Journal of Applied Sciences

ISSN 1812-5654

science alert

ANSInet
an open access publisher
http://ansinet.com

OPEN ACCESS Journal of Applied Sciences

ISSN 1812-5654
DOI: 10.3923/jas.2016.594.599

Research Article
Neural Network for Farm Household Output Prediction

[1]Muhd. Khairulzaman Abdul Kadir, [2]Arjunan Subramanian, [1]Nadaraj Miniappan and [1]Mohd. Zaki Ayob

[1]British Malaysian Institute, Universiti Kuala Lumpur, Kuala Lumpur, Malaysia
[2]Adam Smith Bussiness School, University of Glasgow, United Kingdom

Abstract
In ensuring that food security is at an acceptable level, all of its range of indicators needs to be monitored and maintained. In this study, farm household behaviour such as the mode of labour, type of fertilizer being used; cost of each component for the farm and so forth can be used to predict the farm household output of crops. The crop output is one indicator of food security for household level. The dataset used is based on the Village Level study (VLS) by the International Crops Research Institute for Semi Arid Tropics (ICRISAT). This dataset consists of 37 features and 29 samples of households in 1975. For the prediction model, the Optimum Weight and Threshold Neural Network (OWTNN) is proposed on 37 inputs and one output of crops for each household and compared with the Artificial Neural Network (ANN) performance. The result of the proposed model shows that OWTNN gives a better performance than ANN, which is 99% compared to 86%. Hence, it proves that the proposed model can be offered as one of the best predictors for farm household output. The model also shows that farm household behaviour can affect farm crop output.

Key words: ANN, neural network, optimization, food security, household behaviour, farm output, crops output

Received: April 11, 2016 Accepted: June 29, 2016 Published: November 15, 2016

Citation: Muhd. Khairulzaman Abdul Kadir, Arjunan Subramanian, Nadaraj Miniappan and Mohd. Zaki Ayob, 2016. Neural network for farm household output prediction. J. Applied Sci., 16: 594-599.

Corresponding Author: Muhd. Khairulzaman Abdul Kadir, British Malaysian Institute, Universiti Kuala Lumpur, Kuala Lumpur, Malaysia

Competing Interest: The authors have declared that no competing interest exists.

Data Availability: All relevant data are within the paper and its supporting information files.

INTRODUCTION

In 2010 UK Department of Food and Rural Affairs (DEFRA) report[1], one of the key features of food security was based on household food security where everyone in all places should be able to access and afford a healthy food[1]. The headline indicator for the food security theme is the proportion of income that a household spends on food. India, in the case of villages such as Aurepalle, the village considered here, households both plant their own crops and work in other farms to provide food. In some instances, the farms are owned and operated for consumed food and profitable sales where some farm owner sold the crops for its profits. Thus, the food supply in this case was based mostly on crops from the farm or other farm household outputs within the same village[2].

The International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) is one organization which studies the socioeconomics, agro-biological, institutional constraints to agricultural development in Semi-Arid Tropical (SAT) areas and they also help in testing and modifying the new technologies generated by themselves for the benefits of people in Indian villages[2]. For our case study, one dataset for the year 1975 from Aurepalle village in India was taken based on the ICRISAT study data to predict the farm household output of crops. The 'farm household' term being used rather than 'farm' because of the categories on the dataset. This dataset was based on the research of small farmer, medium farmer and large farmer[2]. Previous studies have shown that many factors can reflect farm household crop outputs such as the types of labour being used, the type of fertilizer being used, the cost of each component for the farm and so on[3-8]. Here, based on the village dataset, the farm household crops output will be referred to as grains, vegetables or fruits, these are also for fodder and seedlings use as well as human consumption[2].

In summary, this research studies farm household crop output predictions, making use of the ICRISAT data for Aurepalle village in the year of 1975. The model utilised is an Optimized Weight and Threshold Neural Network (OWTNN) the performance of which will be compared with the traditional Artificial Neural Network (ANN).

LIMITATIONS IN THE STATE OF THE ART

The ANN is one of the best and well known tools for prediction because of its fast and effectiveness in solving non-linear problems[9,10]. Generally, ANNs have 3 layers, the input layer (I), the hidden layer (can be more than one layer, J and K) and the output layer (L) as shown in Fig. 1. The interconnection between each layer consists of weights (W_{IJ}, W_{JK} and W_{KL}) and at each neuron there are also thresholds connection (B_I, B_K and B_L). An ANN can be used with either multiple inputs to single output or multiple inputs to multiple outputs[11].

Although, an ANN offers a good tool in prediction it does travel to local extrema and the convergence is slow[9]. Moreover, when the ANN is not able to generalize, over-fitting and under-fitting can occur especially when the dataset is divided into training, validation and testing parts[12,13]. Most of these problems happen because the final interconnecting weight and threshold of the network are constantly and rapidly changing in an uncontrolled manner during the training phase[14]. In solving these problems, a Genetic

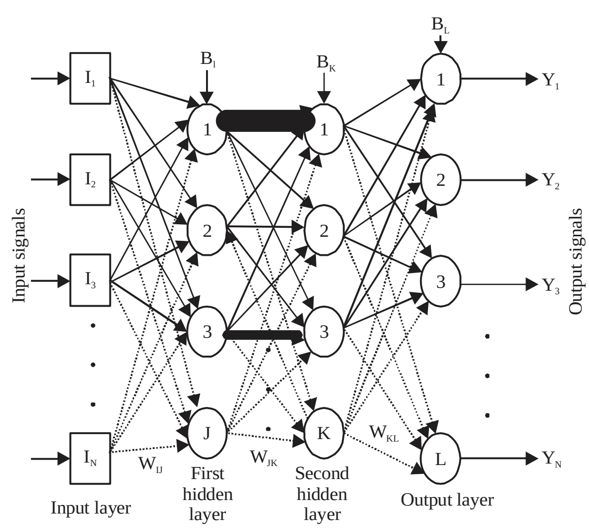

Fig. 1: Basic ANN architecture

Algorithm (GA) is used to optimize the weight and the threshold of ANN and the process will be explained in later sections.

MATERIALS AND METHODS

The GA is a well-known technique for its capability in optimizing certain datasets or any network architecture. Furthermore, it offers an effective search technique based on the principle of genetics and natural selection[11,15]. The advantages of adding a GA to an ANN are that it can avoid local minima, which it finds rapidly and discards. It also can search every region simultaneously with great efficiency through the GA operations and parameters of selection, crossover probability and mutation probability[9,10,15,16].

In this study, ANN had multiple inputs and one output with one hidden layer with 30 neurons. The number of hidden neurons was selected based on multiple ANN training and was taken from the best regression value. The data was divided to 70% for training, 15% for validation and 15% for testing. Then, the GA will optimize the weights and the thresholds of the ANN in Fig. 1 by its aforementioned bio-inspired evolutionary operators. The process of the OWTNN is shown in the flow diagram of Fig. 2.

This model was simulated in the MATLAB™ 2010 environment where each parameter was selected as Fig. 2. The inputs and the output of the ANN were as in Table 1, consisting of 37 features with 29 samples. As explain previously, the dataset was based on Aurepalle village in India for the year 1975.

RESULTS AND DISCUSSION

From Fig. 2, GA will randomly initialize all the weight and threshold values for the ANN based on the size of chromosomes for each population. Chromosomes consist of

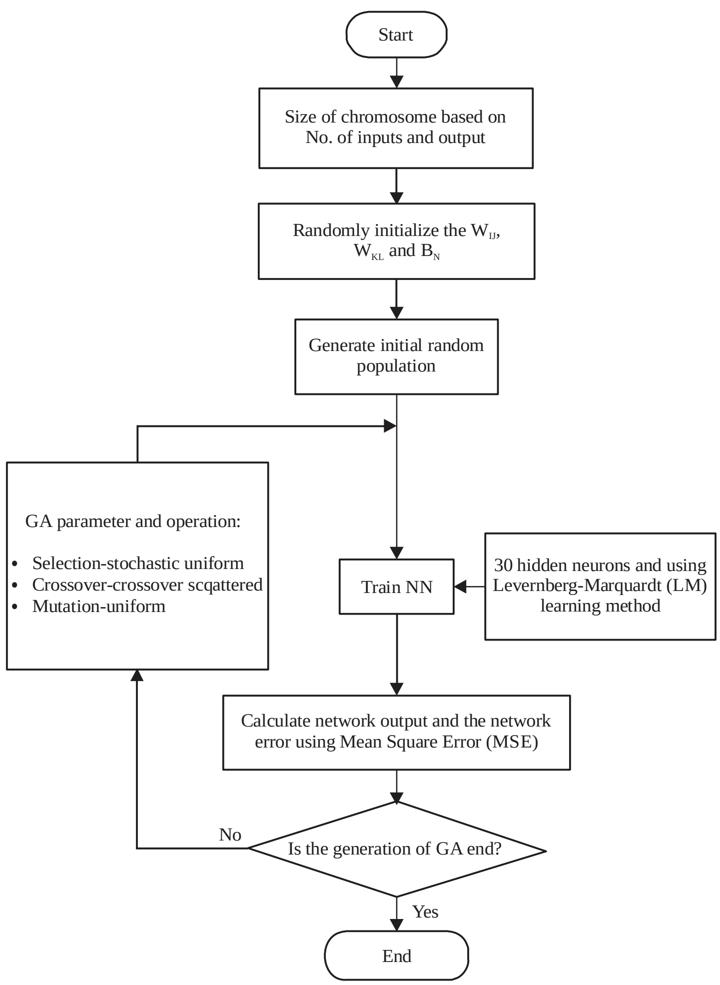

Fig. 2: OWTNN process using GA

a number of genes which is an array of variable values to be optimized[11,15]. The chromosomes sizes will depend on the total interconnections between the number of inputs, number of hidden neurons and number of output neurons in ANN. After that, the ANN will be trained iteratively by the GA based on the population size for each chromosome until it achieves the best Mean Square Error (MSE) as Eq. 1 for the fitness function.

In Eq. 1, NN_{out} is the ANN output and T is the actual output based on the dataset. The result of the optimization process is shown in Fig. 3 where the plot are based on MSE (fitness value) and the average distance between individual chromosomes at each generation:

$$MSE = \frac{\sum_{i=1}^{N_{total}} (NN_{out_i} - T)^2}{N_{total}} \tag{1}$$

where, N_{total} is the total number of outputs.

When the GA optimization is performed, the best weight and threshold values generate by the GA is applied to the ANN and it will calculate the farm household output prediction as in Fig. 4. This shows that OWTNN gives better performance than ANN (Fig. 5), R = 0.99 compared to 0.87. The performance comparison of ANN is not good because of the over-fitting and under-fitting problems for validation and testing, as

Table 1: Inputs and output variable

Features	Output
Land related	
Plot value (Rs), crop areas (Acres), soil type, irrigated area (Acres)	
Crops related	
Total seed value (Rs), total main product (Rs), total by product value (Rs), total output value (Rs), total input value (Rs), net income (Rs), net return (Rs)	
Fertilizer and pesticide related	
Total fertilizer value (Rs), total FYM quantity (Quintals)	
Total FYM value (Rs), Sheep penning value (Rs), tank silt/soil adding (Rs), all organic manure value (Rs), nitrogen inorganic (kg), phosporus (kg), potash (kg), total N (kg), total P_2O_5 (kg), total K_2O (kg), all pesticides value (Rs), crop output	
Manpower/animal power related	Quantity (kg)
Family male (h), family female (h), family child (h), hired male (h), hired female (h), hired child (h), owned bullock (h), hired bullock (h), total family labor value (Rs), total hired labor value (Rs), total owned bullock labor value (Rs), total hired bullock labor value	

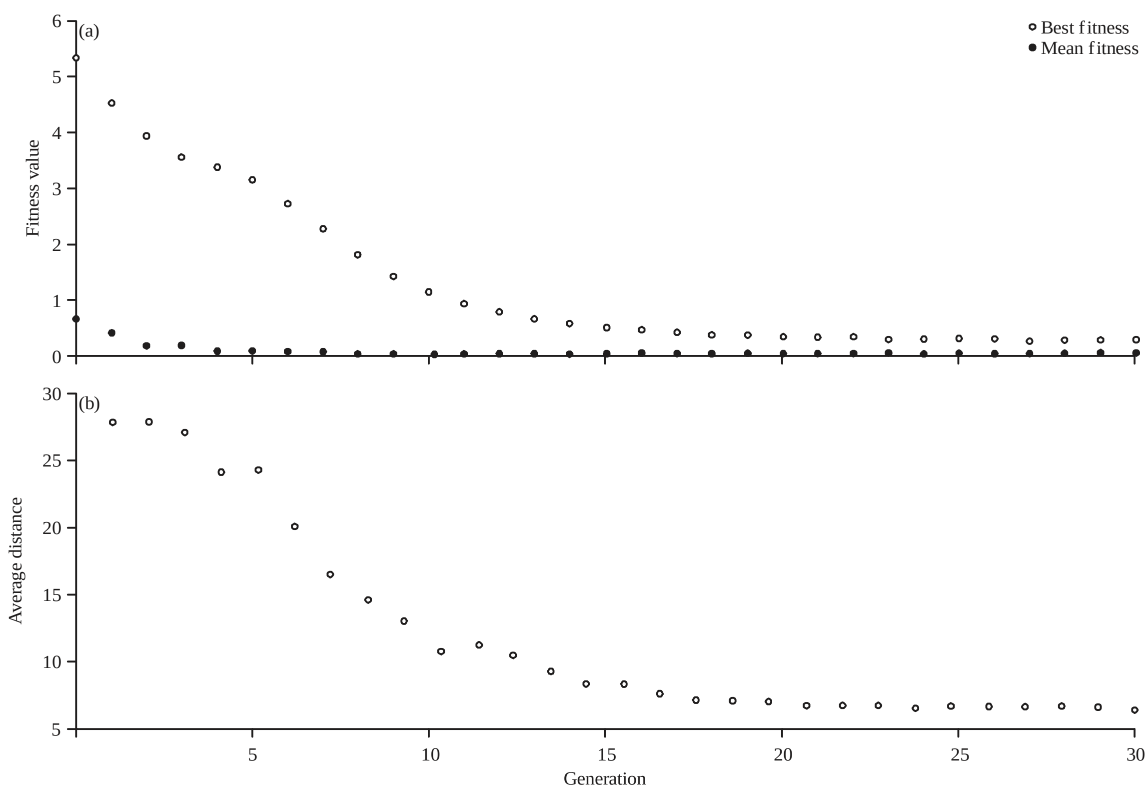

Fig. 3(a-b): GA optimization result, (a) Best: 0.029793, Mean: 0.27315 fitness values and (b) Average distance between individuals

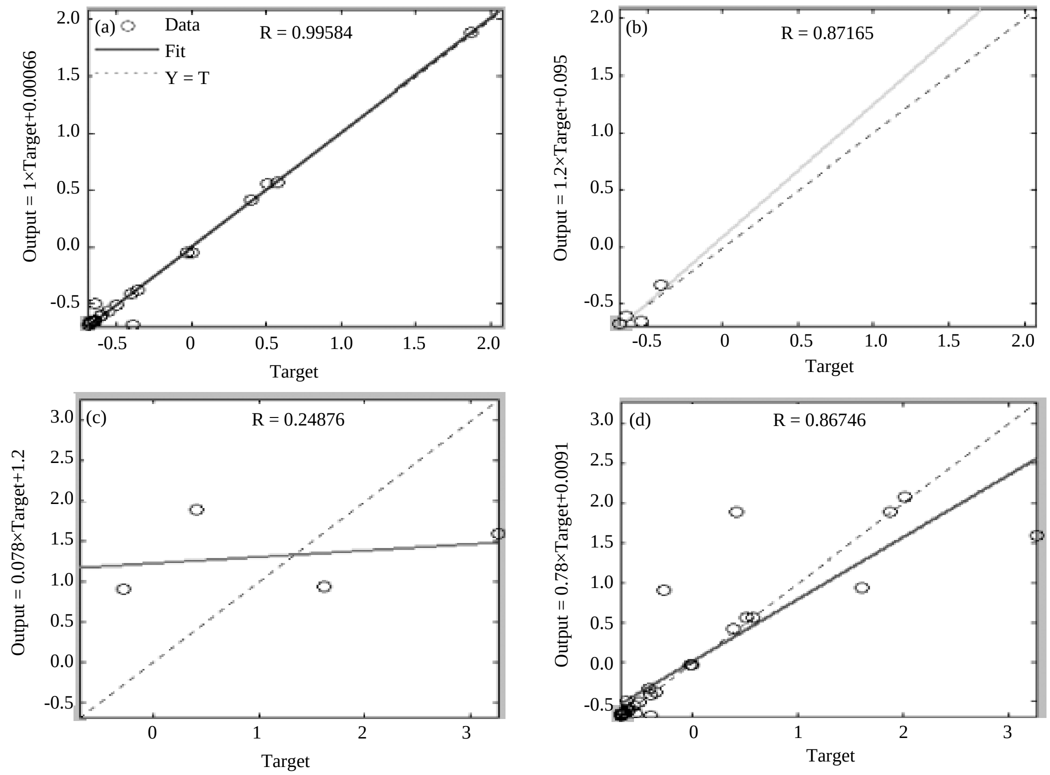

Fig. 4(a-d): ANN results, (a) Training, (b) Validation, (c) Test and (d) All

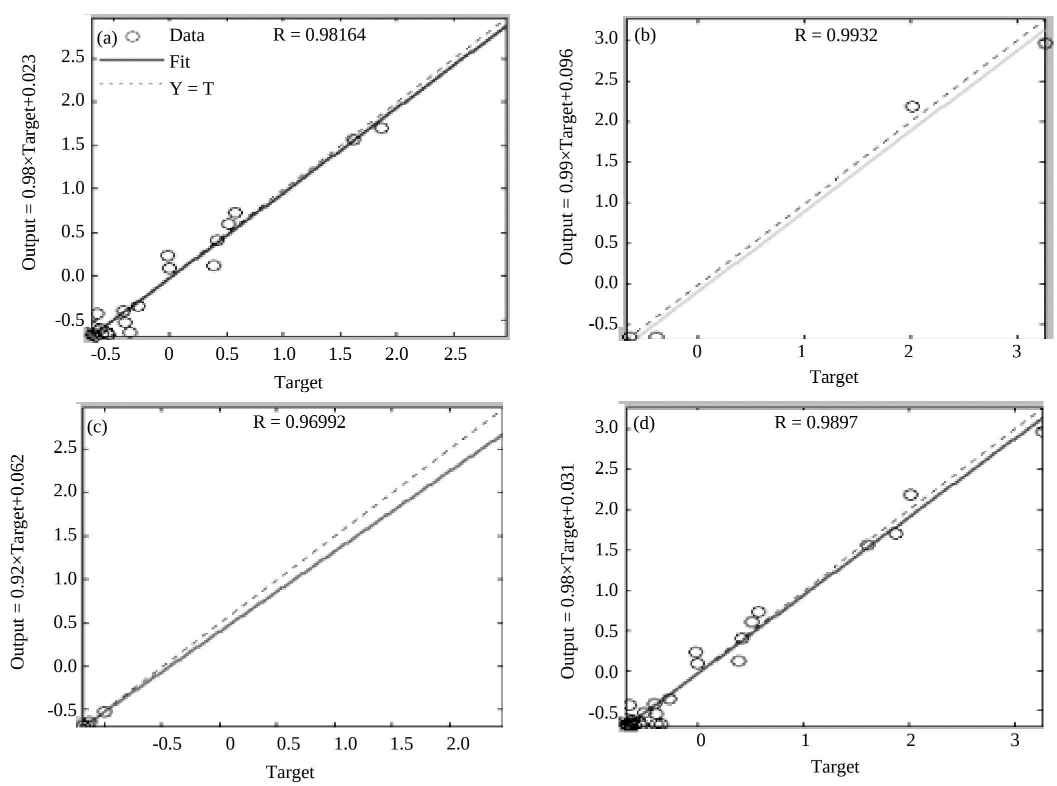

Fig. 5(a-d): OWTNN results, (a) Training, (b) Validation, (c) Test and (d) All shown in Fig. 4, which have been pointed out. If compare with Fig. 5, all data division part (training, validation and testing) show considerable generalization ability.

CONCLUSION

In this study, OWTNN has been successfully applied to predict the farm household crop output for each household in Aurepalle, India with a 14% improvement in the regression value compares to original ANN methods (to 0.99). It also shows that each activity or component being used and the value of each component on the farm can affect the crops output of a household.

In addition, the proposed model shows how to achieve good generalizations of the ANN network by optimizing the weight and the threshold for each neuron. The model offers the prospect of an excellent prediction tool for farm household crop output. This method offers the option of timely analysis of household food security in order to inform future modifications of farming activity patterns to achieve better crops outputs.

ACKNOWLEDGMENT

We would like to thank the School of Engineering, University of Warwick for their support. Thank you also to Arjunan Subramanian for gaining permission to use the data and to the ICRISAT in making the data available.

REFERENCES

1. DEFRA., 2010. UK food security assessment: Detailed analysis. Department for Environment Food and Rural Affairs (DEFRA), January 2010.
2. Singh, R.P., H.P. Binswanger and N.S. Jodha, 1985. Manual of instructions for economic investigators in ICRISAT's village level studies (revised). International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Andhra Pradesh, India, May 1985.
3. Singh, R.P., B.L. Jain and V.B. Rao, 1984. Feature of traditional farming systems in selected villages of Madhya Pradesh. Working Paper, Economics Progress Report No. 61, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India.
4. Skoufias, E., 1993. Seasonal labor utilization in agriculture: Theory and evidence from agrarian households in India. Am. J. Agric. Econ., 75: 20-32.
5. Skoufias, E., 1994. Risk and seasonality in empirical model of the farm household. J. Econ. Dev., 19: 93-116.
6. Bhende, M.J. and J.V. Venkataram, 1994. Impact of diversification on household income and risk: A whole-farm modelling approach. Agric. Syst., 44: 301-312.
7. Walker, T.S. and K.V. Subba Rao, 1982. Yield and net return distributions in common village cropping systems in the semi-arid tropics of India. International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Andhra Pradesh, India, August 1982.
8. Jha, D. and R. Sarin, 1981. An analysis of levels, patterns and determinants of fertilizer use on farms in selected regions of semi-arid tropical India. Progress Report No. 25, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Andhra Pradesh, India, October 1981.
9. Karimi, H., F. Yousefi and M.R. Rahimi, 2011. Correlation of viscosity in nanofluids using Genetic Algorithm-Neural Network (GA-NN). Heat Mass Transfer, 47: 1417-1425.
10. Rajendra, M., P.C. Jena and H. Raheman, 2009. Prediction of optimized pretreatment process parameters for biodiesel production using ANN and GA. Fuel, 88: 868-875.
11. Negnevitsky, M., 2005. Artificial Intelligence: A Guide to Intelligent Systems. 2nd Edn., Addison-Wesley, Boston, MA USA., ISBN: 9780321204660, Pages: 415.
12. Betechuoh, B.L., T. Marwala and T. Tettey, 2007. Using inverse neural networks for HIV adaptive control. Int. J. Comput. Intell. Res., 3: 11-15.
13. Demuth, H.B. and M.H. Beale, 2004. Neural Network Toolbox for Use with MATLAB: User Guide. The Math Works Inc., USA., Pages: 474.
14. Zhang, Q. and C. Wang, 2008. Using genetic algorithm to optimize artificial neural network: A case study on earthquake prediction. Proceedings of the 2nd International Conference on Genetic and Evolutionary Computing, September 25-28, 2008, Hubei, pp: 128-131.
15. Haupt, R.L. and S.E. Haupt, 2004. Practical Genetic Algorithms. 2nd Edn., John Wiley and Sons, New York, USA., ISBN-13: 9780471455653, Pages: 253.
16. Goldberg, D.E., 1997. Genetic Algorithm in Search, Optimization and Machine Learning. Addison-Wesley, USA.