DETERMINATION OF PERFORMANCE METRIC OF A CLASSIFICATION SYSTEM FOR MEDICAL DISEASES

by

Abdu MasanawaSagir

Hassan Usman Katsina Polytechnic, Katsina. Katsina State. Nigeria.

E-mail: amsagir@yahoo.com

Abstract

The objective of this study is to determine the performance metric of a classification system on medical diseases. Accuracy, sensitivity and specificity of a classification system were determined. The proposed classifier maximises the correctly classified data and minimises the number of incorrectly classified patterns. For robustness, the proposed method was tested with two different sets of data obtained from University of California at Irvine's (UCI) machine learning repository. In comparison, the proposed method achieves superior performance when compared to conventional approach ANFIS based gradient descent algorithm as well as to some related existing methods. The software used for the implementation is MATLAB R2016a (version 9.0) and executed in PC Intel Pentium Quad Core N3700 processor with 4.0 GB of RAM.

Keywords*: artificialneural network, classification, diagnosis.*

INTRODUCTION

Heart attack disease is among the main cause of death rate worldwide. The world health organization (WHO) estimated 17.5 million people died from cardiovascular diseases in 2012, representing 31% of all global deaths. An estimate of about 16 million deaths under the age of 70 was due to non-communicable diseases, 82% of which are in low and middle income countries. An estimated 17.7 million people died from CVDs in 2015, representing 31% of all global deaths. Of these deaths, an estimated 7.4 million were due to coronary heart disease and 6.7 million were due to stroke (WHO, 2017). A single photon emission computed tomography (SPECT) is a nuclear medicine imaging test. It uses radioactive tracers that are injected into the blood to produce pictures of your heart. Doctors use Single Photon Emission Computed Tomography (SPECT) scan of the heart to diagnose coronary artery disease and find out if a heart attack has occurred (Asadiet al., 2009).

Classification is a process that is used to find a model that describes and differentiate data classes or concepts, for the purpose of using the model to predict the class of objects whose class label is unknown (Han &Kamber, (2011); Tan, 2006). Some authors have written on classification approach such include (Carneiro et al., 2012; Gong et al., 2014; Westlake et al., 2016; Coli, 2007).

Fuzzy logic was conceived by Zadeh (1965) is a form of many valued logics in which a truth values of variables may be any real number between 0 and 1. In fuzzy logic, everything is allowed to be a matter of degree, imprecise, linguistic and perception based. Its aim is at formalisation of modes of reasoning that are approximate rather than exact (Delgada et al 2004).

Adaptive network based fuzzy inference system (ANFIS) semantically called adaptive neuro fuzzy inference system (ANFIS) is one of the hybrid neuro fuzzy inference expert systems that have the potential to capture the benefits of both artificial neural network learning rules to conclude and adjust the fuzzy inference systems (Jang, 1993; Sugeno& Kang, 1988; Takagi &Sugeno, 1985).

In the literatures about the use of classification methods, the medical sector had a number of

related work in which Takagi-Sugeno-Kang fuzzy inference system was applied.
Asadi et al. (2009) developed a new supervised multi-layer feed forward neural network (SMFFNN) model to accelerate classification with high accuracy based on SPECT and SPECTF heart disease data sets.

Rani & Deepa (2010) developed an optimal fuzzy classifier system using particle swarm optimization using Cleveland heart disease data set.
Others include (Palanisany&Kanmani, 2012; Uzer et al., 2013; El-Rafaie et al., 2012; Abushariah et al., 2014; Vadicheria&Sonawane, 2013).
This research work introduces a modified hybrid approach for training the adaptive network based fuzzy inference system (ANFIS). Sagir and Saratha (2017) emphasised the gradient base method with Levenberg-Marquardt algorithm using finite difference. In this study, Levenberg-Marquardt algorithm was modified using sparse storage technique.
This paper involved diagnosis of two medical diseases, SPECT-heart and Cleveland heart obtained from University of California at Irvine's (UCI) machine learning repository Lichman (2013). The remaining parts of this paper are organized as follows: in second part, materials and methods are presented. This led us to third part in which experiments and results were recorded and analyzed. Fourth part concludes the work, and lastly future research work was presented.

MATERIALS AND METHODS

This section deals with the materials and methods used in developing the proposed classifier for the diagnosis of medical diseases.

Description of Input Attributes

The dataset is available from UCI website, centre for machine learning and intelligent systems (Lichman 2013). SPECT-Heat dataset describes diagnosing of cardiac Single Proton Emission Computed Tomography (SPECT) images. Each of the patients is classified into two categories: normal and abnormal. The database of 267 SPECT image sets (patients) was processed to extract features that summarize the original SPECT images. As a result, 44 continuous feature patterns were created for each patient. The pattern was further processed to obtain 22 binary feature patterns. The dataset contains 267 samples (212 normal and 55 abnormal). 80 sample used in training the network while 187 samples used in testing the network. The CLIP3 algorithm was used to generate classification rules from these patterns. In this proposed method, 5 input attributes F1 to F5 are used.

Cleveland Heart disease database contains 76 attributes, but all published experiments refer to using a subset of 14 of attributes including class attribute with 303 instances but there are missing value. The class field refers to the presence of heart disease in the patient. It is integer valued from 0 (no presence) to 4. Experiments with the Cleveland database have concentrated on simply attempting to distinguish presence (values 1,2,3,4) from absence (value 0). The types of attributes include real numbers, ordered, binary and nominal. In this research work, 2-class problem using yes (1) or no (0) was proposed as in (Bhatia et al., 2008). Six input attributes are selected namely age, chest pain type, resting blood pressure, cholesterol resting blood sugar, resting electrocardiographic and maximum heart rate, which are really important for prediction in order to reduce the cost of diagnosis by avoiding many tests.

Design of a Proposed Classification System

Adaptive neuro fuzzy inference system (ANFIS) was first introduced by Jang (1993). The ANFIS is a framework of adaptive techniques to assist learning and adaptation. To illustrate the ANFIS structure, two fuzzy IF-THEN rules, according to a first order Sugeno model are to be considered.

According to Jang et al [23], if $f(x,y)$ is a first order polynomial, then the Takagi Sugeno Kang Fuzzy model is given as:

$$IF\ \mathrm{x} = \mathrm{A}_i\ and\ \mathrm{y}\ is\ B_i\ THEN\ z_i\ = f\left(x,y\right) \quad (1)$$

where A_i and B_i are fuzzy sets in the rule antecedent part, while $z = px + qy + r = f(x,y)$ is a crisp function in the rule consequent part, and *p, q & r* are the optimal consequent parameters. Usually $f(x,y)$ is a polynomial in the input variables *x* and *y*.

Hybrid Learning Algorithm of Proposed ANFIS-MLM Method

In designing this Neuro fuzzy model, parameters are calculated by the hybrid learning technique based on Least squares estimate and Modified Levenberg-Marquardt algorithm. The analytical derivation method is used for computation of the Jacobian matrix. In the ANFIS-MLM algorithm, S_1 and S_2 represent the antecedent (non linear) and consequent (linear) parameters, respectively.

Let *n* represent the number of inputs,

Let *p* represent the number of membership functions of each input

Let *l* represent the layers of the ANFIS-MLM, where $l = \{1,2,3,4,5\}$.

Let the output of node i of layers *l* be O_i^j

Forward Pass

Least squares estimate (LSE) was used at the very beginning to get the initial values of the consequent parameters $S_2 = \{p_i, q_i, r_i\}$.

Layer 1: Calculate the Membership Functions values for inputs. The Gaussian activation function was used as fuzzification nodes. The output of this layer is O_i^1, where $i = \{1,2,...,p.n\}$. O_i^1 is a membership function that satisfies the degree to which the given input satisfies the fuzzy sets A_i for $i = \{1,2,...,p\}$. The fuzzy sets are represented as a membership functions. The functions are expressed as $\mu_{A_i}(x_t;\{c,\sigma\})\ for\ \mathrm{t}\ = \{1,2,...,n\}$, where input *n* features are grid partitioned into *p* membership functions.

For simplicity, the output of two fuzzy membership grade of inputs, which are given by:

$$O_i^1 = \mu_{Ai}(x) = e^{-\frac{1}{2}\left(\frac{x-c_i}{\sigma_i}\right)}, i = 1,2 \quad (2)$$

$$O_i^1 = \mu_{Bi}(y) = e^{-\frac{1}{2}\left(\frac{y-m_i}{\beta_i}\right)}, i = 1,2 \quad (3)$$

where *x* and *y* are the inputs to node i, $[c_i, \sigma_i, m_i, \beta_i]$ is a parameter set, represents the membership function's centres and widths of $\mu_{A_i}(x)$ and $\mu_{\beta_i}(x)$ respectively. The membership functions representing the antecedent parameters of the ANFIS-MLM are described as $S_1 = \{a_i, b_i, c_i, ...\}$.

Layer 2: Calculate the rule firing strengths. Each node in this layer corresponds to a single Takagi-Sugeno type fuzzy rule. The output of this layer is O_i^2 where node $i = \{1,2,...,p^n\}$. The conjunction of rules antecedent are evaluated by either of the operator AND (minimum of incoming signals) or an OR (maximum of incoming signals).

Let R represent the rule choice of second layer nodes

$$R = \{\min\ [AND]\}\ or\ R\ =\ \{\max\ [OR]\} \quad (4)$$

For simplicity, the output O_i^2 for two fuzzy IF-THEN rules is given by:

$$O_i^2 = w_i = rule\{A_i\} = \mu_{Ai}(x) * \mu_{Bi}(y) = e^{-\frac{1}{2}\left(\frac{x-c_i}{\sigma_i}\right)} * e^{-\frac{1}{2}\left(\frac{y-m_i}{\beta_i}\right)}, i = 1, 2 \quad (5)$$

where the value w_i represents the firing strength or weights from the rule node.
Layer 3: Determine the normalised firing strengths. The ratio of the firing strength of a given rule to the sum of firing strengths of all rules is called the normalized firing strength.
Let N represent the normalisation of the node in layer 3. The output of this layer is O_i^3, where node $i = \{1, 2, ..., p^n\}$.

Let $\overline{w_i}$ represent the normalised weight of each rule

$$O_i^3\ =\ \overline{w_i}\ =\ \frac{w_i}{w_1 + w_2}\ =\ \frac{\mu_{Ai}(x) * \mu_{Bi}(y)}{\mu_{Ai}(x) + \mu_{Bi}(y)},\ i = 1, 2 \quad (6)$$

Layer 4:This is a defuzzification layer. It calculates the rules outputs for rule consequent layer. Each node in this layer is an adaptive node. The output in this layer is simply the product of the normalised firing strength and a first order polynomial (for a first order TSK model). Thus, the output of this layer O_i^4, where node $i = \{1, 2, ..., p^n\}$

Let $S_2 = [p_i, q_i, r_i]$ be the consequent parameters, which can be identified using least square estimation. A linear function f_i is expressed as a multiplication of the inputs with the corresponding consequent parameters. The output O_i^4 is the product of normalised firing strength $\overline{w_i}$ of layer 3 with the linear function f_i given by

$$O_i^4 = \overline{w_i} f_i = w_i\left(p_i x + q_i y + r_i\right),\ i = 1, 2, ...p^n \quad (7)$$

Layer 5:This is the summation layer. It is designed to calculate the sum of the output of all incoming signal, that's to compute the overall output as the summation of all incoming signals. The output of this layer is O_i^5, where node *i = {1}*. Since there is only one output, the ANFIS is a binary classifier. The output is the aggregation of all defuzzified outputs O_i^4 from layer 4, and thus it follows the weighted average.

$$O_i^5 = \sum_i \overline{w_i} f_i = \frac{\sum_i w_i f_i}{\sum_i w_i},\ i = 1, 2, ..., p^n \quad (8)$$

The least squares estimator is used to minimise the squared error $\|\mathbf{AX\text{-}B}\|^2$, where **A** = Output produced by O_i^3,**y** = Target output and **X** = Unknown consequent values related to the set of consequent parameters $p_i, q_i \& r_i$, which can be obtained using pseudo-inverse of **X**.
Following equation (10), we can develop an expression involving the normalized weights $\overline{w_i}$,.equation (7) multiplied by the inputs x_t, (layer 1), gives:

$$O_i^{5*} = \sum_i [(\overline{w_i} x) p_i + (\overline{w_i} y) q_i + \overline{w_i} r_i]\ \ for\ i\ =\ \{1, 2, ..., p^n\} \quad (9)$$

After the consequent parameters S_2 are identified, the network output can be computed and the error measure E_k represents an objective function for kth of the training data can be obtained as:

$$\mathbf{E}_k = (\mathbf{t}_k - \mathbf{a}_k)^2 \qquad (10)$$

where $\mathbf{t}_k$ and $\mathbf{a}_k$ represent the target output vector and actual output vector, and N is the number of total points. The overall error measure E of the training data set can be computed using performance measure, mean square error (MSE) defined as:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\mathbf{E}_k \qquad (11)$$

Backward Pass

In the backward pass, error signals are propagated and antecedent parameters $S_1 = \{\sigma_i, c_i\}$ are to be updated by Modified Levenberg-Marquardt algorithm.

The performance index to be optimised is defined by Madsen et al., (2004) and presented as:

$$F(\mathbf{w}) = \frac{1}{2}\mathbf{e}^T\mathbf{e} \qquad (12)$$

*F(***w***)* is the total error function, $\mathbf{w} = [w_1, w_2, ..., w_k]$ comprising of all weights of the network, **e** is the error vector comprising the error of all the training samples.

With the Modified Levenberg-Marquardt method, the increment of the parameter in training will be obtained as:

$$\Delta\mathbf{X}_k = \left(\mathbf{J}_k^T\mathbf{J}_k + \eta\mathbf{I}\right)^{-1}\mathbf{J}_k^T\mathbf{e} \qquad (13)$$

where**X** is parameter vector, $\mathbf{J}_k$ is the Jacobian matrix, and k is the index of iterations.

Results and Discussion

Based on the two data sets obtained from UCI machine learning repository, throughout the experiments, ten-fold cross validation method was used instead of hold-out validation method as applied in (Sagir& Saratha 2017). The obtained results with two data sets for robustness were compared based on machine learning process in terms of performance metric (Kahramanli&Allahverdi, 2008; Saed, 2015).

Table 1. Comparison of test accuracy results with some related existing models for SPECT Heart data set

Methodology adopt	Accuracy (%)	Sensitivity (%)	Specificity (%)	MSE	Elapsed Time
Proposed method	96.34	98.29	89.61	0.02919	2.99
Conventional method	93.46	90.44	88.03	0.02930	9.06
RS, (Palanisamy&Kanmani, 2012)	93±3.8	95.00	85.00	-------	-------
SMFFNN, (Asadi et al., 2009)	92.00	-------	-------	-------	-------
SBPN, (Asadi et al., 2009)	87.00	-------	-------	-------	-------
BPN+PCA, (Asadi et al., 2009)	73.30	-------	-------	-------	-------

The accuracy, sensitivity and specificity of the proposed classifier were obtained as 96.34 %, 98.29% and 89.61%, respectively. The MSE was obtained as 0.02919 with elapsed time as 2.99 seconds, for SPECT Heart data set, as presented in Table 2. Hyphen (---) indicates that there is no such type of result in the respective existing classifiers.

Table 2. Comparison of test accuracy results with some related existing models for Cleveland Heart data set

Methodology adopted	Accuracy (%)	Sensitivity (%)	Specificity (%)	MSE	Elapsed Time (s)
Proposed method	79.71	67.06	80.38	0.13084	0.64
Conventional method	75.56	71.05	78.85	0.40027	1.95
ANFIS, (Uzer et al. (2013)	75.93	------	------	------	------
ANN, (El-Rafaie, 2012)	65.00	------	------	------	56.50
ANN, (Rani & Deepa, 2010)	76.00	------	------	------	------

The proposed model yields better results than other models with much faster convergence speed, as presented in Table 3 for Cleveland heart data set. The accuracy, sensitivity and specificity of the proposed classifier were obtained as 79.71 %, 67.06% and 80.38%, respectively. The MSE was obtained as 0.13084 with elapsed time as 0.64 seconds. Hyphen (---) indicates that there is no such type of result in the respective existing classifiers.

Graphs of MSE Vs No. of Iterations

In each of the figure there are two curves, which show the proposed method in blue colour and the conventional method in red colour, both are trained simultaneously with the same number of epoch.

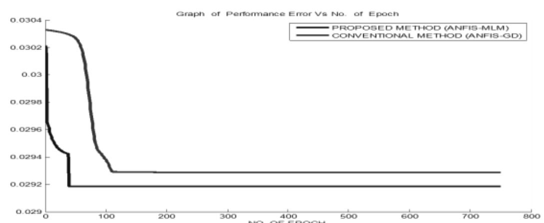

Fig. 1. Graph of MSE Vs No. of Epoch for SPECT-Heart

In Fig. 1, it appears that the proposed method outperformed conventional method as error decrease per training examples of the network and continuous to drop. The error starts stabilizing after 50 epochs for the proposed method and after 175 epochs for conventional method. The training process does not over fit the training data and the proposed method has gained and produced the estimation of generalisation in final error achieved of 0.02919 at 2.99 seconds as against 0.02930 at 9.06 seconds of conventional method, both after 750 epochs.

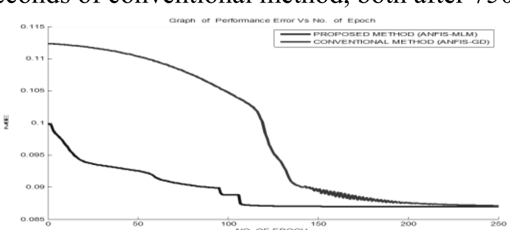

Fig.2. Graph of MSE Vs No. of Epoch for Cleveland Heart

In Fig. 2, it appears that the proposed method outperformed conventional method as error decrease per training examples of the network and continues to drop. The error starts stabilizing after 150 and 230 epochs for the proposed method and conventional method, respectively. The training process does not over fit the training data and the proposed method has gained and produced the estimation of generalisation in final error achieved of 0.13084 at 0.64 seconds as against 0.40027 at 1.95 seconds of conventional method, both after 250 epochs.

CONCLUSION AND FUTURE RESEARCH WORK

This research paper developed an effective model which incorporates the capability of fuzzy logic and artificial neural network learning algorithm that can be used by physicians to accelerate diagnosis process. For robustness, the applicability of the proposed method in data classification using two benchmark data sets in the area of medical diagnosis was demonstrated. There are some improvements of the accuracy results of the proposed method than the results of existing methods.

The proposed method could be enhanced in the future towards the improvement by applying adaptive neuro fuzzy inference system based metaheuristic algorithms like the genetic algorithm, firefly algorithm or cat swarm optimisation.

References

Abushariah, M. A., Alqudah, A. A., Adwan, O. Y., and Yousef, R. M. (2014). Automatic Heart Disease Diagnosis System Based on Artificial Neural Network (ANN) and Adaptive Neuro-Fuzzy Inference Systems (ANFIS) Approaches. Journal of SoftwareEngineering and Applications, *7*(12), 1055.

Asadi, R., Mustapha, N. Sulaiman, N. and Shiri, N.(2009). New supervised multi-layer feed forward neural network (SMFFNN) model to accelerate classification with high accuracy based on SPECT and SPECTF heart disease data sets, European Journal of Scientific Research33(1), 163-178.

Bhatia, S., Prakash, P., & Pillai, G. (2008). *SVM based decision support system for heart disease classification with integer-coded genetic algorithm to select critical features.* Paper presented at the Proceedings of the World Congress on Engineering and Computer Science, WCECS.

Carneiro G, da Silva NP, Del Bue A, Costeira JP. (2012). Artistic image classification: An analysis on the printart database. InEuropean Conference on Computer Vision, 143-157. Springer, Berlin, Heidelberg.

Coli P, Marcialis GL and Roli F. (2007). Vitality detection from fingerprint images: a critical survey. In International Conference on Biometrics, 722-731.Springer, Berlin, Heidelberg.

Delgada, M., Molina, C., Sanchez, D., Rodrigue, L and Villa, M.A. (2004). A Flexible Approach to the Multidimensional Model: The Fuzzy Datacube, Springer-Verlag Berlin Heidelberg, 26-36.

El-Rafaie, S., Salem, A.B.M., and Revett, K. (2012). *On the use of SPECT imaging datasets for automated classification of ventricular heart disease*. Informatics and Systems (INFOS), 8th International Conference on,IEEE, BIO-195.

Gong Y, Wang L, Guo R and Lazebnik S. (2014). *Multi-scale orderless pooling of deep convolutional activation features*. In European conference on computer vision, 392-407.Springer, Cham.

Han, J., Pei, J., &Kamber, M. (2011). *Data mining: concepts and techniques*: Elsevier.

Jang, J.-S.R. (1993). ANFIS: adaptive-network-based fuzzy inference system. *Systems, Man and Cybernetics, IEEE Transactions on, 23*(3), 665-685.

Jang, J.-S.R., Sun, C.-T., &Mizutani, E. (1997). Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence. Upper Saddle River, N.J: Prentice-Hall.

Kahramanli, H., and Allahverdi, N. (2008). Design of a hybrid system for the diabetes and heart diseases.Expert Systems with Applications, 35(1), 82-89.

Lichman, M. UCI Machine Learning Repository (2013), http://archive.ics.uci.edu/ml/datasets.html, last accessed 2018/02/10.

Madsen, K., Nielsen, H. B., & Tingleff, O. (2004). Methods for non-linear least squares problems [Lecture notes]. Retrieved 28February, 2018 from http://soe.rutgers.edu/~meer/GRAD561/ADD/nonlinadvanced.pdf

Palanisamy, S., & Kanmani, S. (2012). Artificial bee colony approach for optimizing feature selection. *International Journal of Computer Science 9*(3), 432-438.

Rani, C., & Deepa, S. (2010). *Design of optimal fuzzy classifier system using particle swarm optimization.* Paper presented at the 2010 International Conference on Innovative Computing Technologies (ICICT).

Saed, S. (2015). An Introduction to Data Mining. Retrieved 04 April, 2018 from http://www.saedsayad.com/model_evaluation.htm

Sagir, A.M., and Saratha, S. (2017). A Novel Adaptive Neuro Fuzzy Inference System Based Classification Models for Heart Disease Prediction. Pertanika Journal of Science and Technology, 25(1) 43-56.

Sugeno, M., & Kang, G. (1988).Structure identification of fuzzy model. *Fuzzy Sets and Systems, 28*(1), 15-33.

Takagi, T., & Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *Systems, Man and Cybernetics, IEEE Transactions on*(1), 116- 132.

Tan, Pang-Ning (2006). Introduction to data mining. Pearson Education India.

Uzer, M. S., Yilmaz, N., & Inan, O. (2013). Feature selection method based on artificial bee colony algorithm and support vector machines for medical datasets classification.

The Scientific World Journal, 2013. Vadicherla, D., and Sonawane, S. (2013). Classification Of Heart Disease Using SVM And ANN. *IJRCCT*, *2*(9), 693-701.

Westlake N, Cai H and Hall P. (2016). Detecting People in Artwork with CNNs.InEuropean Conference on Computer Vision, 825-841.Springer International Publishing.

World Health Organization. (2017). Key Facts. Retrieved 10 July, 2018 from http://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

Zadeh, L. A. (1965). Fuzzy sets. *Information and control, 8*(3), 338-353.