulas are used\[16\]:

\[
v_i^{\text{new}} = v_i^{\text{old}} + (x_i - x_{\text{best}}) \times f_i
\]

(1)

\[
x_i^{\text{new}} = x_i^{\text{old}} + v_i^{\text{new}}
\]

(2)

\[
f_i = f_{\text{min}} + \varphi_1 \times (f_{\text{max}} - f_{\text{min}})
\]

(3)

\(x_{\text{best}}\) is the position of the best bat, \(\varphi_1\) is a random value in the range \([0, 1]\), \(f_{\text{max}}\) and \(f_{\text{min}}\) are maximum and minimum values of the frequency that are assumed in this paper as being 1 and 0, respectively. The initial value of the frequency of each bat is selected from the range \([f_{\text{min}}, f_{\text{max}}]\). \(f_i\) is applied to control the velocity and scope of movement of the bats.

Afterwards in the local search, each bat uses a random walk in order to generate a new solution. To accomplish this, each bat produces a random number \(\beta\). If \(\beta\) is greater than the rate of the pulse emission, the new solution is generated by equation 4, otherwise it is generated by equations 5 to 8.

\[
x_i^{\text{new}} = x_i^{\text{old}} + \mu \times A_m^{\text{old}}
\]

(4)

\(\mu\) is a random value in the range \([-1, 1]\) and \(A_m^{\text{old}}\) is the mean value of the loudness of all bats. Here, in order to optimize the generated solution in the case that \(\beta\) is not greater than \(A_i\), a modification method is proposed. The main objective of this modification is to increase the diversity of the bat population by utilizing mutation and crossover which help to improve the search efficiency. Therefore for each bat \(X_i\), three bats \((X_{k1}, X_{k2}, X_{k3})\) are selected randomly in which \(i \neq k1 \neq k2 \neq k3\). Now by using the mutation and crossover operators, two below improved solutions are produced:

\[
X_{\text{opt1}} = X_{k1} + \alpha_1 (X_{k2} - X_{k3})
\]

(5)

\[
X_{\text{opt1}} = [X_{\text{opt1,1}}, X_{\text{opt1,2}}, \ldots, X_{\text{opt1,n}}]
\]

(6)

\(n\) is the dimension of this problem.

\[
X_{\text{opt2}} = \begin{cases} 
  X_{\text{best,i}}, & \text{if } \alpha_2 < \alpha_3 \\
  X_i, & \text{otherwise}
\end{cases}
\]

(7)
\[ X_{best} = [X_{best,1}, X_{best,2}, \ldots, X_{best,n}] \]  

\[ \alpha_1, \alpha_2 \text{ and } \alpha_3 \text{ are random numbers in the range [0,1]. The best solution among } X_{opt1}, X_{opt2} \text{ and } X_i \text{ is replaced with } X_i. \]

The new generated solution is acceptable if the following two conditions are met:

\[ (\beta < A_i) \land (f(x_i) < f(x_{best})) \]  

\[ f(x_i) \] is the objective function value for \( x_i \). If the new solution is acceptable, the loudness and the rate of pulse emission are updated through the formulas presented below:

\[ A_i^{new} = \alpha \ast A_i^{old} \]  

\[ r_{i}^{new} = r_{i}^{0} \ast [1 - \exp(-\gamma \ast t)] \]

\( \alpha \) and \( \gamma \) are constant values, \( r_{i}^{0} \) is the initial rate of the pulse emission and \( t \) is the number of each iteration. This study employs the explained BA in order to modify the weight matrix of ANN. In bat-ANN, at first the weight matrix is initialized with the initial population of bats, and it is passed to ANN in order to start the training phase; in the next step, BA specifies the best solution by means of the results obtained from the neural network. Following that, a local search is performed to discover new solutions. Each new acceptable solution will be replaced with the best solution if it is optimal and this process continues until the termination criteria are met. Thus, the optimal values for the weight matrix are found. Fig.2 shows the pseudo code of this procedure.

![Begin](Begin)

BAT is initialized then passes its first population to ANN as weight’s values

Load data

ANN starts training and computes the accuracy of the model

Bat finds the initial best solution by means of the ANN’s results

While \( i < \) Max number of iterations

![Fig. 2](Fig. 2. Pseudo code of Bat-ANN)

**IV. Experimental results and discussion**

**A. Data Preprocessing**

This study has used 78 months of data from a private bank in Iran which has been registered on the stock exchange in Tehran (from March 2005 to August 2011) as the main data set. This data set contains nine features as follows: date, last day price, high price, low price, close price, volume, number of transactions, difference compared to the day before and the percent of difference compared to the day before.
In data pre-processing, this study has checked missing values and outliers, normalized the data and reduced attributes by feature selection technique[25]. Missing values and outliers are omitted from the data set. Additionally, Feature selection has been performed by means of Weka. After that, only five features were chosen, which had greater effect on the output: last day price, high price, low price and volume (as input variables) and closing price (as a class variable). The data have been normalized using the following formula:

\[ x_n = \frac{x_i - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \]  

\( x_i \) is the actual datum, \( x_n \) is the normalized datum, \( x_{\text{min}} \) and \( x_{\text{max}} \) represent minimum and maximum values for each attribute respectively.

The data set is divided into three sections: 70% for training, 15% for testing and 15% for validation. Finally we have tested the proposed method on another two data sets, the petrochemical and the cement, in order to demonstrate the satisfying performance of the method.

B. Experimental Design

This study has used MATLAB to simulate the three discussed methods. In order to compare the error rate we have used equations 13 and 14:

\[ \sigma_i \% = \left( \frac{|P_i - T_i|}{T_i} \right) \times 100 \]  

\[ \text{MAPE}\% = \frac{1}{N} \sum_{i=1}^{N} \sigma_i \]  

\( \sigma_i \) is the relative error percentage, \( P_i \) and \( T_i \) are the predicted and the actual values of the ith data respectively, \( \text{MAPE} \) is the mean absolute percentage error and \( N \) is the number of data that must be forecasted. Equation 14 is the objective function which should be minimized in this problem using a bias and weight values adjustment. It is evident that a lower MAPE value represents a more accurate method.

The goal of this study is to predict the next day price of the private bank's stocks. In order to accomplish this goal, a two layer network with five hidden neurons was presumed experimentally. The network was trained by three explained algorithms separately using a training data set, after that it was tested using a test data set and the error rate was computed. Finally, a performance evaluation of the three methods (BP-ANN, PSO-ANN and Bat-ANN) was performed using a validation data set. Table 1 shows random thirty-day observations and their relative error as a sample. It should be mentioned that the validation phase has been performed using data from 15% of total, but to show some of the results as a sample, we have chosen thirty of outcomes randomly. As can be seen, relative errors decreased with the use of the suggested method on most days. For each method the MAPE was computed using the training data set which is shown in table 2. In the bat-ANN method, the MAPE was 2.9002 which had a lower error rate than the BP-ANN method with 3.6581 and the PSO-ANN method with 3.2150. As a consequence, bat-ANN forecasted the next day’s price more precisely.
TABLE I. Relative error comparison in three methods (the observations are about thirty random days.)

<table>
<thead>
<tr>
<th>Day</th>
<th>BP-ANNσij %</th>
<th>PSO-ANNσij %</th>
<th>BAT-ANNσij %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.4267</td>
<td>3.0936</td>
<td>2.0508</td>
</tr>
<tr>
<td>2</td>
<td>4.1568</td>
<td>3.3837</td>
<td>3.0469</td>
</tr>
<tr>
<td>3</td>
<td>0.26</td>
<td>0.6128</td>
<td>0.0409</td>
</tr>
<tr>
<td>4</td>
<td>6.7955</td>
<td>5.3523</td>
<td>4.175</td>
</tr>
<tr>
<td>5</td>
<td>1.7977</td>
<td>1.6045</td>
<td>1.3358</td>
</tr>
<tr>
<td>6</td>
<td>7.6094</td>
<td>7.0315</td>
<td>7.224</td>
</tr>
<tr>
<td>7</td>
<td>3.282</td>
<td>3.0383</td>
<td>2.7116</td>
</tr>
<tr>
<td>8</td>
<td>3.2549</td>
<td>2.3727</td>
<td>2.1732</td>
</tr>
<tr>
<td>9</td>
<td>5.6145</td>
<td>4.0329</td>
<td>3.889</td>
</tr>
<tr>
<td>10</td>
<td>1.5221</td>
<td>1.5334</td>
<td>1.0839</td>
</tr>
<tr>
<td>11</td>
<td>1.6223</td>
<td>1.93</td>
<td>1.0432</td>
</tr>
<tr>
<td>12</td>
<td>2.1756</td>
<td>2.2535</td>
<td>2.0668</td>
</tr>
<tr>
<td>13</td>
<td>3.0119</td>
<td>3.3209</td>
<td>2.0425</td>
</tr>
<tr>
<td>14</td>
<td>2.9987</td>
<td>1.0179</td>
<td>1.2112</td>
</tr>
<tr>
<td>15</td>
<td>3.169</td>
<td>2.7221</td>
<td>2.0677</td>
</tr>
<tr>
<td>16</td>
<td>4.2106</td>
<td>4.1577</td>
<td>4.0002</td>
</tr>
<tr>
<td>17</td>
<td>4.439</td>
<td>3.9112</td>
<td>3.012</td>
</tr>
<tr>
<td>18</td>
<td>4.9584</td>
<td>3.2076</td>
<td>2.6815</td>
</tr>
<tr>
<td>19</td>
<td>6.6789</td>
<td>4.7654</td>
<td>4.0198</td>
</tr>
<tr>
<td>20</td>
<td>3.8327</td>
<td>2.0919</td>
<td>1.9557</td>
</tr>
<tr>
<td>21</td>
<td>1.8905</td>
<td>2.0221</td>
<td>1.3113</td>
</tr>
<tr>
<td>22</td>
<td>6.2315</td>
<td>4.2378</td>
<td>4.0319</td>
</tr>
<tr>
<td>23</td>
<td>2.6667</td>
<td>1.2134</td>
<td>1.0019</td>
</tr>
<tr>
<td>24</td>
<td>2.3347</td>
<td>2.7569</td>
<td>2.1008</td>
</tr>
<tr>
<td>25</td>
<td>1.9454</td>
<td>1.4377</td>
<td>0.9523</td>
</tr>
<tr>
<td>26</td>
<td>3.1442</td>
<td>2.7125</td>
<td>2.122</td>
</tr>
<tr>
<td>27</td>
<td>3.7178</td>
<td>2.2344</td>
<td>2.1239</td>
</tr>
<tr>
<td>28</td>
<td>3.7228</td>
<td>1.1954</td>
<td>1.032</td>
</tr>
<tr>
<td>29</td>
<td>4.2218</td>
<td>1.5803</td>
<td>3.0122</td>
</tr>
<tr>
<td>30</td>
<td>3.9158</td>
<td>2.0115</td>
<td>2.0006</td>
</tr>
</tbody>
</table>

TABLE II. Comparison in three neural network methods

<table>
<thead>
<tr>
<th>Methods</th>
<th>BP-ANN</th>
<th>PSO-ANN</th>
<th>BAT-ANN</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAPE</td>
<td>3.6581</td>
<td>3.2150</td>
<td>2.9002</td>
</tr>
</tbody>
</table>
Fig. 3 shows a comparison of MAPE among the three methods using the training data from three different data sets. Case 1 is the bank data set, in case 2, a petrochemical stock data set is used, and in case 3 we have used a cement company data set. As it can be seen bat-ANN has the minimum MAPE in all cases.

![Fig. 3. Performance comparison in three neural network methods in validation phase](image1)

Fig. 4 depicts a comparison of the actual close price and the close price predicted by three methods using the validation data set. As it illustrates, the proposed method (Bat-ANN) predicts the next day’s price closer to the actual next day’s price.

![Fig. 4. MAPE comparison among the three methods using three different datasets.](image2)
Conclusion

Choosing a neural network with an appropriate training algorithm and structure can provide a powerful tool for forecasting a time series. This study suggested a predictive method based on artificial neural networks trained by a meta-heuristic bat algorithm in order to predict the next day’s price of the stock exchange. ANNs have the ability to solve problems with nonlinear variables. This ability of ANNs along with the powerful search and fast convergence of the bat algorithm were used to detect optimum solutions in the search space. Furthermore a modification was applied to empower the search ability of the bat algorithm. In order to investigate the desired goal, all three algorithms were applied to the dataset. The results showed that the proposed method forecasted the next day’s price with less error than the two other methods.

Future Works

As a basis for future research we can consider the modification and optimization of the bat algorithm in order to perform better searches in more complex problem spaces and creatediversity in the bat population to generate better solutions, which can help achieve a better performance in a time series prediction.

References


\(^1\) Autoregressive integrated moving average

\(^2\) Smooth transition autoregressive