CSCanada
Progress in Applied Mathematics
Vol. 3, No. 2, 2012, pp. 28-34
DOI: 10.3968/j.pam.1925252820120302.1255

ISSN 1925-251X [Print]
ISSN 1925-2528 [Online]
www.cscanada.net
www.cscanada.org

On the Power Efficiency of Artificial Neural Network (ANN) and the Classical Regression Model

Chukwu, A. U[1,*]; Adepoju K. A[1]

[1]Department of Statistics, University of Ibadan, Ibadan, Nigeria.
[*]Corresponding author.
Address: Department of Statistics, University of Ibadan, Ibadan, Nigeria.

Received December 2, 2011; accepted April 20, 2012

Abstract

This research work presents new development in the field of natural science, where comparison is made theoretically on the efficiency of both classical regression models and that of artificial neural network models, with various transfer functions without data consideration. The results obtained based on variance estimation indicates that ANN is better which coincides with the results of Authors in the past on the efficiency of ANN over the traditional regression models. The certain conditions required for ANN efficiency over the conventional regression models were noted only that the optimal number of hidden layers and neurons needed to achieve minimum error is still open to further investigation.

Key words

Artificial neural network models; Transfer functions; Hidden layers; Regression models

Chukwu, A. U, Adepoju K. A (2012). On the Power Efficiency of Artificial Neural Network (ANN) and the Classical Regression Model. *Progress in Applied Mathematics*, *3*(2), 28-34. Available from: URL: http://www.cscanada.net/index.php/pam/article/view/j.pam.1925252820120302.1255 DOI: http://dx.doi.org/10.3968/j.pam.1925252820120302.1255

1. INTRODUCTION

Neural networks are being widely used in many fields of study. This can be attributed to the fact that these networks are attempts to model the capabilities of human brains. Since the last decade, neural networks have been used as a theoretically sound alternative to traditional regression models.

Although neural networks (NNs) originated in mathematical neurology, the rather simplified practical models currently in use have moved steadily towards the field of statistics. A number of researchers have illustrated the connection of neural networks to traditional statistical methods. For example Gallinari, Thiria, Badran and Fogelman-Soullie (1991) have presented analytical results that establish a link between discriminant analysis and multilayer perceptions (MLP) used for classification problems. Cheng and Titterington ((1994) made a detailed analysis and comparison of various neural network models with traditional statistical methods.

Neural networks are being used in the areas of prediction and classification, areas where regression models and other related statistical techniques have traditionally been used. Ripley (1994) discusses the statistical aspects of neural networks and classifies neural networks as one of a class of flexible non-linear regression methods.

2. AIM AND OBJECTIVES OF THE STUDY

The main purpose of this study is to compare efficiency of traditional regression model and that of artificial neural network models with an attempt to recognize the one with better discriminating and predictive power.

For the realization of the above intention, the following measures are the underlined objectives:

(i) To compare theoretically a ANN model with a logistic transfer function and a logistic regression model.

(ii) To compare analytically the ANN model with linear transfer function and linear regression model.

(iii) To estimate the means of the estimates of parameters of various classes of non linear regression and that of the ANN.

(iv) To estimate and compare variances of the parameters of both ANN and that of traditional regression model.

(v) To identify a better model based on the result obtained from the comparison.

3. SIGNIFICANCE OF THE STUDY

Methodological disputes that arise in practice often turn on questions of the nature interpretation and justifications of methods and models that relieved on to learn from incomplete and often "observational" (or non experimental) data, the methodology of statistical inference and statistical modeling. This research work is of very high significance as it attempts to unravel the truth and settle the scores of methodological disputes in the field of mathematical statistics, Accounting and finance, Health and medicine, Engineering and Manufacturing, Marketing and general body of knowledge with regards to using classical regression and artificial neural modeling.

As the performance of a particular technique in comparison to other technique depend on various factors like the size of the sample, among others, in this study, attempt is comparing both techniques (ANN and logistic models) analytically without data consideration.

4. METHODOLOGY

The purpose of the study is to have the theoretical explanation of both techniques, which include their variance analysis. The classical regression model includes:

(i) Log-linear model

$lnY = \alpha_0 + \alpha_1 x + e_i$

(ii) Linear-log model

$Y = \alpha_0 + \alpha_1 lnx + e_i$

(iii) $Y_{LR} = \hat{\alpha}_0 + \hat{\alpha}_1 (1 + e^{-x})^{-1}$ Logistic regression

The ANN class of models

(i) $Y = \alpha_0 + \alpha_1 \left[1 + e^{-(\gamma_0 + \gamma_{1x})}\right]^{-1} + e_i$ Logistic transfer function

(ii) $Y = \alpha_0 + \alpha_1 \left[\frac{e^{\gamma_0 + \gamma_{1x}} - e^{-(\gamma_0 + \gamma_{1x})}}{e^{\gamma_0 + \gamma_{1x}} + e^{-(\gamma_0 + \gamma_{1x})}}\right] + e_i$ Hyperbolic transfer function

$Y_{LNN} = \hat{\alpha}_0 + \hat{\alpha}_1 \hat{\gamma}_0 + \hat{\alpha}_1 \hat{\gamma}_1 x$ linear transfer function

5. COMPARATIVE ANALYSIS OF THE VARIANCES OF TRADITIONAL REGRESSION MODEL AND ANN MODELS

5.1 Linear-log Versus ANN with Logistic as Transfer Function

In the section, the variances of the model under review will be compared.

Let Y_{LG} represents a lines-log model and Y_{LANN} denotes variable of a ANN with logistic as transfer function.

$$Var\,[Y_{LG}] = Var\,[(Y_{LG} - Y_{LANN}) + Y_{LANN}]$$

Recall that *Var(Z+M) = Var(Z) + Var(M) + 2Cov(Z, M)*

Let

$$Z = Y_{LG} - Y_{LANN} \text{ and } M = Y_{LANN}$$

$$Var\,[Y_{LG}] = Var\,(Y_{LG} - Y_{LANN}) + Var\,(Y_{LANN}) + 2Cov\,(Y_{LG} - Y_{LANN})\,(Y_{LANN})$$

Since $Y_{LG} = \hat{\alpha}_0 + \hat{\alpha}_1 \ln x$ and $Y_{LANN} = \hat{\alpha}_0 + \hat{\alpha}_1 \left(1 + e^{-(\gamma_0+\gamma_{1x})}\right)^{-1}$

We have

$$Var\,(Y_{LG}) = Var\left[\hat{\alpha}_0 + \hat{\alpha}_1 \ln x - \hat{\alpha}_0 - \hat{\alpha}_1 \left(1 + e^{-(\gamma_0+\gamma_{1x})}\right)^{-1}\right] + Var\left[\hat{\alpha}_0 + \hat{\alpha}_1 \left(1 + e^{-(\gamma_0+\gamma_{1x})}\right)^{-1}\right] + 2Cov\left[\hat{\alpha}_0 + \hat{\alpha}_1 \ln x, \hat{\alpha}_0 + \hat{\alpha}_1 \left(1 + e^{-(\gamma_0+\gamma_{1x})}\right)^{-1}\right]$$

$$Var\,(Y_{LG}) = Var\,(Y_{LANN}) + Var\left[\hat{\alpha}_1 \left(\ln x - \hat{\alpha}_0 - \hat{\alpha}_1 \left(1 + e^{-(\gamma_0+\gamma_{1x})}\right)^{-1}\right)\right] + 2Cov\left[\hat{\alpha}_0 + \hat{\alpha}_1, \ln x - \hat{\alpha}_0 + \hat{\alpha}_1 \left(1 + e^{-(\gamma_0+\gamma_{1x})}\right)^{-1}\right]$$

Considering the nature of independency of the variables, one expect *Cov(Z,M) = 0*

$$\Rightarrow \quad E\,(ZM) = E\,(Z) \cdot E\,(M)$$

$$E\left[(\hat{\alpha}_0 + \hat{\alpha}_1 \ln x)\left(\hat{\alpha}_0 + \hat{\alpha}_1 \left(1 + e^{-(\gamma_0+\gamma_{1x})}\right)^{-1}\right)\right] = (\alpha_0 + \alpha_1 \ln x)\left[\alpha_0 + \alpha_1 \left(1 + e^{-(\gamma_0+\gamma_{1x})}\right)^{-1}\right]$$

$$E\,(\hat{\alpha}_0 + \hat{\alpha}_1 \ln x) = \alpha_0 + \alpha_1 \ln x \text{ since } E\,(\hat{\alpha}_0) = \alpha_0$$

$$E\left[\hat{\alpha}_0 + \hat{\alpha}_1 \left(1 + e^{-(\gamma_0+\gamma_{1x})}\right)^{-1}\right] \text{ and } = \alpha_0 + \alpha_1 \left(1 + e^{-(\gamma_0+\gamma_{1x})}\right) \quad E\,(\hat{\alpha}_1) = \alpha_1$$

Therefore

$$Cov(ZM) = 0$$

$$Var\,(Y_{LG}) = Var\,(Y_{LANN}) + Var\left[\hat{\alpha}_1 \left(\ln x - \left(1 + e^{-(\gamma_0+\gamma_{1x})}\right)^{-1}\right)\right]$$

$$Var\,(Y_{LG}) = Var\,(Y_{LANN}) + \left[\ln x - \left(1 + e^{-(\gamma_0+\gamma_{1x})}\right)^{-1}\right]^2 Var\,(\hat{\alpha}_1)$$

$Var\,(\hat{\alpha}_1)$ is non negative function

$[\ln x - (1 + e^{-(\gamma_0+\gamma_{1x})})^{-1}]^2$ is also a non-negative function for every x.

Therefore,

$$Var\,(Y_{LG}) \geq Var\,(Y_{HANN})$$

5.2 Linear-log versus *ANN* with Hyperbolic as transfer function

If Y_{LG} and Y_{LANN} denote linear-log and *ANN* with hyperbolic as transfer function respectively, the variance of linear-log is as follows

$$Var\left[Y_{LG}\right] = Var\left[\left(Y_{LG} - Y_{HANN}\right) + Y_{HANN}\right]$$

$$Var\left(Y_{LG}\right) = Var\left(Y_{LG} - Y_{HANN}\right) + V\left(Y_{ANN}\right) + 2Cov\left[\left(Y_{LG} - Y_{HANN}\right), Y_{ANN}\right]$$

With $Y_{LG} = \hat{\alpha}_0 + \hat{\alpha}_1 \ln x$ and

$$Y_{HANN} = \hat{\alpha}_0 + \hat{\alpha}_1 \left[\frac{e^{\gamma_0+\gamma_{1x}} - e^{-(\gamma_0+\gamma_{1x})}}{e^{\gamma_0+\gamma_{1x}} + e^{-(\gamma_0+\gamma_{1x})}}\right]_i$$

$$\begin{aligned} Var\left(Y_{LG}\right) = Var\left[\hat{\alpha}_0 + \hat{\alpha}_1 \ln x - \hat{\alpha}_0 - \hat{\alpha}_1\left[\frac{e^{\gamma_0+\gamma_{1x}} - e^{-(\gamma_0+\gamma_{1x})}}{e^{\gamma_0+\gamma_{1x}} + e^{-(\gamma_0+\gamma_{1x})}}\right]\right] + V\left[\hat{\alpha}_0 + \hat{\alpha}_1\left[\frac{e^{\gamma_0+\gamma_{1x}} - e^{-(\gamma_0+\gamma_{1x})}}{e^{\gamma_0+\gamma_{1x}} + e^{-(\gamma_0+\gamma_{1x})}}\right]\right] \\ + 2Cov\left[\hat{\alpha}_0 + \hat{\alpha}_1 \ln x\hat{\alpha}_0 - \hat{\alpha}_1\left[\frac{e^{\gamma_0+\gamma_{1x}} - e^{-(\gamma_0+\gamma_{1x})}}{e^{\gamma_0+\gamma_{1x}} + e^{-(\gamma_0+\gamma_{1x})}}\right], \hat{\alpha}_0 + \hat{\alpha}_1\left[\frac{e^{\gamma_0+\gamma_{1x}} - e^{-(\gamma_0+\gamma_{1x})}}{e^{\gamma_0+\gamma_{1x}} + e^{-(\gamma_0+\gamma_{1x})}}\right]\right] \end{aligned}$$

$$\begin{aligned} Var\left(Y_{LG}\right) = Var\left[\hat{\alpha}_1\left(\ln x - \frac{e^{\gamma_0+\gamma_{1x}} - e^{-(\gamma_0+\gamma_{1x})}}{e^{\gamma_0+\gamma_{1x}} + e^{-(\gamma_0+\gamma_{1x})}}\right)\right] + Var\left[\hat{\alpha}_0 + \hat{\alpha}_1\left(\frac{e^{\gamma_0+\gamma_{1x}} - e^{-(\gamma_0+\gamma_{1x})}}{e^{\gamma_0+\gamma_{1x}} + e^{-(\gamma_0+\gamma_{1x})}}\right)\right] \\ + 2Cov\left[\hat{\alpha}_1 \sqrt{b^2 - 4ac}\left(\ln x - \left(\frac{e^{\gamma_0+\gamma_{1x}} - e^{-(\gamma_0+\gamma_{1x})}}{e^{\gamma_0+\gamma_{1x}} + e^{-(\gamma_0+\gamma_{1x})}}\right), \hat{\alpha}_0 + \hat{\alpha}_1\left(\frac{e^{\gamma_0+\gamma_{1x}} - e^{-(\gamma_0+\gamma_{1x})}}{e^{\gamma_0+\gamma_{1x}} + e^{-(\gamma_0+\gamma_{1x})}}\right)\right)\right] \end{aligned}$$

Similarly Covariance structure can be shown to be equal to zero. Then we have

$$Var\left(Y_{LG}\right) = Var\left(Y_{HAAN}\right) + Var\left[\hat{\alpha}_1\left(\ln x - \left(\frac{e^{\gamma_0+\gamma_{1x}} - e^{-(\gamma_0+\gamma_{1x})}}{e^{\gamma_0+\gamma_{1x}} + e^{-(\gamma_0+\gamma_{1x})}}\right)\right)\right]$$

$$Var\left(Y_{LG}\right) = Var\left(Y_{HANN}\right) + \left[\ln x - \left(\frac{e^{\gamma_0+\gamma_{1x}} - e^{-(\gamma_0+\gamma_{1x})}}{e^{\gamma_0+\gamma_{1x}} + e^{-(\gamma_0+\gamma_{1x})}}\right)\right]^2 Var\left(\hat{\alpha}_1\right)$$

Since $Var\left(\hat{\alpha}_1\right)$ is a non-negative function, then

$Var\left(Y_{LG}\right) \geq Var\left(Y_{HANN}\right)$ if and only if $\left[\ln x - \frac{e^{\gamma_0+\gamma_{1x}-e^{-(\gamma_0+\gamma_{1x})}}}{e^{\gamma_0+\gamma_{1x}+e^{-(\gamma_0+\gamma_{1x})}}}\right]^2 \geq 0$

5.3 Log-linear Versus ANN with Logistic Transfer Function

Similarly, by letting Y_{LL} and Y_{LANN} denote the Log-linear and *ANN* respectively, the variance of one can be obtained in terms of the other.

The log-linear model is defined as follows

$$Y_{LL} = \ln Y = \hat{\alpha}_0 + \hat{\alpha}_1 x$$

$$Var\left(Y_{LL}\right) = Var\left[\left(Y_{LL} - Y_{LANN}\right) + Y_{LANN}\right]$$

$$\begin{aligned} Y_{LL} - Y_{LANN} &= \hat{\alpha}_0 + \hat{\alpha}_1 - \hat{\alpha}_0 - \hat{\alpha}_1\left(1 + e^{-(\gamma_0+\gamma_{1x})}\right)^{-1} \\ &= \hat{\alpha}_1\left[x - \left(1 + e^{-(\gamma_0+\gamma_{1x})}\right)^{-1}\right] \end{aligned}$$

$$\begin{aligned} Var\left(Y_{LL}\right) = Var\left[\hat{\alpha}_1\left(x - \left(1 + e^{-(\gamma_0+\gamma_{1x})}\right)^{-1}\right)\right] + Var\left[\hat{\alpha}_0 + \hat{\alpha}_1\left(1 + e^{-(\gamma_0+\gamma_{1x})}\right)^{-1}\right] + 2Cov \\ \left[\hat{\alpha}_1\left(x - \left(1 + e^{-(\gamma_0+\gamma_{1x})}\right)^{-1}\right), \hat{\alpha}_0\hat{\alpha}_1\left(1 + e^{-(\gamma_0+\gamma_{1x})}\right)^{-1}\right] \end{aligned}$$

Since covariance structure is zero, we have

$$Var\left(Y_{LL}\right)=\left[x-\left(1+e^{-(\gamma_0+\gamma_{1x})}\right)^{-1}\right]^2 Var\left(\hat{\alpha}_1\right)+Var\left(Y_{LANN}\right)$$

Since $Var\left(\hat{\alpha}_1\right)$ is a non-negative function,

$Var\left(Y_{LL}\right)\geq Var\left(Y_{LANN}\right)$ iff $\left[x-\left(1+e^{-(\gamma_0+\gamma_{1x})}\right)^{-1}\right]^2$ is equal to or greater than zero.

5.4 ANN Model with a Logistic Transfer Function Versus Logistic Regression Model

The logistic regression model of the form $Y_{LR}=\hat{\alpha}_0+\hat{\alpha}_1\left(1+e^{-x}\right)^{-1}$ is considered

A perception model with a logistic transfer function is given as

$$Y_{LTMN}=\hat{\alpha}_0+\hat{\alpha}_1\left(1+e^{-(\gamma_0+\gamma_{ix})}\right)^{-1}$$

$$Y_{LOR}-Y_{LTNN}=\hat{\alpha}_1\left[\left(1+e^{-x}\right)^{-1}-\left(1+e^{-(\gamma_0+\gamma_{ix})}\right)^{-1}\right]$$

$$V\left(Y_{LOR}\right)=V\left(Y_{LOR}-Y_{LTNN}\right)+V\left(Y_{LTNN}\right)+2Cov\left[\left(Y_{LOR}Y_{LTNN}\right)Y_{LTNN}\right]$$

Since $Cov\left[Y_{LOR}-Y_{LTNN},Y_{LTNN}\right]=0$, we have

$$V\left(Y_{LOR}\right)=V\left(Y_{LOR}-Y_{LTNN}\right)+V\left(Y_{LTNN}\right)$$

$$V\left(Y_{LOR}\right)=V\left(Y_{LTNN}\right)+V\left(\hat{\alpha}_1\right)\left[\left(1+e^{-x}\right)^{-1}-\left(1+e^{-(\gamma_0+\gamma_{ix})}\right)^{-1}\right]^2$$

$V\left(Y_{LOR}\right)\geq V\left(Y_{LTNN}\right)$ iff

$V\left(\hat{\alpha}_1\right)\left[\left(1+e^{-x}\right)^{-1}-\left(1+e^{-(\gamma_0+\gamma_{ix})}\right)^{-1}\right]^2\geq 0$ and $e^{-x}\neq e^{-(\gamma_0+\gamma_{1x})}$

5.5 ANN Model with a Linear Transfer Function Versus Univariate Linear Regression Model

$$\begin{aligned}
Y_L &= \hat{\alpha}_0+\hat{\alpha}_1 x\\
Y_{LNN} &= \hat{\alpha}_0+\hat{\alpha}_1\hat{\gamma}_0+\hat{\alpha}_1\hat{\gamma}_1 x\\
Y_L-Y_{LNN} &= \hat{\alpha}_1 x-\hat{\alpha}_1\hat{\gamma}_0-\hat{\alpha}_1\hat{\gamma}_0 x\\
&= \hat{\alpha}_1\left(x-\hat{\gamma}_0-\hat{\gamma}_0 x\right)\\
Y_L-Y_{LNN} &= \hat{\alpha}_1\left[x\left(1-\hat{\gamma}_0\right)-\hat{\gamma}_0\right]\\
V\left(Y_L\right) &= V\left[Y_L-Y_{LNN}\right]+V\left(Y_{LNN}\right)+2Cov\left[Y_L-Y_{LNN},Y_{LNN}\right]
\end{aligned}$$

Since $Cov\left[Y_L-Y_{LNN},Y_{LNN}\right]=0$

$$V\left(Y_L\right)=V\left(Y_{LNN}\right)+V\left(\alpha_1\right)\left[x\left(1-\hat{\alpha}_1\right)-\hat{\alpha}_0\right]^2$$

$V\left(Y_L\right)\geq V\left(Y_{LNN}\right)$ iff

$$V\left(\alpha_1\right)\left[x\left(1-\hat{\gamma}_1\right)-\hat{\gamma}_0\right]^2\geq 0$$

6. CERTAIN CONDITIONS

From the critical analysis of the variances of the models under review, the following are noted.

(i) ANN model with logistic function out-performed the linear-log if and only if

$$\left[\ln x - \left(1 + e^{-(\gamma_0+\gamma_{1x})}\right)^{-1}\right]^2 > 0$$

(ii) ANN model with Hyperbolic as transfer function will be better than that of logistic regression iff

$$\left[\ln x - \left(\frac{e^{\gamma_0+\gamma_{1x}} - e^{-(\gamma_0+\gamma_{1x})}}{e^{\gamma_0+\gamma_{1x}} + e^{-(\gamma_0+\gamma_{1x})}}\right)\right]^2 > 0$$

(iii) ANN model with logistic as transfer function posses a better efficiency over log-linear model iff

$$\left[x - \left(1 + e^{-(\gamma_0+\gamma_{1x})}\right)^{-1}\right]^2 > 0$$

7. SUMMARY, CONCLUSION AND RECOMMENDATION

7.1 Summary

This research work has presented an analytically the efficiency of both traditional regression model of various classes and the ANN of various transfer function. Analysis of variance of each model was also conducted without data consideration.

7.2 Conclusion

Based on the result obtained from the variable analysis, the following measures are hereby concluded:

(i) Without data consideration, the results obtained from variance analysis confirmed the competence of artificial neural network model of various transfer functions over the conventional regression model, which coincided with the earlier work by different Authors on the competence of ANN over the traditional regression models using data across fields.

7.3 Recommendation

This research work is conducted to put end to methodological dispute among the users of both classical regression and Artificial/neural network model. Therefore, from the conclusion arrived at, the following measures are recommended objectively:

(i) That World body of statistics should incorporate the teaching of artificial neural network into their Curriculum.

(ii) That ANN should be welcomed as very good replacement for alternative approach to the classical regression model.

REFERENCES

[1] Akaike H. (1973). Information Theory and an Estimation of the Maximum Likelihood Principle. *Second International Symposium on Information Theory. Budapest*.

[2] Amemiya T. (1980). Selection of Regressors. *International Economics Review, 21*, 331-354 leisten.

[3] Anders U. (1996). Was Neuronale Netze Wirklich Leisten. *Die Bank, 3*, 162-165.

[4] Davidson R. Mackinnon J.G. (1993). *Estimation and Inference in Econometrics*. Oxford University Press.

[5] Efron B., Tibshirani R. (1986). Boostrap Methods for Standard Errors Confidence Intervals and Other Measures of Statistical Accuracy. *Statistical Science, 1*(1), 54-77.

[6] Hardle W. (1993). *Applied NonParametrics Regression*. Cambridge University Press.

[7] Hornik K., Stinchcombe M., White H. (1989). Multilayer Feedforward Networks are Universal Approximators. *Neural Networks, 2*, 356-366.