JBAAR
Journal of Basic and Applied Research in Biomedicine
ISSN 2413-7014

Original Article

Sparse Matrix Approach in Neural Networks for Effective Medical Data Sets Classifications

Sagir Abdu Masanawa [1] and Hamza Abubakar[2*]

[1]Department of Mathematical Sciences, Federal University Dutsin-Ma, Nigeria
[2]Department of Mathematics, Isa Kaita College of Education, Dutsin-Ma, Katsina, Nigeria

*Corresponding Author: zeehamza4u2c@yahoo.com

Received: 13-8-2020
Revised: 15-9-2020
Published: 5-10-2020

Keywords:
Fuzzy logic,
Hybrid Learning Algorithm,
Adaptive neuro fuzzy
Inference system
Feed forward
Neural network

Abstract: In this paper, a hybrid intelligent system that consists of the sparse matrix approach incorporated in neural network learning model as a decision support tool for medical data classification is presented. The main objective of this research is to develop an effective intelligent system that can be used by medical practitioners to accelerate diagnosis and treatment processes. The sparse matrix approach incorporated in neural network learning algorithm for scalability, minimize higher memory storage capacity usage, enhancing implementation time and speed up the analysis of the medical data classification problem. The hybrid intelligent system aims to exploit the advantages of the constituent models and, at the same time, alleviate their limitations. The proposed intelligent classification system maximizes the intelligently classification of medical data and minimizes the number of trends inaccurately identified. To evaluate the effectiveness of the hybrid intelligent system, three benchmark medical data sets, *viz.*, Hepatitis, SPECT Heart and Cleveland Heart from the UCI Repository of Machine Learning, are used for evaluation. A number of useful performance metrics in medical applications which include accuracy, sensitivity, specificity. The results were analyzed and compared with those from other methods published in the literature. The experimental outcomes positively demonstrate that the hybrid intelligent system was effective in undertaking medical data classification tasks.

Cite this article as: Sagir, A, M. and Abubakar, H. (2020). Sparse Matrix Approach in Neural Networks for Effective Medical Data Sets Classifications. Journal of basic and applied Research in Biomedicine, 6(2): 90-97

INTRODUCTION

Many factors have also led to unhealthy behaviours, inappropriate diets and a lack of physical activity in the developed countries, which has intensified the development of chronic diseases also known as non-communicable diseases (NCDs). These NCDs are now the main contributors to the health burden causes of disability and death in low- and middle-income (LMICs) countries and disaster-prone areas accounting for 63% of all annual deaths globally(Prakash, 2017). It is responsible for about 80% of all deaths that occur in low and middle-income countries. It reported to have recorded for 71% (41 million) of the world's 57 million deaths every year and about 75% of prematurely adult deaths (in those between the ages of 30 and 69) were caused by NCDs, revealing that NCDs are not just a problem for older age groups. Detection, identification, screening, diagnosis of NCDs and emergency care are essential components of NCD response (WHO, 2019). Cardiovascular diseases (CVDs) are NCDs that causes the largest mortality worldwide, and much attention has been focused to unravel the mechanisms and optimize the treatment regimens (Li et al., 2020). Cardiovascular diseases (CVDs) are global health concerns as they cause the most deaths worldwide, and the morbidity and mortality are still on the rise (Kaptoge et al., 2019). It has been predicted that, by the year 2030, 40.5% of individuals in the United States will have CVDs, leading to approximately $818 billion for medical costs and $276 billion in indirect costs (due to lost productivity) (Tseng et al., 2019).

Medical practitioners (doctors) need an accurate diagnosis system that can accurately assess and diagnose the state of the patient, the ability to provide explanation and justification for the prediction is of paramount importance, to convince domain users (i.e., medical practitioners) with the outcome given by a computerized decision support system. This ability is essential in safety-critical applications, such as medical diagnosis and prognosis, whereby domain users need to understand and be convinced of, how the computerized system arrives at such a prediction (Seera & Lim, 2014). The elucidated rules in the form of a decision tree from the hybrid model are, therefore, important in practice, whereby the rules could serve as a source of second opinions in medical diagnostic situations (Kovalerchuk et al., 2000). Secondly, the accuracy of a decision support system is very crucial in medical applications. As stated in (Kuhn et al., 2012), a high false-negative rate of a screening system would increase the risk of patients by depriving them of getting the necessary medical attention, while a high false alarm rate would cause unnecessary worry and stress in patients as well as increase the demand on medical resources. One the other hand, as indicated in (Baty et al., 2003), a decision support system with high specificity and variable sensitivity could save medical costs and improve scheduling of vestibular patients in an otolaryngology clinic. Besides that,(Giger, 2018)also recognized the usefulness of machine learning models in reducing cost and saving time for undertaking medical diagnostic tasks. As shown in the experimental study, the proposed hybrid model not only can achieve high accuracy, sensitivity, and specificity rates but also to explain its predictions in the form of a decision tree; hence demonstrating its usefulness as a decision support system in practical environments.

Researchers have employed various data mining techniques to diagnose and treat different diseases, such as diabetes (Kavakiotis et al., 2017), liver disorder (Rajeswari & Reena, 2010), Parkinson's (Ramani & Sivagami, 2011), and breast cancer (Delen et al., 2005) and Heart disease(Luo et al., 2017). The artificial neural network (ANN) is commonly used in the classification and forecasting of disease mining(Kuruvilla & Gunavathi, 2014).

System modelling based on conventional mathematical tools (e.g., differential equations) is not well suited for dealing with ill-defined and uncertain systems. By contrast, a fuzzy inference system employing fuzzy if-then rules can model the qualitative aspects of human knowledge and reasoning processes without employing precise quantitative analyses. This fuzzy modelling or fuzzy identification, first explored systematically by Takagi and Sugeno (Takagi & Sugeno, 1985), has found numerous practical applications in control (Scharf, 1987; Rutkowski *et al.*, 2013), prediction and inference(Kazeminezhad et al., 2005). Adaptive Neuro-Fuzzy Inference System (ANFIS) has been described as one of the hybrid neuro-fuzzy inference expert systems capable of capturing the advantages of both artificial neural network learning principles to complete and modify the Fuzzy Inference Systems (Ahmed & Shah, 2017). Neuro-fuzzy is a prototype artificial intelligence technique, a fusion of fuzzy logic and artificial neural networks. Through ANFIS, cognitive capabilities and the conceptual framework of artificial neural networks are paired with the decision-making process of fuzzy logic. ANFIS is learning with experiments using train data set as in artificial neural networks. In this way, the most optimal ANFIS configuration is obtained to solve the related problem. The structure collected is exposed to a test protocol in order to see its effect on materials that it has never seen before. One of the most important drawbacks of the artificial neural network is that it has not been possible to explain the weight values obtained. The fuzzy inference system discovered in the ANFIS structure eliminates this drawback.

The health sector has had several related articles in the literature on the use of diagnosis and procedures for the identification of diseases based on Takagi-Sugeno-Kang's fuzzy inference system. A hybrid hepatitis disease diagnostic model has been used to predicts or dictates whether hepatitis patients will either lives or dies(Nilashi et al., 2019). A lot of models were developed based on ANFIS structure for detection and classification of data used in various disciplines. This includes student simulation method (Dimitriou et al., 2008), health care system (Reddy & Rothschild, 2002), financial and economic model (Chen & Chen, 2015) and intrusion detection(El-Semary et al., 2005) are some of the fields in which ANFIS model for their analysis been applied. Other researches on that involve integrating ANFIS model with another model as a single network to improve the performance for data classification (Dash, Nayak and Behera, 2015; AlMuhaideb and Menai, 2014; Tahmasebi and Hezarkhani, 2012; Asadi *et al.*, 2009; El Rafaie, Salem and Revett, 2012; Neshat *et al.*, 2012; Abubakar et al., 2020).

Many scholars have successfully combined the advantage of fuzzy logic and ANFIS. However, there is no recent effort to training the adaptive network-based fuzzy inference system for Cardiovascular diseases detection. This study proposes an updated hybrid approach for training the Fuzzy inference system based on the adaptive network (ANFIS). We emphasised on work of (Sagir and Sathasivam, 2017; Abushariah *et al.*, 2014) based on the gradient base method with Levenberg-Marquardt algorithm using finite-difference. The Levenberg-Marquardt algorithm was modified using sparse storage technique and vectorisation technique was introduced towards better performance results.

This paper associated the diagnosis of three medical issues, hepatitis, SPECT-heart, and Cleveland-heart obtained from Irvine's (UCI) machine learning repository at the University of California(Asuncion & Newman, 2007). The remaining parts of this paper are organized as follows: materials and methods are presented in the second part of this work. This led us to the third part of the recording and analyzing experiments and results. The work is finished in the fourth part, and future research work has been presented.

MATERIALS AND METHODS

This section addresses the materials and methods used in the development of the proposed model for an intelligent classification system. All the data sets mentioned below can be found in the UCI machine learning repository.

Design of a Proposed Intelligent System

For the first time, (Abushariah et al., 2014) introduced the Adaptive Neuro-Fuzzy Inference Framework (ANFIS). ANFIS can be described as an adaptive strategy framework to encourage growth and adaptation. To explain the structure of ANFIS, two fuzzy IF-THEN principles are to be interpreted in conjunction with a first-order model of Sugeno**.** According to Sagir and Sathasivam (2017), it $f(r,l)$ is described as the first order polynomial, given the Takagi Sugeno Kang Fuzzy model in Eq.(1):

$$IF\ \mathrm{r} = \mathrm{A}_i\ and\ \mathrm{l}\ is\ B_i\ THEN\ z_i\ = f\left(r,l\right) \qquad (1)$$

where A_i and B_i represent the fuzzy sets in the rule antecedent part, while given by $z = pr + ql + x = f(r,l)$ is identified as a crisp function in the rule consequent part, and *p, q & x* represent the optimal consequent parameters. Usually $f(r,l)$ described a polynomial in the input variables *r* and *l.*

Hybrid Learning Algorithm

All Parameters can be determined by the hybrid learning framework based on the estimation of Least-squares and the Modified Levenberg-Marquardt while modelling the Neuro-fuzzy model. The analytical method of derivation is being applied to calculate the Jacobian matrix. In the ANFIS-MLM algorithm, S_1 and S_2 described the antecedent (non-linear) and consequent (linear) parameters, respectively.

Let *n, p and l* define the number of membership functions of each input, the number of inputs, and the layers of the ANFIS-MLM respectively, where is given as $l = \{1,2,3,4,5\}$.

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

Forward Pass

At the very beginning, Least Squares Estimate (LSE) was used to derive the initial values of the parameters. $S_2 = \{p_i, q_i, x_i\}$.

Layer 1: compute 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 represent a membership function that satisfies the degree to which the given input achieved the fuzzy sets A_i for $i = \{1,2,...,p\}$. The fuzzy sets are represented as a membership functions. The functions are expressed as $\mu_{A_i}(r_t;\{c,\sigma\})\ for\ \mathrm{t} = \{1,2,...,n\}$, where input *n* features are grid partitioned into *p* membership functions.

The output of two fuzzy membership grade of inputs will be given by:

$$\mathrm{O}_\mathrm{i}^1 = \mu_{\mathrm{Ai}}\left(r\right) = \mathrm{e}^{-\frac{1}{2}\left(\frac{r-\mathrm{c_i}}{\sigma_\mathrm{i}}\right)}, \mathrm{i} = 1,2 \qquad (2)$$

$$\mathrm{O}_\mathrm{i}^1 = \mu_{Bi}\left(l\right) = \mathrm{e}^{-\frac{1}{2}\left(\frac{l-m_\mathrm{i}}{\beta_\mathrm{i}}\right)}, \mathrm{i} = 1,2 \qquad (3)$$

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

Let A_p represent the number of antecedent parameters for each membership function $\mu_{A_i}(r)$. The total number of antecedent parameters T_A is presented as follows:

$$T_A\ = A_p * (p * n) \qquad (4)$$

Layer 2: compute 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 the node $i = \{1,2,..., p^n\}$. The conjunction of rules antecedents is 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

$$R = \{\min\ [AND]\}\ or\ R\ =\ \{\max\ [OR]\} \qquad (5)$$

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

$$O_i^2 = w_i = rule\{A_i\} = \mu_{Ai}\left(r\right) * \mu_{Bi}\left(l\right) = e^{-\frac{1}{2}\left(\frac{r-c_i}{\sigma_i}\right)} * e^{-\frac{1}{2}\left(\frac{l-m_i}{\beta_i}\right)}, i = 1,2 \qquad (6)$$

where the value $\mathbf{w_i}$ described the firing strength or weights from the rule node.

Layer 3: Determine the normalized firing strengths. The ratio of the firing force of a given

rule to the total of all rules firing forces is named the normalized firing force.
Let *N* represent the normalisation of the node in layer 3. The output of this layer is defined as O_i^3, where a node $i = \{1, 2, ..., p^n\}$

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

$$O_i^3 = \overline{w_i} = \frac{w_i}{w_1 + w_2} = \frac{\mu_{Ai}(r) * \mu_{Bi}(l)}{\mu_{Ai}(r) + \mu_{Bi}(l)},\ i = 1, 2 \qquad (7)$$

Normalisation is done to ensure stable convergence of weights and biases to avoids the time-consuming process of defuzzification.

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

Let $S_2 = [p_i, q_i, x_i]$ be the consequent parameters that can be defined using the least square estimate. 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 normalized firing strength $\overline{w_i}$ of layer 3 with the linear function f_i given by

$$O_i^4 = \overline{w_i} f_i = w_i\left(p_i r + q_i l + x_i\right),\ i = 1, 2, ...p^n \qquad (8)$$

and the total number of consequent parameters T_C represented as follows,

$$T_C = (n+1) * p^n \qquad (9)$$

Layer 5: This is the summation layer. It is designed to calculate the sum of the output of all incoming signal, that's to calculates the total output as a description of all incoming signals. The output of this layer is defined as 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.

$$O_i^5 = \sum_i \overline{w_i} f_i = \frac{\sum_i w_i f_i}{\sum_i w_i},\ i = 1, 2, ..., p^n \qquad (10)$$

The least-squares estimator is used to minimize the squared error $||\mathbf{AX}-\mathbf{B}||^2$, where **A** = Output produced by O_i^3, **y** = Target output and **X** = Unknown consequent values associated to the set of consequent parameters $p_i,\ q_i\ \&\ x_i$, which can be obtained using pseudo-inverse of **X**.

Following equation (10), we can develop an expression involving the normalized weights $\overline{w_i}$,.eq. (7) multiplied by the inputs r_t , (layer 1), gives:

$$O_i^{5*} = \sum_i [(\overline{w_i} r) p_i + (\overline{w_i} l) q_i + \overline{w_i} x_i]\ \ for\ i = \{1, 2, ..., p^n\} \qquad (11)$$

After the corresponding (consequent) parameters S_2 are defined, the network output can be measured and determine the error, E_k described an objective function for kth of the training data can be computed as:

$$\mathbf{E}_k = (\mathbf{t}_k - \mathbf{a}_k)^2 \qquad (12)$$

where $\mathbf{t}_k$ and $\mathbf{a}_k$ defined the target output vector and actual output vector, and N described the number of total points. The overall error measure E of the training data set can be measured using Mean Square Error (MSE) as the performance indicator defined as follows

$$MSE = \frac{1}{N} \sum_{i=1}^{N} \mathbf{E}_k \qquad (13)$$

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 optimized as follows.

$$F(\mathbf{w}) = \frac{1}{2} \mathbf{e}^T \mathbf{e} \qquad (14)$$

where *F(***w***)* represents the total error function, $\mathbf{w} = [w_1, w_2, ..., w_k]$ consisting of all weights vector of the network, **e** defined the error vector consisting of the error of all the training samples.
The parameters of the present fuzzy inference system (FIS)'s unique membership functions are to be collected, which is a novel approach that enables the program to run more quickly described as:

$$\mathbf{v} = I(\mathbf{R}_{ij}) \qquad (15)$$

where **v** described as the rules index vector that keeps track of the unique membership functions MFs, *I* is considered as the index table of the unique MF applied in the rules, $\mathbf{R}_{ij}$ defined a size number matrix of the rule by the number of inputs representing the membership functions of the rule *ith* rule and *jth* input.
The Jacobian matrix is constructed-up column-wise, which includes partial system error derivatives of first order using analytical derivation and the model for chain rule is given by:

$$\mathbf{J}_k = \frac{\partial f_i}{\partial \rho_j} = \begin{bmatrix} \partial y_1/\partial\sigma_1 & \partial y_1/\partial\beta_1 & \dots & \partial y_1/\partial\sigma_{N_f} & \partial y_1/\partial\beta_{N_f} \\ \partial y_2/\partial\sigma_1 & \partial y_2/\partial\beta_1 & \dots & \partial y_2/\partial\sigma_{N_f} & \partial y_2/\partial\beta_{N_f} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \partial y_{N_t}/\partial\sigma_1 & \partial y_{N_t}/\partial\beta_1 & \dots & \partial y_{N_t}/\partial\sigma_{N_f} & \partial y_2/\partial\beta_{N_f} \end{bmatrix} \quad (16)$$

A novelty approach was introduced to transform Jacobian into a sparse Jacobian matrix to Speed the process up. With scattered(Sparse) matrix storage, storing the J_k rows in a compacted form is generally practical, i.e. without zero entries.

$$\mathbf{J}_k = Sparse\left(\frac{\partial f_\mathrm{i}}{\partial \rho_\mathrm{j}}\right) \quad (17)$$

Applying the chain rule, the gradient and Hessian of f(**x**) can be represented in terms of the Jacobian,

$$F^{'}(\mathbf{X}) = \mathbf{g} = \mathbf{J}_k^T\mathbf{e} \quad (18)$$

The approximate Hessian matrix representing second-order partial network error derivative employing Jacobian's cross product can be derived as

$$F^{''}(\mathbf{X}) = \mathbf{H} \approx \mathbf{J}_k^T\mathbf{J}_k \quad (19)$$

This approximation is applied in both Gauss-Newton and Levenberg-Marquardt methods. Hence, it is possible to avoid the explicit computation of second-order derivatives(Pujari et al. 2016). A hessian matrix as defined in the Gauss-Newton method is not invertible. To ensure that the equation is invertible, a further approximation of the Hessian matrix is implemented and can be modified as,

$$\mathbf{H}^* = \mathbf{J}_k^T\mathbf{J}_k + \eta\mathbf{I} \quad (20)$$

where η is described as a combination coefficient or learning rate parameter, **I** defined the sparse identity matrix.

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

$$\Delta\mathbf{X}_\mathrm{k} = \left(\mathbf{J}_k^T\mathbf{J}_\mathrm{k} + \eta\mathbf{I}\right)^{-1}\mathbf{J}_\mathrm{k}^\mathrm{T}\mathbf{e} \quad (21)$$

RESULTS AND DISCUSSION

Based on the three data sets extracted from the University of California Irvine (UCI) machine learning repository, throughout the experiments, the ten-fold cross-validation method was applied instead of the hold-out validation method as applied in our previous method (Abushariah et al., 2014). The results obtained with three data sets for reliability purposes were compared in terms of accuracy, sensitivity, specificity and execution time based on the machine learning mechanism.

The proposed classifier yields better results with faster convergence speed than other classifiers, as presented in Table 1. The accuracy, sensitivity and specificity of the proposed classifier for Hepatitis datasets were obtained as 95.18 %, 95.04% and 97.00%, respectively. The MSE was found to be 0.0743, standard deviation (SD) 0.0525 with an elapsed time of 12.63 seconds. The MLM algorithm is more effective and achieves a lower MSE and higher mapping precision. This is presented in Table 1 – Table 3. The asterisk (*) indicates that there is no such type of result in the respective existing classifiers.

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, standard deviation (SD) 0.02064 with an elapsed time of 2.99 seconds, for SPECT Heart data set, as presented in Table 2. The asterisk (*) indicates that there is no such type of result in the respective existing classifiers.

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

Methodology	Accuracy	Sensitivity	Specificity	MSE	SD	Time (Secs)
Proposed method	95.18	95.04	97.00	0.0743	0.0525	12.63
Conventional method	93.33	93.20	94.26	0.0817	0.0578	27.46
ABC-SVM (El Rafaie, 2012)	94.92	97.13	88.33	*	*	*
CBR-PSO (Neshat et al., 2012)	94.58	*	*	*	*	*
ABC-Boosting (Asadi et al., 2009)	83.44	*	*	*	*	*

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

Methodology	Accuracy	Sensitivity	Specificity	MSE	SD	Time (Secs)
Proposed method	96.34	98.29	89.61	0.0291	0.0206	2.99
Conventional method	93.46	90.44	88.03	0.0293	0.0207	9.06
RS (Asadi et al., 2009)	93±3.8	95.00	85.00	*	*	*
SMFFNN(Neshat et al., 2012)	92.00	*	*	*	*	*
SBPN (Neshat et al., 2012)	87.00	*	*	*	*	*
BPN+PCA (Neshat et al., 2012)	73.30	*	*	*	*	*

Table 3. Comparison of test accuracy results with some related existing models for the Cleveland Heart data set

Methodology	Accuracy	Sensitivity	Specificity	MSE	SD	Time (Secs)
Proposed method	79.71	67.06	80.38	0.13084	0.09252	0.64
Conventional method	75.56	71.05	78.85	0.40027	0.28303	1.95
ANFIS(Talei et al., 2010)	75.93	*	*	*	*	*
ANN(Delgado et al., 2004)	76.00	*	*	*	*	*

The model proposed gives better results compared with other models with a much faster convergence speed, as shown in y 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 0.1308, standard deviation (SD) 0.0925 with an elapsed time of 0.64 seconds. The asterisk (*) indicates that there is no such type of result in the respective existing classifiers.

Graphs of MSE Vs No. of Iterations

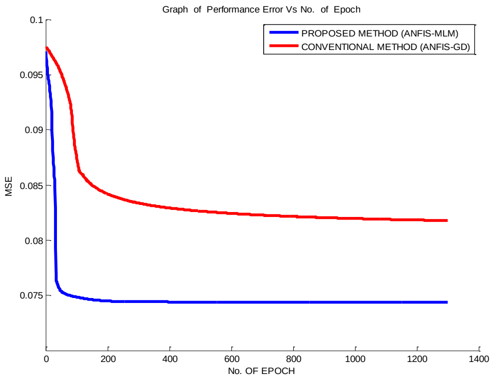

Figure 1. Graph of MSE Vs No. of Iteration for Hepatitis Data set

In Figure 1, it appears that the proposed method outperforms the conventional method as error decreases per training examples of the network and continues to drop. The error starts stabilizing after 200 iterations for the proposed method and after 350 iterations for the conventional method. This is because the conventional method has weaker convergence rates as the number of iterations is very high. The training process does not overfit the training data and the proposed method has gained and produced the estimation of generalisation in a final error achieved of 0.07435 at 12.63 seconds as against 0.08177 at 27.46 seconds of the conventional method, both after 1300 iterations.

In Figure 2, it appears that the proposed method outperforms the conventional method as error decreases per training examples of the network and continues to drop. The error starts stabilizing after 50 iterations for the proposed method and after 175 iterations for the conventional method. The training process does not overfit the training data and the proposed method has gained and produced the estimation of generalisation in a final error achieved of 0.02919 at 2.99 seconds as against 0.02930 at 9.06 seconds of the conventional method, both after 750 iterations.

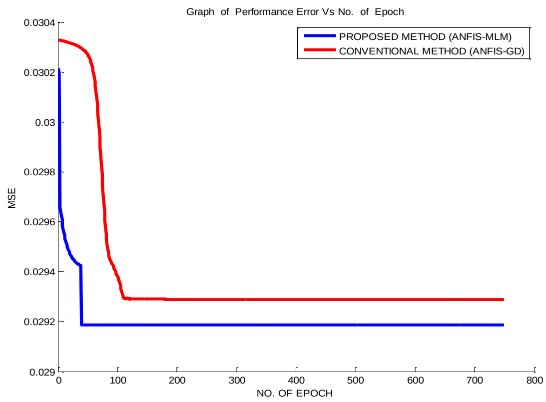

Figure 2. Graph of MSE Vs No. of Iteration for SPECT-Heart

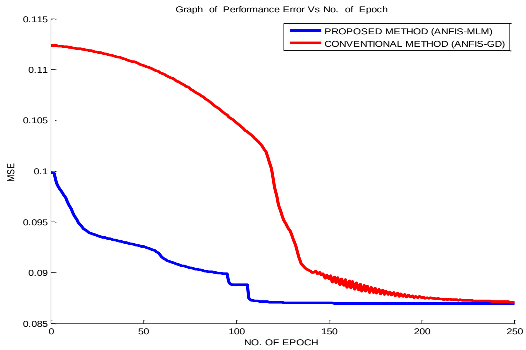

Figure 3. Graph of MSE Vs No. of Iteration for Cleveland Heart

In Figure 3, it appears that the proposed method outperforms the conventional method as error decreases per training examples of the network and continues to drop. The error starts stabilizing after 150 and 230 iterations for the proposed method conventional method, respectively. The training process does not overfit the training data and the proposed method has gained and produced the estimation of generalisation in a final error achieved of 0.13084 at 0.64 seconds as against 0.40027 at 1.95 seconds of the conventional method, both after 250 iterations.

CONCLUSION AND FUTURE RESEARCH WORK

The objective of this study is to develop an effective intelligent classification system. This study proposed a major novelty technique by the incorporative sparse matrix in neural network learning algorithm for speed up the processing of the data and reduces the system storage space. For robustness, the applicability of the proposed method in data classification using three benchmark data sets in the area of medical diagnosis was demonstrated. There are improvements to the accuracy results of the proposed method than the results of the existing methods. The proposed method was able to cope up with erroneous data as shown in figure 1- 3.

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

REFERENCES

Abubakar, H., Rijal, S., Sabri, M., Masanawa, S. A., & Yusuf, S. (2020). Modi fi ed election algorithm in hop fi eld neural network for optimal random k satisfiability representation. *International Journal for Simulation and Multidisciplinary Design Optimization*, 11(16).

Abushariah, M. A., Alqudah, A. A., Adwan, O. Y., & Yousef, R. M. (2014). Automatic heart disease diagnosis system based on artificial neural network and adaptive neuro-fuzzy inference systems (ANFIS) approaches. *Journal of software engineering and applications*, *7*(12), 1055.

Abilash, N., & Sivapragash, M. J. K. S. U. (2016). Optimizing the delamination failure in bamboo fiber reinforced polyester composite. *Journal of King Saud University-Engineering Sciences*, *28*(1), 92-102.

AlMuhaideb, S., & Menai, M. E. B. (2014). A new hybrid metaheuristic for medical data classification. *International Journal of Metaheuristics*, *3*(1), 59-80.

Asadi, R., Mustapha, N., Sulaiman, N., & Shiri, N. (2009). New supervised multi layer feed forward neural network model to accelerate classification with high accuracy. *European Journal of Scientific Research*, *33*(1), 163-178

Asuncion, A., & Newman, D. J. (2007). *UCI Machine Learning Repository: Data Sets*. University of California Irvine School of Information. *http://www.ics.uci.edu/~ mlearn/MLRepository. html.*

Baty, B. J., Kinney, A. Y., & Ellis, S. M. (2003). Developing culturally sensitive cancer genetics communication aids for African Americans. *American Journal of Medical Genetics*. *118*(2), 146-155.

Chen, M. Y., & Chen, B. T. (2015). A hybrid fuzzy time series model based on granular computing for stock price forecasting. *Information Sciences*. *294*, 227-241.

Pujari, D., Yakkundimath, R., & Byadgi, A. S. (2016). SVM and ANN based classification of plant diseases using feature reduction technique. *IJIMAI*, *3*(7), 6-14.

Dash, T., Nayak, S. K., & Behera, H. S. (2015). Hybrid gravitational search and particle swarm based fuzzy MLP for medical data classification. In *Computational Intelligence in Data Mining-Volume 1* (pp. 35-43). Springer, New Delhi.

Delen, D., Walker, G., & Kadam, A. (2005). Predicting breast cancer survivability: a comparison of three data mining methods. *Artificial intelligence in medicine*, *34*(2), 113-127.

Delgado, M., Molina, C., Sánchez, D., Ariza, L. R., & Vila, M. A. (2003, November). A flexible approach to the multidimensional model: The fuzzy datacube. In *Conference on Technology Transfer* (pp. 26-36). Springer, Berlin, Heidelberg.

Dimitriou, L., Tsekeris, T., & Stathopoulos, A. (2008). Adaptive hybrid fuzzy rule-based system approach for modeling and predicting urban traffic flow. *Transportation Research Part C: Emerging Technologies*, *16*(5), 554-573.

El, A., Edmonds, J., Gonzalez, J., & Papa, M. (2005, May). A framework for hybrid fuzzy logic intrusion detection systems. In *The 14th IEEE International Conference on Fuzzy Systems, 2005. FUZZ'05.* (pp. 325-330). IEEE.

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

Giger, M. L. (2018). Machine learning in medical imaging. *Journal of the American College of Radiology*, *15*(3), 512-520.

Kaptoge, S., Pennells, L., De Bacquer, D., Cooney, M. T., Kavousi, M., Stevens, G., ... & Amouyel, P. (2019). World Health Organization cardiovascular disease risk charts: revised models to estimate risk in 21 global regions. *The Lancet Global Health*, *7*(10), e1332-e1345.

Kavakiotis, I., Tsave, O., Salifoglou, A., Maglaveras, N., Vlahavas, I., & Chouvarda, I. (2017). Machine learning and data mining methods in diabetes research. *Computational and structural biotechnology journal*, *15*, 104-116.

Kazeminezhad, M. H., Etemad-Shahidi, A., & Mousavi, S. J. (2005). Application of fuzzy inference system in the prediction of wave parameters. *Ocean Engineering*, *32*(14-15), 1709-1725.

Kovalerchuk, B., Vityaev, E., & Ruiz, J. F. (2000). Consistent knowledge discovery in medical diagnosis. *IEEE Engineering in Medicine and Biology Magazine*, *19*(4), 26-37.

Kuhn, M. J., Mahfouz, M. R., Rowe, N., Elkhouly, E., Turnmire, J., & Fathy, A. E. (2012, January). Ultra wideband 3-D tracking of multiple tags for indoor positioning in medical applications requiring millimeter accuracy. In *2012 IEEE Topical Conference on Biomedical Wireless Technologies, Networks, and Sensing Systems (BioWireleSS)* (pp. 57-60). IEEE.

Kuruvilla, J., & Gunavathi, K. (2014). Lung cancer classification using neural networks for CT images. *Computer methods and programs in biomedicine*, *113*(1), 202-209.

Li, H., Sureda, A., Devkota, H. P., Pittalà, V., Barreca, D., Silva, A. S., ... & Nabavi, S. M. (2020). Curcumin, the golden spice in treating cardiovascular diseases. *Biotechnology advances*, *38*, 107343.

Luo, Y., Li, Z., Guo, H., Cao, H., Song, C., Guo, X., & Zhang, Y. (2017). Predicting congenital heart defects: A comparison of three data mining methods. *PLoS One*, *12*(5), e0177811.

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

Nilashi, M., Ahmadi, H., Shahmoradi, L., Ibrahim, O., & Akbari, E. (2019). A predictive method for hepatitis disease diagnosis using ensembles of neuro-fuzzy technique. *Journal of infection and public health*, *12*(1), 13-20.

Prakash, S. (2017). Non-communicable diseases (NCDs): a global challenge. *Biomed J Sci Tech Res*, *1*(2), 325-326.

Rajeswari, P., & Reena, G. S. (2010). Analysis of liver disorder using data mining algorithm. *Global journal of computer science and technology 10*(14):.48-52

Ramani, R. G., & Sivagami, G. (2011). Parkinson disease classification using data mining algorithms. *International journal of computer applications*, *32*(9), 17-22.

Reddy, N. P., & Rothschild, B. M. (2002, October). Hybrid fuzzy logic committee neural networks for classification in medical decision support systems. In *Proceedings of the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society][Engineering in Medicine and Biology* (Vol. 1, pp. 30-31). IEEE.

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

Scharf, E. M. (1987). Industrial Applications of Fuzzy Control: North-Holland, Amsterdam, 1985, viii+ 269 pages, Dfl. 150.00.

Seera, M., & Lim, C. P. (2014). A hybrid intelligent system for medical data classification. *Expert Systems with Applications*, *41*(5), 2239-2249.

Tahmasebi, P., & Hezarkhani, A. (2011). Application of a modular feedforward neural network for grade estimation. *Natural resources research*, *20*(1), 25-32.

Takagy, T., & Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on System, Man and Cybernetics*, *15*(1), 116-132.

Talei, A., Chua, L. H. C., & Quek, C. (2010). A novel application of a neuro-fuzzy computational technique in event-based rainfall–runoff modeling. *Expert Systems with Applications*, *37*(12), 7456-7468.

Tseng, Z. H., Salazar, J. W., Olgin, J. E., Ursell, P. C., Kim, A. S., Bedigian, A., ... & Vittinghoff, E. (2019). Refining the World Health Organization definition: predicting autopsy-defined sudden arrhythmic deaths among presumed sudden cardiac deaths in the POST SCD Study. *Circulation: Arrhythmia and Electrophysiology*, *12*(7), e007171.

Shojaei, E., Wong, A., Rexachs, D., Epelde, F., & Luque, E. (2020). A Method for Projections of the Emergency Department Behaviour by Non-Communicable Diseases from 2019 to 2039. *IEEE journal of biomedical and health informatics*, *24*(9), 2490-2498.

.