

Optimization Artificial Neural Network Using Artificial Bee Colony in Letter Recognition Classification

I Gusti Ngurah Alit Indrawan^{a1}, I Made Widiartha^{b2}

^{a,b}Computer Science Major in Mathematics and Science Faculty of Udayana University
Unud Street, Bukit Jimbaran, Badung, Bali. Postal Code: 80364. Indonesia
¹alitindrawan71@gmail.com
²madewidiartha@unud.ac.id

Abstract

Artificial Neural Networks or commonly abbreviated as ANN is one branch of science from the field of artificial intelligence which is often used to solve various problems in fields that involve grouping and pattern recognition. This research aims to classify Letter Recognition datasets using Artificial Neural Networks which are weighted optimally using the Artificial Bee Colony algorithm. The best classification accuracy results from this study were 92.85% using a combination of 4 hidden layers with each hidden layer containing 10 neurons.

Keywords : Artificial Neural Network, Letter Recognition, Feedforward, Artificial Bee Colony

Keywords: Image Processing, Expert System, Information

1. Introduction

Artificial neural network or commonly abbreviated as ANN is one branch of science from the field of Artificial Intelligence. ANN is a tool for solving various problems, especially in areas that involve grouping and pattern recognition [1].

Many things can be classified using this ANN, one of which is the classification of letters. The introduction of letters needs to be done because it will be very useful in the current era of technological progress. Old documents that are still in paper format will certainly be damaged within a certain time so that documents in digital form are also needed. In addition, letter recognition can also be used to create letter recognition applications in children.

The data to be used in this study is a secondary dataset, the letter recognition dataset. This dataset consists of 20,000 data with 16 attributes with integer values between 0 to 15 and 1 class of letter classification results.

Peter W. Frey and David J. Slate wrote a study entitled "Letter recognition using Holland-style adaptive classifiers" which uses the same dataset as this study. In this study the classification was carried out using the Holland-style adaptive classifier algorithm by dividing 16,000 data as training data and 4,000 data as training data and obtained an accuracy of 80% [2].

Hussein Salim Qasim also conducted research using the same dataset as this research. The study, entitled Letter Recognition Data Using Neural Networks, successfully implemented Artificial Neural Networks to classify letters. Of the 20,000 data, only 1,000 data were used as samples and 1,000 of these data were divided into 3, namely 60% of training data, 20% of test data, and 20% of validation data. The accuracy of this study is 60% [3].

In this research, the classification of letters using Artificial Neural Networks will be carried out with optimization in weighting using the Artificial Bee Colony algorithm.

2. Reseach Methods

Data Collection

The dataset used in this research is secondary data, namely the letter recognition dataset. This dataset was obtained from UCI Machine Learning Repository [4]. This dataset has 20,000 data and consists of 16 attributes and 1 class. Each attribute in this dataset has an integer value between 0 and 15. While the classes in this dataset are between 'A' to 'Z'. The following is a table of the letter recognition dataset attributes :

Tabel 1. Dataset attributes

Attribute	Explanation
x-box	Horizontal position of the box
y-box	Vertical position of the box
width	Square Width
height	Box height
onpix	# Of pixels
x-bar	x pixel on the box
y-bar	y pixel on the box
x2bar	x variance
y2bar	y variance
xybar	Correlation of x with y
x2ybr	$x * x * y$
xy2br	$x * y * y$
x-ege	Point from left to right
xegvy	Correlation of x-ege with y
y-ege	Point from top to bottom
yegvx	Correlation of y-ege with x

The class in the dataset is initially of the character data type which is then converted to an integer. The 'A' class will be changed to 1, 'B' to 2, and so on until 'Z' becomes 26.

Dataset

In this research, not all datasets are used because some data intersect with each other because some letters have similar shapes, making it difficult for the classification process later. Only 5 letter classes are used, namely the letters 'A', 'E', 'I', 'O', and 'U'. The five letters are used because they have a level of similarity far enough so that they will be classified properly later. In total there are 3,878 data used by the division of 2,714 training data and 1,164 test data.

Neural Network

Artificial Neural Network is one of the artificial representations of the human brain that always tries to simulate the learning process in the human brain. The term artificial here is used because this neural network is implemented using a computer program that is able to complete a number of calculation processes during the learning process [5].

In this research, the network training that will be made will be carried out by using 4 combinations of the number of different hidden layers between 1 hidden layer to 4 hidden layers with each neuron on the same hidden layer of 10 neurons.

Artificial Bee Colony

Bonabeau et al. define swarm intelligence as an experiment to design algorithms or to distribute problem solving tools inspired by the collective behavior of the social colonies of insects and other animal societies. The term swarm generally refers to a collection of controls from interacting agents and individuals [6]. One example of the swarm intelligence algorithm is the Artificial Bee Colony Algorithm.

Artificial Bee Colony (ABC) is an artificial intelligence algorithm that can be used for optimization in artificial neural network training [7]. The Artificial Bee Colony algorithm has a strong ability to find optimal global results or global best [8]. The ABC Algorithm model has 3 groups of bees, namely: employed bees, onlookers, and scouts in the artificial bee colony.

There are 3 main stages on basic ABC. The first stage, generate initial solutions from random food sources. To update possible solutions, each employed bee chooses a candidate for a new food source position, which is different from the previous position. The new position of the food source is calculated by the following equation:

$$x_{ij} = \theta_{ij} + \varnothing(\theta_{ij} - \theta_{kj}) \quad (1)$$

Where:

x_i : solutio from θ_i .

θ_i : employed bee position i

θ_k : Neighbor employed bee from θ_i .

\varnothing : Random number from $[-1,1]$ where number from $i \neq k$

n : number employed bee

D : Dimension.

In the second stage, each onlooker bee chooses one food source obtained from the employed bee. The probability of a food source being chosen can be obtained from the equation below:

$$P_{ij} = \frac{F(\theta_i)}{\sum_{k=1}^S F(\theta_k)} \quad (2)$$

Dimana :

P_{ij} : chance to choose employed bee- i

$F(\theta_i)$: Fitness Value from employed bee- i

S : number employed bee

θ_i : position from employed bee

After choosing a food source, the onlooker bee goes to the selected food source and selects a new prospective food source. Then in the final stage, the limit is a limit that has been set in the ABC Algorithm cycle and controls the number of certain solutions that are not updated. Any food sources that do not increase beyond the limit will be abandoned and replaced with new positions and employed bees become scout bees. The new random position chosen by scout bee will be calculated via the equation below:

$$\theta_{ij} = \theta_{j \min} + \text{rand}(\theta_{j \max} - \theta_{j \min}) \quad (3)$$

Where :

rand: a random number between $[0,1]$

$\theta_j \max$: upper limit of the position source in the dimension j

$\theta_j \min$: the lower boundary of the position source in the dimension j

Weight Optimization

The initial weight of each neuron in the hidden layer network will be optimized using the Artificial Bee Colony algorithm. At first each weight on the hidden layer will be generated randomly by a neural network, then later it will be optimized by the Artificial Bee Colony algorithm.

3. Result and Discussion

After each of the 4 experiments performed with the number of hidden layers that differ with each hidden layer containing 10 neurons in the artificial condition network without optimization and a neural network with optimization, the following results are obtained.

Tabel 2. Result testing without optimization

Hidden Layer	1	2	3	4
--------------	---	---	---	---

Epocs	30	50	18	30
Time	00:03	00:10	00:12	00:19
Accuracy	39.63%	50.47%	82.11%	90.91%

Tabel 3. Result testing with optimization

Hidden Layer	1	2	3	4
Epocs	23	28	12	20
Time	00:12	00:07	02:15	01:19
Accuracy	42.99%	77.81%	89.59%	92.85%

From table 2 and table 3, it can be seen the results of testing artificial neural networks without optimization and with optimization using artificial bee colony. From the test results it can be seen that the number of training iterations or epocs on the optimized neural network is less than the artificial neural network requirements that are not optimized at the same number of hidden layers. In addition, the classification accuracy of optimized neural networks is always higher than the accuracy of classification of artificial neural networks that are not optimized.

4. Conclusion

The conclusions obtained from this study are as follows:

- Artificial Terms Network can be used to classify letter recognition datasets well depending on the number of hidden layer combinations and neurons used.
- The Artificial Bee Colony algorithm is proven to be capable of optimizing the weighting of artificial neural networks thereby reducing the number of training iterations and also improving classification accuracy.
- The best classification accuracy results obtained by 92.85% using a combination of 4 hidden layers with each hidden layer containing 10 neurons

The suggestions from the authors are:

- Using all datasets available in the letter recognition dataset, but it is necessary to pre-process the data first, thereby reducing noise in the data so that the data can be used optimally in future studies.
- Can use other Swarm Intelligence Optimazion for optimization on the artificial neural network, such as Ant Colony Optimazion, Particle Swarm Optimazion, and others.

References

- [1] D. Puspitaningrum, "Pengantar Jaringan Syaraf Tiruan," 2006.
- [2] P. W. Frey and D. J. Slate, "Letter recognition using Holland-style adaptive classifiers," *Machine learning*, vol. 6, pp. 161-182, 1991.
- [3] H. S. Qasim, "Letter Recognition Data Using Neural Network," *International Journal of Scientific & Engineering Research*, vol. 4, 2013.
- [4] D. a. G. C. Dua. (2017). [4] *Machine Learning Repository*. Available: <http://archive.ics.uci.edu/ml>
- [5] M. Andrijasa and M. Mistianingsih, "Penerapan Jaringan Syaraf Tiruan Untuk Memprediksi Jumlah Pengangguran di Provinsi Kalimantan Timur Dengan Menggunakan Algoritma Pembelajaran Backpropagation," *Informatika Mulawarman: Jurnal Ilmiah Ilmu Komputer*, vol. 5, pp. 50-54, 2016.
- [6] J. A. Bullinaria and K. AlYahya, "Artificial bee colony training of neural networks: comparison with back-propagation," *Memetic Computing*, vol. 6, pp. 171-182, 2014.

- [7] C. OZTURK AND D. KARABOGA, "HYBRID ARTIFICIAL BEE COLONY ALGORITHM FOR NEURAL NETWORK TRAINING," IN *2011 IEEE CONGRESS OF EVOLUTIONARY COMPUTATION (CEC)*, 2011, PP. 84-88
- [8] C. Ozturk and D. Karaboga, "Hybrid artificial bee colony algorithm for neural network training," in *2011 IEEE congress of evolutionary computation (CEC)*, 2011, pp. 84-88.