IJERMS

International Journal of Engineering Researches and Management Studies

A FAMILY OF RECURRENCE GENERATING ACTIVATION FUNCTIONS BASED ON GUDERMANN FUNCTION

Anna Malinova[*1], Angel Golev[2], Anton Iliev[3] and Nikolay Kyurkchiev[4]
[*1,2,3&4]Faculty of Mathematics and Informatics, University of Plovdiv PaisiiHilendarski, 24, TzarAsen Str., 4000 Plovdiv, Bulgaria

ABSTRACT
In this note we construct a family of recurrence generating activation functions based on Gudermann function. We prove lower estimate for the Hausdorff approximation of the sign function by means of this family. Numerical examples, illustrating our results are given.

Keywords: Family of GudermannActivation Functions, Sign Function, HausdorffDistance, LowerBound.

1. INTRODUCTION

Sigmoidal functions (also known as "activation functions") find multiple applications to neural networks [8]–[18], [43], [44], [46]–[49].

The modified hyperbolic tangent is a special S-shaped function constructed on the basis of the hyperbolic tangent function, which is expressed in terms of the exponent.

We study the distance between the sign function and a special class of family of recurrence generating activation function based on modified half Gudermann function (FMHGUDAF).

The distance is measured in Hausdorff sense, which is natural in a situation when a sign function is involved. Precise lower bound for the Hausdorff distance is reported.

Any neural net element computes a linear combination of its input signals, and uses a logistic function to produce the result; often called "activation" function [19]– [20].

2. PRELIMINARIES

The following are common examples of activation functions:

- logistic

$$\varphi_1(t) = \frac{1}{1+e^{-t}}; \quad (1)$$

- Parametric Hyperbolic Tangent Activation (PHTA) function

$$\varphi_2(t) = \frac{e^{\beta t}-e^{-\beta t}}{e^{\beta t}+e^{-\beta t}} = 1 - \frac{2e^{-\beta t}}{e^{\beta t}+e^{-\beta t}}, \quad t \in \mathbb{R}, \quad \beta \geq 1; \quad (2)$$

- Parametric Half Hyperbolic Tangent Activation (PHHTA) function

$$\varphi_3(t) = \frac{1-e^{-\beta t}}{1+e^{-\beta t}}, \quad t \in \mathbb{R}, \quad \beta \geq 1; \quad (3)$$

- Parametric Fibonacci hyperbolic tangent activation function (FHTAF) [38] based on the Fibonacci hyperbolic tangent function [7]

$$\varphi_4(t) = \frac{\Psi^{\beta t}-\Psi^{-\beta t}}{\Psi^{\beta t}+\Psi^{-\beta t}}, \quad t \in \mathbb{R}, \quad \beta \geq 1 \quad (4)$$

where $\Psi = 1 + \phi = \frac{3+\sqrt{5}}{2} \approx 2.61$ and ϕ is the "Golden Section";

International Journal of Engineering Researches and Management Studies

A survey of new mathematical models of Nature is presented based on the Golden Section and using a class of hyperbolic Fibonacci and Lucas functions in [6].

- Parametric Soboleva' modified hyperbolic tangent activation function [39] based on Soboleva' modified hyperbolic tangent function [1]–[3]

$$\varphi_5(t) = m(t; c, d, c, d) = \frac{e^{ct} - e^{-dt}}{e^{ct} + e^{-dt}}; \qquad (5)$$

The function find application to approximate the current-voltage characteristics of light-emitting diodes [4].

In [21] the authors create the binary logistic regression model as to find the optimal vector $\beta = [\beta_0, \beta_1, \dots, \beta_n]$ that best fits

$$y = \begin{cases} 1, & \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon > 0 \\ 0, & \text{otherwise} \end{cases}$$

here ε represents the error.

Evidently, in (1) t can be regarded as a variable, which is a linear weighted combination of independent variable $x = [x_1, \dots, x_n]$ as

$$t \leftarrow \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n.$$

Thus, the binary logistic model is [21]:

$$F(x) = \frac{1}{1 + e^{-t(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n)}} \qquad (6)$$

where $F(x)$ represents the probability of dependent variable $y = 1$.

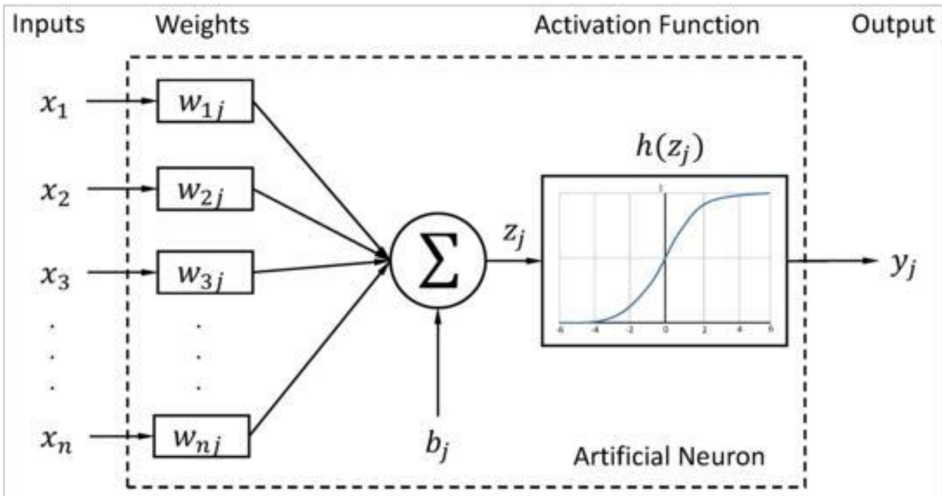

Figure 1: Nonlinear, parametrized function with restricted output range [45].

Training a multilayer perceptron with algorithms employing global search strategies has been an important research direction in the field of neural networks.

Multi–layer perceptrons are feed forward neural networks featuring universal approximation properties used both in regression problems.

The standard feed forward networks with only a single hidden layer can approximate any continuous function uniformly on any compact set and any measurable function to any desired degree of accuracy [22]–[25], [5], [40].

The nonlinear, parametrized function with restricted output range is visualized on Fig.1.

IJERMS
International Journal of Engineering Researches and Management Studies

It is straightforward to extend this analysis to networks with multiple hidden layers.
For recurrent neural networks are typical:
a) stable outputs may be more difficult to evaluate;
b) unexpected behavior (chaos, oscillation).
A survey of neural transfer activation functions can be found in [26].
Moreover, the nodes in the hidden layer are supposed to have a sigmoidal activation function which may be one of the following:

a) logistic sigmoid

$$\varphi_1(net) = \frac{1}{1+e^{-\beta net}}; \qquad (7)$$

b) hyperbolic tangent

$$\varphi_2(net) = \frac{e^{\beta net}-e^{-\beta net}}{e^{\beta net}+e^{-\beta net}}; \qquad (8)$$

c) half hyperbolic tangent

$$\varphi_3(net) = \frac{1-e^{-\beta net}}{1+e^{-\beta net}}; \qquad (9)$$

d) Parametric Fibonacci hyperbolic tangent

$$\varphi_4(net) = \frac{\Psi^{\beta net}-\Psi^{-\beta net}}{\Psi^{\beta net}+\Psi^{-\beta net}}; \qquad (10)$$

e) Parametric Soboleva' modified hyperbolic tangent

$$\varphi_5(net) = \frac{e^{cnet}-e^{-dnet}}{e^{cnet}+e^{-dnet}}; \qquad (11)$$

where net denotes the input to a node and β, c and d are the slope parameters of the sigmoids.

Definition 1. *The sign function of a real number t is defined as follows:*

$$sgn(t) = \begin{cases} -1, & \text{if} \quad t < 0, \\ 0, & \text{if} \quad t = 0, \\ 1, & \text{if} \quad t > 0. \end{cases} \qquad (12)$$

Definition 2. *[27], [28] The Hausdorff distance (the H–distance) [27] $\rho(f,g)$ between two interval functions f, g on $\Omega \subseteq \mathbb{R}$, is the distance between their completed graphs $F(f)$ and $F(g)$ considered as closed subsets of $\Omega \times \mathbb{R}$. More precisely,*

$$\rho(f,g) = \max\{ \sup_{A\in F(f)} \inf_{B\in F(g)} ||A-B||, \sup_{B\in F(g)} \inf_{A\in F(f)} ||A-B||\}, \qquad (13)$$

wherein $||.||$ is any norm in $\mathbb{R}^2$, e. g. the maximum norm $||(t,x)|| = max\{|t|,|x|\}$;
hence the distance between the points $A = (t_A, x_A)$, $B = (t_B, x_B)$ in $\mathbb{R}^2$ is
$||A-B|| = max\,(|t_A - t_B|, |x_A - x_B|)$.

In [29]–[34], [38] the authors consider some families of recurrence generated parametric activation functions on the base of (7)–(11).

3. MAIN RESULTS. A FAMILY OF RECURRENCE GENERATING ACTIVATION FUNCTIONS BASED ON GUDERMANN FUNCTION

A definition for the Gudermann function is [42]:

$$gd(x) = \int_0^x \frac{1}{cosh(t)} dt.$$

IJERMS
International Journal of Engineering Researches and Management Studies

The Gudermannian is named after *Christoph Gudermann (1798–1852).*
We define the following family of modified half Gudermann activation functions (FMHGUDAF):

$$\begin{aligned} g_{i+1}(t) &= \frac{4}{\pi} arctg\left(e^{\frac{\pi}{2}(t+g_i(t))}\right) - 1; \;\; i = 0,1,2,\ldots, \\ g_0(t) &= \frac{4}{\pi} arctg\left(e^{\frac{\pi}{2}t}\right) - 1; \;\; g_0(0) = 0. \end{aligned} \quad (14)$$

Evidently, $g_{i+1}(0) = 0$ for $i = 0,1,2,\ldots$.

Approximation Issues
In this Section we prove lower estimate for the Hausdorff approximation of the sign function by means of this family.

Denote the number of recurrences by p.
The Hausdorff distance $d_p(sgn(t), g_p(t))$ between the sgn function and the function $g_p(t)$ satisfies the following nonlinear equation:

$$g_p(d_p) = \frac{4}{\pi} arctg\left(e^{\frac{\pi}{2}(t+g_{p-1}(d_p))}\right) - 1 = 1 - d_p. \quad (15)$$

The following Theorem gives lower bound for d_p

Theorem 3.1. *For the Hausdorff distance d_p between the sgn function and the function $g_p(t)$ the following bound hold for $p \geq 0$:*

$$d_{l_p} < d_p, \quad (16)$$

where

$$d_{l_p} = \frac{1}{2}\left(p^3 + 7p^2 + 16p + 12 - \sqrt{p^6 + 14p^5 + 81p^4 + 248p^3 + 420p^2 + 364p + 120}\right). \quad (17)$$

Proof. We define the functions

$$F_p(d_p) = g_p(d_p) - 1 + d \quad (18)$$

and

$$G_p(d_p) = -1 + (2+p)d_p - \frac{1}{(p+2)(p+3)} d_p^2. \quad (19)$$

From Taylor expansion we find (see, Fig.2 and Fig.3)

$$F_p(d_p) - G_p(d_p) = O(d_p^2)$$

i.e. the function G_p approximates the function F_p with $d_p \to 0$ as $O(d_p^2)$.

IJERMS
International Journal of Engineering Researches and Management Studies

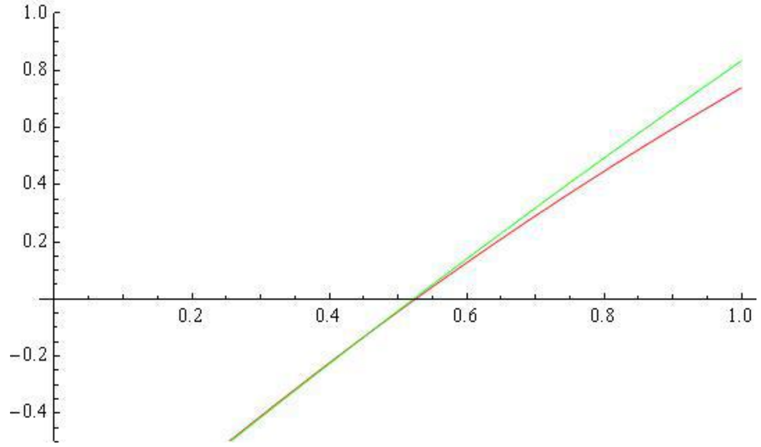

Figure 2: The functions $F_p(d_p)$ ***– (red) and*** $G_p(d_p)$ ***– (green) for*** $p = 0$.

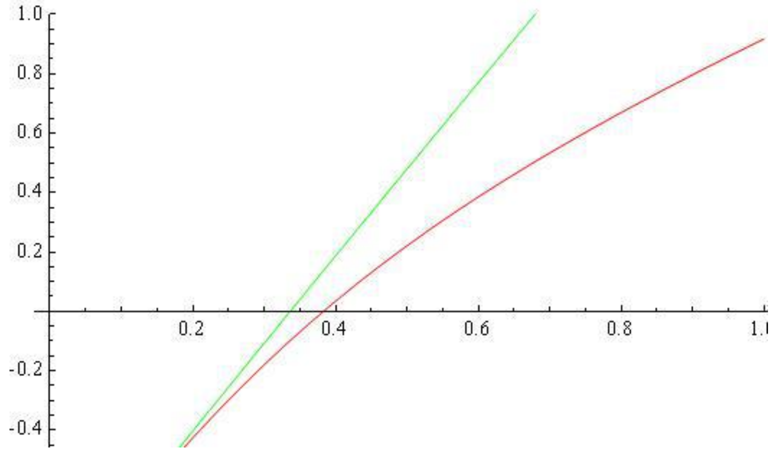

Figure 3: The functions $F_p(d_p)$ ***– (red) and*** $G_p(d_p)$ ***– (green) for*** $p = 1$.

In addition $G'_p(d_p) > 0$ and the second derivative $G''_p(d_p) = -\frac{2}{(p+2)(p+3)} < 0$ has a constant sign on $(0,1)$.

Evidently, the smallest positive root d_{l_p} of the quadratic equation

$$-1 + (2+p)d_p - \frac{1}{(p+2)(p+3)} d_p^2 = 0$$

is lower bound for d_p.

This completes the proof of the Theorem.

Some computational examples are presented in Table 1.

The last column of Table 1 contains the values of d computed by solving the nonlinear equation (15).

The recurrence generated (FMHGUDAF)–functions: $g_0(t)$, $g_1(t)$, $g_2(t)$ and $g_3(t)$ are visualized on Fig.4.

IJERMS
International Journal of Engineering Researches and Management Studies

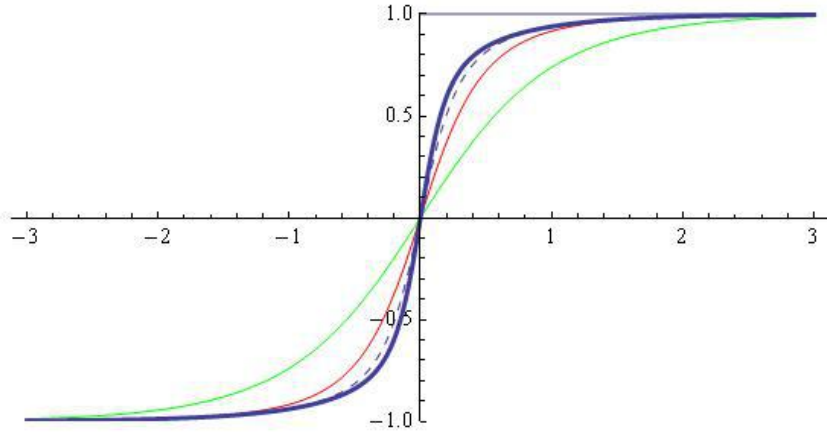

Figure 4: Approximation of the $sgn(t)$ by (FMHGUDAF); $g_0(t)$ (green) – Hausdorff distance: $d = 0.525598$; $g_1(t)$ (red) – Hausdorff distance: $d = 0.383098$; $g_2(t)$ (dashed) – Hausdorff distance: $d = 0.320107$; $g_3(t)$ (thick) – Hausdorff distance: $d = 0.289012$.

Table 1: Bounds for d_p for various p.

p	d_{l_p} from (17)	d_p from (15)
0	0.5227744	0.5255982
1	0.3364783	0.3830983
2	0.2507862	0.3201070
3	0.2002674	0.2890117
4	0.1667770	0.2736509
5	0.1429092	0.2664330
6	0.1250271	0.2632390
	. .	. .
20	0.0454547	0.2609638

APPENDIX

We define the following family of modified parametric Gudermann activation functions for $b > 0$:

$$\begin{aligned} h_{i+1}(t) &= \frac{4}{\pi} arctg\left(e^{\frac{\pi}{2}\frac{1}{b}(t+h_i(t))}\right) - 1; \quad i = 0,1,2,\ldots, \\ h_0(t) &= \frac{4}{\pi} arctg\left(e^{\frac{\pi}{2}\frac{1}{b}t}\right) - 1; \quad h_0(0) = 0. \end{aligned} \tag{20}$$

Evidently, $h_{i+1}(0) = 0$ for $i = 0,1,2,\ldots$.

Denote the number of recurrences by p.

The H-distance $d_p(sgn(t), h_p(t))$ between the sgn function and the function $h_p(t)$ satisfies the following nonlinear equation:

ISSN: 2394-7659
IMPACT FACTOR- 2.789

International Journal of Engineering Researches and Management Studies

$$h_p(d_p) = \frac{4}{\pi} arctg\left(e^{\frac{\pi}{2}\frac{1}{b}(t+h_{p-1}(d_p))}\right) - 1 = 1 - d_p. \qquad (21)$$

The recurrence generated functions are visualized on Fig.5 – Fig.6.

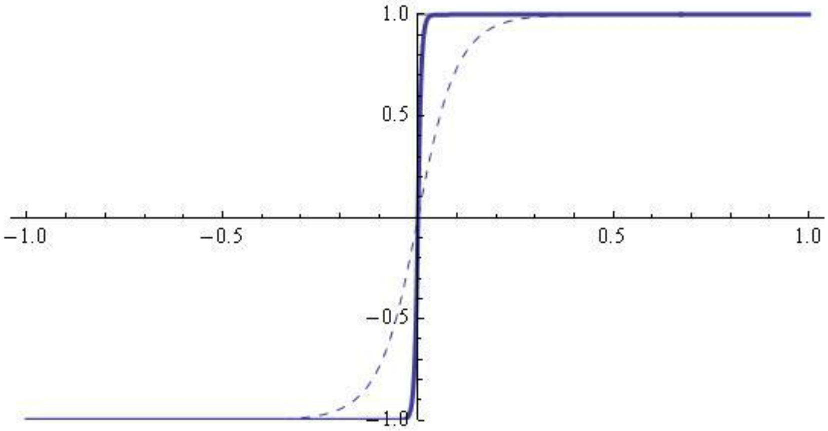

Figure 5: Approximation of the $sgn(t)$ by family (20) for $b = 0.1$; $h_0(t)$ (dashed) – Hausdorff distance: $d = 0.140197$; $h_1(t)$ (thick) – Hausdorff distance: $d = 0.0235623$.

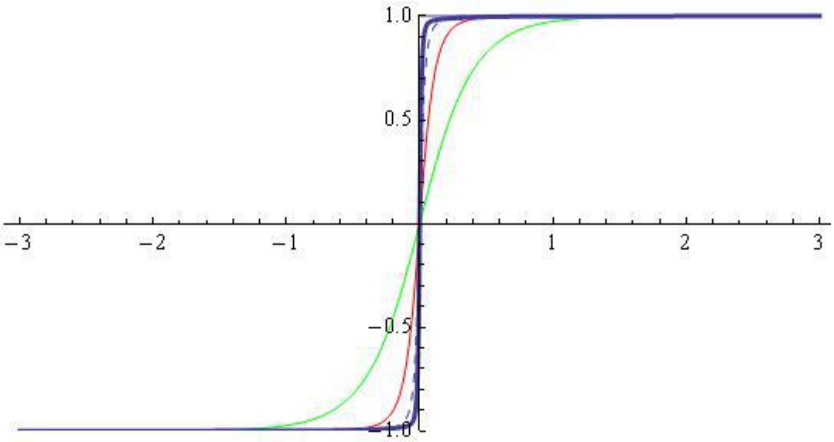

Figure 6: Approximation of the $sgn(t)$ by family (20) for $b = 0.4$; $h_0(t)$ (green) – Hausdorff distance: $d = 0.33446$; $h_1(t)$ (red) – Hausdorff distance: $d = 0.158046$; $h_2(t)$ (dashed) – Hausdorff distance: $d = 0.0842263$; $h_3(t)$ (thick) – Hausdorff distance: $d = 0.0488646$.

From the graphics it can be seen that the ”saturation” is faster.
Based on the methodology proposed in the present note, the reader may formulate the corresponding approximation problems on his/her own.
For the Hausdorff distance d_p for fixed $b = 0.1$ from (21) we have:

Table 2: Bounds for d_p from (21) for various p.

p	d_p
0	0.1401974
1	0.0235623
2	0.0035771
3	0.0005039
4	0.0000686
5	0.0000093
6	0.0000013
7	0.0000003

4. CONCLUSION

A family of modified parametric Gudermann activation functions based on Gudermann function is introduced finding application in neural network theory and practice.

Theoretical and numerical results on the approximation in Hausdorff sense of the sgn function by means of functions belonging to the family are reported in the paper.

We propose a software module within the programming environment *CAS Mathematica* for the analysis of the considered family of recurrence generated (FPSMHTAF) functions.

```
Clear[b]
Manipulate[Dynamic@Show[Plot[f[t], {t, -2, 2},
    LabelStyle → Directive[Green, Bold],
    PlotLabel → 4 / Pi * ArcTan[Exp[Pi / 2 * 1 / b * t]] - 1],
   PlotRange → {Automatic, {-1, 1}}], {{b, 0.01}, 0.01, 10,
  Appearance → "Open"},
 Initialization :→ (f[t_] := 4 / Pi * ArcTan[Exp[Pi / 2 * 1 / b * t]] - 1)]
```

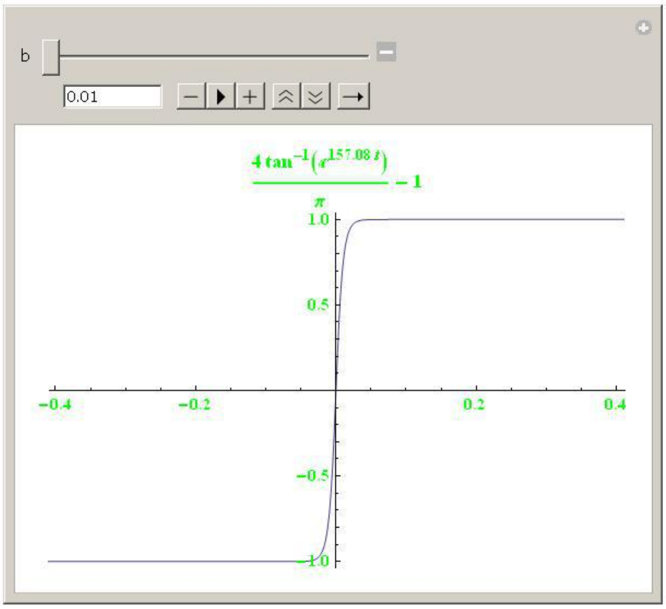

Figure 7: Software module.

The module offers the following possibilities:

- generation of the activation functions under user defined values of the parameter p – number of recursions;
- calculation of the H-distance d_p, $p = 0,1,2, \ldots$, between the sgn function and the activation functions $g_0, g_1, g_2, \ldots, g_p$ and $h_0, h_1, h_2, \ldots, h_p$ respectively;
- software tools for animation and visualization.

For other results, see [35]–[39].

We will explicitly say that the results have independent significance in the study of issues related to neural networks.

Some techniques for recurrence generating of families of activation functions can be found in [41].

5. ACKNOWLEDGMENTS

This work has been supported by the project FP17–FMI–008 of Department for Scientific Research, PaisiiHilendarski University of Plovdiv.

REFERENCES

1. *E. V. Soboleva, V. V. Beskorovainyi, The utility function in problems of structural optimization of distributed objects, Kharkiv University of Air Force, 121, 2008.*
2. *E. V. Soboleva, The S-shaped utility function of individual criteria for multi-objective decision–making in design, Kharkiv National University of Radioelectronics, 247, 2009.*
3. *E. V. Soboleva, V. V. Beskorovainyi, Identification of utility functions in multi–objective choice modelling by using S–shaped functions (PDF), Kharkiv National University of Radioelectronics, 2010, 50–54.*
4. *V. I. Tuev, M. V. Uzhanin, Using modified hyperbolic tangent function to approximate the current–voltage characteristics of field–effect transistors, Tomsk Polytechnic University, 2009, 135–138.*
5. *S. Haykin, Neural networks and learning machines, – 3rd ed., Copyright by Pearson Education, Inc., Upper Saddle River, New Jersey 07458, 2009.*
6. *A. Stakhov, I. Tkachenko, Hyperbolic Fibonacci Trigonometry, Dokl. Akad. NaukUkrainy, 7, 1993, 9–14.*
7. *Z. Trzaska, On Fibonacci hyperbolic trigonometry and modified numerical triangles, Fib. Quart., 34, 1996, 129–138.*
8. *N. Guliyev, V. Ismailov, A single hidden layer feedforward network with only one neuron in the hidden layer san approximate any univariate function, Neural Computation, 28, 2016, 1289–1304.*
9. *D. Costarelli, R. Spigler, Approximation results for neural network operators activated by sigmoidal functions, Neural Networks, 44, 2013, 101–106.*
10. *D. Costarelli, G. Vinti, Pointwise and uniform approximation by multivariate neural network operators of the max-product type, Neural Networks, 2016, doi:10.1016/j.neunet.2016.06.002*
11. *D. Costarelli, R. Spigler, Solving numerically nonlinear systems of balance laws by multivariate sigmoidal functions approximation, Computational and Applied Mathematics 2016; doi:10.1007/s40314-016-0334-8*
12. *D. Costarelli, G. Vinti, Convergence for a family of neural network operators in Orlicz spaces, MathematischeNachrichten, 2016, doi:10.1002/mana.20160006*
13. *J. Dombi, Z. Gera, The Approximation of Piecewise Linear Membership Functions and Lukasiewicz Operators, Fuzzy Sets and Systems, 154 (2), 2005, 275–286.*
14. *I. A. Basheer, M. Hajmeer, Artificial Neural Networks: Fundamentals, Computing, Design, and Application, Journal of Microbiological Methods, 43, 2000, 3–31, doi:10.1016/S0167-7012(00)00201-3*
15. *Z. Chen, F. Cao, The Approximation Operators with Sigmoidal Functions, Computers & Mathematics with Applications, 58, 2009, 758–765, doi:10.1016/j.camwa.2009.05.001*
16. *Z. Chen, F. Cao, The Construction and Approximation of a Class of Neural Networks Operators with Ramp Functions, Journal of Computational Analysis and Applications, 14, 2012, 101–112.*
17. *Z. Chen, F. Cao, J. Hu, Approximation by Network Operators with Logistic Activation Functions, Applied Mathematics and Computation, 256, 2015, 565–571, doi:10.1016/j.amc.2015.01.049*
18. *D. Costarelli, R. Spigler, Constructive Approximation by Superposition of Sigmoidal Functions, Anal.*

Theory Appl., 29, 2013, 169–196, doi:10.4208/ata.2013.v29.n2.8

19. *D. Elliott, A better activation function for artificial neural networks, the National Science Foundation, Institute for Systems Research, Washington, DC, ISR Technical Rep. TR–6, 1993.*
20. *K. Babu, D. Edla, New algebraic activation function for multi-layered feed forward neural networks. IETE Journal of Research, 2016, doi:10.1080/03772063.2016.1240633*
21. *S. Wang, T. Zhan, Y. Chen, Y. Zhang, M. Yang, H. Lu, H. Wang, B. Liu, P. Phillips, Multiple Sclerosis Detection Based on Biorthogonal Wavelet Transform, RBF Kernel Principal Component Analysis, and Logistic Regression, IEEE Access, Special section on advanced signal processing methods in medical imaging, 4, 2016, 7567–7576.*
22. *G. Cybenko, Approximation by superposition of a sigmoidal function, Math. of Control Signals and Systems, 2, 1989, 303–314.*
23. *K. Hornik, M. Stinchcombe, H. White, Multi-layer feed forward networks are universal approximations, Neural Networks, 2, 1989, 359–366.*
24. *V. Kreinovich, O. Sirisaengtaksin, 3–layer neural networks are universal approximations for functionals and for control strategies, Neural Parallel and Scientific Computations, 1, 1993, 325–346.*
25. *H. White, Connectionist nonparametric regression: multilayer feedforward networks can learn arbitrary mappings, Neural Networks, 3, 1990, 535–549.*
26. *W. Duch, N. Jankowski, Survey of neural transfer functions, Neural Computing Surveys, 2, 1999, 163–212.*
27. *F. Hausdorff, Set Theory (2 ed.) (Chelsea Publ., New York, (1962 [1957]) (Republished by AMS-Chelsea 2005), ISBN: 978–0–821–83835–8.*
28. *B. Sendov,Hausdorff Approximations,Kluwer, Boston, 1990, doi:10.1007/978-94-009-0673-0*
29. *N. Kyurkchiev, A family of recurrence generated sigmoidal functions based on the Verhulst logistic function. Some approximation and modeling aspects, Biomath Communications, 3 (2), 2016, 18 pp.*
30. *A. Iliev, N. Kyurkchiev, S. Markov, A family of recurrence generated parametric activation functions with applications to neural networks, International Journal on Research Innovations in Engineering Science and Technology, 2 (1), 2017, 60–68.*
31. *N. Kyurkchiev, S. Markov, Hausdorff Approximation of The Sign Function by a Class of Parametric Activation Functions, Biomath Communications, 3 (2), 2016, 14 pp., doi:10.11145/bmc.2016.12.217*
32. *N. Kyurkchiev, A. Iliev, S. Markov, Families of recurrence generated three and four parametric activation functions, Int. J. Sci. Res. and Development, 4 (12), 2017, 746-750.*
33. *V. Kyurkchiev, N. Kyurkchiev, A family of recurrence generated functions based oh Half-hyperbolic tangent activation functions, Biomedical Statistics and Informatics, 2 (3), 2017, 87-94.*
34. *V. Kyurkchiev, A. Iliev, N. Kyurkchiev, On some families of recurrence generated activation functions, Int. J. of Sci. Eng. And Appl. Sci., 3 (3), 2017.*
35. *N. Kyurkchiev, S. Markov, Sigmoid functions: Some Approximation and Modelling Aspects, LAP LAMBERT Academic Publishing, Saarbrucken, 2015, ISBN 978-3-659-76045-7.*
36. *N. Kyurkchiev, S. Markov, On the Hausdorff distance between the Heaviside step function and Verhulst logistic function. J. Math. Chem., 54 (1), 2016, 109–119, doi:10.1007/S10910-015-0552-0*
37. *N. Kyurkchiev, A. Iliev, On the Hausdorff distance between the Heaviside function and some transmuted activation functions, Mathematical Modelling and Applications, 2 (1), 2016, 1-5.*
38. *N. Kyurkchiev, A. Iliev, A note on the new Fibonacci hyperbolic tangent activation function, Int. J. of Innovative Science, Engineering and technology, 4 (5), 2017, 364–-368.*
39. *A. Golev, A. Iliev, N. Kyurkchiev, A Note on the Soboleva' Modified Hyperbolic Tangent Activation Function, International Journal of Innovative Science Engineering and Technology, 4 (6), 2017, 177–182.*
40. *R. Rojas, Neural Networks. A Systematic Introduction, Springer, 1996.*
41. *N. Kyurkchiev, A. Iliev, S. Markov, Some techniques for recurrence generating of activation functions, LAP LAMBERT Academic Publishing, Balti, 2017, ISBN 978-3-330-33143-3.*
42. *F. W. J. Olver, D. W. Lozier, R. F. Boisvert, C. W. Clark, NIST Digital Library of Mathematical Functions, Cambridge University Press, 2010, 951 pp.*
43. *K. Goulianas, A. Margaris, I. Refanidis, K. Damantaras, Solving polynomial systems using a fast adaptive back propagation-type neural network algorithm, European Journal of Appl. Math., 2017, doi:10.1017/S0956792517000146*
44. *V. Kurkova, M. Sanguineti, Probabilistic lower bounds for approximation by shallow perceptron networks, Neural Networks, 91, 2017, 34–41.*
45. *S. Ryu, J. Noh, H. Kim, Deep neural networks based demand side short term load forecasting, Energies, 10*

IJERMS
International Journal of Engineering Researches and Management Studies

(1), 2017.
46. *P. Petrushev, Approximation by ridge functions and neural networks, SIAM J. Math. Anal., 30, 1999, 155–189.*
47. *S. B. Lin, F. L. Cao, Z. B. Hu, Essential rate for approximation by spherical neural networks, Neural Networks, 24, 2011, 752–758.*
48. *Shao-Bo Lin, Limitations of shallow nets approximation, Neural Networks, 2017, doi:10.1016/j.neunet.2017.06.016*
49. *R. Anguelov, M. Borisov, A. Iliev, N. Kyurkchiev, S. Markov, On the chemical meaning of some growth models possessing Gompertzian-type property, Math. Meth. Appl. Sci., 2017, doi:10.1002/mma.4539*