Diabetes Prediction by using Bacterial Foraging Optimization Algorithm And Artificial Neural Network

Rabina
Research Scholar (C.S.E)
S.S.C.E.T, Badhani.
Pathankot,India.
rabinabagga@gmail.com

Prof.Meenakshi Sharma
Associate Professor (C.S.E)
S.S.C.E.T, Badhani.
Pathankot,India.
sharma.minaxi@gmail.com

***Abstract*-Diagnosis of any diseases earlier is always preferable .diabetes is one such diseases.it reaches to the epidemic proportional in many developing countries and also reaching to industrialized nations. In this study, we investigate a programmable method to analyze diabetes diseases based on BFO and ANN.ANN is a powerful tool and solves the problems such as classification and prediction. Its performance is highly dependent on its architecture and weights. To gain high efficiency in ANN, an appropriate Architecture and learning Algorithm is preferable. The experiments on dataset occupied from undertaking repository to verify the potential of BFO algorithm and showed that BFO-ANN is more accurate than traditional ANN in terms of classification.**

General Terms-Artificial Neural Network, BacterialForaging Optimization Algorithm.

I.INTRODUCTION

In today's world, technology is a crucial ingredient of health care. Indeed, all health care consists of either mortal collaboration, the solicitation of tools, or, most commonly, both. Consideration of technology is important in any examination of the organization and functioning of health care services and systems for many reasons: Knowledge is a major module of modern health care costs and possibly the key driver of future cost. Major regulatory frameworks and organizations exist solely to cope with the overview and use of safe, effective and efficient technology in health care. Improvements in health care knowledge have the potential to be major sources of economic wealth as well as forces for change in the organization of health care [1]. Technology transforms the way that help the healthcare practitioners can continue to find ways to improve the effective management of chronic disease such as diabetes. Diabetes grows when glucose can't pass in the body's cells to be used as fuel. This is because your pancreas doesn't produce any insulin, or not enough insulin, to support glucose to move into your body's cells – or the insulin that is formed does not work appropriately (known as insulin resistance).

The two most common forms of diabetes, Type 1 diabetes and Type 2 diabetes. In Type 1 diabetes there is no insulin to expose the cells. Type 2 diabetes arises when there is not enough insulin or the insulin is there but not functioning appropriately.it is usually enduring disease, also called chronic, and needs consistent treatment and regular consideration. Diabetes can be fatal if not treated quickly and regularly. In this paper we activist the use of Feature selection

methods that play an important role in classifying systems such as Bacterial foraging optimization. To increase the system performance, the system segregates the irrelevant and terminated structures from a dataset; thus, due to this dimension of dataset will be reduced. Subsequently, its complication is decreased, and its recital can be increased. There are many general methods for dealing this problem such as: principal component analysis (PCA) linear discriminant analysis (LDA), forward feature selection (FFS) and backward feature selection (BFS), Particle swarm optimization**.** Furthermore, to accommodate inconsistently high service procedure demand, our resource arrangement must spontaneously increase or contract with the changing load.

To ensure this, we augment our system with bacterial foraging algorithm that scales the underlying resource pool in accordance with the predicted resource procedure configurations.

A. ARTIFICIAL NEURAL NETWORK

A neural network is an integrated group of artificial neurons parallel made up of simple processing units, which has a tendency for storing tentative knowledge and making it available for use.

The stimulation for the field of neural networks came from the desire to produce artificial system capable of sophisticated perhaps intelligent computation similar to those that the human brain routinely performs and thereby improving the understanding of brain [2].

Artificial neural network involve three kinds of layers: input layer,hidden layer and output layer.

- Input Layer: This layers comprises of input units which signifies the unprocessed information provided for the networks.
- Hidden Layer: This layer is characterized by hidden units which are inclined by the behaviour of the input units and the weight that join these input and the hidden components.
- Output Layer:The output units behaviour is dependent on the particularity of the hidden units and the weights joining the hidden and output units [3].

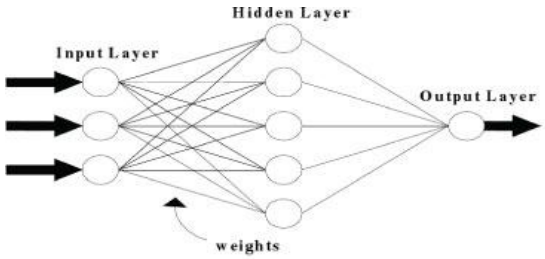


Figure1.Artificial Neural Network with layered architecture.

B. ANN AND ANN TRAINING

A neural network has to be designed such that the application of a set of inputs creates the desired set of inputs. Various methods to set the strong point of the connections that occur. One way is to train the neural network by serving it training patterns and letting it changes its weight according to some learning rule.

C. PARADIGMS OF LEARNING:

We can categories the learning situation into two distinct sorts:

- Supervised learning in which the network is trained by presuming it with input and identical output patterns. These input output pairs can be provided

by external teacher or by a system which contain a network.

- Unsupervised learning in which an output unit is trained to respond to cluster of patterns with in the input. In this paradigm the system is supposed to discover statistically designed salient features of the input population.

D. BFO ALGORITHM:

Bacterial Foraging Optimization (BFO) Algorithm is suggested by Kevin Passino, the family of nature inspired optimization algorithms. These new algorithms give a new idea to the application of a swarm of E.coli bacteria in multi-optimal function optimization. Bacteria search for nutrients is a manner to maximize energy obtained per unit time. Individual bacterium also interconnects with others by sending signals. A bacterium takes foraging decisions after considering two preceding factors. The process, by which searching for nutrients by bacterium take place through small steps called chemotaxis. The main idea of BFOA is simulating chemotactic measure of virtual bacteria in the difficult exploration space.
D_s: Dimension of the search space,
T_p: Overall number of bacteria in the population,
C: Chemotactic steps,
S: Swimming length.
R: Reproduction steps,
E_{ed}: Elimination-dispersal events,

P_{ed}: Elimination-dispersal probability,
Foraging theory is based on the assumption that animals search for and obtain nutrients in a way that exploits their energy intake *E* per unit time *T* spent foraging. Hence, they try to exploit a function like E/T (or the average rate of energy intake maximize). Maximization of such a function provides nutrient sources to stay alive and supplementary time for other important activities (e.g., fighting, fleeing, mating, and reproducing). Mate finding activities sometimes bear similarities to foraging. Clearly, foraging is very different for different species. Herbivores normally discover food easily but must eat a lot of it. Carnivores normally find it hard to detect food but do not have to eat as much since their food is of high energy value. The "surroundings" establishes the outline of nutrients that are available and it places constraints on obtaining that food (e.g., small portions of food may be separated by large distances). During foraging there can be threats due to predators, the prey may be transportable so it must be chased and the physiological characteristics of the forager constrain its capabilities and ultimate success. E.coli bacteria undergo four stages during the foraging process i.e.

Chemotaxis: This method pretends the movement of an E.coli cell through swimming and tumbling via flagella. Biologically an E.coli bacterium can transfer in two dissimilar ways. It can swim for a period of time in the same direction or it may tumble and alternate between these two modes of operation for the entire lifetime. Suppose $\theta^i(j, k, l)$ represents *i*th bacterium at *j*th chemotactic, *k*th reproductive and *l*th elimination-dispersal step. *C*(*i*) specify the extent of the steps taken in the random direction specified by the tumble (run length unit). Then in computational chemotaxis the movement of the bacterium may be represented by

$$\theta^i(j+1,k,l) = \theta^i(j,k,l) + \frac{C(i)\Delta(i)}{\sqrt{\Delta^T(i)\Delta(i)}} \quad (1)$$

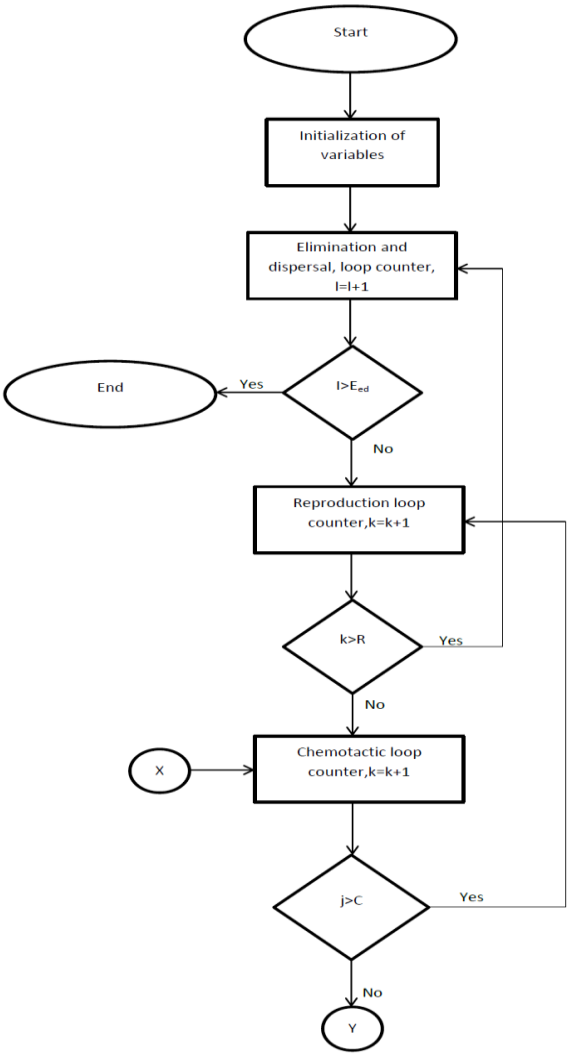
Where Δ indicates a vector in the random direction whose elements lie in [-1, 1].

Swarming
E.coli offers activities in which convoluted and stable–spatio-temporal patterns in a semisolid nutrient medium. The cells when moved using a high level of succinate .it will

release an attractment a spertate, that help to combine into groups and move into a concentric arrangement of swarms with high bacterial density.

Reproduction

After evolving through various stages of chemotaxis bacteria reached at reproduction stage which it is divided into two forms. The healthier half replaces the other half of bacteria which gets eliminated due to their poor foraging capabilities. This replacement gives a constant bacteria population.



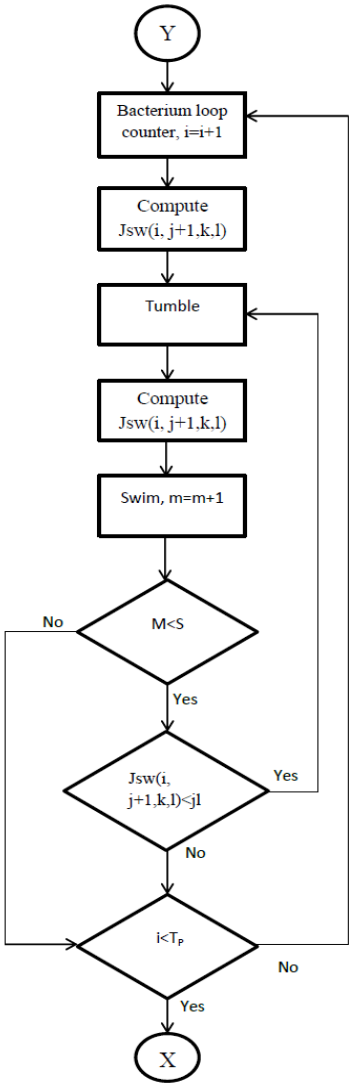


Figure 2. Flowchart of BFO [5]

Elimination and Dispersal

In this stage, some unpredicted event occurs such as bacteria in a region are killed or a group of bacteria dispersed to a new environment which may significantly alter the smooth process of evolution. From a wide outlook, elimination and dispersal are parts of the population-level long-distance motile behavior [4].

II. DIABETES DATASET USED

In this study, we use the UCI Diabetes diseases dataset introduced by Black C.L. (Blake C. L., 1998). This dataset contains 768 illustrations, where each illustration has 8 features which are eight clinical findings:

(i). Number of times pregnant

(ii). Plasma glucose concentration 2 hours in an oral glucose tolerance test.

(iii). Diastolic blood pressure (*mm Hg*)

(iv). Triceps skin fold thickness (mm)

(v). 2-hour serum insulin (*mu U/ml*)

(vi). Body mass index (*kg/m^2*)

(vii). Diabetes pedigree function

(viii). Age (years)

The class target variable contains values '0' or '1'. While '1' specify positive test for Diabetes, '0' specify negative test. At hand 268 cases in class '1' and 500 cases in class '0' [5].

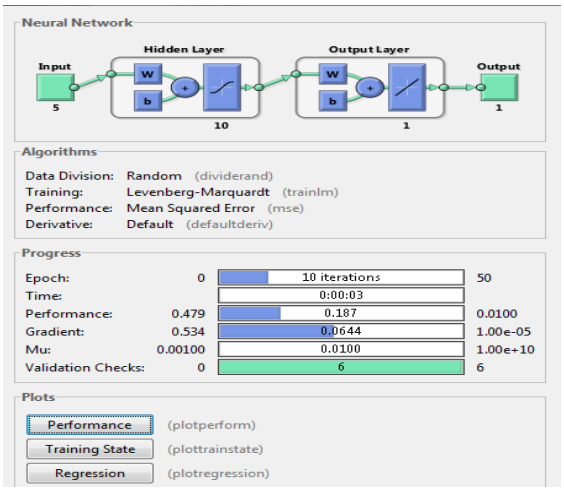
III. RESULTS & DISCUSSION

Feature selection using BFO

First of all we implement BFO algorithm on diabetes dataset for feature selection, it selects three features:

1. 'Number of times pregnant'
2. 'Plasma glucose'
3. 'Serum insulin'.

The neural network is applied on whole 768 samples of diabetes data as the train samples and amongst them, 55 samples are used as test for testing(accuracy).



A. Result Analysis

The predictive accuracy of different neurons is shown in the table below:

Neurons	Accuracy (%)
10	94.54
20	95.08
30	93.78
40	94.59
50	94.90
60	96.23
70	93.87
80	94.39
90	94.06
100	95.01

Table 1 – Predictive Accuracy of different neurons

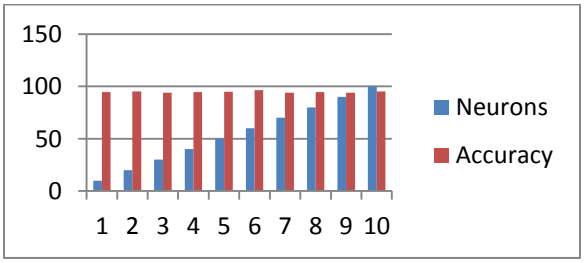


Figure 3. Predictive Accuracy of different neurons.

The performance plots of ANN Training are shown from Figure 4 to Figure 6. In Figure 4, the performance plot of trained ANN is shown respectively.

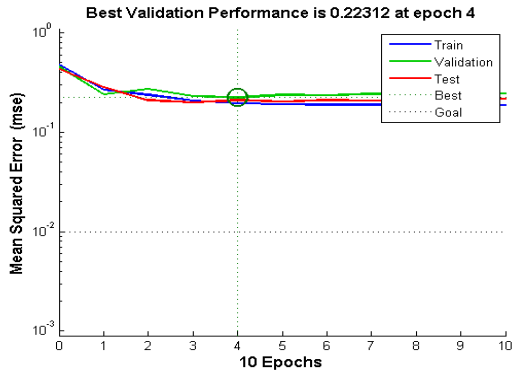


Figure 4. Performance plot of trained ANN.

In Figure 5, the regression plot of trained ANN is shown and in Figure 6, the training state plot of trained ANN is shown respectively.

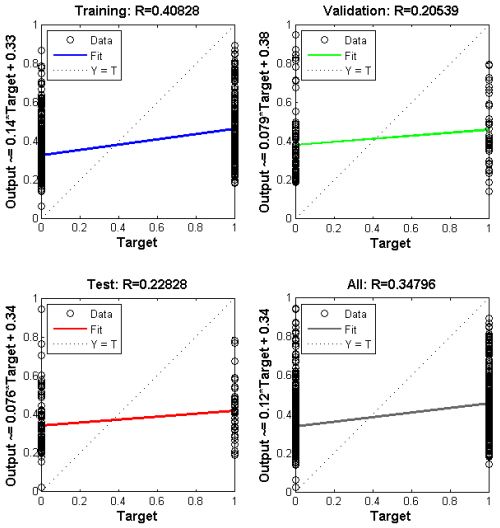


Figure 5. Regression plot of trained ANN.

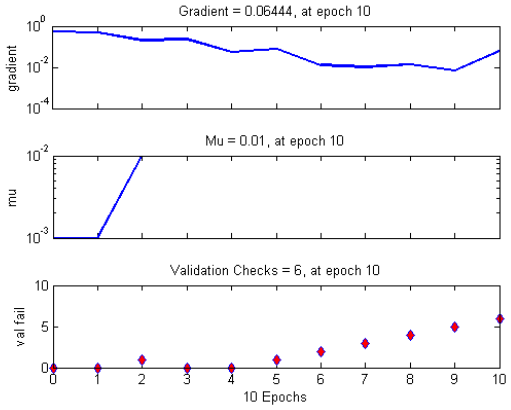


Figure 6. Training state plot of trained ANN.

IV. CONCLUSION & FUTURE WORK

The work presented in this paper is an innovative automatic model to diagnose the diabetes disease based on BFO & ANN. To discard the irrelevant features BFO was used, and then ANN was applied to the selected features to find out the classification accuracy. Results shows that it improve the accuracy and demonstrate that the proposed model is faster and take less time. Therefore BFO is regarded as an effective training algorithm for ANN.

FUTURE WORK

The next step of this work is to use Genetic Algorithm to improve optimization and use other SVM's kernel or other classifiers like: MLP, Random Forrest etc. Such an approach is expected to improve the flexibility of the system. In addition to the above we have identified that we can use Adaptive BFO in the place of BFO. It contains some advanced parameter for feature selection.

REFERENCES:

[1] A. Care, "International Technology , approaches to funding health care Care Health and Health," vol. 5.

[2] I. Koprinska, "Introduction to Neural Networks and Machine Learning What is this course about ? OMP4302 / 5322 Neural Networks , w1 , s2 Special Considerations Due to Illness or Misadventure University and School Policies," 2003.

[3] R. Kaur, "Artificial Neural Network Learning Enhancement using Bacterial Foraging Optimization Algorithm," vol. 102, no. 10, pp. 27–33, 2014.

[4] I. Nowadays, "CHAPTER 5 BACTERIAL FORAGING OPTIMIZATION ALGORITHM," pp. 62–73.

[5] D. Rani, "An Efficient Technique for Disease Diagnosis using Bacterial Foraging Optimization and Artificial Neural Network," 2013.

[6] G. Mahapatra, "A Study of Bacterial Foraging Optimization Algorithm and its Applications to Solve Simultaneous Equations," vol. 72, no. 5, pp. 2–8, 2013.

[7] Z. Hongwei and T. Liwei, "Cooperative Approaches to Bacterial Foraging Algorithm for Clustering," vol. 8, no. 4, pp. 81–90, 2015.

[8] N. D. Dementia, "Chapter 1 Introduction of Neurodegenerative Diseases (Dementia) and their diagnosis," pp. 13–28.

[9] R. Kaur and A. N. N. Training, "Bacterial Foraging Optimization Algorithm for Evolving Artificial Neural Networks," vol. 8, no. 5, pp. 16–19, 2015.

[10] S. Das, A. Biswas, S. Dasgupta, and A. Abraham, "Bacterial Foraging Optimization Algorithm : Theoretical Foundations , Analysis , and Applications."

[11] B. Krose and P. Van Der Smagt, "introduction to Neural," no. November, 1996.

[12] Ismail Ahmed Al-Qasem Al-Hadi, Siti Zaiton Mohamed Hashim, "Bacterial Foraging Optimization Algorithm for Neural Network Learning Enhancement", International Journal of Innovative Computing, Vol. 1, No. 1, 2012.

[13] Alireza Askarzadeh, "Artificial Neural network training using a new efficient optimization algorithm", Applied Soft Computing, Volume 13, Issue 2, February 2013, Pages 1206-1213.

[14] N. Kushwaha, V.S. Bisht and G. Shah, Genetic algorithm based bacterial foraging approach for optimization, National Conference on Future Aspects of AI in Industrial Automation (NCFAAIIA 2012), Proceedings published by International Journal of Computer Applications (IJCA), (0975-8887), vol. 17, pp. 11-14, September, 2012.

[15] Jun Li, Jianwu Dang,Feng Bu, Jiansheng Wang, "Analysis and Improvement of the Bacterial Foraging Optimization Algorithm",

Journal of Computing Science and Engineering,

Vol. 8, No. 1, March 2014, pp. 1-10.