International Journal of Computational Intelligence Research
ISSN 0973-1873 Volume 13, Number 4 (2017), pp. 595-614

http://www.ripublication.com

The novel hybrid Modified Particle Swarm Optimization – Neural Network (MPSO-NN) Algorithm for classifying the Diabetes

Karamath Ateeq[1], and Dr. Gopinath Ganapathy[2]

[1]Research Scholar, [2]Professor & Head, School of Computer Science, Engineering and Applications, Bharathidasan University, Tiruchirappalli, India.

Abstract

Diabetes Mellitus (DM) is a chronic, lifelong metabolism disorder. It affects the ability of the body system to use the energy found in food. The improper management of the disease will lead to Heart disease, kidney disease, eye disease, nerve disease and pregnancy complications. Classification model helps physicians to improve their prognosis, diagnosis or treatment planning procedures. It is most widely utilized in medical domain to explore patient's data and extract the knowledge from it. This paper aims in developing a novel hybrid classification algorithm for diabetes classification. The Particle Swarm Optimization algorithm is modified and combined with Multi Layer Perceptron Network with back propagation learning and is named as novel Modified Particle Swarm Optimization-Neural Network (MPSO-NN) Algorithm to classify the diabetic patients as positive as binary 0 and negative as binary 1. The metrics like Sensitivity, Specificity, False Positive Rate, Accuracy, Mean Squared Error, Regression, and Percentage of Error is used to evaluate the performance of the algorithm. It is found that the proposed novel hybrid algorithm is better in overall performance in classifying the data set.

Keywords: Diabetes, Classification, Modified PSO, Neural Network, Hybrid Algorithm, Convergence, Accuracy.

INTRODUCTION

Diabetes is a metabolic disease which results when the body does not produce enough Insulin. It is associated with increased risk of blindness, blood pressure, heart disease, and kidney disease and nerve damage [1, 2]. So, it is also named as silent killer. It is of two types namely Type I and Type II. Early diagnosis of the disease was improved but about half of patients with Type II are unaware of this disease at the onset. It is because the symptoms for Type II diabetes is common symptoms shared with other diseases. The only way for the diabetes patient to live with this disease is to keep the blood sugar as normal as possible without serious high or low blood sugars, and this is achieved when the patient uses a correct management or therapy which may include diet and exercising, taking oral diabetes medication or using some form of insulin. Hence, in addition to the evaluation of test results, physicians must pay attention to previous decisions which made for patients in the same conditions [2]. It is estimated that about 194 million people around the world is suffering from diabetes. It is also estimated that it will increase to 333 million or 6.3% by 2025 [4].

Therefore, with regard to importance of problem and to ensure make fast, accurate and meaningful decisions, classification systems (or pattern recognition systems) could be used. From the machine learning perspective, classification is the problem of identifying a set of observations into several categories, based on the training result of a subset of observations whose belonging category is known. The Pima is one of the most studied populations regarding diabetes, not only among American Indians, but in the world [3]. This paper aims in developing a novel hybrid classification algorithm for classifying the diabetes data.

This paper is organized as follows: Section 2 explains the general Particle Swarm Optimization Algorithm. Section 3 briefs about the Artificial Neural Network. Section 4 explicates the literature reviewed for this work. Section 5 describes the proposed hybrid algorithm for classification. Section 6 is about the experimentation and results and Section 7 concludes the work.

PARTICLE SWARM OPTIMIZATION ALGORITHM

Particle Swarm Optimization (PSO) is based on the social behaviour associated with bird's flocking for optimization problem. A social behaviour pattern of organisms that live and interact within large groups is the inspiration for PSO [8]. In PSO, each particle adjusts its own flying memory and its companion's flying involvement in order to fly in the search space with velocity.

In the basic PSO, Particle Swarm consists of 'n' particles. The position of each particle stands for potential solution in D-dimensional space. Individuals, potential solutions, flow through hyper dimensional search space. The experience or acquired knowledge about its neighbours influences the changes in a particle within the swarm. The PSO algorithm involves of just three steps, which are being replicated until stopping condition. They are as follows.

(i) Evaluate the fitness of each particle.
(ii) Update individual and global best functions.
(iii) Update velocity and position of each particle.

The position of each particle is influenced by the best-fit particle of the entire swarm [6]. Each individual particle i € [1… n] where n > 1, has current position in search space x_i, a current velocity v_i and a personal best position $P_{best,\, i}$ where i is the smallest value determined by objective function f. By using the $P_{best,\, i}$ the global best position G_{best} is calculated, which is the buck value obtained by comparing all the $P_{best,\, i}$.

The $P_{best,i}$ is calculated by using the formula

$$p_{best,i} = \{ \begin{matrix} P_{best,i} \; if \; f(x_i) > P_{best,i} \\ x_i \; if \; f(x_i) \le P_{best,i} \end{matrix} \qquad \text{..(1)}$$

The formula used to calculate Global Best Position G_{best} is

$$G_{best} = \{\min \{ P_{best,i} \}, where \; i \in [1, \dots\dots, n] \text{ where } n > 1 \qquad \text{.................(2)}$$

Velocity can be updated by using the formula

$$V_i = w_i v_i \; (t) + c_1 r_1 [p_i(t) - x_i(t)] + c_2 r_2 [g_i(t) - x_i(t)] \qquad \text{.................. (3)}$$

where v(t) is the velocity, w, c_1 and c_2 are user supplied co-efficient. The r_1 and r_2 are random values x (t) is the individual best solution; g (t) is the swarm's global best candidate solution. $w_i v_i$ (t) is known as inertia component. The inertia component is responsible for keeping the particle moving in same direction that it was originally heading. It lies between 0.8 and 1.2. Lower value of inertia component speeds up the convergence of swarm to optima. Higher values encourage the exploration of entire search space. $c_1 r_1$ [p_i (t)-x_i(t)] is known as cognitive component. It pretends as a particle's memory and it verges to return to the region of search space, where it experiences high individual factors. $c_2 r_2$ [(g_i(t)-x_i(t)] is known as social component, which causes the particle to move to the best region the swarm has found so far.

ARTIFICIAL NEURAL NETWORK

The indication for mounting the Neural Network was obtained from scrutinizing the clear pre-eminence of brain over conventional computers [7]. The reckoning speed of Neural Network is more over than the conventional computer because of its parallel computing nature. The encumbrance of the network can be concocted by using the efficient algorithm. Neural Network architecture generally falls under two categories namely Feed Forward and Feedback architecture. The three layers like input layer, one or more hidden layers and an output layer contribute to modelling a structure for Neural Network [9, 10]. The complexity of the system and study defines the number of hidden layers and neurons. The below Figure 1 shows a simple Neural Network diagram.

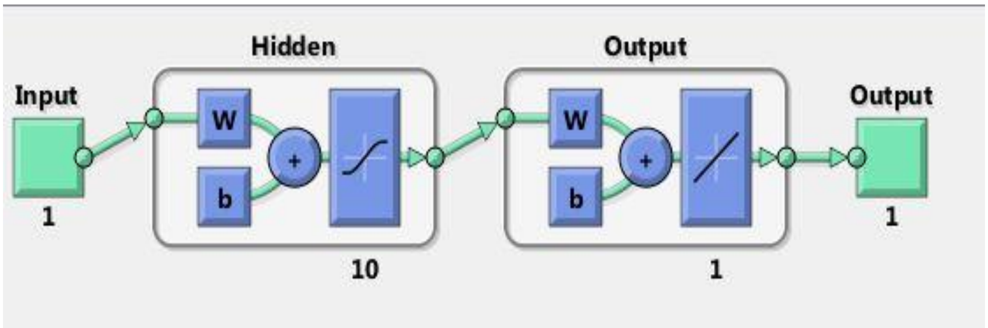


Figure 1: A sample Neural Network diagram

LITERATURE REVIEW

[11] used Ant Colony based classification system to extract a set of fuzzy rules for diagnosis of diabetes disease. This classification system was named as ANTMINER. The Classification Accuracy obtained was 84.24%. The results reveal that the FCS-ANTMINER outperforms several methods in classification accuracy for diabetes disease diagnosis. [12] applied data gravitation based classification algorithm and obtained 76.56% of classification accuracy. [13] developed a cascade learning system which was based on Generalized Discriminate Analysis (GDA) and Least Support Vector Machine (LSVM) to detect the Diabetes. A classification accuracy of 78.21% was obtained when this (LS-SVM) was applied on Pima Indian diabetic data set.

[14] used Fuzzy Neural Network and obtained a classification accuracy of about 81.8%. [15] has introduced an instance counting algorithm ARTMAP-IC. An accuracy of 81% was obtained by this method. [16] used hybrid algorithm which combines Artificial Neural Network and Fuzzy Neural Network. The accuracy obtained was 84.2%. [3] developed an intelligent machine learning technique for classifying the diabetes. Expectation Maximization Algorithm was used for clustering, Principal Component Analysis was used for noise removal and Support Vector Machine was applied for classification. The proposed method has improved the prediction accuracy and reduced the computation time. It was developed to assist medical practitioners in health care practice. [5] proposed a Hybrid Prediction Model (HPM) for the diabetes prediction. The filtrated feature selection method is used to select the most discriminatory predictors. Then two layered classification approach is applied by combining Support Vector Machine (SVM) and Neural Network to enhance the overall recognition rate of the model. The overall accuracy gained in this model is 96.09%

[17] used General Regression Neural Network (GRNN) for diabetes classification. This proposed method was also tested on Pima Indian Diabetes (PID). The accuracy obtained was 80.21%. [18] developed a method for diabetes classification using Genetic Programming. The classification was done through three stages namely Feature Selection, Feature generation and testing. K-Nearest Neighbour (KNN) and Support Vector Machine (SVM) was used for evaluating the selected features. [19] proposed an intelligent system based on Small World Feed Forward Artificial Neural

Network (SW-FFANN). It predicted with an accuracy of 91.66%. [20] developed an intelligent diagnosis system based on Linear Discriminant Analysis – Adaptive Neuro Fuzzy Inference System (LDA-ANFIS). The classification accuracy was around 84.61%.

PROPOSED METHOD

The Figure 2 explains the method proposed for the classification of the diabetes disease. The required dataset is collected from UCI website. The datasets used are PID data set and US diabetic data set [21]. The data set is initially pre-processed using the Genetic-Relative Reduct Algorithm. The reduced data is then subject to classification. The two types of ANN, Radial Basis Function Network (RBFN) and Multi Layer Perceptron Network (MLPN) are compared with each other to figure out which network is best apt network for the classification. Then PSO algorithm is modified as Modified Particle Swarm Optimization (MPSO) Algorithm.

The comparison and results obtained while training the MLPN and RBFN on PID and US data set is tabulated in Table 1. Figure 3 to Figure 7 depicts the results obtained using the PID and US data set. From the results it is inferred that the MLPN network performs well in classifying the data set. Hence, the MLPN network is adopted in this work for classification.

The MPSO and ANN are combined and the novel hybrid MPSO-NN is developed. The developed algorithm is then tested with the data set and the results are validated and the output will be the classified dataset as healthy (0) and diseased (1).

MAE- Mean Squared Error, RMSE-Root Mean Squared Error, RAE-Relative Absolute Error, RRSE-Root Relative Squared Error, TPR-True Positive Rate, FPR-False Positive Rate

Table 1: Performance Evaluation

Parameters taken	**Pima Indian Diabetic Data**		**US Diabetic Data**	
	MLPN	**RBFN**	**MLPN**	**RBFN**
Time	0.13	0.38	0.07	0.09
Kappa Statistic	0.3	0.3	0.50	0.49
MAE	0.3	0.4	0.17	0.16
RMSE	0.5	0.5	0.33	0.32
RAE	72.44	78.23	52.97	53.08
RRSE	105.32	105.34	81.24	84.22
TPR	0.7	0.6	0.8	0.8
FPR	0.3	0.4	0.2	0.2
Precision	0.67	0.68	0.849	0.841
Recall	0.67	0.68	0.858	0.848
F-Measure	0.67	0.68	0.852	0.844
ROC Value	0.7	0.6	0.9	0.89
Accuracy	68.84	67.48	85.77	84.77

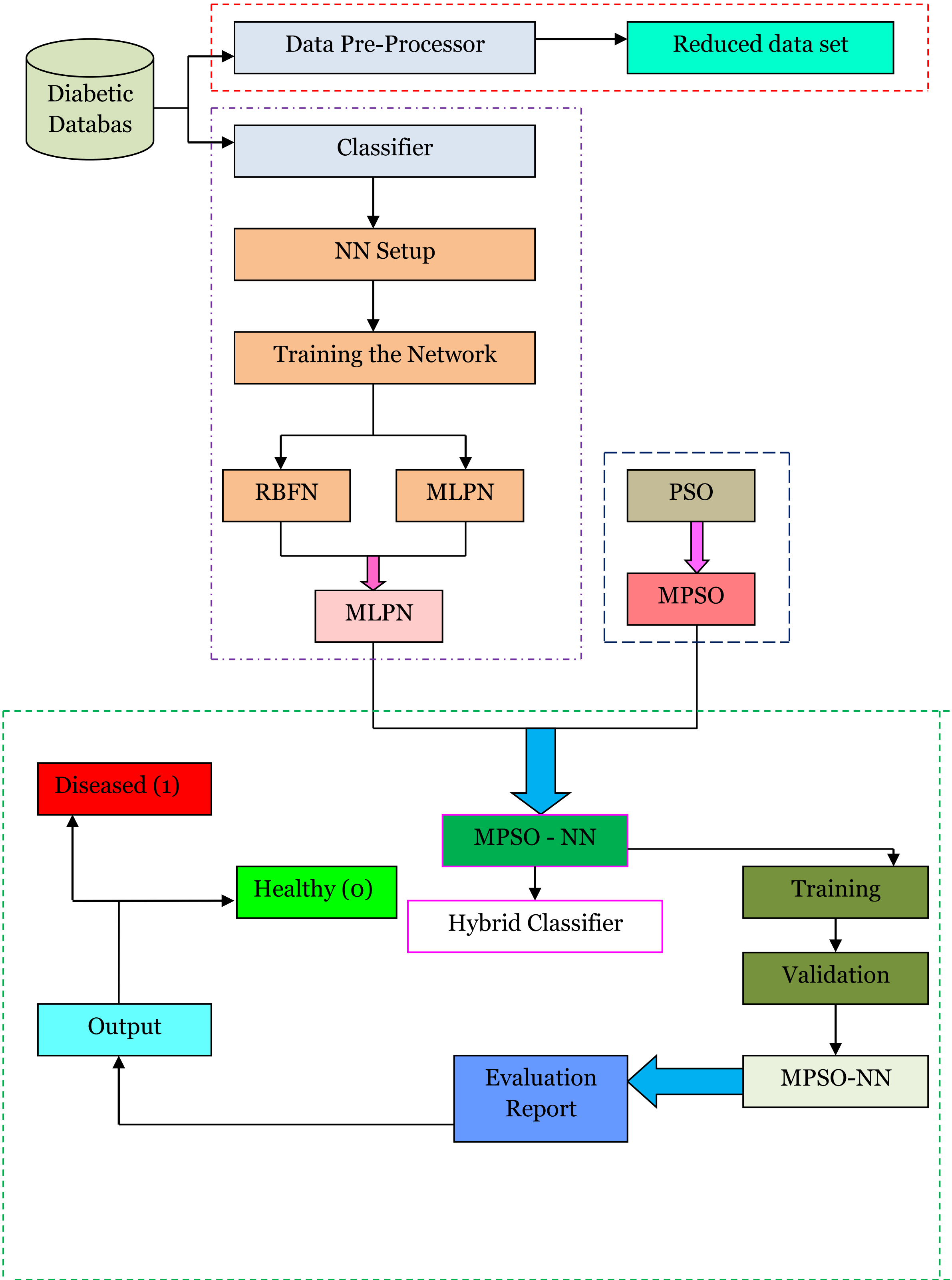


Figure 2: Overall Framework of the Classification Process

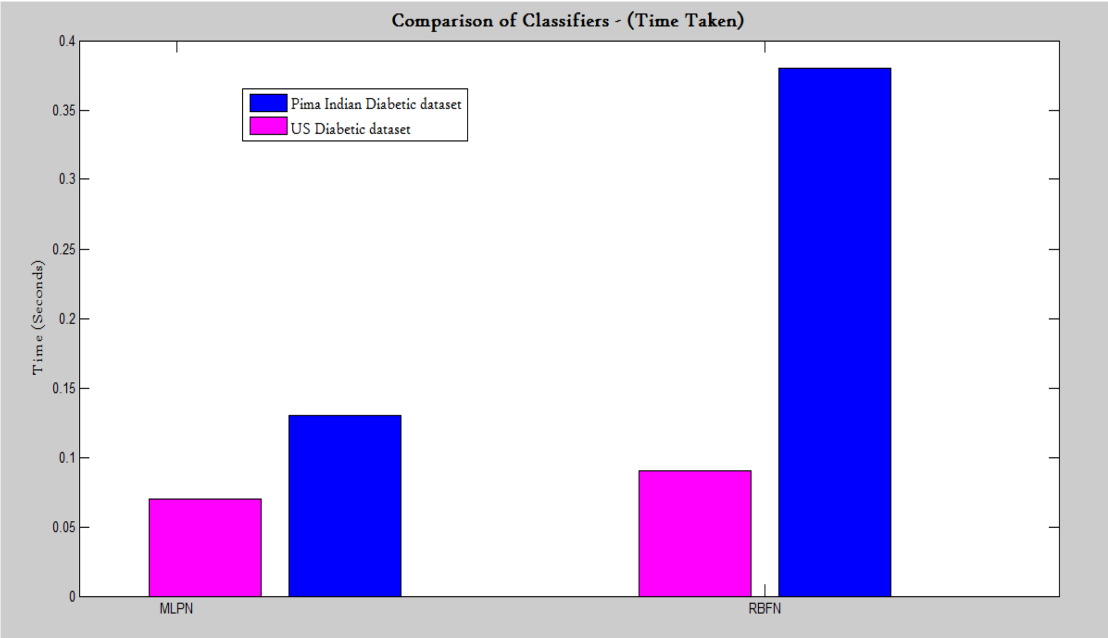


Figure 3: Comparison of Classifier – Time

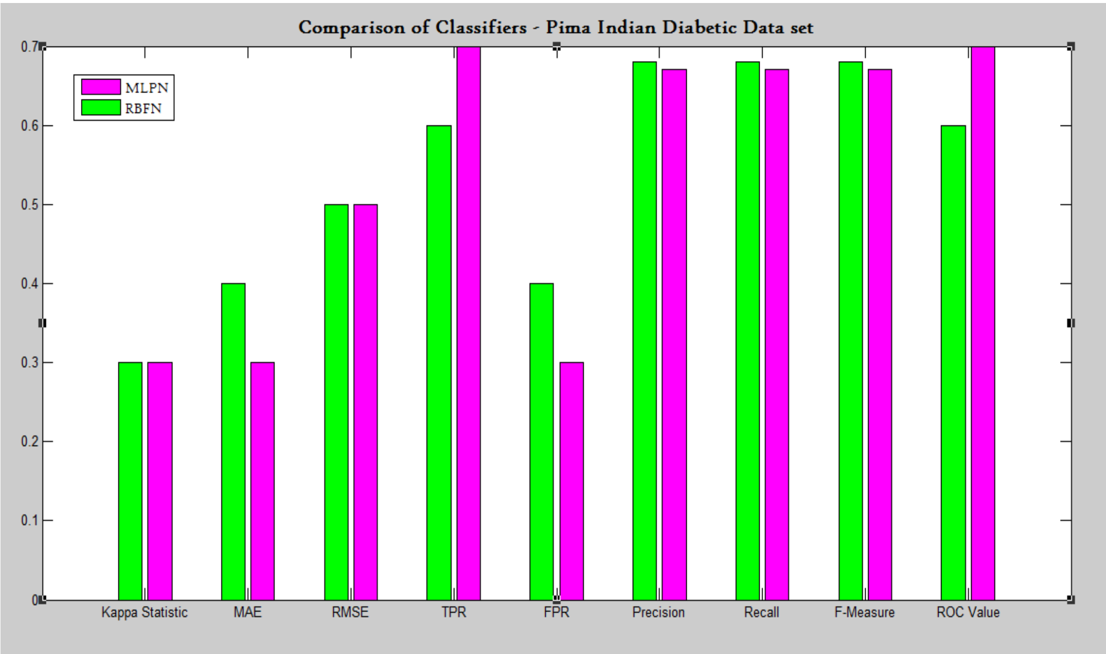


Figure 4: Comparison of Classifiers on PID

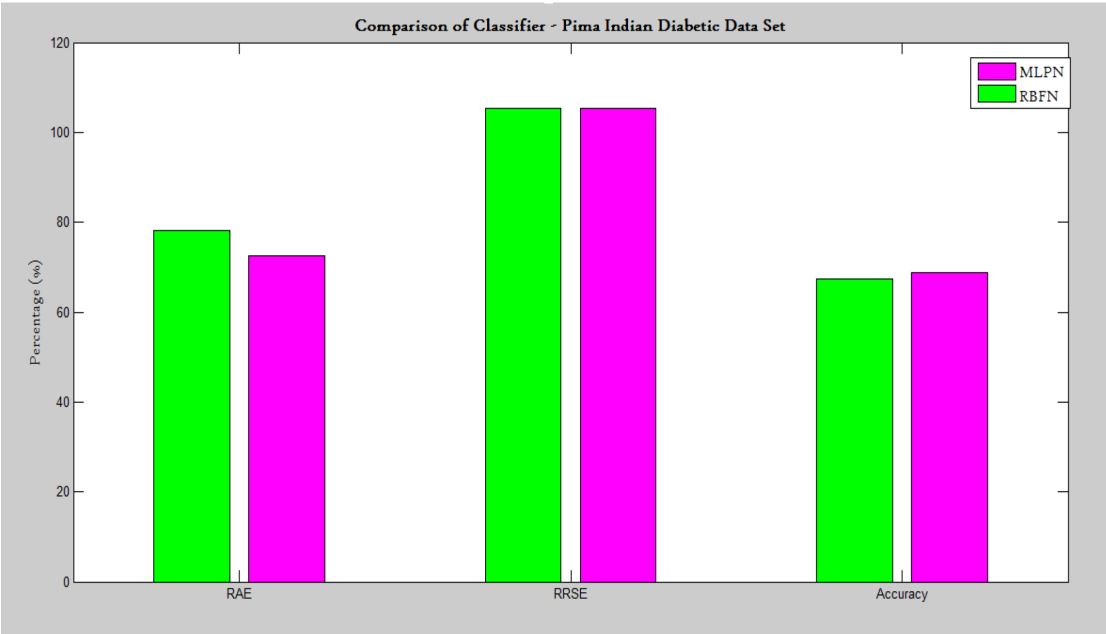


Figure 5: Comparison of Classifiers on PID

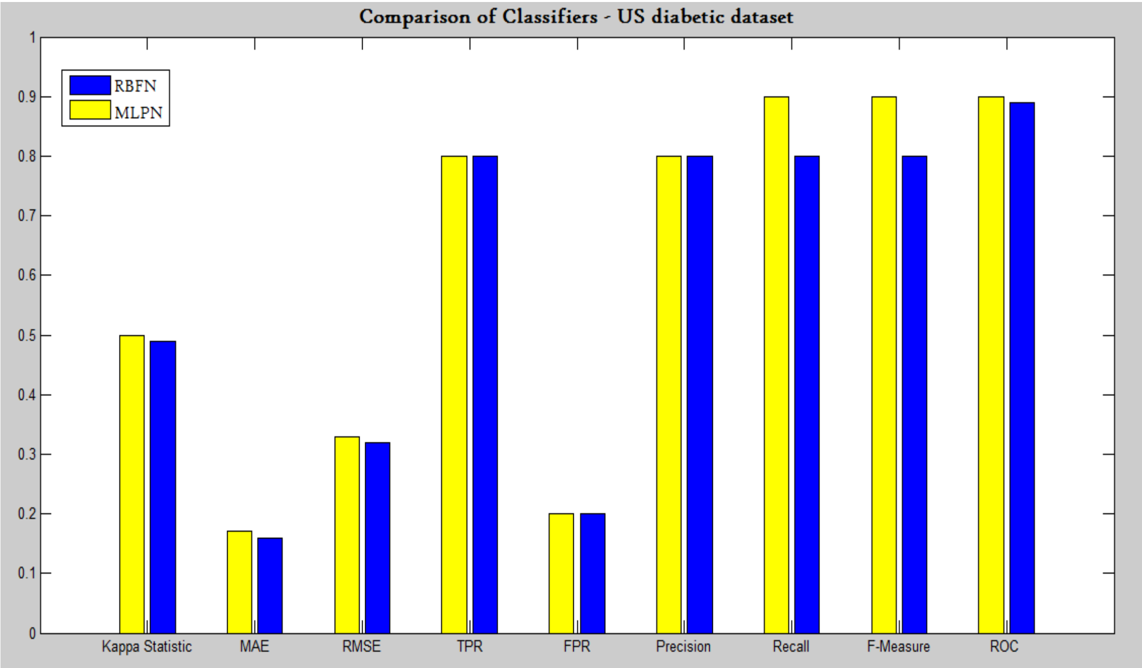


Figure 6: Comparison of Classifiers on US Dataset

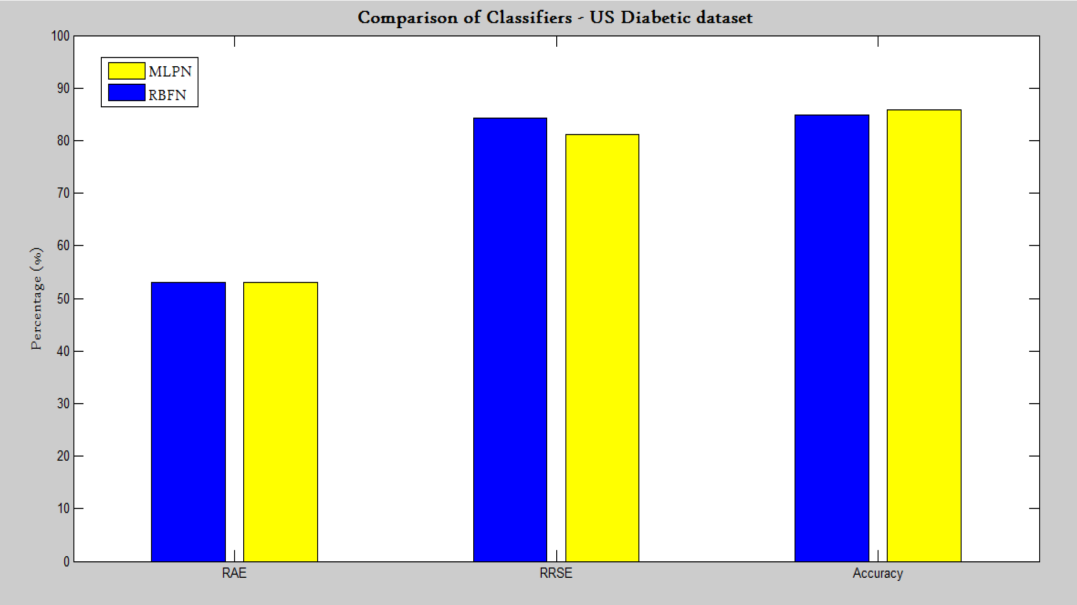


Figure 7: Comparison of Classifiers on US Dataset

***The novel hybrid Modified Particle Swam Optimization – Neural Network (MPSO-NN) Algorithm*:**

Input*:* *Reduced Data set*

Algorithm:

Step 1: *Initialize the population*

Step 2: *Evaluate the fitness of the attribute*

Step 3: *For each attribute, find the maximum fitness and compare it with the best found so far*

Step 4: *Pbest$_i$ is equal to the location of maximum fitness function*

Step 5: *Compare Fitness evaluation with population overall Pbest*

Step 6: *If particle best is greater than g_{best}, then reset $g_{best(i)}$ is equal to the current Pbest's array index and value*

Step 7: *Calculate Convergence factor*

Step 8: *Calculate Inertia weight*

Step 9: *Update velocity and position and new population is generated*

Step 10: *Adjust acceleration*

Step 11: *If the data set is optimized, Go to Step 12*
Else

Go to Step 2.

Step 12: *Choose the initial weight*

Step 13: *If the error is minimum*
Go to step 21
Else
Go to Step 14

Step 14: *Apply the optimized dataset to the network*

Step 15: *Calculate output for every neuron through hidden layer(s) to output layer*

Step 16: *Calculate Error value at the output layer*

Step 17: *Update the weight and bias at the output layer*

Step 18: *Calculate the Error value at the hidden layer*

Step 19: *Update the weight and bias at the hidden layer*

Step 20: *Check if the maximum number of epochs reaches*
If yes Go to Step 21
Else
Go to Step 14

Step 21 *Evaluate the network performance*

Step 22: *Classified output*

Output: *Classified diabetic data set*

At the preliminary stage, the population is initialized and the fitness of the population is calculated. The maximum fitness of each particles is compared with the best found so far as given in (1). The $P_{best,i}$ obtained will be equal to the location of the maximum fitness function. The fitness evaluation is then compared with the population's overall best as in (2). The convergence factor is calculated as

$$\lambda = \frac{2}{\left\| 2 - C - \sqrt{C^2 - 4C} \right\|} \qquad where\ C = C1 - C2 \qquad \text{.................. (4)}$$

C1 is the cognitive learning parameter and C2 is the social collaboration parameter. The C1 and C2 always lie between 0 and 2. Then the inertia value is calculated. The inertia value provides the balance between the exploration and exploitation. Generally the inertia value lies between 0.8 and 1.2.

$$\omega_i = \omega_{\max} - \frac{\omega_{\max} - \omega_{\min}}{iter_{\max}} \qquad \text{....................(5)}$$

The velocity and position of the particle is updated using the (6).

$$\begin{aligned} V_i &= \lambda(\omega_i V_i + c_1 r_1 (p_i - x_i) + c_2 r_2 (p_g - x_i)) \\ x_i &= x_i + V_i \end{aligned} \qquad \text{.........................(6)}$$

V_{id} is the momentum of the particle and r1, r2 is random numbers (0, 1). The acceleration of the particle is adjusted as

$$V_i = \begin{cases} V_{\max} \ if \ V_i > V_{\max} \\ -V_{\max} \ if \ V_i < V_{\max} \end{cases} \qquad \text{....................................(7)}$$

If the reduced data set is optimized, then the optimized data set is feed as input to Multi Layer Perceptron Network. Back propagation learning is used train the network. The output for every neuron is calculated in the hidden unit and output unit. It is calculated as

$$Z_{inj} = V_{oj} + \sum_{i=1}^{n} x_i v_{ij}$$

$$Z_j = f(Z_{inj}) \ where \ f(Z_{inj}) = \frac{1}{1 + e^{-\lambda Zinj}} \qquad \text{.. (8)}$$

The sigmoid activation function is applied in every layer. The number of neurons in the hidden layer is fixed by

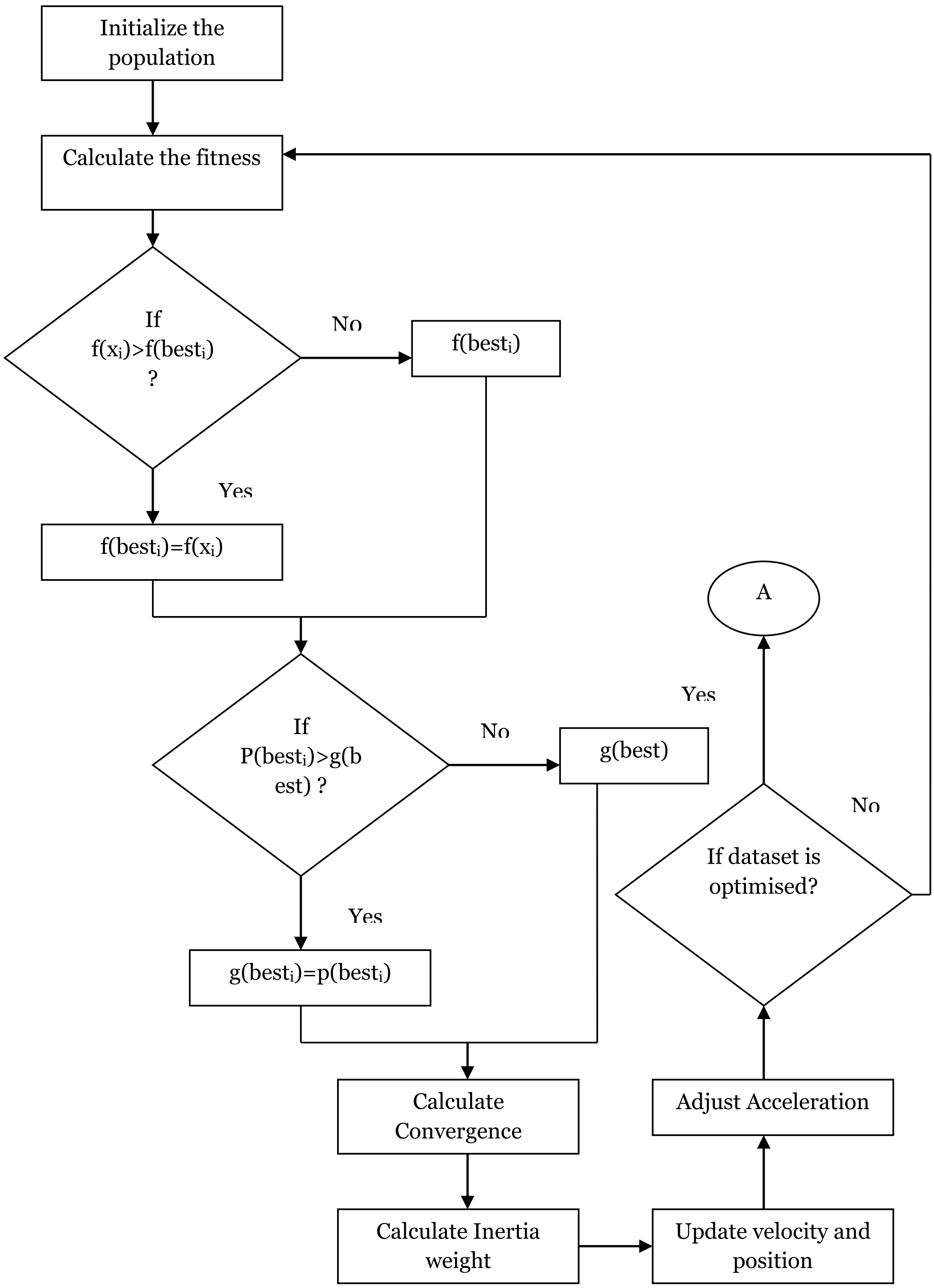
$$N_h = \sqrt{N_i N_o} \qquad \text{..(9)}$$

where N_h is the number of hidden neurons in the hidden layer. N_i and N_o represent the number of neurons in the input layer and output layer. The error at the output layer is calculated using (10) and the weight and bias are updated using (11).

$$\delta_k = (t_k - y_k) f'(y_{ink}) \ where \ f'(Y_{ink}) = \lambda f(Y_{ink})[-f(Y_{ink})] \qquad \text{.........................(10)}$$

$$\Delta W_{jk} = \alpha \delta_k z_j \ \Delta w_{ok} = \alpha \delta_k \qquad \text{........................(11)}$$

If the maximum number of epochs is reached and the error is minimized the training process is stopped, and the classified output and results are taken. The Figure 8 shows the framework of the proposed algorithm in classifying the diabetic data set.


Initialize the population
Calculate the fitness
If f(xᵢ)>f(bestᵢ) ?
No
f(bestᵢ)
Yes
f(bestᵢ)=f(xᵢ)
If P(bestᵢ)>g(best) ?
No
g(best)
Yes
g(bestᵢ)=p(bestᵢ)
A
Yes
No
If dataset is optimised?
Calculate Convergence
Adjust Acceleration
Calculate Inertia weight
Update velocity and position

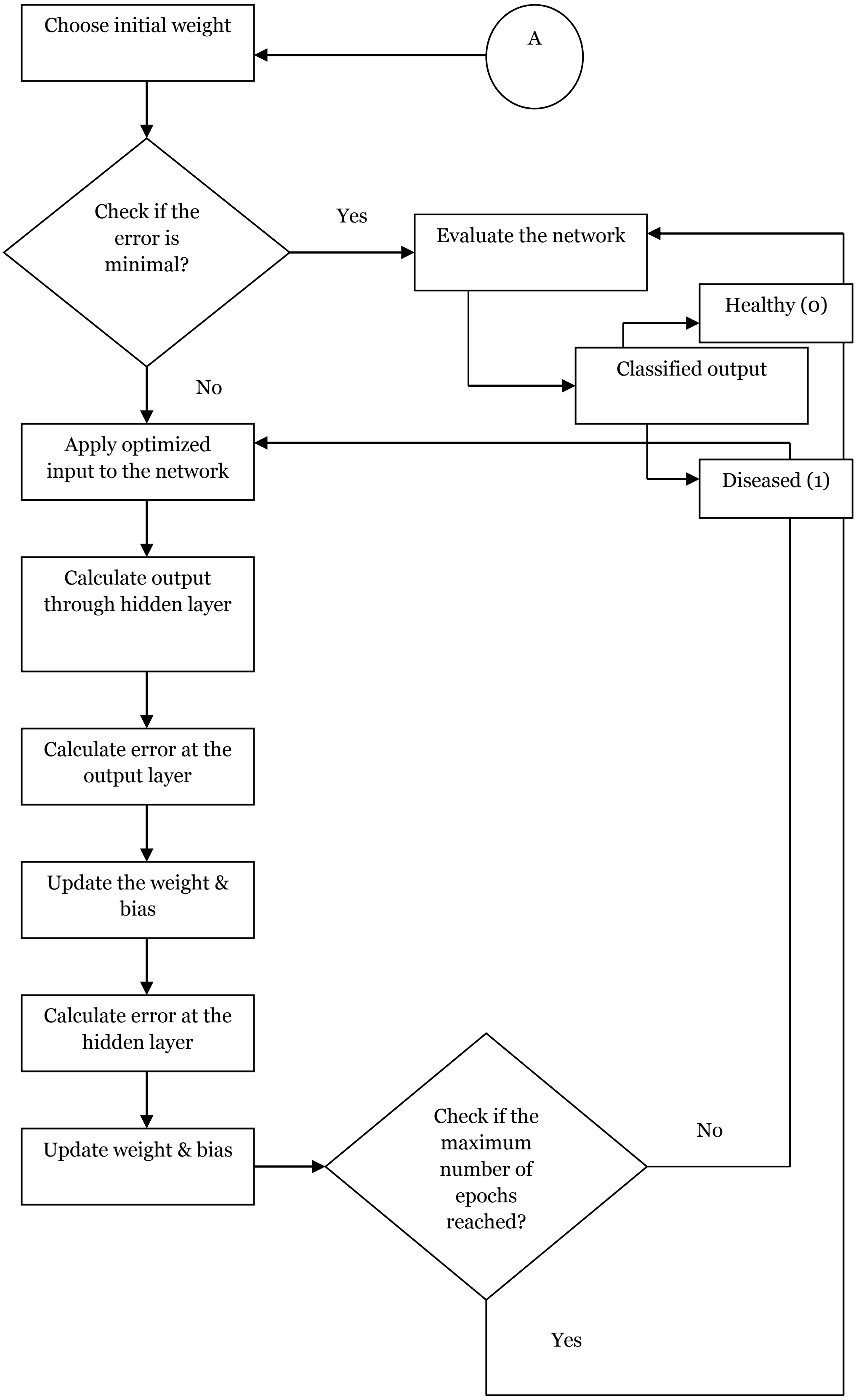


Figure 8: Working of the Proposed MPSO-NN Algorithm

RESULTS AND DISCUSSION

The performance of the proposed hybrid MPSO-NN Algorithm is measured with metrics like Mean Squared Error (MSE), Regression (R) and Percentage of Error (%E). MSE is the average squared difference between the input and output. Lower values of the MSE are the best one for classification. Regression is the correlation between output and target. 0 means a random relationship and 1 means a close relationship between output and target. The values obtained after experimentation is tabulated in Table 2 and Table 3. Figure 9 and 10 depicts the same results obtained.

Table 2: Performance of the hybrid MPSO-NN algorithm with PID and US data set

	MSE		**R**		**%E**	
	Pima Indian diabetic data	**US diabetic data**	**Pima Indian diabetic data**	**US diabetic data**	**Pima Indian diabetic data**	**US diabetic data**
Training	1.6	2.08	4.52	5.62	22.66	23.66
Testing	1.3	1.95	4.51	4.89	19.85	19.96

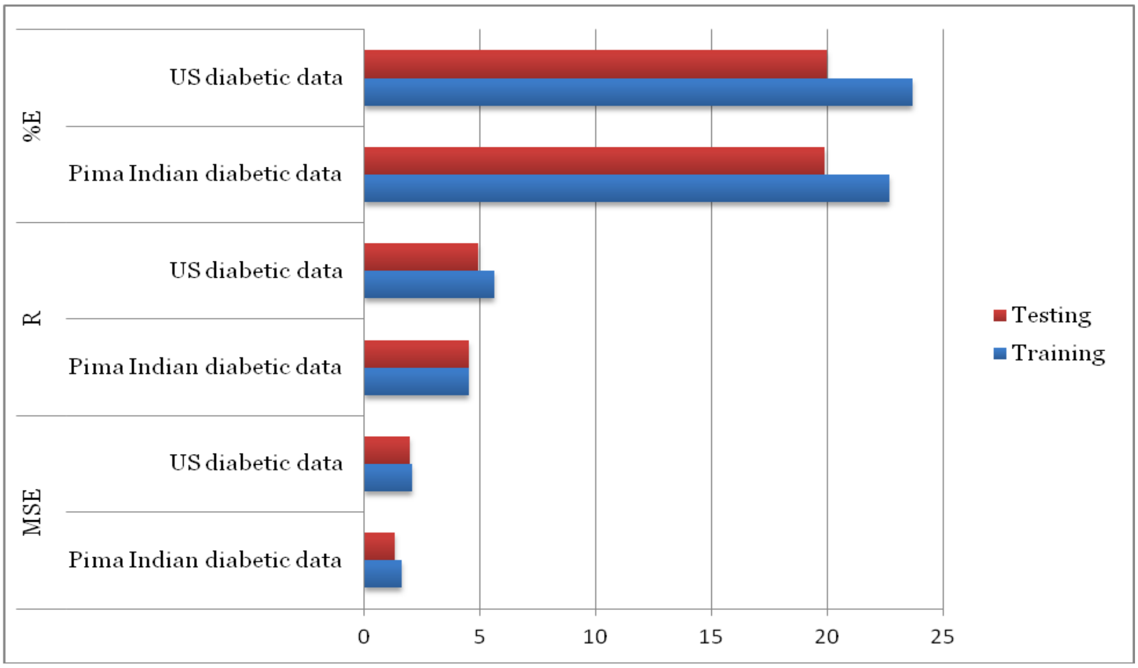


Figure 9: Performance of the hybrid MPSO-NN algorithm with PID and US data set

Table 3: Performance of the hybrid MPSO-NN algorithm with optimized PID and US data set

	MSE		**R**		**%E**	
	Pima Indian diabetic data	**US diabetic data**	**Pima Indian diabetic data**	**US diabetic data**	**Pima Indian diabetic data**	**US diabetic data**
Training	1.4	1.61	7.26	8.2	18.18	15.76
Testing	1.2	1.29	8.96	8.0	10.45	13.45

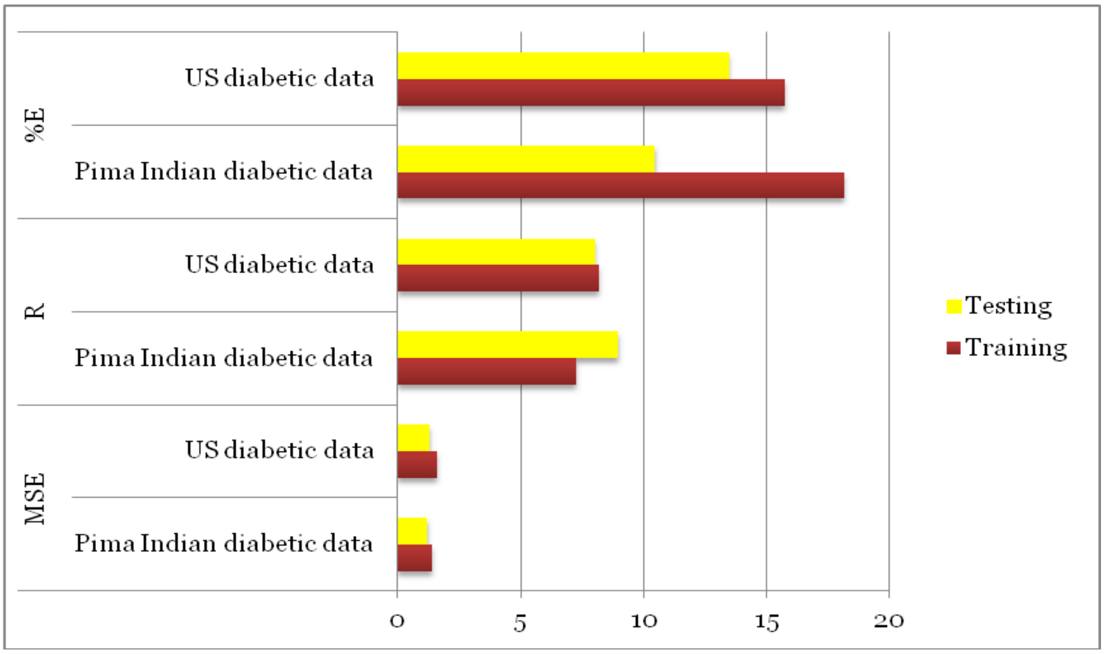


Figure 10: Performance of the hybrid MPSO-NN algorithm with optimized PID and US data set

The metrics like Sensitivity, Specificity, False Positive Rate (FPR) and Accuracy is used to evaluate the performance of the network with different number of neurons. Sensitivity is also called as True Positive Rate which measures the proportions of positivies that are correctly identified.

$$Sensitivity = \frac{TP}{(TP + FN)} \quad \text{......................... (12)}$$

Specificity is defined as the measure of proportions of negatives that are correctly identified. It is also denoted as True Negative Rate.

$$Specificit\,y = \frac{TN}{(TN + FP)} \quad \text{....................(13)}$$

False Positive Rate (FPR) (α) is defined as

$$FPR = \frac{FP}{(FP + TN)}$$
$$or$$
$$FPR = 1 - Specificity \qquad \text{.....................................(14)}$$

The classified accuracy is calculated as

$$Accuracy = \frac{(TP + TN)}{Total\ no.of\ ins\tan ces} \qquad \text{.................................(15)}$$

Where,
TP- True Positive
FN- False Negative
FP- False Positive
TN- True Negative

Figure 11 and 12 shows the Confusion Matrix obtained by the proposed MPSO-NN Algorithm. The Confusion Matrix shows the correctly classified instance obtained using the proposed algorithm.

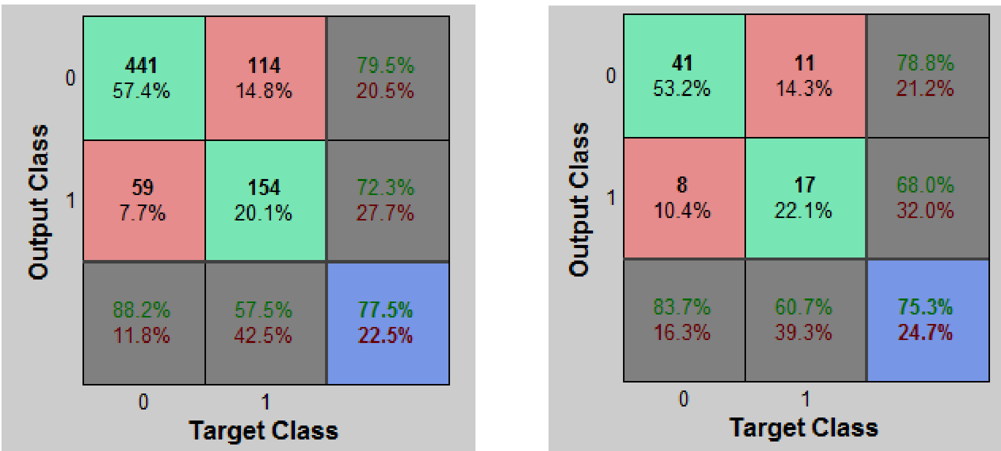
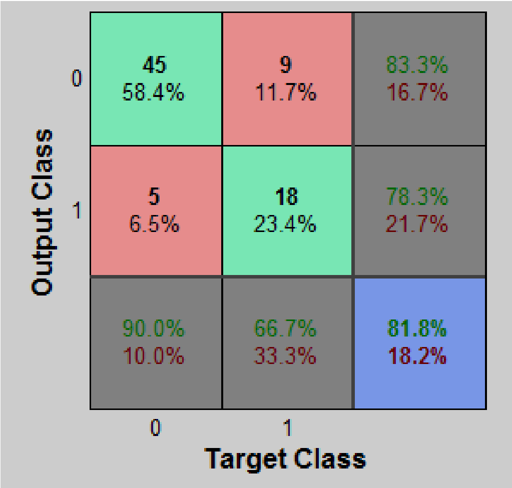


Figure 11: Confusion Matrix for PID and US data



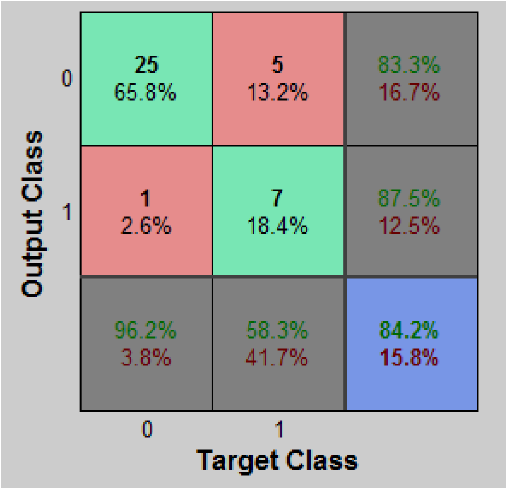


Figure 12: Confusion Matrix for optimized PID and US data

Table 4 shows the metric value obtained while experimenting the proposed Algorithm. The obtained values are represented using graphs in Figures 13 and 14.

Table 4: Performance Evaluation of Proposed MPSO-NN Algorithm

	Diabetic Data set		**Optimized Diabetic data set obtained from MPSO-NN Algorithm**	
	PID	**US Data**	**PID**	**US Data**
Sensitivity	0.79	0.78	0.83	0.83
Specificity	0.72	0.68	0.78	0.87
FPR	0.28	0.32	0.22	0.13
Accuracy	77.5	75.3	81.8	84.2

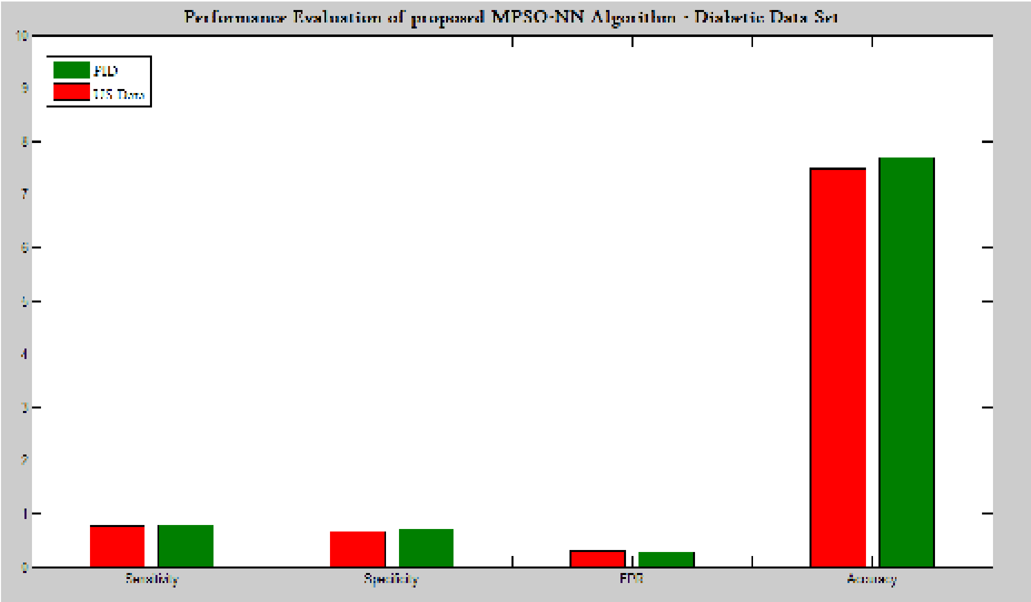


Figure 13: Performance Evaluation of proposed MPSO-NN Algorithm-Diabetic data set

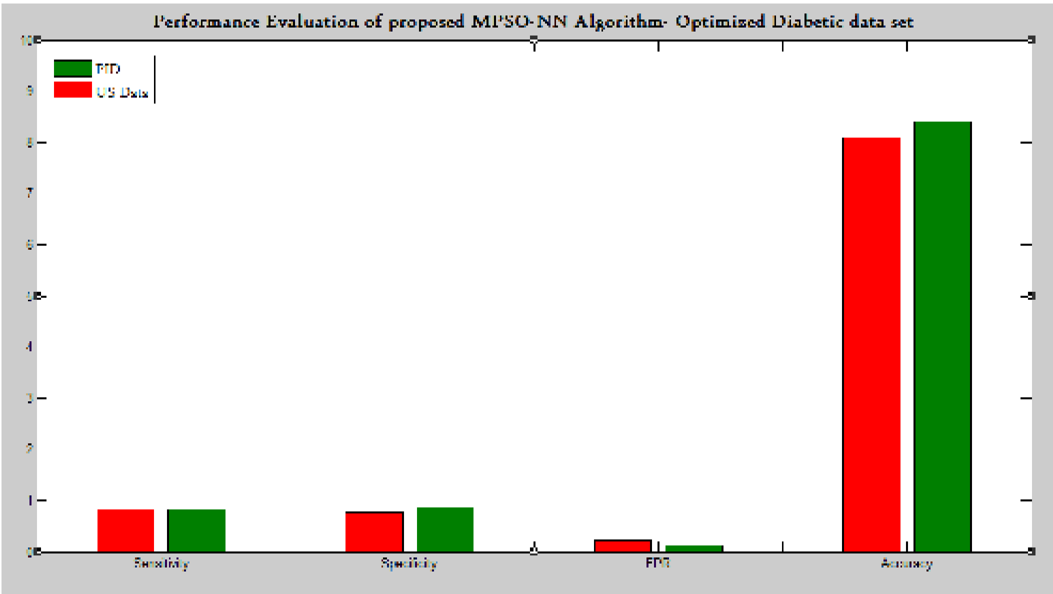


Figure 14: Performance Evaluation of proposed MPSO-NN Algorithm- optimized Diabetic data set

The accuracy was increased 81.8% for PID and 84.2 for US data set. The values obtained for other metrics also supports the proposed MPSO-NN algorithm in diabetic disease classification.

CONCLUSION

This paper presents a new hybrid classification algorithm called MPSO-NN, which was obtained by the modifying the Particle Swarm Optimization Algorithm and combining it with Multi Layer Perceptron Network. The results obtained have confirmed that the proposed novel hybrid MPSO-NN Algorithm is best fit for classifying the diabetic data set. The reduced data set obtained from Genetic – Relative Reduct Algorithm is applied into the proposed work. The reduced data set is optimized and classified according. The accuracy level was promising. However, there is still plenty of work going on in developing a best algorithm with more classification accuracy. The proposed hybrid algorithm is been applied only on data set with minimum instances. The future direction is to test the same algorithm with the increased number of instances and incorporating cloud computing in such a circumstance.

REFERENCES

[1] American diabetes association (2010). http://www.diabetes.org/diabetes-basics (last accessed February 2017)

[2] Temurtas H, Yumusak N, & Temurtas F, "A comparative study on diabetes disease diagnosis using neural networks", Expert Systems with Applications, Vol 36(4), pp 8610–8615, 2009.

[3] Mehrbakshsh Nilashi, Othman Bin Ibrahim, Abbas Mardani, Ali Ahani and Ahmad Josh, "A soft computing approach for diabetes disease classification", *Health Informatics Journal,* pp 1-15, 2016.

[4] Gan D, "Diabetes Atlas, 2nd Edition", Brussels: International Diabetes Federation, 2003.

[5] Nabib Singh Gill and Pooja Mittal, "A Computational Hybrid Model with two level classification using SVM and Neural Network for predicting the diabetes disease", Journal of Theoretical and Applied Information Technology*, Vol* 87(1), pp 1-10, 2016.

[6] Xiangyang Wang, Jie Yang, Xialong Tens and Weijan Xia, Richard Jension, " Feature selection basedon Rough Set and Particle Swarm Optimization", Pattern Recognition Letters, Vol 28(4), pp 459-471, 2007.

[7] Sivagowry S, Dr.Durairaj M, "PSO - An Intellectual Technique for Feature Reduction on Heart Malady Anticipation Data", International Journal of Advanced Research in Computer Science and Software Engineering, Vol 4(9), pp 610-621, 2014.

[8] Sivagowry S, Dr. Durairaj M, "An Intelligent Hybrid Quick Reduct Particle Swarm Optimization Algorithm for Feature Reduction in Cardiac Disease Prediction", International Journal of Emerging Technologies in Computational and Applied Sciences, Vol 12(2), pp 163-173, 2015.

[9] Preethi Gupta, Putinam Bajaj, " Heart Disease Diagnosis based on Data Mining and Neural Network", International Journal of Engineering Science and Research Technology, Vol 3(6), pp 172-176, 2014.

[10] Punam Bajaj and Prethi Gupta, "Review on Heart Disease Diagnosis based on Data Mining Techniques", International Journal of Science and Research, Vol 3(5), pp 1593-1596, 2014

[11] Mostafa Fathi Ganji and Mohammad Saniee Abadeh, "A fuzzy classification system based on Ant Colony Optimization for diabetes disease diagnosis", Expert Systems with Applications, Vol 38, pp 14650-14659, 2011.

[12] Peng L, Chen Y, Yang B, & Chen, "A novel classification method based on data gravitation. Neural Networks and Brain", Proceedings of International Conference on Neural Network and Brain, Beijing, China, pp 809-819, 2005.

[13] Polat K, Günes S, & Arslan, A, "A cascade learning system for classification of diabetes disease: Generalized discriminate analysis and least square support vector machine", Expert Systems with Applications, Vol 34, 482–487, 2008.

[14] Leon W. D, "Enhancing pattern classification with relational fuzzy neural networks and square BK-products", Ph.D dissertation in computer science. FL, USA: Springer, pp. 71–74, 2006.

[15] Jaganathan P, Thangavel K, Pethalakshmi A, & Karnan M, "Classification rule discovery with ant colony optimization and improved quick reduct algorithm", IAENG International Journal of Computer Science, Vol 33(1), pp 125-131, 2007.

[16] Kahramanli H, & Allahverdi N, "Design of a hybrid system for the diabetes and heart diseases", Expert Systems with Applications, Vol 35, 82–89, 2008.

[17] Kayaer K and Yildirim T, "Medical diagnosis on Pima Indian diabetes using general regression neural networks", Proceedings of the international conference on artificial neural networks and neural information processing, Istanbul, pp 181-184, 2003.

[18] Aslam MW, Zhu Z and Nandi AK, "Feature generation using genetic programming with comparative partner selection for diabetes classification", Expert System and Application, Vol 40(13), pp5402–5412, 2013.

[19] Erkaymaz O and Ozer M, "Impact of small-world network topology on the conventional artificial neural network for the diagnosis of diabetes", Chaos Soliton Fractals, Elsevier,Vol 83, 178–185, 2016.

[20] Dogantekin E, Dogantekin A, Avci D, "An intelligent diagnosis system for diabetes on linear discriminant analysis and adaptive network based fuzzy inference system", Proceedings of Digit Signal Processing, Vol 20(4), pp 1248–1255, 2010.

[21] UCI Website https://archive.ics.uci.edu/ml/datasets/Diabetes (last accessed November 2016).