Early Detection of Dengue Hemorrhagic Fever (DHF) using Feed Forward Neural Network with Gravitational Search Algorithm Optimization

Nur Azmi Prasetyo, Hasbi Yasin, Budi Warsito
Department of Statistics – FSM, Universitas Diponegoro
Jl. Prof. Soedarto, SH., Tembalang, Semarang, 50275, Indonesia
nurazmiprasetyo@gmail.com

Abstract. Health is an important part of life, it is also an indicator for progress of country development. But, there are still some health problems found in Indonesia, one of it is dengue infection. Data from the Department of Health states that in 2016 there were 77.96 per 100,000 population (201,885 cases) with mortality rate is 0.79% (1,585 death cases). Clinical symptoms of dengue infection is so resemble with some other fevers, then the right diagnosis is very needed. Feed Forward Neural Network (FFNN) can be used to do an early detection to Dengue Hemorrhagic Fever (DHF) sufferer with training process to the existing data. In this research, we experiment the Gravitational Search Algorithm (GSA) as the Optimization Method to find the optimal weights of the FFNN. We use the medical record data of DHF sufferers who are getting treatment at Tugurejo District Hospital, start from January 2016 to December 2016. The accuracy of this research is 0.73188.

1. Introduction
Dengue Hemorrhagic Fever (DHF) is one of the health problems we could found in Indonesia. Data from the Department of Health states that in 2016 there were 77.96 per 100,000 population (201,885 cases) with mortality rate is 0.79% (1,585 death cases). This problem should be solved right and efficiently, then the country development could be faster. The challenge is that the success of the dengue infection handling is very affected by the right diagnose, prompt treatment, and the right handling. The clinic symptoms of dengue infection is very resemble with the other fevers, so the right diagnose is really needed [2].

Early detection of DHF could be conducted using Feed Forward Neural Network (FFNN). FFNN could be used because it can save the existing information to make it experience and is saved for a useful purpose. Various optimization methods have been collaborated to train FFNN, such as Genetic Algorithm (GA), Particle Swarm Optimization (PSO, etc). Lately, Gravitational Search Algorithm (GSA) comes up as a novel optimization method with some advantages. Rashedi et al. [3] state that GSA could find more optimal result than the other methods. Besides that, Kumar et al. [1] state that GSA has smaller memory usage than the other algorithms.

In this research, the use of GSA is as an optimization method for training FFNN to do the early detection on a DHF case. The result of this research is expected to get information from the existing data, and help the government to move forward in DHF handling.

2. Dengue Hemorrhagic Fever (DHF)
DHF is caused by the dengue virus from the Arbovirus B group, namely arthropod-borne viruses or viruses spread by arthropods. This virus belongs to the Flavivirus genus from the Flaviviridae family
3. Gravitational Search Algorithm (GSA)

Rashedi et al. [3] introduce a novel heuristic optimization algorithm called Gravitational Search Algorithm (GSA) to find the best solution using physical law. GSA is inspired by the Newton theory about gravitation: “Every particle in the universe attracts every other particle with a force that is directly proportional to the product of their masses and inversely proportional to the square of the distance between them” [4]. GSA can be considered as collection of agents (candidate solutions) which have masses proportional to their value of fitness function [5]. In generation process, the masses attract each other by the law of gravity around them.

Agent in GSA is defined as:

\[ X_i = (x_i^1, \ldots, x_i^d, \ldots, x_i^n), \quad i = 1, 2, \ldots, N \]  

where \( N \) number of dimension, and \( x_i^d \) defines the position of the \( i \)th agent at the \( d \)th dimension.

GSA is started by placing all the agents in a search space randomly. During the searching process, gravitation forces from agent \( j \) on agent \( i \) at the certain time \( t \) are defined as:

\[
F_{ij}^d = G(t) \frac{M_{pj}(t) \times M_{aj}(t)}{R_y(t) + \varepsilon} (x_j^d(t) - x_i^d(t))
\]  

where \( G(t) \) is gravitational constant, \( M_{pj}(t) \) is passive gravitational mass related to agent \( i \) at the time \( t \), \( M_{aj}(t) \) is active gravitational mass of the agent \( j \) at the time \( t \), \( R_y(t) \) is absolute distance between agent \( i \) and \( j \) at the time \( t \), and \( \varepsilon \) is small constant. \( G(t) \) and \( R_y(t) \) are calculated as follow:

\[
G(t) = G_0 \times e^{(-\alpha \times \text{iter} / \text{max iter})}
\]  

\[
R_y(t) = |X_i(t) - X_j(t)|
\]

where \( \alpha \) is descending coefficient, \( \text{iter} \) is the current iteration, and \( \text{max iter} \) is the maximum iteration.

Total force and acceleration of an agent are calculated as below:

\[
F_i^d(t) = \sum_{j=1, j \neq i} \text{rand} \cdot F_{ij}^d(t)
\]

\[
a_i^d(t) = \frac{F_i^d(t)}{M_i(t)}
\]

where \( \text{rand} \) is random number in the interval of \([0,1]\).

Then just like the other object, agents have the updated velocity and position, they are calculated as follow:

\[
v_i^d(t+1) = \text{rand} \times v_i^d(t) + a_i^d(t)
\]

\[
x_i^d(t+1) = x_i^d(t) + v_i^d(t+1)
\]

where \( \text{rand} \) random number form interval of \([0,1]\).

The mass of the agents is calculated as follow:

[7]. The main vectors of dengue fever are the mosquitoes Aedes aegypti (in urban areas) and Aedes albopictus (in rural areas) [7]. The Aedes mosquito can transmit the dengue virus to humans either directly, namely after biting a person who is experiencing virema, or indirectly after going through an incubation period in his body for 8-10 days (extrinsic incubation period). In humans, it takes 4-6 days (intrinsic incubation period) before becoming sick after the virus enters the body. In mosquitoes, once the virus can enter and multiply in the body, the mosquito will be able to transmit the virus during its life (infection). Transmission to humans can only occur when the body is in a viremia state, which is between 3-5 days [6].
\begin{align}
m_i(t) &= \frac{\text{fit}_i(t) - \text{worst}(t)}{\text{best}(t) - \text{worst}(t)} \tag{9}
\end{align}

where \( \text{fit}_i(t) \), \( \text{best}(t) \), and \( \text{worst}(t) \) defined as follow:

\begin{align}
\text{fit}_i(t) &= \sum_{k=1}^{m} \frac{q}{q} (o_i^k - d_i^k) \tag{10}
\end{align}

\begin{align}
\text{best}(t) &= \min_{j=1,\ldots,N} \text{fit}_j(t) \tag{11}
\end{align}

\begin{align}
\text{worst}(t) &= \max_{j=1,\ldots,N} \text{fit}_j(t) \tag{12}
\end{align}

where \( o_i^k \) is the final output of FFNN from the \( k \) th training sample, \( d_i^k \) is the desired output of FFNN from the \( k \) th training sample, \( q \) is the number of the training sample.

The mass should be normalized, defined as follow:

\begin{align}
M_i(t) &= \frac{m_i(t)}{\sum_{j=1}^{N} m_j(t)} \tag{13}
\end{align}

Here is the Figure 1 shows GSA steps to see it holistically:

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{gsa_steps.png}
\caption{GSA steps}
\end{figure}

4. Methodology
The data used in this research is obtained from Tugurejo District Hospital of Semarang, Indonesia. The data contains 414 medical records of DHF sufferer in 2016, with the label proportion of Dengue Fever : DHF : Dengue Shock Syndrome is 69.81% : 24.64% : 5.56% respectively. This data is including 5 features affecting the labels, those are : Age, Platelets, Initial Hematocrit, Enforcement Hematocrit, and Enforcement Hemoglobin.

This research using 3-Folds Cross-Validation method to make sure that the obtained result is the best result of GSA usage as optimization method. This research does not do any handling related to unbalanced data, it is conducted to measure the effectiveness of GSA as optimization method for FFNN specifically. The result is measured by classical Confusion Matrix : Accuracy, Precision, and Recall.
5. Result and Discussion
The FFNN architecture built in this research is considered by the character of the data. In this research, it is built with 5 units input layer for the 5 existing input variables, 15 units hidden layer, and three units output layer for the existing three labels. Figure 2 shows the architecture of the used FFNN.

![FFNN architecture](image)

**Figure 2. FFNN architecture**

This research conducting training result and testing result. The training result is shown by Table 1, and the testing result is shown by Table 2.

| Fold | Fitness (MSE) |
|------|---------------|
| 1    | 0.38552       |
| 2    | 0.40854       |
| 3    | 0.39554       |

Table 1. Training result each fold

Training result shows that the average of obtained MSE is 0.39653, with the smallest MSE is 0.38552 by the fold 1.

| Fold | Accuracy | DF Precision | DHF Precision | DSS Precision | DF Recall | DHF Recall | DSS Recall |
|------|----------|--------------|---------------|---------------|-----------|------------|------------|
| 1    | 0.71014  | 0.76423      | 0.26667       | NaN           | 0.93069   | 0.16       | 0          |
| 2    | 0.63768  | 0.74545      | 0.21429       | NaN           | 0.81188   | 0.24       | 0          |
| 3    | 0.73188  | 0.76378      | 0.36364       | NaN           | 0.96040   | 0.16       | 0          |
| Average | 0.69324  | 0.75782      | 0.28153       | NaN           | 0.90099   | 0.18667    | 0          |

Table 2. Testing result each fold

Table 2 gives us information that the best accuracy is when the FFNN is tested by fold 3, with the accuracy value of 0.73188. The rest value like precision and recall are affected by the unbalanced data as shown in the section 3.

6. Conclusion
The best accuracy value of the DHF early detection using FFNN optimized with GSA is 0.73188. The accuracy value is concluded as good value. The accuracy value proves that GSA could be used to optimize the classification problem using FFNN well. We suggest to do advanced research using GSA and to combine it with unbalanced data handling, or to combine it with hybrid method to obtain better result.

References
[1] Kumar V, Chhabra J K and Kumar D 2014 Automatic cluster evolution using gravitational search algorithm and its application on image segmentation *Engineering Applications of Artificial Intelligence* 29 93-103
[2] Purwanto P 2002 Pemeriksaan Laboratorium Pada Penderita Demam Berdarah Dengue *Media Litbang Kesehatan* 12(1) 14-19

[3] Rashedi E, Nezamabadi S and Saryazdi S 2009 GSA: a gravitational search algorithm *Information Sciences* 179(13) 2232–2248

[4] Newton I 1729 *In Experimental Philosophy Particular Propositions are Inferred from The Phenomena and Afterwards Rendered General by Induction* (Andrew Motte’s English Translation Published) 2

[5] Mirjalili S, Hashim M Z and Moradian S H 2012 Training feed forward neural networks using hybrid particle swarm optimization and gravitational search algorithm *Applied Mathematics and Computation* 218 pp 11125-11137

[6] Hadinegoro S R and Satari H I 2004 *Demam Berdarah Dengue* (Jakarta: Balai Penerbitan FKUI)

[7] Widoyono 2009 *Penyakit Tropis: Epidemiologi, Penularan, Pencegahan, dan Pemberantasannya* (Jakarta: Erlangga)