

Predictive Intelligent Decision Support Model in Forecasting of the Diabetes Pandemic Using a Reinforcement Deep Learning Approach

Arnold Adimabua Ojugo

Department of Computer Science, Federal University of Petroleum Resources Effurun, Delta State, Nigeria
Email: ojugo.arnold@fupre.edu.ng, arnoldojugo@gmail.com, maryarnoldojugo@gmail.com

Elohor Ekurume

Department of Computer Science, Delta State University, Abraka, Delta State, Nigeria
Email: elohorogaga@gmail.com, elohorogaga@delsu.edu.ng,

Received: 11 December 2020; Revised: 04 January 2021; Accepted: 26 January 2021; Published: 08 April 2021

Abstract: Diabetes has since become global pandemic – which must be diagnosed early enough if the patients are to survive a while longer. Traditional means of detection has its limitations and defects. The adoption of data mining tools and adaptation of machine intelligence is to yield an approach of predictive diagnosis that offers solution to task, which traditional means do not proffer low-cost-effective results. The significance thus, is to investigate data feats rippled with ambiguities and noise as well as simulate model tractability in order to yield a low-cost and robust solution. Thus, we explore a deep learning ensemble for detection of diabetes as a decision support. Model achieved a 95-percent accuracy, with a sensitivity of 0.98. It also agrees with other studies that age, obesity, environ-conditions and family relation to the first/second degrees are critical factors to be watched for type-I and type-II management. While, mothers with/without previous case of gestational diabetes is confirmed if there is: (a) history of babies with weight above 4.5kg at birth, (b) resistant to insulin showing polycystic ovary syndrome, and (c) have abnormal tolerance to insulin.

Index Terms: Diabetes, Type-I, Type-II, Gestational, deep neural network, modular learning, Silent killer

1. Introduction

Diabetes is a group of metabolic diseases characterized by high levels sugar in blood when the pancreas does not make enough insulin (hormone produced by the pancreas to control blood sugar) – to help its cells respond to insulin normally. The chronic hyperglycemia of diabetes is associated with long-term damage, dysfunction, and failure of different organs, especially the eyes, kidneys, nerves, heart, and blood vessels. Symptoms of marked hyperglycemia include polyuria, polydipsia, weight loss, sometimes with polyphagia, and blurred vision [1-3].

Diabetes set today as a global epidemic and now affects some 300-million people worldwide. It is also estimated that most people who have diabetes know their condition. There are 4 million deaths per year related to diabetes and its complications, representing 9% of world mortality [4]. Diabetes is associated with major socio-economic impact for both the individual and society. Their high costs are mainly related to a high frequency of acute and chronic complications, such as higher incidence of cardiovascular and cerebrovascular diseases, blindness, kidney failure and non-traumatic amputations of lower limbs, which are causes of hospitalization, greater needs for medical care, disability, lost productivity and premature death of life [5-7].

Diabetes prevalence continues to grow and people with this chronic disorder requires continued medical care, support to prevent acute complications, reduce the risks associated with long-term complications and ongoing patient self-management education,. On average, their medical expense approximates to twice higher than regular patient, not including indirect costs due to factors such as absenteeism, reduced productivity, and lost productive capacity due to early mortality [8-11]. Here in this study, we are propose a deep learning model to effectively classify with consistent outcome, early diagnosis of diabetes, combining stochastic rules of the various dataset feats carefully selected via a multi-criteria decision agent-based model to yield a solution that effectively classifies dataset (as well as dataset not present from the outset) into the various diabetes classes (i.e. type-I, type-II and gestational) with marginal error rates.

A. Literature Review

Diabetes mellitus (also known as the *silent sugar killer*) is a metabolic disorder characterized by presence of hyperglycemia from defective insulin secretion, action or both. This metabolic disease features high glucose level in a body with insufficient insulin to breakdown glucose, or in a body that is resistant to the effects of insulin [2]. To improve its early diagnosis, various studies have used data mining tools to help experts effectively classify the disorder with criteria based on glycemia thresholds associated with micro-vascular disease and especially with the retinopathy. Individuals with diabetes are grouped into: **chronic** hyperglycemia (a relatively specific long-term micro-vascular disorder 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 these classes, places patient at a high risk of developing diabetes [1-2, 12].

Also, [1, 3-7] classified diabetes into: (a) **Type-I** (as prone to ketoacidosis and results from pancreatic beta cell destruction as measured from etiology and cases due to autoimmune process), (b) **Type-II** ranges from predominant insulin resistance of a body 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 of pregnancy. Other types [13] are a variety of relative uncommon conditions and specific-gene-type, some of which are associated with other diseases or drug use. Its diagnostic criteria are based on glucose threshold, is as measured from its etiologic classification so that differentiating **type-1** from **type-2** is critical due to management. This is quite a difficult task in some cases [14]. Studies reveal that signs such as insulin resistance and the use of autoimmune markers antibodies (anti-glutamic acid decarboxylase, or anti-islet cell antibody) can help in its early diagnosis [9-11, 15] – though, these have not been adequately studied as diagnostic tests. Low C-peptide levels measured after months of clinical stabilization favors **type-I** (but not for acute hyperglycemia). Ultimately, clinical judgment and follow-up has always been a prudent method to treat and manage the disorder [13-15].

The complex nature of diabetes, its complications along with its varied types – often makes early diagnosis critical so as to help with drug-use administering and aid faster treatment. This has made manual diagnostic somewhat redundant, difficult, often inconclusive and time-wasting [refer to studies in 3-15]. Early prediction of diabetes thus, is a complex task due to the chaotic nature of its classification [16-17]. Studies continue to advance early and accurate detection of diabetes – even though, it is a challenging task [18]. Predictions are only an improvised means via 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 [19-20]. We thus explore a deep learning reinforcement model to enhance accurate classification aimed at optimal solution of the 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 as its new language, a qualitative knowledge and experience [21-24].

B. Study Motivation and Objective

Our study is motivated as thus [6-8]:

1. Many studies had previously aimed at production of a single heuristic to globally classify various diabetes-types. Our proposed model employs rule-based linguistic universe discourse to generate over 82-rules that are trained via the memetic deep learning neural network as in Section II.
2. Neural network models have often been known to employ hill climbing methods that often gets their solution trapped at local minima because their speed shrinks as such models often approaches its optima. However, this is curbed with the use of deep learning methods as in Section II with results as in Section III. Deep learning allows faster convergence of dataset classification – even though model as a decision support is a hybrid of genetic algorithm. This, allows better rule optimization, greater flexibility and adaptation of the rules as well as improved navigation of the model.
3. It has been noted that many stochastic models are riddled with (false-positive and true-negative) errors in diagnostic results. These may have been generated due to drug use and in some instance, symptoms of related disorders that mimics symptoms of the various diabetes-classes. To reduce such errors in classification (as in Section II/III), we employ a modular design that improves network intelligence, learning time of the network and its computational efficiency.
4. With hybrids (for this reinforced model), there issue of resolving conflicts in structured learning and from statistical dependencies imposed by data encoding as data signals are transmitted within the model is achieved via modularization that allows for greater efficiency as in Section II – though quite a tedious feat.

Thus, the goal of the study is to model selected data features of interest using a deep modular network design that simulates the generation of rules that will effectively classify data as well as reduce errors in result classifications.

2. Materials and Methods

A. Data Gathering / Sampling

We extend the study [2] via frameworks modelled in [25-27], we use dataset as presented in table 1 – obtained via a survey of questionnaires, consisting of two phases: (a) demographic data, and (b) tele-medical data. A total of a hundred questionnaires were distributed to various medical (diabetic) professional across forty teaching hospitals in six (6) Geo-political zones in Nigeria. Even with the insurgency, the sixth geo-political zone (North-North) gave a fair representation of the dataset retrieved. To generate the rule-based, selected dataset feats were tuned utilizing the proposed equation:

$$PFCMUDE = \sum(A, B, C, D, E) * X \quad (1)$$

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

Table 1. Rule-based Dataset Encoded For each class of Diabetes Dataset Values

Code	Fuzzy Set Linguistic Variable	Membership Function Degree		
		Type-1	Type-2	Gestational
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/numb hands/feet	0.00	0.50	0.00
P11	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

(Source: Ojugo et al. 2015)

B. Hybrid Reinforcement Learning Ensemble

The Modular Neural Network (MNN) as detailed in [25-27] is an improved deep learning neural network with learning that features an independent series of intermediary components and module operating under a certain architecture. It advances a model that receives individual network module output as input that helps compute final output, resolved via tangent activation function. The large network is divided into potentially, smaller and more manageable network [25] with enhanced efficiency via connected units that exponentially increases, as independent networks are added. This complicates the network structure; But, also improves its computational efficiency, reduces time convergence on individual task assigned to segmented modules, and allows tasks to be executed in parallel with module that are re-organized to improve flexibility and adaptability [26].

The network enhances intelligence and increases time efficiency by reducing the network's learning time – achieved via an independent training algorithms applied at each module with training dataset implemented independently and more quickly. This makes the model more flexible, adaptable and robust as rules can be re-used independently at various networks. Re-usability of rules has been a tedious experienced with such large and complex networks. But, with appropriate data encoding and carefully selected feats – network experiences improved performance, improves flexibility of compartmentalization via removal of partitions of the interfaces and eliminates redundancy [27].

The MNN architecture is a larger network, comprising of smaller network. Its modularization allows for easy learning and understandability of the underlying feats of interest, grants model greater flexibility via task execution parallelism via compartmentalization, flexibility, eases code reuse and adaptability [28]. MNN passes data input via task decomposition and training modules via a multi-objective, multi-agent and multi-region support module that aids effective classification. MNN can be implemented using the multilayered perceptron, adaptive resonance theory and self-organizing maps. The network is trained via either the supervised, unsupervised or reinforcement learning [29].

For the study, our hybrid framework is divided into three (3) components: (a) a supervised cultural genetic algorithm, (b) an unsupervised Kohonen network, and (c) a knowledgebase – as in figure 1, as explained in the section below for the supervised genetic algorithm and the Kohonen self-organizing map neural network.

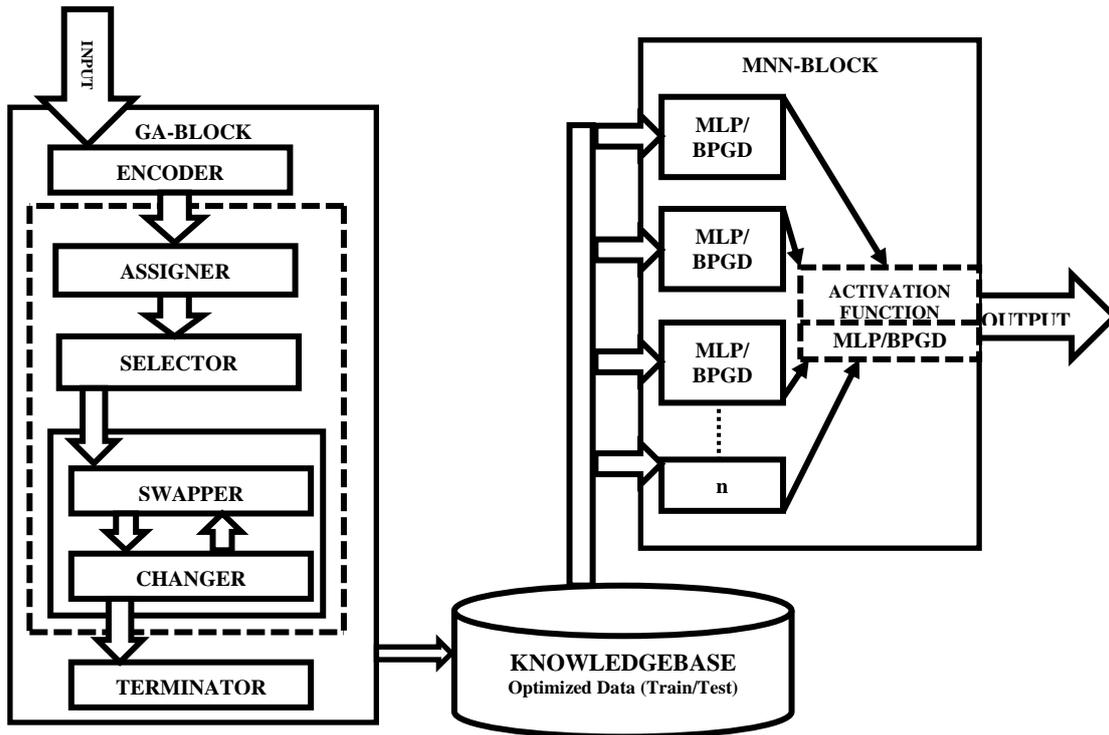


Fig. 1. Schematics Diagram of Genetic Algorithm Trained Modular Neural Network

- ✓ Supervised Cultural Genetic Algorithm (CGA): GA model is inspired by Darwinian evolution of survival of fittest, it consists of a chosen population with potential solutions to specific task. Each potential optimal solution is found via four operators [36] and individuals with genes close to its optimal, are said to be fit. Fitness function determines how close an individual is to optimal solution. Basic operators of GA includes initialize, fitness function and select, crossover and mutation [28-30]. Cultural GA as a variant, has some belief spaces defined thus: (a) *normative* belief (has specific value ranges to which an individual is bound), (b) *domain* belief (has data about task domain), (c) *temporal* belief (has data about events' space is available), and (d) *spatial* belief (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 [31-33].
- ✓ Kohonen Self-Organizing Map Neural Network is a feed-forward 2-layer network – arranged like in a grid. The first layer receives the initial input and transmits it unbound to second layer, which then provides competitive computation via the activation of its transfer function. Also, similarities among patterns are mapped into relations on the competitive layer. After training, the pattern relations are observed from this layer which are used as the result determination [26-27].

However, to achieve deep learning – we adapt the selected parameters and carefully constructing our Kohonen multi-layer network using a deep architecture at its input, hidden and output layers. We employ a hidden layer that transforms non-linearly from a previous layer to the next. We adopt the deep neural network approach in [34], which is trained via two phases: (a) pre-trained, and (b) fine-tuned processes.

The Auto-Encoder is an unsupervised multi-layered neural network consisting both an encoder and a decoder network. Its encoder seeks to transform inputs data-points from a high unto a low-dimension via an encoding function $f_{encoder}$ as in Eq. 2 – where x^m is a data point, and h^m is the encoding vector obtained. Conversely, its decoder network seeks to reconstruct the function using $f_{decoder}$ as in Eq. 3 with x^m as decoding vector from h^m . Thus, reverts the operations of the encoder. Ojugo and Eboka [34] in Gilrot and Bengio [35] details specific algorithms for encoding as well as decoding functions respectively as in Eq. 2 and Eq. 3 respectively.

$$h^m = f_{encoder}(X^m) \quad (2)$$

$$X^m = f_{decoder}(h^m) \quad (3)$$

At pre-training phase, N autoencoders can be stacked on to an N -hidden-layer so that with input accepted, the input layer and first hidden layer acts an encoders of the first auto-encoder. They are trained to minimize reconstruction error. Training parameter(s) of the encoder are used to initialize first hidden layer before proceeding to second hidden layer. There, the first and second hidden layers are selected as encoder(s) and as in earlier stage, the second hidden layer is

initialized by the second trained auto-encoder. This process continues till N th auto-encoder is trained to initialize final hidden layer. With all hidden layers stacked in the auto-encoder at each training N -times, they are thus regarded as pre-trained. This feat is significantly better than random initialization. It also achieves better generalization [34-39].

Fine-tuning is a supervised phase to enhance deep learning by retraining the optimized data labels. It computes errors as using back-propagated stochastic gradient descent (SGD), which randomly selects data, and iteratively update gradient direction with the weight parameters. It converges faster and does not require the entire dataset to simulate tractability – making it suitable for such complex structure. Eq. 4 yields E as a loss function, y is label and t is output for the proposed deep learning network [34, 36-38]:

$$E = \frac{1}{2} \sum_{j=1}^M (y_i - t_i)^2 \quad (4)$$

The gradient of the weight w is obtained as a derivative of the error equation – so that an updated SGD is given by Eq. 5 with η is step-size, h is number of hidden layers [29, 47]:

$$W_{ij}^{new} = W_{ij}^{old} - \eta \cdot (y_j - t_j) \cdot y_j (1 - y_i) \cdot h_i \quad (5)$$

C. Experimental Proposed GAMNN Model

The experimental model is trained as thus:

- a. Input is received via GA-block (basic operators such as encoder, selector, swapper, changer, and CGA terminator is used to train dataset. On completion, trained (optimized) dataset feats are held within knowledge base [28].
- b. The MNN receives optimal dataset from knowledgebase with successive labeled/unlabeled rules perceived learnt instances [25-27]. The optimized-rules are propagated as IF-THEN rules (enhanced data with predefined variables) classified into various diabetes classes. Rules are modeled as a production system with four components: (i) **rule set** contains in each rule how to apply rule(s), and task(s) to be performed, (b) **knowledgebase** of optimized selected data stored as IF-THEN rules (diabetes) classes, (c) a **control** strategy to specify the order in which rules are compared to those in the knowledgebase to find a match and it seeks also a way to resolve conflicts that arise when several rules are matched at the same time, and (d) a rule **applier**. Also, the MNN provides a self-learning capability and acts as the principal component analyzer with rules optimized by CGA's recombination and mutation operators so that the trained model or network can effectively, autonomously classify transaction into both class types.
- c. Last stage of the network acts as a decision support and recognition system, with predicted values (output) and the automatic update of rules-knowledgebase, as transactions are diagnoses on encounter of new data, and consequently classified into class-types.

3. Result Findings and Discussion

A. Parametric Estimation, Tuning and Findings

Dataset is divided into 60:40 ratio. 60% for training and 40% for testing. The prowess cum predictive capability of the model is identified via data labels amongst the rule-based optimized dataset. Our deep learning Kohonen map uses 14-neurons at its input layer (a neuron for each feat). We also use 3-neurons for output layer (a neuron for each possible class of type-I, type-II and gestational). Other parameters are: (a) number of epochs, (b) activation function, (c) learning rate and (d) hidden layer topology. We used the Rectified Linear Unit (ReLU) Activation Function with 500-epochs. Optimal values were reached at 100, 300 and 500 epochs taking into account accuracy and training time. We seek minimum training error that will result in best fit, selecting the number of hidden layers was done via a trail-and-error method, and examining the results [25].

Figure 2 shows fitness score values for the generated rules. The rules generated that were fit was generally 82-transaction rules. These were partitioned into row values, each of which corresponds to an array of chromosomes for rules classified into diabetes classes.

Predictive Intelligent Decision Support Model in Forecasting of the Diabetes Pandemic Using a Reinforcement Deep Learning Approach

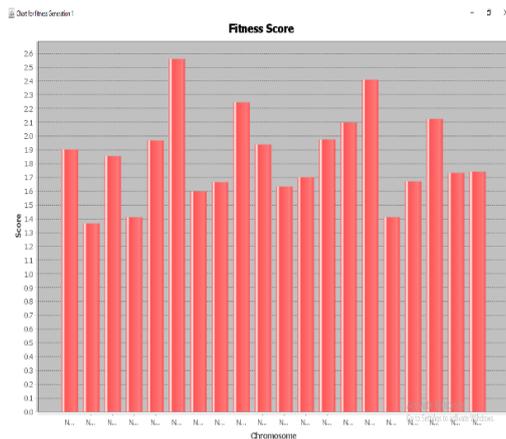


Fig.2. Fitness Score graph of the model

B. Discussion of Findings

The best possible number of layers was achieved by running tests on a single layer with 1 to 14-neurons at first instances – to yield greatest f-score with the least amount of training loss time. Addition of a second hidden layer of neurons from 1 to 17 yielded scores also; And, addition of a third hidden layer using the best possible number of neurons produced the greatest f-score. Thus, we selected the overall best possible hidden layer configuration. Results of first hidden layer with configuration of 14-neurons and f-score of 92% at 9th-iteration with a training loss of 1.140. F-score shows accuracy of each run – since we used an unbalanced dataset to train/test model with more records in normal class than in benign class [26-38].

Table 2. First hidden layer configuration analysis

Hidden Layer	Precision	Recall	F1-Score	Iteration	Training Loss	Epoch
1	0.84	0.92	0.88	44	0.294	500
2	0.84	0.92	0.87	24	0.278	500
3	0.84	0.92	0.88	26	0.293	500
4	0.84	0.92	0.88	9	0.501	500
5	0.89	0.55	0.64	19	1.496	500
6	0.94	0.94	0.92	18	1.400	500
7	0.86	0.53	0.63	4	2.230	500
8	0.90	0.84	0.86	16	2.071	500
9	0.92	0.93	0.92	18	1.140	500
10	0.92	0.92	0.90	16	1.779	500
11	0.88	0.91	0.89	7	2.134	500
12	0.91	0.92	0.89	8	2.320	500
13	0.87	0.87	0.87	13	2.006	500
14	0.92	0.92	0.90	8	1.970	500

Table 3 shows first layer with 14-neurons and others neurons varying from 1 to 17. Extra neurons cater for extra processing for optimal. Hidden layers of 9 and 11 neurons yielded f-score of 93% and training loss of 0.39. The second hidden layer is favored as it yields greater f-score.

Table 3. Second hidden layer configuration analysis

Layer	Precision	Recall	F1	Iteration	Train-Loss	Epoch
9, 1	0.84	0.92	0.88	25	0.293	500
9, 2	0.84	0.92	0.88	29	0.292	500
9, 3	0.91	0.92	0.91	15	0.583	500
9, 4	0.87	0.87	0.87	5	1.058	500
9, 5	0.92	0.92	0.90	13	1.628	500
9, 6	0.91	0.92	0.89	10	1.996	500
9, 7	0.84	0.92	0.88	24	0.281	500
9, 8	0.93	0.93	0.92	11	1.884	500
9, 9	0.92	0.92	0.89	12	1.590	500
9, 10	0.90	0.92	0.90	12	1.731	500
9, 11	0.95	0.94	0.93	14	0.390	500
9, 12	0.93	0.93	0.91	12	1.130	500
9, 13	0.91	0.92	0.91	20	1.929	500
9, 14	0.92	0.93	0.90	13	2.237	500
9, 15	0.94	0.94	0.92	7	1.765	500
9, 16	0.85	0.52	0.62	7	2.010	500
9, 17	0.94	0.94	0.94	6	1.620	500

Table 4 shows third configuration with first and second layer having 9 and 11 nodes and varying third hidden layer. Best configuration is 9-11-14 neurons, yielding f-score of 92% with a training loss at 0.560.

Table 4. Third hidden layer configuration analysis

Layer	Precision	Recall	F1	Iteration	Train-Loss	Epoch
9, 11,1	0.83	0.91	0.87	32	0.287	500
9, 11,2	0.91	0.92	0.89	6	1.592	500
9, 11, 3	0.83	0.91	0.87	29	0.280	500
9, 11, 4	0.90	0.91	0.90	16	1.564	500
9, 11, 5	0.92	0.92	0.90	18	0.741	500
9, 11, 6	0.93	0.92	0.89	21	0.282	500
9, 11, 7	0.92	0.93	0.90	6	1.322	500
9, 11, 8	0.90	0.86	0.88	6	1.239	500
9, 11, 9	0.90	0.91	0.90	7	1.886	500
9, 11, 10	0.88	0.91	0.89	8	0.623	500
9, 11, 11	0.92	0.93	0.91	5	2.000	500
9, 11, 12	0.86	0.83	0.85	11	2.370	500
9, 11, 13	0.86	0.83	0.84	8	2.350	500
9, 11, 14	0.93	0.92	0.92	15	0.560	500
9, 11, 15	0.93	0.93	0.91	8	1.204	500
9, 11, 16	0.94	0.94	0.92	8	1.730	500
9, 11, 17	0.87	0.54	0.63	12	1.730	500
9, 11, 18	0.93	0.94	0.93	6	1.850	500
9, 11, 19	0.93	0.93	0.90	9	0.660	500
9, 11, 20	0.92	0.92	0.90	28	1.180	500

C. Model Performance Evaluation

From our confusion matrix, (a) sensitivity measures how likely a model will predict the presence of all classes of diabetes symptoms when it is present, (b) specificity measures how likely model will detect the absence of diabetes symptoms when it is not present, and (c) accuracy measures the proportion of the true results seen as the degree of truth of a prediction. Thus, we compute the sensitivity, error rate and accuracy to evaluate model performance as in Eq. 6 to Eq. 8 – given that the values TP = 53, TN = 7, FP = 1 and lastly, FN = 1. Thus, we have:

$$\begin{aligned} \text{Sensitivity} &= \frac{TP}{TP+FN} \\ &= \frac{53}{53+1} * 100 = 98.1\% \end{aligned} \quad (6)$$

$$\begin{aligned} \text{Specificity} &= \frac{TN}{TN+FP} \\ &= \frac{7}{7+1} * 100 = 87.5\% \end{aligned} \quad (7)$$

$$\begin{aligned} \text{Accuracy} &= \frac{TP+TN}{TP+TN+FP+FN} \\ &= \frac{53+7}{53+7+1+1} = 95\% \end{aligned} \quad (8)$$

The model is found to have a sensitivity of 98%, specificity value of 87.5% and prediction accuracy of 95% (0.95) for data inclusion that were not originally used to train the mode. This is in agreement with [26-27, 30-31].

4. Summary and Conclusion

After testing the model, the results were also compared to other benchmark model as thus: (a) GA took 43seconds to find the solution after 98-iterations (at best). Recall that for deep learning, we observed earlier that our best forecast and optimal values were reached at 100, 300 and 500 epochs. CGA was run 25 times (to eradicate non-biasness), and it found optima each time – though, convergence time varied between 0.89seconds and 56seconds. We later observed that the model convergence time depended on how close the initial population is to the solution as well as on the mutation applied to the individuals in the pool. The use of the rule-based solution was adopted using fuzzy variable dataset (as preprocessor). Our MNN architecture is a larger network, comprising of smaller network. Its modularization allows for easy learning and understandability of the underlying feats of interest, grants model greater flexibility via parallel task execution and compartmentalization, greater adaptability and eases code reuse. MNN passes input via task decomposition and training modules via a multi-objective, multi-agent and multi-region support module that aids effective classification.

References

- [1] A.A. Ojugo., D.O. Otakore., Improved early detection of gestational diabetes via intelligent classification models: a case of Niger Delta region of Nigeria, *J. of Computer Science & Application*, 6(2): pp82-90, doi: 10.12691/jcsa-6-2-5, 2018
- [2] A.A. Ojugo., A. Eboka., R.E. Yoro., M. Yerokun., F.N. Efozia., Hybrid model for early diabetes diagnosis, *Maths. and Computers in Science & Industry*, 50, pp207-217, 2015, [web] www.semanticscholar.org/paper/Hybrid-Model-for-Early-Diabetes-Diagnosis-Ojugo-Eboka/662ce32a1f353eca02391a4a0cfe684499ad4448
- [3] M. I. Harris., Diabetes in America: Epidemiology and scope of the problem. *Diabetes Care*, 21(3), pp. 11-14, 1998
- [4] A.C. Menezes., P.R. Pinheiro., M.C. Pinheiro., T. Pequeno., Towards applied hybrid model in decision making: support the early diagnosis of type 2 diabetes, *In Proc of Maths & Computer in Science and Engineering*, doi: 10.1007/978-3-642-34062-8_84
- [5] D.M. Holmes, The person and diabetes psychosocial context. *Diabetes Care*, 9(2), pp.194-206, 1986.
- [6] The Expert Committee., On the diagnosis and classification of diabetes mellitus. Report of the Expert Committee on the diagnosis and classification of diabetes mellitus. *Diabetes Care*, 20, pp. 1183-1197, 1997
- [7] American Diabetes Association. Standards of Medical Care in Diabetes – 2009. *Diabetes Care*, 32: S13-61.
- [8] N.H. Barakat, A.P. Bradley, M.N.H Barakat, Intelligible support vector machines for diagnosis of diabetes mellitus, *IEEE Transactions on Information Technology in Biomedicine*, 14(4), pp 1114-1120, 2010
- [9] Canadian Diabetes Association. Standards of Medical Care in Diabetes 2014, *Journal of Diabetes Care*, 32, S13 – 16.
- [10] G. Berks, D. Keyserlingk, J. Jantzen., et al., Fuzzy clustering: versatile means to explore medical database, ESIT, Aachen, Germany, 2000
- [11] S. Chinenye, E. Young, State of diabetes care in Nigeria: a review, *The Nigerian Health Journal*, 11(4), pp101-106, 2011.
- [12] R. Goldenberg, Z. Punthakee., Definition, classification and diagnosis, prediabetes and metabolic syndrome, 37(1), S8-S11, 2013
- [13] A. Edo, G.O. Edo, O.A. Ohehen, N.P. Ekhaton, W.C. Ordiah, Age and diagnosis of type-2 diabetes seen in Benin City Nigeria, *African J. Diabetes Medicine*, 23(1), pp18, 2015
- [14] M. Khashei, S. Eftekhari, J. Parvizian, Diagnosing diabetes type-II using a soft intelligent binary classifier model, *Review of Bioinformatics and Biometrics*, 1(1), pp 9 – 23, 2012
- [15] O. Vaarala, M. Knip, J. Paronen, A.M. Hamalainen, P. Muona, P., Vaatainen, M., Ilonen, J., Simell, O and Akerblom, H.K., Cow's milk formula feeding induces primary immunization to insulin in infants at genetic risk for type-1 diabetes, *Diabetes*, 8(7), pp1389-1394, 1999
- [16] A.A. Ojugo., O.D. Otakore., Empirical evaluation for intelligent predictive models in the prediction of potential cancer problematic cases in Nigeria, *Journal of Mathematical and Computational Science*, 2020
- [17] A.A. Ojugo., A.O. Eboka., Modelling behavioural evolution as social predictor for coronavirus contagion and immunization in Nigeria, *J. Applied Sci., Engr., Tech. & Education*, 3(2): pp37–45, 2021, doi: 10.35877/454RI.asci130
- [18] A.A. Ojugo., O.D. Otakore., Forging an optimized Bayesian network model with selected parameter for detection of Coronavirus in Delta State, *J. Appl. Sci., Engr., Tech. & Edu.*, 3(1): pp37–45, 2020, doi: 10.35877/454RI.asci2163
- [19] R.E. Yoro., A.A. Ojugo., Quest for prevalence rate of Hepatitis-B infection in Nigeria: comparative study of supervised versus unsupervised model, *American Journal of Modeling & Optimization*, 7(2): 42-48, doi: 10.12691/ajmo-7-2-2, 2019
- [20] A.A. Ojugo., I.P. Okobah., Prevalence rate of hepatitis-B virus infection in Niger Delta region of Nigeria using graph-based diffusion heuristic model, *IJCAOnline Int. Journal of Computer Application*, 179(39): pp27 –33, 2018
- [21] A.A. Ojugo., D.A. Oyemade., D. Allenotor., R.E. Yoro., C.N. Anujeonye., Immunization problem for Ebola virus in rural Sierra-Leone, *African J. of Comp. & ICT.*, 8(1): pp1–10, 2015
- [22] A.A. Ojugo., F.O. Aghware., R.E. Yoro., M.O. Yerokun., A.O. Eboka., C.N. Anujeonye., F. Efozia., Predict behavioral evolution on graph model, *Adv. in Net.*, 3(2): pp8-21, 2015.
- [23] A.A. Ojugo., J. Emudianughe., R.E. Yoro., E. Okonta., A.O Eboka., A hybrid neural network gravitational search algorithm for rainfall runoff modeling and simulation in hydrology, *Progress in Intelligence Computing and Applications*, 2(1): 22-33, doi: 10.4156/pica.vol2.issue1.2, 2013
- [24] A.A. Ojugo., R.E. Yoro., Computational intelligence in stochastic solution for Toroidal N-queen, *Progress in Intelligence Computing and Applications*, 2(1): 46-56, 2013
- [25] B. Ghazale, Reasoning using modular neural network – innovative solution to address question answering AI tasks, retrieved from [web]: <https://towardsdatascience.com/reasoning-using-modular-neural-networks-f003cb6109a2?gi=7dbcd12eb7c>, July 18, 2020
- [26] R.E. Yoro., Machine learning optimization model for network intrusion detection, Unpublished Doctoral thesis submitted to Babcock University Ilishan-Remo, Ogun State, Nigeria, November 2020
- [27] A.A. Ojugo., R.E. Yoro., Forging a deep learning neural network intrusion detection framework to curb the distributed denial of service attack, *Int. Journal of Electrical and Computer Engineering*, 11(2): pp1498-1509, doi: 10.11591/ijece.v11i2.pp1498-1509, 2021
- [28] A.A. Ojugo., E. Ekurume., Towards a more satisfied user framework through a dependable-secured hybrid deep learning ensemble for detection of credit-card fraud, Submitted to *Int. J. of Emerging Trends in Engineering Research*, 2020
- [29] A.A. Ojugo., E. Ekurume., *Spatio-temporal solution for credit-card fraud using a genetic algorithm trained modular neural network ensemble*, Submitted to *Int. J. of Emerging Trends in Engineering Research*, 2020
- [30] A.A. Ojugo., A. Eboka., E. Okonta., R.E. Yoro., F.O Aghware., Genetic algorithm rule-based intrusion detection system, *J. Emerging Trends in Comp. Info. Sys.*, 3(8): pp1182-1194, 2012
- [31] M. Perez, T. Marwala, Stochastic optimization approaches for solving Sudoku, *IEEE Transaction on Evol. Comp.*, pp.256–279, 2011
- [32] R. Reynolds, Introduction to cultural algorithms, *Transaction on Evolutionary Programming*, pp.131-139, 1994.

- [33] R. Ursem, T. Krink, M. Jensen, Z. Michalewicz, Analysis and modeling of controls in dynamic systems. *Transaction on Memetic Systems and Evolutionary Computing*, 6(4), pp.378-389, 2002
- [34] A.A. Ojugo, A.O., Eboka, Modeling solution of market basket associative rule mining approaches using deep neural network, *Digital Technologies*, 3(1): pp1–8, 2018, doi: 10.12691/dt-3-1-1, [web]: www.sciepub.com/dt/content/3/1
- [35] Y. Bengio., A. Courville, P. Vincent., Representation learning: a review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence* , 35 (8), pp1798-1828, 2013.
- [36] G. Hinton, L. Deng, D. Yu, G.E. Dahl, A.R. Mohamed et al., Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Process. Mag.*, Vol. 29, pp82–97, 2012.
- [37] X. Glorot, Y. Bengio, Understanding the difficulty of training deep feedforward neural networks, In Proc. of 13th Int Conf. on Artificial Intelligence and Statistics, Sardinia, Italy, 13–15 May 2010; pp. 249–256.
- [38] D. Erhan, Y. Bengio, A. Courville, P.A. Manzagol, P. Vincent, S. Bengio, Why does unsupervised pre-training help deep learning?, *Journal of Machine Learning Res.*, Vol. 11, pp625–660, 2010
- [39] A.A. Ojugo., R.E. Yoro., Empirical solution for an optimized machine learning framework for anomaly-based network intrusion detection, *Technology Report of Kansai University, TRKU-13-08-2020-10996*, 62(10): pp6353-6364, 2020
- [40] D.A. Oyemade., A.A. Ojugo., A property oriented pandemic surviving trading model, *Int. J. Advanced Trends in Computer Science and Engineering*, 9(5): pp7397-7404, 2020
- [41] A.A. Ojugo., A. Eboka., Empirical evaluation on comparative study of machine learning techniques in detection of distributed denial of service attack, *J. Applied Sci. Eng. Tech. & Edu.*, 2(1): pp18–27, 2020, doi: 10.35877/454RI.asci2192
- [42] A.A. Ojugo., D.A. Oyemade., Predicting diffusion dynamics of coronavirus in Nigeria through ties-strength threshold on a cascading SI-graph, *Technology Report of Kansai University, TRKU-13-08-2020-10998*, 62(8): pp4313-4323, 2020
- [43] A.A. Ojugo., A.O. Eboka., Signature-based malware detection using approximate Boyer Moore string matching algorithm, *International Journal of Mathematical Sciences and Computing(IJMSC)*, 5(3): pp49-62, doi: 10.5815/ijmsc.2019.03.05, 2019
- [44] R.E. Yoro., A.A. Ojugo., Quest for prevalence rate of Hepatitis-B infection in Nigeria: comparative study of supervised versus unsupervised model, *American Journal of Modeling & Optimization*, 7(2): 42-48, doi: 10.12691/ajmo-7-2-2, 2019
- [45] A.A. Ojugo., E. Ben-Iwhiwhu, et al., Malware propagation on social time varying networks: comparative study of machine learning frameworks, *International Journal of Modern Education and Computer Science (IJMECS)*, 6(8): pp25-33, doi: 10.5815/ijmeecs.2014.08.04, 2014

Authors' Profiles



Arnold Adimabua Ojugo received his BSc in 2000, MSc in 2005 and PhD in 2013 – all in Computer Science from The Imo State University Owerri, The Nnamdi Azikiwe University Awka, and The Ebonyi State University Abakiliki respectively. He is an Associate Professor currently at Department of Computer Science (Federal University of Petroleum Resources Effurun) in Delta State, Nigeria. His research interests are in: Intelligent Systems, Machine-Learning, Performance and Ubiquitous Computing, Data Security and Graph Theory. He is also an Editor with the Progress for Intelligent Computation and Application, SciencePG Journals, and others. He is also a member of: Nigerian Computer Society, Computer Professionals of Nigeria and International Association of Engineers (IAENG).



Elohor Ekurume, in 2009, received her BSc in Computer Science from Benson Idahosa University, Benin-City in Edo State, and her MSc in Information Technology from the University of Bradford, Bradford, United Kingdom in 2013. She currently lectures with the Department of Computer Science at the Delta State University Abraka in Delta State, Nigeria. Her research interests are in: Intelligent Systems, Data Mining, Machine-Learning, and Ubiquitous Computing. She is a member of: Computer Professionals of Nigeria.

How to cite this paper: Arnold Adimabua Ojugo, Elohor Ekurume, " Predictive Intelligent Decision Support Model in Forecasting of the Diabetes Pandemic Using a Reinforcement Deep Learning Approach", *International Journal of Education and Management Engineering (IJEME)*, Vol.11, No.2, pp. 40-48, 2021. DOI: 10.5815/ijeme.2021.02.05