On Some Properties of a Heterogeneous Transfer Function Involving Symmetric Saturated Linear (SATLINS) with Hyperbolic Tangent (TANH) Transfer Functions

Christopher Godwin Udomboso
University of Ibadan, Ibadan, Nigeria, cg.udomboso@gmail.com

Follow this and additional works at: http://digitalcommons.wayne.edu/jmasm

Recommended Citation
Available at: http://digitalcommons.wayne.edu/jmasm/vol12/iss2/26
Emerging Scholars:
On Some Properties of a Heterogeneous Transfer Function Involving Symmetric Saturated Linear (SATLINS) with Hyperbolic Tangent (TANH) Transfer Functions

Christopher Godwin Udomboso
University of Ibadan
Ibadan, Nigeria

For transfer functions to map the input layer of the statistical neural network model to the output layer perfectly, they must lie within bounds that characterize probability distributions. The heterogeneous transfer function, SATLINS_TANH, is established as a Probability Distribution Function (p.d.f), and its mean and variance are shown.

Keywords: Statistical neural network, SATLINS, TANH, SATLINS_TANH, mean, variance

Introduction
Anders (1996) proposed a statistical neural network model given as

\[ y = f(X, w) + u \]  \hspace{1cm} (1)

where \( y \) is the dependent variable, \( X = (x_0 \equiv 1, x_1, \ldots, x_I) \) is a vector of independent variables, \( w = (\alpha, \beta, \gamma) \) is the network weight: \( \alpha \) is the weight of the input unit, \( \beta \) is the weight of the hidden unit, and \( \gamma \) is the weight of the output unit, and \( u_i \) is the stochastic term that is normally distributed (that is, \( u_i \sim N(0, \sigma^2 I_n) \)).

Basically, \( f(X, w) \) is the artificial neural network function, expressed as

\[ f(X, w) = \alpha X + \sum_{h=1}^{H} \beta_h g \left( \sum_{i=0}^{I} \gamma_{hi} x_i \right) \]
where \( g(.) \) is the transfer function.

The proposed convoluted form of the artificial neural network function used in this study is

\[
\begin{align*}
  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] \\

eq \sum_{i=0}^{I} \gamma_{hi} x_i
\end{align*}
\]

and thus, the form of the statistical neural network model proposed is

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

where \( y \) is the dependent variable, \( X = (x_0 \equiv 1, x_1, \ldots, x_I) \) is a vector of independent variables, \( w = (\alpha, \beta, \gamma) \) is the network weight: \( \alpha \) is the weight of the input unit, \( \beta \) is the weight of the hidden unit, and \( \gamma \) is the weight of the output unit, \( u_i \) and \( u_j \) are the stochastic terms that are normally distributed (that is, \( u_i, u_j \sim N(0, \sigma^2 I_n) \)) and \( g_1(.) \) and \( g_2(.) \) are the transfer functions.

The distributional properties of the heterogeneous model arising from the convolution of SATLINS and TANH is investigated here. Let \( g_1(.) = \text{Symmetric Saturated Linear function (SATLINS)} \), defined as

\[
satlins = g_1(n) = f_1(n) = \begin{cases} 
  -1, & n < -1 \\
  n, & -1 \leq n \leq 1 \\
  1, & n > 1
\end{cases}
\]

Let \( g_2(.) = \text{Hyperbolic Tangent function (TANH)} \), defined as

\[
tanh = g_2(n) = f_2(n) = \frac{e^n - e^{-n}}{e^n + e^{-n}}
\]
Symmetric Saturating Linear and Hyperbolic Tangent

(i) Let
\[ f(n) = f_1(n) \otimes f_2(n) = \int_a^n f_1(n-m)f_2(m)\,dm \quad \text{(5)} \]

For \( n < -1 \), \( f_1(n) = -1 \), which implies also that \( f_1(n-m) = -1 \).

\[ f_2(m) = \frac{e^m - e^{-m}}{e^m + e^{-m}} \]

Therefore,
\[ f_1(n) \otimes f_2(n) = \int_{-r}^n (-1) \left( \frac{e^m - e^{-m}}{e^m + e^{-m}} \right) dm, \quad r < n < -1 \]
\[ = \log \left( \frac{e^m + e^{-m}}{e^{n} + e^{-n}} \right) \quad \text{(6)} \]

(ii) Similarly, for \(-1 \leq n \leq 1\), \( f_1(n) = n \), which implies that \( f_1(n-m) = n-m \), such that \(-1 \leq m \leq n\).

Therefore,
\[ f_1(n) \otimes f_1(n) = \int_{-1}^n f_1(n-m)f_2(m)\,dm \]
\[ = \int_{-1}^n (n-m) \left( \frac{e^m - e^{-m}}{e^m + e^{-m}} \right) dm \quad \text{(7)} \]

Using integration by part, and noting that
\[ \int uv' = uv - \int u'v \]

Let \( u = n - m \). This implies that \( du = -dm \).
and $v' = -\frac{d\left[e^m + e^{-m}\right]}{e^m + e^{-m}}$. This implies that $v = \log(e^m + e^{-m})$.

Thus,

$$f_1(n) \otimes f_2(n) = (n - m)\log(e^m + e^{-m}) + \int_{-1}^{n} \log(e^m + e^{-m}) \, dm \quad (8)$$

In (6), let $I = \int_{-1}^{n} \log(e^m + e^{-m}) \, dm$

Now, let $x = \log(e^m + e^{-m})$, which implies that $e^x = e^m + e^{-m}$

But $x = k \in \mathbb{N}$ for $-1 \leq m \leq 1$. Hence $I = 0$.

Therefore,

$$f_1(n) \otimes f_2(n) = -(n + 1)\log(e + e^{-1}) \quad (9)$$

(iii)

Also, for $n > 1$, $f_1(n) = a = 1$. This implies that $f_1(n - m) = 1$

Therefore,

$$f_1(n) \otimes f_2(n) = \int_{1}^{n} f_1(n - m) f_2(m) \, dm$$

$$= \int_{1}^{n} e^m - e^{-m} \, dm = \log\left(\frac{e^n + e^{-n}}{e + e^{-1}}\right) \quad (10)$$

The summary of the derived function is given as
\[ g_1 \left( \sum_{i=0}^{l} y_{hi}x_i \right) g_2 \left( \sum_{i=0}^{l} y_{hi}x_i \right) = f(n) = \begin{cases} 
\log \left( \frac{e^n + e^{-n}}{e^n + e^{-n}} \right), & \text{for } n < -1 \\
(n+1) \log \left( e + e^{-1} \right)^{-1}, & \text{for } -1 \leq n \leq 1 \\
\log \left( \frac{e^n + e^{-n}}{e + e^{-1}} \right), & \text{for } n > 1 
\end{cases} \]  

(11) is the derived transfer function for the Symmetric Saturated Linear transfer function and the Hyperbolic Tangent transfer function.

**Distributional Properties of the SATLINS\_TANH SNN Model**

Next it is shown that the derived transfer functions are probability density functions. By definition, the probability density function \( p.d.f \) of function \( f(x) \) of a random variable \( X : \Omega \rightarrow \mathbb{R} \) is said to be a proper \( p.d.f \) if for \( x \in (-\infty, +\infty), x \in X \), thus,

\[ \int_{-\infty}^{\infty} f(x) \, dx = 1, \ x \in X \]

From the derived transfer function in (11),

\[ \int_{-\infty}^{\infty} f_1(n) \otimes f_2(n) \, dn = \int_{-\infty}^{-1} \log \left( \frac{e^n + e^{-n}}{e^n + e^{-n}} \right) \, dn + \int_{-1}^{1} (n+1) \log \left( e + e^{-1} \right)^{-1} \, dn \\
+ \int_{1}^{\infty} \log \left( \frac{e^n + e^{-n}}{e + e^{-1}} \right) \, dn \]
PROPERTIES OF SATLINS WITH TANH TRANSFER FUNCTIONS

\[
\begin{align*}
= & \int_{-\infty}^{-1} \left[ \log(e^r + e^{-r}) - \log(e^n + e^{-n}) \right] dn \\
& + \log(e + e^{-1})^{-1} \int_{-1}^{1} (n+1) dn \\
& + \int_{1}^{\infty} \left[ \log(e^n + e^{-n}) - \log(e + e^{-1}) \right] dn \\
= & \int_{-\infty}^{-1} \log(e^r + e^{-r}) dn + \log(e + e^{-1})^{-1} \int_{-1}^{1} (n+1) dn - \int_{1}^{\infty} \log(e + e^{-1}) dn \\
= & \left[ n \log(e^r + e^{-r}) \right]_{-\infty}^{-1} + \left. \log(e + e^{-1}) \left( \frac{n^2}{2} + n \right) \right|_{1}^{1} - \left[ n \log(e + e^{-1}) \right]_{1}^{\infty} \\
= & \infty + 2 \log(e + e^{-1}) - \infty \\
= & 2 \log(e + e^{-1})
\end{align*}
\]

The mean and variance of the derived transfer functions are obtained next.

For \( f_1(n) \otimes f_2(n) \)

\[
f_1(n) \otimes f_2(n) = \begin{cases} \\
\log \left( \frac{e^r + e^{-r}}{e^n + e^{-n}} \right), \quad \text{for } n < -1 \\
(n+1) \log \left( e + e^{-1} \right)^{-1}, \quad \text{for } -1 \leq n \leq 1 \\
\log \left( \frac{e^n + e^{-n}}{e + e^{-1}} \right), \quad \text{for } n > 1 \\
\end{cases}
\]

\[
E(n) = \int_{-\infty}^{\infty} n \left( f_1(n) \otimes f_2(n) \right) dn \\
= \int_{-\infty}^{-1} n \log \left( \frac{e^r + e^{-r}}{e^n + e^{-n}} \right) dn + \int_{-1}^{1} n (n+1) \log(e + e^{-1})^{-1} dn + \int_{1}^{\infty} n \log \left( \frac{e^n + e^{-n}}{e + e^{-1}} \right) dn
\]
\[
= \int_{-\infty}^{-1} n \log(e^{-r} + e^{r}) \, dn - \int_{-\infty}^{-1} n \log(e^{n} + e^{-n}) \, dn \\
+ \int_{-1}^{1} n(n+1) \log(e^{n+1}) \, dn \\
+ \int_{-\infty}^{\infty} n \log(e^{n} + e^{-n}) \, dn - \int_{1}^{\infty} n \log(e^{n+1}) \, dn \\
= \log(e^{-r} + e^{r}) \int_{-\infty}^{-1} n \, dn - \log(e^{n+1}) \int_{-1}^{1} (n^2 + n) \, dn - \log(e^{n+1}) \int_{1}^{\infty} n \, dn \\
= \log(e^{-r} + e^{r}) \left[ \frac{n^2}{2} \right]_{-\infty}^{1} - \log(e^{n+1}) \left[ \frac{n^3}{3} + \frac{n^2}{2} \right]_{-1}^{1} - \log(e^{n+1}) \left[ \frac{n^2}{2} \right]_{1}^{\infty} \\
= \log(e^{-r} + e^{r}) \left( \frac{1}{2} \right) - \log(e^{n+1}) \left[ \frac{1}{3} + \frac{1}{2} \right] - \log(e^{n+1}) \left( \frac{3}{2} - \frac{1}{2} \right) \\
\]

Hence, the mean of derived transfer function in is

\[
E(n) = \frac{2}{3} \log(e^{n+1})^{-1} \\
\text{(13)}
\]

Similarly,

\[
E(n^2) = \int_{-\infty}^{\infty} n^2 \left( f_1(n) \otimes f_2(n) \right) \, dn \\
= \int_{-\infty}^{-1} n^2 \log\left( \frac{e^{-r} + e^{r}}{e^{n} + e^{-n}} \right) \, dn + \int_{-1}^{1} n^2 (n+1) \log(e^{n+1})^{-1} \, dn + \int_{1}^{\infty} n^2 \log\left( \frac{e^{n+1}}{e^{n} + e^{-n}} \right) \, dn \\
= \log(e^{-r} + e^{r}) \int_{-\infty}^{-1} n^2 \, dn - \int_{-\infty}^{-1} n^2 \log(e^{n} + e^{-n}) \, dn \\
+ \log(e^{n+1})^{-1} \int_{-1}^{1} (n^3 - n^2) \, dn \\
+ \int_{1}^{\infty} n^2 \log(e^{n} + e^{-n}) \, dn - \log(e^{n+1}) \int_{1}^{\infty} n^2 \, dn 
\]

433
PROPERTIES OF SATLINS WITH TANH TRANSFER FUNCTIONS

\[
= \log \left( \frac{e^r + e^{-r}}{3} \right) \left[ \frac{n^3}{3} \right]^{-1} + \log \left( e^r + e^{-r} \right)^{-1} \left[ \frac{n^4 - n^3}{4} \right]^{-1} + \log \left( e^r + e^{-r} \right) \left[ \frac{n^3}{3} \right]
\]

\[
= \log \left( e^r + e^{-r} \right) \left( \frac{\infty}{3} - \frac{1}{3} \right) + \log \left( e^r + e^{-r} \right)^{-1} \left[ \left( \frac{1}{4} + \frac{1}{3} \right) - \left( \frac{1}{4} - \frac{1}{3} \right) \right] - \log \left( e^r + e^{-r} \right) \left( \frac{\infty}{3} - \frac{1}{3} \right)
\]

\[
= \frac{2}{3} \log \left( e^r + e^{-r} \right)^{-1}
\]

Therefore, variance of \( \left( f_i(n) \otimes f_z(n) \right) \) is

\[
\text{var}(n) = E(n^2) - \left[ E(n) \right]^2
\]

\[
= \frac{2}{3} \log \left( e^r + e^{-r} \right)^{-1} - \left[ \frac{2}{3} \log \left( e^r + e^{-r} \right)^{-1} \right]^2
\]

\[
= -\frac{2}{3} \log \left( e^r + e^{-r} \right) + \frac{4}{9} \left( \log \left( e^r + e^{-r} \right) \right)^2
\]

\[
= \log \left( e^r + e^{-r} \right) \left[ \frac{4}{9} \log \left( e^r + e^{-r} \right) - \frac{2}{3} \right]
\]

Thus,

\[
g_1 \left( \sum_{i=0}^{\hat{i}} \gamma_{hi} x_i \right) g_2 \left( \sum_{i=0}^{\hat{i}} \gamma_{hi} x_i \right) = f(n) = \begin{cases} 
\log \left( \frac{e^r + e^{-r}}{e^n + e^{-n}} \right), & \text{for } n < -1 \\
(n+1) \log \left( e^r + e^{-r} \right)^{-1}, & \text{for } -1 \leq n \leq 1 \\
\log \left( \frac{e^n + e^{-n}}{e + e^{-1}} \right), & \text{for } n > 1
\end{cases}
\]

with mean, \( E(n) = \frac{2}{3} \log \left( e^r + e^{-r} \right)^{-1} \)

and variance, \( \text{var}(n) = \log \left( e^r + e^{-r} \right) \left[ \frac{4}{9} \log \left( e^r + e^{-r} \right) - \frac{2}{3} \right] \).

434
References


