Journal of Science and Technology, Vol. 40, No. 1 (2022), pp 27 - 37

RESEARCH PAPER

A DERIVED HETEROGENEOUS TRANSFER FUNCTION FROM CONVOLUTION OF SYMMETRIC HARDLIMIT AND HYPERBOLIC TANGENT SIGMOID TRANSFER FUNCTIONS

***Udomboso, C.G. and Ilori O.O.**
Department of Statistics, University of Ibadan, Ibadan, Nigeria

***Corresponding author: cg.udomboso@gmail.com**

ABSTRACT

This study derived a new heterogeneous transfer function of the Statistical Neural Network from a convolution of two transfer functions: the Symmetric Hard Limit and Hyperbolic Tangent Sigmoid, showing their various mathematical forms. The properties of the derived function were examined. Results show that it is a proper probability distribution with distributional properties shown to exist with mean 0, and variance $\sum_{p=1}^{\infty}\frac{1}{2p^4}$. ***Numerical illustrations showed that the derived heterogeneous model is more efficient than its homogeneous forms, as indicated from their respective predictive performances. From the foregoing, the use of homogeneous models of the statistical neural networks in solving empirical problems is encouraged, for effective outcomes.***

Keywords*:* Probability distribution, transfer functions, convolution, homogeneous statistical neural networks, heterogeneous statistical neural network, symmetric hard limit, hyperbolic tangent Sigmoid,

INTRODUCTION

Artificial Neural Networks (ANN) has been used over the years as empirical models to outperform most conventional statistical models. Its similarities to statistical models have been extensively discussed by Ripley (1993), Sarle (1994) and Sarle (1996). Anders (1996) was arguably the first to build an ANN model in statistical view, which he termed, Statistical Neural Network (SNN) model. In the same vein, Medeiros et al. (2006) developed an SNN model for time series models. Over the years, the Homogeneous Statistical Neural Networks (HOMSNN) models had been widely used in various studies. Literature is scarce on the Heterogeneous Statistical Neural Network (HETSNN) models. In Udomboso (2013a,b), it has been shown analytically and empirically, that the HETSNN is found to be more efficient than the HOMSNN model. Neural networks models, generally obtain their structure for the transfer functions used in the architecture. The choice of transfer functions in neural networks is of crucial importance to their performance.

Duch and Jankowski (1999) made a survey of transfer functions. They noted that before then, literature had been scarce on the subject matter, and had not been reviewed. Transfer functions may be used in the input pre-processing stage or as an integral part of the network (Duch and Jankowski, 2001). Testing several networks with different transfer functions and selecting the best one has been suggested as the simplest approach. Furthermore, they suggested that using heterogeneous transfer functions, also known as mixed transfer functions, in one network has the tendency to result into better effects. This may be introduced in two ways. Starting from a network with several types of transfer function one may train it, possibly using pruning techniques to drop functions that are not useful. Also, a constructive method that selects the most promising function from a pool of candidates, and selecting the one that is the best performer, adding it to the network, has been introduced (Duch, Adamczak and Dierksen, 2001; Jankowski and Duch, 2001). Other constructive algorithms, such as the cascade correlation (Fahlman and Lebiere, 1990), may also be used for this purpose. Each candidate node using different transfer function should be trained and the most useful candidate added to the network. Initially the network may be too complex but at the end only the functions that are best suited to solve the problem are left out.

This study centers on the Multi-Layer Perceptron (MLP) which happens to be the most used type of ANN (Resop, 2006). The choice of MLP is because it is the only ANN type that allows for statistical inference. The mathematical rigors of SNN model, which have been sparse in literature in the past, have been established in recent times (Anders, 1996; Medeiros et al., 2006, Udomboso, 2013a, b). Few studies have reported the use of one and/or combination of two transfer functions. For example, Adepoju et al. (2007), Adewole et al. (2011), Akaike (1974), Akinwale et al. (2009), Anders (1996), Anderson (2003) and Ashigwuike (2012), used the sigmoid transfer function, while Battiti (1992) compared logistic and hyperbolic tangent transfer functions. Adeyiga et al. (2011) used both the sigmoid and the family of tangent transfer functions, while Carling (1992) as well as Falode & Udomboso (2016) used the symmetric saturated linear transfer function. Udomboso (2013a,b) used a convolution of both the symmetric saturated linear and hyperbolic tangent sigmoid (satlins*tansig) as well as symmetric saturated linear and hyperbolic tangent (satlins*tanh) transfer functions.

This study endeavors to analytical derive a heterogeneous transfer function using the symmetric hard limit (HARDLIMS) as well as the hyperbolic tangent sigmoid (TANSIG) transfer functions. It further showed that the derived transfer function is a proper probability density function (pdf), with existing mean and variance.

A major limitation to the use of neural networks is the vanishing gradient problem, which happens in the event of large number of hidden layers. This is, sometimes, due to the need for a high number of feed-forward or -backward propagation from the input layer to the output layer in between many hidden layers. This can result into actual information getting lost (that is, vanishing) in the process. In order to reduce this limitation, this study does not go beyond a single hidden layer.

MATERIALS AND METHODS

The Statistical Neural Network Model and its Identification

The Statistical Neural Network (SNN) model proposed by Anders (1996) is given as

$$y = f(X, w) + u \tag{1}$$

where y is the dependent variable, $X = (x_0 \equiv 1, x_1, \dots, x_I)$ is a vector of independent variables, w = (α, β, γ) is the network weight: 'α' is the weight of the input unit, 'β' is the weight of the hidden unit, and 'γ' is the weight of the output unit, and u is the stochastic term that is normally distributed as $u \sim N(0, \sigma^2 I_n)$. Basically, $f(X, w)$ is the artificial neural network function, expressed as

$$f(X, w) = \alpha X + \sum_{h=1}^{H} \beta_h g\left(\sum_{i=0}^{I} \gamma_{hi}\, x_i\right) \tag{2}$$

where $g(.)$ is the Homogeneous Transfer Function (HOMTF),$g: X \to y$. Putting (2) in (1) gives

$$y = \alpha X + \sum_{h=1}^{H} \beta_h g\left(\sum_{i=0}^{I} \gamma_{hi}\, x_i\right) + u \tag{3}$$

which is known as the Homogeneous SNN (HOMSNN) model.

In this study, a convoluted form of the artificial neural network function given by Udomboso (2013) using product convolution is derived:

$$f(X, w) = \alpha X + \sum_{h=1}^{H} \beta_h \left[g_1\left(\sum_{i=0}^{I} \gamma_{hi}\, x_i\right) g_2\left(\sum_{i=0}^{I} \gamma_{hi}\, x_i\right)\right] \tag{4}$$

Putting (4) in (1), then (3) can be rewritten as

$$y = \alpha X + \sum_{h=1}^{H} \beta_h \left[g_1\left(\sum_{i=0}^{I} \gamma_{hi}\, x_i\right) g_2\left(\sum_{i=0}^{I} \gamma_{hi}\, x_i\right)\right] + u_i u_j \tag{5}$$

where u_i and u_j are the stochastic terms that are also normally distributed as $u_i, u_j \sim N(0, \sigma^2 I_n)$, and $g_1(.)$ and $g_2(.)$ are the transfer functions, $g_1(.).\, g_2(.)$ is known as Heterogeneous Transfer Function (HETTF), and other terms are as defined earlier. Equation (5) is known as Heterogeneous SNN (HETSNN).

In this paper, two HOMSNN models are convoluted to derive a new HETSNN model. The HOMSNN models consist of models having the HOMTFs: Symmetric HardLimit (*hardlims*) and Hyperbolic Tangent Sigmoid (*tansig*) transfer functions, respectively. Moreover, the distributional properties of the resulting HETSNN were investigated as follows:

Now,

1. Let $g_1(.) = \ hardlims$, defined as

$$hardlims = f_1(x) = \begin{cases} -1 & x < 0 \\ 1 & x > 0 \end{cases} \tag{6}$$

For this transfer function, the HOMSNN model can thus be written as,

(i) for 2 variables:

$$y = \alpha_0 + \alpha_1 x_1 + \sum_{h=1}^{H} \beta_h (\gamma_{h0} + \gamma_{h1} x_1) \ + u \tag{7}$$

So that at 1 hidden unit, (7) can be written as,

$$y = \alpha_0 + \alpha_1 x_1 + \beta_1 (\gamma_{10} + \gamma_{11} x_1) \ + u \tag{7.1}$$

and at 2 hidden units,

$$y = \alpha_0 + \alpha_1 x_1 + \beta_1 (\gamma_{10} + \gamma_{11} x_1) \ + \beta_2 (\gamma_{20} + \gamma_{21} x_1) \ + u \tag{7.2}$$

(ii) Similarly, for 3 variables:

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \sum_{h=1}^{H} \beta_h (\gamma_{h0} + \gamma_{h1} x_1 + \gamma_{h2} x_2) \ + u \tag{8}$$

So that at 1 hidden unit, (8) can be written as,

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \beta_1 (\gamma_{10} + \gamma_{11} x_1 + \gamma_{12} x_2) \ + u \tag{8.1}$$

and at 2 hidden units,

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \beta_1 (\gamma_{10} + \gamma_{11} x_1 + \gamma_{12} x_2) + \beta_2 (\gamma_{20} + \gamma_{21} x_1 + \gamma_{22} x_2) + u \tag{8.2}$$

2. Let $g_2(.) = tansig$, defined as

$$tansig = f_2(x) = (2/(1 - e^{-2\mathrm{x}})) - 1 \tag{9}$$

For this transfer function, the HOMSNN model can thus be written as,

(i) for 2 variables:

$$y = \alpha_0 + \alpha_1 x_1 + \ \sum_{h=1}^{H} \beta_h \big(2/(1 - e^{-2(\gamma_{h0} + \gamma_{h1} x_1)})) - 1\big) \ + u \tag{10}$$

So that at 1 hidden unit, (10) can be written as,

$$y = \alpha_0 + \alpha_1 x_1 + \beta_1 \big(2/(1 - e^{-2(\gamma_{10} + \gamma_{11} x_1)}) - 1\big) + u \tag{10.1}$$

and at 2 hidden units,

$$y = \alpha_0 + \alpha_1 x_1 + \beta_1 \big(2/\big(1 - e^{-2(\gamma_{10} + \gamma_{11} x_1)}\big) - 1\big) + \beta_1 \big(2/\big(1 - e^{-2(\gamma_{10} + \gamma_{11} x_1)}\big) - 1\big) + u \tag{10.2}$$

(i) for 3 variables:

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \sum_{h=1}^{H} \beta_h \left(2/(1 - e^{-2\left(\gamma_{h0} + \gamma_{h1} x_{1 +} \gamma_{h2} x_2\right)})) - 1\right) \ + u \tag{11}$$

So that at 1 hidden unit, (11) can be written as,

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \beta_1\big(2/(1 - e^{-2(\gamma_{10}+\gamma_{11}x_1+\gamma_{12}x_2)}) - 1\big) + u \qquad (11.1)$$

and at 2 hidden units,

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \beta_1\left(2/\left(1 - e^{-2(\gamma_{10}+\gamma_{11}x_{1+}\gamma_{12}x_2)}\right) - 1\right) +$$

$$\beta_1\big(2/(1 - e^{-2(\gamma_{10}+\gamma_{11}x_1+\gamma_{12}x_2)}) - 1\big) + u \qquad (11.2$$

3. The HETSNN is derived by convoluting the HOMTFs to obtain a new HETTF, and putting it into (5).

$$hardlims * tansig = f_1(x) * f_2(x) \qquad (12)$$

Let

$$f(x) = f_1(x) * f_2(x) = \int_a^b f_1(x-m) f_2(m)dm \qquad (13)$$

For $x < 0, f_1(x) = -1 \Rightarrow f_1(x-m) = -1$, then

$$f_1(x) * f_2(x) = \int_{-x}^{0}(-1)\,((2/(1 - e^{-2x})) - 1)dm, \quad -x < m < 0 \qquad (14)$$

$$= \int_{-x}^{0}(1 - 2/(1 - e^{-2x}))dm \qquad (15)$$

$$= \int_{-x}^{0} 1\,dm - \int_{-x}^{0}(2/(1 - e^{-2x}))dm \qquad (16)$$

$$= \int_{-x}^{0} dm - 2\int_{-x}^{0}(1 - e^{-2x})^{-1}dm \qquad (17)$$

$$= x - 2\int_{-x}^{0}(1 + e^{-2x} + e^{-4x} + e^{-6x} + \cdots)dm \qquad (18)$$

$$= x - 2(m - (e^{-2x}/2) - (e^{-4x}/4) - (e^{-6x}/6) + \cdots)\big|_{-x}^{0} \qquad (19)$$

$$= x - 2(0 - (1/2) - (1/4) - (1/6) + \cdots) + (x + (e^{-2x}/2)$$
$$+ (e^{-4x}/4) + (e^{-6x}/6) + \cdots) \qquad (20)$$

$$= x + \textstyle\sum_{p=1}^{\infty}(e^{2px}/p) - \sum_{p=1}^{\infty}(1/p) \qquad (21)$$

(ii) Similarly, for $x > 0, f_1(x) = 1 \Rightarrow f_1(x-m) = 1$, then

$$f_1(x) * f_2(x) = \int_0^x (1)\,(2/(1 - e^{-2x}) - 1)dm, \quad 0 < m < x \qquad (22)$$

$$= 2\int_0^x (1 - e^{-2x})^{-1}dm - \int_0^x 1\,dm \qquad (23)$$

$$= 2\int_{-x}^{0}(1 + e^{-2x} + e^{-4x} + e^{-6x} + \cdots)dm - \int_0^x 1\,dm \qquad (24)$$

$$= 2(m - (e^{-2x}/2) - (e^{-4x}/4) - (e^{-6x}/6) + \cdots)\big|_0^x - x \qquad (25)$$

$$= 2(x + (e^{-2x}/2) + (e^{-4x}/4) + (e^{-6x}/6) + \cdots) +$$
$$2(0 - (1/2) - (1/4) - (1/6) + \cdots) - x \qquad (26)$$

$$= x - \textstyle\sum_{p=1}^{\infty}(e^{-2px}/p) + \sum_{p=1}^{\infty}(1/p) \qquad (27)$$

The summary of the derived function is given as

$$f_1(x) * f_2(x) = \begin{cases} x + \sum_{p=1}^{\infty} (e^{2px}/p) - \sum_{p=1}^{\infty} (1/p) \ , & x < 0 \\ x - \sum_{p=1}^{\infty} (e^{-2px}/p) + \sum_{p=1}^{\infty} (1/p) \ , & x > 0 \end{cases} \tag{28}$$

Equation (28) is the derived HETTF for the symmetric hard limit and hyperbolic tangent sigmoid transfer functions.

For this derived transfer function, the HETSNN model can thus be written as

(i) for 2 variables,

$$y = \alpha_0 + \alpha_1 x_1 + \sum_{h=1}^{H} \beta_h \big((2(\gamma_{h0} + \gamma_{h1} x_1)^2) + 3(\gamma_{h0} + \gamma_{h1} x_1) + (1/2)\big)$$
$$- (\gamma_{h0} + \gamma_{h1} x_1)(\sum_{p=1}^{x} (e^{-2p(\gamma_{h0}+\gamma_{h1}x_1)}/p) - \sum_{p=1}^{x} (e^{-2p}/p))) + u_1 u_2 \tag{29}$$

So that at 1 hidden unit,

$$y = \alpha_0 + \alpha_1 x_1 + \beta_1 (2(\gamma_{10} + \gamma_{11} x_1)^2 + 3(\gamma_{10} + \gamma_{11} x_1) + (1/2))$$
$$-(\gamma_{10} + \gamma_{11} x_1)\big(\sum_{p=1}^{x} (e^{-2p(\gamma_{10}+\gamma_{11}x_1)}/p) - \sum_{p=1}^{x} (e^{-2p}/p))\big) + u_1 u_2 \tag{30}$$

at 2 hidden units,

$$y = \alpha_0 + \alpha_1 x_1 + \beta_1 (2(\gamma_{10} + \gamma_{11} x_1)^2 + 3(\gamma_{10} + \gamma_{11} x_1) + (1/2)$$
$$-(\gamma_{10} + \gamma_{11} x_1)(\sum_{p=1}^{x} (e^{-2p(\gamma_{10}+\gamma_{11}x_1)} \Big/ p) - \sum_{p=1}^{x} (e^{-2p}/p)) + \beta_2((2(\gamma_{20} + \gamma_{21} x_1)^2$$
$$+3(\gamma_{20} + \gamma_{21} x_1) + (1/2) - (\gamma_{20} + \gamma_{21} x_1)(\sum_{p=1}^{x} (e^{-2p(\gamma_{20}+\gamma_{21}x_1)}/p)$$
$$- \sum_{p=1}^{x} (e^{-2p}/p)) + u_1 u_2 \tag{31}$$

(ii) Similarly, for 3 variables,

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \sum_{h=1}^{H} \beta_h (2(\gamma_{h0} + \gamma_{h1} x_1 + \gamma_{h2} x_2)^2 + 3(\gamma_{h0} + \gamma_{h1} x_1 + \gamma_{h1} x_2)$$
$$+(1/2)) - (\gamma_{h0} + \gamma_{h1} x_1 + \gamma_{h1} x_2)(\sum_{p=1}^{x} (e^{-2p(\gamma_{h0}+\gamma_{h1}x_1+\gamma_{h1}x_2)}/p) - \sum_{p=1}^{x} (e^{-2p}/p))$$
$$+u_1 u_2 \tag{32}$$

at 1 hidden unit,

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \beta_1 (2(\gamma_{10} + \gamma_{11} x_1 + \gamma_{12} x_2)^2 + 3(\gamma_{10} + \gamma_{11} x_1 + \gamma_{12} x_2)$$
$$+ \sum_{p=1}^{x} (e^{-2p}/p)) + u_1 u_2 \tag{33}$$

at 2 hidden units,

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \beta_1 (2(\gamma_{10} + \gamma_{11} x_1 + \gamma_{12} x_2)^2 + 3(\gamma_{10} + \gamma_{11} x_1 + \gamma_{11} x_2)$$
$$+(1/2)) - (\gamma_{10} + \gamma_{11} x_1 + \gamma_{12} x_2)(\sum_{p=1}^{x} \big(e^{-2p(\gamma_{10}+\gamma_{11}x_1+\gamma_{12}x_2)}/p\big) - \sum_{p=1}^{x} (e^{-2p}/p))$$
$$+\beta_2 \big(2(\gamma_{20} + \gamma_{21} x_1 + \gamma_{22} x_2)^2 + 3(\gamma_{20} + \gamma_{21} x_1 + \gamma_{21} x_2) + (1/2)\big)$$
$$-(\gamma_{20} + \gamma_{21} x_1 + \gamma_{22} x_2)(\sum_{p=1}^{x} \big(e^{-2p(\gamma_{20}+\gamma_{21}x_1+\gamma_{22}x_2)}/p\big) - \sum_{p=1}^{x} (e^{-2p}/p)) + u_1 u_2 \tag{34}$$

2.2 Distributional Properties of the *Hardlims*Tansig*Function

- **Definition:**

  Let X be a continuous random variable with the probability distribution $f(x)$, which is said to be a probability density function (pdf) such that:

  (i) $f(x) \geq 0, \forall x$.

  (ii) $f(x)$ has, at most, a finite number of discontinuities in every finite interval on the real line.

  (iii) $\int_{-\infty}^{\infty} f(x)dx = 1$,

  (iv) For every interval $[a, b]$, $P\{a \leq X \leq b\} = \int_a^b f(x)dx$

- Now, to show that the derived HETTF in (28) is pdf:

$$\int_{-\infty}^{\infty} f_1(x) * f_2(x)dx = \int_{-\infty}^{0}(x + \textstyle\sum_{p=91}^{\infty}(e^{2px}/p) - \sum_{p=1}^{\infty}(1/p))\,dx$$

$$+ \int_0^{\infty}\left(x - \textstyle\sum_{p=1}^{\infty}(e^{2px}/p) + \sum_{p=1}^{\infty}(1/p)\right)dx \qquad (35)$$

$$= \int_{-\infty}^{0} x\,dx + \int_{-\infty}^{0}\textstyle\sum_{p=1}^{\infty}(e^{2px}/p) - \int_{-\infty}^{0}\sum_{p=1}^{\infty}(1/p)\ + \int_0^{\infty} x\,dx$$

$$- \int_0^{\infty}\textstyle\sum_{p=1}^{\infty}(e^{-2px}/p)\ dx + \int_0^{\infty}\sum_{p=1}^{\infty}(1/p)\,dx \qquad (36)$$

$$= \textstyle\sum_{p=1}^{\infty}(e^{2px}/2p^2)\,\Big|_{-\infty}^{0} + \sum_{p=1}^{\infty}(e^{2px}/2p^2)\,\Big|_{0}^{\infty} \qquad (37)$$

$$= \textstyle\sum_{p=1}^{\infty}(1/2p^2) + \sum_{p=1}^{\infty}(1/2p^2) \qquad (38)$$

$$= 1 + c \qquad (39)$$

where $c = 2\sum_{p=2}^{\infty}(1/2p^2)$ is negligible (0) as $p \to \infty$.

- Hence, we obtain the mean and variance of the derived HETTF as follows:

$$f_1(x) * f_2(x) = \begin{cases} x + \sum_{p=1}^{\infty}(e^{2px}/p) - \sum_{p=1}^{\infty}(1/p) & x < 0 \\ x - \sum_{p=1}^{\infty}(e^{2px}/p) + \sum_{p=1}^{\infty}(1/p) & x > 0 \end{cases}$$

$$E(x) = \int_{-\infty}^{\infty} x(f_1(x) * f_2(x))dx \qquad (40)$$

$$= \int_{-\infty}^{\infty} x(f_1(x) * f_2(x))dx = \int_{-\infty}^{0} x(x + \textstyle\sum_{p=1}^{\infty}(e^{2px}/p) - \sum_{p=1}^{\infty}(1/p))\,dx$$

$$+ \int_0^{\infty} x(x - \textstyle\sum_{p=1}^{\infty}(e^{2px}/p) + \sum_{p=1}^{\infty}(1/p))dx \qquad (41)$$

$$= (x^3/3)\Big|_{-\infty}^{0} + \textstyle\sum_{p=1}^{\infty}(1/p)\int_{-\infty}^{0} xe^{2px} - \sum_{p=1}^{\infty}(1/p)\int_{-\infty}^{0} x\,dx$$

$$+(x^3/3)\Big|_0^{\infty} - \textstyle\sum_{p=1}^{\infty}(1/p)\int_0^{\infty} xe^{-2px} + \sum_{p=1}^{\infty}(1/p)\int_0^{\infty} x\,dx \qquad (42)$$

$$= \textstyle\sum_{p=1}^{\infty}(e^{-2px}/p)\,(x - (1/2p))\Big|_{-\infty}^{0} - \sum_{p=1}^{\infty}(e^{-2px}/p)\,(x - (1/2p))\Big|_0^{\infty} \qquad (43)$$

$$= -\textstyle\sum_{p=1}^{\infty}(1/2p^3)\ +\ \sum_{p=1}^{\infty}(1/2p^3)\ = 0 \qquad (44)$$

$$E(x^2) = \int_{-\infty}^{\infty} x^2(f_1(x) * f_2(x))dx \qquad (45)$$

$$= \int_{-\infty}^{\infty} x^2(f_1(x) * f_2(x))dx = \int_{-\infty}^{0} x^2(x + \textstyle\sum_{p=1}^{\infty}(e^{2px}/p) - \sum_{p=1}^{\infty}(1/p))\,dx$$

$$+\int_0^{\infty} x^2\big(x-\textstyle\sum_{p=1}^{\infty}(e^{2px}/p)+\sum_{p=1}^{\infty}(1/p)\big)dx \qquad (46)$$

$$=(x^4/3)\big|_{-\infty}^{0}+\textstyle\sum_{p=1}^{\infty}(1/p)\int_{-\infty}^{0}x^2e^{2px}-\sum_{p=1}^{\infty}(1/p)\int_{-\infty}^{0}x^2\,dx+\ (x^4/3)\big|_0^{\infty}$$

$$-\textstyle\sum_{p=1}^{\infty}(1/p)\int_0^{\infty}x^2e^{-2px}+\sum_{p=1}^{\infty}(1/p)\int_0^{\infty}x^2\,dx \qquad (47)$$

$$=\textstyle\sum_{p=1}^{\infty}(1/p)\,(x^2(e^{2px}/2p)(x(e^{2px}/2p^2)-(e^{2px}/4p^3))\big|_{-\infty}^{0}$$

$$-\textstyle\sum_{p=1}^{\infty}(1/p)\,(-x^2(e^{-2px}/2p)-x(e^{-2px}/2p^2)+(e^{2px}/4p^3))\big|_0^{\infty} \qquad (48)$$

$$=\textstyle\sum_{p=1}^{\infty}(1/p)(1/4p^3)+\sum_{p=1}^{\infty}(1/p)(1/4p^3\ )\ =2\sum_{p=1}^{\infty}(1/4p^4) \qquad (49)$$

$$=\textstyle\sum_{p=1}^{\infty}(1/2p^4)=((1/2)+(1/32)\ +\cdots) \qquad (50)$$

Therefore, variance of $f_1(x)*f_2(x)$ is

$$var(x)=E(x^2)-[E(x)]^2 \qquad (51)$$

$$=\textstyle\sum_{p=1}^{\infty}(1/2p^4)-0 \qquad (52)$$

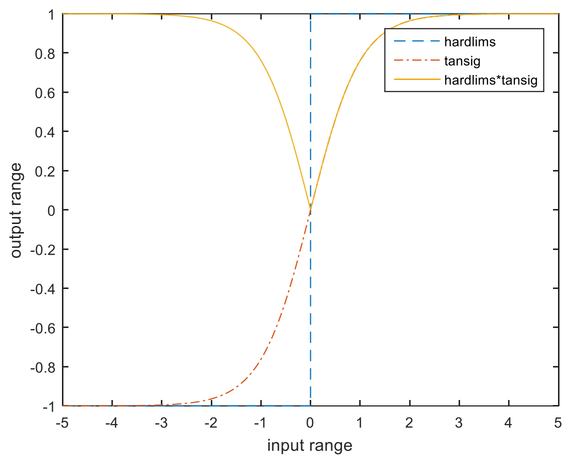
$$=\textstyle\sum_{p=1}^{\infty}(1/2p^4) \qquad (53)$$

Hence, equations (44) and (53) are the mean and the variance of the derived HETTF, respectively.

Thus, HETTF

$$g_1(\textstyle\sum_{i=0}^{I}\gamma_{hi}x_i)g_2(\sum_{j=0}^{J}\gamma_{hi}x_i)=\begin{cases}x+\sum_{p=1}^{\infty}(e^{2px}/p)-\sum_{p=1}^{\infty}(1/p), & x<0\\ x-\sum_{p=1}^{\infty}(e^{2px}/p)+\sum_{p=1}^{\infty}(1/p), & x>0\end{cases}$$

has a mean value of 0 and variance $\sum_{p=1}^{\infty}\frac{1}{2p^4}$.



Adepoju, G. A., Ogunjuyigbe, S. O. A., & Alawode, K. O. (2007). Application of neural network to load forecasting in Nigerian electrical power system. The Pacific

Figure 1 is the plot of the *hardlims*, *tansig*, and the *hardlims*tansig* transfer functions. The plot shows that the *hardlims* is vertically linear at $x=0$ and $-1\geq y\geq +1$, *tansig* is sigmoidal at $-5\geq x\geq +5$ and $-1\geq y\geq +1$, while *hardlims*tansig* is sigmoidal v-shaped at $x\geq -1$, $x\leq +1$, and $0\geq y\geq +1$.

Figure 1: Plot of the homogeneous (hardlims and transig) and heterogeneous (hardlims*tansig) transfer functions

RESULTS AND DISCUSSIONS

To demonstrate the models (HOMSNNs and HETSNN), the data used were generated from the normal distribution with mean of 5 and variance of 1. The data were subjected to each of the modelling approaches. A fixed hidden neuron was used in the models. Analyses were done using MATLAB R2015a at 1000 iterations. The mean and variance were computed for each prediction and their generated errors. Also, computed is their Network Information Criterion (NIC).

Table 1: Predictive performance of the HOMSNN and HETSNN transfer functions

Model	Transfer Function	Predicted		Error		NIC
		Mean	Variance	Mean	Variance	
HOMSNN	***Hardlims***	5.0999	0.1564	0.1001	0.5744	1.0854
	Tansig	5.1373	0.1776	0.0627	0.4357	1.1899
HETSNN	***Hardlims*Tansig***	4.8648	0.0395	0.3352	0.5989	0.6938

Table 1 presents the mean and variance of predicted values and the error generated. It is noticed that for the predicted values, the HETSNN model with hardlims*tansig transfer function had the least mean and variance, while in the case of the generated error, the HOMSNN model with tansig transfer function had the least mean and variance. Model selection based on the NIC shows that the HETSNN model with hardlims*tansig transfer function is the better preferred model, while the tansig function is the least preferred.

CONCLUSION

This study derived a heterogeneous transfer function involving the symmetric hardlimit and hyperbolic tangent sigmoid transfer functions. The mathematical functions of the homogeneous and heterogeneous transfer functions were incorporated into the structural form of the statistical neural network models, respectively. Hence, the empirical form was described, showing the number of terms ensuing from the number of input and hidden neurons, respectively. It went further to show that the derived transfer function is a proper probability distribution function (pdf), having mean and variance to be and $\sum_{p=1}^{\infty}\frac{1}{2p^4}$, respectively. The numerical illustrations showed that the heterogeneous model is more efficient than its homogeneous forms, given from their predictive performances. This study suggests the use of homogeneous functions of the statistical neural networks in solving empirical problems for effective outcomes.

ACKNOWLEDGEMENTS

The authors greatly acknowledge valuable comments from colleagues and experts in the field including the anonymous reviewers and editors that have made this work valuable.

REFERENCES

Journal of Science and Technology, 8 (1), 68-72.

Adewole, A. P., Akinwale, A. T., &Akintomide, A. B. (2011). Artificial neural network model for forecasting foreign exchange rate. World of Computer Science and Information Technology Journal, 1 (3), 110-118.

Adeyiga, J. A., Ezike, J. O. J., Omotosho, A.,&Amakulor, W. (2011). A neural network based model for detecting irregularities in e-Banking transactions. African Journal of Computer and ICTs, 4 (2), 7-14.

Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19, 716-723.

Akinwale, A. T., Arogundade, O. T. and Adekoya, A. F. (2009). Translated Nigeria stock market prices using artificial neural network for effective prediction. Journal of Theoretical and Applied Information Technology, 36-43.

Anders, U. (1996). Statistical model building for neural networks. Nunberg,Germany: AFIR Colloqium.

Anderson, J. A. (2003). An introduction to neural networks. Prentice Hall.

Ashigwuike, C. E. (2012). Estimation of solar power generation in some Nigerian cities using artificial neural network. Journal of Chemical, Biological and Physical Sciences, 2 (2), 929-936.

Battiti, R. (1992). First- and second-order methods for learning: Between steepest descent and Newtons method. Neural Computation, 4, 141-166.

Carling, A. (1992). Introducing neural networks. England: Sigma Press.

Duch W. and Jankowski N. (1999). Survey of Neural Transfer Functions. Neural Computing Surveys 2, 163-212. Retrieved at http://www.icsi.berkeley.edu/_jagota/NCS.

Duch, W.; Jankowski, N. (2001). Transfer functions: hidden possibilities for better neural networks. 9th European Symposium on Artificial Neural Networks (ESANN), Brugge 2001. De-facto publications, pp. 81-94.

Duch W., Adamczak R. and Diercksen G.H.F. (2001). Constructive density estimation network based on several different separable transfer functions. In 9th European Symposium on Artificial Neural Networks, Bruges, Belgium, April 2001.

Fahlman S. E. and Lebiere C. (1990). The cascade-correlation learning architecture. In D. S. Touretzky, editor, Advances in Neural Information Processing Systems, 2, pages 524–532. Morgan Kaufmann.

Falode, O.A. and Udomboso, C.G. (2016). Predictive Modeling of Gas Production, Utilization and Flaring in Nigeria using TSRM and TSNN: A Comparative Approach. Open Journal of Statistics.Vol. 6: 194-207.

Jankowski N. and Duch W. (2001). Universal neural networks. In 9th European Symposium on Artificial Neural Networks, Bruges, Belgium, April 2001.

Medeiros M. C., Terasvirta T. and Rech G. (2006). “Building neural network models for time series: a statistical approach”, Journal of Forecasting, Vol. 25, pp. 49-75.

Resop, J. P. (2006). A comparison of artificial neural networks and statistical regression with biological resources applications. (M.S. Thesis). University of Maryland, College Park, USA.

Ripley B. D. (1993). Statistical aspects of neural networks, in Barndorff-Nielsen, O.E., Jensen, J.L. and Kendall, W.S., eds., Networks and Chaos: Statistical and Probabilistic Aspects, London: Chapman & Hall.

Sarle W. S. (1994). Neural networks and statistical models, Proceedings of the

Nineteenth Annual SAS Users Group International Conference, Cary, NC: SAS Institute, USA, pp. 1538-1550.

Sarle W. S. (1996). “Neural network and statistical jargon?”, April 29, 1996, [Online], Archive-name:ai-faq/neural-nets/part3, Available: ftp://ftp.sas.com/pub/neural/jargon.

Udomboso C. G. (2013a). “On some properties of a hetereogeneous transfer function involving Symmetric Saturated Linear (SATLINS) with Hyperbolic Tangent Sigmoid (TANSIG) transfer functions”, Progress in Applied Mathematics, Vol. 5, No. 2, pp. 40-47.

Udomboso C. G. (2013b). “On some properties of a hetereogeneous transfer function involving Symmetric Saturated Linear (SATLINS) with Hyperbolic Tangent (TANH) transfer functions”, Journal of Modern Applied Statistical Methods.Vol.12. No. 2, pp. 427-435.