Artificial Neural Network and Fuzzy Logic Approach to diagnose Autism Spectrum Disorder

Nguyen Van Hieu[1], Ngo Le Huy Hien[2]

[1]*Doctor, Dept. of Information and Technology, The University of Science and Technology, Danang, Vietnam*
[2]*Researcher, Dept. of Computer Science, VN-UK Institute for Research and Executive Education, Danang, Vietnam*

---***---

Abstract – *Autism Spectrum Disorder (ASD) is becoming a big issue in numerous countries around the world which can even negatively affect human natural evolution. Even though autism can be diagnosed early - before 2 years old, most children were not diagnosed with ASD until the age of 4 because of its complex symptoms and ambiguous manifestation in infant's disorders. Applying science and technology into early autism diagnosis is of vital importance, especially when data mining branches and decision-making support systems are developing and achieving many accomplishments in various fields, medicine included. Contributing to those developments, the combination between the Artificial Neural Network (ANN) and Fuzzy logic has triggered a huge revolution in data mining and is able to solve a variety of problems. This paper is the elaboration on the method of employing this combination to facilitate the early diagnosis of ASD. The result of the paper shows that the aforementioned approach has the potential to be the fundamental basis of the supporting decision-making system in ASD researching and diagnosing.*

Key Words: Diagnosis of autism, Artificial Neural Network, Fuzzy logic, Medicine support systems

1. INTRODUCTION

Autism or Autism Disorder Spectrum is the general term that refers to the complex disorders in the development of human brain, by which people with ASD may have some troubles in communicating, studying, touching and interacting with other people. According to the report from Autism and Developmental Disabilities Monitoring (ADDM) Network in America, the percentage of children diagnosed with ASD is 1/68, while in 2009 it was 1/110[1]. ASD is becoming a top concern of many countries in the world. The percentage of ASD children is increasing over years regardless of their economic backgrounds and origins. ASD is not only the burden of families who have children suffering from this syndrome but also poses challenges to the national economy and could disrupt the human evolutionary process.

Because of its indefinite forms of manifestation, ASD covers a large spectrum of symptoms, from mild to severe. ASD's undefined causes make it hard prevent; therefore, early recognition is the linchpin of repelling ASD. The earlier the early ASD recognition is implemented, the more opportunity ASD victims have to integrate into our society and normally develop as healthy children do. With the advance of the data mining field, the process of applying machine learning algorithm has made remarkable steps in disease diagnosis. For instance, the percentage of cancer diagnosis carried by doctors was 79.9%, while this number reached 91.1% with the presence of science and technology [2]. There is no similar statistics in ASD diagnosing, but the high number of inaccurate diagnosed cases testifies for the lack of clinical quality and medical staff quantity. Most of the current ASD diagnosing methods are only effective for children after 2 years old. Early ASD diagnosis for children before 2 years old is usually based on the prediction from the combinations of many simple symptoms, which are sometimes not evident and even hard to implement in many cases. It is understandable to claim that early diagnosis of ASD is a very complicated nonlinear function and hard to observe. Therefore, we propose an approach using the combination of artificial neural network and fuzzy logic in order to establish a system of early diagnosis of ASD. The effectiveness of this method has been verified in many fields, including finance, geology, physics, medicine, as well as intelligent decision-making systems [5, 6, 7, 8]. Concentrating on this approach, this paper will be able to address the complexity of autism symptoms thanks to the ability to comprehend the complicated relations between inputs and outputs of the artificial neural network. Moreover, it is able to deal with multiple levels of ASD manifestation as well as the accuracy of the collected information from parents. The result of this paper shows the potential to apply artificial neural network and fuzzy logic in diagnosing ASD early and the accuracy of the diagnosis by using this combination. The later of this paper will be categorized into 5 parts: part 2 presents documentation overview, part 3 diagnosing model base on trial data, part 4 the establishment of the trial system, and part 5 the conclusion and development.

2. LITERATURE REVIEW

2.1 Diagnosis of ASD methods

ASD diagnosis method with the complex characteristics of ASD, several methods have been used to detect autism.

The first method is the classic one, which diagnoses autism based on a series of strange repetitive symptoms. However, this method has a low level of accuracy and is not suitable for diagnosing autism in infants before 24 months of age because the expression of autistic symptoms in younger children is unclear.

The screening questionnaire is widely used for diagnosis of ASD, some of which are the CARS (Childhood Autism Rating Scale), M-CHAT, M-CHAT-R, etc. [3, 4]. Those ASD rating scales usually have the common ground. They receive information about autistic symptoms via questions, then sum up the points and eventually conclude classify types and levels of autism. Take the CARS for example. By this scale, there are 3 levels of disorder: mild, moderate, severe. As one of the most popular scale, CARS is a system of 15 sections, each one has 4 primary levels and 3 secondary levels. Evaluators, who are children's caregivers like parents or teachers, need to appraise kids' behaviors and then score them corresponding to those described symptoms in 15 listed section by 7 levels, from 1, 1.5, 2..., to 4. Later, the sum of 15 sections' points will determine the level of autism which is measured by this scale: smaller than 30 – mild, from 30 to 36 – moderate, greater than 36 – severe. The current ASD diagnosis methods share some general features: they are easy to use and be able to recognize many cases of autism. However, these methods are incapable of grasping the complexity of ASD leading to a lot of false positive cases and false negative cases, which is mainly caused by the property of the rating scale and the accuracy of the evaluators' answers. Moreover, diagnosing using rating scale like the CARS is not suitable to apply to infants. In order to fix those drawbacks, data mining algorithms, namely classification techniques, will be applied to support improve the accuracy of early autism diagnosis. In this paper, the artificial neural network will be combined with fuzzy logic to handle these affairs.

2.2 Fuzzy logic, Fuzzy set

In human life in general and data mining field in particular, we often deal with ambiguous data which is incomplete and not entirely accurate, but it still needs to be analyzed to make decisions. Classical logic with truth value true – false is weak to solve above problems, therefore, in 1965, professor Lotfi Zadeh – Electronic dean of the California University have successfully developed fuzzy set theory and fuzzy logic system.

Classical set: $\forall \boldsymbol{u} \in \boldsymbol{U}, \boldsymbol{\mu_A}(\boldsymbol{u}) = \begin{cases} \mathbf{1}\ if\ \boldsymbol{u} \in \boldsymbol{A} \\ \mathbf{0}\ if\ \boldsymbol{u} \notin \boldsymbol{A} \end{cases}$ in which U is the universe of discourse.

If fuzzy logic extends the concept of classical logic, the Fuzzy set will be the extended version of the classical set, by that, if the truth value set of classical logic is {0, 1}, the truth value of fuzzy logic will be: $\{a \mid a \in \boldsymbol{R}\ and\ 0 < \mathbf{a} \leq 1\}$.

Each element of fuzzy set A needs a membership function to represent a degree of truth which defines how it belongs to set A: $\forall \boldsymbol{u} \in \boldsymbol{U}, 0 \leq \boldsymbol{\mu_A}(\boldsymbol{u}) \leq 1$, U is the universe of discourse and $A \subset U$.

Therefore, the Fuzzy set is able to present more problems than classical set, moreover, it can even present classical set when the value of membership function is 0 or 1. Similar to classical set, the Fuzzy set has various operators like union, intersection, complement, etc.

In the framework of this paper, the Fuzzy set is used in defining the input set and its accuracy.

2.3 Artificial Neural Network

Inspired by the operating principle of the biological neural network in animal's brain, the artificial neural network (ANN) is an estimate nonlinear function with the input is an unformatted or unclassified dataset. The aim of ANN is to figure out the rules of the dataset and solve classification problems, which are the core of data mining field or scientific recognition. Figure 2.1 represents the nature of ANN.

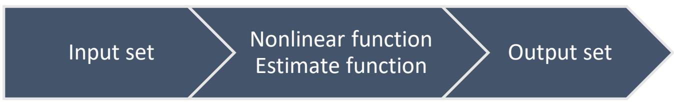

Figure 2.1. The nature of ANN

Depending on the purpose of the ANN, there are many types of different neural network, the most common of which is the multiple-layer straightforward ANN. This model is constructed from three main layers: input, hidden, and output. With the base components of the neural network are "neural cells", each layer contains a certain number of cells, and synapse is responsible for connecting the signals between them.

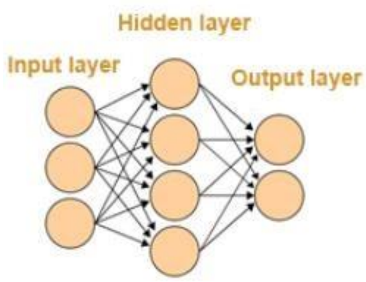

Figure 2.2. The construction of feed-forward ANNs

In order to manage the process of transmitting the signals between cells, activation functions will be applied because they are able to show the ability to transmit signals of a neural cell. Activation functions also determine whether one signal will be delivered to the next cell or not and how great the intensity of that transmission should be.

The common activation function used in this research is Sigmoid:

$$f(x) = \frac{1}{1+e^{-x}}$$

Depending on the number of output signals to determine the number of thresholds, with a simple binary ANN, there is one threshold. If f(x) is greater than threshold (0.5 commonly), the output of the neural cell is 1, if f(x) < 0.5, output is 0 oppositely.

Demonstrating ANN in mathematics, we design a simple ANN having 3 layers, the number of neural cell in each layer input, hidden, output respectively is: v, a, r. Activation function is f(x), we have:

- Content Convention:
  - J and k are the index used for layer hidden and output correspondingly.
  - I_{Xy} is the input of neural cell which is in layer X with the index of y.
  - Similarly, O_{Xy} is the out of neural cell which is in layer X with the index of y.
- Input set: I = {i1, i2, …, iv}.
- Weight set w^1 of the synapses between the input layer and hidden layer corresponding to hidden neural cell H_j:
  $w_j^1 = \{w_{1j}^1, w_{2j}^1, \dots, w_{vj}^1\}$.
- Weight set w^2 of the synapses between the hidden layer and output layer corresponding to output neural cell O_k:
  $w_k^2 = \{w_{1k}^2, w_{2k}^2, \dots, w_{ak}^2\}$.
- $IH_j = [w_{1j}^1, w_{2j}^1, \dots, w_{vj}^1] * \begin{bmatrix} i_1 \\ i_2 \\ \dots \\ i_v \end{bmatrix}$ is the input signal of the hidden neural cell H_j.
- $OH_j = f(IH_j)$ is the output signal of the hidden neural cell H_j.
- $IO_k = [w_{1k}^2, w_{2k}^2, \dots, w_{ak}^2] * \begin{bmatrix} O_{H1} \\ O_{H2} \\ \dots \\ O_{Ha} \end{bmatrix}$ is the input signal of the output neural cell O_k.
- $OO_k = f(IO_k)$ is the output of the output neural cell Ok.

With the weight set after training, each input set $I=\{i_1,i_2,\dots,i_v\}$ provide an output set $R=\{OO_1,OO_2,\dots,OO_r\}$.

In order to apply ANN, the most important step is training and it decides how successful the network is going to be. Various training methods exist and can be divided into three types: supervised training, unsupervised training, and reinforcement training. Depending on the dataset, structure, and purpose of the ANN, we can pick the appropriate training method. For the diagnosis of ASD mentioned in this paper, supervised training method will be the most suitable. By using this method, the dataset includes both input and output sets, which means the dataset covers both sets of symptoms and result of the diagnosis. Hundreds or even thousands of sample data will be used to train the ANN for the purpose of making ANN more intelligent so the result of the prediction will be more exact. The supervised training method that will be used in this research is the backpropagation method.

2.4 Fuzzy Artificial Neural Network

Fuzzy ANN is the ANN using fuzzy input set, fuzzy output, fuzzy weights or operations on the Fuzzy set. Fuzzy ANN can be separated into 4 type:

- Type 1: Input is real set, weights are fuzzy.
- Type 2: Input is fuzzy set, weights are real.
- Type 3: Both inputs and weights are fuzzy.
- Type 4: Using t-norm, t-conorm operations.

In this paper, the type of fuzzy ANN is type 3, both inputs and weights are fuzzy.

3. PROPOSING DIAGNOSIS MODEL

By using the combination between the ANN and fuzzy logic, the diagnosis model will be constructed for the purpose of researching and confirming the possibility of the approach in this paper. Because of the difficulty in finding real-life dataset while the training process of ANN requires a large number of data samples, we decided to use a temporary dataset which is generated randomly from the CARS. The requirements and the process of generating this dataset will be presented in part 4 of this paper. The following sections of this part present proposing diagnosis model's factors, input set, dealing with the fact that input values are not absolutely correct by applying parameters of certainty.

3.1 Fuzzy input signal

The intense of the children's expression display the fuzziness of the input set. For example, the connection between the child and people:

- Level 1: No evidence of difficulty or abnormality in relating to people. The child's behavior is appropriate for his or her age. Some shyness, fussiness, or annoyance at being told what to do may be observed, but not to an atypical degree.
- Level 2: Mildly abnormal relationships. The child may avoid looking the adult in the eye, avoid the adult or become fussy if an interaction is forced, be excessively shy, not be as responsive to the adult as is typical, or cling to parents somewhat more than most children of the same age.
- Level 3: Moderately abnormal relationships. The child shows aloofness (seems unaware of the adult) at times. Persistent and forceful attempts are necessary to get the child's attention at times. Minimal contact is initiated by the child.
- Level 4: Severely abnormal relationships. The child is consistently aloof or unaware of what the adult is doing. He or she almost never responds or initiates contact with the adult. Only the most persistent attempts to get the child's attention have any effect.

The ascending order of the level means, the more severe the symptoms become. The point will be evaluated in 1-to-4 score scale with the step of 0.5. Therefore, the input set I of the ANN is the fuzzy set having 15 elements which are a count of questions besides the value of each element is the intensity of the children's expression:

$$\mathbf{I} = \{\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_{15}} \mid \mathbf{x_i} = 1.0; 1.5; \ldots; 3.5; 4.0\}$$

3.2 Fuzzy weights

After completing the test under some existed measures like CARS, M-CHAT… parents usually express skepticism of the accuracy of the test and their evaluations themselves. A woman after finishing the test for her kid with CARS and seeing that the result falls into the beginning of the scale of mild to moderate (sum up to 31-32 while 30-36 is a scale of mild to moderate) usually has two opposite attitudes:

- Negative attitude: afraid of the ability of her child to be suffered from ASD. Moreover, the ASD extent may be worse than the of the test result.
- Positive attitude: stress relieved, she motivated herself to believe that her child does not carry the symptoms of ASD.

Those two flows of thinking are opposite but they do not override each other, thereby still bringing the mother feeling of unease and anxiety. The cause is rooted in her perception that she, to a certain extent, does not trust the accuracy of the test scale as well as her evaluations. In this paper, we propose a parameter representing the status of certain of the parents. Specifically, besides the intensity of children's expression, evaluators need to provide the certainty of their own answers from "uncertain" to "absolutely sure."

3.3 Status of certainty

Since the certainty is fuzzy, the best way to deal with it is modeling in the Fuzzy set by that we can set 3 main levels of certainty:

Because the certainty of an answer cannot be absolutely unsure so the number which presents this parameter cannot be zero. Call the parameter which presents the status of certainty c, we have $0<c\leq1$. Corresponding to "not quite sure", c equals 0.2. Similarly, c equals 0.6 for "sure" and 1.0 for "absolutely sure". As previously mentioned, those statuses of certainty are the main one, there are numerous extra levels in scoring scale for this parameter which is specified with the step of 0.05, we have a total of 16 levels for the status of certainty from 0.2, 0.25,… to 1.

3.4 Applying status of certainty

As indicated above, ANN is the multiple elements network, each element is composed of cells. The signal is transferred from this cell to another through the synapse, the intensity of the signal directly affects the results or the network's output. On the other hand, the accuracy of the input affects the output of a network as well. Therefore, the signal intensity and the accuracy of input are evenly related to the results. The level of accuracy is also the status of certainty.

In this paper, we propose a parameter called status of certainty which is collected from respondent's evaluations. Each evaluation is an input of the input set and has its own status of certainty which stands for the accuracy. This parameter will be applied with the weight correspondingly to change the ANN's results. As the consequence of the proposal status of certainty, we propose an additional parameter called practical weight. Refer $\mathbf{c_i}$ to the status of certain of the input $\mathbf{i_i}$, $\mathbf{w_i}$ is the weight respect to $\mathbf{i_i}$ that we get after training the network, $\mathbf{W_i}$ is the practical weight with respect to weight $\mathbf{w_i}$. We call the exchange from $\mathbf{w_i}$ to $\mathbf{W_i}$ is the reduction of the signal. Figure 3.1 is the demonstration for this:

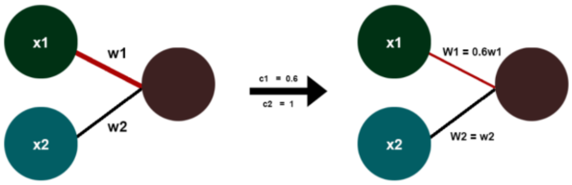

Figure 3.4. The reduction of signal

Practical weight $\mathbf{W_i}$ does not involve in the training process and is calculated by the following formula:

$$\mathbf{W_i = w_i * c_i}$$

The essence of the training process is the change of weights **w**, through it $\mathbf{W_i}$ is modified indirectly. Status of certainty of this input set is different from the other, thus weight set **w** after training process is constant, practical weight set is not the same indistinct input set.

4. RESEARCH SYSTEM

For the purpose of reinforcing the potential of the approach described above, which is the process of using the combination between ANN and fuzzy logic to diagnose ASD, the research system is based on proposed system and uses temporary dataset. Figure 4.1 shows the comparison between these two systems.

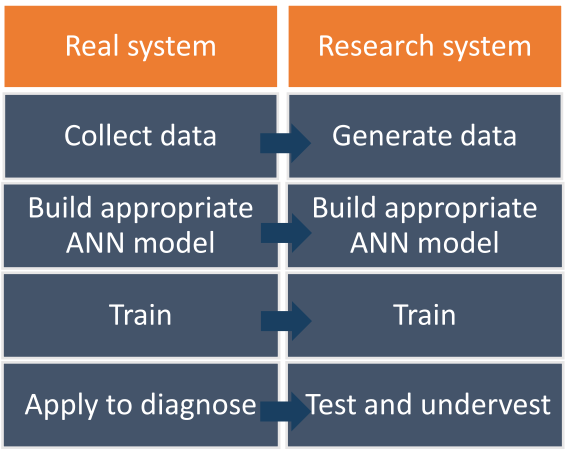

Figure 4.1. From real system to research system

4.1 Generating temporary dataset

The temporary dataset is randomly generated from CARS, which has 18.000 samples in three types representing three levels of ASD:

- Sum up is less than 30 points: normal, 6000 samples.
- Sum up from 30 to 36: mild to moderate, 6000 samples.
- Sum up equals 36 or higher: severe, 6000 samples.

Each sample of dataset contains sixteen fields separated into two parts: the first one contains first fifteen fields which are the points of the evaluations, the other contains one field which is the representation of the diagnosis results. These two parts are the input and output set of the ANN in training process. In this paper, the test dataset has been generated using python notebook and has been saved as a .csv file (Figure 4.2).

Input set (columns 0–14) | Output (column 15)

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	3.0	1.5	1.0	4.0	3.5	2.0	1.5	1.5	2.5	3.5	4.0	3.0	2.0	3.5	2.5	3.0
1	4.0	3.0	3.5	2.5	2.5	1.5	1.0	1.0	2.0	1.5	2.0	1.0	2.0	4.0	1.5	2.0
2	3.0	1.0	3.5	4.0	1.5	3.5	2.5	3.0	4.0	1.0	2.0	2.5	3.0	3.0	3.5	3.0
3	3.5	2.5	1.0	1.0	1.0	4.0	4.0	2.5	2.0	3.5	2.5	2.0	2.0	4.0	3.0	3.0
4	1.0	4.0	4.0	3.5	3.5	2.0	1.0	3.0	4.0	4.0	2.0	3.5	1.0	3.5	2.5	3.0
5	1.0	2.5	3.0	4.0	1.0	1.5	3.0	1.5	2.0	1.5	4.0	3.0	2.5	2.0	4.0	3.0
6	1.0	1.0	1.5	2.5	1.5	2.5	3.5	1.5	2.0	1.5	2.0	3.0	3.0	1.5	1.0	1.0

Figure 4.2. Temporary dataset

4.2 Deploying and training Artificial Neural Network

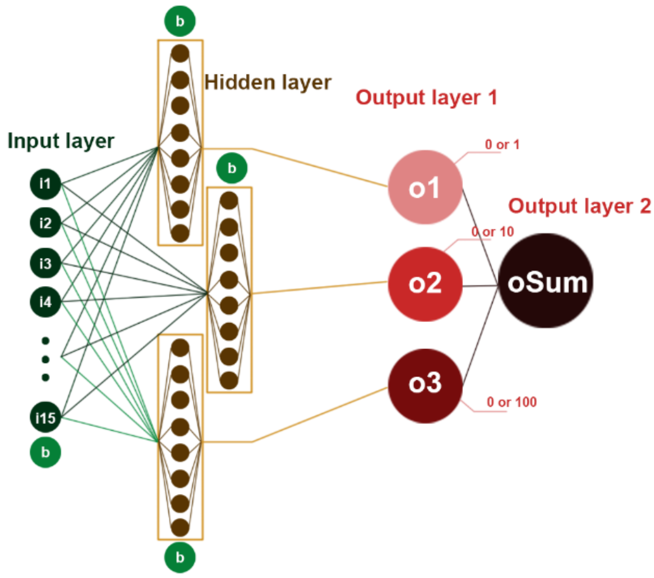

Figure 4.3. Proposed ANN model

The ANN which was described in this research is shown in picture 4.3. This network contains three binary ANNs which are described in part 2 of the paper and an output 2 layer, besides, biases are used in this model as extra cells with a value of 1. Each binary ANN has three layers: input layer contains fifteen cells, hidden layer contains eight cells and output layer contains one cell. The output of these binary ANNs is 0 (false) and 1 (true). From **o1** to **o3** is the ascending disorders level from normal to severe.

The training process of each binary ANN is executed separately while the last layer, layer output 2 is not used in the training process, it combines the results from three binary ANN for the final result in the output of oSum neural cell. oSum retrieve the signal from **o1**, **o2** and **o3** and correspondingly multiplied with **w1** = 1, **w2** = 10 and **w3** = 100. In that case, the signal from **o1** to **oSum** will be 0 if its output is 0, will be 1 if its output is 1. Similarly, the signal from **o2**, **o3** to **oSum** are 10, 100 correspondingly in case the outputs equal to 1 and are 0 in case outputs equal 0.

The output of neural cell oSum is sum up inputs, refer **S** to the output of **oSum**:

- If **S** > 100: severe.
- If 10 < **S** < 100: mild to moderate.
- If 0<**S**<10: normal.
- If **S** = 0: Unable to diagnose.

The training method for those binary ANNs is backpropagation.

4.3 Training result

After the process of training with the temporary dataset with 100 epochs and testing with test dataset which was created just like temporary dataset we have the following table:

Output	Quantity of input	Quantity of correct sample	Rating
oSum	18000	17111	95.79%

Table 4.1. Result of testing

5. CONCLUSION AND FUTURE

The result of the training process above shows that the approach of using ANN and fuzzy logic to diagnose ASD is feasible. Even though our research is conducted based on the temporary dataset, using the model above to improve the accuracy of existing rating scale like CARS is valuable.

The application of ANN and fuzzy logic in general and the model above, in particular, can be developed in two directions: the first one is to improve the accuracy for exist rating scales previously mentioned, the other is to create a new rating scale depending on the standardized ANN. The second development supposes to be more potential because ANN does not require a formal input like CARS. Using even minor symptoms for the input of the ANN could be a chance to detect ASD in infants, which is also one of the main goals in this paper. The result of the training process can show a lot more important symptoms which have never been used before to detect ASD.

To sum up, through this paper we wish to convey the following contents:

- The potential approach in diagnosing ASD by using the artificial neural network combining with fuzzy logic.
- Proposing using the reduction of synapse signal to represent the accuracy of an input, thereby determining the status of certainty in this paper and showing the certainty of evaluators in each question and dealing with ensuing problems.
- Confirming the accuracy of ANNs in dealing with problems similar to the diagnosis of autism by research system.

Although the approach in this paper is potential, the diagnosis system built based on that model needs to be invested early for long-term development and need the contribution of the community to build an abundant and trusted dataset.

ACKNOWLEDGMENT

This research is funded by Funds for Science and Technology Development of the University of Danang under grant number B2016-DNA-37-TT.

REFERENCES

[1] Autism and Developmental Disabilities Monitoring Network, Community report on Autism 2016, 2016.

[2] R.W. Brause, "Medical Analysis by Neural Networks" in proceeding ISMDA '01 proceeding of the Second International Symposium on Medical Data Analysis, 2001, pp 1-13.

[3] D. Robins, D. Fein and M. Barton. "Modified Checklist for Autism in Toddlers", Revised with Follow-Up (M-CHAT-R/F) TM,

[4] Wikipedia, Childhood Autism Ratin Scale definition, https://en.wikipedia.org/wiki/Childhood_Autism_Rating_Scale, Dec. 2017.

[5] M. Chattopadhyay, P. K. Dan, S. Majumdar and P. S. Chakraborty, "Application of artificial neural network in market segmentation: a review of recent trends", 2012.

[6] A. Ultsch, D. Korus and A. Wehrmann, "Neural networks and their rules for classification in marine geology", 1995, pp. 676-693.

[7] C. P. Utomo, A. Kardiana, "Breast Cancer Diagnosis using Artificial Neural Networks with Extreme Learning Techniques", International Journal of

Advanced Research in Artificial Intelligence, Vol. 3, No. 7, 2014.

[8] L.Teodorescu, "Artificial neural networks in high-energy physics", Brunel University, United Kingdom, 2008.

BIOGRAPHIES

Nguyen Van Hieu

Department of Information and Technology, Danang University of Technology

Ngo Le Huy Hien

Department of Computer Science, VN-UK Institute for Research and Executive Education Computer Science