Digital Innovations & Contemporary Research in SCIENCE, ENGINEERING & TECHNOLOGY
Vol. 5 No. 4. December, 2017

Article Citation Format
A.A. Ojugo & I.P. Okobah (2017): Hybrid Fuzzy-Genetic Algorithm Trained Neural Network Stochastic Model for Diabetes Diagnosis and Classification.
Journal of Digital Innovations & Contemp Res. In Sc., Eng & Tech.
Vol. 5, No. 4. Pp 69-90.

Article Progress Time Stamps

Article Type: Research Article
Manuscript Received: 25rd October, 2017
Review Type: Blind
Final Acceptance:: 15th December, 2017
DOI Prefix: 10.22624

Hybrid Fuzzy-Genetic Algorithm Trained Neural Network Stochastic Model for Diabetes Diagnosis and Classification

A.A. Ojugo
Department of Mathematics/Computer
Federal University of Petroleum Resources
Effurun, Nigeria
ojugo.arnold@fupre.edu.ng

I.P. Okobah
Department of Computer Science Education
Federal College of Education Technical
Asaba, Delta State, Nigeria
okobah.ify@gmail.com

ABSTRACT

Soft-computing yields predicted possible solution to dynamic and complex task for which conventional method do not have cost-effective, complete results. It analyses observed data by investigating feats of interest via its underlying probabilities. It simulates tractability, robustness, effective and low cost solution with high tolerance to ambiguity, uncertainties, partial truth and noise when applied to its input via stochastic models (heuristics) from an evolutionary point. Successfully applied to various disciplines, our study adopts hybrid fuzzy-genetic algorithm trained neural network model to aid diabetes diagnosis (as decision support model) for proper treatment and diabetes classification. Adopted data is split into: training (50%), cross validation (25%) and testing (25%) to aid model validation with appropriate weights and biases set for each variable. Results indicate that age, obesity and family relations (first and second degree), environmental conditions are critical factors to be watched for type-I and type-II; While in gestational diabetes, mothers with or without a previous case of GDM is confirmed if there is: (a) history of babies with weight > 4.5kg at birth, (b) resistant to insulin showing polycystic ovary syndrome, and (c) have abnormal tolerance to insulin.

Keywords—Diabetes, Gestational, fuzzy classifiers, linguistics variable, membership function, mutation,

1. INTRODUCTION

Diabetes mellitus (sugar killer) is a metabolic disorder characterized by presence of hyperglycemia from defective insulin secretion, action or both. Diabetes Mellitus (silent killer or sugar disease) is a metabolic disease characterized by high glucose levels, either in a body with insufficient insulin to breakdown glucose, or body that is resistant to effects of insulin. To improve early diagnosis, data mining tools are used to help physicians effectively classify the disease. Its diagnostic criteria based on thresholds of glycemia associated with micro-vascular disease especially with retinopathy. Individuals with such defects are grouped into: **chronic** hyperglycemia (relatively specific long-term micro-vascular complications affecting the eyes, kidneys and nerves with increased risk for cardiovascular defect) and **prediabetes** (a practical, convenient term for impaired fasting glucose, impaired glucose tolerance or glycated hemoglobin of 6.0% to 6.4%) – both of which places patient at a high risk of developing diabetes (Goldenberg and Punthakee, 2013).

Ekoe, Punthakee, Ransom, Prebtani and Goldenberg (2013) and Goldenberg and Punthakee (2013) classified diabetes into: (a) **Type 1** (is prone to ketoacidosis, results from pancreatic beta cell destruction as measured from etiology and cases due to autoimmune process), (b) **Type 2** ranges from predominant insulin resistance with relative insulin deficiency to a more serious case of the predominant, secretory defect with insulin resistance, and (c) Gestational diabetes refers to glucose intolerance with onset or first recognition during pregnancy.

Other types include variety of relative uncommon conditions as well as specific genetically defined types of diabetes, some of which are associated with other diseases or drug use. Its diagnostic criteria based on glucose threshold, is as measured from its etiologic classification so that differentiating type 1 from type 2 is critical due to management, which is a difficult task at diagnosis in some cases. Studies reveal that physical sign such as insulin resistance and use of autoimmune markers antibodies (like anti-glutamic acid decarboxylase or anti-islet cell antibody) are helpful in early classification – though, have not been adequately studied as diagnostic tests. Very low C-peptide levels measured after months of clinical stabilization favours **type 1** – though, not helpful in acute hyperglycemia. Clinical judgment with safe management and ongoing follow-up is a prudent method (Ransom, Goldenberg, Michalachki, Prebtani and Punthakee, 2013).

The complex nature of diabetes and its complications along with its varying types makes precision diagnosis as critical so as to help in effective administering of the proper drugs to aid faster treatment. This difficulty has made manual diagnostic means somewhat redundant, inconclusive and time-wasting since early detection is a critical task. Modeling and prediction is also, quite a complex, difficult task due to its chaotic nature (Sivakumar, 2001) – since researches continues to advance precision diagnosis for early detection and accurate prediction and prevention of **type 2**, though still a challenging operation (Sarke and Wolfe, 1985). Predictions are only an improvised means through which a model allows propagation of a set of observed dataset as the user seeks feats of interest. The dataset often contains ambiguities, noise and assumptions as inputted, so that the model yields an output of outcomes simulated via optimization methods and taxonomy (Ojugo et al, 2013a).

Researchers continue to advance evolutionary, stochastic models via machine learning optimization methods, which has been successfully applied to enhance accurate identification via simulated prediction (that aims at an optimal solution in a task, chosen from a set of possible solution space) – to yield output guaranteed of high quality and void of ambiguities. Models are tuned to be robust so they can perform quantitative processing to ensure qualitative knowledge and experience, as its new language (Gaume and Gosset, 2003).

2. SOFT COMPUTING (SC) PARADIGM / FRAMEWORK

Evolutionary models utilize SC as synergy between AI and other discipline via machine learning methods, dedicated to solve problems. It exploits historic data and explores human knowledge via mathematic models and symbolic reasoning – to yield an output system that is tolerant to imprecision, noise, uncertainty and partial truth as applied to its input (Coello et al, 2004; Abbot et al, 1986). Thus, it evolves into meta-rules for constraint satisfaction problem that use intelligent agents in vector space as it seeks optimal fitness. They are inspired by evolution, behavioural patterns in biological populations and natural laws; And they mimics agents seeking food, which have proven efficient in complex optimization (French et al, 1992). These include Genetic Algorithm, Artificial Neural Network, Fuzzy set, Annealing etc (Ojugo et al, 2012).

Thus, evolutionary algorithms attempt to explore dynamic processes via the exploitation of observed datasets exhibits 3-feats (robustness, continuous adaptation and flexibility) – while displaying the underlying probabilities of data feats of interest. Thus, it considers output feats – with uncontrollable constraints modeled within the models input that may not be explicitly present in the search space but confined to real parameters as well as limited by boundary values (Campolo et al, 1999; Ojugo et al, 2013b).

The proposed model framework hinges on 3-basic methods: (a) fuzzy logic, (b) genetic algorithm and (c) artificial neural network. These are explained as thus:

2.1 Artificial Neural Network (ANN)

ANN data processing model is inspired by neurons in the human brain. Thus, consists of interconnected neurons (nodes) with capability to learn by example that makes them universal estimators (Abraham, 2005). As it processes data, its nodes shares data signals as well as adjust the network weights and bias adjustments representing the synapse axons and dendrites that indicates the connection strength between these synapses respectively (Caudill, 1987 and Fausett, 1994). The signals are converted so that learning occurs to adjust weight, summed by an **adder**. Depending on task, its activation function limits its output (Mandic and Chambers, 2001) to modulate associated inputs and nonlinear feats exhibited via transfer or activation function as in Eq. 1 below:

$$\emptyset = f(net) = f\sum_{i=1}^{m} Xi * Wij \qquad (1)$$

ANN are trials in an attempt to translate into mathematical models, principles of biological processing so as to generate in the fastest time, implicit and predictive evolution outcomes of a task. ANN derives its possible outcomes from experience and is able to recognize feats and behaviours of interest from historic dataset – to suggest optimal solutions of high quality and void of **over-fitting**, irrespective of modification via other approximations with multiple agents. These also, constantly affects the quality of any solution (Dawson and Wilby 2001a). Its configuration depends on the area to be applied, captured data feats and system requirement. Its connections are set as either **explicit** (apriori knowledge) and/or implicit (post-priori knowledge) to allow learning so that the net is trained to learn patterns that change its weight and bias based on a rule (Beven and Binley, 1992; Bishop, 1995). Learning is grouped into:

a. **Supervised** in which an input vector has a set of desired responses, one for each node as relayed to the output. A forward pass is done to measure errors between **desired** and **actual** response for each node in the output, which is used to determine weight changes in the net based on the learning algorithm (Gupta et al, 1998; Hall, 2001). Thus, desired signal on output is provided by external teacher via back-propagation, delta rule and perceptron rule

b. **Unsupervised** or self-organized allows its output trained to respond to a clusters of patterns at its input so that the model discovers statistically, salient feats within the input dataset – such that the model has no prior knowledge how patterns are grouped. Rather, the model develops its own representation of the input dataset (Hsu et al, 1995; Heppner and Grenander, 1990).
c. **Reinforcement** in which network learns what to do, map states to actions to help maximize a numerical reward data. Network must discover the actions that yield most reward by trying them. Sometime, such actions affect not only the immediate data, but also the rest states (Jang, 1993; Kennedy et al, 2001).

ANN are adapted in most tasks as either: (a) trial/error search and (b) delayed reward. ANN are encoded into 3-layers: input, hidden and output. ANN configurations are: (a) **feedforward** net where data flows directly from input to output, and extends over multiple layers, and (b) **recurrent** net has a feedback with dynamic feats to evolve the net as it undergoes relaxation to a stable state where its activation values and output changes no more. In some tasks, output change is significant and dynamic behavior constitutes its output (Ojugo et al, 2013b).

Nature of diabetes diagnosis requires previous knowledge. Thus, we adopt recurrent (Jordan) net to incorporate previous dataset feats of interest as input variables into model. Thus, allows the previous dataset and previous output to be feedback as input into the model's hidden units (Rajurkar et al, 2004; Karunanthis et al, 1994) to yield next output. Its correlated weights are interconnected in that with W_{ij} as weight between input and hidden layers, $W_{o}j$ is bias and x_i is diabetes input data sent – its yields an output via tangent/sigmoid transfer function, which sums its weighted input as in Eq. 2 and Eq. 3 (Minns, 1998; Chakraborty, 2010).

The network also resolves the structural dependencies imposed on the model by dataset and hybrid methods used via its ability to store earlier data as generated from previous layer(s) (Kuan, 1994). Feed-forward nets must be expanded and extended to represent complex dynamic patterns and cases such as this, since it treats all data as new – so that previous data signals do not help to identify data feats, even if such observed datasets exhibits temporal dependence; And, causes practical difficulty as large nets are not easily implemented. **Jordan** net overcomes such difficulty via its internal feedbacks – making it appropriately suitable for dynamic, non-linear and complex tasks such as these. Its output unit is fed-back as input to hidden unit with a time delay, so that its outputs at time t−1, is also input at time t.

$$Z_{ij} = w_{o}j + \sum_{i=1}^{m} x_i * w_{ij} \quad (2)$$

$$F(Z_{ij}) = \frac{2}{1 + e^{-2*Zij}} - 1 \quad (3)$$

Jordan's net is more plausible and computationally more powerful than other adaptive models - via the backpropagation in time learning (advanced training algorithm) so that its output at time t is used along with a new input to compute the network's output at time t+1 in response to dynamism (Mandic and Chambers, 2001). Thus, output is computed via Tansig activation function given as y^k, sums input, receives target value of input training pattern, computes error data, weight correction updates in layers (c_i^k) and bias weights correction updates (c_o^k). This error is sent from output layer back to input nodes via error backpropagation, to correct weights. Backpropagation is used – so as to find weights that approximate target output with selected accuracy. Weights are modified by minimizing error between target and computed outputs at the end of each forward pass. If the error is higher than selected value, process continues with a reverse pass; else, training stops. Weights are updated via mean square error until a minimal error is achieved (Ursem et al, 2002).

Our Jordan recurrent net is constructed by modifying the multilayered feedforward with addition a **context** layer to help retain data between observations. At each move, new inputs are fed to the net. Previous contents of hidden layer is passed into context layer and later fed back into the hidden layer in the next time step (Regianni and Rientjes, 2005). The context layer contains nothing initially. Output from the hidden layer after the first input will be same as if there is no context layer (Perez and Marwala, 2011). Weights are calculated same way for the new connections from and to the context layer from the hidden layer. Training aim at best fit data weights computed via Tansig function that assumes approximation influence of data points at the center – so that function decreases with distance from its center.

Euclidean length (r_j) is yields distance between datum vector $y = (y_1,...,y_m)$ and center $(w_{1j},...,w_{mj})$ as:

$$r_j = \left\| y - Y^j \right\| = \left\{ \sum_{i=1}^{m} (y_i - w_{ij})^2 \right\}^{1/2} \quad (4)$$

The suitable transfer function is applied to r_j:

$$\emptyset(r_j) = \emptyset \left\| y - Y^j \right\| \quad (5)$$

Finally, output k receives weighted combination as:

$$y^k = w_o + \sum_{j=1}^{n} \left(c_j^k * \emptyset(r_j) \right) = w_o + \sum_{j=1}^{n} \left(c_j^k * \emptyset \| y - Y^j \| \right) \quad (6)$$

2.2 Fuzzy Logic (Rule-Based Variables)

A fuzzy system chooses between different control actions and transforms such into a fuzzy set value. It is divided into:

a. Fuzzy classifier model assigns a class label to an object based on an object's description, so that it can also predict each class label. Object descriptions are vector values with feats/attributes relevant for such classification task (Ludmila, 2008). Classifier learns to predict a class labels via training algorithm and its accompanying dataset. If a training data set is not available, classifier is designed to learn apriori (prior knowledge) so that trained, it classifies objects. Thus, the rule-based classifier focuses on if-then rules with actions and possible outcomes, constructed as a user specifies its class rules and linguistic variables (fuzzy set) that helps tune a fuzzy set in line with such class rules (Ludmila, 2008). For example:
If Math Error is medium and is small, Then Class 1
If Math Error is medium and is large, Then Class 2 etc

b. Fuzzy Cluster algorithm groups the data points (linguistic variables) into homogeneous classes known as clusters so that items in the same class are as similar as possible and vice-versa (Yang and Wang 2001). Clustering is a data compression technique, where a large number of samples are converted into small number of representative clusters (Giles and Draeseke, 2001). Depending on data and task, different types of similarity measures are used to identify classes, to control how clusters are formed. Examples of values that can be used as similarity measures include distance, connectivity, and intensity (Ojugo et al, 2012).

In some cases called non-fuzzy (hard clusters), data is divided into crisp clusters such that each data point belongs to exactly one cluster; While, in fuzzy clusters, it is quite possible for the data points to belong to more than one cluster, and associated with each of the points are membership grades which indicate the degree to which the data points belong to these different clusters (Nascimento, 1991). Thus, we employ a Fuzzy Cluster Means (FCM) algorithm which attempts to partition a finite collection of elements into a collection of fuzzy clusters with respect to some given criterion.

FCM linguistic description can be implemented by fuzzy logic (Berks et al 2000) as thus:

a. Select the number of clusters c ($2 \leq c \leq n$), exponential weight μ ($1 < \mu < \infty$), initial partition matrix U^0, and the termination criterion ε. Also, set the iteration index 1 to 0.
b. Calculate fuzzy cluster centers $\{V_i^l \mid i=1, 2,..., c\}$ using U^l.
c. Calculate new partition matrix U^{l+1} via $\{V_i^l \mid i=1, 2,..., c\}$.
d. Compute new partition matrix = $\| U^{l+1} - U^l \| = | U_{ij}^{l+1} - U_{ij}^l |$. If > ε, set l = l + 1. Go to step (b); Else stop if $\leq \varepsilon$.

Initial cluster centers are computed via (a) arithmetic means of all the data points, or (b) running FCM several times each starting with different initial cluster centers. For this study, we adopt the first method. To implement the fuzzy logic system, we perform the following:

a. Define control goals/criteria: What do I wish to control? How do I achieve the control? What response is needed? What are the possible (probable) system failure modes?
b. Define input/output relationships and choose minimum number of input variables to the system as well as define the error rate and rate-of-change-of-error.
c. Using the rule-based structure, define the control problem into a series of IF X AND Y THEN Z rules that define the desired model output response for given input condition or case. The number and complexity of rules depends on the number of input parameters that are to be processed and the number variables associated with each parameter. If possible, use atleast a variable and its time derivative (though, possible to use an instantaneous error parameter without knowing its rate of change, which can cripples the model's ability to minimize overshoot for a step inputs).
d. Create Fuzzy Logic membership functions that define the meaning (values) of Input/output terms used in the rules.
e. Create necessary pre- and post-process Fuzzy set routines if implementing as software; Else, hardwire the rules into the Fuzzy Logic hardware engine.
f. Test the system, evaluate the results, tune the rules and membership functions, and retest until satisfactory result is obtained.

Linguistic variables are non-precise variables used to convey a surprising amount of data about our environment or an object under observation (Inan and Elif, 2005). In common usage, linguistic variables often overlap. Linguistic variables require a formal way of describing a linguistic variable in crisp terms that the computer can deal with. For example, to indicate the relationship between measured distance and linguistic term far. Each individual may have slightly differing ideas about the exact distance measurement that far actually represents – though, the said distance be consistent (Berks et al 2000).

At some point, some individuals agree that it is not far and at some point, they also all do agree that it is far. The space between far and not far indicates a distance measure that is to some degree, a bit of both. Thus, the horizontal axis of such a graph shows the measured or crisp value of distance; while, its vertical axis describes degree to which the linguistic variable fits with the crisp measured data (Inan and Elif, 2005).

2.3 Genetic Algorithm (GA)

GA is inspired by Darwinian genetic evolution (survival of fittest) consists of population (data) chosen for selection with potential solutions to a specific task. Each potential solution is an individual for which optimal is found using four operators: initialize, select, crossover and mutation (Coello et al, 2004 and Reynolds, 1994). Individuals with genes close to optimal, is said to be fit. Fitness function determines how close an individual is to optimal solution. Ojugo et al (2013a, 2013b) notes the operators as:

a. Initialize – Individual data are encoded into forms suitable for selection. Each encodings type used has its merit. Binary encodings are computationally more expensive. Decimal encoding has greater diversity in chromosome and greater variance of pools generated; float-point encoding or its combination is more efficient than binary. Thus, it encode as fixed length vectors for one or more pools of different types. The fitness function evaluates how close a solution is to its optimal – after which they are chosen for reproduction. If solution is found, function is good; else, is bad and not selected for crossover. The fitness function is the only part with knowledge of task. If more solutions are found, the higher its fitness value.
b. Selection – best fit individuals close to optimal are chosen to mate. The larger the number of selected, the better the chances of yielding fitter individuals. This continues until one is chosen, from the last two/three remaining solutions, to become selected parents to new offspring. Selection ensures the fittest individuals are chosen for mating but also allows for less fit individuals from the pool and the fittest to be selected. A selection that only mates the fittest is elitist and often leads to converging at local optima.
c. Crossover ensures best fit individual genes are exchanged to yield a new, fitter pool. There are two crossover types (depends on encoding type used): (a) **simple** crossover for binary encoded pool. It allows single- or multi-point cross with all genes from a parent, and (b) **arithmetic** crossover allows new pool to be created by adding an individual's percentage to another.
d. Mutation alters chromosomes by changing its genes or its sequence, to ensure new pool converges to global minima (instead of local optima). Algorithm stops if optimal is found, or after number of runs if new pools are created (though computationally expensive), or when no better solution is found. Genes may change based on probability of mutation rate. Mutation improves the much needed diversity in reproduction and its algorithm is as thus:

Cultural GA is a variants of GA with a belief space define as thus: (a) Normative (has specific value ranges to which an individual is bound), (b) Domain (has data about task domain), (c) Temporal (has data about events' space is available), and (d) Spatial (has topographical data). In addition, an influence function mediates between belief space and the pool – to ensure and alter individuals in the pool to conform to belief space. CGA is chosen to yield a pool that does not violate its belief space and helps reduce number of possible individuals GA generates till an optimum is found (Reynolds, 2004).

3. MATERIALS AND METHODS

Retrieved dataset is presented in table 1 below, obtained via research survey, utilizing questionnaires as the research tool. The quantitative and qualitative questionnaires comprises of two phases: (a) demographic data, and (b) tele-medical data. A total of a hundred questionnaires were distributed to various medical (diabetic experts) professional spread across fifteen teaching hospitals in five Geo-political regions in Nigeria. The sixth geo-political zone was not covered due to the insurgency issue. All questionnaires administrated were retrieved without mutilation. In other to generate a fuzzy Linguistic variable Universe of discourse, all questionnaires responds were tuned utilizing the proposed equation:

$$PFCMUDE = \sum (A, B, C, D, E) * X \qquad (7)$$

A,B,C,D,E = picked questions option; X(0.02) = Assigned questions option fuzzy range value, X(0.00) = unpicked option

Table 1: Fuzzy Encoded Universe of Discourse for each class of diabetes dataset Values

Code	Fuzzy Set (Parameters)	Membership Function Degree for Diabetes		
		Type 1	Type 2	Gest.
P01	Frequent Urination	0.50	0.00	0.00
P02	Unusual Thirst	0.50	0.00	0.50
P03	Extreme Hunger	0.50	0.00	0.00
P04	Unusual Weight Loss	0.50	0.50	0.00
P05	Extreme Fatigue	0.50	0.00	0.50
P06	Serious Irritation	0.00	0.00	0.50
P07	Frequent Infection	0.00	0.00	0.50
P08	Blurred Vision	0.00	0.50	0.00
P09	Slow Healing of bruises/cuts	0.00	0.50	0.00
P10	Tingle/numbness in hands/feet	0.00	0.50	0.00
P11	Regular skin/bladder infection	0.00	0.50	0.50
P12	Nausea/vomiting	0.00	0.00	0.50
P13	Haemoglobin test > 10	0.20	0.20	0.20
P14	Leg cramp	0.20	0.20	0.20

3.1 Model Design Problems

Issues to be resolved in the model design include:

a. Many studies aim at single heuristic to globally classify various diabetes types. Also, false-positives and true-negatives resulting from drug use, related diseases and in some cases (not even diabetes) – shows symptoms quite similar or mimics any of the diabetes types or class.
b. Such models employ hill climbing methods that often gets their solution trapped at local minima because their speed shrinks as such models often approaches its optima.
c. Resolving conflict issues in structured learning and from statistical dependencies imposed by data and the use of multiple methods adopted/adapted, is quite a tedious.

3.2 Model Design Goals and Objectives

The proposed system aims to solve the existing problem of diabetes diagnosis utilizing the following properties:

a. Perform repetitive tasks without emotional defects
b. Embody the knowledge of human experts with the help of special software tools, manipulate data to solve problems and make decisions in that domain.
c. Processes are better formalized and defined on machines.
d. Updating the knowledgebase is automatic, with or without the help of a human
e. Processes are better formalized and defined on machines.

4. EXPERIMENTAL MODEL FRAMEWORK

From Fig 1, the proposed model design employs these:

a. Fuzzy system consists of a classifier which propagates if-then rule values of selected data, enhanced as predefined linguistic variables classification into the diabetes classes, and the fuzzy cluster means universe discourse equation which enhances the linguistic variables partitioning it into data-point that cumulate into the universal of discourse generated via a survey exercise utilizing questionnaires.
b. Jordan neural network provides a machine, self-learning ability as manipulated and optimized through a cultural genetic algorithm optimizer that recombines and mutates the rule-based fuzzy dataset to train and test the system to autonomously classify diabetes into its varying class type.
c. Genetic algorithm helps train ANN so as to optimize our collated-answers within the tuned fuzzy dataset linguistic variable (symptoms) universe of discourse values in other to yield a centralized, fuzzy-scaled function boundary in determining high/low degree membership function.

The proposed model consists of four parts:

a. Knowledgebase – consists of historic structured data feats and the database of diabetes symptoms, the fuzzy if-then rules, its global universe discourse linguistic variables and optimized membership functions. Its houses the optimized universe discourse values as represented by fuzzy-if-then, linguistic variables (rule-based) as selected data feats.
b. Inference engine – consists of hybrid fuzzy logic (rule-based) genetic algorithm trained neural network model. Inference engine infers conclusion derived from genetic algorithm trained neural network from the selected data feats encoded as fuzzy-if-then conditions with possible outcomes and consequent action upon criteria being met.
c. Decision support – consists of the predicted output and the output database that is updated automatically in time as patients are diagnoses as long as it encounters and read sin new data. The decision support predicts system output based on the cognitive and the emotional filers as display by the output device. This is seen in fig 1.

Model starts off with Proposed Fuzzy Classifier and Cluster Mean Universe Discourse Equation (PFCMUDE) in which the answers from the collected questionnaire are expressed thus:

4.1 Genetic Algorithm Trained Neural Network (GANN)

GANN is initialized with the fuzzy (if-then rules) linguistic variables. Individual fitness is computed as 30-individual are selected as new pool via tournament method. It determines mating individuals and solutions. Crossover and mutation is applied to help network learn dynamic and non-linear feats in the dataset and feats of interest using a multi-point crossover. With mutation, data between 1 and 30 is randomly generated using Gaussian distribution corresponding to crossover points (all genes are from single parent). As new parents contribute the rest to yield new individuals whose genetic makeup is combination of both parents, mutation is also applied to yield 3-random genes that also undergoes another mutation as they are allocated new random values that still conforms to belief space.

Number of mutation applied depends on how far CGA is progressed on the network (how fit is the fittest individual in the pool), which equals fitness of the fittest individual divided by 2. New individuals replace old with low fitness so as to create a new pool. Process continues until individual with a fitness value of 0 is found – indicating that the solution has been reached (Branke, 2001). Initialization/selection via ANN ensures that first 3-beliefs are met; mutation ensures fourth belief is met. Its influence function influences how many mutations take place, and the knowledge of solution (how close its solution is) has direct impact on how algorithm is processed.

Algorithm stops when best individual has fitness of 0 (Campolo et al, 1999; Dawson and Wilby 2001b). GANN model is as thus:

INPUT:

1. Pool size (k), crossover (c), mutation (v), influence functn (Ifnc) and n;
   // Initialization and Selection
2. Randomly generate K possible solution
3. Save solution in population Kok;
   // Loop till terminal point
4. For m = 1 to n do;
   // Crossover
5. Number of crossover nc = (k – Ifnc)/2;
6. For u = 1 to n do;
7. Select two solutions randomly E_A and F_G for K;
8. Generate G_V and H_N by 2-point crossover to E_A and F_G;
9. Save G_V and H_N to K2;
10. End For;
    //Mutation
11. For u = 1 to n do;
12. Selection a solution Y_h from K2;
13. Mutate each bit of Y_h under Ifnc
14. Generate a new solution Y_h^i
15. If Y_h^i is impossible
16. Recompute Y_h^i with possible solution by modifying Y_h^i
17. End if
18. Recompute Y_h with Y_h^i in K2
19. End for
    //Recompute
20. Recompute K = K2;
21. Return **Best** solution in Y

Model stops if stop criterion is met. GANN utilizes number of epochs to determine stop criterion. Initial selection is given as:

1. **R1:** If R01 Then C1 = 0.50
2. **R2:** If R01 AND R02 Then C1 = 0.50
3. **R3:** If R01, R02 AND R03 Then C1 = 0.50
4. **R4:** If R01, R02, R03 AND R04 Then C2 = 0.13
5. **R5:** If R01, R02, R03, R04 AND R05 Then C3 = 0.17
6. **R6:** If R01, R02, R03, R04, R05 AND R06 Then C3 = 0.03
7. **R7:** If R01, R02, R03, R04, R05, R06 AND RO7 Then C3=0.03
8. **R8:** If R01, R02, R03, R04, R05, R06, RO7 AND R08 Then C3 = 0.17
9. **R9:** If R01, R02, R03, R04, R05, R06, RO7, R08 AND R09 Then C3 = 0.17
10. **R10:** If R01, R02, R03, R04, R05, R06, RO7, R08, R09 AND R10 Then C3 = 0.17
11. **R11:** If R01, R02, R03, R04, R05, R06, RO7, R08, R09, R10 AND R11 Then C3 = 0.17
12. **R12:** If R01, R02, R03, R04, R05, R06, RO7, R08, R09, R10, R11 AND R12 Then C3 = 0.17
13. **R13:** If R01, R02, R03, R04, R05, R06, RO7, R08, R09, R10, R11, R12 AND R13 Then C3 = 0.20
14. **R14:** If R01, R02, R03, R04, R05, R06, RO7, R08, R09, R10, R11, R12, R13 AND R14 Then C3 = 0.20

Fitness function (f) is resolved with initial pool (Parents) as:

R1:50	R2:50	R3:50	R4:13	R5:17	R6:3
R7:3	R8:17	R9:17	R10:17	R11:17	R12:17
R13:20	R14:20				

Table 2: 1st and 2nd Generation of population from Parents

S/N	Selection	Chromosomes (Binary 0 or 1)			Fitness Function
		Parent 1st Gen	Crossover	Parent 2nd Gen	
1	50	110010	1 and 9	110**00**	48
2	50	110010	2 and 14	110**11**	51
3	50	110010	3 and 13	110**00**	48
4	13	001101	4 and 12	001**101**	13
5	17	010001	5 and 6	0100**11**	19
6	3	000011	5 and 6	0000**01**	1
7	3	000011	7 and 8	0000**01**	1
8	17	010001	7 and 8	0100**11**	19
9	17	010001	mutation	0100**10**	18
10	17	010001	mutation	0100**10**	18
11	17	010001	mutation	0100**10**	18
12	17	010001	4 and 12	0100**01**	17
13	20	010100	3 and 13	0101**10**	22
14	20	010100	1 and 14	0101**10**	22

Table 3: 2nd and 3rd Generation of population from Parents

SN	Selection	Chromosomes (Binary 0 or 1)			Fitness Functn
		Parent 2nd Gen	Crossover	Parent 3rd Gen	
1	48	110000	1 and 2	110011	51
2	51	110011	1 and 2	110000	48
3	48	110000	3 and 13	110010	50
4	13	001101	4 and 12	001101	13
5	19	010011	5 and 9	010000	16
6	1	000001	6 and 8	000011	3
7	1	000001	mutation	000010	2
8	19	010011	6 and 8	010001	17
9	16	010000	5 and 9	010011	19
10	16	010000	10 and 15	010010	18
11	16	010000	mutation	010010	18
12	17	010001	4 and 12	010001	17
13	22	010110	3 and 13	010100	20
14	22	010110	10 and 14	010100	20

The Proposed Fuzzy Classifier Diabetes training Algorithm:

INPUT:

1. Diabetes Types (Type 1, Type 2, Gestational, MODY, LADA)
   No. of Symptoms (P1, P2,..., Pn)
   P = Fuzzy parameters (Symptoms Codes)
   N = 15
   More than five symptoms = Serious
   Exactly four symptoms = Moderate
   Three symptoms and below = Minor
2. Glucose Level (125md/dl) = High
3. Age Range (R)
   1 – 21yrs of age = teenager
   30 – 40yrs of age = pre-Adult
   > 41yrs of age = post-Adult
   > 50 = pre-menopause
4. Origin (descent)
   Caucasians: Americans, Europeans, Asians, North-Africa
   Blacks: African, African-Americans, Blacks Indians etc
   Plus; either Caucasians or blacks
5. // INITIALIZATION
6. Randomly pick a patient K;
7. Save identification (diagnosis) Result in Knot;
8. // Loop till terminal point
9. For P = 1 to n do;
10. // Type 1 diabetes
11. Diagnose for Type 1 Diabetes;
12. If Type I symptoms is serious, age is teenager, glucose level is high, patient origin is Plus and pancreas destruction is swift THEN Type 1;
13. Else: May be Type 1;
14. Else: May Not Type 1;
15. // Type 2 diabetes
16. Diagnose for Type 2 Diabetes;
17. If TYPE 2 symptoms is serious, patient age is post-Adult, glucose level is high and patient origin is black THEN Type 2;
18. Else: Might be Type 2;
19. Else: May Not Type 2;
20. //Gestational diabetes
21. Diagnose for Gestational Diabetes;
22. If Gestational symptoms is serious, patient age is pre-menopause, glucose level is high, patient origin is plus and patient is pregnant THEN Gestational diabetes;
23. Else: May be Gestational Diabetes;
24. Else: May Not Gestational Diabetes;
25. // MODY diabetes
26. Diagnose for MODY Diabetes;
27. If MODY symptoms is serious, patient age is teenager, glucose level is high, patient origin is plus and parents (siblings) diagnosed with MODY THEN MODY diabetes;
28. Else: May be MODY diabetes;
29. Else: May Not MODY diabetes;
30. // LADA diabetes
31. Diagnose for LADA Diabetes;
32. If LADA symptoms is serious; patient age is pre-Adult, glucose level is high, patient origin is Caucasians and pancreas destruction is progressive THEN LADA diabetes;
33. Else: May be LADA diabetes;
34. Else: May Not LADA diabetes;
35. //Save results in Knot;
36. Return diabetes result for patient K

Table 4: 3rd and 4th Generation of population from Parents

	Selection	Chromosomes (Binary 0 or 1)			Fitness Function
		Parent 3rd Gen	Crossover	Parent 4th Gen	
1	51	110011	1 and 2	110**00**	48
2	48	110000	1 and 2	110**11**	51
3	50	110010	3 and 13	110**00**	48
4	13	001101	4 and 12	001**101**	13
5	16	010000	5 and 9	010**11**	19
6	3	000011	6 and 8	0000**01**	1
7	2	000010	Mutation	0000**01**	1
8	17	010001	6 and 8	010**11**	19
9	19	010011	5 and 9	010**00**	16
10	18	010010	10 and 14	010**00**	16
11	18	010010	Mutation	0100**01**	17
12	17	010001	4 and 12	0100**01**	17
13	20	010100	3 and 13	0101**10**	22
14	20	010100	10 and 14	0101**10**	22

Table 2, 3 and 4 respectively shows generation of optimized fuzzy set (linguistic variable/symptoms) – from single parents by choosing first and second bits from the left as our crossover and mutation points. Fourth generation is our stop criterion with best fitness function of 51 (in row 2) that implies – the clusters of the various universe of discourse variable values as searched has been optimized to 0.51 (with its base value as 0.50). Thus, the parameter combination yields a membership function **< 0.50 = Low Degree Membership Function**; while those of **≥ 0.50 = High Degree Membership Function**. Based on optimization carried-out, utilizing 0.50 as our fuzzy scaled central value alongside the predefined fuzzy rules, it results in table 5 as generated below:

Table 5: Optimized Dataset for Diabetes (Scale 0.00- 1.00)

Code	Fuzzy Set (Parameters)	Membership Function Degree Diabetes				
		T-1	T-2	Gest.	MODY	LADA
P01	Frequent Urine	0.50	0.00	0.00	0.00	0.50
P02	Unusual Thirst	0.50	0.00	0.00	0.00	0.50
P03	Extreme Hunger	0.50	0.00	0.00	0.00	0.50
P04	Weight Loss	0.50	0.00	0.00	0.00	0.50
P05	Extreme Fatigue	0.50	0.00	0.00	0.00	0.50
P06	Irritability	0.00	0.00	0.50	0.50	0.00
P07	Freq. Infection	0.00	0.00	0.50	0.50	0.00
P08	Blurred Vision	0.00	0.50	0.00	0.50	0.00
P09	Slow Healing	0.00	0.50	0.50	0.00	0.00
P10	Tingle/numbnes	0.00	0.50	0.00	0.50	0.00
P11	Regular skin and bladder infect	0.00	0.50	0.50	0.00	0.00
P12	Nausea/vomit	0.00	0.50	0.50	0.00	0.00
P13	Haemoglobin test > 10	0.20	0.20	0.20	0.20	0.20
P14	Leg cramp	0.20	0.20	0.20	0.20	0.20
	Result	T-I diag	T-II diag	Gest. Diag	MODY diag	LADA diag

Table 6: Fuzzy Classifier Diabetes Training Algorithm Class

Code	Fuzzy Set (Parameters)	Membership Function Degree Diabetes				
		T-1	T-2	Gest.	MODY	LADA
P01	Frequent Urination	S	M	M	M	S
P02	Unusual Thirst	S	M	M	M	S
P03	Extreme Hunger	S	M	M	M	S
P04	Weight Loss	S	M	M	M	S
P05	Extreme Fatigue	S	M	M	M	S
P06	Irritability	M	M	S	S	M
P07	Freq. Infection	M	M	S	S	M
P08	Blurred Vision	M	S	M	S	M
P09	Slow Healing of bruises/cuts	M	S	S	M	M
P10	Tingle/numbness in hands/feet	M	S	M	S	M
P11	Regular skin and bladder infect	M	S	S	M	M
P12	Nausea/vomit	M	S	S	M	M
P13	Haemoglobin test > 10	M	M	M	M	M
P14	Leg cramp	M	M	M	S	M
	Result	T-I diag	T-II diag	Gest. diag	MODY diag	LADA diag

As below, S = serious and M = minor case of diabetes, T-1 = Type 1 diabetes, T-II = type 2 diabetes, Gest. = Gestational.

5. RESULT FINDINGS AND DISCUSSION

Table 6 assigns linguistic labels to class based on the fuzzy classifier diabetes training algorithm; While, table 5 highlights the 5-types of diabetes as thus:

a. **Type I –** is any patient in age range 5-15yrs that exhibits symptoms (P01-P05), glucose ≥ 125mg (7.0mmol/L) with swift and rapid destruction of the pancreas.
b. **Type II –** patients above 40years with symptoms (P08 - P12) and glucose level of more than 125mg (7.0mmol/L).
c. **Gestational –** patients with symptoms (P06, P07, P09, P11, P12), glucose level ≥ 125mg (7.0mmol/L) during or after gestation (incubation, maturation or pregnancy).
d. **MODY –** patients in age range (1 - 21yrs) with symptoms (P06-P08, P10-P14) from birth and who siblings/parents have MODY.
e. **LADA –** patient above 35yrs with symptoms (P01 - P05), a glucose level of more than 125mg (7.0mmol/L) with a slow, gradual and progressive destruction of the pancreas.

Each symptom falls into P_1 - P_{14} with five clusters: Type-I, Type-II, Gestational, MODY and LADA diabetes – makes up the degree of membership/intensity. Example, P13 in cluster 1, 2, 3, 4 and 5 respectively has 0.20 value, which implies that the degree of symptoms of P13 matches **20% of Type I, 20% of Type II diabetes, 20% of Gestational diabetes, 20% in MODY and 20% in LADA.** The fuzzy partition for each input feat consists of clinical symptoms of the varied forms of diabetes (frequent urination, unusual thirst, extreme hunger, unusual weight loss, extreme fatigue, irritability, frequent infection, blurred vision, slow to heal cuts/bruises, tingling/numbness in hands/feet, regular skin/bladder/gum infection, nausea/vomiting, hemoglobin A1c test (HbA1c) >10 and leg cramp).

However, it can occur that a fuzzy partition for a diabetes type is not correctly set up, or the number of linguistic terms for the input feats is not large enough – to result in some patterns being misclassified.

a. Not exhibiting any class diabetes (C_1)
b. Might be exhibiting a class of diabetes (C_2)
c. Exhibiting a class of diabetes (C_3)

If a patient exhibits at least three or less of symptoms of a class of diabetes THEN (C_1), if patient exhibits exactly four of the symptoms of a class of diabetes THEN (C_2) and if the patients is exhibits five or more of the symptoms of a class of diabetes THEN (C_3). The IF-THEN Rules generated from fuzzy partitions of classification of varied for diabetes is thus:

1. **R1**: IF patient exhibit frequent urination and its serious THEN class C_1.
2. **R2**: IF patient exhibit frequent urination and unusual thirst and both symptoms are serious THEN class C_1.
3. **R3**: IF patient exhibit frequent urination, unusual thirst and severe hunger and symptoms serious THEN class C_1.
4. **R4**: IF patient exhibit frequent urination, unusual thirst, extreme hunger and unusual weight loss and symptoms are serious THEN C_2.
5. **R5**: IF patient exhibit frequent urination, unusual thirst, extreme hunger, unusual weight loss and extreme fatigue and the symptoms are serious THEN class C_3.
6. **R6**: IF patient exhibit frequent urination, unusual thirst, extreme hunger, unusual weight loss, extreme fatigue and irritability and the symptoms are serious THEN class C_3.
7. **R7**: IF patient exhibits frequent urination, unusual thirst, extreme hunger, unusual weight loss, extreme fatigue, irritability and frequent infection and these symptoms are serious THEN class C_3.
8. **R8**: IF patient exhibits frequent urination, unusual thirst, extreme hunger, unusual weight loss, extreme fatigue, irritability, frequent infection and blurred vision and these symptoms are serious THEN class C_3.
9. **R9**: IF patient exhibit frequent urination, unusual thirst, extreme hunger, unusual weight loss, extreme fatigue, irritability, frequent infection, blurred vision and slow to heal cuts and these symptoms is serious THEN class C_3.
10. **R10**: IF patient exhibit frequent urination, unusual thirst, extreme hunger, unusual weight loss, extreme fatigue, irritability, frequent infection, blurred vision, slow to heal cuts/bruises and tingling/numbness in hands/feet and these symptoms are serious THEN class C_3.
11. **R11**: IF patient exhibit frequent urination, unusual thirst, extreme hunger, unusual weight loss, extreme fatigue, irritability, frequent infection, blurred vision, slow to heal cuts/bruises, tingling/numbness in hands/feet and regular skin/bladder/gum infection and these symptoms are serious THEN class C_3.
12. **R12**: IF patient exhibit frequent urination, unusual thirst, extreme hunger, unusual weight loss, extreme fatigue, irritability, frequent infection, blurred vision, slow to heal cuts/bruises, tingling/numbness in hands/feet, regular skin/bladder/gum infection and nausea/vomiting, and these symptoms are serious THEN class C_3.
13. **R13:** IF patient exhibit frequent urination, unusual thirst, extreme hunger, unusual weight loss, extreme fatigue, irritability, frequent infection, blurred vision, slow to heal cuts/bruises, tingling/numbness in hands/feet, regular skin/bladder/gum infection, nausea/vomiting, and hemoglobin >10 and symptom is serious THEN class C_3.
14. **R14:** IF patient exhibits frequent urination, unusual thirst, extreme hunger, unusual weight loss, fatigue, irritability, frequent infection, blurred vision, slowly heal cut/bruise, tingling/numbness in hands/feet, regular skin/bladder/gum infection, nausea/vomiting, hemoglobin >10 and Leg Cramp and symptoms are serious THEN class C3.

6. CONCLUSION

Our hybrid fuzzy logic (rule-base) genetic algorithm trained neural network model for proper identification and accurate diagnosis of diabetes – adapts the genetic algorithm to help speed up the final stages of ANN so as to yield robust optima in the shortest amount of time for such dynamic and complex task. The fuzzy system will help better represent variables and data values in the model. Hybrids (though quite difficult to implement, exploit and explore) yields better solutions though appropriate parameter selection must be encoded through the model's structured learning (which in turn will address issues of statistical dependencies between the various methods used, resolve the conflicts within data feats and variables of interest and highlight the implications of such multi-agent populated model as agents create their own behavioural rules on the data set used) – to display underlying probabilities of data feats of interest. GA helps to yield better generation via its process of recombination and mutation as applied.

REFERENCES

[1] Abraham, A., (2005). Handbook of Measuring System Design, John Wiley and Sons Ltd, ISBN: 0-470-02143-8, pp. 901 – 918.
[2] American Diabetes Association. Standards of Medical Care in Diabetes – 2009. Diabetes Care, 32: S13-61.
[3] Barakat, N.H., Bradley, A.P., Barakat, M.N.H., (2010). Intelligible support vector machines for diagnosis of diabetes mellitus, IEEE Transactions on Information Technology in Biomedicine, 14(4), pp 1114-1120
[4] Berks, G., Keyserlingk, D., Jantzeen J., Dotoli, M and Axer H. (2000). Fuzzy clustering: versatile means to explore medical database, ESIT, Aachen, Germany
[5] Beven, K. J and Binley, N.J., (1992). Knowledge driven models and computations, Wiley and Sons, Chichester, ISBN: 0-465-08245-3.
[6] Bishop, C. M., (1995). Neural networks for pattern recognition, Oxford University Press, ISBN: 0-342-45230, pp.253-294.
[7] Branke, J., (2001a). Evolutionary Optimization in Dynamic Environments, Kluwer.
[8] Branke, J., (2001b). Reducing the sampling variance when searching for robust solution, Proc. of Man and Cybernetics, 14(6), pp 234-242.
[9] Canadian Diabetes Association. Standards of Medical Care in Diabetes 2014, Journal of Diabetes Care, 32, S13 – 16.
[10] Caudill M., (1987). Neural Networks Primer, AI Expert December, pp.46-52.
[11] Chinenye, S and Young, E., (2011). State of diabetes care in Nigeria: a review, The Nigerian Health Journal, 11(4), pp101-106
[12] Coello, C., Pulido, G and Lechuga, M., (2004). Handling multiple objectives with particle swarm optimization, In Proceedings of Evolutionary Computing, Vol. 8, pp 256–279.
[13] Conway, A. J., Macpherson, K and Brown, J., (1998). Delayed time series predictions with ANN, J. of Neurocomputing, 18, pp.81–89.
[14] Dawson, C and Wilby, R., (2001a). Comparison of neural networks in river flow forecasting, J. of Hydrology and Earth Science, SRef-ID: 1607-7938/hess/2001-3-529.

[15] Dawson, C and Wilby, R., (2001b). Hydrology modeling using neural network, J. Programming in Physics and Geography, 25, pg 80–108
[16] Edo, A.E., Edo, G.O., Ohehen, O.A., Ekhator, N.P and Ordiah, W.C., (2015). Age and diagnosis of type-2 diabetes seen in Benin City Nigeria, African Journal of Diabetes Medicine, 23(1), pp18-19.
[17] Fausett L., (1994). Fundamentals of Neural Networks, New Jersey: Prentice Hall, pp.240.
[18] Giles, D and Draeseke, R., (2001). Economic Model Based on Pattern recognition via fuzzy C-Means clustering algorithm, Economics Dept Working Paper, EWP0101, University of Victoria, Canada.
[19] Goldenberg, R and Punthakee,, Z (2013). Definition, classification and diagnosis, prediabetes and metabolic syndrome, 37(1), S8-S11.
[20] Guo, W.W and Xue, H., (2011) An incorporative statistic and neural approach for crop yield modelling and forecasting, Neural Computing and Applications, 21(1), pp109-117
[21] Hall, M. J., (2001). How well does your model fit data?, J. Hydroinformatics, 3(1), pp.49–55.
[22] Haykin, S., (1999). Neural networks, comprehensive foundation, Prentice Hall: New Jersey, pp. 23 – 89.
[23] Heppner, H and Grenander, U., (1990). Stochastic non-linear model for coordinated bird flocks, In Krasner, S (Ed.), The ubiquity of chaos (pp.233–238). Washington: AAAS.
[24] Hjemfelt, A and Wang, M., (1993). Artificial neural networks as unit hydrograph applications, Proc. of Engineering Hydrology, pp.754–759
[25] Inan, G and Elif, D., (2005). Adaptive neuro-fuzzy inference system for classification of EEG using wavelet coefficient, retrieved from http://www.rorylewis.com/PDF/04
[26] Jain, A. and Srinivasulu, S., (2004). Development of effective, efficient rainfall-runoff models using integration of deterministic, real-coded GA and ANN, Water Resources, 40(4), pp. 23 – 45.
[27] Jang, J.S., (1993). Adaptive fuzzy inference systems, IEEE Transactions on Systems, Man and Cybernetics, vol.23, pp.665–685.
[28] Khashei, M., Eftekhari, S and Parvizian, J (2012). Diagnosing diabetes type-II using a soft intelligent binary classifier model, Review of Bioinformatics and Biometrics, 1(1), pp 9 – 23.
[29] Kuan, C and White, H., (1994). Artificial neural network: econometric perspective", Econometric Reviews, Vol.13, Pp.1-91 and Pp.139-143.
[30] Ludmila I. K. (2008), Fuzzy classifiers Scholarpedia, 3(1), pp2925
[31] Mandic, D and Chambers, J., (2001). Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability, Wiley & Sons: New York, pp56-90.
[32] Minns, A., (1998). Artificial neural networks as sub-symbolic process descriptors, published PhD Thesis, Balkema, Rotterdam, Netherlands
[33] Nascimento L., (1991). Comparative Genomus of two Leptospira interrogans Serovirs, http://jb.asm.org/cgi/content/full/186/7/2164
[34] Ojugo, A., (2012). Artificial neural networks gravitational search algorithm for rainfall runoff modeling in hydrology, unpublished PhD Thesis, CS Department: Ebonyi State University Abakiliki, Nigeria.
[35] Ojugo, A., Eboka, A., Okonta, E., Yoro, R and Aghware, F., (2012). GA rule-based intrusion detection system, Journal of Computing and Information Systems, 3(8), pp 1182 - 1194.
[36] Ojugo, A.A., and Yoro, R., (2013a). Computational intelligence in stochastic solution for Toroidal Queen, Progress in Intelligence Computing and Applications, 2(1), pp 46 – 56.
[37] Ojugo, A.A., Emudianughe, J., Yoro, R.E., Okonta, E.O and Eboka, A., (2013b). Hybrid neural network gravitational search algorithm for rainfall runoff modeling, Progress in Intelligence Computing and Application, 2(1), doi: 10.4156/pica.vol2.issue1.2, pp 22 – 33.

[38] Ojugo, A.A., Ben-Iwhiwhu, E., Kekeje, D.O., Yerokun, M.O and Iyawa, I.J.B., (2014). Malware propagation on time varying network, Int. J. Modern Edu. Comp. Sci., 8, pp25 – 33.

[39] Ojugo, A.A., Eboka, A.O., Yoro, R.E., Yerokun, M.O and Efozia, F.N.,(2015). Hybrid model for early diabetes diagnosis, Mathematics and Computers in Science and Industry (A Mathematics and Computers in Science and Engineering Series), 50: 176-182, ISBN: 978-1-61804-327-6, ISSN: 2227-4588

[40] Perez, M and Marwala, T., (2011). Stochastic optimization approaches for solving Sudoku, IEEE Transaction on Evol. Comp., pp.256–279.

[41] Reynolds, R., (1994). Introduction to cultural algorithms, Transaction on Evolutionary Programming (IEEE), pp.131-139.

[42] Ursem, R., Krink, T., Jensen, M.and Michalewicz, Z., (2002). Analysis and modeling of controls in dynamic systems. IEEE Transaction on Memetic Systems and Evolutionary Computing, 6(4), pp.378-389

[43] Vaarala, O., Knip, M., Paronen, J., Hamalainen, A.M., Muona, P., Vaatainen, M., Ilonen, J., Simell, O and Akerblom, H.K., (1999). Cow's milk formula feeding induces primary immunization to insulin in infants at genetic risk for type-1 diabetes, Diabetes, 8(7), pp1389-1394.

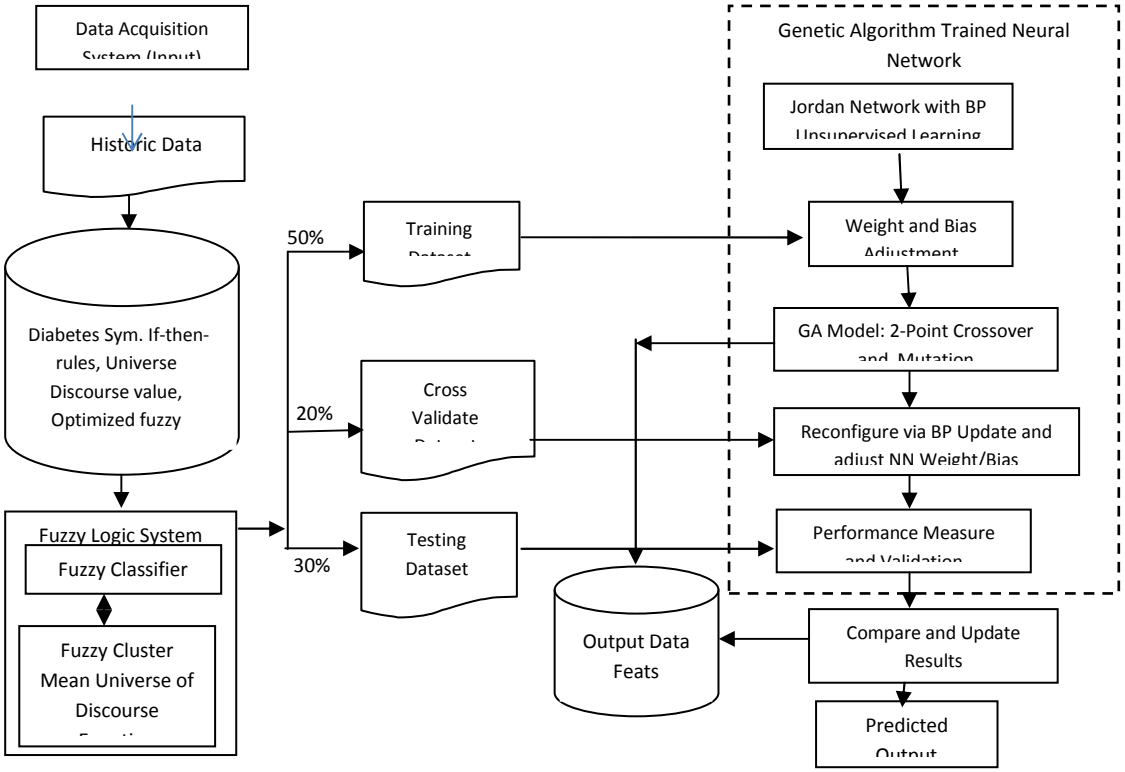


Figure 1: Dataflow Diagram of the Hybrid Model Fuzzy Genetic Algorithm Trained Neural

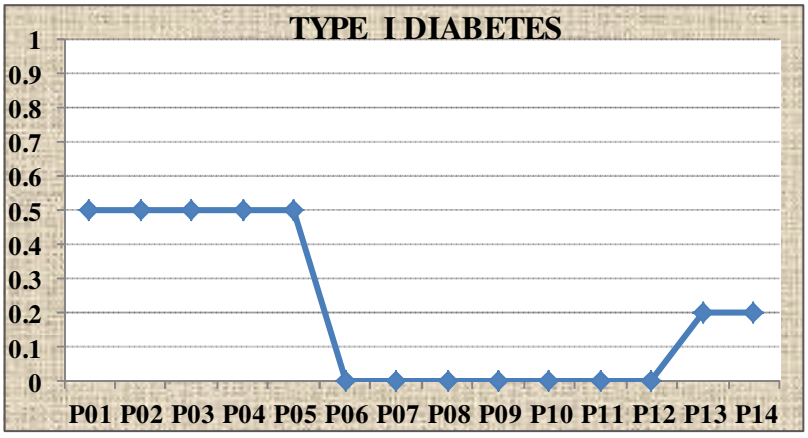


Figure 2: Cluster I representation clearly show five symptoms with high degree membership function of Type-1 Diabetes.

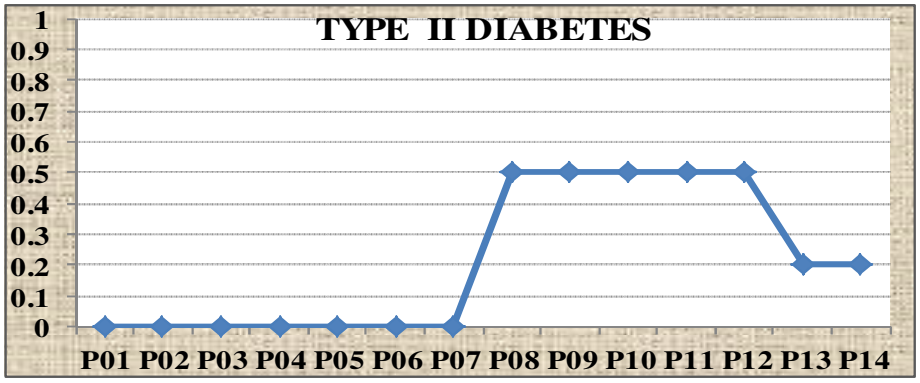


Figure 3: Cluster II representation clearly show five symptoms with high degree membership function of Type-II Diabetes.

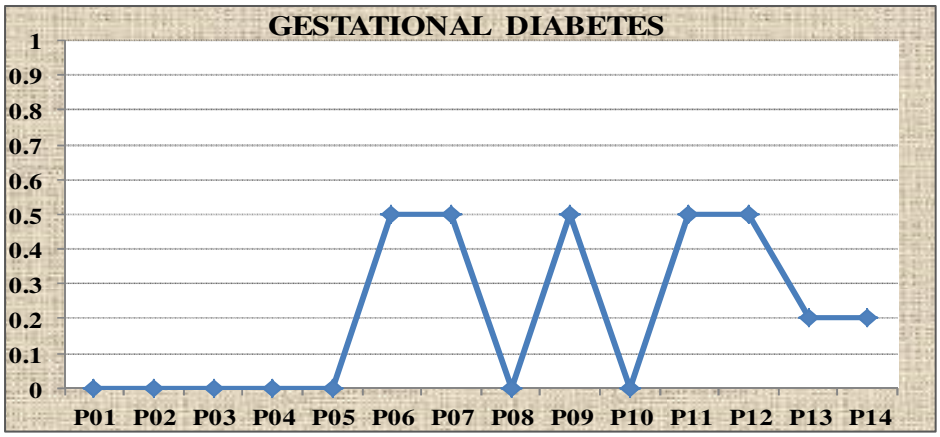


Figure 4: Cluster III representation clearly show five symptoms with high degree membership function of Gestational Diabetes.

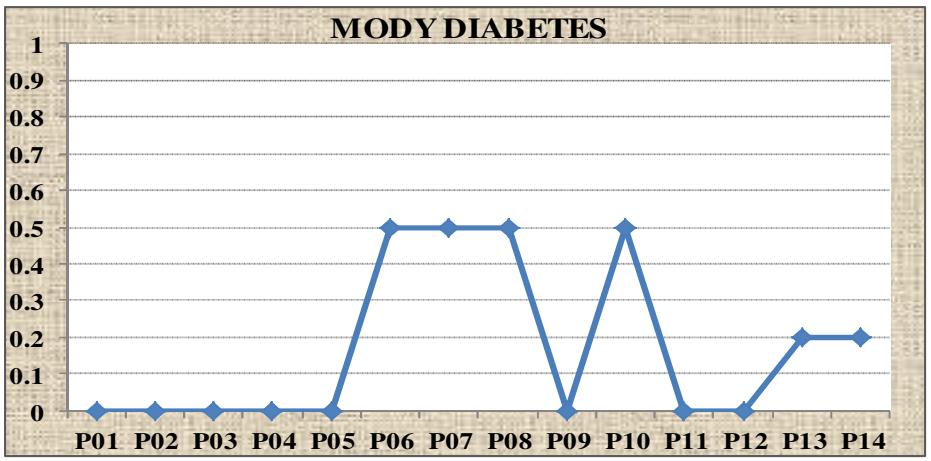


Figure 5: Cluster IV representation clearly show five symptoms with high degree membership function of MODY Diabetes.

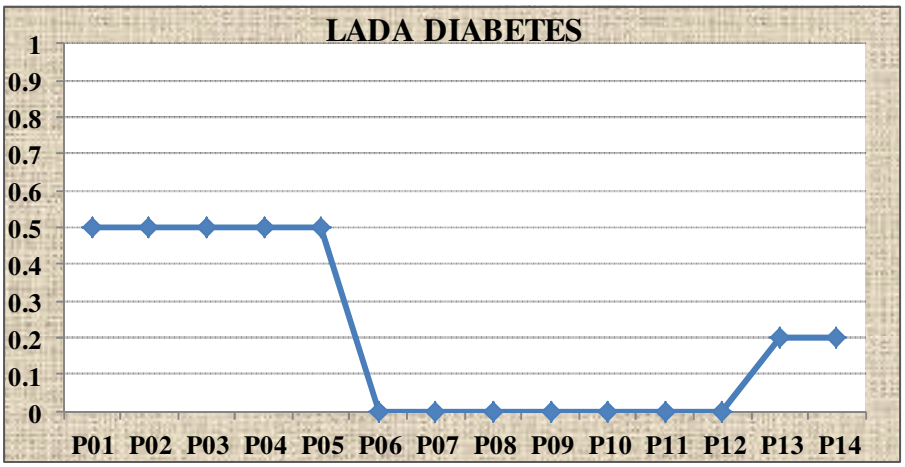
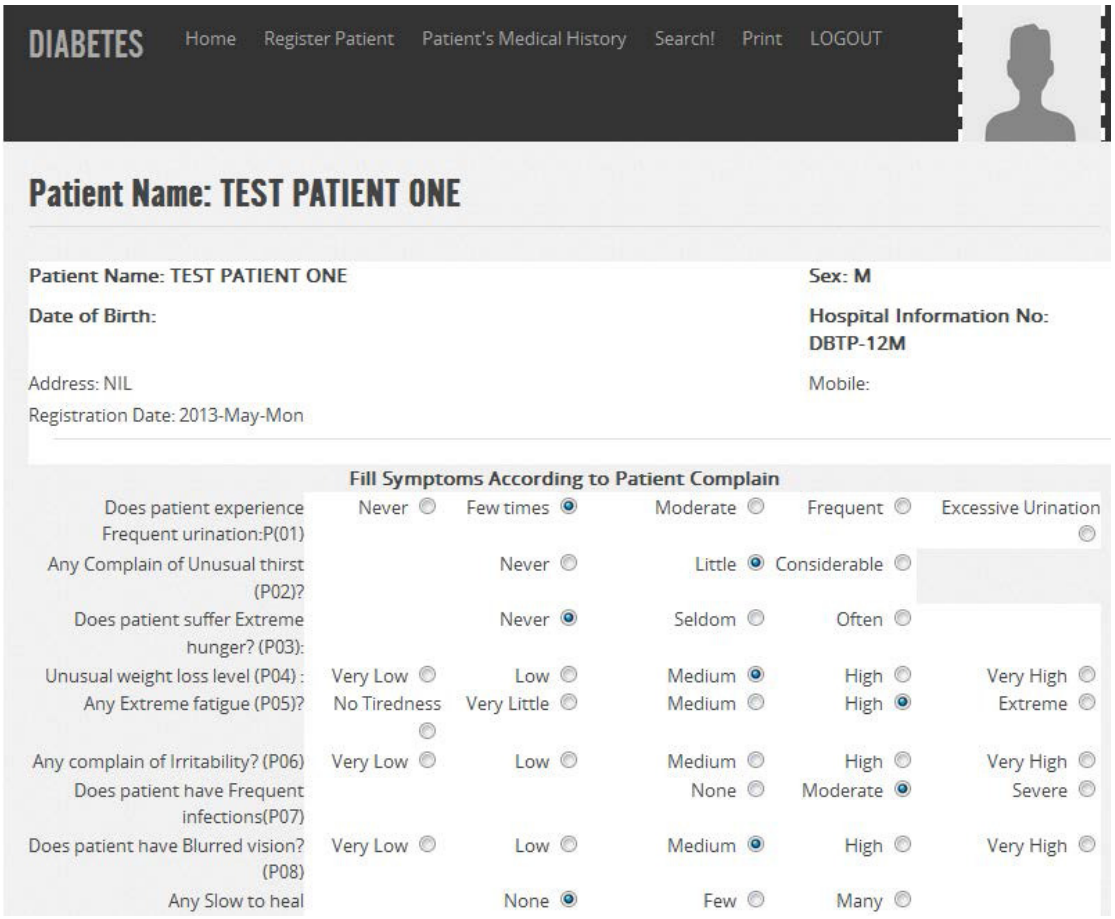


Figure 6: Cluster III representation clearly show five symptoms with high degree membership function of LADA Diabetes.

SAMPLE OUTPUT OF DECISION PAGE



SAMPLE OUTPUT OF REPORT PAGE