# EXPERT SYSTEM FOR THE IDENTIFICATION OF CHRONIC KIDNEY DISEASE

**SAGIR, A. M.**
Department of Basic Studies,
Hassan Usman Katsina Polytechnic, Katsina State, Nigeria
**E-mail:** amsagir@yahoo.com **Phone No:** +234-803-947-4856

**Abstract**
*Chronic Kidney Diseases (CKDs) are mainly caused by diabetes and high blood pressure. For early detection, certain tests have to be taken such as blood tests, urine tests and imaging tests. The decision about presence or absence of chronic kidney disease depends on the physician's intuition, experience and skill for comparing current indicators with previous one than on knowledge rich data hidden in a database. This measure is a very critical and challenging task. The objective of this paper is to predict patient condition by using an Adaptive Neuro Fuzzy Inference System (ANFIS) pre-processed by grid partitioning. A framework describes methodology for developing and evaluation of classification performances of the proposed system using hybrid learning algorithm least square estimates with Levenberg-Marquardt algorithm that can be used by physicians to accelerate identification process. The proposed method's performance was evaluated with Chronic Kidney Data set obtained from benchmarked datasets of University of California at Irvine's (UCI) machine learning repository. The performance measuring total accuracy, sensitivity and specificity were examined. In comparison, the proposed method achieves superior performance when compared to conventional expert based gradient descent algorithm and some related existing methods. The software used for the implementation is MATLAB R2014a (version 8.3) and executed in PC Intel Pentium IV E7400 processor with 2.80 GHz speed and 2.0 GB of RAM.*

**Keywords:** *Adaptive Neuro Fuzzy Inference System, Chronic Kidney Disease, Grid Partition method, Identification, Levenberg-Marquardt Algorithm.*

## Introduction

The rising prevalence of Chronic Kidney Disease (CKD) is emerging as a major global health problem (Wachukwu, Emem-Chioma, Wokoma & Oko-Jaja, 2015). In the developed countries like Nigeria, the cause of the rise of CKD appears not to be due to intrinsic renal disease but to the dramatic rise in systemic diseases that damage the kidney, such as high blood pressure, hypertension and type 2 diabetes. CKD is a prevalent and potentially escalating disease across sub-Saharan Africa with risk factors that include both communicable and non-communicable diseases (Afolabi, Abioye-Kuteyi, Arogundade & Bello, 2009).

The National Kidney Foundation (NKF) estimates that 20 million Americans have chronic kidney disease and at least a further 20 million people have an increased risk (NKF, 2002; Johnson, Levey, Coresh, Levin, Lau & Eknoyan, 2004). In developing countries like Nigeria, Nwankwo, Wudiri and Akinsola (2015) reported an incidence of 45.5% of impaired kidney function among hospitalised hypertensive patients in some part of north east.

In 2012, an estimated 1.5 million deaths were directly caused by diabetes and other 2.2 million deaths were attributed to high blood glucose. The global prevalence of diabetes among adults over the age of 18 years has risen from 4.7% in 1980 to 8.5% in 2014 (World Health Organization, 2016). Two-Thirds of Chronic Kidney Diseases (CKDs) are caused by diabetes and high blood pressure. CKD includes conditions that damages someone's Kidney.

If kidney disease progresses, it may eventually lead to Kidney failure, which requires dialysis or Kidney transplant.

The fuzzy expert systems can be designed to deal with the uncertainty and imprecision of real world problems. Some components of the system are human-like, adaptable and explanations. Two popular and most powerful soft computing techniques of fuzzy logic (Zadeh, 1965) and neural networks (McCulloch & Pitts, 1990), which are complementary to each other rather than competitive for system identification and has the ability to recognize patterns and adapt themselves to cope with changing environment.

Physicians make use of computerised technologies to assist in identification and give suggestion as medical diagnosis is full of uncertainty. Identification of most of the diseases is very expensive as many tests are required for predictions. Neuro fuzzy systems are multilayer connectionist networks that realize the basic elements and functions of traditional fuzzy logic decision systems (Jang, Sun, & Mizutani, 1997).

Nowdays, most of the systems may be considered to be complex in nature. They may be linear or non-linear, predictable or unpredictable. Developing an expert system is of importance in almost all fields, but especially so in Medical disease diagnosis, Transportation, Signal processing and Telecommunication, Engineering (Jang et al., 1997).

Classification techniques are being used in different field of studies to easily identify the type and group to which a particular tuple belongs. 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 target class of each data point (Han, Pei, & Kamber, 2011; Pang-Ning, Steinbach, & Kumar, 2006).

**Literature Review**

In the literatures about the use of ANFIS approach, the medical sector had a number of related works in which Sugeno fuzzy inference system was applied. (Ziasabounchi & Askerzade, 2014) developed ANFIS model based on hybrid learning algorithm, least squares estimate and gradient descent algorithm. (Ramya & Radha, 2016) developed a method of diagnosis of chronic kidney disease using machine learning algorithm. Back propagation algorithm and Radial Basis Function were used. (Jena & Kamila, 2015) described the distributed data mining classification algorithm for prediction kidney disease based on different algorithms such as Naïve Bayes, Support Vector Machine, Multilayer Perceptron. Settouti, Saidi and Chikh (2012) generates fuzzy classification model for diagnosis of diabetes disease. The combination of fuzzy c-means and neuro-fuzzy rule-based classification technique had been established. The FCM-clustering was adopted to reduce the dimension of the classifier and its training time.

The remaining parts of this paper are organized as follows: in second part materials and methods of this work are presented. This led us to third part in which experiments and results were recorded and analyzed. Fourth part concludes the work.

**Materials and Methodology**

This section deals with the materials and method used in developing the proposed expert system for the identification of chronic kidney disease.

## Descriptive Statistics of the Data Sets

**Table 1: Information about input variables for Chronic Kidney Disease Data set**

| Description of Input variable | Type of Attributes | No. of MF | Min | Max | Std | Mean |
|---|---|---|---|---|---|---|
| Age | Numerical | 3 | 6 | 90 | 15.7 | 53 |
| Blood Pressure | Numerical | 3 | 50 | 180 | 13.7 | 76.1 |
| Blood Glucose Random | Numerical | 3 | 22 | 490 | 79.4 | 145 |
| Hemoglobin | Numerical | 3 | 3.1 | 17.8 | 2.9 | 13 |
| Packed Cell Volume | Numerical | 3 | 9 | 54 | 9 | 39.2 |

Chronic Kidney Disease Data set is obtained from UCI machine learning repository (Lichman, 2013). This dataset can be used to predict the chronic kidney disease. This data set contains 25 attributes including class attribute with 400 instances. It contains missing values and the characteristics of attributes are real values. The class field refers to the presence of chronic kidney disease of the patient as 1 or absence as 0. In this research work, five input attributes are selected as described in Table 1.

### Data Pre-processing

Data pre-processing is a data mining technique that involves transforming raw data into an understandable format. Data pre-processing includes (i) feature selection of the subset of the original or existing features without transformation, and (ii) missing data usually called missing values are nearly universal in statistical practice. Data pre-processing is necessary because real world data are usually incomplete, lacking attribute values, lacking certain attributes of interest, containing only aggregate data or containing discrepancies in codes or names (Roy & Mohapatra, 2013; Scholar, 2015).

### Feature Selection

Feature selection is an important step in data pre-processing technique in classification, which is related to dimensionality reduction and can be used to identify the significant attributes (Rajeswari, Vaithiyanathan, & Pede, 2013). In the processing medical data, which is often very high dimensional, choosing the optimal subset features is very important, not only to reduce the computational cost or reduce the dimensionality of large datasets, but also to improve the usefulness and gain good classification performance of the model built from the selected data (Ghazavi & Liao, 2008). In this research, feature selection method was applied with the aid of data mining tool, known as Waikato Environment for Knowledge Analysis (WEKA) 3-7-4 software (Frank, Hall, Trigg, Holmes, & Witten, 2004; Sanders, Bridges, McCarthy, Nanduri, & Burgess, 2007; Sharma & Jain, 2013; Yadav, Malik, & Chandel, 2014) based on supervised selecting technique.

### Missing Data

Missing data usually called missing values may be due to inconsistency with other record data, data not entered due to misunderstanding or certain data may not be considered important at the time of entry. In this research, listwise deletion was adapted (Allison, 2003; Sauro, 2015; Widaman, 2006). Soley-Bori (2013) highlights the important of listwise deletion as it can be used with any kind of statistical analysis and no special computational methods are required. In other words, listwise deletion is the simplest approach to missing data.

### Membership Function

The membership function, often given the designation of *μ*, as the essence of fuzzy sets. A membership function is a curve that defines how each point in the input space is mapped to a degree of membership usually taken as a real number in the interval [0,1]. The input space is sometimes referred to as the *universe of discourse*, a fancy name for a simple concept. For example, fuzzy set *A* on the universe of discourse *X* is defined as $\mu_A : X \to [0,1]$, where each element of *X* is mapped to a value between 0 and 1. The selection of membership function type for fuzzy sets is usually determined by experts or chosen depending on its suitability such as simplicity, convenience, speed, and efficiency (Hamdan, 2013; Jang & Sun, 1995). The significant of MFs is to show that each point in the input space is mapped to a membership value (or degree of membership) between 0 and 1. In this research, Gaussian MF was used because of its lower number of parameters. The high the number of MF parameters the more complexity in the fuzzification process.

### Proposed Expert System for the Identification of Chronic Kidney Disease

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, 2017).

According to (Jang et al., 1997), if $f(x,y)$ is a first order polynomial, then the Takagi Sugeno Kang Fuzzy model is given as:

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

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

In designing this new expert system, a hybrid learning technique based on Least squares estimate and the Levenberg-Marquardt algorithm was used. The central difference scheme was applied for computation of the Jacobian Matrix.

### Forward Pass

Least squares estimate (LSE) was used at the very beginning to get the initial values of the consequent parameters $S_2 = \{p_i, q_i, r_i\}$ details can be found in (Jang et al., 1997), then at backward pass the Levenberg-Marquardt algorithm (Marquardt, 1963; Yu & Wilamowski, 2011) is to update all parameters. After the consequent parameters $S_2$ are identified, the network output can be computed and the error measure $\mathbf{E_k}$ represents an objective function for $kth$ of the training data can be obtained as:

$$\mathbf{E}_k = (\mathbf{T}_k - \mathbf{O}_k)^2 \tag{2}$$

where $\mathbf{T}_k$ and $\mathbf{O}_k$ 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, root mean square error (RMSE) (Ho, Tsai, Lin, & Chou, 2009) defined as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\mathbf{E_k}} \tag{3}$$

**Backward Pass**

In the backward pass, error signals are propagated and antecedent parameters $S_1 = \{\sigma_i, c_i\}$ are to be updated by Levenberg-Marquardt algorithm. The performance index to be optimised is defined by (Madsen, Nielsen, & Tingleff, 2004) and presented as:

$$F(\mathbf{x}) = \frac{1}{2}\mathbf{e}^T\mathbf{e} \tag{4}$$

*F(**x**)* is the total error function, $\mathbf{x} = [\mathrm{x}_1, \mathrm{x}_2, ..., \mathrm{x}_k]$ comprising of all parameter of the network, **e** is the error vector comprising the error of all the training samples.

The parameters of unique membership functions of current fuzzy inference system (FIS) is to be obtained, which is a novel approach that allows program to run faster, defined as:

$$\mathbf{v} = I(\mathbf{R}_{ij}) \tag{5}$$

where **v** is the index vector that keeps track of the unique MFs, $I$ is the index table of the unique MF used in the rules, $\mathbf{R}_{ij}$ is a matrix of size number of rule by number of input, that identifies the membership functions for the *i*th rule and *j*th input.

The Jacobian matrix is to be built column–wise, which contains first order partial derivatives of network error using central difference scheme

$$f'(x_0) = \frac{f_1 - f_{-1}}{2h} + E_{trunc}(f, h), \tag{6}$$

where $f_1 = f(x+h)$. Therefore,

$$\mathbf{J}_{\mathrm{i,j}} = \frac{\partial \mathrm{f_i}}{\partial \mathrm{x_j}} \tag{7}$$

The Hessian matrix is to be approximated, which contains second order partial derivative of network error using the cross product of Jacobian matrix, defined as:

$$\mathbf{H}_{\mathrm{i,j}} \approx \frac{\partial^2 \mathrm{f_i}}{\partial \mathrm{x_i} \partial \mathrm{x_j}} \tag{8}$$

To ensure that equation is invertible, another approximation to Hessian matrix is introduced, that is

$$\mathbf{H}^* = \mathbf{J}_k^T \mathbf{J}_k + \Psi \mathbf{I} \tag{9}$$

where $\Psi$ is called combination coefficient or learning parameter, **I** is the identity matrix.

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

$$\Delta \mathbf{X}_{\mathrm{k}} = \left(\mathbf{J}_k^T \mathbf{J}_{\mathrm{k}} + \Psi \mathbf{I}\right)^{-1} \mathbf{J}_{\mathrm{k}}^{\mathrm{T}} \mathbf{e} \tag{10}$$

The network parameter needs to be updated as:

$$\mathbf{X}_{\mathrm{k+1}} = \mathbf{X}_{\mathrm{k}} - \left(\mathbf{J}_{\mathrm{k}}^{\mathrm{T}} \mathbf{J}_{\mathrm{k}} + \Psi \mathbf{I}\right)^{-1} \mathbf{J}_{\mathrm{k}}^{\mathrm{T}} \mathbf{e}_k \tag{11}$$

where **X** is parameter vector, $\mathbf{J}_{\mathrm{k}}$ is the Jacobian matrix, $\Psi$ is called combination coefficient or learning parameter, **I** is the identity matrix and k is the index of iterations.

Recalculate the RMSE using equation (3). And adjust learning parameters, if the current RMSE is less than the previous RMSE, accept and update the consequent parameters.

Initialised k=0, $\Psi = \text{user define}$, $\xi = \text{user define}$ , $\Psi\max = \text{user define}$ (Max learning rate)

IF $\Psi = \max(\Psi * \xi)$ decrease $\Psi$, accept. Go back to (3)

ELSE $\Psi = (\Psi * \xi)$, increase $\Psi$, reject. Go back and recalculate the RMSE

Equation (11) is the update rule of Levenberg-Marquardt algorithm. Depending on the magnitude of $\Psi$, a method interpolates smoothly between the Gauss-Newton ($\Psi \to \infty$)

and gradient descent ($\Psi \to \infty$). Usually, the Gauss-Newton method is more efficient but less stable; the gradient descent method is more stable but less efficient. By properly setting the value of $\Psi$, the Modified Levenberg-Marquardt method can be efficient and well stable (Marquardt, 1963).

**Results and Discussion**

Based on the data set obtained from UCI machine learning repository, throughout the experiments, two-thirds of examples were selected as training examples and the remaining one-third as test examples. Comparison was made between proposed method and conventional method (ANFIS) based on machine learning process, accuracy, sensitivity and specificity (Colman et al., 1999; Goodall, Colman, Schneider, McLean, & Barker, 2007; Saed, 2015; Waziri, Ozovehe & Isah, 2016) between the models as presented in Table 2.

**Table 2: Comparison of results based on performance measures, RMSE for Chronic Kidney Disease Data set**

| Methodology adopted | Accuracy (%) | Sensitivity (%) | Specificity (%) | RMSE |
|---|---|---|---|---|
| Proposed method (ANFIS-LM) | 89.47 | 98.03 | 88.24 | 0.15994 |
| Conventional method (ANFIS) | 85.26 | 94.21 | 83.53 | 0.26448 |
| ANFIS, (Settouti et al. 2012) | 83.85 | 82.05 | 84.62 | ------ |
| BP, (Ramya & Radha, 2016) | 80.40 | 83.00 | 89.00 | ------ |
| RBF, (Ramya & Radha, 2016) | 85.30 | 87.00 | 92.00 | ------ |

The test performance of the classifier was determined by the computation of sensitivity, specificity, performance error and total classification accuracy, as shown in Table 2 for Chronic Kidney Data set. Based on a comparison of the results, the proposed system produced reasonable results in identifying the possible presence of chronic kidney disease in patients. This assertion is based on the following observation: in Table 2, it is clearly confirmed that none of the cited research work had success rates higher than 85.30% accuracy. The accuracy of the proposed system was obtained as 89.47 % with RMSE 0.15994. The Levenberg-Margquardt algorithm is more effective and it achieved a lower RMSE and thus has higher mapping precision. The dash line (---) indicates that there is no such type of results in the respective existing models.

**Conclusion**

The objective of this study is to develop a fuzzy system for the identification of chronic kidney disease. This study proposed one major novelty techniques by indexing unique membership functions in a row-wise vector using the vectorisation technique. The computation of Jacobian matrix was built based on standard method, central difference scheme. The applicability of the proposed technique in data classification using one benchmark data set in the area of medical diagnosis was demonstrated. Comparison was made between proposed method, conventional method (gradient descent algorithm) and some of the other related existing methods. The results of the proposed system are better than the results of existing methods and have potential in identifying chronic kidney disease.

**References**

Afolabi, M. O., Abioye-Kuteyi, E. A., Arogundade, F. A., & Bello, I. S. (2009). Prevalence of chronic kidney disease in a Nigerian family practice population. *SA Fam Pract., 51*(2), 132 - 137.

Allison, P. D. (2003). Missing data techniques for structural equation modeling. *Journal of Abnormal Psychology, 112*(4), 545 - 557.

Colman, P. G., Thomas, D. W., Zimmet, P. Z., Welborn, T. A., Garcia-Webb, P. & Moore, M. (1999). New classification and criteria for diagnosis of diabetes mellitus. Position statement from the australian diabetes society, New zealand society for the study of diabetes, royal college of pathologists of australasia and australasian association of clinical biochemists. *The Medical Journal of Australia, 170*(8), 375 - 378.

Frank, E., Hall, M., Trigg, L., Holmes, G., & Witten, I. H. (2004). Data mining in bioinformatics using Weka. *Bioinformatics, 20*(15), 2479 - 2481.

Ghazavi, S. N., & Liao, T. W. (2008). Medical data mining by fuzzy modeling with selected features. *Artificial Intelligence in Medicine, 43*(3), 195 - 206.

Goodall, I., Colman, P. G., Schneider, H. G., McLean, M., & Barker, G. (2007). Desirable performance standards for HbA1c analysis–precision, accuracy and standardisation Consensus statement of the australasian association of clinical biochemists (AACB), the australian diabetes society (ADS), the royal college of pathologists of australasia (RCPA), endocrine society of australia (ESA), and the australian diabetes educators association (ADEA). *Clinical Chemical Laboratory Medicine, 45*(8), 1083 - 1097.

Hamdan, H. (2013). *An exploration of the adaptive neuro-fuzzy inference system (ANFIS) in modelling survival.* PhD Thesis University of Nottingham, United Kingdom.

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

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.

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. (1995). Neuro-fuzzy modeling and control. *Proceedings of the IEEE, 83*(3), 378 - 406.

Jang, J. S. R., Sun, C. T., & Mizutani, E. (1997). Neuro-fuzzy and soft computing: *A computational approach to learning and machine intelligence.*

Jena, L. & Kamila, N. K. (2015). Distributed data mining classification algorithms for prediction of chronic-kidney-disease. *International Journal of Emerging Research in Management & Technology, 4(*11), 110 - 118.

Johnson, C. A., Levey, A. S., Coresh, J., Levin, A., Lau, J. & Eknoyan, G. (2004). Clinical practice guidelines for chronic kidney disease in adults. Part I: Definition, disease stages, evaluation, treatment and risk factors. *Am Fam Physician* 70, 869 – 76.

Lichman, M. (2013). *UCI machine learning repository*. Retrieved 12 November, 2015 from http://archive.ics.uci.edu/ml/datasets.html

Madsen, K., Nielsen, H. B., & Tingleff, O. (2004). *Methods for non-linear least squares problems* [Lecture notes]. Retrieved 12 January, 2016 from http://soe.rutgers.edu/~meer/GRAD561/ADD/nonlinadvanced.pdf

Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics, 11*(2), 431 - 441.

McCulloch, W. S., & Pitts, W. (1990). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biology, 52*(1-2), 99 - 115.

Nwankwo, E. A., Wudiri, W. W. & Akinsola, A. (2007). Risk factors for development of chronic kidney disease among Nigerians with essential hypertension. *Journal of Medical Sciences*, *7*(1), 579 - 584.

Pang-Ning, T., Steinbach, M., & Kumar, V. (2006). *Introduction to data mining.* Addison Wesley, 1st edition.

Rajeswari, K., Vaithiyanathan, V., & Pede, S. V. (2013). Feature selection for classification in medical data mining. *International Journal of Emerging Trends and Technology in Computer Science (IJETTCS), 2*(2), 492 - 497.

Roy, S. & Mohapatra, A. (2013). International Journal of Software and Web Sciences (IJSWS) www. iasir. net. *International Journal of Software and Web Sciences (IJSWS), 4*(1), 20 - 25.

Ramya, S. & Radha, N. (2016). Diagnosis of chronic kidney disease using machine learning algorithms. *International Journal of Innovative Research in Computer and Communication Engineering, 4*(1), 812 - 820.

Saed, S. (2015). *An introduction to data mining*. Retrieved 31 May, 2016 from http://www.saedsayad.com/model_evaluation.htm

Sagir, A. M. & 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.

Sanders, W. S., Bridges, S. M., McCarthy, F. M., Nanduri, B., & Burgess, S. C. (2007). Prediction of peptides observable by mass spectrometry applied at the experimental set level. *BMC Bioinformatics, 8*(7), 1 - 9.

Sauro, J. (2015, 5 May). *7 Ways to handle missing data*. Retrieved 5 May, 2016 from http://www.measuringu.com/blog/handle-missing-data.php

Scholar, P. (2015). Energy efficient maximization in OFDM multi-user MIMO systems. *International Journal for Trends in Engeering & Technology, 3*(1), 79 - 84.

Settouti, N., Saidi, M. & Chikh, M. A. (2012). Interpretable classifier of diabetes disease. *International Journal of Computer Theory and Engineering, 4*(3), 438 - 442.

Sharma, T. C. & Jain, M. (2013). WEKA approach for comparative study of classification algorithm. *International Journal of Advanced Research in Computer and Communication Engineering, 2*(4), 1925 - 1931.

Soley-Bori, M. (2013). *Dealing with missing data: Key assumptions and methods for applied analysis*. Retrieved 17 April, 2016 from http://www.bu.edu/sph/files/2014/05/Marina-tech-report.pdf

Wachukwu, C. M., Emem-Chioma, P. C., Wokoma, F. S. & Oko-Jaja, R. I. (2015). Prevalence of risk factors for chronic kidney disease among adults in a university community in southern Nigeria. *Pan African Medical Journal*, *21*(1).

Waziri, V. O., Ozovehe, U. A., & Isah, A. (2016). Comparison of neural network, J48 and random tree based algorithm for anormally intrusion detection. *Journal of Science, Technology, Mathematics and Education (JOSTMED),* 12(1), 201 - 211.

Widaman, K. F. (2006). III. Missing data: What to do with or without them. *Monographs of the Society for Research in Child Development, 71*(3), 42 - 64.

World Health Organization. (2016). *Fact Sheets*. Retrieved 19 August, 2016 from http://www.who.int/mediacentre/factsheets/fs312/en/

Yadav, A. K., Malik, H. & Chandel, S. (2014). Selection of most relevant input parameters using WEKA for artificial neural network based solar radiation prediction models. *Renewable and Sustainable Energy Reviews, 31*, 509 - 519.

Yu, H. & Wilamowski, B. M. (2011). Levenberg–Marquardt training. *Industrial Electronics Handbook, 5*(12), 1 - 16.

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

Ziasabounchi, N., & Askerzade, I. (2014). ANFIS based classification model for heart disease prediction. *Intenational Journal of Electrical and Computer Sciences. IJECS-Ijens, 14*(02).