ISSN 2250-3153

Classifying the population as BPL or non-BPL using Multilayer Neural Network

Preksha Pareek, Dr. Prema K.V.

Computer Science and Engineering, FET-MITS, Lakshmangarh, India

Abstract- Below Poverty Line (BPL) is an economic benchmark and poverty threshold used by the government to indicate economic disadvantage and to identify individuals and households in need of government assistance and aid. In this paper a methodology is shown to classify as BPL or non BPL using Artificial Neural Network. Multi-Layer Perceptron (MLP) network will be used for classification. This classification will take very less time and provides efficient results. Further the effect of network architecture on performance is shown on this application.

Index Terms- Artificial Neural Network, Below Poverty Line, multi-Layer Perceptron

I. INTRODUCTION

Below Poverty Line (BPL) is an economic benchmark and poverty threshold used by the government to indicate economic disadvantage and to identify individuals and households in need of government assistance and aid.. BPL analysis is based on the degree of deprivation in respect of 13 parameter[8]. Measurement and analysis of poverty, Inequality is necessary .firstly it is vital to know what the situation is, that means what is poor and where is the poor located in country? So These Parameters for classifying poverty are: (1)Land holdings:- values contained in this may be one among (i) nil holding (ii) less than1 ha (iii) more than 1 and less than 2 ha (2)Types of house: (i) nil house (ii) non-cemented house (iii) cemented house (3) Availability of clothing :(i)Less than 2 pairs(ii)More than 2 but less than 4 pairs(iii)More than 4 but less than 6 (4)Food security: (i)Less than 1 meal per day in major part of year(ii)Normal 1 meal but sometimes less(iii)Normal 1 meal throughout the year (5)sanitation: (i) open (ii) group bathrooms with irregular water supply (iii) group bathrooms with regular water supply. (6)consumable durables(T.V., electric fan, kitchen appliances): (i)nil (ii)Any one item (iii) Any 2 item (7)literacy status of highest literate: (i) illiterate (ii) up to primary(iii) completed secondary(8)status of household labour:(i) bonded labour(ii) women and child labour(iii) adult males (9)means of livelihood:(i)Casual labour(ii) subsistence cultivation(iii) artisan (10)status of children:(i) not going to school and working(ii) going to school and working (iii)going to school and not working (11) types of indebtedness: (i)daily consumption purpose from normal sources(ii) production purpose from normal sources(iii)for other purpose from normal sources (12) reason for migration :(i)casual work(ii) other form of livelihood (iii) non migrant (13)preference of assistance: (i)wage employment (ii)self employment(iii)Housing[1]. From each parameter one value is selected from three of its sub parameters .Here values for each sub parameter are taken as 0.3 for first one, 0.5 for second one and 0.7 for third one .Such 13 values for each person will be collected. In this paper classification on the basis of these parameters is done by using Neural Network. Neural network can learn few input and target patterns (combinations of above values) they can determine the output of unseen patterns or new patterns of above parameters. So this will give optimization with respect to timings. Here performance for this classification will also depend on network structure this network structure will constitute of number of hidden layers, nodes in each layer if we change number of hidden layers network will learn in different manner and respond differently and number of node in the hidden layer also affect the performance measure and misclassified inputs.

II. NEURAL NETWORKS – AN INTRODUCTION

Neural networks are massively parallel processors composed of simple elements or processing units . In this knowledge is acquired by network from its environment through learning process and this knowledge is stored by using synaptic weights. Model of each neuron constitute a set of synapse or connecting link, which is characterized by weights then adder used for summing input signal weighted by respective synaptic weight and an activation function for limiting the output of neuron[6]. These Processing elements are inspired by biological nervous systems. Each neuron does some amount of information processing.

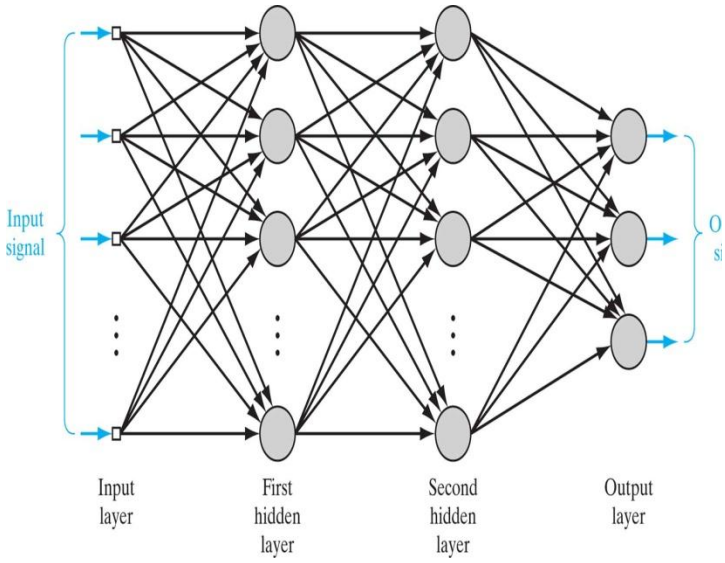
Artificial Neural Network learns by training on past experience which modifies the interconnection weight links So Neural Network is trained to perform a particular function by adjusting the values of the connections (weights) between elements. A neuron is a single processing unit which performs the weighted sum of its inputs. Here output depends on cooperation of each and every neuron. Neural network are characterizes into feed forward and recurrent networks between them feed forward network is used for classification purpose. Neural Networks are trained, so that a particular input leads to a corresponding target output. They can also be trained to solve problems that are difficult for conventional computers or human beings [2].

This paper deals with Multi-layer Perceptron network for performing classification task. . A multilayer (feed forward) network has one or more hidden layers whose neurons are called

as hidden neurons. Here network will be trained using supervised learning algorithm.

Back propagation is most widely used supervised algorithm for learning feed forward networks. It learns the weights for a multilayer feed forward networks, given a network with a fixed set of weights and interconnections. Below figure is architectural graph of MLP [6].

Back propagation employs gradient descent which minimizes the squared error between the networks *output values* and *desired values* for those outputs.



The goal of gradient descent learning process is to minimize the sum of squared errors by propagating those error signals backward through the network architecture upon the presentation of all training samples from the training set. Here error signals are used to calculate the *weight* updates that will represent the knowledge learnt within the network.. The generalization ability of neural networks is an important measure of its performance as it indicates the accuracy of the trained network when presented with data not present in the training set. A poor choice of the network architecture i.e. the number of neurons in the hidden layer will result in poor generalization even with optimal values of its weights after training.

III. METHODOLOGY

Now for performing classification of BPL using Neural Network following tasks are needed to be done:-

1. **DATA COLLECTION**: Initially data will be collected which contains pattern of values. Here supervised learning is used so all the inputs and targets which are required for training needs to be collected. This pair is called as Training Set .In training set number of inputs are 13 as the person will be classified as BPL or not based on these 13 features for this as mentioned above that each 13 features will contain 3 sub features so among them 0.5 is considered as threshold values and output will be 1 or 0 based on number of threshold values. If output will be 1 then the person will belong to BPL otherwise that will be 0 which indicates that person does not belong to BPL. Here testing data will also be prepared in this phase, this data should not overlap with training data this is called as Testing Set.

2. **NETWORK STRUCTURE**: In this step network to be used and then network structure is determined. In this case network will be Multi-Layer Perceptron (MLP). The single-layer perceptron can only classify linearly separable problems. For non-linearly separable problems it is necessary to use more layers .In MLP the model of each neuron in the network includes a nonlinear activation function that is differentiable. The network contains one or more layers that are hidden from both the input and Output nodes. The network exhibits high degree of connectivity, the extent of which is determined by synaptic weight of network.

3. **TRAINING PHASE:** Training in MLP is done by 2 phases: forward phase and backward phase [6]. Back propogation algorithm is widely used algorithm for training multi-layer perceptron network. Steps for back propogation algorithm is firstly initialization of weights are done by assigning random values to weights then input signal receives its input signal and transmits its signal to all the layers above it. Each hidden unit sums its weighted input signal then after applying activation function it is transmitted to output unit each output unit sums its weighted signal then its activation is also calculated. Each output units receives a target pattern to corresponding to input signal and calculates error then weights are updated then this same calculation is performed on hidden layer.[7]

4. **TESTING PHASE:** Finally testing is done on unseen data is done using testing set. This will determine accuracy of our constructed model.

IV. RESULTS AND ANALYSIS

For simulation for above classification problem data set containing 293 training data and 1047 testing data was created. The classification task is done by using Matlab neural network toolbox. Neural network tool box consists of tools for designing, implementing, visualizing and simulating neural networks. Here the coding is performed simply in m file or in command window. [2]

Here performance of classification depends heavily on network structure. In this we have trained the network for

Number of hidden layer	**Mse**	**Gradient**	**Percentage of sample misclassified(%)**
1	**0.17**	**0.00518**	**4.1**
2	**0.0129**	**0.00581**	**5.4**
3	**0.229**	**0.00595**	**33.11**

different hidden layers above table shows the result after simulation. Algorithm used is gradient descent Back propogation. This table shows that for above application (BPL) if

we increase hidden layers then the samples misclassified will also increase.

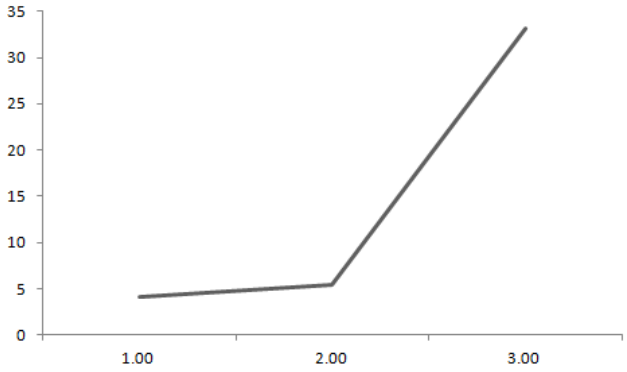


Fig: Relation between Number of hidden layer and samples misclassified is shown. In above table only one hidden layer is taken but by increasing number of nodes in hidden layer percentage of misclassification also increases.

Number of nodes in hidden layer	Mse	Percentage of sample misclassified (%)
20	0.0138	3.07
50	.0100	2.39
88	.0100	1.02

Above table shows the relation between number of nodes in hidden layer and percentage misclassification .

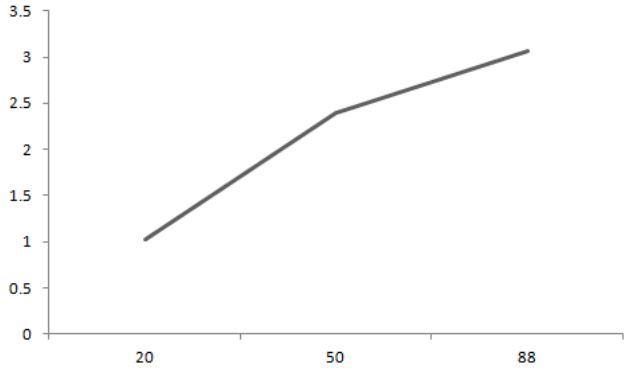


Fig: Relation between nodes in hidden layer and percentage of misclassification. Here we have taken only single hidden layer. As we increase number of nodes in hidden layer then percentage of misclassification decreases.

Activation Function	mse	Percentage of sample misclassified (%)
Purelin	.139	15
Tansig	.116	19
Logsig	.0129	5.46

Above table shows effect of activation function on percentage samples misclassified. Here we have used 2 hidden layers with 10 and 1 hidden node respectively. Here when we use logsigmoid activation function then misclassification rate is least.

V CONCLUSION

Below Poverty Line classification by using Neural Network provides a new method for classifying a person as BPL or not. Neural Network learns the sequences of input-target pair by using learning algorithm through updation of weights and when tested on unseen patterns or samples it provides efficient and quick results and network performs differently when architecture is different.

ACKNOWLEDGMENT

I would like to thank Dr. Prema K. V. for helping me in this work and also all the other people who have encouraged me for this research.

REFERENCES

[1] Indira Hirway(2003)," Identification of BPL Households for Poverty Alleviation Programmes" Economic and Political Weekly, vol 38,no 45,pp 4803-4808.

[2] V. Arulmozhi (2011)"Classification task by using Matlab Neural Network Tool Box – A Beginner's View" International Journal of Wisdom Based Computing, Vol. 1 (2).

[3] Frauke Günther and Stefan Fritsch(2010)" neuralnet: Training of Neural Networks" The R Journal Vol. 2/1

[4] Guoqiang Peter Zhang (2000) "Neural Networks for Classification: A Survey"IEEE transaction on systems,man,and cybernatics—part C: Applications and reviews, VOL. 30, NO. 4.

[5] Industrial Automation – Artificial Neural networks" , *by: Shady Gadoue:*

[6] "Neural Network and learning machine" Simon haykin.

[7] "Introduction to Neural Networks using Matlab 6.0"S N Sivanandam ,S. Sumuthi,S N Deepa.

[8] "Report of expert group to advise ministry of rural development on the methodology for conducting BPL census for 11th 5 year plan",2009.

AUTHORS

First Author –Preksha Pareek, M.Tech from Mody Institute of Engineering and Technology. prekshapareek@gmail.com
Second Author –Dr. Prema K. V., Professor at Mody Institute of Engineering and Technology, premakv.et@mitsuniversity.ac.in