Forecast of Diabetes using Modified Radial basis Functional Neural Networks

G.Magudeeswaran
Research scholar of Computer science
Sree saraswathi Thyagaraja college
Pollachi-642 107, Tamil Nadu

D.Suganyadevi, Ph.D.
Director of MCA
Sree saraswathi Thyagaraja College
Pollachi-642 107, Tamil Nadu

ABSTRACT

The paper entitled "Prediction of Diabetes using Modified Radial basis Functional Neural Networks" is used to predict the diabetes for the patients. Nowadays Data Mining techniques are used to predict the diseases of health care industry. This technique is to find out the information which is hidden in the dataset. Modified Radial basis Functional Neural Networks is the Data Mining technique used to predict the diabetes disease, Modified Radial basis Functional Neural Networks is a Data Mining technique based classification model as one of the powerful method in intelligent field for classifying diabetic patients. This new modified method is used to predict the blood glucose level for the diabetes patients. The proposed approaches are evaluated by the Pima Indian Diabetes data sets, were the Pima Indian Diabetes data set is a data mining dataset. It is observed from the experimental results that the modified RBF obtained better results than the exiting RBF method and other neural network.

Key words:

Data Mining, Artificial Neural Network, Diabetes, MRBF, RBF.

I. INTORDUCTION

Diabetes is one of the most deadly, disabling, and costly diseases observed in many of the nations at present , and the disease continues to be on the rise at an alarming rate. Diabetes Mellitus (DM) is a chronic and progressive metabolic disorder according to the World Health Organization there are approximately million people in this world suffering from diabetes. Common manifestations of diabetes are characterized by insufficient insulin production by pancreas

General Symptoms of Diabetes:

- Excessive thirst
- Frequent urination
- Weight loss
- Blurred vision
- Increased hunger
- Frequent skin, bladder or gum infections
- Irritability
- Tingling or numbness in hands or feet
- Slow to heal wounds
- Extreme unexplained fatigue
- Sometimes there are no symptoms (type 2 diabetes

Diabetes mellitus occurs when the pancreas doesn't make enough or any of the hormone insulin, or when the insulin produced doesn't work effectively [1]. In diabetes, this causes the level of glucose in the blood to be too high.

Complications of Diabetes

- People with diabetes are two to four more times more likely to develop heart disease or have a stroke than those who don't have diabetes
- Diabetes is the leading cause of new blindness among adults between 20 and 74 years old.
- Diabetes is the leading cause of treated end-stage kidney disease.
- More than 60 percent of the limb amputations occur among people with diabetes.
- About 60-70 percent of the people with diabetes have mild to severe nerve damage.

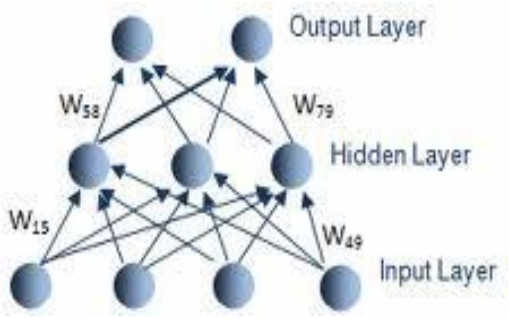

Fig1.1: Feed forward neural network.

In MLP, the weighted sum of the inputs and bias term are passed to activation level through a transfer function to produce the output, and the units are arranged in a layered feed-forward topology called Feed Forward Neural Network (FFNN). The schematic representation of FFNN with *n* inputs, *m* hidden units and one output unit along with the bias term of the input unit and hidden unit is given in Figure 1.1.

II. DATASET DESCRIPTION

In this work Pima Indian Diabetes data sets [11] has been taken for training and testing the neural network model.

2.1 PIMA INDIAN DIABETES Dataset:

Several constraints were placed on the selection of these instances from a larger database. In particular, all patients here are females at least 21 years old of Pima Indian heritage.

Attribute Information:

1. Number of times pregnant

2. Plasma glucose concentration a 2 hours in an oral glucose tolerance test

3. Diastolic blood pressure (mm Hg)

4. Triceps skin fold thickness (mm)

5. 2-Hour serum insulin (mu U/ml)

6. Body mass index (weight in kg/(height in m)^2)

7. Diabetes pedigree function

8. Age (years)

9. Class variable (0 or 1)

The Pima Indian diabetes database, donated by Vincent Sigillito, is a collection of medical diagnostic reports of 768 examples from a population living near Phoenix, Arizona, USA. The paper dealing with this data base uses an adaptive learning routine that generates and executes digital analogs of perceptron-like devices, called ADAP. They used 576 training instances and obtained a classification of 76% on the remaining 192 instances. The samples consist of examples with 8 attribute values and one of the two possible outcomes, namely whether the patient is *tested positive for diabetes* (indicated by output one) or not (indicated by two). The database now available in the repository has 512 examples in the training set and 256 examples in the test set.

III. FACTORS AFFECTING PERFORMANCE

The difficulty of neural networks is that there are a lot of parameters that affect the performance of the model [5]. Taking different values for one parameter might cause the optimal value of another parameter to change. Changing the value for the other parameter, the optimal value for the first may change yet again. For example, two neural networks with a different number of hidden nodes will very likely have different optimal values for the learning rate. The result is a difficult optimisation problem and likes to select the optimal values for all parameters in a systematic way. The high complexity of the problem limits us in this search for the neural network with the best parameters, because it is simply not feasible to try all possible combinations of parameter values. Instead we will fix a number of parameters in advance (with the necessary argumentation of course) and determine the optimal values of those that are left.

A lot of tweaking, tuning and experimenting had already been done and the general direction of the optimal parameters was clear to some extent [6, 7]. It may be assumed that the optimal input selection will not be dramatically different for the optimal neural network.

For the proposed work Pima Indian Diabetes data sets [2, 3] has been taken for training and testing the neural network model. This dataset is used to illustrate the effect of the topology (in terms of the number of bins per attribute) on the generalization ability of the proposed network. With a twelve fold cross validation and special pre-processing, the test result reported with the dataset is 77.7% using the LogDisc algorithm. The generalization obtained on network for the same dataset without any pre-processing. The first column indicates the number of bins used for each attribute and is followed by the classification success percentage for the training and test sets.

Performance of proposed model can be seen in the below figures

1. Modified Radial Basis Functional Neural Network

2. Regression Neural Network

Select your Choice: 1

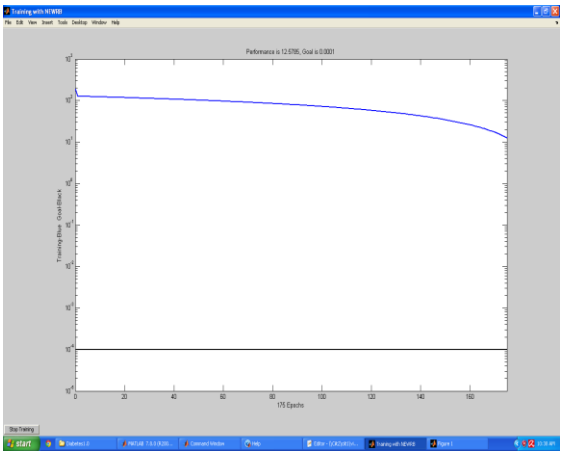

Figure 3.3 Performance of the dataset

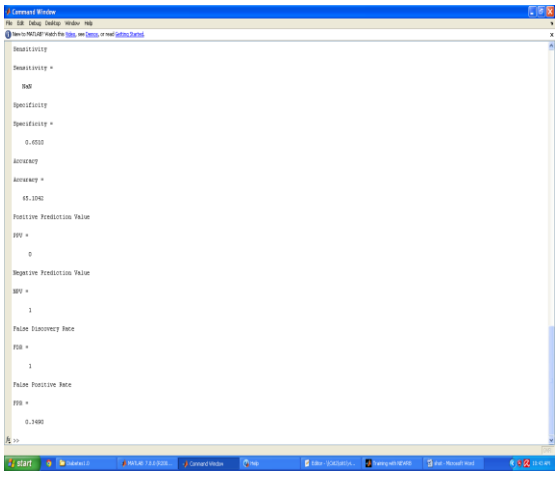

Figure 3.4: MRBFNN values for sensitivity, specificity, Accuracy is obtained.

1. Radial Basis Functional Neural Network

2. Regression Neural Network

Select your Choice: 2

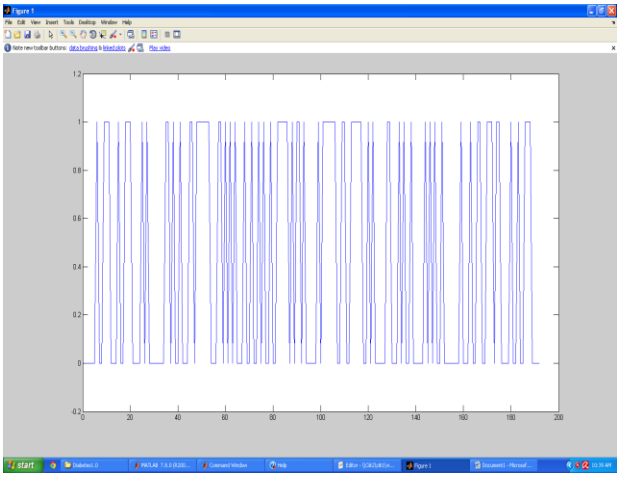

Figure 3.5 Performance of the dataset in regression

Neural Network

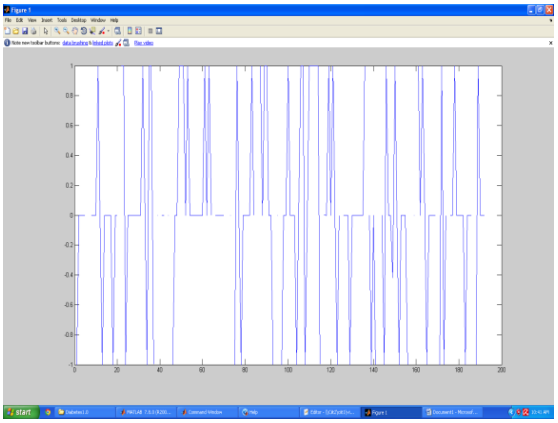

Figure 3.6: Performance of the dataset in regression

Neural Network

Figure 4.2.1: Comparison of the Performance of Various Models

From the experimental results is can be concluded that the modified RBF obtained a better results than the exiting RBF method and other neural network that are compared in this chapter[11, 12].

V. EXPERIMENTAL RESULTS

1. Modified Radial Basis Functional Neural Network

2. Regression Neural Network

Select your Choice: 1

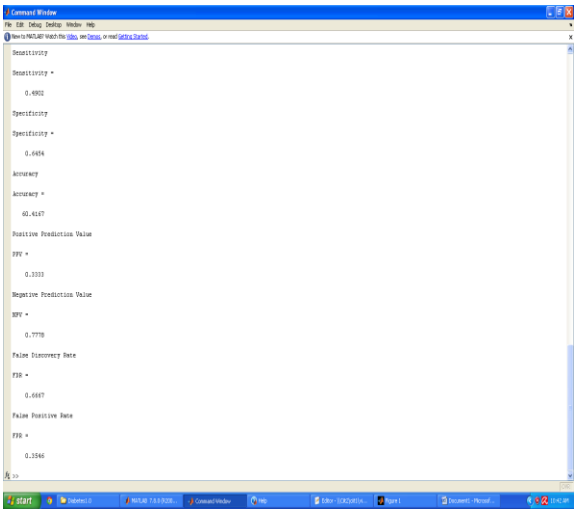

Figure 3.7: Regression values obtained for sensitivity, specificity, accuracy.

IV. COMPARISON WITH OTHER MODELS:

The results obtained for the Pima Indian Diabetes Database dataset were compared with the results where the performance of several models is presented: NN(nearest neighbor), kNN (k-nearest neighbor), BSS(nearest neighbor with backward sequential selection of feature, MFS1(multiple feature subset) , MFS2(multiple feature subset).Table 4.3 presents the results obtained by the various different models.

Table 4.3: Comparison of the average performance of several other classification systems

Models	Pima Indian Diabetes Dataset
KNN	70.1%
BSS	67.9%
NN	68.2%
MFS1	72.4%
MFS2	74.6%
RBF	75.7%
MRBF	78.8%

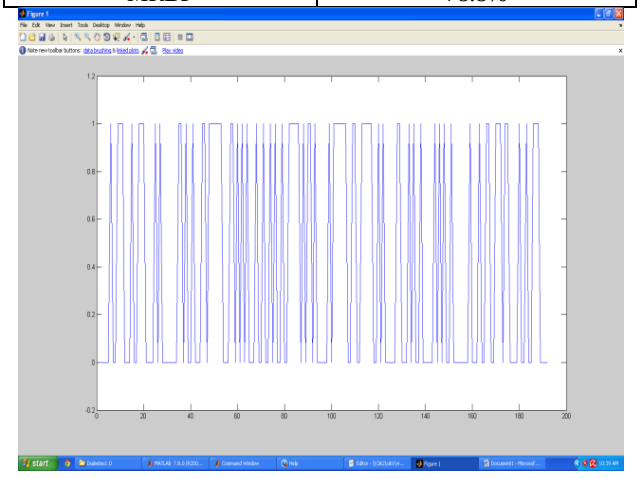

Figure 4.3.1: Performance of MRBFNN

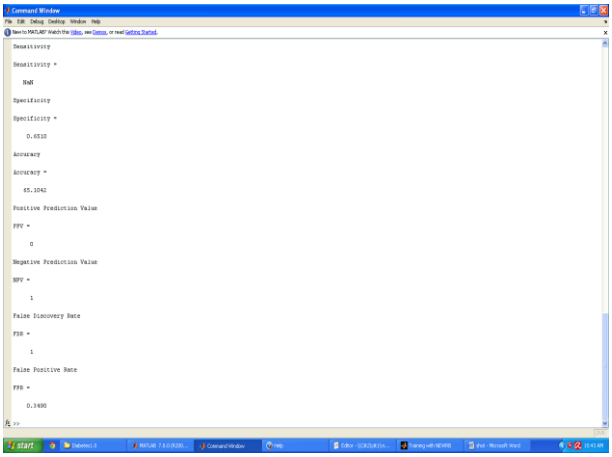

Figure 4.3.2: Sensitivity, Specificity, Accuracy for MRBFNN

1. Radial Basis Funtional Neural Network

2. Regression Neural Network

Select your Choice: 2

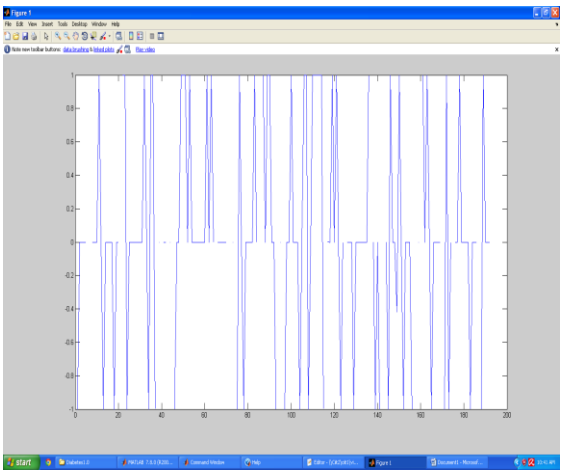

Figure 4.3.3: Performance of Regression NN

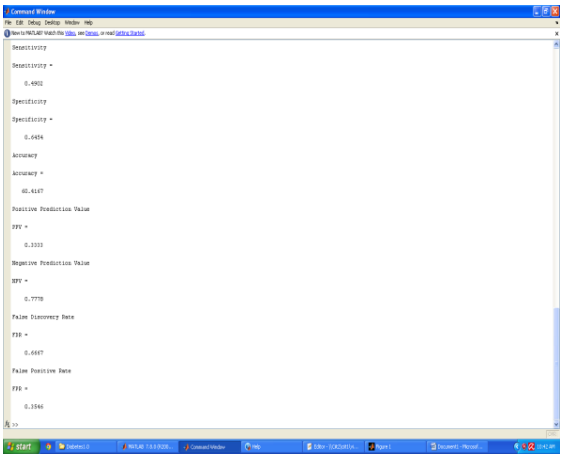

Figure 4.3.4: Sensitivity, Specificity, Accuracy for regression NN

This chapter deals with the experimental results of the proposed approach. The algorithm has been successfully applied in the prediction of diabetes, providing very competitive results. The performance was accurate in all of them, showing thus the robustness and adaptability of the proposed approach for time series of different nature.

6. Conclusion

In recent years several researches have been conducted to classify and show that who is diabetic or not. For Prediction of diabetic person several researchers used data mining techniques. Neural networks (NN) is a data mining technique is used to predict the diabetic person and also showed that neural networks obtained a better accuracy which was higher than other methods like logistic regression, feature selection, decision tree etc.

In this paper propose an Modified BRF neural network approach is a data mining approach which provides better and fastest result compare to the existing BRF neural network for predicting the diabetes patients. The MBRF is envisioned by using GA for optimally deciding the number of neurons in single hidden layer architecture. The proposed approaches are evaluated on the data mining dataset which is a PIMA INDIAN DIABETES Dataset. Again, this gives a very clear impression of the simplicity of the model without sacrificing at the cost of accuracy. This model proved to be better than other models and algorithms with which it was compared.

The performance of this model is remarkable in terms of processing time, which is treated as one of the crucial aspect in data mining. In proposed designed NN (nearest neighbor), kNN (k-nearest neighbor), BSS(nearest neighbor with backward sequential selection of feature, MFS1(multiple feature subset) , MFS2(multiple feature subset)and BRF in the dataset is used to compare with proposed models. Due to less computational cost at the hidden layer the MRBF can have applications in the real time domain.

Diabetes can be predicted easily with the help of proposed model and can be controlled in the early stages itself.

VI. REFERENCES

[1] World Health Organization. Available: http://www.who.int

[2] American Diabetes Asscociation. Available: http://www.diabetes.org

[3] Gan, D. editor. Diabetes atlas, 2nd ed. Brussels: International Diabetes Federation,2003.Available at http://www.eatlas.idf.org/webdata/docs/Atlas%202003-Summary.pdf2]

[4] Siti Farhanah Bt Jaafar and Dannawaty Mohd Ali, “Diabetes mellitus forecast using artificial neural networks”, Asian conference of paramedical research proceedings, 5-7, September, 2005, Kuala Lumpur, MALAYSIA

[5] T.Jayalakshmi and Dr.A.Santhakumaran, “A novel classification method for classification of diabetes mellitus using artificial neural networks”. 2010 International Conference on Data Storage and Data Engineering.

[6] Lehmann, E. D.; Deutsch, T., ‘Compartmental models for glycaemic prediction and decision support in clinical

diabetes care: promise and reality'. Computer Methods and Programs in Biomedicine. May 1998; 56(2): 193-204

[7] Trajanoski, Z.; Regittnig, W.; Wach, P., 'Simulation studies on neural predictive control of glucose using the subcutaneous route'. Computer Methods and Programs in Biomedicine. May 1998; 56(2): 133-9

[8] Eng Khaled Eskaf, Prof. Dr. Osama Badawi and Prof. Dr. Tim Ritchings," Predicting blood glucose levels in diabetes using feature extraction and artificial neural networks".

[9] [24]Gregory Hastings, Nejhdeh Ghevondian, "A selforganizing estimator for hypoglycemia monitoring in diabetic patients", 20 th annual international conference of IEEE engineering in medicine and biology society, Vol. 20, No 3, 1998\

[10] UCI machine learning repository and archive.ics.uci.edu/ml/datasets.html

[11] Andreassen, S.; Benn, J. J.; Hovorka, R.; Olesen, K. G.; Carson, E. R.,'A probabilistic approach to glucose prediction and insulin dose adjustment: description of metabolic model and pilot evaluation study'. Computer Methods and Programs in Biomedicine. Jan. 1994; 41(3-4): 153-65

[12] A.K. El-Jabali, "Neural network modeling and control of type I diabetes mellitus," Bioprocess Biosystem Engineering, vol. 27, pp. 75-79, 2005.

[13] S.A. Quchani, and E. Tahami, "Comparison of MLP and Elman neural network for blood glucose level prediction in type I diabetics," Proceedings of the 3rd International Federal of Medical and Biological Engineering, Kuala Lumpur, 2007, pp. 54-58.