Neural Comput & Applic
DOI 10.1007/s00521-017-2889-8

ORIGINAL ARTICLE

Neural network rule extraction for gaining insight into the characteristics of poverty

Arnulfo Azcarraga[1] · **Rudy Setiono**[2]

Received: 30 June 2016 / Accepted: 10 February 2017

Abstract Nearly one in five families in the country was poor in 2012, according to the Philippine Statistics Authority. While this proportion is lower than the corresponding figures from 2006 and 2009, the absolute number of poor families has actually grown from 3.8 million in 2006 to 4.2 million in 2012 due to the increase in population. Using data samples that have been collected from 69,130 households through a comprehensive community-based monitoring survey conducted in one of the cities that comprise Metro Manila, we attempt to identify the characteristics that differentiate between poor and non-poor households. Using back-propagation neural networks, we are able to correctly predict 73% of the poor households and 60% of the non-poor households. Moreover, the rules extracted from one of these networks provide concise description of how households are classified as poor based on their demographic characteristics and information pertaining to their surrounding living conditions.

Keywords Neural networks · Back-propagation · Pruning · Rule extraction · Poverty

✉ Rudy Setiono
disrudy@nus.edu.sg

Arnulfo Azcarraga
arnie.azcarraga@delasalle.ph

1 College of Computer Studies, De La Salle University, 2401 Taft Avenue, Manila, Philippines

2 School of Computing, National University of Singapore, 13 Computing Drive, Singapore 117417, Republic of Singapore

1 Introduction

Fighting poverty remains one of the top priorities for many governments in the developing world. In the Philippines, the National Anti-Poverty Commission was established in 1998 as the agency that coordinates the implementation of poverty alleviating programs run by local and national governments as well as the private sector. To help it in its mission, the National Anti-Poverty Commission has adopted the Community-Based Monitoring System (CBMS) as the local monitoring system. Designed as a tool to facilitate the implementation of poverty reduction programs in the country, the CBMS collects individual and household data from several provinces in the Philippines covering a majority area of the archipelago.

The CBMS database is made available to all interested users, in particular the local area planners, policy makers and non-governmental organizations. It is hoped that the availability of these data could better serve the planning and management needs of the local government and non-government agencies, particularly for the delivery of basic services such as clinics and hospitals, school, new roads and bridges and provision of clean water and electricity. The poverty datasets that have been collected in the Philippines contain in-depth information about individual households, clustered into small social and political units or districts, referred to locally as *barangays*. Each barangay has from a few hundreds to a few thousand households.

The availability of these datasets provides us with an opportunity to conduct a quantitative study on the incidence of poverty in the Philippines by employing one of the most effective machine learning techniques for pattern classification, namely neural networks. In particular, back-propagation networks are used to predict whether a household is poor or non-poor, based on their demographic characteristics and various other information pertaining to the physical dwellings and the surrounding living conditions. More importantly, rule extraction from the back-propagation networks provides concise descriptions of the major conditions that correlate highly with poverty, as well as the conditions that generally correlate with households that can be considered non-poor.

Most existing quantitative studies on poverty data employ traditional statistical approaches such as least squares regression and logistic regression to identify the most relevant predictors of poverty. In this paper, we present the application of a more novel data mining technique, i.e., rule extraction from back-propagation neural networks that have been trained on the CBMS poverty data. The goals of the data analysis are as follows: (1) to see whether it is possible to build a neural network model that distinguishes between poor and non-poor households in the data with acceptable prediction accuracy and (2) to discover the characteristics of households that make them more likely to be poor.

For many applications in business and economics, neural networks have been shown to be very effective tools for data analysis. Neural networks are applied to analyze international debt problems using data from the World Bank [39]. The classification problem is to differentiate between countries that reschedule their international debt and those that do not. The results from the study show that neural networks trained with variants of the conjugate gradient method for error minimization predict with higher accuracy than logit and probit models. The superior performance of neural networks is also demonstrated in the application of time-delay neural network (TDNN) for predicting the monthly wholesale price of oilseeds in India [18]. The authors find that TDNN provides better predictive accuracy than the traditional time series forecasting approach, namely the autoregressive integrated moving average (ARIMA) model due to the nonlinear nature of the time series. Other successful business and economics applications of neural networks include an early warning prediction system for high inflation [27], an analysis of service satisfaction in web auction logistic [21], and a study of the relationship between corporate reputation and market values of Spanish listed companies [13].

The rest of this paper is structured as follows. In the next section, we present related works on quantitative approaches for poverty data analysis. The majority of these papers describe analyses of poverty data that have been collected in developing countries in Asia and Africa. In the third section, the details of the Philippines poverty data used in our study are introduced. In this section, we also present the results from our study applying artificial neural networks on this dataset. We discuss our finding and present our conclusion in the last two sections.

2 Related works

Two main purposes for studying poverty are (1) targeting interventions and (2) designing programs and policies [5, 31]. For targeting interventions, the household poverty data are used to evaluate if certain poverty alleviating programs have positive impact on their targeted group. As for designing programs and policies, we seek to help policy makers develop programs that can reduce poverty over a sufficiently long period of time. Data for quantitative studies of poverty are normally obtained via household survey where factors such as access to clean water and medical care and demographic information (e.g., age and gender of the head of the household) are collected.

An abundance of research papers on quantitative data analysis related to poverty and poverty alleviation can be found in the literature. The data analysis reported in these papers normally involves the use of traditional statistical approaches such as principal component analysis, least squares regression and logistic regression [1, 8, 11, 16, 25, 37].

Sy [35] builds a probit regression model to predict poverty in Senegal. Instead of measuring poverty in monetary terms, the author chooses to measure it in terms of how the households themselves feel about their socioeconomic situation. Hence, for each household, the value for the dependent variable in the model of either poor (value = 1) or non-poor (value = 0) is determined by the head of the household himself. Ten independent variables are included in the model. These variables represent demographic information about the head of the household (gender, age, education level, place of residence), economic situation in the household and in the district residence, the perceptions on policies implemented by the state, and the household per capita expenditure. The probit regression model built using a total of 13,584 data samples finds all the independent variables to be statistically significant at the 1% level, except for gender which is significant at the 5% level.

The subjective assessment of whether a household is poor is also studied by Alem et al. [3]. In this study, the authors analyze data collected from urban population in Ethiopia over a 15-year period from 1994 to 2009. They note that the proportion of the population who regards themselves as poor has not changed despite increases in income and material consumption. The data for the study were collected in 1994, 1997, 2000, 2004 and 2009 from households in four major urban centers in Ethiopia including its capital Addis Ababa. The subjective poverty measure is obtained by asking the survey participants whether they consider themselves as rich, middle-income or poor. The objective poverty measure is also computed using food and non-food expenditure data. A household is classified as poor when its real consumption expenditure per adult equivalent units is below the capital city's poverty line in 1994. Probit models for the probability of being in poverty reveal that households that are poor in one period are more likely to be poor in the next period for both types of poverty. Among other interesting findings, the study also discovers that the probability of being poor decreases with increasing education level of the head of the household. A household headed by one who has completed tertiary education has a 20% less probability of being poor compared to another household headed by an illiterate.

Poverty headcount or the proportion of the population whose level of welfare is below the poverty level is the dependent variable in a study to find determinants of poverty prevalence in rural Malawi [7]. The aim of the study is to identify spatial determinants that could explain the different levels of poverty headcount. A total of 26 independent variables are selected as potentially significant determinants. They include agroclimatological data such as average rainfall, natural hazards information such as areas subject to flooding and accessibility to hospital and market. Demographic information that is also used in the regression model includes the ratio of men to women between 20 and 49 years old, the proportion of the population under 15 years old and above 65 years old, the percentage of households headed by women, and population density. The total number of data samples of 3004 corresponds to the number of aggregated Enumeration Area (EA), where each EA has a population count of just above 500 households. Two different analyses are carried out, a single global model and a geographically weighted regression local models. The global model fits one regression model using all 3004 data samples, while the local models produce one set of regression coefficients for each EA. Eight independent variables are found to be significant in the global model, but all independent variables are considered significant in at least one of the 3004 local models. From the results of this study, it is suggested that poverty reduction efforts in rural Malawi should be designed for and targeted at local levels.

The relationship between poverty- and disease-related mortality in countries with high HIV prevalence is the subject of a recent study by Chapoto et al. [9]. Using data that have been collected in Zambia and Kenya, probit regression models are generated to predict the probability an individual surveyed has died of disease-related causes between 2001 and 2004 in the case of Zambia, and 1997 and 2004 in the case of Kenya. The regression results show that while there is no correlation between wealth status and the probability of mortality in Zambia, there exists a positive correlation between HIV infection and wealth status in this country. On the other hand, a negative correlation between wealth status and the probability of mortality is observed in the regression model built using the data from Kenya.

Clustering and principal component analysis are applied to identify household groups with varying livelihood in Rwanda [4]. The study is done using a dataset consisting of 1220 cases that are representative of rural inhabitants of Rwanda. There are 11 variables that capture information relating to the 5 types of livelihood capitals (i.e., natural, physical, human, social and financial capital). An additional set of 4 variables is also included to provide regional context of the cases. Principal component analysis reduces the data dimensionality to just 6. The new dataset is then clustered using K-mean procedure where 7 clusters are found to provide the optimal balance between parsimony and homogeneity. Based on the statistics of the cases in each cluster, the authors identify these clusters as (1) rural entrepreneurs, (2) association type, (3) resource rich, (4) resource poor in fertile regions, (5) resource poor—centrally located, (6) isolated and (7) female-headed clusters. Targeted poverty alleviation policies are recommended based on different livelihood profiles and pathways of the households in the various clusters.

Education is shown to have a positive and significant influence on the tendency of the people in Fiji to engage in health care and wellness activities and in acquiring good housing facilities [14]. Health-promoting activities such as acquiring life-accident insurance and improvement in sanitation are two dependent variables used as proxy that measure the impact of education and other independent variables in the non-monetary models of poverty reduction proposed by the authors. The other independent variables in the models are age, sex, ethnicity and years of schooling of the household head, as well as whether the household is living in rural areas. The results from a monetary model where the dependent variable is the household income show that a household's total income is expected to increase by 3.5% with an additional year of schooling completed by its head. Total income also increases as the age of the household head increases. Total income is less for households that are headed by female, ethnic Indo-Fijian, and living in rural areas.

Several poverty data analyses that illustrate the applications of non-traditional methods such as fuzzy logic and machine learning techniques have also been reported more recently. Fuzzy set logic is suggested as the solution to the identification and aggregation problems of poverty measurements [28]. Identification problem refers to the problem of identifying who is poor and who is not poor according to one or more indicators, e.g., income, consumption or the combination of life expectancy, education and living standard. Aggregation problem is the problem of determining the units in the poverty study, whether the unit of analysis is household, social groups, or other bigger groupings. Datasets collected from 187 households in two villages in India are used to illustrate how a fuzzy scale multidimensional poverty index can be developed. The multidimensional poverty index is derived by first transforming each variable in the data into a crisp set or a fuzzy set. In the case of fuzzy set, the range of the variable value is now between 0 and 1 with infinite steps in between to indicate the degree of membership in the set. Indicators that are in the same domain are then aggregated using different fuzzy set aggregation functions *AND* or *OR*. For example, a household is considered deprived if it does not have a house or a toilet, or if it owns a house but not a toilet if the indicators 'own a house' and 'own a toilet' have been aggregated using the *AND* function. The multiscale poverty index is finally derived on a scale from 0 to 1 with an equal increment of 0.11 with the fuzzy scale value of 0 indicating full membership in the set of non-poor and the value of 1 indicating full membership in the set of fully poor.

A class of one-dimensional fuzzy measures is proposed by aggregating three levels of poverty: strong poverty (SP), medium poverty (MP) and weak poverty (WP) [6]. The approach is applied to analyze poverty of different regions in Tunisia and also to compare poverty in rural versus urban regions according to the activity and the educational level of the household head. The membership functions of SP, MP and WP are calculated for 7734 households. The study reveals that strong poverty is a mostly rural phenomenon in Tunisia and it affects the interior region of the country more severely. Farm laborers and illiterate people also show high degree of membership in the SP fuzzy subset.

The incidence of poverty in Vietnam which represents the percentage of the population that falls below the poverty line is predicted by using both a statistical method (linear regression) and a machine learning technique (support vector machine) by Senf and Lakes [32]. Twelve independent variables are collected. These variables describe land use (open and close forest resources, crop density), the remoteness of the locale, the literacy rates and household as well as demographic characteristics. The models are generated using a training dataset consisting of 61 samples. Their predictive accuracy is tested on a test dataset containing 641 samples. Each data sample corresponds to an area in the country. The results show that linear regression (LR) model has lower prediction error than support vector machine (SVM). However, the SVM errors are mostly below one standard deviation and their average deviation is lower than the average deviation of the LR errors.

In the next section, we present the results from our own study on determining factors that affect poverty in the Philippines. Neural networks are used for data analysis as they have been shown to be effective alternatives to statistical approaches for solving classification and regression problems. Practical applications of artificial neural networks can be found in a wide variety of fields ranging from resource management [19, 26, 38], business [17, 30, 36], engineering/science [2, 20, 40] and health/medicine [12, 22, 23]. A comprehensive study that looks into research papers reporting performance comparison between neural networks and statistical approaches such as regression and discriminant analysis reveals interesting findings [29]. The authors review 73 papers which report 96 cases in total showing a comparison between neural networks and statistical approaches. Out of 96 cases, neural networks are found to perform as well as or better than the statistical approaches in 82% of these cases.

3 Predicting poverty: neural network models

3.1 Dataset

The Philippines CBMS database is comprehensive and includes information that captures in great detail the various living conditions as well as demographic profile of the members of each household. The dataset pertaining to individual households has numerous data fields and comprises two separate subfiles. The first file contains information referring to the various poverty-related information concerning the entire household's living conditions, such as access to clean water, type of toilet facilities, type of materials used for dwelling walls and roof, and various other household information. The other subfile incorporates demographic information of each member of the household, i.e., head of household, spouse, children, grandparents, in-laws and relatives.

"Poverty" is not a single-faceted concept. More and more, poverty is expressed as a multidimensional phenomenon. For example, a household may not be poor in terms of salaries and income, but may be poor because of the prevalence of crime and violence that has directly affected the household ("security poor"). Another example is the inaccessibility of basic utilities and sanitation facilities that would expose the household to health issues, especially the children. In this case, the household may be classified as "health poor." For the purpose of this research, households have been classified as "poor" or "non-poor," based on their annual per capita income. This is computed based on the annual income of the entire household divided by the number of members in the household. A household is considered "poor" when the per capita annual income is less than 2 US$ / day $\times$ 365 days. Based on this threshold, just slightly over 35% of the 69,130 households in our dataset are considered poor.

As we are building a predictive neural network model to distinguish between poor and non-poor households, we divide the available data samples randomly into training

(10%), cross-validation (10%) and test (80%) sets. Table 1 shows the number of each type of households in the three sets.

3.2 Data preparation and preprocessing

We merge the two subfiles containing living conditions and demographics information, so that data about members of a given household are assembled together with the rest of the household information. The merging of files has to deal with some errors in the encoding of the household identification number or key. Some other issues with inconsistencies and missing data, typical in live, real-world datasets, have to be dealt with. Once merged, cleaned, and preprocessed, the data fields are uniformly encoded based on the elaborate coding scheme adopted by the survey instrument.

For the experiments discussed in this paper, we have chosen to retain only a small fraction of the data fields—concentrating only on (a) the physical attributes of the house or dwelling, such as the materials used for the walls and roof; (b) various community-, health-, hygiene-related conditions of a household that may be generally associated with poverty, such as access to clean drinking water, access to proper toilet facilities and garbage collection; and (c) various other indicators affecting the household such as being hit by a calamity (e.g., typhoon, flood, fire) in the past 12 months or having had no food to eat for at least one instance during the last 3 months. Non-poverty-related fields, such as those that refer to the demographic profile of specific members of the household, e.g., gender, marital status, education, age and religion of the head of the household, are also used for training the neural networks. Appendix A provides a complete listing of all the variables.

Prior to training the neural networks, we preprocess the input variables with non-binary values as follows.

1. Age: we discretize the values of this variable using 5-year sub-intervals and represent the discretized values as a binary string of length 11 as shown in Table 2.
2. Civil status and Religion: we represent these variables using binary strings of length 6 and 7, respectively. Exactly one bit in the string has value equal to 1.

Table 1 The number of samples in the training, cross-validation and test datasets

	Poor	Non-poor
Training set	1258	2199
Cross-validation set	1163	2294
Test set	21902	40314
Total	23323	44807

Table 2 Binary representation of the continuous variable age

Age	Binary representation										
≤ 20	0	0	0	0	0	0	0	0	0	0	0
(20, 25]	0	0	0	0	0	0	0	0	0	0	1
(25, 30]	0	0	0	0	0	0	0	0	0	1	1
(30, 35]	0	0	0	0	0	0	0	0	1	1	1
(35, 40]	0	0	0	0	0	0	0	1	1	1	1
(40, 45]	0	0	0	0	0	0	1	1	1	1	1
(45, 50]	0	0	0	0	0	1	1	1	1	1	1
(50, 55]	0	0	0	0	1	1	1	1	1	1	1
(55, 60]	0	0	0	1	1	1	1	1	1	1	1
(60, 65]	0	0	1	1	1	1	1	1	1	1	1
(65, 70]	0	1	1	1	1	1	1	1	1	1	1
> 70	1	1	1	1	1	1	1	1	1	1	1

3. Education: we represent this variable as a binary string of length 4 as shown in Table 3.

3.3 Predicting poverty using neural networks

We employ the standard feedforward neural networks with one hidden layer (Fig. 1) to distinguish between poor and non-poor households in our data. Let I, H and T be the number of input, hidden and output units, respectively. The number of input units I is equal to the dimensionality of the input data samples, while the number of output units T is usually set equal to the number of classes in the data. The goal of network training is to obtain a set of weights $(\mathbf{W}, \mathbf{V})$ that produces the correct output given the sth input data sample $\mathbf{X_s}$. Here $\mathbf{W}$ is an $(I \times H)$ matrix of weights for the connections between the input units and the hidden units, and $\mathbf{V}$ is an $(H \times T)$ matrix of weights for the connections between the hidden units and the output units. The objective function to be minimized during network training is a function that measures the difference between the correct (actual) output $\mathbf{y_s}$ and the neural network predicted output $\mathbf{p_s}$. A widely used error function is the quadratic error function:

$$E(\mathbf{W}, \mathbf{V}) = \frac{1}{2}\sum_{s=1}^{N}\|\mathbf{y_s} - \mathbf{p_s}\|^2 \qquad (1)$$

Table 3 Binary representation of the categorical variable education

Education level	Binary representation			
Elementary school	0	0	0	0
Secondary school	0	0	0	1
Post-secondary school	0	0	1	1
College graduate	0	1	1	1
Post-graduate	1	1	1	1

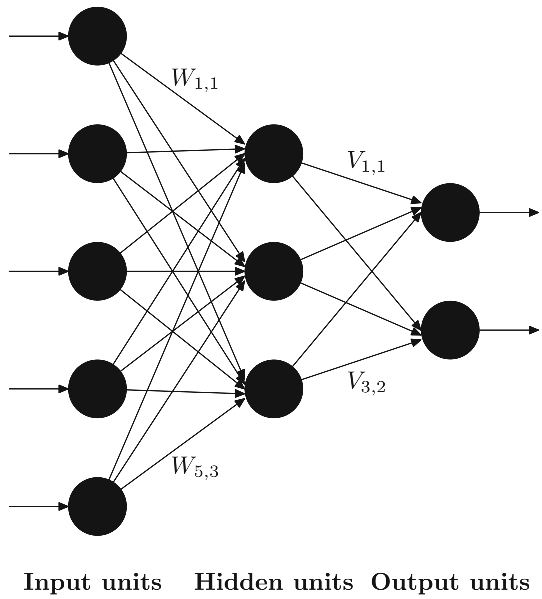

Fig. 1 A three-layer feedforward neural network with 5 input units, 3 hidden units and 2 output units

where N is the number of training data samples, and the norm is the Euclidean norm. The T-dimensional vector of predicted output $\mathbf{p_s}$ is computed as a function of the hidden unit activation values:

$$\mathbf{p_s} = F(\mathbf{a_s}^t\mathbf{V}) \tag{2}$$

where $\mathbf{a_s}$ is the H-dimensional vector of activation values at the network's hidden layer given input $\mathbf{X_s}$ and t indicates the transpose of a vector. The function F computes its output by transforming each of the T components of the vector $\mathbf{a_s}^t\mathbf{V}$ via the sigmoid function:

$$f(\alpha) = 1/(1 + e^{-\alpha}) \tag{3}$$

The computation of the outputs using the sigmoid function (3) results in predicted outputs that have values between 0 and 1. For a dataset where there are samples from T classes, T output units in the network are required and the target output y_s for each sample is a string of 0's except for one 1 indicating the class membership. In our case $T = 2$, we assign the target output (1, 0) for poor households and (0, 1) for non-poor households.

The vector of hidden unit activation values $\mathbf{a_s}$ is similarly computed as the sigmoid of the product between the inputs and the weights from the input units to the hidden units:

$$\mathbf{a_s} = F(\mathbf{X_s}^t\mathbf{W}) \tag{4}$$

The training process starts from an initial random weight $(\mathbf{W_0}, \mathbf{V_0})$, and a sequence of weights $(\mathbf{W_k}, \mathbf{V_k})$ is then computed to minimize the error function (1). Any unconstrained optimization methods can be applied. The error back-propagation method updates the weights by moving along the negative gradient of the error function [15]. It is the most popular method for neural network training; however, this method is quite slow to converge. Other optimization techniques such as the quasi-Newton method and conjugate gradient method have been proposed to speed up convergence. We apply the BFGS method [10] instead of the back-propagation method to find a local minimum of the error function.

Network pruning removes redundant and irrelevant network connections from a trained network. By removing such connections, the generalization capability of the network can be improved. That is, the network predictive accuracy on new data samples is expected to be better when data overfitting is reduced. In order to facilitate the pruning process, an error term that penalizes the use of large weights is added to the error function (1). We add the following penalty term:

$$p(w) = \gamma\left(w^2 + \frac{w^2}{1 + w^2}\right) \tag{5}$$

for each weight w in the network with the penalty parameter γ set to a small positive value. By adding this penalty, network connections that are redundant or irrelevant can be detected by their small magnitude. Such connections can be removed without sacrificing too much the accuracy of the network on the training dataset [33].

Since there are many pieces of information collected in the survey and included as inputs for network training, we expect the majority of these inputs to be not useful for classification. It is important that such inputs can be identified and their network connections can be removed efficiently. Our iterative network pruning algorithm identifies a network connection for potential removal by checking the network predictive accuracy when the connection weight w is set to 0. Starting from the connection from the first input unit to the first hidden unit, we check if it is possible to remove it. This connection is removed if the accuracy of the network is still above a preset minimum threshold. Otherwise, the next possible connection is checked and so on. When there is no more connection that can be removed without suffering too large a drop in the accuracy, the connection that causes the smallest drop in accuracy is removed, and the network is retrained. The process is repeated if the retrained network recovers its accuracy above the threshold; otherwise, the next connection producing the second smallest accuracy drop is removed and the network is retrained. The network pruning process is terminated when there is no remaining connection that can be removed without decreasing the accuracy below the threshold.

We train 30 neural networks starting from different initial random weights $(\mathbf{W_0}, \mathbf{V_0})$. Each value of $(\mathbf{W_0}, \mathbf{V_0})$ is generated randomly in the interval $[-1, 1]$. The number of hidden units H in the networks is set to 3. As the number of poor households is only about 35% of the total households, each data sample in the training set and cross-validation set representing a poor household is duplicated in order to improve the accuracy of the networks in identifying such households. When the training process terminates, i.e., a local minimum of the error function has been found, the accuracy rates of the fully connected networks on the training and cross-validation sets are computed and averaged. This average is then used as the threshold to terminate the network pruning process. Network pruning is stopped when removal of a connection would cause the training set accuracy to drop below this average.

Table 4 summarizes the results. In this table, we show the average and standard deviation from 30 pruned neural networks. The accuracy rate reflects the percentage of samples that are correctly classified. True positives and true negatives are the percentages of poor households and non-poor households that are correctly classified, respectively. Overall, the results indicate that the pruned neural networks can correctly identify more than 7 out of 10 poor households and about 6 out of 10 non-poor households, giving an overall classification accuracy of 65%. In the next subsection, we present the rules that are extracted from one of these pruned networks. The rules provide a clear description of how the households covered in the CBMS survey are determined to be poor or non-poor based on the responses they provided.

3.4 Neural network rule extraction

Algorithms that extract comprehensible *if-else-then* rules from neural networks have been developed in part to address the main criticism about neural networks: While they provide good predictions, they are essentially black boxes. Because the network predictions are computed as a rather complex nonlinear function of the input data, it is difficult to understand how exactly a sample is classified as belonging to one class or the other. Neural network rule extraction algorithms convert this complex input–output mapping of the network into a set of *if-else-then* rules that explicitly states how a sample is classified according to the values of its input attributes.

Table 4 Percentage average accuracy rates, true positives and true negatives of 30 pruned neural networks on the combined training+cross-validation datasets and on the test dataset

	Training + cross-validation	Test
Accuracy rate	67.95 ± 0.36	65.08 ± 1.98
True positives	73.40 ± 4.26	73.12 ± 3.93
True negatives	62.09 ± 5.18	60.71 ± 5.11

We select one of the pruned neural networks for rule extraction. This network has an accuracy rate that is average and has one of the fewest number of network connections. Only 7 input units and 1 hidden unit are still present after pruning. The remaining inputs and their corresponding original input variables are:

- $x_5 = 1$ if and only if Age is older than 50 years.
- $x_{13} = 1$ if and only if Civil status is single.
- $x_{29} = 1$ if and only if education level is post-secondary school or higher.
- $x_{34} = 1$ if and only if there is a family member who is working overseas.
- $x_{36} = 1$ if and only if there is a family member who is a board exam passer.
- $x_{39} = 1$ if and only if the family experienced a calamity in the past year.
- $x_{65} = 1$ if and only if dwelling walls are made of strong material

Poor households and non-poor households differ in their proportions that have a value of $+1$ for the above 7 input variables. For example, among the 2421 poor households in the training dataset, only 324 households (or 13.38%) are headed by someone who is older than 50 years. On the other hand, among the 4493 non-poor households, 979 households (or 21.79%) are headed by someone older that 50 years. Table 5 summarizes the differences in the responses provided by the survey participants between poor and non-poor households in the combined training and cross-validation datasets as well as the test dataset.

A closer inspection of the comparative percentages in the test set, as shown in the table, reveals that the education-related variables figure very differently among poor and non-poor households. For variable x_{29}, we can see that more than 50% of the heads of non-poor households have completed post-secondary school or higher, compared to only 28.69% among the poor households. The same trend can be seen for variables x_{34} and x_{39}. About 5 times more households who are not poor have professional board exam passers among the members of the household or have a family member who is an overseas Filipino worker.

It may be quite surprising that as much as 28.69% of the poor households have heads who have actually completed post-secondary school or higher. This may imply that their training in school does not land them jobs that allow them to earn enough for the family as to provide more than two US dollars per member of the household per day.

The neural network rule extraction algorithm NeuroRule [34] extracts if-then-else rules from neural networks that have only discrete inputs. The steps of NeuroRule are as follows:

Table 5 Summary of the responses provided by survey participants on the 7 relevant inputs

Input i	Poor households			Non-poor households		
	$x_i = 0$	$x_i = 1$	% of $(x_i = 1)$	$x_i = 0$	$x_i = 1$	% of $(x_i = 1)$
Training + cross-validation data						
5	2097	324	13.38	3514	979	21.79
13	2292	129	5.33	3610	883	19.65
29	1693	728	30.07	2094	2399	53.39
34	2373	48	1.98	4048	445	9.90
36	2403	18	0.74	4242	251	5.59
39	2076	345	14.25	3980	513	11.42
65	921	1500	61.96	1107	3386	75.36
Test data						
5	18,895	3007	13.73	31,762	8552	21.21
13	20,747	1155	5.27	32,271	8043	19.95
29	15,619	6283	28.69	19,441	20,873	51.78
34	21,356	546	2.49	36,306	4008	9.94
36	21,678	224	1.02	38,027	2287	5.67
39	18,800	3102	14.16	35,793	4521	11.21
65	8647	13,255	60.52	10,160	30,154	74.80

1. Cluster the hidden unit activation values of samples that are correctly classified by the pruned network.
2. Generate rules that explain the network outputs in terms of the clustered activation values.
3. Generate rules that explain each cluster of activation values in terms of the original attributes of the input data.
4. Merge the two sets of rules generated in Steps 3 and 4 to obtain a set of rules that explain the network outputs in terms of the original attributes of the input data.

As there is only one hidden unit left in the pruned network, Steps 1 and 2 of the algorithm above yield the simple classification rule:

- If $\alpha_s < 0.3765$, then (Output 1 > Output 2),
- Otherwise (Output 1 $\leq$ Output 2).

In the rule above, α_s is the hidden unit activation value of a data sample (c.f. Eq. 4). Step 3 requires that the condition in this rule to be replaced by a rule set involving the inputs that would generate an activation value less than 0.3765. The rule generating method X2R [24] is applied to generate the following rule:

- if $x_5 = 0, x_{13} = 0, x_{29} = 0, x_{34} = 0, x_{36} = 0,$ then $\alpha_s < 0.3765$,
- else if $x_5 = 0, x_{13} = 0, x_{34} = 0, x_{36} = 0, x_{65} = 0,$ then $\alpha_s < 0.3765$,
- else if $x_5 = 0, x_{13} = 0, x_{34} = 0, x_{36} = 0, x_{39} = 1,$ then $\alpha_s < 0.3765$,
- else if $x_{13} = 0, x_{29} = 0, x_{34} = 0, x_{36} = 0, x_{39} = 1,$ then $\alpha_s < 0.3765$,
- else if $x_{13} = 0, x_{29} = 0, x_{34} = 0, x_{65} = 0,$ then $\alpha_s < 0.3765$,
- otherwise, $\alpha_s \geq 0.3765$.

Merging the two sets of rules, we have the final rule set that describes when the network's output 1 is greater than its output 2, i.e., when the network predicts that a data sample belongs to the class of poor households. In terms of the original survey data attributes, the rules are as follows:

Rule 1: $x_5 = 0, x_{13} = 0, x_{29} = 0, x_{34} = 0, x_{36} = 0$ is true when:

- Age is 50 years old or younger
- Civil status is not single
- Education level is secondary school or lower
- There is no family member currently working overseas
- There is no board exam passer in the family

Rule 2: $x_5 = 0, x_{13} = 0, x_{34} = 0, x_{36} = 0, x_{65} = 0$ is true when:

- Age is 50 years old or younger
- Civil status is not single
- There is no family member currently working overseas
- There is no board exam passer in the family
- Dwelling walls are not made of strong material

Rule 3: $x_5 = 0, x_{13} = 0, x_{34} = 0, x_{36} = 0, x_{39} = 1$ is true when:

- Age is 50 years old or younger
- Civil status is not single
- There is no family member currently working overseas
- There is no board exam passer in the family
- The family experienced a calamity in the past year

Rule 4: $x_{13} = 0, x_{29} = 0, x_{34} = 0, x_{36} = 0, x_{39} = 1$ is true when:

- Civil status is not single
- Education level is secondary school or lower
- There is no family member currently working overseas
- There is no board exam passer in the family
- The family experienced a calamity in the past year

Rule 5: $x_{13} = 0, x_{29} = 0, x_{34} = 0, x_{65} = 0$ is true when:

- Civil status is not single
- Education level is secondary school or lower
- There is no family member currently working overseas
- Dwelling walls are not made of strong material

The true positive, true negative and accuracy rates of the above rule on the combined training and cross-validation set are 70.63, 64.43 and 66.60%, respectively. The corresponding figures for the test set are 70.76, 63.19 and 65.85%.

4 Discussion

The extracted rules from the trained neural network are able to capture, in a succinct manner, what truly characterizes poverty in the Philippines—in the specific case of 69,130 households in one of the big cities that comprise the highly and densely populated greater Manila area of more than 12 million people. Even more importantly, the extracted rules reveal which factors seem to correlate highly with lifting households out of poverty. All told, the key insight is that *education* appears to be the most important factor separating poor households from non-poor households.

The education level of the head of the household (x_{29}) is one of seven variables that remain in the pruned network with a predictive accuracy of 65.85%, a true positive rate of as high as 70.63%, and a true negative rate of 63.19%. In three of the five rules that were extracted (Rules 1, 4, 5), households that are determined to be poor tend to have heads of household who have only finished up to high school (secondary school), or even lower. In the Philippines, most high paying jobs require a college/university (also referred to as tertiary level) degree and so many of those who only studied up to high school would not get employed in decent paying jobs.

With the K to 12 reform that has been recently implemented by the Philippine government, a large sector of the high school population would be encouraged to follow a senior high school program that is oriented toward specific competencies and targeted for specific industry sectors that have a large demand for employment. This departs from the current expectation that high school students must go to college in order to be gainfully employed. Part of the reform is to add 2 years (senior high school) to the existing 4 year (junior) high school of the public school system—and the curricula have been revamped to integrate many practical, industry-oriented training to bridge the gap between competencies learned in school and labor demand of industry. By 2018, the first graduates of the K to 12 education reform would have finished senior high school, and it is expected (and hoped) that many of these graduates can land jobs and be able to pull numerous households out of poverty.

The presence of a professional board exam passer among the members of the household (x_{36}) is another education-related variable that remains in the pruned neural network. This variable appears in Rules 2, 3 and 4 and it is also a salient predictor of poverty, but unlike the low education level of the head of household which is an indicator of poor households, the presence of a professional board exam passer is an indicator of a non-poor household. A true product of education, passing a professional board exam to become a licensed teacher, nurse, accountant, chemist, pharmacist, dentist, seaman, engineer, medical doctor or lawyer is such a huge achievement in the Philippines and is generally expected to provide a way out of poverty.

Another input variable that is related to education and is also a predictor for a non-poor household is the presence of a family member who is currently working abroad. Although it can be argued that many of the so-called overseas Filipino workers (OFW), that mainly compose the very large Filipino diaspora in the world, are not necessarily well-educated, education remains a key ingredient for any of these OFWs to be able to land jobs abroad. Preference is always given to college graduates, even for jobs that do not demand such academic qualifications among the locals of the host country. This is why it is normal to have Filipina nannies and domestic helpers who are college graduates, and sometimes are even licensed professionals (e.g., teachers).

As for the other input variables that are retained, we note that a very small set of predictors can already correctly predict 7 out of 10 poor households. These are variables that can be explained only when the social conditions of city households are taken into account. When the head of the household is 50 years old or younger (variable x_5), a variable used in Rules 1, 2 and 3, the household tends to be poor. This might be reflective of the fact that in general, urban poor households would be composed of younger families mostly coming from elsewhere in the country who have tried (and failed) to seek a better life in the metropolis. Perhaps, those families with a head of household who is above 50 years old would tend to be non-poor because they are more settled, with stable jobs, and with children who have been able to go to school and have themselves also been able to succeed in life.

Input variable x_{39}, whether the family has experienced a calamity in the past year, can be easily understood as a predictor for poor households. Indeed, even if the same flooding or typhoon can strike an entire swath of the metropolis, it is really the urban poor households who will more likely experience the calamity. Metro Manila regularly experiences more than 20 typhoons every year, but for those dwellings built with strong materials, these are passing events and not regarded as calamities at all. For shanties, however, a relatively strong typhoon and accompanying floods can be devastating. Other calamities, like a major fire that can engulf an entire neighborhood of hundreds of dwellings in shantytowns, would again tend to afflict more the poor than the non-poor.

Variable x_{13}, the civil status of the head of the household, is the hardest to explain and would need further probes. According to all the rules extracted, from Rules 1 to 5, civil status of *not single* is one predictor of non-poor. In other words, when the head of the household is still single, the household tends to be non-poor. Note that non-single civil status would include legally married, widowed, separated, common law marriage (not married in church). One explanation might be that such heads of households are in general employed or have some stable source of income and have decided not to get married, possibly by choice in order to provide better for the siblings, or by sexual orientation.

Finally, variable x_{65}, whether walls of dwellings are made of strong materials, is naturally associated with non-poor households. It is worth mentioning that there are many other variables, especially those under Groups 3–7, such as *access to toilet* and *access to water source*, that can be generally associated with poor households. Examples are open or closed pits (i.e., holes dug in the ground) that are used as toilets, or river, stream, lake, spring and other bodies of water as source of water. And yet, these variables do not appear among those variables that can help predict poverty. The reason is that these variables have been removed by the network pruning process as the pruning algorithm attempts to remove as many input variables as possible using network's accuracy as the yardstick. In other words, keeping them in the network would not increase the classification accuracy of the network.

5 Conclusion

Poverty is a multi-faceted notion, but based on a definition of absolute poverty, using the per capita income of a household converted to US dollars, households can be generally classified as either poor or not poor. Close to 70,000 households in the city of Pasay, one of the large cities that comprise greater Metro Manila in the Philippines has been the subject of a very detailed household survey under the Community-Based Monitoring System (CBMS) project, funded by the International Development Research Centre (IDRC) of Canada, which is under a larger Partnership for Economic Policy (PEP) program.

Using a multilayered perceptron (MLP), as much as 7 out of 10 poor households are positively identified among more than 62,000 households that are randomly included in the test set. The MLP, a feedforward neural network with one hidden layer, is trained with just a very small subset of the entire dataset. Less than 7000 households are randomly selected to train the neural network, and yet more than 70% of the poor households not encountered during training are positively identified.

What is more remarkable is that only seven variables are needed to predict whether a household is poor or not. And even more importantly, these seven variables are assembled in just five simple rules that were extracted from the neural network, which could help explain and characterize poverty in the Philippines. Three of these variables are related to education, and very clearly, it can be seen that education is one single instrument that can move entire households out of poverty. This echoes similar results reported in other developing countries, such as Ethiopia [3]. In the Philippines, a household tends to be poor when its head has finished only up to high school, or even less. On the other hand, whenever there is a professional board exam passer in the household, or when a member of the family is working abroad as an overseas Filipino worker, the household tends to be not poor.

Acknowledgements The authors wish to thank the Angelo King Institute (AKI) of De La Salle University, under whom the CBMS project is lodged, for sharing the raw survey data from Pasay City.

Compliance with ethical standards

Conflicts of interest The authors declare that they have no conflict of interest.

Appendix

List of variables included in the study to distinguish between poor and non-poor households.

Group 1: Demographic information of the head of the household

1. Gender: 1 = male, 0 = otherwise.
2. Age: continuous.
3. Civil status: single, legally married, widowed, separated, common law marriage, unknown.
4. Religion: 1 = Roman Catholic, 2 = Protestant, 3 = local non-Catholic church, 4 = local Catholic-rite church, 5 = Islam, 6 = others, 7 = no religion, 8 = unknown.
5. Education: 1 = elementary school, 2 = secondary school, 3 = post-secondary, 4 = college graduate, 5 = post-graduate.
6. Literacy: 1 = literate, 0 = otherwise.
7. Job: 1 = currently employed, 0 = otherwise.
8. Phil Health: 1 = member of Philippine Health, 0 = otherwise.

Group 2: Other information about the household

1. Foreign worker in family: 1 = at least one member of the family is currently working overseas, 0 = otherwise.
2. Single parent: 1 = single parent, = 0 otherwise.
3. Board passer: 1 = if there is a board exam passer in the family (e.g., nurse, engineering, architecture), 0 = otherwise.
4. Disable: 1 = if there is a disable family member, 0 = otherwise.
5. Death-indicator: 1 = if there was a death in the family in the past year, 0 = otherwise.
6. Calamity-indicator: 1 = if the family experienced a calamity (e.g., typhoon) in the past year, 0 = otherwise.
7. FoodShortage-indicator: 1 = the family experienced hunger in the past year, 0 = otherwise.
8. Garbage: 1 = if there is an organized (city, town, subdivision) garbage collection, 0 = otherwise.

Group 3: Access to water source (binary variable)

1. Community water system—own use
2. Community water system—shared with other households
3. Artesian deep well—own use
4. Artesian deep well—shared with other households
5. Artesian shallow well—own use
6. Artesian shallow well—shared with other households
7. Dug/shallow well—own use
8. Dug/shallow well—shared with other households
9. River, stream, lake, spring and other bodies of water
10. Bottled water/purified/distilled water
11. Tanker truck/peddler
12. Others

Group 4: Access to water distribution (binary variable)

1. Unknown
2. Within premises
3. Outside premises but 250 m or less
4. Outside premises 251 m or more
5. don't know

Group 5: Access to toilet (binary variable)

1. Water sealed flush to sewerage system/septic tank—own use
2. Water sealed flush to sewerage system/septic tank—shared with other households
3. Closed pit
4. Open pit
5. No toilet
6. Others

Group 6: Dwelling wall (binary variable)

1. Strong materials
2. Light materials
3. Salvaged/makeshift materials
4. Mixed but predominantly strong materials
5. Mixed but predominantly light materials
6. Mixed but predominantly salvaged materials

Group 7: Dwelling roof (binary variable)

1. Strong materials
2. Light materials
3. Salvaged/makeshift materials
4. Mixed but predominantly strong materials
5. Mixed but predominantly light materials
6. Mixed but predominantly salvaged materials

References

1. Achia TBO, Wangombe A, Khadioli N (2010) A logistic regression model to identify key determinants of poverty using demographic and health survey data. Eur J Soc Sci 13(1):38–45
2. Adeli H (2001) Neural networks in civil engineering. Comput-Aided Civil Infrastruct Eng 16(2):126–142

3. Alem Y, Köhlin G, Stage J (2013) The persistence of subjective poverty in urban Ethiopia. World Dev 56:51–61
4. Ansoms A, McKay A (2010) A quantitative analysis of poverty and livelihood profiles: the case of rural Rwanda. Food Policy 35:584–598
5. Babu SC, Sanyal P (2009) Food security, poverty and nutrition policy analysis: statistical methods and applications. Academic Press, Burlington
6. Belhadj B (2011) New fuzzy indices of poverty by distinguishing three levels of poverty. Res Econ 65:221–231
7. Benson T, Chamberlin J, Rhinehart I (2005) An investigation of the spatial determinants of the local prevalence of poverty in rural Malawi. Food Policy 30:532–550
8. Bigman D, Srinivasan PV (2002) Geographical targeting of poverty alleviation programs: methodology and applications in rural India. J Policy Model 24:237–255
9. Chapoto A, Kirimi L, Kadilaya S (2012) Poverty and prime-age mortality in eastern and southern Africa: evidence from Zambia and Kenya. World Dev 40(9):1839–1853
10. Dennis JE Jr, Schnabel RE (1983) Numerical methods for unconstrained optimization and nonlinear equations. Prentice Halls, Englewood Cliffs
11. Duy VQ, D'Haese M, Lemba J, Hau LL, D'Haese L (2011) Determinants of household access to formal credit in the rural areas of the Mekong delta, Vietnam. Afr Asian Stud 11:261–287
12. Erguzel TT, Sayar GH, Tarhan N (2016) Artificial intelligence approach to classify unipolar and bipolar depressive disorders. Neural Comput Appl 27(6):1607–1616
13. Fernández-Gámez MA, Gil-Corral AM, Galán-Valdivieso F (2016) Corporate reputation and market value: evidence with generalized regression neural networks. Expert Syst Appl 46:69–76
14. Gounder R, Xing Z (2012) Impact of education and health on poverty reduction: monetary and non-monetary evidence from Fiji. Econ Model 29:787–794
15. Haykin S (1999) Neural networks: a comprehensive foundation, 2nd edn. Prentice Hall, Englewood Cliffs
16. Hoogeveen JG (2005) Measuring welfare for small but vulnerable groups: poverty and disability in Uganda. J Afr Econ 14(4):603–631
17. Huang W, Lai KK, Nakamori Y, Wang S (2004) Forecasting foreign exchange rates with artificial neural networks: a review. Int J Inf Technol Decis Mak 03(1):145–165
18. Jha GK, Kanchan S (2014) Time-delay neural networks for time series prediction: an application to the monthly wholesale price of oilseeds in India. Neural Comput Appl 24(3):563–571
19. Kalogirou SA (2001) Artificial neural networks in renewable energy systems applications: a review. Renew Sustain Energy Rev 5(4):373–401
20. Kiakojoori S, Khorasani K (2016) Dynamic neural networks for gas turbine engine degradation, health monitoring and prognosis. Neural Comput Appl 27(8):2157–2192
21. Lin S-M (2013) Analysis of service satisfaction in web auction logistics service using a combination of Fruit fly optimization algorithm and general regression neural network. Neural Comput Appl 22(3):783–791
22. Lisboa PJG (2002) A review of evidence of health benefit from artificial neural networks in medical intervention. Neural Netw 15(1):11–39
23. Lisboa PJ, Taktak AFG (2006) The use of artificial neural networks in decision support in cancer: a systematic review. Neural Netw 19(4):408–415
24. Liu H, Tan ST (1995) X2R: a fast rule generator. In: Proceedings of IEEE international conference on systems, man and cybernetics, pp 1631–1635
25. Maharjan KL, Joshi NP (2011) Determinants of household food security in Nepal: a binary logistic regression analysis. J Mt Sci 8:403–413
26. Maier HG, Dandy GC (2000) Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environ Model Softw 15(1):101–124
27. Mitra S (2013) Early warning prediction system for high inflation: an elitist neuro-genetic network model for the Indian economy. Neural Comput Appl 22(S1):447–462
28. Neff D (2013) Fuzzy set theoretic applications in poverty research. Policy Soc 32:319–331
29. Paliwal M, Kumar UA (2009) Neural networks and statistical techniques: a review of applications. Expert Syst Appl 36:2–17
30. Pany PK, Ghoshal SP (2015) Dynamic electricity price forecasting using local linear wavelet neural network. Neural Comput Appl 26(8):2039–2047
31. Ravallion M (1998) Poverty lines in theory and practice. Living Standard measurement study working paper no. 133. World Bank, Washington DC
32. Senf C, Lakes T (2012) Comparing support vector regression and statistical linear regression for predicting poverty incidence in Vietnam. In: Bridging the geographic information sciences, lecture notes in geoinformation and cartography, pp 251–265
33. Setiono R (1997) A penalty-function approach for pruning feedforward neural networks. Neural Comput 9(1):185–204
34. Setiono R, Liu H (1996) Symbolic representation of neural networks. IEEE Comput 29(3):71–77
35. Sy I (2013) The subjective approach as a tool for understanding: The case of Senegal. Proc Econ Financ 5:336–345
36. Vellido V, Lisboa PJG, Vaughn J (1999) Neural networks in business: a survey of applications. Expert Syst Appl 17(1):51–70
37. Woolard I, Klasen S (2005) Determinants of income mobility and poverty dynamics in South Africa. J Dev Stud 41(5):865–897
38. Yaseen ZM, El-Shafie A, Afan HA, Hameed M (2016) RBFNN versus FFNN for daily river flow forecasting at Johor River, Malaysia. Neural Comput Appl 27(6):1533–1542
39. Yazici B, Memmedli M, Aslanargun A, Sama S (2010) Analysis of international debt problem using artificial neural networks and statistical methods. Neural Comput Appl 19(8):1207–1216
40. Zhang Z, Friedrich K (2003) Artificial neural networks applications to polymer composites: a review. Compos Sci Technol 63(14):2029–2044