Pertanika J. Sci. & Technol. 25 (2): 425 - 440 (2017)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

A Hybridised Intelligent Technique for the Diagnosis of Medical Diseases

Abdu Masanawa Sagir* and Saratha Sathasivam

School of Mathematical Sciences, Universiti Sains Malaysia, 11800 Pulau Pinang, Malaysia

ABSTRACT

Medical diagnosis is the process of determining which disease or medical condition explains a person's determinable signs and symptoms. Diagnosis of most diseases is very expensive as many tests are required for predictions. This paper aims to introduce an improved hybrid approach for training the adaptive network based fuzzy inference system (ANFIS). It incorporates hybrid learning algorithms least square estimates with Levenberg-Marquardt algorithm using analytic derivation for computation of Jacobian matrix, as well as code optimisation technique, which indexes membership functions. The goal is to investigate how certain diseases are affected by patient's characteristics and measurement such as abnormalities or a decision about the presence or absence of a disease. In order to achieve an accurate diagnosis at this complex stage of symptom analysis, the physician may need efficient diagnosis system to classify and predict patient condition by using an adaptive neuro fuzzy inference system (ANFIS) pre-processed by grid partitioning. The proposed hybridised intelligent technique was tested with Statlog heart disease and Hepatitis disease datasets obtained from the University of California at Irvine's (UCI) machine learning repository. The robustness of the performance measuring total accuracy, sensitivity and specificity was examined. In comparison, the proposed method was found to achieve superior performance when compared to some other related existing methods.

Keywords: Adaptive neuro fuzzy inference system, classification, Levenberg-Marquardt algorithm, diagnosis of medical diseases

INTRODUCTION

Heart attack disease remains the main cause of death worldwide. The World Health Organisation estimated that 17.5 million people died from cardiovascular diseases in 2012, representing 31% of all deaths globally. An estimate of about 16 million deaths under the age of 70 was due to non-communicable

Article history:
Received: 09 March 2016
Accepted: 02 August 2016

E-mail addresses:
ams13_mah013@student.usm.my (Abdu Masanawa Sagir),
saratha@usm.my (Saratha Sathasivam)
*Corresponding Author

ISSN: 0128-7680

diseases, 82% of which occurred in low- and middle-income countries. Meanwhile, about 7.4 million deaths were due to coronary heart disease, whereas 6.7 million were due to stroke (World Health Organisation, 2015).

Hepatitis is an injury to the liver with inflammation of the liver cell. There are five main different types of the hepatitis virus, which are referred to as type A, B, C, D, E and possibly G. Hepatitis A and E are symptomatically acute types, whereas Hepatitis B, C and D lead to chronic diseases (Davis, 2016). The chronic hepatitis types lead to cirrhosis, which causes destruction of liver parenchymal cells.

These five types are of the greatest concern because of the burden of illness and death they cause and the potential for outbreaks and epidemic spread. Infection from these viruses results in approximately 1.45 million deaths each year (World Health Organisation [WHO], 2015).

Eighty one per cent (81%) of world's infants are vaccinated and protected from hepatitis B infection and about 2 million people with hepatitis B and C infections occur yearly through unsafe injections. These viruses are transmitted through contaminated water and food, as well as by contact with blood or bodily fluids through unsafe injections or transfusions. Infection also occurs from a mother to a child, or through sexual contact (WHO, 2005).

In order to investigate the misfortune of medical diseases, certain features must be observed. Some factors make the physician's work even more difficult to be analysed when evaluating existing test results of patients because some complicated measures are not easy to perform when considering a large number of factors. As described by Anooj (2012), Hedeshi and Abadeh (2014), the decision about the presence or absence of a patient with certain diseases depends on the physician's intuition, experience and skill for comparing current indicators with previous ones than on knowledge-rich data hidden in a database. This measure is a challenging task with regards to the large number of factors that must be considered. To achieve an accurate diagnosis at this complex stage of symptom analysis, the physician may need accurate and efficient hybridised intelligent systems that can classify and predict the likelihood of a patient getting a medical disease problem and to help in diagnosing disease.

Fuzzy logic was conceived by Zadeh (1965); it is a form of many valued logics in which a true value 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. Fuzzy logic provides a foundation for the development of new tools for dealing with natural languages and knowledge representation. Its aim is formalisation of the modes of reasoning that are approximate rather than exact. Fuzzy logic has four principal facets: logical, set theoretic, relational and epistemic (Zadeh, 2004).

The literature in the medical field includes a number of research on the use of intelligent methods, among which, the Takagi-Sugeno-Kang fuzzy inference system was applied.

Nahato, Harichandran, and Arputharaj (2015) designed an intelligent system using rough set theory with a backpropagation neural network to classify the clinical dataset. The 10-fold cross validation technique is used in their research for validation of the training phase of the classifier, including 10 attributes that were selected but then reduced to six attributes in order to produce better results. An accuracy of 90.4% was achieved in their proposed work.

Subbulakshmi and Deepa (2015) developed a hybrid methodology based on a machine learning paradigm that integrates the self-regulated learning capability of particle swarm optimisation (SRLPSO) with an extreme learning machine (ELM) classifier. Five benchmarked datasets of the UCI machine learning repository were used for handling medical data classification, namely, Heart-Statlog, Cleveland Heart Disease, Hepatitis, Wisconsin Breast Cancer and Pima Indians Diabetes. For Heart-Statlog, number of instances, number of features and number of classes used are 270, 13 and 2, respectively. The parameters used for the proposed algorithms are 500 for the maximum iterations for both SRLPSO and ELM. The classification accuracy achieved was 85.88%.

Jaganathan and Kuppuchamy (2013) described feature relevant measures based on fuzzy entropy values and devised three feature selection strategies known as Mean Selection, Half Selection and Neural Network Threshold Selection with an RBF Network classifier. The fuzzy entropy with Neural Network Threshold Selection has achieved the maximum accuracy. Five benchmark datasets are used for evaluation. These data sets are: Breast Cancer, Pima Indians Diabetes, Heart-Statlog, Hepatitis and Cleveland heart diseases datasets from the UCI Machine Learning Repository. Feature selection that is currently being proposed results in accuracies that are as good as or better than when using the entire feature set without any feature selection.

Anooj (2012) developed a heart disease risk prediction technique using a weighted fuzzy rule. The researcher used three heart disease datasets from the UCI repository, namely, Cleveland data, Hungarian data and Switzerland data. The researcher used a pre-processing stage for removing missing values and other noisy information from the selected dataset. The researcher grouped the instances based on the class label and then changed numerical data type to categorical type data based on the equi-width technique. After discretisation, the researcher selected attributes for the fuzzy rule base based on the occurrence of the attribute value in each class.

Neshat, Sargolzaei, Nadjaran Toosi, and Masoumi (2012) developed a hybrid method for diagnosis of hepatitis disease that predicts or determines whether patients with hepatitis will either live or die. The case base weighted cluster algorithm for clustering and Particle Swarm Optimisation (PSO) for classification are combined to diagnose the risky hepatitis disease. The Case – Based Reasoning (CBR) was used to diagnose the disease, while PSO was adopted to assemble a decision-making system based on the selected features and disease recognised. The method CBR-PSO has been tested to ascertain the best accuracy results. The obtained results showed that the lowest, average and best accuracy results are 77.16%, 92.83% and 94.58%, respectively.

Uzer, Yilmaz, and Inan (2013) described the feature selection method based on a hybrid approach, which are the Artificial Bee Colony (ABC) algorithm and Support Vector Machine (SVM) for three different datasets, namely, Hepatitis, Liver disorders and Diabetes obtained from UCI Machine learning repository for classification purposes. A 10-fold cross validation method was used for a reliable performance of the classifier. The classification accuracy of the proposed systems obtained was 94.92%, 74.81% and 79.29% for Hepatitis dataset, Liver disorders dataset and Diabetes dataset, respectively.

This research work introduces a modified hybrid approach for training the adaptive network based fuzzy inference system (ANFIS). Sagir and Saratha (2015) emphasised the gradient base method and Levenberg-Marquardt algorithm using finite difference. In this study, the incorporation of the hybrid learning algorithm least square estimates with the Levenberg-Marquardt algorithm using analytic derivation and chain rule for the computation of a Jacobian matrix is used.

The remaining part of this paper is organised as follows: In section 2, the design of the newly hybridised intelligent system is presented. This leads us to section 3, in which simulation results are described. Meanwhile, discussion and conclusion of the work are given in section 4.

METHODS AND MATERIALS

According to Jang, Sun, and Mizutani (1997), if f(x,y) is a first order polynomial, then the Takagi Sugeno Kang Fuzzy model is given as:

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

where A and B 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. When f(x,y) is a first–order polynomial, the resulting fuzzy inference system is called a first – order model, which was employed in this research work.

Degree the input matches' number of rule is typically computed using min operator,

$$\mathrm{w}_i = \min\left(\mu_{\mathrm{A}_i}(\mathrm{x}), \mu_{\mathrm{B}_i}(\mathrm{y})\right) \tag{2}$$

where μ_{A_i} & μ_{B_i} are membership functions that define the fuzzy sets A and B, respectively, on the universe x and i = 1,..., n.

The final output of the system is the weighted average of all rule outputs, computed as,

$$Final\ output = \frac{\sum_{i=1}^{N} \mathrm{w}_i z_i}{\sum_{i=1}^{N} w_i} \tag{3}$$

These rules are combined to get a function, which is defined as:

$$R(\mathrm{x}) = \frac{\mathrm{A}_1(\mathrm{x})\mathrm{f}_1(\mathrm{x}) + \mathrm{A}_2(\mathrm{x})\mathrm{f}_2(\mathrm{x}) + \ldots + \mathrm{A}_n(\mathrm{x})\mathrm{f}_n(\mathrm{x})}{\mathrm{A}_1(\mathrm{x}) + \mathrm{A}_2(\mathrm{x}) + \ldots + \mathrm{A}_n(\mathrm{x})} \tag{4}$$

This Takagi Sugeno Kang (TSK) fuzzy model produces a real-valued function.

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 for simplicity (Sagir & Saratha, 2015).

Image of the ANFIS Structure

Figure 1 shows the structure of the Adaptive Neuro Fuzzy Inference System (ANFIS). The structure of the proposed model contains five layers: the input and output layers and three hidden layers that represent membership functions and fuzzy rules.

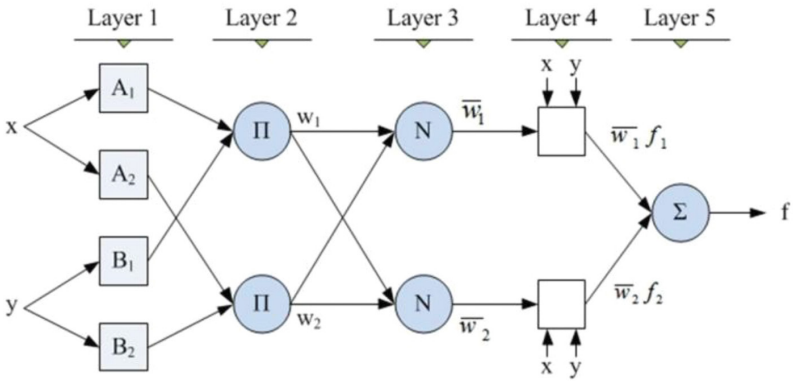

Figure 1. ANFIS-LMAD Structure

Rules Index Vector

Index Membership function is the index vector that keeps track of the unique Membership Functions (MF_s). This function determines the unique MFs in the ANFIS structure and indexes them row-wise. Therefore, the final index vector collects the indices found in MF "row-wise" according to rules. For example,

$$R = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 \end{bmatrix} \quad (5)$$

$$MF = [4\ 4\ 2\ 3\ 3\ 3\ 3] \quad (6)$$

Where, R is the rule list matrix (Jamsandekar & Mudholkar, 2014).

Finally, the index vector is:

$$ix = [1\ 5\ 9\ 11\ 14\ 17\ 20\ 26\ 10\ 12\ 15\ 18\ 21] \quad (7)$$

The Proposed ANFIS-LMAD Model Design

In designing this new ANFIS model, a hybrid learning technique, based on Least squares estimate and the Levenberg-Marquardt algorithm using the analytic difference method for computing the Jacobian Matrix, was used.

Forward Pass

Least squares estimate (LSE) was used at the very beginning to get the initial values of the conclusion parameters and then at backward pass for the Levenberg-Marquardt algorithm to update all parameters. Detail of this is found in Sagir and Saratha (2015).

Backward Pass

For the Levenberg-Marquardt algorithm, the performance index to be optimised is defined as,

$$F(w) = \frac{1}{2} E^T E \tag{8}$$

Where, E is the total error function

Error signals are propagated and the premise parameters are to be updated by the Levenberg-Marquardt algorithm:

$$\mathbf{W}_k(t+1) = \mathbf{W}_k - \left(J_k^T J_k + \mu I\right)^{-1} J_k^T E(w) \tag{9}$$

$$\Delta \mathbf{W}_k = (J_k + \mu I)^{-1} J_k^T E(w) \tag{10}$$

Get the parameters of unique MF_s of current FIS
Obtain Cumulative Current Error vector and MSE

$$\Delta \mathbf{W}_k = (J_k + \mu I)^{-1} J_k^T E(w) \tag{11}$$

where $e_{kp} = d_{kp} - o_{kp},\ k = 1,2, \dots, K,\ p = 1, \dots, P$

e_{kp} is the training error at output p when applying k pattern, d_{kp} is the desire output vector & o_{kp} is the actual output vector.

$$mse = \sqrt{\frac{E_p}{p}} \tag{12}$$

If the current total error is increased as a result of the update, then retract the step and increase the learning parameter.

Built up Jacobian matrix column – wise, which contains 1st order partial derivatives of network error by establishing the derivatives with respect to MF_s parameters ρ using analytical derivation and chain rule.

$$\frac{\partial y_i}{\partial \sigma_i} \text{ and } \frac{\partial y_i}{\partial \beta_i} \tag{13}$$

where y_i is the number of output of the network, σ_i & β_i are parameters of MF_s, i = 1, ..., n

Therefore,

$$J_{i,j} = \frac{\partial f_i}{\partial \rho_j} \tag{14}$$

Transform Jacobian into sparse matrix to speed things up

$$J_{i,j} = \text{Sparse}\left(\frac{\partial f_i}{\partial \rho_j}\right) \tag{15}$$

Approximate Hessian matrix which contains 2nd order partial derivative of network error using the cross product of Jacobian

$$H \approx J^T J \tag{16}$$

Therefore,

$$H_{i,j} = \frac{\partial^2 f_i}{\partial \rho_i \, \partial \rho_j} \tag{17}$$

Compute the error gradient

$$g = J^T E \tag{18}$$

Update the Hessian matrix

$$H^* = [H + \Phi I] \tag{19}$$

where I is the sparse identity matrix and $\Phi = 0.1$ is the learning parameter, and the network parameter needs to be updated using (10).

Then, recalculate the mse using the updated parameters

$$mse = \sqrt{\frac{E_p}{p}} \tag{20}$$

Adjust the learning parameters, if the total error is decreased as a result of the update then go to the next epoch. Then accept, otherwise reject.

Description of Input attributes for Datasets

The Statlog-Heart and Hepatitis datasets were obtained from the University of California at Irvine (UCI) machine learning repository (Bache & Lichman, 2013). Detailed information about the input variables is shown in Table 1 and Table 2.

Table 1
Information about input variables for the Statlog-Heart Disease dataset

Variable Name	Min	Max	No. of MF	Description of Input Variable	Type
CP	1	4	4	Chest pain type (1-typical angina, 2- atypical angina, 3- non-anginal pain, 4-asymtomatic)	Nominal
CHOL	126	564	4	Cholesterol (low, medium, high & very high)	Real
FBS	0	1	2	Resting blood sugar (0=false, 1=true) it is true when fbs>120	Binary
RESTECG	0	2	3	Resting electrocardiographic (0-normal, 1-having ST-T & 2-showing definite left VH)	Nominal
OPK	0	6	3	Oldpeak (low, risk, terrible)	Real
CA	0	3	3	No. of major Vessels (one, two, three)	Real
THAL	3	7	3	Thal (3 =normal, 6 = fixed defect, 7 = reverse defect)	Nominal

Table 2
Information about input variables for the Hepatitis Disease dataset

Variable Name	Min	Max	No. of MF	Description of Input Variable	Type
MAL	1	2	2	Malaise (no, yes)	Integer
SPI	1	2	2	Spiders (no, yes)	Integer
ASC	1	2	2	Ascites (no, yes)	Integer
VAR	1	2	2	Varices (no, yes)	Integer
BRB	0.3	8	3	Bilirubin (low, medium, high)	Continuous
ALB	2.1	6.4	3	Albumin (low, medium, high)	Real
HIS	2	2	2	Histology (no, yes)	Integer

RESULTS AND DISCUSSION

Membership function curves

In the Rule Based applications of Fuzzy Logic, membership functions are associated with terms that appear in the antecedents or consequents of rules. The Gaussian membership function provides smooth and non-linear functions that can be used by the learning systems like Neural Networks (Hagras, 2004).

The membership function curves for the input variables of optimised model (ANFIS-LMAD) of two datasets are shown in Figure 2 and Figure 3 respectively. The curves define how each point in the input space is mapped to the membership value (or degree of membership) between 0 and 1.

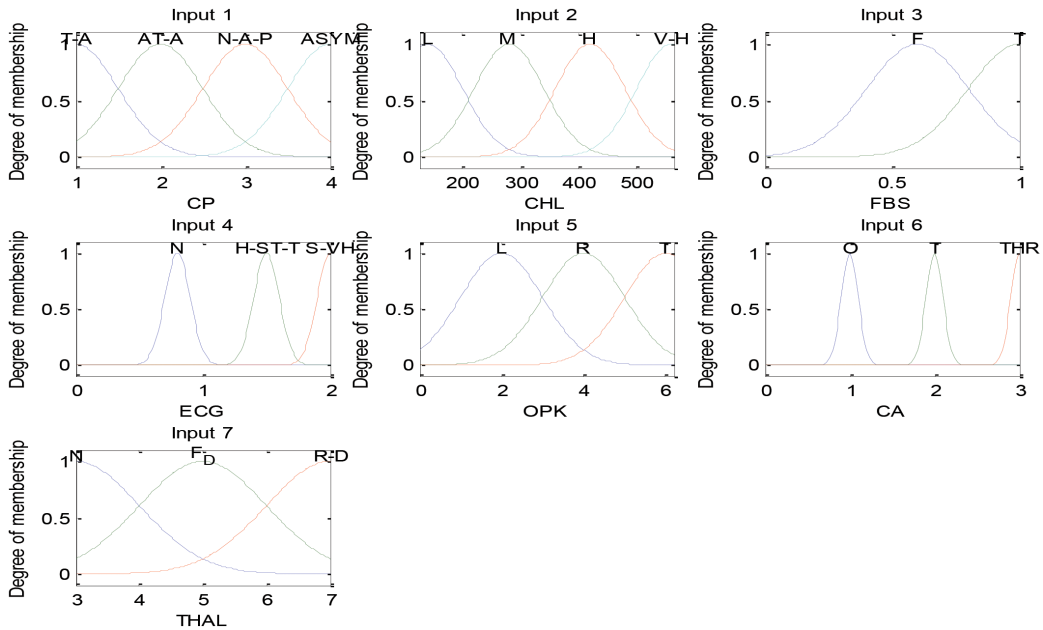

Figure 2. Membership Function Curve of Optimised Model for the Statlog-Heart Disease Dataset

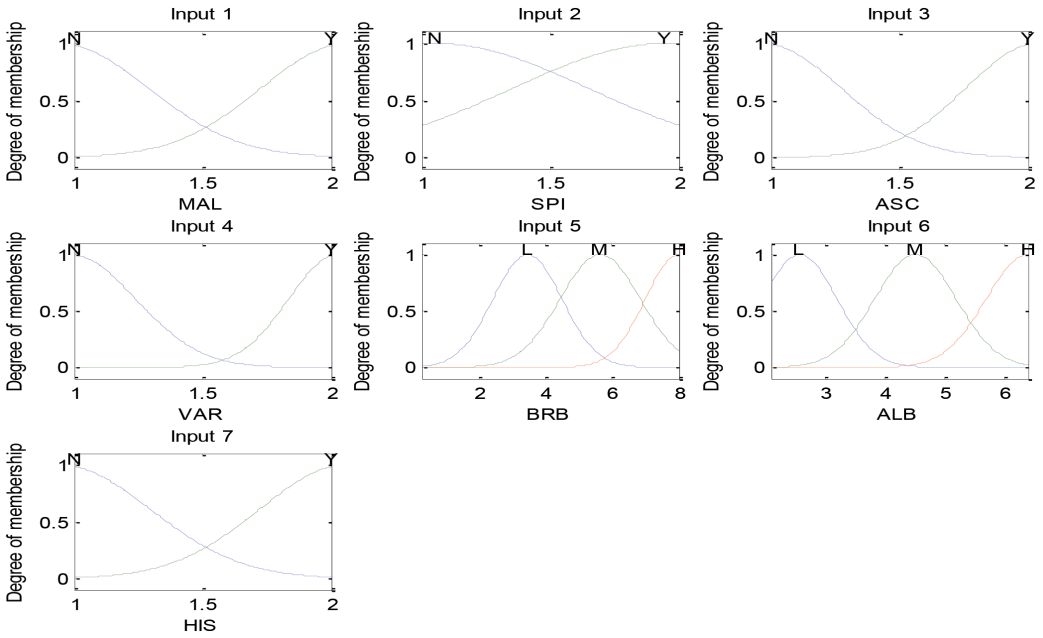

Figure 3. Membership Function Curve of Optimised Model for the Hepatitis Disease Dataset

Accuracy, Sensitivity and Specificity

The measure of the ability of the classifier to produce an accurate diagnosis is determined by accuracy. The measure of the ability of the model to identify the occurrence of a target class accurately is determined by sensitivity. The measure of the ability of the model to separate the target class is determined by specificity (Kahramanli & Allahverdi, 2008).

Table 3
Performance measure of the proposed model

	Statlog-Heart Disease Data set		Hepatitis Disease Data set	
	Train (%)	Test (%)	Train (%)	Test (%)
Sensitivity	81.71	84.21	96.97	97.67
Specificity	87.76	90.38	70.83	100.00
F-measure	83.23	85.33	93.43	98.82
Precision	84.81	86.49	90.14	100.00
Accuracy	85.00	87.78	90.00	97.86

Performance Error

Mean Square Error is one of the most acceptable indicators that describe the differences between the actual data and the predicted values. The values of premise and consequent parameters can be obtained after network training by directly minimising the MSE performance criterion (Ho, Tsai, Lin, & Chou, 2009).

$$\mathrm{MSE} = \frac{1}{\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{N}}(\mathrm{y}-\mathrm{y}')^2 \quad (20)$$

where y and y' are the desired output and predicted output respectively; and N is the number of total points.

Hence, in line with equation (21), the result for the values of performance error (mean square error), number of iterations and elapsed time of the proposed method for Statlog-Heart disease and Hepatitis disease data sets were found to be 0.10167, 1800 & 18.56762 seconds and 0.07263, 1300 & 11.30284 seconds, respectively.

Comparison of the Results

According to Kathy (2013), the train or test error rates can be obtained based on the calculation of the ratio of number of incorrect predicted values to the total number of train or test of dataset.

Table 4
Comparison of the results based on our recent work for the Statlog-Heart Disease Data set (Sagir & Saratha, 2015)

Author	Methodology Adopted	Accuracy (%)	Performance Error	Elapsed Time (second)
Proposed method	**ANFIS-LMAD**	**87.78**	**0.10167**	**18.56762**
Sagir & Saratha, (2015)	ANFIS_LSLM	76.67	0.40327	260.13590
Sagir & Saratha, (2015)	ANFIS_LSGD	75.56	0.40344	15.67020

Table 5
Comparison of the results with previous works based on performance measures and selected features for the Statlog-Heart Disease Dataset.

Methodology adopted	Accuracy (%)	Sensitivity (%)	Specificity (%)	Selected features
Proposed method	**87.78**	**84.21**	**90.38**	**3,5,6,7,10,12, 13**
PSO + ELM, Subbulakshmi & Deepa (2015)	85.88	86.00	86.03	3,11,12,13
ESUNN, Subbulakshmi & Deepa (2015)	83.22	84.32	81.65	3,8,9,11,12
EPUNN, Subbulakshmi & Deepa (2015)	81.89	83.67	84.91	8,9,11,12
NN for threshold selection, Jaganathan & Kuppuchamy (2013)	85.19	85.00	86.00	3,11,12,13
MS method, Jaganathan & Kuppuchamy (2013)	84.44	85.00	84.00	3,8,9,11,12,13
WFR+FL, Anooj (2012)	57.85	45.22	68.75	1,4,5,8,10,13

Table 6
Comparison of results with previous works based on performance measures and selected features for the Hepatitis Disease Dataset

Methodology adopted	Accuracy (%)	Sensitivity (%)	Specificity (%)	Selected features
Proposed method	**97.86**	**100.00**	**98.82**	**7,12-15,18,20**
PSO + ELM, Subbulakshmi & Deepa (2015)	97.43	93.65	95.71	6,11-14,17,19
RS-BPNN, Nahato et al. (2015)	97.30	98.32	97.28	9-17
ABC-SVM, Uzer et al. (2015)	94.92	97.13	88.33	1,5,12-15,17,18
CBR-PSO, Neshat et al. (2012)	94.58	____	____	Full features
PSO, Neshat et al. (2012)	89.46	____	____	Full features

Discussion of the Results

This research introduces an improved hybrid approach for training the adaptive network based fuzzy inference system (ANFIS). The approach incorporates hybrid learning algorithms least square estimates with the Levenberg-Marquardt algorithm, in addition to using analytic derivation for computation of Jacobian matrix instead of using the central difference method as applied in our previous research work (Sagir & Saratha, 2015).

In our previous work (Sagir & Saratha, 2015), ANFIS_LSLM classifier yields better results when compared with the ANFIS_LSGD classifier but with slightly lower convergence speed, as presented in Table 4. In the present research, the ANFIS-LMAD classifier yields a convergence time faster than the previous classifiers presented in Table 4 through Table 6.

In Table 5 and Table 6, each number in column 5 represents the selected feature for the datasets used. Detailed descriptions of the selected features of the two benchmark data sets can be found in Bache and Lichman (2013).

The present work evaluates the performance of the proposed model using two benchmark datasets for Medical diseases diagnosis. The proposed classifier was justified by the performance measures used in the present study. The test performance of the classifier was determined by the computation of sensitivity, specificity, performance error and total classification accuracy, as shown in Table 3 through Table 6, for the Statlog-Heart Disease and Hepatitis Disease datasets. The developed ANFIS-LMAD classifier yields better results than other classifiers, as presented in Table 5 and Table 6. Based on a comparison of the results, the developed technique produced reasonable results in diagnosing the possible presence of medical disease inpatients.

Hold-out validation is used by dividing the training and data sets into 70% to 30%, respectively (Schneider, 1997; Thacker et al., 2004; Fortmann-Roe, 2012). The test result is better than the training because we have not repeatedly used a holdout set to test a model during development. In other words, test data were not used during training. This shows that the classifier has learned things well and given a better desired output. The results provided good performance in accuracy and generalisation.

Compared to the literature discussed of the existing methods described above, the proposed hybridised intelligent system outperformed other existing methods. This assertion is based on the following observations: in Table 5, it is clearly confirmed that none of the research studies has success rates higher than 85.88% and selected features greater than six. In the grid partition method, the higher the input variables are, the less accuracy is to be obtained because it generates rules by enumerating all possible combinations of membership functions of all inputs. Nonetheless, our proposed method was able to produced better accuracy results. Employing an indexing membership function in a row-wise vector, which is a new way of code optimisation, however was used to make convergence time faster, and the related mentioned existing algorithms lacks this new technique. Nonetheless, unlike other existing methods, the Jacobian matrix is computed via analytic derivation and chain rule, which also contributed to a faster convergence speed by using sparse structure and produced results with better accuracy.

There are limitations associated with this study that constrain the generalisation of the results. The main limitations are the selection of two benchmark medical data sets from UCI machine learning repository and the use of the grid partitioning method, which caused by a high partition of fuzzy rules. Other approaches were taken in developing and testing the proposed model in this study such as the feature selection, learning algorithm, validation method and performance measure.

The proposed hybridised intelligent technique could be enhanced in the future by applying another means of derivation in computation of Jacobian matrix in order to increase convergence speed and accuracy of the results. The use of intelligent algorithms such as Genetic Algorithm, Particle Swarm Optimization and Ant Bee Colony algorithm with k-fold cross validation method could be used.

CONCLUSION

The objective of this study is to design a hybridised intelligent technique for the diagnosis of medical diseases. This study proposed three major novelty techniques by using and modifying the Levenberg-Marquardt algorithm, indexing membership functions using the vectorisation technique and employed an effective way to compute a sparse Jacobian matrix using analytical derivation.

The proposed technique provides a good representation of the training and test samples, input and output mapping and convergence time by using two different datasets, which make the model more robustness.

The classifier was able to learn how to classify input data by class level and correctly maximise classified data, while minimising instances of incorrectly classified patterns. Its performance was evaluated based on machine learning process.

The applicability of the proposed technique in data classification using two benchmark datasets in the area of medical diagnosis was demonstrated. The results of the proposed technique are better than the results of existing methods in the literatures and have potential in diagnosing medical diseases.

ACKNOWLEDGEMENTS

This research is partly financed by FRGS grant (203/ PMATHS/6711368) from the Ministry of Higher Education, Malaysia.

REFERENCES

Anooj, P. (2012). Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules. *Journal of King Saud University-Computer and Information Sciences, 24*(1), 27-40.

Bache, K., & Lichman, M. (2013). *UCI machine learning repository*. School of Information and Computer Sciences, University of California, Irvine.

Davis, C. P. (2016). Hepatitis (Viral Hepatitis, A, B, C, D, E, G). *MedicineNet*. Retrieved June 1, 2016, from http://www.medicinenet.com/viral_hepatitis/article.htm

Fortmann-Roe, S. (2012). *Accurately measuring model prediction error*. Retrieved December 05, 2015, from http://scott.fortmann-roe.com/docs/MeasuringError.html

Hagras, H. A. (2004). A hierarchical type-2 fuzzy logic control architecture for autonomous mobile robots. *Fuzzy Systems, IEEE Transactions on, 12*(4), 524-539.

Hedeshi, N. G., & Abadeh, M. S. (2014). Coronary artery disease detection using a fuzzy-boosting PSO approach. *Computational Intelligence and Neuroscience, 2014*(6), 1-12.

Ho, W. H., Tsai, J. T., Lin, B. T., & Chou, J. H. (2009). Adaptive network-based fuzzy inference system for prediction of surface roughness in end milling process using hybrid Taguchi-genetic learning algorithm. *Expert Systems with Applications, 36*(2), 3216-3222.

Jaganathan, P., & Kuppuchamy, R. (2013). A threshold fuzzy entropy based feature selection for medical database classification. *Computers in Biology and Medicine, 43*(12), 2222-2229.

Jamsandekar, S. S., & Mudholkar, R. R. (2014). Fuzzy Classification System by Self Generated Membership Function Using Clustering. *International Journal of Information Technology, 6*(1), 697-704.

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*. Prentice Hall, Upper Saddle River.

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

Kathy, L. T. (2013). *Data Mining Concepts*. Retrieved September 10, 2015, from http://docs.oracle.com/cd/E11882_01/datamine.112/e16808/classify.thm#DMCON035

Nahato, K. B., Harichandran, K. N., & Arputharaj, K. (2015). Knowledge mining from clinical datasets using rough sets and backpropagation neural network. *Computational and Mathematical Methods in Medicine, 2015*, 1-13.

Neshat, M., Sargolzaei, M., Toosi, A. N., & Masoumi, A. (2012). Hepatitis disease diagnosis using hybrid case based reasoning and particle swarm optimization. *ISRN Artificial Intelligence, 2012*, 1-6.

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

Schneider, J., & Moore, A. W. (2000). *A locally weighted learning tutorial using vizier 1.0.* Carnegie Mellon University, the Robotics Institute.

Subbulakshmi, C., & Deepa, S. (2015). Medical Dataset Classification: A Machine Learning Paradigm Integrating Particle Swarm Optimization with Extreme Learning Machine Classifier. *The Scientific World Journal, 2015*(2015), 1-12.

Thacker, B. H., Doebling, S. W., Hemez, F. M., Anderson, M. C., Pepin, J. E., & Rodriguez, E. A. (2004). *Concepts of model verification and validation*. SciTech Connect. United States, USA: Los Alamos National Lab.

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*(2013), 1-10.

WHO. (2015). *Fact sheets*. World Health Organisation. Retrieved August 28, 2015, from http://www.who.int/mediacentre/factsheets/en/

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

Zadeh, L. A. (2004). *Fuzzy Logic Systems: Origin, Concepts, and Trends*. Distinguished Lecture at Hong Kong Baptist University, Hong Kong.