ANFIS Based Classification Model for Heart Disease Prediction

Negar Ziasabounchi*, Iman Askertzade
Department of Computer Engineering, Ankara University, Ankara, Turkey
Email: n.z.sabounchi@gmail.com, imasker@eng.ankara.edu.tr

Abstract-- Heart disease diagnosis procedure is very vital and critical issue for the patient’s health. Furthermore, it will help to decrease disease to a more specific level. The role of using machine learning techniques and data mining algorithms in diagnosis of heart disease is very considerable. The aim of this study was to develop a method of classifying for heart disease degree of patient based characteristic data using adaptive neuro fuzzy inference system. The data were obtained from the University of California at Irvine (UCI) machine learning repository. Seven variables are used as input of prediction model. To test the ability of the trained anfis models to recognize heart disease diagnosis, we used k-fold cross validation method. The experimental results demonstrate that the model successfully forecasts the patient's heart disease degree with an accuracy rate of 92.30%.

Index Term-- ANFIS; Adaptive Neuro; Fuzzy Interface System; Classification; Heart Disease.

1. INTRODUCTION
As stated in [1], heart and blood vessel diseases called cardiovascular diseases contain several problems, many of which are related to atherosclerosis process which is a condition develops when a substance called plaque builds up in the walls the arteries. This process shrinks these arteries and cause problems for blood flowing through the vessels. Therefore, as a bad result, the potential risk of having heart diseases starts to increase. The World Health Organization report that world people die each year from heart diseases than from any other causes and it is estimated that more than 80% death from cardiovascular disease take place in low middle income countries and it is on the rise [2]. Hence, the main question comes that whether or not we can predict it before it comes true. As stated in [1, 3], with the advent of computer technology, the use of intelligent methods and algorithms (e.g. neural network, fuzzy logic and genetic algorithm) has started to play crucial role in complex and uncertain medical tasks such as diagnosis of diseases.

In last decade, the literature about the use of intelligent methods in medicine domain has seen enormous number of related studies [3, 4, 5, 6, 7]. As reported in [6], computer assisted applications and tools for diagnosis and treatment of patients seems to be more recent area of interest. Furthermore, the medical practitioners are also employing computerized technologies to assist in diagnosis and opinions as medical diagnosis is full of uncertainty [6]. On the other hand, fuzzy logic and neural networks stand as good methodologies dealing with these uncertainties [1, 6]. As noted in [1], in case of vague data or prior knowledge is involved, both of them serve certain advantages over classical methods. Nonetheless, as reported in [1], individual usage of these two methods can cause some weaknesses. At this point, neuro-fuzzy integration presents a hybrid intelligent system that combines the power of human-like reasoning style of fuzzy logic with the connectionist structure of neural networks [8]. Moreover, Adaptive 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 by Jyh-Shing and Roger Jang in 1993 [9]. As reported in [1], the ANFIS method provides a method for the fuzzy modeling procedure to learn information about a data set, in order to calculate membership function parameters which allow the associated fuzzy inference system to track the given input/output data.

When the literature is investigated, it can be easily seen that there exist diverse types of studies based on fuzzy and ANFIS methodologies [2, 10, 11, 12, 13, 14]. In [15], Malhotra and Malhotra introduced an ANFIS based smart predictive model for screening potential defaulters on consumer loans. Furthermore, in [16] Soyguder and Alli employed ANFIS method to build an expert system for the humidity and temperature control in HVAC systems. When the medical related ones are encountered, some of them can be listed as follow: In [11], Sungging et al. designed and developed an ANFIS based artificial intelligence system for lung cancer diagnosis. On the other hand, in [17] a fuzzy rule based expert system was implemented for asthma diagnosing. Likewise, Ucar et al. [18] carried out a study about tuberculosis disease diagnosis by using adaptive neuro fuzzy inference system and rough sets.

In this study we applied ANFIS method to Cleveland Clinic Foundation heart disease dataset which has been obtained from the well-known UCI machine learning data repository. The dataset consists of 303 subjects. The followed analysis methodology and the experiment results are given in the respective section. According to results, ANFIS method stands as a good and flexible predicting mechanism for heart disease prediction.

Remaining of this paper is organized as follows: The theoretical background of ANFIS is demonstrated in section 2. Section 3 presents the properties and structural information of dataset. The experiments and results of followed methodology is given in section 4 and whole study is summarized in section 5.

2. OVERVIEW OF ADAPTIVE NEURO FUZZY INFERENCE SYSTEM
ANFIS is a network structure which represents Takagi-Sugeno-type fuzzy inference system. ANFIS computed initial membership functions by training itself with training data,
afterwards adjusting membership functions using either a back propagation algorithm or a hybrid-learning algorithm (a combination of back propagation and the least squares method) to minimize error measure. Actually, ANFIS benefits Artificial Neural Network’s learning ability and Fuzzy-Logic decision making capability together. In [9], the learning rules of ANFIS have been described in detail.

2.1 ANFIS Architecture
In the following lines we express the five layers of ANFIS system.

Layer 1:
The first layer is fuzzification layer. Every node i in this layer is a square node, Which are given by:
\[ O_i^1 = \mu_A(x), \quad \text{for } i=1,2 \]  
\[ O_i^1 = \mu_B(y), \quad \text{for } i=3,4 \]
where \( x \) and \( y \) are the inputs to node \( i \) and outputs are fuzzy membership grade of inputs [9]. In order to calculate the degree of membership of the input, every node uses Gaussian membership function.
\[ O_i^1 = \mu_A(x) = e^{-\frac{(x-c)^2}{2\sigma^2}}, \quad \text{for } i=1,2 \]
where \( \{c,\sigma\} \) is a parameter set. \( C \) represents the membership function’s center and \( \sigma \) determines the membership function’s width. These parameters called premise parameters.

Layer 2:
Rule layer is the second layer. In this layer the input values are the membership functions and each node multiplies inputs and gives output which represents the firing strength of rule. The output of this layer given in equation
\[ O_i^2 = \bar{w}_i = \mu_A(x1) \times \mu_B(x2) \times \mu_C(x3) \times \mu_D(x4) \times \mu_E(x5) \times \mu_F(x6) \times \mu_G(x7) \quad i=1,2,...,7 \]

Layer 3:
Here the \( i \)-th node is calculated by the ratio of the \( i \)-th rules firing strength to the sum of the rule’s firing strengths.
\[ O_i^3 = \bar{w}_i = \frac{w_i}{\sum_i w_i} \quad i=1,2,...,7 \]

Layer 4:
In this layer the nodes are adaptive nodes.
\[ O_i^4 = \bar{w}_i f_i = \bar{w}_i (p_1 x_1 + q_1 y_1 + r_1) + \bar{w}_2 (p_2 x_2 + q_2 y_2 + r_2) \]
\[ = (\bar{w}_1 x_1) p_1 + (\bar{w}_1 y_1) q_1 + (\bar{w}_2 x_2) p_2 + (\bar{w}_2 y_2) q_2 \quad i=1,2,...,7 \]
where \( \{w, p, q, r \} \) is the parameter set which referred to as consequent parameters.

Layer 5:
This layer there is single fixed node. In this layer computes the overall output as the summation of the all incoming signals, given in equation
\[ O_i^5 = y = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \quad i=1,2,...,7 \]  

Fig. 1 shows the ANFIS architecture:

2.1 Hybrid-learning Algorithm
There are two passes for hybrid algorithm, forward pass and backward pass. In the forward pass of the hybrid algorithm when the values of premise algorithm are fixed, the overall output can be expressed as a linear combination of the consequent parameters. Output \( y \) can be written as
\[ y = f(i,s) \]
where \( i \) is the vector of input variables, \( S \) is the set of parameters. If there is a function \( H \) such that the composite function \( H\circ f \) is linear in some of the elements of \( S \), then these elements can be identified by the least-squares method [15]. If the parameter set \( S \) can be divided into two sets \( S_1 \) and \( S_2 \), which \( S_1 \) is set of premise parameters and \( S_2 \) is set of consequent parameters we have
\[ S = S_1 \bigoplus S_2 \]
(\( \bigoplus \) represents the direct sum) such that \( H \circ f \) is linear in the elements of \( S_2 \). After substituting training data in (11) a matrix equation is obtained
\[ A\theta = Y \]
where \( \theta \) is an unknown parameter vector whose elements are parameters in \( S_2 \). Let \( S_2 = M \), and dimensions of \( t \) is \( M \times 1 \) parameter vector, \( A \) is \( p \times M \) matrix and \( Y \) is \( p \times 1 \) output vector.
This is a standard linear least-squares problem, and the best solution for \( \theta \), which minimizes \( \|A\theta - Y\|^2 \), is the least-squares estimator (LSE)\( \theta^* \) [15].

\[
\theta^* = (A^T A)^{-1} A^T y,
\]

(13)

In the back propagation learning rule, consequent parameters are constant and a gradient model is applied to update premise parameters with output of network distribute through back. The following equation is used for updating premise parameters

\[
\Delta x = -\eta \frac{\partial E}{\partial x},
\]

(14)

where \( x \) is premise parameters, \( \eta \) training rate, \( E \) is error value at out of the network.

3. Dataset

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Min</th>
<th>Max</th>
<th>Number of MF</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>age</td>
<td>29</td>
<td>77</td>
<td>3</td>
<td>Age in years</td>
</tr>
<tr>
<td>cp</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>Chest pain type(4 type)</td>
</tr>
<tr>
<td>trestbps</td>
<td>94</td>
<td>200</td>
<td>3</td>
<td>resting blood pressure</td>
</tr>
<tr>
<td>chol</td>
<td>126</td>
<td>564</td>
<td>3</td>
<td>cholesterol</td>
</tr>
<tr>
<td>fbs</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>resting blood sugar (0=false, 1=true) it is true when fbs&gt;120</td>
</tr>
<tr>
<td>restecg</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>resting electrocardiographic(ECG)</td>
</tr>
<tr>
<td>talach</td>
<td>71</td>
<td>202</td>
<td>2</td>
<td>maximum heart rate</td>
</tr>
</tbody>
</table>

4. Experiments and Results

4.1 Data Preparation
At the first stage, our crisp data which was collected from Cleveland heart disease dataset was converted in order to be used in MATLAB environment and preprocessed by normalization method. After that seven input variables are introduced to input of ANFIS with their membership functions. Gaussian membership function is used for fuzzy set description.

4.2 ANFIS Model Design
ANFIS or adaptive neuro fuzzy inference system is a class of adaptive networks that are functionally equivalent to fuzzy inference systems [9]. Our network has premise parameters and consequent parameters for the network structure illustrated in section 2. In this study, hybrid learning algorithm is used for our ANFIS model to identify parameters. There are two passes for hybrid algorithm, forward pass and backward pass. After presentation of premise parameters, in the forward pass a node outputs move ahead until layer 4 and the consequent parameters are calculated with least square estimate (LSE) then error measure is calculated for each node. In the backward pass, the error signals distribute backward to update premise parameters with gradient descent. In this section, we used MATLAB software to develop ANFIS model. In the first stage, it is necessary to generate a fuzzy inference system for the heart diseases. To create an initial set of membership functions we used Grid Partition method. At the beginning of training, this method divides the data space in to rectangular sub-spaces using axis-paralleled partition based on predefined number of membership functions and their types in each dimension [19]. Seven variables were used with their membership functions. We used Gaussian membership function for each of the input variables. The number of these membership functions is shown in Table 1. In the training process, consequent parameters are learnt during forward pass, when least squared error estimate approach is employed and premise parameters are learnt, when gradient descent method is applied during backward pass. After repeating forward and backward passes, premise and consequent parameters are determined. For our ANFIS model, training error tolerance set to 0.01. Three Gaussian membership function were used for blood pressure and cholesterol, The graph of these membership functions before training are shown in Fig. 2. and modified membership functions after designed ANFIS controller assume the form shown in Fig. 3.
All the rules in the fuzzy inference system include all the seven variables. These rules are shown in Fig 4.

4.3 System validation

For testing results we used k-fold cross validation. In k-fold cross validation, data set are separated in to k equal size subsets randomly and method is repeated k times. Each time one of the subsets used for testing data and remaining of them used for training data [20]. We used 10-fold cross validation.

4.4 Performances evaluation

For determining how well the ANFIS model is worked, accuracy was calculated this defined as correct classified instances divided by the total number of instances [21].

4.5 Analysis of results

To random sub sampling we used 10 cross validation for estimating performance of a predictive model. In this study, 80% (243 samples) of Cleveland data set were used for training data and 20% (60 samples) in Cleveland dataset were used for testing data. Our proposed system used two or more Gaussian membership functions for each input 0.01 was obtained as smallest error. The convergence curve of ANFIS achieved RMSE values of 0.01 (error goal) as shown in Fig. 5. The proposed system was tested with 60 samples and the rate of testing error was 0.15 which was smaller too. The result of validation training and testing data was shown in Table II.
Surface viewer is a three dimensional curve that demonstrate the mapping with two input parameters to one output for obtaining heart disease degree. Fig. 5. and Fig. 6. show result surface after training.

![Fig. 5. Surface view of max heart rate versus blood pressure](image1)

![Fig. 6. Surface view of cholesterol versus chest pain](image2)

To determine accurate diagnosis, the accuracy was obtained using this model and the result was compared with classify results of previous methods which used Cleveland heart disease dataset. The reported classification accuracy on this database [22] varies between 46.2% and 90% [20]. In this database researchers used varies data mining techniques in the diagnosis of heart disease to identify which data mining technique can provide more reliable accuracy. Table III illustrates a sample of these data mining techniques.

### Table III
Comparison of proposed method accuracy with other classifiers for the Cleveland heart disease

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy %</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suggested approach in this study</td>
<td>92.3</td>
<td>This study</td>
</tr>
<tr>
<td>Inc Net</td>
<td>90</td>
<td>Nobert Jankowski</td>
</tr>
<tr>
<td>Hybrid system</td>
<td>86.8</td>
<td>Humar Kahramanli &amp; Novruz Allahverdi</td>
</tr>
<tr>
<td>28-NN, stand, Euclid, 7 features</td>
<td>85.1± 0.5</td>
<td>WD KG</td>
</tr>
<tr>
<td>LDA</td>
<td>84.5</td>
<td>Ster&amp;Dobnikar</td>
</tr>
<tr>
<td>Fisher discriminate analysis</td>
<td>84.2</td>
<td>Ster&amp;Dobnikar</td>
</tr>
<tr>
<td>K=7, Euclid, std</td>
<td>84.2± 6.6</td>
<td>WD Gostminer</td>
</tr>
<tr>
<td>16-NN, stand, Euclid</td>
<td>84± 0.6</td>
<td>WD KG</td>
</tr>
<tr>
<td>FSM, 82.4-84% on test only</td>
<td>84.0</td>
<td>Rafal Adamczak</td>
</tr>
<tr>
<td>K=1:10, Manhattan, std</td>
<td>83.8±5.3</td>
<td>WD, GostMiner</td>
</tr>
<tr>
<td>Native Bayes</td>
<td>82.5−83.4</td>
<td>Rafal; ster, Dobnikar</td>
</tr>
<tr>
<td>SNB</td>
<td>83.1</td>
<td>Ster&amp;Dobnikar</td>
</tr>
</tbody>
</table>

5. CONCLUSION
This paper suggests a technique for classification of heart diseases for helping patients to early predict and reliable diagnosis. Adaptive Neuro-Fuzzy Inference System has been used for classification which has both the advantages of neural network and fuzzy logic. In proposed model, training and average testing errors are 0.01 and 0.15 which are very satisfying. The experimental results show that the proposed
technique has high accuracy especially when compared with other studies that use the same database of heart disease. Therefore this model is an appropriate model to classification of heart diseases. In the future an attempt will be made to develop this simulation with more input parameters for different databases.

REFERENCES
