

A Study of Psychosocial and Built Environment factors for Diabetes using MLP Neural Network

M. R. Narasinga Rao
Department of CSE
KLUniversity
Vaddeswaram,Guntur

P. Srinivasa Rao
Department of CS&SE
AndhraUniversity
Visakhapatnam

V. Ramakrishna
Department of CSE
KLUniversity
vaddeswaram,Guntur

ABSTRACT

Background: Diabetes Mellitus is a common metabolic disorder and which is prevalent among many people in India. Although there are many factors which influence the onset of diabetes, the relationship between built environment and psycho-social factors is one such factor which play a crucial role for the onset of diabetes. Hence, in this research, the authors try to make a comparative study between built environment and psychosocial factors by using multilayer feed-forward neural networks using back propagation and Modified Feed-Forward Neural Network Constructive (MFNNC) algorithm.

Material and Methods

In this research, we tried to compare the performance of two individual neural networks, one is an MLP with back-propagation algorithm and the other is also an MLP with a constructive algorithm known as multilayer feed forward neural network constructive (MFNNC) algorithm. For using an MLP with back-propagation algorithm prior knowledge of neural network architecture is necessary for proper training and testing the network. Usually, trial and error method is used to predict the exact architecture of the network. Hence, in this connection we have proposed a constructive algorithm called multilayer feed-forward neural network constructive algorithm (MFNNC) with which one can predict the exact architecture of the network along with predictive capability. The MFNNC algorithm starts with one hidden neuron in the hidden layer and gradually increases the hidden neurons in the hidden layer if the error generated by the network is more than the value given in the interactive mode or the error is the same for three consecutive examples. The performances of the MLP with back-propagation algorithm and the other with MFNNC algorithm were compared and tested.

Results

Males who are having lesser primary activity (PMA) have more food intake (FI) and more basic mobility (BMB) than their female counterparts. Males who are having more activity by health (ABH), have more professional activity (PRA), almost the same rest (RES) values, more food intake (FI) and have more basic mobility (BMB) compared to their female counter parts. Males who are having more professional activity (PRA) have lesser leisure (LEI) values, almost the same rest (RES) values, more food intake (FI), more basic mobility (BMB), when compared to their female counterparts.

Males who are having similar leisure (LEI) have the same rest (RES), when compared to their female counterparts.

Conclusion

We have developed two prototype models using neural networks one is an MLP with back propagation algorithm and other one is a feed forward neural network with MFNNC algorithm. We have observed that both the networks have performed well in assessing the parameters for the onset of diabetes by giving the inputs as biological and biographical variables and by considering error values of both of the networks. It is also observed that, using the MFNNC algorithm is better if one wants to know the architecture of the trained network rather than just using MLP with back propagation algorithm.

Keywords

MLP Neural Network, Back-propagation, modified feed-forward neural network constructive algorithm (MFNNCA), Primary Activity, Food Intake, Transport, Basic Mobility, Activity By Health, Professional Activity, Leisure, Rest.

1. INTRODUCTION

Diabetes Mellitus (DM) is an epidemic throughout the world and it is more prevalent in India. It is also rampant in India and has shown a worsening trend in the past few years⁽¹⁻³⁾. Studies have shown that in the developed countries, the built environment contributes to obesity and diabetes mellitus which indicates that the structure of the urban living conditions may be a highly cost-effective way to curb the rise of these ailments⁴. The outcome of the Built environment of a region is dependent on some complex interactions of its geographical setup, socio-cultural, economic, technological advancements, natural resources, and other factors. Thus it is imperative to identify elements of the built environment which impact the behavior of the people in a given region⁵. This research is about trying to find how the built environment variables and psychosocial variables play an important role for the onset of diabetes. There are very wide applications of neural networks in medical diagnosis⁶⁻¹⁰. Feed forward neural networks are widely used for medical diagnosis and the network connections are allowed only between the nodes in one layer and those in the other layer¹¹.

We have developed two types of prototype models, one is an MLP with back propagation algorithm and the other is a multilayer feed forward neural network with MFNNC (multilayer feed forward Neural Network Constructive) algorithm. One of the major drawbacks of traditional back propagation algorithm is that, we have to know the topology of the network prior to training. This takes lot of time for identifying the exact network architecture. To overcome this difficulty, a new constructive algorithm called MFNNC algorithm, that constructs a network dynamically have been proposed. This algorithm starts with a near minimal architecture starting with minimum number of hidden units¹⁰. Later, it constructs the network by adding nodes one after the other in the hidden layer based on the performance of the network on the training data¹¹.

2. NETWORK TOPOLOGY

The topology of the network is basically about the size of the feed forward network and which will be determined based on the number of neurons present in the input layer, hidden layer and output layer. The number of neurons present in the input layer depends on the number of input vectors and the number of neurons present in the output layer depends on the output vectors of the network. The neurons present in the hidden layer plays a very crucial role for making the network to learn the data and hence for the decision making of the network. The determination of exact number of neurons plays a vital role for the performance of the entire neural network. The number of neurons in the hidden layer is obtained based on the trial and error method in the case of a MLP neural network with back propagation algorithm whereas in a feed forward network, we have used MFNNC algorithm.

3. METHODS

3.1 Data Collection and variables

3.1.1 Study Area

In this research, we have collected data from a region called Muvvalavanipalem colony (MVP Colony) and which is located in the neighbourhood of visakhapatnam with Bay of Bengal on the East and Adavivaram Hill Range (Eastern Ghats) on the North. Most of the Visakhapatnam city is located to the South and West of the MVP Colony⁵.

3.1.2 Participant Selection:

The selection of individuals in the age group of 16-69 and who had lived in the area for atleast 3 years were selected in this research. These individuals were randomly selected households in all the residential plots of the two sectors.⁵

3.1.3 Procedure:

A Self report questionnaire is used for getting to know the daily activity of people living in the colony. Data on the energy consumed and energy utilized by the participants were obtained for the purposes of correlating them with Body Mass Index (BMI).⁵

3.1.4 Dependent Variable:

Body Mass Index was calculated by dividing bodyweight in kilograms by the square of the height in meters.⁵

3.1.5 Proxy Variables

The five proxy variables that were derived from the questionnaire are as follows. Nature of employment (home based or non-home based), professional activity (PRA), Shift Work, Sleep, and nature of diet (vegetarian or non-vegetarian).⁵

3.1.6 Nature of employment

The participants who have been considered for this research include home based (HB) or Non-Home Based (NHB). The latter included participants who had to travel some distance from home either for work or educational purposes.⁵

3.1.7 Professional activity

Professional activity was defined as the major activity a person engages in during the day—that is, the activity that takes up most of the person's time either at workplace or at home. The participants were divided into four categories based on their metabolic levels of activity as follows: Sedentary (90–150 KCAL/H); Moderately Active (151–300 KCAL/H); Active (301–480 KCAL/H); and Very Active (481–750 KCAL/H).⁵

3.1.8 Shift work

Participants were categorized as working in General Shift from 8.00/9.00 AM to 4.00/5.00 PM and those working in shifts with constant change of timings.⁵

3.1.9 Siesta

Afternoon naps were rated as 2 for less than 1 hour (considered healthy) and 1 for any other duration of time. Participants were divided into two categories—those who engaged in afternoon siestas and those who did not.⁵

3.1.10 Nature of diet

Participants were categorized as vegetarians or non-vegetarians.⁵

The descriptive statistics of the data which was used in this research has been given in Table1.

For a feed forward network with MFNNC algorithm, we have considered the learning rate as 0.5, number of iterations as 500, maximum error as 0.01, and number of hidden layers is 1 and the number of hidden units considered is 4. The topology 2-4-1 is considered for both the above said networks.

3.2 Steps of the Modified Feed Forward NN Constructive Algorithm for a feed forward network:

To build and train the network the following step are followed.

1.Create an initial neural network with number of hidden unit $h=1$. Set all the initial weights of the network randomly with in a certain range. Here, the range considered is between (-0.5 to +0.5).¹¹

2. Train the network on training set by using a training algorithm(Back Propagation algorithm) for a certain number of epochs that minimizes the error.¹¹

3. If the error function on validation set is acceptable (i.e., if the error is below the desired value which is set initially for constructing the network) or converge(the error is constant for newly created consecutive hidden nodes), then stop.

4. Add one hidden unit to hidden layer. Randomly initialize the synaptic weights of the arcs connecting this new hidden unit with input node and out put units. Set $h=h+1$ and go to step2¹¹. The error function is usually defined as the mean squared error

$$\xi(n) = \frac{1}{2} \sum e_k^2(n) \dots \dots \dots (1)^{11}$$

$$\xi_{av} = \frac{1}{N} \sum_{n=1}^N \xi(n) \dots \dots \dots (2)^{11}$$

where k denotes the k th output unit, n denotes the n th iteration, N is the total number of patterns, d_k denotes the desired output for the k th neuron, y_k denotes the actual output of the k th neuron and e_k denotes the error term for the k th output unit.

3.3 Methodology of MLP

The MLP network is constructed with back propagation algorithm with multiple hidden layers. The summary of the operation of the MLP with back propagation algorithm is given below.¹²

The operation of the typical MLP with back propagation algorithm is as follows.¹²

The operation of the typical back propagation network occurs as follows.

1. After presenting input data to the input layer , information propagates through the network to the output layer (forward propagation). During this time input and output states for each neuron will be set.¹²

Table1: The Descriptive statistics of the data

Variable	N	Minimum	Maximum	Mean	SD
Input					
Primary Activity	Male(120)	0.50	7.50	2.02	1.66
	Female(60)	1.00	13.00	3.46	2.80
Activity By Health	Male(120)	0.00	8.00	1.15	1.34
	Female(60)	0.00	6.00	0.89	1.50
Professional Activity	Male(120)	8.00	32.00	20.80	5.15
	Female(60)	16.00	24.00	18.97	3.90
Leisure	Male(120)	0.00	2.75	0.60	0.73
	Female(60)	0.00	3.00	0.58	0.90
Rest	Male(120)	4.00	8.00	5.88	0.71
	Female(60)	2.00	8.00	5.88	0.90
Transportation	Male(120)	0.00	116.07	20.34	20.55
	Female(60)	0.00	100.73	19.72	22.10
Food Intake	Male(120)	3.14	18.29	10.12	3.79
	Female(60)	4.14	18.29	10.05	3.80
Basic Mobility	Male(120)	0.00	153.00	8.31	17.18
	Female(60)	0.00	31.00	6.27	7.9
Output					
BMI	Male(120)	17.3	30.1	25.0	3.0
	Female(60)	19.2	30.5	25.7	2.9
Weight	Male(120)	49.00	85.00	66.81	8.58
	Female(60)	48.00	82.00	67.25	8.27

$$x_j^{[s]} = f(I_j^{[s]}) = f(\sum (w_{ij}^{[s]} * x_i^{[s-1]}))^{13}$$

Where $x_j^{[s]}$ denotes the current input state of the j th neuron in the current $[s]$ layer. $I_j^{[s]}$ Denotes the weighted sum of inputs to the j th neuron in the current layer $[s]$. f is conventionally the sigmoid function. $w_{ij}^{[s]}$ denotes the connection weight between the i th neuron in the current layer $[s]$ and j th neuron in the previous layer $[s-1]$.¹²

2. Global error is generated based on the summed difference of required and calculated output values of

each neuron in the output layer. The Normalized System error $E(\text{glob})$ is given by the equation $E(\text{glob}) = 0.5 * \sum (r_k - o_k)^2$ and $(r_k - o_k)$ denotes the difference of required and calculated output values.¹²

- Global error is back propagated through the network to calculate local error values and delta weights for each neuron. Delta weights are modified according to the delta rule that strictly controls the continuous decrease of synaptic strength of those neurons that are mainly responsible for the global error. In this manner the regular decrease of global error can be assured.¹²

$$E_j^{[s]} = x_j^{[s]} * (1.0 - x_j^{[s]}) * \sum (e_k^{[s+1]} * w_{kj}^{[s+1]})^{12}$$

Where $E_j^{[s]}$ is the scaled local error of the j th neuron in the current layer $[s]$ layer.

$$\Delta w_{ji}^{[s]} = \text{lcoef} * e_j^{[s]} * x_i^{[s-1]} \quad 12$$

Where $\Delta w_{ji}^{[s]}$ Denotes the delta weight of the connection between the current neuron and the joining neuron. Here, lcoef denotes the learning coefficient/ learning constant of the training parameters.¹²

- Synaptic weights are updated by adding delta weights to the current weights.¹²

Neural network simulate neurotransmission by changing the strength of inter neural connections.¹² Positive synaptic weights provide amplified neural signal and stronger effect to the joining neuron.¹² No modification in the information flow is modeled by zero weight. Negative weights mean inhibition.¹²

The learning process of a neural network is similar to the learning function of the human brain¹². The learning takes place by providing data for both inputs and outputs¹³. The calculated output value is compared to the required value that is also given in the training set¹³. Depending on the difference between the required and the calculated output values, the network adjusts the synaptic weights whose distribution constitutes the basis of the problem-solving algorithm¹². The network processes the elements of the training set in cyclical order until the difference becomes lower than a given value.

In the second part of the training process, the system is tested. The test set is fundamentally similar to the training set, but it contains different data¹³. If testing fails network structure or the learning parameters are then modified¹².

3.4 Assigning the synaptic weights for both of the networks:

Initially, the synaptic weights between input layer to hidden layer and from hidden layer to output layer have been taken in the form of an array. All the weights have been taken between -0.5 to +0.5. The weights have been assigned to the connections by using the random function called `random()`.

These values are assigned initially small random values and later on gets modified as the training phase of the network progresses. The neural network assigns the synaptic weights to the connections every time we provide an example to the network. The crucial part of the neural network is that, the knowledge acquired by the network through training is stored in the form of synaptic weights. Once the training phase is complete, the synaptic weights are stored in a file. These synaptic weights represent the knowledge acquired by the network through the environment

3.5 Format for presenting the input to the network for MLP Network with BP algorithm:

The input to the network is the data pertaining to built environment and psychosocial factors of a region named MLP Colony in Visakhapatnam. The data has been collected from a cohort of 180 individuals of which 120 are males and 60 are females. The data consist of values related to 14 variables namely Age, Sex, BMI, primary activity, Activity By Health, leisure, professional Activity, Rest, Basic Mobility, transport, food intake, weight, public transport and private transport respectively. The network variable bias, which has always been set to the value +1 is given as input to the MLP network whereas this is fixed for MFNN Network. The input variables considered are primary activity, activity by health, professional activity, leisure, rest, transportation, food intake and basic mobility whereas the output variables considered are BMI and Weight. Among these input variables, which in turn, we have predicted the relationship between the variables in between as well. Actual values of these variables been considered for both input and output during training and testing phases of the network.

3.6 Training of MLP and MFCNN network models:

In the case of MFCNN network, A training set of inputs and outputs were given to train both of the networks. Initially, the weights were initialized with a set of random numbers generated through random function for both of the networks. The weights were calculated from input layer to hidden layer and from hidden layer to output layer. These weights were adjusted (back propagation) from the required output values for both the networks. A single hidden layer is considered for both of the networks. Before the MFCNN network is tested, this network is trained with certain examples and with certain epochs. After, sufficient training has been performed with certain examples, the network is validated. Here, the test set is fundamentally different from the validation set. In this validation, different samples are given as inputs to the network, and find the error for each of the samples. If the error obtained is more than the error we have set previously during the interactive session, a new node will be created in the hidden layer dynamically and again the network is validated. If the error for three successive samples remains constant or the error generated by the network is less than the error we have set initially during the interactive session, then validation is terminated and go for testing of the network. We observe the architecture of the network. During the interactive session, we give values to various parameters for network like number of samples, learning rate, number of iterations, maximum error allowed, etc.,

Similar training has been performed for MLP network with back propagation algorithm. In this model, the architecture is

fundamentally similar to MFCNN network i.e., The number of neurons in the hidden layer are same. In this model also, there will be an interactive session between user and the network in which a number of network parameters like number of input samples, number of output samples, number of hidden neurons, learning rate, maximum individual error, maximum total error, number of iterations etc., need to be given as input to the network. The transfer function implemented was a sigmoidal transfer function. Once the training was performed with different samples, the network was tested by providing the input samples. Based on the input samples, error (Normalized System Error) was calculated and this error is compared with the error of MFCNN network. It is observed that, Both the networks have shown similar variations in the error when built environment, psychosocial variables were taken as inputs.

3.7 Projected Outcomes in Neural Network

An extensive search revealed that no valid tools were available to assess the variables of built environment and psychosocial aspects of diabetes with a reasonable accuracy. The models developed in this research was a first step in determining these factors with a reasonable accuracy.

These are prototype models we had developed in assessing the variables of built environment and psychosocial aspects of diabetes with a reasonable accuracy.

3.8 Resource Requirements

The resources that were required to complete this research involved usage of the Visual Studio software using higher level programming language VC++ under windows XP environment for MLP Network and Java language under windows XP environment for MFCNN network for building the network.

4. RESULTS

The data is obtained from a questionnaire and contains 13 elements viz...primary activity, activity by health, professional activity, leisure, rest, transportation, food intake, basic mobility, BMI and weight. The input variables considered are primary activity, activity by health, professional activity, leisure, rest, transportation, food intake, basic mobility where as the output variables are BMI and weight. The statistics of the data collected is given in the above table1.

4.1 Result Analysis

The following is the result obtained from both of the networks i.e., MLP with back propagation algorithm and feed forward network with MFNNC algorithm. The results are depicted in tabular format. The results are compared based on the error generated from both of the networks. **For MLPNN** with back propagation algorithm, we have considered momentum rate as

0.9, learning rate as 0.5, maximum total error as 0.01, maximum individual error as 0.001, number of iterations as 500, number of hidden layers is 1, and the number of hidden units considered are 4.

4.1.1 PMA VS FI

Table 2 :Primary Activity Vs Food Intake

	MEAN PMA	SD PMA	No of samples for MLP	No of samples for MFNNCA	Error for MLP	Error for MFNNC Alg	T - value
males	2.02	1.66	90	90	49.68	2.17	13.49
			120	120	49.51	2.17	
fe ma les	3.46	2.8	40	40	45.74	2.017	9.65
			60	60	47.89	2.017	

Inference: males who are having less PMA have more FI compared to females and which is evident from the both the tables.

4.1.2 PMA vs BMB

Table 3: primary Activity Vs Basic Mobility

	Mean PMA	SD PMA	No of samples for MLP	No of samples for MFNNCA	Error for MLP	Error for MFNNCA	T -value
males	2.02	1.66	90	90	89.65	16.95	13.49
			120	120	177.70	16.95	
fe ma les	3.46	2.8	40	40	37.03	2.31	9.65
			60	60	44.66	2.31	

Inference: males who are having lesser PMA values have more BMB w.r.t their female counterparts and which is evident from the table.

4.1.3 ABH VS PRA

Table 4: Activity By Health Vs Professional Activity

	Mean ABH	SD ABH	No of samples for MLP	No of samples for MFNN CA	Error for MLP	Error for MFNN CA	T-value
males	1.15	0.73	90	90	196.18	9.89	9.50
			120	120	212.23	9.89	
females	0.89	1.50	40	40	174.10	6.11	4.69
			60	60	168.76	6.11	

Inference: males who are having more ABH have more PRA compared to their female counterparts, and which is evident from the both of the networks.

4.1.4 ABH vs RES

Table 5: Activity By Health Vs Rest

	Mean ABH	SD ABH	No of samples for MLP	No of samples for MFNN CA	Error for MLP	Error for MFNN CA	T-value
males	1.15	0.73	90	90	12.17	0.46	9.50
			120	120	12.13	0.46	
females	0.89	1.50	40	40	12.30	0.47	4.69
			60	60	12.32	0.47	

Inference: Males who are having more ABH have almost the same RES values as w.r.t female counterparts.

4.1.5 ABH vs FI

Table 6: Activity By Health Vs Food Intake

	Mean ABH	SD ABH	No of samples for MLP	No of samples for MFNN CA	Error for MLP	Error for MFNN CA	T-value
males	1.15	0.73	90	90	46.68	2.172	9.50
			120	120	49.51	2.172	
females	0.89	1.50	40	40	45.14	2.0171	4.69
			60	60	47.89	2.0171	

Inference: Males who are having more ABH have more FI compared to female counterparts and which is evident from the both of the networks.

4.1.6 ABH vs BMB

Table 7: Activity By Health Vs Basic Mobility

	Mean ABH	SD ABH	No of samples for MLP	No of samples for MFNN CA	Error for MLP	Error for MFNN CA	T-value
males	1.15	0.73	90	90	89.65	16.987	9.50
			120	120	177.10	16.987	
females	0.89	1.50	40	40	37.03	2.31	4.69
			60	60	44.66	2.31	

Inference: males who are having more ABH values have more BMB values compared to the female counterparts.

4.1.7 PRA VS LEI:

Table 8 : Professional Activity Vs Leisure

	Me an PRA	SD PRA	No of sampl es for MLP	No of sampl es for MFNN CA	Erro r for MLP	Error for MFNN CA	T - value
ma les	20. 80	5.15	90	90	0.30 5	0.010 3	44.83
			120	120	0.28 1	0.010 3	
fe ma les	18. 97	3.9	40	40	0.46 5	0.013	37.96
			60	60	0.45 2	0.013	

Inference: Males who are having more PRA have lesser LEI values when compared to their female counter parts.

4.1.8 PRA VS RES

Table 9 : Professional Activity Vs Rest

	MEAN PRA	SD PRA	No of sampl es for MLP	No of sampl es for MFNN CA	Err or for MLP	Err or for MFNN CA	T -value
ma les	20.80	5.15	90	90	12. 17	0.4 7	44.82
			120	120	12. 13	0.4 7	
fe ma les	18.97	3.9	40	40	12. 30	0.4 7	37.96
			60	60	12. 32	0.4 7	

Inference: males who are having more PRA values have almost the same RES values when compared to their female counter parts, which is evident from the both of the networks.

4.1.9 PRA vs FI

Table 10: Professional Activity Vs Food Intake

	Mean PRA	SD PRA	No of sampl es for MLP	No of sampl es for MFNN CA	Error for MLP	Error for MFNN CA	T - value
ma les	20.80	5.15	90	90	46.68	2.172	44.83
			120	120	49.51	2.172	
fe ma les	18.97	3.9	40	40	45.74	2.017	37.96
			60	60	47.89	2.017	

Inference: males who are having more PRA values have more FI values w.r.t their female counter parts and which is evident from both of the network.

4.1.10 PRA vs BMB

Table 11: professional Activity Vs Basic Mobility

	Mea n PRA	SD PRA	No of sampl es for MLP	No of sampl es for MFNN CA	Erro r for MLP	Erro r for MFNN CA	T -value
ma les	20.80	5.15	90	90	89.6 5	16.9 5	44.83
			120	120	177. 10	16.9 5	
fe ma les	18.97	3.9	40	40	37.0 3	2.31	37.96
			60	60	44.6 6	2.31	

Inference: males who are having more PRA values have more BMB values compared to their female counter parts.

4.1.11 LEI vs RES

Table 12: Leisure Vs Rest

	Mean LEI	Standard Deviation LEI	No of samples for MLP	No of samples for MFN NCA	Error for MLP	Error for MFN NCA	T - value
males	0.60	0.73	90	90	12.17	0.4748	9.17
			120	120	12.13	0.4748	
females	0.58	0.90	40	40	12.30	0.4783	5.19
			60	60	12.32	0.4783	

Inference: Males who are having similar leisure have the same rest with respect to the female counterparts and which is evident from both of the networks

5 DISCUSSION AND CONCLUSION

This research is a first step to better understand the elements of the built environments that have an impact on the onset of diabetes using the Neural Networks. The current generation now faces its challenges of the impact of built environment on health, especially the chronic diseases like diabetes. This research gives an insight for future communities that promote physical and mental health.¹⁴ by considering the psychosocial and built environment factors for the prevention of diabetes. Prevention of Type 2 diabetes will require measures to promote physical activity and stress reduction measures in the built environment they live in, and to reduce obesity in adults and children¹³. Finally, in this research, The authors have tried to find the relationship between built environment and diabetes mellitus using neural network technology.

REFERENCES

[1] Sridhar GR, Venkata P, Lakshmi G. Time trends in the prevalence of diabetes mellitus: ten year analysis from Southern India (1994–2004) on 19,072 subjects with diabetes. *J Assoc Physicians India*. 2010;58:290–4.
 [2] Mohan V, Sandeep S, Deepa R, Shah B, Vardhese C. Epidemiology of type 2 diabetes: Indian scenario. *Indian J Med Res*. 2007;125:217.

[3] Sridhar GR, Rao PV, Ahuja MMS. Epidemiology of diabetes and its complications. In: Ahuja MMS, Tripathy BB, Moses SGP, Chandalia HB, Das AK, Rao PV, editors. *RSSDI Textbook of Diabetes*, Hyderabad. Research Society for the Study of Diabetes in India; 2002; pp 95–112.
 [4] Ewing R, Schmid T, Killingsworth R, Zlot A, Raudenbush S. Relationship between urban sprawl and physical activity, obesity, and morbidity. *Am J Health Promotion*. 2003;18:47–57.
 [5] Gumpeny R Sridhar, Pasala Sudhir Kumar, Putcha Venkata, AppaRao Allam, Vijay Kishore Durai, Madhu Kosuri, Manda R NarasingaRao, V K Kumar, Zibin jiang, Gumpeny Lakshmi. Built Environment Factors, Psychosocial Factors and Diabetes Mellitus:A South Indian Study; *Indian Journal of Clinical Medicine* 2010;1 15–22.
 [6] W.David Aha and Kibler,” Instance-based prediction of heart diseases presence with Cleveland database”
 [7] J.W.Everhart,J.E.,Dickson,W.C.,Knowler,W.C.,Johannes , R.S.,” Using the ADAP learning algorithm to forecast the onset of diabetes mellitus”, *Proc. Symp.on Computer Applications and Medical Care*, pp.261-5,1988.
 [8] R.W.Brause, “ Medical Analysis and Diagnosis by neural networks”, *J.W.Goethe-University, Computer Science Dept., Frankfurt a M., Germany*.
 [9] W.H.Wolberg and O.L. Mangasarian, “ Multisurface method of pattern separation for medical diagnosis applied to breast cytology”, *Proceedings of the National Academy of Sciences, USA, Volume 87,pp 9193-9196, December 1990*.
 [10] Kornel Papik, Bela Molnar, Rainer Schaefer, Zalan Dombovari, Zsolt Tulassay, Janos Feher. Application of neural networks in medicine—a review. *Medical Science Monthly* 1998;4:538-46.
 [11] S.M.Kamruzzaman, Ahmad Ryadh Hasan, Abu Baker Siddiquee and Md.Ehsanul Hoque Mazumder. Medical Diagnosis Using Neural Network, 3rd International Conference on Electrical and Computer Engineering, ICECE 2004,28-30 , December 2004,Dhaka,Bangladesh.
 [12] Narasinga Rao MR, Sridhar GR, Madhu K, Rao AA. “ A clinical decision support system using multilayer perceptron neural network to Assess wellbeing in diabetes”, *J Assoc Physicians India* 2009; 57:127-33.
 [13] Manda R NarasingaRao, GR Sridhar, K.Madhu, Allam AppaRao, “ A clinical decision support system using multilayer perceptron neural network to predict quality of life in diabetes”, *Diabetes and Metabolic Syndrome: Clinical Research and Reviews*, 4(2010),57-59.
 [14] Sudhirkumar Pasala, Allam AppaRao, GR Sridhar, “ Built environment and diabetes”, *Int J Diabetes Dev Ctries*. 2010 Apr-Jun; 30(2): 63–68. doi: 10.4103/0973-3930.62594