Vol. 1, No. 2, 2019, pp. 131-143.

Hybrid Artificial Neural Network with Extreme Learning Machine Method and Cuckoo Search Algorithm to Predict Stock Prices

Piping Prabawati¹, Auli Damayanti^{2,*} & Herry Suprajitno³

^{1,2,3}Department of Mathematics Faculty of Science and Technology Universitas Airlangga

*Corresponding author: auli-d@fst.unair.ac.id

Abstract. This research aims to predict the stock prices, using artificial neural network with extreme learning machine (ELM) method and cuckoo search algorithm (CSA). Stock is one type of investment that is in great demand in Indonesia. The portion ownership of stock is determined by how much investment is invested in the company. In this case, stock is an aggressive type of investment instrument, because stock prices can change over time. In this case, ELM is used to determine forecasting values, while CSA is applied to compile and optimize the values of weights and biases to be used in the forecasting process. After obtaining the best weights and biases, the validation test process is then carried out to determine the level of success of the training process. The data used is the daily data of the stock price of PT. Bank Mandiri (Persero) Tbk. the total is 291 data. Furthermore, the data is divided into 70% for the training process is a smany as 199 data and 30% for the validation test as many as 87 data. Then compiled pattern of training and validation test patterns is 198 patterns and 82 patterns. Based on the implementation of the program, with several parameter obtained the result of MSE training is 0.001304353, with an MSE of validation test is 0.0031517704. Because the MSE value obtained is relatively small, this indicates that the ELM-CSA network is able to recognize data patterns and is able to predict well.

Keywords: extreme learning machine, cuckoo search algorithm, artificial neural network, forecasting, stock.

1 Introduction

Stock is one type of investment that is in great demand in Indonesia. According to [1], stock can be defined as a sign of participation or ownership of a person or entity in a company. The portion of ownership is determined by how much investment is invested in the company. In this case, stock are an aggressive type of investment instrument, because stock prices can change over time. Changes in stock prices can be influenced by several factors such as fluctuations in the rupiah exchange rate, government policies, and projections of company performance in the future. In addition to being able to make stock investments generate high returns, these price changes are also able to provide high risk if investors are wrong in analyzing stocks. So, to minimize the risk of loss of stocks ownership, it is important for investors to conduct an analysis before investing in stock. According to [2], there are two approaches that can be done to analyze stocks,

fundamental analysis approach and the technical analysis approach. Fundamental analysis is a stock analysis carried out by estimating the intrinsic value of stocks based on fundamental information that the company has already publicized. According to [3], technical analysis is based on mathematical calculations, such as formulas and graphs by combining the movements of a data. This analysis can provide a picture or prediction in the future. In this analysis fundamental research is conducted on repetitive price movement patterns so that predictions can be obtained.

Forecasting can be done by using Artificial Neural Network (ANN). According to [4], forecasting methods involving ANN provide better forecasting results. In 2004, Huang et al. introduced a new training method, namely the Extreme Learning Machine (ELM). Extreme Learning Machine method is one of the new learning method of ANN. The superiority of the ELM method is that it has very fast learning speeds, as well as better generalization performance than traditional learning methods [5].

Cuckoo Search Algorithm is a meta-heuristic algorithm inspired from the nature of some species of cuckoo parasites that lay their eggs in other birds' nests host. There are three rules in the Cuckoo Search Algorithm. First, each cuckoo places one egg at a time, and dumps it in a randomly chosen nest. Second, the best nests with high-quality eggs will be carried over to the next generations. Third, The number of available host nests is fixed, and the egg laid by a cuckoo is discovered by the host bird with a probability of $P_a \in [0,1]$.

According to research conducted by [6], CSA has successfully completed the optimization problem in some cases as in the global job scheduling and optimization. Research on the use of algorithms Cuckoo Search (CS) in the ANN training has been done by [7]. CS algorithm has excellent ability to find the global minimum. In research by [8], Hybrid CSA and Levenberg-Marquadt (LM) able to obtain the best results compared with CS algorithm, and LM when applied individually. This result was obtained because the ANN was trained using the best set of connecting weights resulting from training with the Cuckoo Search Algorithm.

Based on the background above, a research will be conducted on the hybrid artificial neural network extreme learning machine method with a cuckoo search algorithm to predict stock prices. This research is expected to be able to obtain more optimal results in predicting stock prices.

2 Stock Prices

Stock can be defined as a sign of participation or ownership of a person or entity in a company. The portion of ownership is determined by how much investment is invested in the company [1]. Stock is one of capital market products that became one of the instruments for long-term investment. By buying company stocks, we become the owner

of the company. The advantages of the shares include: Capital Gain and the Dividend (Distribution of company profits) (Indonesia Stock Exchange). In addition to these advantages, according [8] there are several risks faced by investors with ownership of shares, such as:

1. Not getting dividends

The Company will pay dividends if its operating profit. Therefore, companies cannot distribute dividends if they suffer losses.

2. Capital loss

In stock trading activities, investors do not always get capital gains or profits on the shares they sell. There are times when investors have to sell shares with a selling price lower than the purchase price of the stock. Sometimes, to avoid potential losses greater and greater as the continued decline in the stock price an investor willing to sell their shares at low prices. This term is known as stop loss (cut loss).

3. The company went bankrupt or liquidated Liquidation risk occurs when a company owned by financier declared bankrupt by a court or is being dismissed. In this case, investor receive last priority after all liabilities of the company repaid from the sale of the company's assets.

3 Extreme Learning Machine

Extreme Learning Machine (ELM) is a new method of training artificial neural networks (ANN) and includes supervised training methods. ELM is a neural network architecture using a single-hidden layer feedforward Netwoks (SLFNs). [5] argue that pre-existing ANN methods have weaknesses, especially in terms of the pace of learning (*learning speed*). The main reason for the low rate of learning is for all of the parameters on the network is determined iteratively using a training method. The parameter in question is the weight and bias input that connects between layer one layer to another.

In the method of ELM, weight feedback and bias initially determined randomly. After that, the output of the hidden layer matrix used in the calculation of the final weights. Computing the final weights obtained by using Moore-Penrose Generalized Inverse. Thus, ELM has a fast learning rate and is capable of producing good generalization performance.



Figure 1. ELM Architecture

In Figure 1 describes a network with single hidden layer. In the network there are nodes in the input layer, nodes in the hidden layer and a node in the output layer. As for the weights between nodes in the input layer and the hidden layer node denoted by a, the bias of the nodes in the hidden layer is represented by b and weights between nodes in the hidden layer and the output layer node denoted by β . The ELM architecture is as follows:

To calculate the output of the hidden layer can use the equation [5]:

$$G(a_{*j}, b_j, x_*) = g(\sum_{i=1}^N a_{ij} x_i + b_j) \qquad j = 1, \dots, L$$
(1)

Whereas to calculate network output, the equation used is:

$$Y = \sum_{j=1}^{L} \beta_j G(a_{*j}, b_j, x_*)$$
 (2)

So for *P* input patterns and hidden layer as *L*, can be arranged a matrix *H* that contains the output of the hidden layer size $P \ge W$, the matrix can be expressed in the following equation :

$$H = \begin{bmatrix} G(a_{*1}, b_{*1}, x^{1}) & G(a_{*2}, b_{*2}, x^{1}) & \cdots & G(a_{*L}, b_{*L}, x^{1}) \\ \vdots & \vdots & \ddots & \vdots \\ G(a_{*1}, b_{*1}, x^{P}) & G(a_{*2}, b_{*2}, x^{P}) & \cdots & G(a_{*L}, b_{*L}, x^{P}) \end{bmatrix}$$
(3)

where:

 $\begin{array}{ll} a_{*j} & : \{a_{1j}, a_{2j}, \dots, a_{ij}, \dots, a_{Lj}\} \\ a_{ij} & : \text{Weights between } i\text{-th input and } j\text{-th hidden} \\ b_j & : \text{The bias of } j\text{-th hidden} \end{array}$

- g(.) : Activation function
- *Y* : The output value
- β_j : Final weight between *Y* and *j*-th the hidden
- *P* : The number of input patterns

The algorithm used during ELM training is as follows [9]:

- 1. Input training normalized data as the input node $(x_1, ..., x_n)$ and define the activation function and the number of neurons in the hidden layer.
- 2. Initialize input weights (a_i) and bias (b_j) randomly.
- 3. Calculate each output from the node in hidden layer and denoted by $G(a_{*j}, b_j, x_*)$. Then set a matrix H consisting of all the output from the node in hiden layer.
- 4. Final weight (β) of the hidden layer to the output layer is calculated using equation:

$$\beta = H^+ T \tag{3}$$

where:

- H^+ : Moore-penrose Generalized invers of the matrix H
- *T* : Target vector
- 5. Calculate all outputs at the node in the output layer with equation (2).
- 6. Calculating the error value at each node in the output layer.
- 7. Check the stopping condition.
- 8. To check the termination conditions can be done in two ways:
 - i. Limiting the number of iterations.
 - ii. Limiting the value of *error*.
- 9. If the error or iteration has not been fulfilled, then do step 2 and so on.

After conducting training using ELM, then will do the testing process with the following steps [5]:

- 1. Input normalized test data as an input node $(x_1, ..., x_n)$.
- 2. Initialize input weights (a_i) and bias (b_j) of the training algorithm.
- 3. Calculate each output from the node in hidden layer $G(a_{*i}, b_i, x_*)$.
- 2. Calculate all outputs node in the output layer with equation (2).

4 Cuckoo Search Algorithm

Cuckoo Search Algorithm is one of Nature-Inspired Algorithm, which is inspired by the nature of some species of cuckoo parasites that lay their eggs in other birds' nests host. There are three rules used in Cuckoo Search Algorithm (CSA) is formulated as follows [7]:

1. Each cuckoo lay one egg at a time, and throw eggs in the nest chosen randomly.

- 2. The best nest with egg quality as the solution to be used for the next generation.
- 3. The number of other birds' nests that are available are fixed, and bird nest owners can find cuckoo eggs with probability $P_a \in [0,1]$. In this case, bird nest owners can dispose of cuckoo egg or nest leaving the nest so the owner can build a new nest in a new location. Assuming that P_a is a small part of n the nest can be replaced by a new nest that has been obtained (with a new random solution in the new location).

According to [10], Cuckoo Search using two random walks, Lévy Flights Random Walk (LFRW) and biased Random Walk (BRW). Lévy Flights Random Walk (LFRW) is a random walk with stepsize based on the Lévy distribution. According to [11], Lévy Flights is a search for new solutions which are located around the best solution and carried out based on equation:

$$x_{i}^{(t+1)} = x_{i}^{t} + \alpha S (x_{best}^{t} - x_{i}^{t})k$$
(4)

where x_i^t is a nest to *i* iteration to *t*, $\alpha > 0$ is a parameter stepsize, *S* is Lévy flights of algorithms Mantegna, κ is a real number of random normal distribution with average 0 and standard deviation 1, x_i^t is the best nest on iteration *t*.

According [12], step length (S) can be calculated using

$$s = \frac{u.\sigma}{|v|^{\overline{\beta}}} \tag{5}$$

where β is a parameter, variables u and v are random real numbers that are normally distributed with an average of 0 and standard deviation 1, and the value of σ is determined based on :

$$\sigma(\beta) = \left(\frac{\Gamma(1+\beta).\sin(\pi.\beta/2)}{\Gamma(\frac{1+\beta}{2}).\beta.2^{\frac{\beta-1}{2}}}\right)^{\frac{1}{\beta}}$$
(6)

where Γ is a Gamma function. According to [13], the function

 $\Gamma: (0, \infty) \to \mathbb{R}$, is defined as

$$\Gamma(z) = \int_0^\infty (e^{-x} \cdot x^{z-1}) dx \tag{7}$$

Biased Random Walk (BRW) is used to find new solutions at random which is positioned further away from the best solution. The method used is to build a new solution using the

current solution as a basis and with two other randomly chosen solutions, and a new solution is obtained from the sum of new solutions with current solutions. BRW can be formulated as follows:

$$x^{t+1} = \begin{cases} x_i^t + R \left(x_{permutj^t} - x_{permutk^t} \right), & \text{if } r_j > P_a \\ x_i^t & \text{, if } r_j \le P_a \end{cases}$$
(8)

permutation random index j and k are nests to j and k of the nest population. rand1, R is a random real number at the interval [0,1] and P_a is the probability of cuckoo eggs found by nesting birds [11].

Steps to implement the Cuckoo Search Algorithm (CSA) is as follows [7]:

- 1. Define the objective function f(x), $x = (x_1, x_2, ..., x_n)^T$.
- 2. Generating initial population as *n* nest xi (j = 1, 2, ..., n).
- 3. Evaluate the objective function F_{i} .
- 4. Get cuckoo randomly (i) / generate a new solution with Lévy Flights.
- 5. Evaluate the objective function F_{i} .
- 6. Make a selection by choosing *j* between *n* nests randomly. If $F_i < F_j$ then replace line *j* with a new solution.
- 7. Turn of the worst nest, if $r_j > P_a$ the cuckoo builds a new nest with a biased random walk.
- 8. Determine and save the best nest.

5 Stock Price Forecasting using Hybrid ELM-CSA

The steps used in this paper are as follows:

- 1. Review literature on stock issues, forecasting (forcasting), artificial neural networks, Extreme Learning Machine (ELM) methods, Cuckoo search Algorithm (CSA), and hybrid ELM-CSA.
- 2. Implement hybrid ELM-CSA with the following steps:
 - a. Designing a data plan, the data used in this study is the daily data stock prices of Bank Mandiri (Persero) Tbk, which is obtained from [14]. In drafting the data, there are two things to be done is to normalize the data, which would then be divided into training data and test data.
 - Data for training is 70% of the total data.
 - Data for validation test as much as 30% of total data.
 - b. Make a pattern design data to be used as an input value in the ELM-CSA network. Because the stock price data obtained is daily except Saturday and Sunday, then the number of selected input nodes is five. The number of input nodes present the number of days per week with a target of 6-*th* day or the

first day of the following week. The data design for the training process (199 data patterns) is presented in Table 1.

Pattern	<i>x</i> ₁	<i>x</i> ₂	<i>x</i> ₃	<i>x</i> ₄	<i>x</i> ₅	Target
1	1st Data	2nd Data	3th Data	4th Data	5th Data	6th Data
2	2nd Data	3th Data	4th Data	5th Data	6th Data	7th Data
:	:	:	:	:	:	:
100	199th	200th	201th	202th	203th	204th
199	Data	Data	Data	Data	Data	Data

 Table 1. Design data for training process

while the data design for the validation test process is compiled using 82 data patterns.

c. Designing ELM neural network architecture

Design an ELM multi layer network using 5 neurons in the input layer with hiden layer neurons is equal to the number of neurons in the input layer as many as 5 neurons. While the output layer (Y) there is only one unit of output due to the forecasting process will only generate one value.

- d. Implement Hybrid ELM-CSA in the network training process. The stages of the training process in hybrid ELM-CSA is as follows:
 - i. Input training data pattern that has been normalized.
 - ii. Determine the parameters ELM-CSA. CSA such as number of nests (*x*), stepsize (α), nest turn probability (P_a), and β_{CSA} .and maximum iteration.
 - iii. Generates initial nest populations, by generating random numbers (0.1) as many as nests.
 - iv. Convert the nest into weights and biases ELM.
 - v. Calculating the value of the output of each hidden layer using equation (3).
 - vi. Activating the output value of the node in the hidden layer using sigmoid binary function.
 - vii. Calculate the final weight of the hidden layer to the output layer (β) using equation (3).
 - viii. Counting all network output value (*Y*) using equation (3).
 - ix. Check stopping condition. There are two stopping conditions used:
 - The number of iterations has been reached.
 - The MSE value obtained from the network training process has been reached.
 - x. If the MSE value obtained is less than the error limits, the nest with the best eggs *xbest* can be used as weights and biases in the process

of validation test. If the iteration has not reached the maximum iteration (max iteration) or MSE obtained is greater than the error limit, then do next process.

- xi. Updating the nest early position with Levy Flight Random Walk (LFRW) performed with the following steps:
 - 1. Take *nonbest* (x_i) nest.
 - 2. Getting a value of *u* by generating a real random number normally distributed N (0,1) to a number of nests.
 - 3. Determining v by generate randomly real number normal distribution N(0,1).
 - 4. Calculating the value of σ .
 - 5. Calculating the value of *step length* (S) using :

$$s = \frac{u.\sigma}{|v|^{\frac{1}{\beta}}}$$

- 6. Geting a value of *k* by generating a real random number normally distributed *N*(0,1) to a number of nests.
- 7. Obtain a new x_i value using the equation:

$$x_i^{(t+1)} = x_i^t + a S (x_{best}^t - x_i^t)k$$

- xii. Calculate the objective function by using ELM for each nest LFRW results
- xiii. Select one nest *j* between n random nests, if $F_{LFRW} < F_j$ then replaces the *j*-nest with a new nest resulting from LFRW.
- xiv. Do the worst nest turn by randomly generating a number [0,1] of a number of nests. Compare rand values where P_a . If $rj > P_a$, replace the worst nest with the Biased Random Walk (BRW) process. The worst turnover with BRW is done by the following steps :
 - 1. Generating R with random numbers (0.1) number of nests.
 - 2. Generates random integers as $permut_j$ and $permut_k$ nest indexes.
 - 3. Getting x_{BRW} value using the equation Mendapatkan :

$$x_i^{(t+1)} = x_i^t + R \left(x_{permutj^t} - x_{permutk^t} \right)$$

- xv. Calculate the objective function using ELM for each nest of BRW results.
- xvi. Comparing the worst nest turnover results with nests from the *j* nest of cuckoo. If it is $F_{BRW} < F_j$, then change the nest.
- xvii. Determining the best nest x_{best} .
- xviii. Repeating the process (2.d.ix) to (2.d.x).

The purpose of the training process the neural network is to determine the weight and bias the weights and biases will be used in the process of validation test..

e. Validation test.

Process validation test can be explained as follows:

- i. Input normalized validation test data.
- ii. Input obtained weights and biases of the training process.
- iii. Calculating the value of the network output (Y) and MSE.
- f. Denormalize data.
- 3. Creating a hybrid ELM-CSA program.
- 4. Implemented the program to predict stock prices.

6 Result and Discussion

This research was conducted by compiling a program using Java programming language. Then performed a simulation to predict stock prices. The initial step in the hybrid program ELM-CSA is input the required parameters. These parameters include the number of nests, stepsize, P_a , the value of beta (Bcsa), the limit of MSE, and Maximum iterations. In the training process input data is done with some variance of parameter values to obtain the smallest MSE value. With a limit of MSE = 0.0001, Maksimiterasi = 100, and Bcsa = $2 \in [1,2]$ [15]. The results of the implementation of the process of training and validation test can be seen in Table 2.

Max	Nest	Stepsize	Training	Validation Test	
neration			MSE	MSE	
100	5	0.1	0.001458338	0.003303124	
		0.25	0.001365264	0.003449715	
	20	0.1	0.001304353	0.003151770	
		0.25	0.001304915	0.003105631	

Table 2. Results of Implementation of the Training and Validation Test Process

Based on Table 2 it is known that the smallest MSE value from the training process is 0.001304353. This value is achieved by iterations of 100 iterations, number of nests = 20, and the value of stepsize = 0.1. From the results of the implementation, the best weights and biases will be used in the validation process. The best weight and bias of program implementation can be seen in Table 3.

Weight (α_{ij})				Bias(h)	Beta	
<i>x</i> ₁	<i>x</i> ₂	<i>x</i> ₃	<i>x</i> ₄	x_5	Dias (D_j)	(β)
-0.69029	0.297098	-0.55881	0.222784	0.880438	0.511819	-0.69029
0.182606	0.190154	0.448638	0.225482	0.169087	0.839029	0.182606
0.282654	0.435734	0.795923	0.297883	0.958096	0.954264	0.282654
0.518253	0.607969	0.244527	1.310136	0.997338	0.222733	0.518253
0.420486	0.361283	0.089238	1.757488	-1.94089	1.079453	0.420486

Table 3. Best Weight and Bias of Program Implementation

Based on the weight and bias values, the predicted value of the validation test results is obtained in Table 4.

Table 4: Vandation Test Results					
Pattern	Date	Prediction Value	Target	Error	
1	7 Mei 2018	6871.371	6600	0.04112	
2	8 Mei 2018	6617.607	6800	0.02682	
3	9 Mei 2018	6813.573	6725	0.01317	
4	11 Mei 2018	6809.287	7250	0.06079	
5	14 Mei 2018	7261.567	7300	0.00526	
6	15 Mei 2018	7312.189	7000	0.0446	
7	16 Mei 2018	6976.157	6700	0.04122	
8	17 Mei 2018	6704.336	6950	0.03535	
9	18 Mei 2018	6972.321	6775	0.02912	
10	21 Mei 2018	6871.5	6725	0.02178	

Table 4. Validation Test Results



Figure 2. Output Value of Test Results

Because of the value of average is good enough, it can be said hybrid ELM-CSA is able to generate forecasting stock price approached the actual value. Furthermore, stock price

prediction is carried out for the first day of the following week, September 17, 2018. The prediction process is carried out using the optimal weight and bias obtained through the training process (training MSE is 0.001304353). Predicted output values are presented in Table 5. To obtain the predicted results on September 17, 2018, performed by using the input data in the form of stock prices as much as five days in the previous week.

Table 5. Predicted Output Values			
Date	Prediction Value		
17 September 2018	6649.5812		

7 Conclusion

Based on the implementation of programs and simulation for several values of parameters, it is known that the smallest MSE value of the training process is equal to 0.001304353 with MSE value validation test of 0.0031517704. This value is achieved by iterations of 100 iterations, number of nests = 20, and the value of stepsize = 0.1. The values of weight and bias obtained from these parameter values are quite good, so that they can be used to predict stock prices in the future.

8 References

- [1] Darmadji, M. and Fakhrudin M., 2001, *Perangkat dan Model Analisis Investasi di Pasar Modal*, Jakarta : PT. Elex Media Komputindo.
- [2] Tandelin, E., 2010, *Portofolio dan Investasi Teori dan Aplikasi*, Edisi Pertama, Kanisius, Yogyakarta.
- [3] Utami T. Endah, 2010, *Cara Cerdas Berinvestasi via Onlne Trading*, Edisi Pertama, Trans Media Pustaka, Jakarta.
- [4] Herawati S., Djunaidy A., 2014, Peramalan Harga Minyak Mentah menggunakan gabungan Metode Ensemble Empirical Mode Decomposition (EEMD) dan Jaringan Syaraf Tiruan. Jurnal SimanteC, Vol.4, 60.
- [5] Huang, G.-B., Zhu, Q.-Y., & Siew, C.-K., 2006, *Extreme learning machine: Theory and applications*. Neurocomputing, *70*(1), 489-501.
- [6] Pentapalli, V.V.G, Varma, P. R. K., 2016, *Cuckoo Search Optimization and its Applications: A Review*, International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE). Vol. 5, 556-562.
- [7] Yang, X.-S., and Deb, S., 2009, *Cuckoo Search via Lévy Flights*, In Nature and Biologically Inspired Computing. World Congress on NaBIC (pp. 201-214). IEEE.
- [8] Purnomo, R., S., and Hariani, I., 2010, *Buku Pintar Hukum Bisnis Pasar Modal*, Jakarta: Transmedia Pustaka.
- [9] Sun, Z-L., Choi, T.-M., Au, K.-F., dan Yu, Y., 2008, Sales Forecasting using Extreme Learning Machine with Applications in Fashion Retailing, Decision Support Systems, 46, 411-419.

- [10] Lin, Y., Zhang, C., and Liang, Z., 2016, Cuckoo Search Algorithm with Hybrid Factor Using Dimensional Distance, Journal of Mathematical Problems in Engineering, Vol. 2016.
- [11] Civicioglu, P. and Besdok, E.,2013, Comparative Analysis of the Cuckoo Search Algorithm, In: Yang XS.(eds) Cuckoo Search and Firefly Algorithm. Studies in Computational Intelligence Vol 516, 85-113.
- [12] Mantegna, R.N., 1994, *Fast, Accurate Algorithm for Numerical Simulation of Levy Stable Stochastic Processes.* Phys. Rev E 49, 4677-4683.
- [13] Diethelm, K., 2004, *The Analysis of Fractional Differential Equation*, Springer, New York.
- [14] Bursa Efek Indonesia, <u>http://www.idx.co.id</u>, 5 November 2018.
- [15] Kaveh, A., Bakhshpoori, T., 2013, *Optimum Design of Steel Frames Using Cuckoo Search Algorithm with Levy Flights*, Struct Des Tall Spec Build 22 (13). 1023-1036.