IOSR Journal of Mathematics (IOSR-JM)
e-ISSN: 2278-5728, p-ISSN: 2319-765X. Volume 12, Issue 5 Ver. VI (Sep. - Oct.2016), PP 26-31
www.iosrjournals.org

A Comparison Between Logistic Regression And Neural Networks In Modeling Mortality Amongst Children Under Five Years In Ghana

Salifu Nanga[1], Dr. Anani Lotsi[2]
[1]*(SANA Research and Consultancy Ghana Limited)*
[2]*(Department of Statistics, University of Ghana)*

Abstract: *Child mortality is regarded as one of the most revealing measures of society's ability to meet the needs of its people. The Millennium Development Goal 4 (MDG 4) advocates a reduction of under-five mortality rate by two-thirds between 1990 and 2015. The main objective of this study was to develop a validated set of statistical models and select the most appropriate model between logistic regression and Neural Networks to predict mortality among children under five and to compare the influence of selected risk factors on the probability of death before the age of 5 years among children in Ghana. The study revealed that the Logistic Regression model was the most efficient in modeling Mortality in Children under five with a CCR of 81% and Neural Network will also do a good job at predicting mortality in children under five with a CCR of 80%. The highest educational level of mother, Age of mother at birth, Type of toilet facility used by family, alcohol consumption and the wealth index of family were discovered as the most important variables in predicting mortality amongst children under five in Ghana across both models.*
Keywords: *Logistic Regression, Neural Networks, Children under five years,*

I. Introduction

One of the most revealing measures of how well a society is performing in meeting the needs of its people is Child mortality (Iram, 2008). Millennium Development Goal 4 (MDG 4) is aimed at reducing under-five mortality rate by two-thirds between 1990 and 2015 (Unicef, 2014).

The past two decades has seen an overwhelming increase in under-five deaths; between the period 1990 and 2013, 223 million children died before age five. There has been a 49 percent decline in the under-five mortality rate globally since 1990.This statistic however is still far below the two-thirds reduction targeted to reach the Millennium Development Goal 4(Unicef, 2014).

Various factors has been empirically proven in previous research to influence a child's health and survival, including place of residence, breastfeeding, place of delivery, access to postnatal care, maternal age and education (Doctor HV, 2011). Variations in child mortality rates has been linked to geographic differences in maternal literacy levels and sociocultural practices within countries (Black et al, 2003). Women who deliver at health facilities tend to have a lower probability of reporting child death as compared to those delivering in home settings (Doctor HV, 2011). Postnatal care access has been associated with a drop in under-five mortality, with a study carried out in Bangladesh showing that postnatal home visits within the first 2 days after birth by skilled healthcare workers was significantly related with a lower likelihood of child death (Baqui, 2009). Maternal education and age has been found to important determinants of child mortality in a study conducted in the developing countries (Deribew et al., 2007). Evidence also shows that child mortality rates is found to be higher among less educated mothers as compared to mothers who have higher levels of education (You, 2011). The importance of education, mother's education in particular, has been confirmed in many subsequent studies (Murthi et al., 1995; Dre`ze and Murthi, 2001).To the extent that, education is able to improve an individual's ability to undertake these changes, more educated mothers are expected to have healthier babies (Iram et al,2008). Mother's employment status has been empirically proven as a significant factor that affects neonatal, infant and child mortality (Arriaga and Hobbs, 1982). The work status of mother determines the amount of time and care she can give to her child, and it may determine the amount of income available to the mother and hence her access to various goods and services.

An advanced data analysis technique used to explore large data sets to identify useful and unexpected patterns and rules that provide relevant knowledge for predicting future outcomes is referred to as data mining (Lungu and Bâra, 2012). The term neural network is a data mining technique that applies to a loosely related family of models, which is characterized by a parameter space that is large in nature and flexible in structure, descending from studies of brain functioning. Logistic regression is also a mathematical modelling approach that can be used to describe the relationship between several independent variables to a dichotomous dependent variable. Yeh proposed the use of artificial neural networks algorithm to model heart beats classification (Yeh, 2012). He first selected in the proposed model, the best features using dynamic programming; then, applying artificial neural networks, heart beats were successfully classified into seven categories. Rajendra et al.,(2011) employed artificial neural networks and Fuzzy equivalence classifier to classify mortality amongst under five. Results from Liang and Liu (1996) and West, Brockett and Golden (1997) also revealed that in most of the applications where neural networks have been used to model business problems in support of finance and marketing decision-making, it outperformed traditional compensatory models such as discriminant and regression analysis. Cervical cancer risk classification was modelled, using artificial neural networks (Qiu et al. (2007). The results revealed a sensitivity and specifity values of 98% and 97%, respectively. Artificial neural network methods was used for prediction of late onset heart failure (Salari et al., 2013). Wilson et al., (2003), looked at the predictive capability of bankruptcy in firms using neural networks and classical multivariate discriminant analysis. Wiginton (1980) was one of the first people who first published accounts of logistic regression applied to credit scoring in comparison to discriminant analysis and drew a conclusion that logistic regression gave a superior result. The purpose of this study was to assess what factors are most important in determining mortality in children under five years in Ghana by comparing Logistic Regression which is a statistical technique and Neural Networks a data mining technique.

II. Materials And Methods

The study was undertaken using the 2008 Ghana Demographic and Health Survey (GDHS). A sample of 11,888 respondents were considered in the study .A logistic Regression model and Neural Network model was applied to the dataset. The variables considered in the study was whether a child was alive or not as the dependent variable. The explanatory variables were Type of place of residence of family, Highest educational level of mother, Source of drinking water, type of toilet facility used by family, Wealth index of family, alcohol consumption of mother, Whether mother is covered by health insurance, Marital status of mother, Sex of child and age of mother at child birth.

Logistic Regression

Logistic regression is a mathematical modelling approach that can be used to describe the relationship of several independent variables to a dichotomous dependent variable. Logistic regression allows the prediction of discrete variables by a mix of continuous and discrete predictors.

Letting Y be the binary response variable, it is assumed that $P(Y=1)$ is possibly dependent on $\bar{x}$, a vector of predictor values. The goal is to model

$$p(\bar{x}) \equiv P(Y=1 \mid \bar{x}) .$$

Logistic Regression –With Dichotomous Responses and numeric and/or categorical explanatory variable(s). Model the probability of a particular as a function of the predictor variable(s) Probabilities are bounded between 0 and 1

$$\pi = \frac{e^{\alpha+\beta_1 x_1+\cdots+\beta_k x_k}}{1+e^{\alpha+\beta_1 x_1+\cdots+\beta_k x_k}}$$

· The P=primary interest in estimating and testing hypotheses regarding Large-Sample test is the Wald Test:

· H_0: $b = 0$ H_A: $b \neq 0$

$$T.S.: X^2_{obs} = \left(\frac{\hat{\beta}}{\hat{\sigma}_{\hat{\beta}}} \right)^2$$

$$R.R.: X^2_{obs} \geq \chi^2_{\alpha,1}$$

$$P-val: P(\chi^2 \geq X^2_{obs})$$

Odds Ratio

$$\frac{odds(x+1)}{odds(x)} = e^{\beta} \qquad \left(odds(x) = \frac{\pi(x)}{1-\pi(x)} \right)$$

Thus e^b shows the change in the odds of the outcome (multiplicatively) by an increase in x by 1 unit

When $b = 0$, the odds and probability are the same at all x levels ($e^b=1$)

. when $b > 0$, the odds and probability increase as x increases ($e^b>1$)

When $b < 0$, the odds and probability decrease as x increases ($e^b<1$)

Testing Regression Coefficients

$$H_0 : \beta_1 = \cdots = \beta_k = 0$$
$$H_A : \text{Not all } \beta_i = 0$$
$$T.S.\ X^2_{obs} = (-2\log(L_0)) - (-2\log(L_1))$$
$$R.R.\ X^2_{obs} \geq \chi^2_{\alpha,k}$$
$$P = P(\chi^2 \geq X^2_{obs})$$

III. Neural Network

A neural network is referred to as a software (or hardware) simulation of a biological brain. It is sometimes called Artificial Neural Network or 'ANN'. A typical neuron in the human brain, gathers signals from others through a host of fine structures called *dendrites*. It sends out spikes of electrical activity which goes through a long thin strand called an *axon*, which in turn splits into thousands of branches. A structure called *synapse* at the end of each branch converts the activities of from the axon into electrical signals which inhibits or excites activity within the connected neurons. The receipt of excitatory input from a neuron that is sufficiently large compared with its inhibitory input, sends a spike of electrical activity down its axon. Learning then occurs by effectively changing the synapses so that the influence of one neuron on another changes. The neuron input value is taken by an activation function which takes and produces a value that becomes the output value for the neuron and it passes through to other neurons in the network. This process is called the multilayer perceptron (MLP). The number of parameters within a MLP with one hidden layer with *h* neurons and *k* inputs is given as;

$$h(k+1)+h+1 = h(k+2)+1$$

When the weights are adjusted on the connections between layers, the perceptron output is "trained" to match a preferred output. The Weights (W_{ij}) can be determined by adding to the old weight an error correction value. By multiplying the difference between the actual output $(\mathrm{x}[\mathrm{j}])$ and target $(\mathrm{t}[\mathrm{j}])$ values by a learning rateconstant C, the amount of correction is determined. The connection weight is adjusted, if the input node output $(\mathrm{a}[\mathrm{j}])$ is a 1, and if it sends 0, it has no effect on the output and therefore no need for adjustment. The process can be represented as:

$$W_{ij(new)} = W_{ij(old)} + C(t_j - X_j)a_i$$

Where, C = learning rate. The training process is repeated until the performance of the network no longer improves. Training performance may be increased by a large number of hidden nodes may increase training, but at the expense of generalizing and computing cost. A number of hidden nodes were experimented to check performance the nodes in a layer were chosen. The initial weights are randomly selected by the software and the error reduction at a convergence tolerance of 0.0001 is used to obtain the best weights. The learning rate was set at 0.001 and the weight decay at 10. Double sigmoid function was activation function shown below:

$$F(sum_j) = \frac{w_1}{1+\exp(sum_j)} + \frac{w_2}{1+\exp(sum_j)}$$

IV. Results And Discussion

4.1 Logistic Regression

Table 1 displays a summary of the model we see that the -2 Log Likelihood statistic is 10768.032. This statistic measures how poorly the model predicts the decisions, the smaller the statistic the better the model. A more useful measure to assess the utility of a logistic regression model is classification accuracy. From Table 2 the overall classification accuracy of the model was 81%.

Table 1 : Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	10768.032	.020	.034

Table 2 : Classification Table

			Child is alive?		Percentage Correct
			Yes	No	
Step 1	Child is alive?	Yes	8187	1566	83.94
		No	683	1400	67.21
	Overall Percentage				81.00

Table 3 displays the results for the omnibus tests of model coefficients. The presence of a relationship between the dependent variable and combination of independent variables is based on the statistical significance of the model chi-square at step 1 after the independent variables have been added to the analysis. In this analysis, the probability of the model chi-square (244.581) was < 0.001, less than or equal to the level of significance of 0.05. The existence of a relationship between the independent variables and the dependent variable was therefore supported

Table 3 : Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	244.581	15	.000
	Block	244.581	15	.000
	Model	244.581	15	.000

Table 4 : Variables in Logistic Regression Model

Variables in the Equation						
	β	**S.E.**	**Wald**	**df**	**P-value**	**Odds**
Place_Of_Residence(1)	.050	.073	.475	1	.491	1.051
Educational_Level			25.751	3	.000	
Educational_Level(1)	.515	.007	4.726	1	.030	1.674
Educational_Level(2)	.634	.009	7.183	1	.007	1.501
Educational_Level(3)	.312	.024	1.797	1	.180	1.366
Source_of_drinking_water(1)	.129	.065	4.002	1	.045	1.138
Toilet_facility			57.315	2	.000	
Toilet_facility(1)	.617	.122	25.431	1	.000	1.854
Toilet_facility(2)	.457	.062	54.250	1	.000	1.579
Wealth_index			30.675	2	.000	
Wealth_index(1)	-.492	.089	30.321	1	.000	.612
Wealth_index(2)	-.244	.075	10.666	1	.001	.784
Age_of_mother	.031	.003	82.199	1	.000	1.031
health_insurance(1)	.136	.052	6.952	1	.008	1.146
Marital_status			.949	2	.622	
Marital_status(1)	-.190	.234	.659	1	.417	.827
Marital_status(2)	-.056	.074	.565	1	.452	.946
Sex_of_child(1)	.180	.049	13.471	1	.000	1.197
alchohol_consumption(1)	.195	.060	10.498	1	.001	1.215
Constant	-3.625	.292	153.985	1	.000	.027

The coefficient of any variable is deemed to be significant if P value ≤ 0.05.Hence from Table 4 the significant values in the model are Mothers educational level, Age of mother at birth, Type of toilet facility used, Wealth index of family, whether mother is registered with health insurance, Sex of child and alcohol consumption.

A standard error larger than 2.0 indicates numerical problems, such as multicollinearity among the independent variables. From Table 4 we can deduce that none of the independent variables had a standard error greater than 2.0 hence there is no evidence of multicollinearity.

Table 5 : Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	3.229	8	.919

Table 5 shows the Hosmer and Lemeshow Test , since the p – value, 0.919, is greater than the significance level, $\alpha = 0.05$, we conclude that there is enough evidence to show that the hypothesized model fits the data set used in predicting mortality among under five children in Ghana.

2.1 Neural Network

The training sample was assigned 69.8% of cases in the dataset in order to obtain a model. The testing sample was assigned 20.5% of cases in the dataset to track errors during training and the holdout sample was assigned 9.7% of cases to give an "honest" estimate of the ability of the model to predict. The error computed across the three samples were approximately the same.This signifies that the model is a good fit in predicting child mortality.

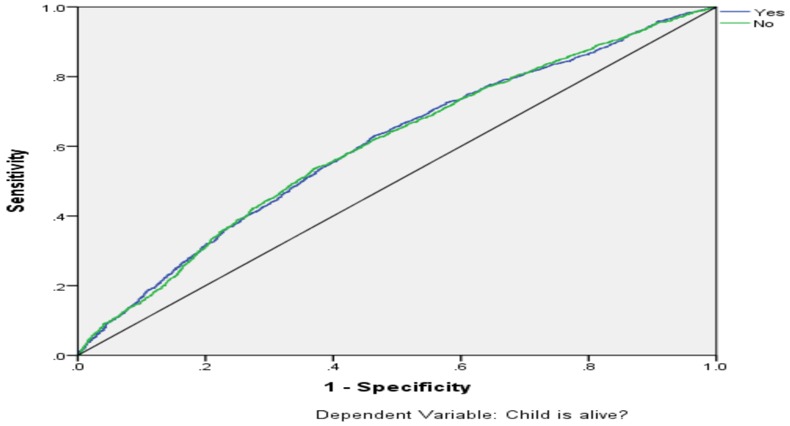

Figure 1 : ROC Curve

Table 6 : Area under Curve

		Area
Child is alive?	Yes	.701
	No	.701

Table 6 shows a numerical summary of the ROC for a randomly selected child who is alive or a randomly selected child who is dead, there is a 0.701 probability that the model-predicted pseudo-probability will be higher for a child who is alive than for a child who is dead.

Table 7 : Independent Variable Importance

	Importance	Normalized Importance
Age of mother	0.304	100.00%
Highest educational level	0.17	55.80%
Type of toilet facility	0.121	39.60%
Alcohol consumption	0.084	27.50%
Wealth index	0.073	24.00%
Source of drinking water	0.068	22.30%
Covered by health insurance	0.061	20.20%
Type of place of residence	0.058	19.20%
Sex of child	0.035	11.40%
Marital status	0.027	8.70%

Table 7 displays the variable importance table. It appears that characteristics of the mother(age of mother, mothers' educational level and mothers alcohol consumption) and sanitation issues (type of toilet facility) have the greatest effect on how the network classifies whether a child live or die; what neural network cannot reveal is the "direction" of the relationship between these variables.

2.2 Comparison of Models

The Correct Classification Rate (CCR) was used to compare the overall classification accuracy of the both models. Table 8 revealed that Logistic Regression was the most efficient in modeling Mortality in Children under five with a CCR of 81% and Neural Network with a CCR of 80%.It is evident that the difference in the CCR amongst the models was quite insignificant. It can therefore be concluded that both models will also do a good job in predicting mortality in children under five years in Ghana.

Table 8 : Correct Classification rate Table

		Actual +	**Actual -**	**CCR**
Logistic Regression	**Predicted +**	8187	1566	**0.81**
	Predicted -	683	1400	
Neural Network	**Predicted +**	8070	1683	**0.80**
	Predicted -	726	1358	
	Predicted -	536	1547	

V. Conclusion

The study sought to compare two different models in the prediction of mortality in children under five years. The models were Logistic Regression and Neural Network. The study revealed that both Logistic regression and Neural Network will also do a good job in predicting mortality in children. The study also sought to identify the predictor variables that were most significant or important in the models. Taking a cursory look at the models, five variables cut across as the most significant or important in predicting mortality in amongst children under five in Ghana. Mothers' highest educational level, Age of mother at birth, Type of toilet facility used, alcohol consumption and wealth index were found to be the most important variables in predicting mortality in amongst children under five in Ghana across all models.

References

[1]. Bâra, A., & Lungu, I. (2012). Improving Decision Support Systems with Data Mining Techniques. INTECH Open Access Publisher.

[2]. Black, R. E., Morris, S. S., & Bryce, J. (2003). Where and why are 10 million children dying every year?. The lancet, 361(9376), 2226-2234.

[3]. Breiman, L. (1996). Bagging predictors. Machine learning, 24(2), 123-140.

[4]. Doctor, H. V., Bairagi, R., Findley, S. E., Helleringer, S., & Dahiru, T. (2011). Northern Nigeria maternal, newborn and child health programme: selected analyses from population-based baseline survey. The Open Demography Journal, 4(11-12), 11.

[5]. Drèze, J., & Murthi, M. (2001). Fertility, education, and development: evidence from India. Population and development review, 27(1), 33-63.

[6]. Hand, D. J., & Henley, W. E. (1997). Statistical classification methods in consumer credit scoring: a review. Journal of the Royal Statistical Society: Series A (Statistics in Society), 160(3), 523-541

[7]. Huang, R., Xi, L., Li, X., Liu, C. R., Qiu, H., & Lee, J. (2007). Residual life predictions for ball bearings based on self-organizing map and back propagation neural network methods. Mechanical Systems and Signal Processing, 21(1), 193-207.

[8]. Iram, U., & Butt, M. S. (2008). Socioeconomic determinants of child mortality in Pakistan: Evidence from sequential probit model. International Journal of Social Economics, 35(1/2), 63-76.

[9]. Johnson, P. D., Hobbs, F., & Arriaga, E. E. (1982). Techniques for Estimating Infant Mortality (No. 8). US Department of Commerce, Bureau of the Census

[10]. Liang, G. E., Liu, D., Woolsey, T. A., Rovainen, C. M., Cox, S. B., Henegar, M. H.,... & Wei, L. (1996). Neuronal units linked to microvascular modules in cerebral cortex: response elements for imaging the brain. Cerebral Cortex, 6(5), 647-660.

[11]. Murthi, M., Guio, A. C., & Dreze, J. (1995). Mortality, fertility, and gender bias in India: A district-level analysis. Population and development review, 745-782.

[12]. Thatte, N., Kalter, H. D., Baqui, A. H., Williams, E. M., & Darmstadt, G. L. (2009). Ascertaining causes of neonatal deaths using verbal autopsy: current methods and challenges. Journal of Perinatology, 29(3), 187-194.

[13]. Weinberger, K. Q., & Saul, L. K. (2009). Distance metric learning for large margin nearest neighbor classification. Journal of Machine Learning Research, 10(Feb), 207-244.

[14]. West, P. M., Brockett, P. L., & Golden, L. L. (1997). A comparative analysis of neural networks and statistical methods for predicting consumer choice. Marketing Science, 16(4), 370-391.

[15]. Wiginton, J. C. (1980). A note on the comparison of logit and discriminant models of consumer credit behavior. Journal of Financial and Quantitative Analysis, 15(03), 757-770.

[16]. Wilson, G. F., & Russell, C. A. (2003). Real-time assessment of mental workload using psychophysiological measures and artificial neural networks. Human Factors: The Journal of the Human Factors and Ergonomics Society, 45(4), 635-644.

[17]. Rajendra, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

[18]. Salari, P Barylski, M., Nikfar, S., Mikhailidis, D. P., Toth, P. P.,., Ray, K. K., ... & Nicholls, S. J. (2013). Statins decrease all-cause mortality only in CKD patients not requiring dialysis therapy—a meta-analysis of 11 randomized controlled trials involving 21,295 participants. Pharmacological research, 72, 35-44.

[19]. World Health Organization, & UNICEF. (2014). Trends in maternal mortality: 1990 to 2013: estimates by WHO, UNICEF, UNFPA, The World Bank and the United Nations Population Division: executive summary.

[20]. Yeh, Y. C., Chiou, C. W., & Lin, H. J. (2012). Analyzing ECG for cardiac arrhythmia using cluster analysis. Expert Systems with Applications, 39(1), 1000-1010.

[21]. You, D., New, J. R., & Wardlaw, T. (2011). Levels and trends in child mortality. Report 2012. Estimates developed by the UN Inter-agency Group for Child Mortality Estimation.