Development of Neuro-fuzzy System for Early Prediction of Heart Attack

Obanijesu Opeyemi
Department of Computer Science and Engineering, LadokeAkintola University of Technology
ooobanijesu@lautech.edu.ng

Emuoyibofarhe O. Justice
Department of Computer Science and Engineering, LadokeAkintola University of Technology
eojustice@gmail.com

Abstract—This work is aimed at providing a neuro-fuzzy system for heart attack detection. Theneuro-fuzzy system was designed with eight input field and one output field. The input variables are heart rate, exercise, blood pressure, age, cholesterol, chest pain type, blood sugar and sex. The output detects the risk levels of patients which are classified into 4 different fields: very low, low, high and very high. The data set used was extracted from the database and modeled in order to make it appropriate for the training, then the initial FIS structure was generated, the network was trained with the set of training data after which it was tested and validated with the set of testing data. The output of the system was designed in a way that the patient can use it personally. The patient just need to supply some values which serve as input to the system and based on the values supplied the system will be able to predict the risk level of the patient.

Index Terms— ANFIS, Adaptative Neuro-Fuzzy System, Fuzzification, Membership Function, Fuzzy Rule, Membership Function

I. Introduction

Heart and blood vessel diseases called cardiovascular diseases include numerous problems, many of which are related to a process called atherosclerosis. Atherosclerosis is a condition that develops when a substance called plaque builds up in the walls of the arteries.

This build-up narrows the arteries making it harder for blood to flow through. This can cause blood clot formation which can cause a heart attack (also called Myocardial Infarction) or stroke [1].

When a heart attack occurs, the speed of detection and quick intervention is highly essential to save the life of heart attack patient and to prevent heart damage. Nowadays, the use of computer technology in the field of medicine has highly increased [2]. The use of intelligent systems such as neural network, fuzzy logic, genetic algorithm and neuro-fuzzy systems has highly helped in complex and uncertain medical tasks such as

diagnosis of diseases [3]. Over the last few decades, neural networks and fuzzy systems have established their reputation as alternative approaches to intelligent information processing systems. Both have certain advantages over classical methods, especially when vague data or prior knowledge is involved. However, their applicability suffered from several weaknesses of the individual models. Therefore, combinations of neural networks with fuzzy systems have been proposed, where both models complement each other.

Neuro-fuzzy hybridization results in a hybrid intelligent system that synergizes these two techniques by combining the human-like reasoning style of fuzzy systems with the learning and connectionist structure of neural networks [4]. The basic idea of combining fuzzy systems and neural networks is to design an architecture that uses a fuzzy system to represent knowledge in an interpretable manner and the learning ability of a neural network to optimize its parameters [5].

The remaining sections of the paper are organized as follows: In Section 2 the architecture of an Adaptative Neuro-fuzzy Inference System (ANFIS) was presented. The design of the neuro-fuzzy model for heart attack detection was presented in Section 3.Section 4 describes the data set used. The experimental results are described in Section 5. The conclusion of the work is given in section 6.

II. Adaptative Neuro-Fuzzy Inference System

Adaptative Neuro-Fuzzy Inference System (ANFIS) is one of the hybrid neuro-fuzzy inference expert systems and it works in Takagi-Sugeno-type fuzzy inference system, which was developed byJyh-Shing and Roger Jang1993[6]. The technique provide a method for the fuzzy modeling procedure to learn information about a data set, in order to compute the membership function parameters that best allow the associated fuzzy inference system to track the given input/output data. This learning method works in a manner similar to that of neural networks.

A. ANFIS Architecture

The ANFIS architecture uses two fuzzy if-then rules based on a first order Sugeno model:

$Rule\ 1 : if\ (x\ is\ A1)\ and\ (y\ is\ B1)\ then\ (f1 = p1x + q1y + r1)$

$Rule\ 2 : If\ (x\ is\ A2)\ and\ (y\ is\ B2)\ then\ (f2 = p2y + q2y + r2)$

where x and y are the inputs, Ai and Bi are the fuzzy sets, fi are the outputs within the fuzzy region specified by the fuzzy rule, pi, qi and ri are the design parameters that are determined during the training process.

The architecture consists of five layers of nodes. Out of the five layers, the first and the fourth layers consist of adaptive nodes while the second, third and fifth layers consist of fixed nodes [7]. The adaptive nodes are associated with their respective parameters, get duly updated with subsequent iteration while the fixed nodes are devoid of any parameters. Figure 1 shows the ANFIS architecture
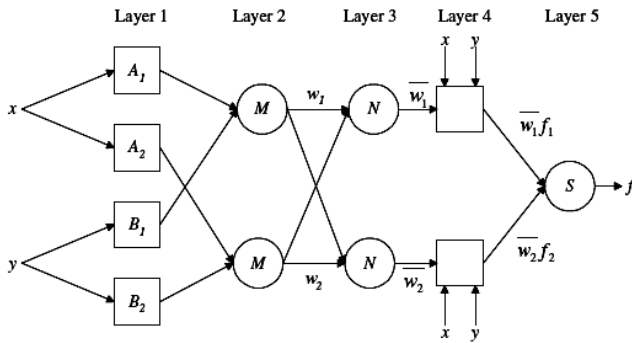

Fig 1 ANFIS ARCHITECTURE

Layer 1: Fuzzification layer: Every node i in the layer 1 is an adaptive node. The outputs of layer 1 are the fuzzy membership grade of the inputs, which are given by:

$$o_i^1 = \mu_{Ai}(x), \qquad \text{for } i = 1,2 \qquad (1)$$

$$o_i^1 = \mu_{Bi-2}(y), \qquad \text{for } i = 3,4 \qquad (2)$$

where x and y are the inputs to node i, where A is a linguistic label (small, large) associated with this node function o_i^1 is the membership function of A_i and it specifies the degree to which the given x,y satisfies the quantifier A_i . $\mu_{Ai}(x)$, $\mu_{Bi-2}(y)$ can adopt any fuzzy membership function. Usually we choose $\mu_{Ai}(x)$ to be bell-shaped with maximum equal to 1 and minimum equal to 0, such as:

$$\mu_{Ai}(x) = \frac{1}{1 + [(\frac{x-c_i}{a_i})^2]^{b_i}} \qquad (3)$$

where (a_i, b_i and c_i) are the parameters of the membership function. Parameters are referred to as premise parameters.

Layer 2: Rule layer a fixed node labeled M which multiples the incoming signals and send the product out. Each node output represent the firing strength of the rule. The outputs of this layer can be represented as:

$$o_i^1 = w_i = \mu_{Ai}(x) \times \mu_{Bi}(y) \qquad i = 1,2 \qquad (4)$$

Layer c x3: Every node in this layer is a circle labeled N. The ith node calculates the ratio of the ith rule's firing strength to the sum of all rule's firing strengths. Output of this layer is called normalized firing strengths

$$o_i^3 = \bar{\omega} = \frac{\omega i}{\omega 1 + \omega}, \qquad i = 1,2 \qquad (5)$$

Layer 4: In this layer, the nodes are adaptive nodes. The output of each node in this layer is simply the product of the normalized firing strength and a first order polynomial (for a first order Sugeno model) , where $\bar{\omega}$ is the output of layer 3 and {p_i,q_i, r_i} is the parameter set. Thus, the outputs of this layer are given by:

$$o_i^4 = \bar{\omega}_i f_i = \bar{\omega}_i(p_i x + q_i y + r_i) \qquad i = 1,2 \qquad (6)$$

Parameters in this layer will be referred to as consequent parameters.

Layer 5: The single node in this layer is circle node labeled $\sum$ that computes the overall output as the summation of all incoming signals, i.e.,

$$o_i^5 = \text{overall output} = \sum_i \bar{\omega}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w i} \quad i = 1,2 \qquad (7)$$

B.   Hybrid Learning Algorithm

Hybrid learning rule used the gradient method and the least squares estimate (LSE) to identify and update parameters. Each epoch of this hybrid learning procedure is composed of a forward pass and a backward pass.

i.   Forward Pass

In the forward pass of the hybrid learning algorithm, functional signal go forward until layer 4 and the consequent parameters are identified by the least squares estimate. In the backward pass, the error rate propagates backward and the premise parameters are updated by the gradient descent. When the values of premise parameters are fixed the overall, output can be expressed as a linear combination of the consequent parameters.

$$F = \frac{w1}{w1+w2} f_1 + \frac{w_1}{w_1+w_2} \qquad (8)$$

$$= \bar{w}_1 f_1 + \bar{w}_2 \qquad (9)$$

$$= (\bar{w}_1 x)p_1 + (\bar{w}_1 y)q_1 + (\bar{w}_1)r_1 + (\bar{w}_2 x)p_2 + (\bar{w}_1 y)q_1 + (\bar{w}_1)r_2 \qquad (10)$$

$which is linear in the consequent parameters$ p_1 , q_1 , r_1, p_2, q_2 and r_2

$$f = XW \tag{11}$$

If X matric is invertible then

$$W = W^{-1}f \tag{12}$$

$$W = (X^T X)^{-1} X^T f \tag{13}$$

Otherwise a pseudo-inverse is used to solve for W .

ii. Backward Pass

In the backward pass, the error signals propagate backward and the premise parameters are updated by gradient descent.

$$a_{ij}(t+1) = a_{ij}(t) - \frac{\eta}{p} \cdot \frac{\partial f}{\partial f} \tag{14}$$

Where η is the learning rate for a_{ij} . The chain rule is used to calculate the partial derivatives used to update the membership function parameters.

$$\frac{\partial E}{\partial a_{ij}} = \frac{\partial E}{\partial f} \cdot \frac{\partial f}{\partial f_i} \cdot \frac{\partial f_i}{\partial w_i} \cdot \frac{\partial w_i}{\partial \mu_{ij}} \cdot \frac{\partial u_{ij}}{\partial a_{ij}} \tag{15}$$

The partial derivatives are derived as

$$E = \frac{1}{2}(f - f^1)^2 \ hence \ \frac{\partial E}{\partial f} = (f - f^1) = e \tag{16}$$

$$f = \sum_{i=1}^{n} f_i \ hence \ \frac{\partial E}{\partial f} = 1 \tag{17}$$

$$f_i = \frac{(p_i x + q_i y + r_i) - f}{y \sum_{i=1}^{n} wi} \tag{18}$$

$$w_i = \prod_{j=1}^{m} \mu_{A_{ij}} hence \ \frac{\partial w_i}{\partial \mu_{ij}} = \frac{w_i}{\mu_{ij}} \tag{19}$$

The last partial derivative depends on the type of the membership function used. The parameters of the membership functions are updated in the same fashion.

The gradient is then obtained as

$$\frac{\partial E}{\partial a_{ij}} = e \frac{(p_i x + q_i y + r_i) - f}{y \sum_{i=1}^{n} wi} \frac{w_i}{\mu_{ij}} \frac{\partial \mu_{Aij}}{\partial a_{ij}} \tag{20}$$

$$\frac{\partial E}{\partial a_{ij}} = e \frac{(p_i x + q_i y + r_i) - f}{y \sum_{i=1}^{n} wi} \frac{w_i}{\mu_{ij}} \frac{\partial \mu_{Bij}}{\partial b_{ij}} \tag{21}$$

III. Neuro-fuzzy Model Development

The first stage in the formation of the neuro-fuzzy system is the development of Fuzzy Inference system (FIS) after which the FIS was trained using neural network algorithm. The algorithm used in training the FIS is back propagation gradient descent in combination with a least squares Estimate method. During the training process, the membership function parameters were adjusted and this allowed the fuzzy system to learn from the data they were been modeled.

Eight variables were used which serves as input to the fuzzy system, they are age, sex, chest pain type , resting blood pressure, cholesterol, fasting blood sugar, maximum heart rate and exercise. Two generalized bell Membership function were used for each of the inputs and linear membership function was used for output. The membership function curve for blood sugar and blood pressure are shown in figures2 and 3 respectively .The curve defines how each point in the input space is mapped to the membership value between 0 and 1.
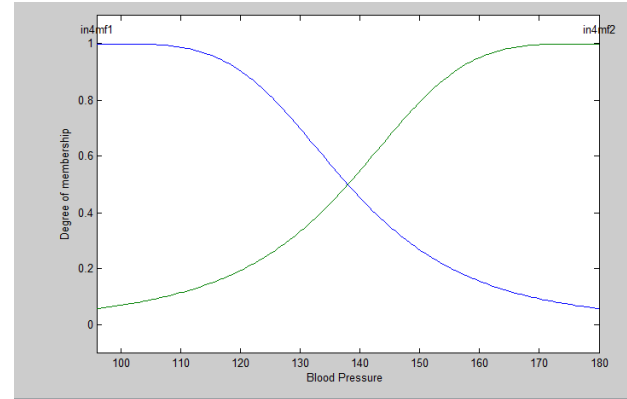

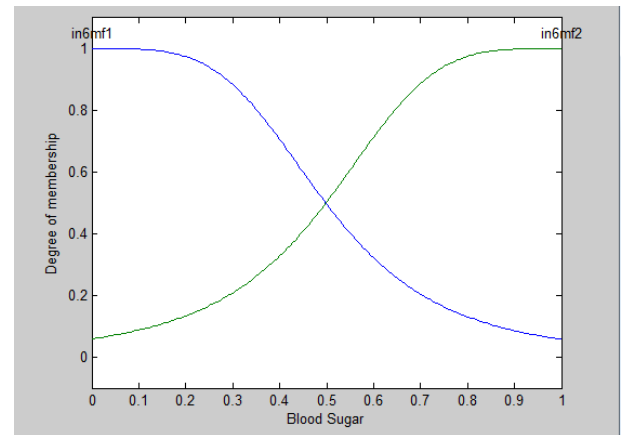Fig 2: Membership Function of Blood Pressure

Fig 3: Membership Function of Blood Sugar

The fuzzy rules for heart disease detection were generated using the set of input and output, the antecedent part of all rules has eight sections. The antecedent parts of the rules were combined using the maximum and product operator. Each of the rule's weight was determined using the prod implication method. The result was aggregated with sum which is the sum of each rule's output set. Because decisions are based on the testing of all of the rules in a FIS, the rules were combined in some manner in other to make a decision; this was done during aggregation process by combining the fuzzy sets that represent the output of each rule into a single fuzzy set.

The input for the defuzzification process is the fuzzy set (the aggregate output fuzzy set) and the output is a single number. The aggregate of a fuzzy set

encompasses a range of output values, and so must be defuzzified in order to resolve a single output value from the set. For the defuzzification process, weighted Average method was used.

The FIS generated was trained using back propagation algorithm in combination with least squares estimate method. During the training process the membership function parameters were adjusted.

System Validation

System validation is the process by which the input vectors from input/output data set on which the FIS was not trained are presented to the trained FIS model to see how well the FIS model predicts the corresponding data set output values. The set of testing data supplied to the system was used to validate the trained neuro-fuzzy system.

Means square error (MSE) was used to evaluate the predictive performance of the model. MSE is a well acceptable indicator that describes the differences between actual data and the predicted values. MSE is calculated below:

$$MSE = \frac{1}{p}\sum_{i=1}^{p}(x_{(i)} - y_{(i)})^2 \qquad (22)$$

Where p the number of data is pairs; $x_{(i)}$ and $y_{(i)}$ are ith desired output and predicted output respectively.

After training the system, the system was tested with a set of testing dataset to verify the accuracy of the predicted values. The degree to which the system output match the actual data set is used to provide an evaluation of the system's predictive abilities. Mean Square Error (MSE) between the actual data set and the predicted values was used to evaluate the system. The validation of the system gave a Mean Square Error, MSE = 0.014463

The system was also tested with another set of data. Out of 12 pairs of data set used there was one misinterpretation which shows an accuracy level of over 90%.

IV.   Data Description

The data set used was extracted from the collection of databases at the University of California collected by David W. Aha from V.A. medical Center, Long Beach and Cleveland Clinic foundation dataset [8]. This dataset is to diagnose the presence or absence of heart disease given the result of various medical tests carried out on the patients. This database contains 76 attributes but all published experiments refer to using a subset of 14 of them and for this project work, we have used 8 attributes for input and 1 attribute for output.

The input variables are :

- Heart Rate: The value of input field is heart rate of man achieved. Heart rate of man decrease in respect to age. There are three fields in these variables which are: low, normal and high.
- Exercise: This input field stands for exercise induced angina.
- Blood Pressure: The blood pressure value used is the systolic blood pressure. The input variable is divided into 3 fuzzy sets which are low, normal and high.
- Age: The input field is divided into 3 fuzzy sets which are very young, young and old.
- Cholesterol: This input field has 4 fuzzy sets which are low, medium, high and very high.
- Chest Pain Type : The chest pain type has 4 values which are : 1 - typical angina, 2 - atypical angina, 3 - non-angina pain, 4 - asymptomatic.
- Blood Sugar (diabetics): The blood sugar field is one of the most important factors in this system. The input field has 3 fuzzy sets which are low, normal and high.
- Sex: The input field has 2 (0, 1) values and sets (female and male).

Anfis Training

Matlab software was used to develop a neuro fuzzy model. A training data set that contains the desired input/ output pairs of the system were presented to ANFIS toolbox for training. The stages involve are listed below:

1) Loading the training and testing data.
2) Generate or load an initial FIS model.
3) View the FIS model structure once an initial FIS has been generated or loaded.
4) Choose the FIS model parameter optimization method: a mixture of back propagation and least squares (hybrid method).
5) Choose the number of training epochs and the training error tolerance for training the system.
6) Train the FIS model by clicking the Train Now button. This training adjusts the membership function parameters and plots the training error plot(s) in the plot region.
7) Validating the trained FIS.

Training data set that contains the desired input / output data of the system to be modeled were loaded to the ANFIS editor. The data set was arranged in an array with the data arranged as column vectors and the output in the last column. Lastly, the model was validated using a testing or checking data.

V.   Results and Discussion

Various experiments were performed and the sizes of the training and testing data sets were determined by taken into consideration the accuracies of system. The data set used was divided into two; the training data and

the testing data, 80% of the data set were used for training while the remaining 20% were used for testing. The training data set was used to train the model while the testing data set was used to verify the accuracy and the effectiveness of the trained Model for diagnosis of heart attack. The training stopped at 10 epochs.

Figures 4 and 5 show the surface view of mapping with input parameters versus output parameters:

Surface View presents a three-dimensional curve that represents the mapping with input parameters versus output parameters. It shows the dependency of output (heart attack rate) on the inputs. Figure 4 shows the dependency of output on the inputs: age and blood pressure. Figure 5 shows the dependency of output on the inputs: cholesterol and chest pain type.
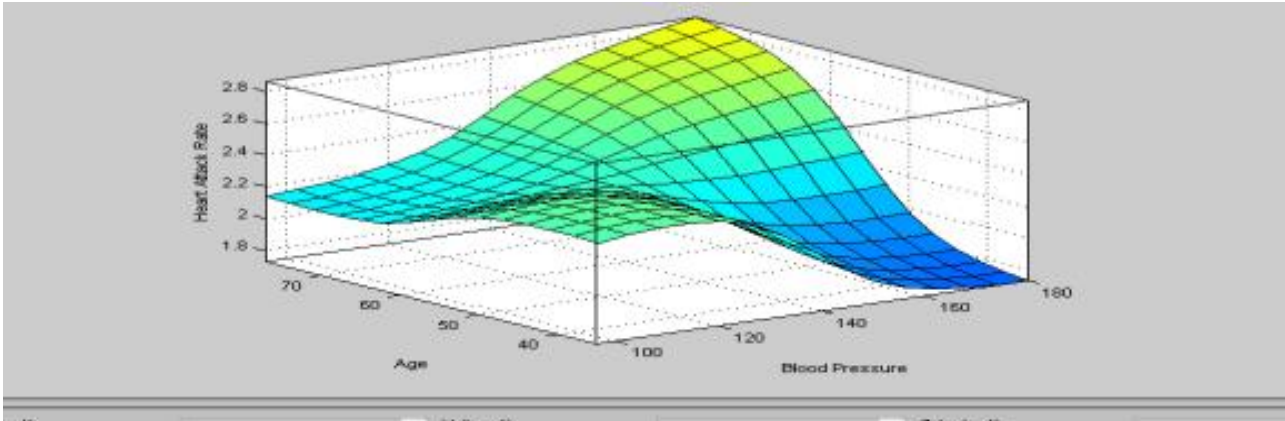
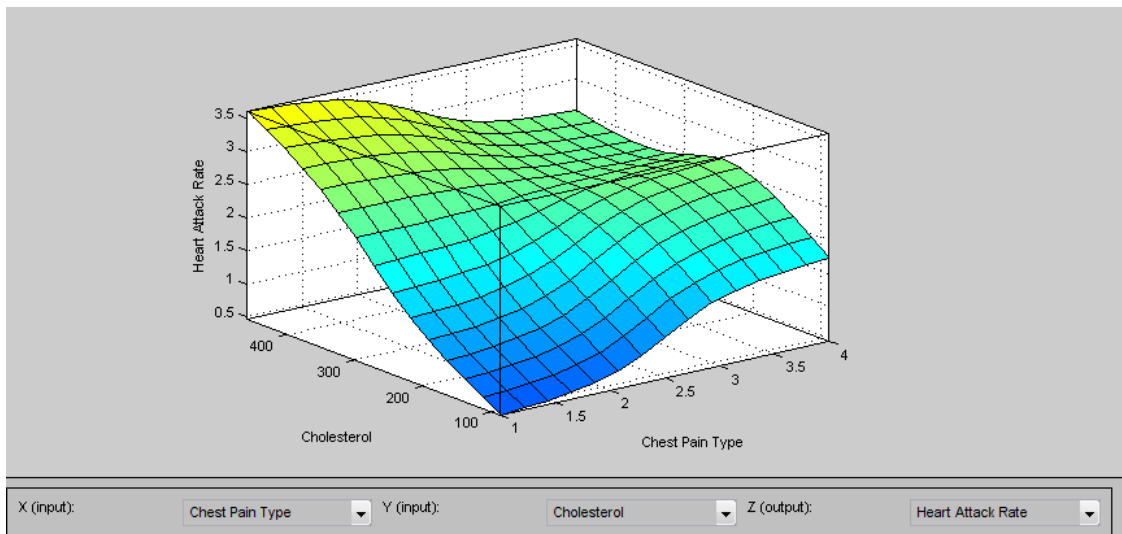Fig 4: Surface View of age versus blood pressure

Fig 5: Surface View of Cholesterol versus chest pain type

A. Fuzzy Inference System for Heart Detection

Figure 6 shows the diagram of the Fuzzy Inference System for heart attack detection. There are 8 input variables which are age, sex, chest pain type, blood pressure, cholesterol, blood sugar, heart rate , exercise and one output variable which shows the risk levels of patient. The system uses 256 rules, each of the input variable uses 2 membership functions.

The names of the input variables are shown are shown at the left. The sample membership function for each of the input variable is displayed in the boxes. The type of the inference system use is displayed at the middle.
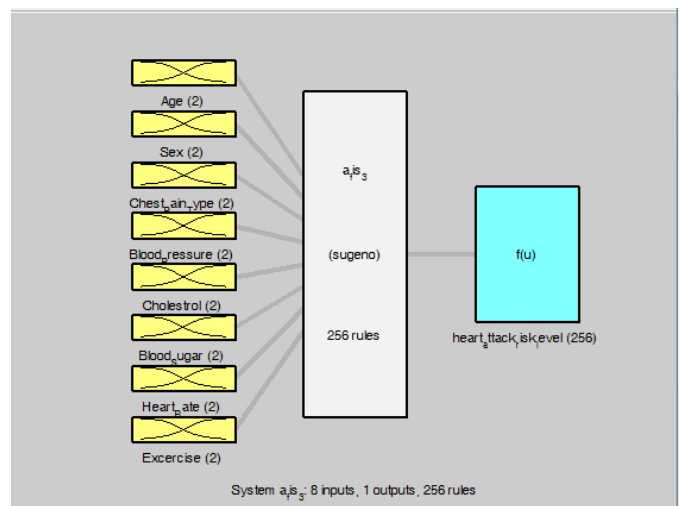

Fig 6: Fuzzy Inference System for Heart Attack Detection

B. Simulated Rule Base for Heart Attack Detection:

Figure 7 shows some part of the simulated rules for heart attack detection. The system used 256 rules.
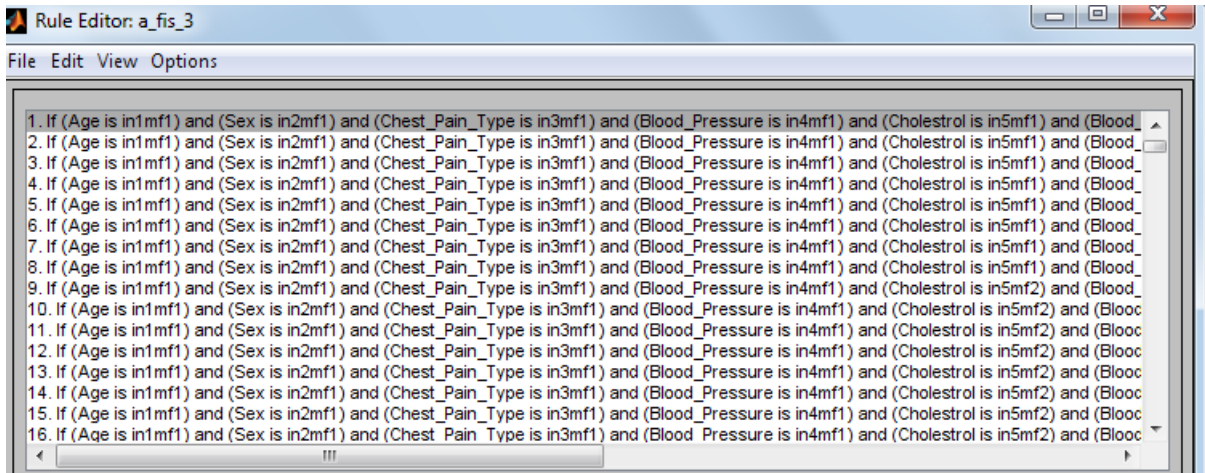

Fig 7:   Simulated View of the Rule-base of heart attack diagnosis

The screenshots of heart attack detection with different risk levels are shown in Figures 8 to 11. There are 8 values that needed to be supplied by the user which are: age, sex which has to values 1 for male and 0 for female, blood pressure the systolic value is used, cholesterol, maximum heart rate, chest pain type which

is classified into four :typical angina, atypical angina, non-angina and asymptomatic , fasting blood sugar which is true if the value is greater than 120 and false if the value is less than 120 note that the patient must fast for about 8hours before taken the reading.
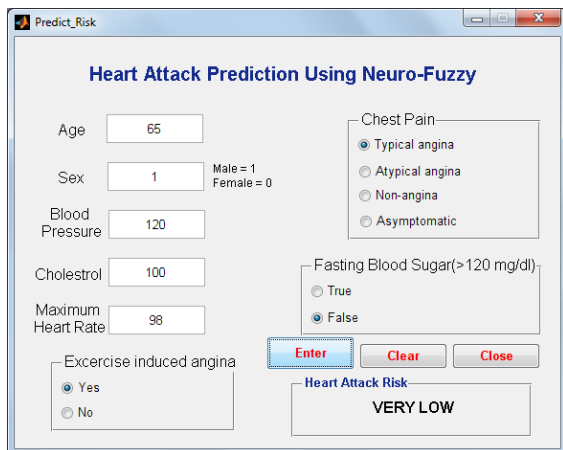

Figure 8: Heart attack prediction with very high risk level
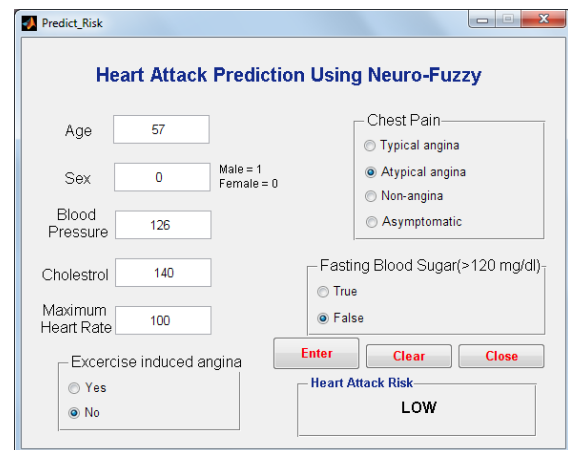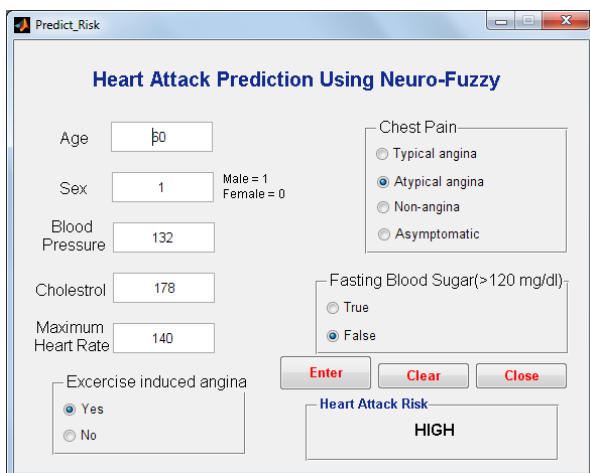

Figure 9: Heart attack predictions with low risk level

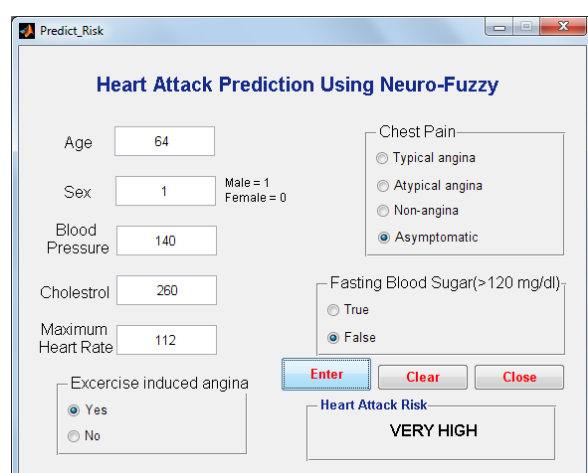
Figure 10: Heart attack prediction with high risk level

Figure 11: Heart attack predictions with very high risk level

VI.   Conclusion

A Neuro-fuzzy system for early detection of heart attack was implemented using ANFIS. The output shows the risk level of patient which has been classified into four different risk levels: very low, low, high and very high. The design of the system can be divided into 3 stages: ANFIS model development, network training and system validation and testing.

The system has been tested and result showed over 90% accuracy, it has been shown that neuro-fuzzy is suitable and feasible to be used as a supportive tool for disease diagnosis. The system is designed in a way that the patient can use it himself.

References

[1] American Heart Association, Inc. (2011), "Sudden Cardiac Arrest" National Heart Lung and Blood Institute U.S. Department of Health and Human Services.

[2] Ali A. and Mehdi N. (2010),"A Fuzzy Expert System for Heart Disease Diagnosis" A Proceeding of the International MultiConference of Engineers and Computer Scientists, vol. 1

[3] ZAPTRON *Systems*, Inc. *Neurofuzzy*(1999),"A Different Type of Neural Nets" Zaptron's High-Order Nonlinear Neural Networks.|Fuzzy logic|Neural.

[4] Mehdi N. and Mehdi Y. (2009)," Designing a Fuzzy Expert System of Diagnosing the Hepatitis B Intensity Rate and Comparing it with Adaptive Neural Network Fuzzy System" A Proceeding of the World Congress on Engineering and Computer Science, vol. II.

[5] Vipul A.S.,(2009),"Adaptive Neuro-Fuzzy Inference System for Effect of Wall Capacitance in a Batch Reactor" Advances in Fuzzy Mathematics ISSN 0973-533X 4;69-75, 69-70

[6] Jyh-Shing and Roger Jang.(1993)," ANFIS: Adaptive-Network-Based Fuzzy Inference System, computer methods and programs in biomedicine", IEEE Transactions on Systems, University of California.

[7] Nazmy T.M., El-messiry H., Al-bokhity B. (2009)," Adaptive neuro-fuzzy inference system for, classification of ECG signals " Journal of Theoretical and Applied Information Technology pp 71-73

[8] Robert D., Matthias P., Williams S., Andras J. ,UCI repository of Machine Learning Databases,University of Califonia. Available online at: www.archive.ics.uci.edu/ml/datasets/ Heart+disease.

Author's Profile

Obanijesu O.O has just completed her masters in Computer Science at adoke Akintola University of Technology and she is currently a staff at the university ICT centre. E-health and telemedicine are her areas of research interest.

Emuoyibofarhe O. Justice received his Ph.D in 2004. He specializes in Neuro-fuzzy computing/ computational optimization. He had post-doctoral fellowship at the centre of excellence for mobile e-service, University of Zululand, South Africa in 2006, He is a member of the (IEEE) computational intelligence society. He is currently an associate professor in the department of computer science and engineering, Ladoke Akintola University of Technology Ogbomoso. He is currently the Deputy Dean of the Faculty of Engineering and Technology of the same Institution. His present reseach area is in the application of mobile computing and wirelss communication to e-health and telemedicine