ISSN No:-2456-2165

Advanced Soft Computing Technique: Introduction of Thousands of Non-linear Specific Activation Functions into the Deep Learning Artificial Neural Networks

Jamilu Auwalu Adamu
Mathematics Programme, 118 National Mathematical Centre, 904105, FCT-Abuja, Nigeria

Correspondence: Jamilu Auwalu Adamu

Abstrct:- Non-linear activation functions play an extremely crucial role in neural networks by introducing non-linearity. This nonlinearity allows neural networks to develop complex representations and functions based on the inputs that would not be possible with a simple linear function. Without a non-linear activation function in the network, a neural network, no matter how many layers it had, would behave just like a single-layer perceptron. So, why is increasing non-linear specific activation functions desired? What effect do they have on the overall performance of the network? The author applied the modified neural network architectures proposed in Jamilu Adamu (2019) to linked the most volatile Chicago City daily maximum temperature data exported from NASA POWER| Data Access from 31st March, 2021 to 31st March, 1981 (roughly 40 years). With the help of Jameel's ANNAF Deterministic and Stochastic Criterion, 86 specific activation functions were generated using Superneunet software prototype developed by the author. With the help of QI Macros, Normality Test confirmed Chicago maximum temperature data has being Non-normal, volatile and unpredictable. To assess the performances between the trial and error traditional activation functions with that of modified, Actual and Initial Predicted values of Chicago City maximum temperature were correlated, Tanh (0.947) outperformed ReLu (0.945) and Sigmoid (0.940) using traditional trial and error activation functions. While, with the modified specific activation functions, Cubic-Exponential-Gaussian-Gaussian (0.928) outperformed Exponential- Exponential-Linear (0.910) and Cubic-Exponential (0.923). Note that with minimum number of layers and nodes, the activation functions were selected arbitrarily and kept constant throughout the hidden layers of Superneunet prototype. With automation, the modified functions can outperform the traditional out rightly. The results of the study suggests that the technique can be used to specifically, accurately and precisely predict any time series application and the prices of over 35 million instruments across all asset classes traded 24/7 in 160 countries aggregated from 330 exchanges of Bloomberg Terminal and other world's largest financial and weather assets terminals as against the current practice that uniformly uses traditional trial and error few activation functions across all areas of application. Now the current version of Superneunet software prototype accommodated up to 100+ specific activation functions and can be extended to 500 or even thousands activation functions per area of application. So, what will happen to our predictions or forecasts if we extend Superneunet to accommodate millions or billions of specific activation functions? Do we need to upgrade Superneunet to accommodate this Quantum Computers application? One of the fundamental benefits of this novel breakthrough, it can waive years of research by researchers around the globe for constructing single desirable activation function, with it can be done by just clicking of a button. This can completely eliminate phobia and psychological trauma experiences when selecting the right activation functions, thus can fantastically save cost, time, energy, simultaneously accelerate AI technology production and global economic growth. To crown it up, the humanity can benefit from the results of this study through Environment, Finance, Education, Health & Pharmaceuticals, Transportation & Security, Science & Technology, and Innovation & Creativity.

Keywords:- *Software, Artificial Intelligence, Chicago City, Superneunet, Traditional, Jameel's ANNAF Criterion.*

I. INTRODUCTION

The world is becoming more chaotic (non-linear), full with big data due to rapid technological and environmental changes to the extent that the few traditional un-specific Sigmoid, Tanh, ReLu among others cannot best model estimates across different areas of application. For instance, same traditional trial and error activation functions are apply for temperature forecasting and also in stock price predictions. In reality, this should not be so. The need for thousands of nonlinear specific activation functions emanating from input data depending on the specific area of application, time change and rules of thumb.

The existing neural network architectures/software uses trial and error activation functions such as Sigmoid, Tanh, ReLU among others. These activation functions rely

ISSN No:-2456-2165

much more on the experts' assumptions not the input datasets. They are arbitrarily use regardless of areas of application, for instance they are the same set apply for temperature forecasts and also in stock price predictions. For instance the top highest rated neural network software such Neural Designer, Neuroph and Darknet uses the trial and error traditional approach.

Jamilu Adamu (2019) proposed that the choice of deep learning artificial neural networks activation functions should basically depend on input datasets, time change, specific area of application and particular rules of thumb, such that each area of application would have its own set of activation functions and to generate larger sets against the current practice thereby increasing non-linearity, speed, accuracy, precision, cost-effectiveness and to drastically reduce computational time consumption.

The proposed study presented an approach named 'modified neural network architectures' that linked the input data with the neural network activation functions with the help of Jameel's ANNAF Deterministic and Stochastic Criterion, this led the author to created Superneunet software prototype to increase non-linearity with increase of more non-linear activation functions to get best optimal results forecasts and predictions. For instance Microsoft and Google stocks have different activation functions; For Microsoft stock's activation functions, the Input Data = Microsoft Stock Data; Area of Application = Microsoft Company; and Time Change = Microsoft Length of Time Series. Similarly, For Google stock's activation functions, the Input Data = Google Stock Data; Area of Application = Google Company; and Time Change = Google Length of Time Series, and so on.

First, to experiment, the author created Topneunet a research-oriented time series forecasting software to test the traditional neural network architectures that uses trial and error activation functions such as Sigmoid, Tanh, ReLU to compare it with the modified specific activation functions of the proposed study. Subsequently, Superneunet software prototype was created to accommodate 100 specific activation functions and can be extended to generate up to 500 or even thousands activation functions per area of application using deterministic component and a lot more with stochastic component.

Superneunet can execute complex predictive computing tasks for specify applications of input datasets strictly without human intervention or trial and error. It generate beautiful activation functions emanate from input datasets, drastically reduce time consumption, energy, resources and cost while selecting the right activation functions, automatically find the best fitted activation functions just by click of a button, provide high level of prediction accuracy, precision and transparency, provide users flexibility to choose probabilistic input datasets for an application whose physical phenomenon is stochastic in nature, and to choose deterministic input datasets for an application whose physical phenomenon is deterministic (non-probabilistic) in nature, automatically select heavily - skewed and fat-tailed activation functions excellent for predicting random movements, consists of two major components, namely; the existing neural network software (standard) plus custom component, Superneunet gave birth to the modern network network architecture, it showcases the high demand for the killer quantum computers applications.

There is a popular quote that says *"if you don't like the weather in Chicago, wait five minutes"*.Chicago weather is so unpredictable even the weatherman has no idea. This paper will use Chicago City of the United States of America historical maximum temperature data exported from POWER| Data Access Viewer as from 31st March, 2021 to 31st March, 1981 (roughly 40 years) as the deep neural network input data. With the help of QI Macros software, a Normality Test will be carried out to check the normality of the input data. Two artificial intelligence time series software were developed by the author, the TOPNEUNET and SUPERNEUNET to assess the performances between the trial and error traditional activation functions and that of modified. Finally, the paper round up with highlights on the proposed study benefits, future research and advantages of the automation of upgrade and automation of SUPERNEUNET software.

II. METHODS

This paper will apply the following criterion as the *'particular rules of thumb'* to generate the modified neural network activation functions.

➢ *Jameel's ANNAF Stochastic Criterion*

ANNAF means Artificial Neural Network Activation Functions. Under this criterion, we run the goodness of fits test on the AI Dataset such that:

- **Distances between AI Datasets:** The goodness of fit test on the Referenced AI Dataset and or Training Dataset shall be conducted across three (3) fundamental AI Dataset Lengths as well as **at least Four (4) Strategic lengths** within the full length of the Referenced AI Dataset. Under this criterion, one could perhaps run the test across different Data scenarios, the daily, weekly, monthly, quarterly, biannually and or annually using the following AI Dataset partitioning strategy diagram for the choice of an optimized Activation Functions:

Therefore, the Set of Distances between the AI Datasets (and or Training Datasets) is given by:

D = {Distance D1, Distance D2, Distance D3, Strategic Distance 1, Strategic Distance 2, Strategic Distance 3, Strategic Distance 4,… , Distance n};

- We accept if the Average of the ranks of Kolmogorov Smirnor, Anderson Darling and Chi-squared is less than or equal to Three (3);

- We must choose the fat-tailed Probability Distribution follows by the AI Dataset Itself regardless of its Rankings;
- If there is tie, we include both the fat-tailed Probability Distributions in the selection;
- At least Two (2) fat-tailed Probability Distributions must be included in the selection;
- We select the most occur Probability Distribution as the qualify candidate in each case of test of goodness of fit on our referenced AI Dataset;
- **Backward Propagation Axiom:** Having chosen the most qualified fat-tailed Probability Distribution in the final round up selection for instance, $f_{final}(x)$, then $f_{final}(x)$ shall be monotone continuously differentiable function. Particularly if $f_{final}(x)$ is not differentiable then discard it and repeat (1) to (6) until we have a monotone differentiable function. Note the Approach is from top-bottom process;
- **Criterion Enhancement Axiom:** Thode (2012) intensively discussed about the Best Goodness of Fit Tests such as Kolmogorov Smirnov (KS) Test, Anderson-Darling Test, Jarque and Bera (JB) Test, Shapiro Wilk (SW) Test, Cramer-Von Mises Test, Pearson $\left(\chi^2\ Godness\ of\ Fit\right)$ Test, Lilliefors Corrected K-S Test, D'AgostinoSkewness Test, Anscombe-Glynn Kurtosis Test, D'Agostino-Pearson and Omnibus Test. Let $\{T_1, T_2, ..., T_n\}$ be the set of such Best Goodness of Fit Tests, $\{x_1, x_2, ..., x_n\}$ be their **RANKS** respectively then the generality of (i) can be re-expressed (or enhanced) if $\frac{(x_1 + x_2 + ... + x_n)}{n} \leq a$, where $0 < a \leq n,\ n \in N$ or equivalently, $x_1 + x_2 + ... + x_n \leq an$,and;
- **Unit Axiom:** let $f_{final}(x)$ be such that it satisfied axioms (1) to (6) and or (7). Let $\{r_1, r_2, ..., r_n\}$ be the ranks of fitness test of $f_{final}(x)$ obtained from the tests $\{T_1, T_2, ..., T_n\}$ respectively then if $\forall i \in \{1, 2, ..., n\}$, $r_i = 1$ regardless of other factors. Consequently, if for all fitness test runs, turn out to be the same $f_{final}(x)$ then $f_{final}(x)$ will gives Deep Learning Artificial Neural Networks super-intelligent capabilities.

➢ *Jameel's ANNAF Deterministic Criterion*

For a Neural Network that require Deterministic Activation Functions can satisfy the following proposed criterion:

- The function $f(x)$ shall emanate from the referenced AI Dataset. The essence of the function $f(x)$ to be emanated from the referenced AI Dataset is to build an incredible and sophisticated Activation Functions that have the Best Match the referenced AI Dataset, since neural network is a system made to learn a function from data;.
- **Distances between AI Datasets:** The goodness of fit test on the Referenced AI Dataset and or Training Dataset shall be conducted across three (3) fundamental AI Dataset Lengths as well as at least Four (4) Strategic lengths within the full length of the Referenced AI Dataset. Under this criterion, one could perhaps run the test across different Data scenarios, the daily, weekly, monthly, quarterly, biannually and or annually using the following AI Dataset partitioning strategy diagram for the choice of an optimized Activation Functions:

Therefore, the Set of Distances between the AI Datasets (and or Training Datasets) is given by:

D = {Distance D1, Distance D2, Distance D3, Strategic Distance 1, Strategic Distance 2, Strategic Distance 3, Strategic Distance 4,… , Distance n};

- A curve fitting for Best Fitted Deterministic Function shall be carried out, we choose the function $f(x)$ whose:

✓ Rank is Unity (1)
✓ Fattiness Standard Error is smaller than any other on the list;

- We must choose the function $f(x)$ follows by our referenced **AI Dataset itself** regardless of its Rankings;
- If there is tie, we include both the functions $f(x)s$ in the selection;
- At least Two (2) functions $f(x)s$ must be included in the selection;
- We select the most occur function $f(x)$ as the qualify candidate in each case of test of goodness of fit on our **AI Dataset**;

For the **Backward Propagation,** having chosen the most qualified function $f(x)$ in the final round up selection for instance, $f_{final}(x)$, then

- The function $f_{final}(x)$ shall be Nonlinear;
- The function $f_{final}(x)$ shall have a Range;
- The function $f_{final}(x)$ shall be Continuously Differentiable;
- The function $f_{final}(x)$ shall be Monotonic;
- The function $f_{final}(x)$ shall be Smooth Function with a Monotonic Derivative;
- The function $f_{final}(x)$ shall Approximates Identity near the Origin.

If these failed Discard the 1st rated function $f_{final}(x)$, repeat (1) to (12) until the qualified Deterministic Activation Function is emanated from our AI Dataset.

Note: Deep Learning Artificial Neural Network's Hidden and output Layers consist of at least one, two or more Best fitted Activation Functions **emanated** from our

ISSN No:-2456-2165

AI-Data Set, therefore, the rank: unity (one) in (1)-(a) and Fattiness Standard Error (1)-(b) of the criterion means when a function whose Real "**Rank =1**" was chosen and it satisfied (1) to (8) then the next function on list whose Real "**Rank=2**" will assume "**New Rank=1**" and will be tested to satisfy all the eight (8) axioms until we have the required number of BEST Activation Functions needed to carry out our Deep Learning Artificial Neural Network.

III. RESULTS

So, how do we use the proposed Advanced Soft Computing Technique to predict the most volatile Chicago City temperature data of the United States of America to enable us introduce *Thousands of Non-linear Specific Activation Functions into the Deep Learning Artificial Neural Networks* with the help of Superneunet software prototype created by the author?

➤ *Chicago City of the United States of America as an Input Weather Data*

We use POWER| Data Access Viewer historical daily temperature data of Chicago City of the United States of America as from 31st March, 2021 to 31st March, 1981 (roughly 40 years) as the deep neural network input data as summarily shown below:

Table 1 Chicago City Maximum Temperature

YEAR	MO	DY	T2M	T2MDEW	T2MWET	TS	T2M_RANGE	T2M_MAX	T2M_MIN
1981	3	31	10.68	6.14	8.4	6.75	12.93	17.36	4.43
1981	4	1	7.56	2.28	4.92	5.23	6.06	10.35	4.29
1981	4	2	9.73	3.11	6.42	6.33	12.05	15.25	3.2
1981	4	3	11.44	9.55	10.5	7.69	5.53	14.15	8.62
1981	4	4	10.3	6	8.15	6.94	12.58	14.56	1.99
1981	4	5	2.57	-3.23	-0.33	3.69	5.34	5.34	0.01
1981	4	6	4.68	-2.4	1.14	4.52	8.35	8.61	0.26
1981	4	7	9.12	3	6.06	6.76	11.73	14.3	2.57
1981	4	8	10.96	8.47	9.72	7.82	10.86	15.46	4.6
1981	4	9	7.97	1.87	4.91	6.51	8.85	13.01	4.15
1981	4	10	10.3	7.76	9.03	7.65	8.88	14.1	5.22
1981	4	11	7.38	6.18	6.78	6.44	7.8	12.82	5.01
1981	4	12	6.61	5.37	5.99	6.17	3.06	8.06	5

Source: POWER| Data Access Viewer (July 09, 2023)

➤ *Training with Chicago City of the United States of America maximum temperature Data and Traditional Activation Functions*

Here we consider Chicago City of the United States of America daily maximum temperature Data (T2M_MAX) then we use LSTM approach and trial and error *ReLu, Sigmoid and Tanh as activation functions*.

Table 2 Normality Test Results of Chicago City Maximum Temperature Data (T2M_MAX)

Anderson-Darling	Non-Normal at 0.01
A-Squared	136.992
P	0.000
95% Critical Value	0.787
99% Critical Value	1.092
Mean	12.048
Mode	24.440
Standard Deviation	9.806
Variance	96.158
Skewedness	-0.127
Kurtosis	-0.951
N	14611.000
Std Err	0.081
Minimum	-21.240
1st Quartile	3.810
Median	12.120
3rd Quartile	20.780
Maximum	34.690
Range	55.930
Confidence Interval	0.159
for Mean (Mu)	11.889
0.95	12.207
For Stdev (sigma)	9.695

ISSN No:-2456-2165

		9.920
for Median		11.820
		12.390
Normal		-7.359
	0.95	31.455
k-Factor One-sided		-4.288
		28.383
k Two-sided		1.979
k One-sided		1.666
Nonparametric		-6.180
	0.95	27.900

Source: QI Macros (July 09, 2023)

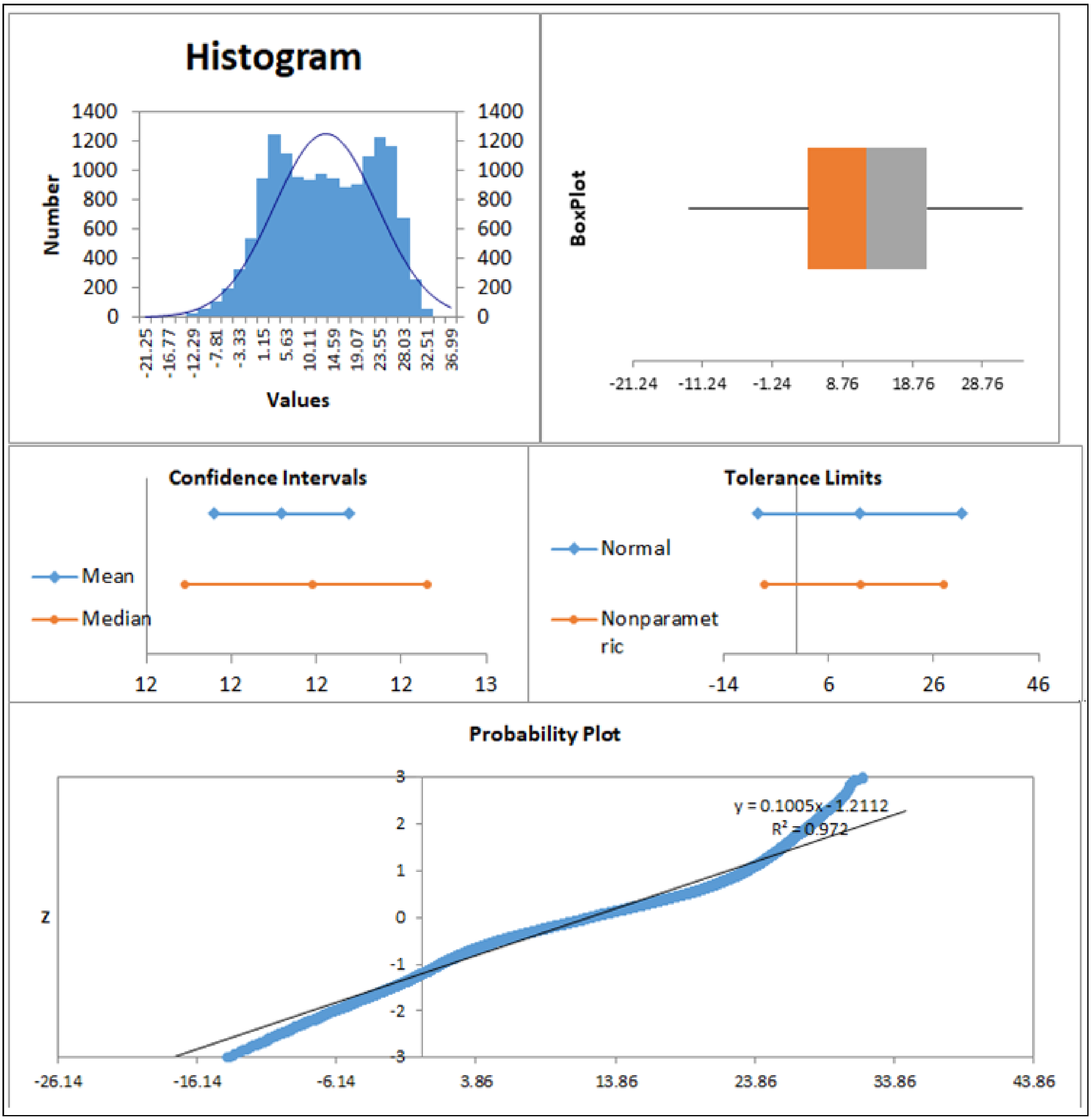

Fig 1 Confirmed that Chicago Daily Maximum Temperature Data is Non-Normal, Volatile and Unpredictable
Source: QI Macros (July 09, 2023)

The results in Table 2 and Figure 1 confirmed that Chicago daily maximum temperature data is Non-normal, volatile and unpredictable.

- *Training with LSTM Approach with ReLu as Activation Function Prediction Results*

Table 3 Training with LSTM Approach with ReLu as Activation Function Prediction Results

	0_actual	**0_prediction**	**1_actual**	**1_prediction**	**2_actual**	**2_prediction**
0	6.92	8.484736	8.68	7.199345	10.86	6.587775
1	0.05	0.155267	6.83	0.618956	4.51	0.801268
2	18.76	11.61664	19.79	11.2219	11.3	10.94524
3	17.73	11.66926	16.76	11.21879	12.9	10.93638
4	9.46	8.957335	2.150001	9.760694	4.2	10.08764
5	8.75	6.666453	2.150001	5.547051	2.43	5.029547
6	0.790001	3.417037	0.730001	3.046682	3.95	2.838913
7	11.88	14.13942	13.72	15.55979	17.65	16.16453
8	14.78	18.68055	14.68	18.51981	14.22	18.38818
9	0.88	0.396488	-2.78	0.356511	-4.08	0.355407
10	11.09	10.46514	11.42	10.97462	12.18	11.17647
11	0.62	1.097413	1.900001	1.804125	3	2.035924
12	2.85	7.063836	2	4.976476	2.4	4.076858
13	14.94	13.54379	8.65	12.9169	7.45	12.59732
14	-2.65	1.956602	-3.02	1.049887	-6.43	0.704674
15	6.58	9.973114	6.61	9.897125	8.38	9.747368
16	28.73	29.42024	33.01	28.80992	31.76	28.44435
17	0.820001	1.85337	-6.77	1.515726	-13.5	1.359329
18	25.58	25.34604	24.54	25.33277	25.61	25.27034
19	9.07	6.523383	8.480001	6.037242	6.58	5.772291

Source: Topneunet Software Prototype (July 09, 2023)

- *Correlation Test between 0_Actual and 0_ Prediction with Relu Activation Function*

Table 4 Correlation Test between 0_Actual and 0_ Prediction with Relu Activation Function

CORREL	**0_autual**	**0_prediction**
0_autual	1.000	0.945
0_prediction	0.945	1.000
p Values	0_autual	0_prediction
Data2	0.000	

Source: QI Macros (July 09, 2023)

- *Training with LSTM Approach with Sigmoid as Activation Function Prediction Results*

Table 5 Training with LSTM Approach with Sigmoid as Activation Function Prediction Results

	0_actual	**0_prediction**	**1_actual**	**1_prediction**	**2_actual**	**2_prediction**
0	25.05	23.85577	22.66	23.21388	21.85	23.04929
1	2.94	7.225395	4.190001	7.101025	3.3	7.175757
2	27.4	22.78512	22.71	22.56461	26.38	22.64298
3	19.12	19.00236	21.17	18.74392	20.48	18.91153
4	-0.63	-0.15477	2.25	0.123631	1.770001	0.52295
5	21.23	20.31094	22.78	20.19241	20.76	20.25626
6	10.85	9.243296	1.73	8.286585	-0.45	8.073393
7	20.26	20.24242	17.01	19.82811	17.4	19.88947
8	23.31	20.60634	23.98	20.37195	22.19	20.30761
9	8.360001	4.790601	10.9	4.15083	8.59	4.191735
10	11.61	10.91484	15.5	10.97463	15.49	11.10027
11	21.83	21.28453	20.58	21.07834	23.77	21.20221
12	25.8	21.94936	29.24	21.52921	30.14	21.41716
13	9.41	11.66666	7.81	12.17102	8.29	12.4581
14	7.9	10.00706	9.51	9.831689	10.32	10.06243
15	25.09	22.12379	18.87	22.21723	19.37	22.3734
16	-8.35	-2.12171	-4.53	-1.7389	-2.69	-1.22676

ISSN No:-2456-2165

17	7.52	14.40508	10.94	14.29342	17.27	14.35114
18	1.73	1.528833	2.37	1.42916	2.86	1.712103
19	4.51	3.38048	5.19	3.606139	8.73	3.910937

Source: Topneunet Software Prototype (July 09, 2023)

➢ *Correlation Test between 0_Actual and 0_ Prediction with Sigmoid Activation Function*

Table 6 Correlation Test between 0_Actual and 0_ Prediction with Sigmoid Activation Function

CORREL	0_actual	0_prediction
0_actual	1.000	0.940
0_prediction	0.940	1.000
p Values	0_actual	0_prediction
0_prediction	0.000	

Source: QI Macros (July 09, 2023)

➢ *Training with LSTM Approach with Tanh as Activation Function Prediction Results*

Table 7 Training with LSTM Approach with Tanh as Activation Function Prediction Results

	0_actual	0_prediction	1_actual	1_prediction	2_actual	2_prediction
0	18.21	12.97369	15.4	11.69404	14.9	11.17294
1	-4.67	-3.7082	-6.12	-2.95023	-6.63	-3.04408
2	23.05	23.85974	26.17	24.01289	26.88	24.13138
3	20.23	23.1936	24.59	22.7925	20.19	22.67771
4	12.73	13.66027	13.4	13.94604	11.83	14.16814
5	2.150001	4.192982	1.92	3.798579	-6.02	3.358323
6	20.69	23.99814	19.57	24.05532	23.41	24.17147
7	14.56	12.19272	11.89	13.63956	12.19	14.17668
8	7.16	6.189599	1.19	6.623359	-1.95	6.694827
9	23.8	23.53896	24.63	23.82232	25.68	23.99921
10	2.139999	2.824577	0.11	1.722054	0.9	0.999502
11	-0.88	1.496425	4.05	1.884069	1.259999	1.813297
12	21.53	21.60676	21.78	22.26805	21.23	22.57961
13	4.63	3.492057	1.7	2.26256	-5.7	1.48843
14	24.45	24.97415	22.94	25.22658	24.09	25.40695
15	9.87	10.51891	14.19	9.693172	11.83	9.346715
16	16.56	11.13028	16.8	11.79261	17.57	12.10385
17	3.270001	3.304918	4.62	4.552684	2.83	4.925143
18	4.97	6.731205	5.500001	5.344305	1.940001	4.560368
19	-0.66	-4.51941	3.26	-3.41612	2.7	-3.34256

Source: Topneunet Software Prototype (July 09, 2023)

➢ *Correlation Test between 0_Actual and 0_ Prediction with Tanh Activation Function*

Table 8 Correlation Test between 0_Actual and 0_ Prediction with Tanh Activation Function

CORREL	0_autual	0_prediction
0_autual	1.000	0.947
0_prediction	0.947	1.000
p Values	0_autual	0_prediction
0_prediction	0.000	

Source: QI Macros (July 09, 2023)

With ***0.945, 0.940 and 0.947*** correlation of ReLu, Sigmoid and Tanh respectively, Tanh outperformed ReLu and Sigmoid with traditional trial and error activation functions.

➢ *Training with Chicago City of the United States of America Maximum Temperature Data with Modified Specific Activation Functions*

Here we consider Chicago City of the United States of America maximum temperature Data then we use Superneunet Software prototype to generate 96 Chicago City daily maximum temperature specific activation functions among which we train with *Cubic-Exponential-Gaussian- Gaussian (65th), Exponential- Exponential-Linear (78th) and Cubic-Exponential (62nd).*

ISSN No:-2456-2165

- *Training with LSTM Approach with Cubic-Exponential-Gaussian- Gaussian (65th) as Chicago City Temperature Specific Activation Function Prediction Results*

Table 9 Training with LSTM Approach with Cubic-Exponential-Gaussian- Gaussian (65th) as Chicago City Temperature Specific Activation Function Prediction Results

	0_actual	0_prediction	1_actual	1_prediction	2_actual	2_prediction
0	22.5	22.52489	21.28	21.2866	20.92	21.231
1	24.5	24.63568	24.1	23.4103	26.28	23.38109
2	20.15	25.41483	21.32	24.14586	19.12	24.10672
3	11.09	13.87306	11.42	12.79602	12.18	12.71551
4	7.3	7.962515	3.56	7.067826	2.150001	6.997157
5	-8.49	0.674926	-1.67	0.118897	-1.42	0.108706
6	15.77	14.59054	18.47	13.49091	10.75	13.40847
7	27.83	27.49851	25.6	26.19706	27.75	26.16745
8	16.7	23.79283	15.95	22.53896	19.91	22.48972
9	2.8	5.718395	1.490001	4.930853	1.16	4.881345
10	20.33	24.57312	23.53	23.32312	25.1	23.28274
11	1.32	7.614257	4.66	6.667022	9.08	6.574102
12	1.07	8.146381	2.050001	7.250598	2.93	7.182032
13	22.33	22.88007	26	21.63523	25.3	21.58094
14	21.85	25.11896	25.81	23.86534	26.66	23.82924
15	0.78	3.1003	3.09	2.437258	3.43	2.411925
16	11.97	6.719452	10.63	5.910301	13.52	5.861311
17	21.55	25.08092	21.45	23.8047	21.73	23.76097
18	24.9	24.4245	22.43	23.17305	20.47	23.13061
19	22.58	22.59311	25.36	21.39988	24.83	21.36255

Source: Superneunet Software Prototype (July 09, 2023)

- *Correlation Test between 0_Actual and 0_ Prediction with Cubic-Exponential-Gaussian- Gaussian (65th) Modified Specific Activation Function*

Table 10 Correlation Test between 0_Actual and 0_ Prediction with Cubic-Exponential-Gaussian- Gaussian (65th) Modified Specific Activation Function

CORREL	**0_actual**	**0_prediction**
0_actual	1.000	0.928
0_prediction	0.928	1.000
p Values	0_actual	0_prediction
0_prediction	0.000	

Source: QI Macros (July 09, 2023)

- *Training with LSTM Approach with Exponential- Exponential-Linear (78th) as Chicago City Temperature Specific Activation Function Prediction Results*

Table 11 Training with LSTM Approach with Exponential- Exponential-Linear (78th) as Chicago City Temperature Specific Activation Function Prediction Results

	0_actual	**0_prediction**	**1_actual**	**1_prediction**	**2_actual**	**2_prediction**
0	6.01	6.55399	7.65	6.790298	5.36	6.202929
1	24.33	24.53516	25.98	24.57447	27.6	23.86232
2	24.36	22.65296	23.71	22.70882	21.4	22.00881
3	26.47	23.8503	22.57	23.88891	21.1	23.17848
4	29.08	29.00157	30.65	29.00088	31.54	28.25818
5	5.84	4.666512	5.36	4.927954	4.63	4.355905
6	8.530001	3.929628	9.85	4.18975	10.87	3.619659
7	13.15	13.94253	16.82	14.09215	16.58	13.45205
8	22.55	25.64491	23.19	25.67654	25.15	24.95783

9	1.97	3.361543	1.03	3.637857	2.730001	3.075764
10	16.1	14.06831	16.62	14.20922	16.28	13.56541
11	25.01	21.70757	21.13	21.7716	20.4	21.07745
12	-2.24	3.305834	-3.98	3.588619	-2.99	3.029306
13	10.44	8.703175	10.55	8.899714	11.58	8.291885
14	3.33	1.622376	4.13	1.921166	0.1	1.37329
15	-1.27	3.544366	-3.11	3.827243	-5.17	3.267333
16	4.18	3.583105	3.070001	3.85523	3.63	3.291023
17	12.1	14.33991	10.33	14.47754	6.06	13.83202
18	-7.4	-0.46543	-6.72	-0.07622	-9.33	-0.57522
19	4.82	4.438962	7.15	4.698238	9.18	4.126372

Source: Superneunet Software Prototype (July 09, 2023)

- *Correlation Test between 0_Actual and 0_ Prediction with Exponential- Exponential-Linear (78th) Modified Specific Activation Function*

Table 12 Correlation Test between 0_Actual and 0_ Prediction with Exponential- Exponential-Linear (78th) Modified Specific Activation Function

CORREL	0_actual	0_prediction
0_actual	1.000	0.910
0_prediction	0.910	1.000
p Values	0_actual	0_prediction
0_prediction	0.000	

Source: QI Macros (July 09, 2023)

- *Training with LSTM Approach with Cubic-Exponential (62nd) as Chicago City Temperature Specific Activation Function Prediction Results*

Table 13 Training with LSTM Approach with Cubic-Exponential (62nd) as Chicago City Temperature Specific Activation Function Prediction Results

	0_actual	0_prediction	1_actual	1_prediction	2_actual	2_prediction
0	19.59	19.15375	25.12	19.14779	23.62	16.26779
1	26.68	25.68838	22.86	25.53882	19.96	22.60839
2	3.56	3.419157	6.68	3.657234	8.2	0.878096
3	1.05	2.998349	-4.4	3.257578	-3.64	0.489279
4	11.79	11.6493	9.15	11.72696	7.94	8.898701
5	12.55	12.7682	7.67	12.76653	4.76	9.916241
6	0.9	5.342593	-2.82	5.49629	0.66	2.690383
7	27.79	25.19748	27.53	25.0667	27.17	22.14016
8	17.35	18.23878	10.56	18.14884	14.6	15.2591
9	23.88	25.68008	24.38	25.53144	28.56	22.59965
10	8.190001	10.02227	2.62	10.09123	5.84	7.251005
11	1.150001	2.803114	2.85	3.055381	1.020001	0.282472
12	7.13	5.077827	-0.08	5.249126	2.35	2.44775
13	14.55	17.92518	11.55	17.91732	11.34	15.07066
14	13.76	12.81074	10.61	12.77315	8.76	9.89728
15	10.95	5.261683	15.14	5.435661	17.58	2.64074
16	5.63	6.150989	9.51	6.277401	9.39	3.454874
17	11.55	14.42338	10.26	14.50778	14.03	11.66212
18	5.32	5.766686	7.08	5.919524	2.45	3.119351
19	14.94	15.34625	11.5	15.25403	7.01	12.36354
20	-0.24	1.95868	-0.13	2.239656	0.43	-0.52451
21	21.01	20.12297	22.4	20.15298	23.57	17.28126

Source: Superneunet Software Prototype (July 09, 2023)

- *Correlation Test between 0_Actual and 0_ Prediction with Cubic-Exponential (62nd) Modified Specific Activation Function*

Table 14 Correlation Test between 0_Actual and 0_ Prediction with Cubic-Exponential (62nd) Modified Specific Activation Function

CORREL	0_actual	0_prediction
0_actual	1.000	0.923
0_prediction	0.923	1.000
p Values	0_actual	0_prediction
0_prediction	0.000	

Source: QI Macros (July 09, 2023)

With *0.928, 0.910 and 0.923* correlation of Cubic-Exponential-Gaussian- Gaussian (65th), Exponential-Exponential-Linear (78th) and Cubic-Exponential (62nd) respectively which are chosen arbitrarily from the generated activation functions of Chicago City daily maximum temperature data by Superneunet software, Cubic-Exponential-Gaussian- Gaussian (65th) outperformed Exponential- Exponential-Linear (78th) and Cubic-Exponential (62nd) with modified specific activation functions.

IV. CONCLUSION AND RECOMMENDATIONS

The chaotic dynamics in technology and climate dramatically navigate extremely high demand for thousands, millions and to billions of nonlinear specific activation functions emanate from input data depending on the specific area of application, time change and rules of thumb in the deep neural network. This paper introduced advanced soft computing technique named Modified Neural Network Architecture in an attempt to eliminate black-box, tremendously increase non-linearity, to add some senses and skills in AI modeling, accuracy, precision, and transparency. In order to do so, a software tool named Superneunet has been developed and used to experiment the proposed technique by comparing the predicted results of the traditional activation functions with that of the modified. The author used POWER| Data Access Viewer historical daily maximum temperature data of Chicago City of the United States of America as from 31st March, 2021 to 31st March, 1981 (roughly 40 years) as an input data of the deep neural network.

With the help of QI Macros, Normality Test confirmed Chicago maximum temperature data has being Non-normal, volatile and unpredictable. To assess the performances between the trial and error traditional activation functions with that of modified, Actual and Initial Predicted values of Chicago City maximum temperature were correlated, Tanh (0.947) outperformed ReLu (0.945) and Sigmoid (0.940) with the traditional trial and error activation functions. While, with the modified specific activation functions arbitrarily chosen, Cubic-Exponential-Gaussian- Gaussian (0.928) outperformed Exponential- Exponential-Linear (0.910) and Cubic-Exponential (0.923). Note that with minimum number of layers and nodes, the activation functions were kept constant throughout the hidden layers of Superneunet prototype. With automation, the modified functions can out rightly outperform the traditional.

The results of the study suggested that the technique can be used to specifically, accurately and precisely predict any time series application and prices of over 35 million instruments across all asset classes traded 24/7 in 160 countries aggregated from 330 exchanges of Bloomberg Terminal and other world's largest financial and weather assets terminals as against the current practice that uniformly uses traditional trial and error few activation functions across all areas of application.

More so, at the times of extreme climate change events such as floods, lightning, hail, tornadoes, hurricanes, ice storms, the proposed study can fascinatedly help World Climate Change Organizations such Climate Action Network (CAN), Climate Cardinals, Climate Collaborative, and Climate Group among others to solve climate crisis. Furthermore, it can also help people to take extra gear to prepare for the unforeseen future climate crisis, save lives and minimize property damage, helping farmers to track and protect their crops, future knowledge about environments (long-term trends and shifts), transportation hazards and Changes Coming. Moreover, as the prediction of complex stock market data requires nonlinear techniques, the proposed work can help to offer the best assets and derivatives prices possible while still turning a profit requires even at the times of financial and economic crises.

The proposed study findings can potentially benefit humanity in the following ways:

- It tremendously adds non-linearity to deep learning artificial neural network prediction;
- One of the fundamental benefits of this novel breakthrough, it can waive years of research by researchers around the globe use for constructing single desirable activation function, with it can be done by just clicking of a button. This completely eliminate phobia and psychological trauma experiences when selecting the right activation functions, thus can fantastically save cost, time, energy and to accelerate AI technology production and global economic growth;
- Create wealth, prevent extreme climate change events and to sharpen AI and Quantum technology;
- It opened up a new fascinating area in developing skillful, reliable and sophisticated time series AI modeling, forecasting and prediction which has greatly potentials to change the world;
- It has opened new area of research and development in deep learning artificial neural network;

ISSN No:-2456-2165

- It has introduced new approach to generate thousands, millions and billions of activation functions into deep learning artificial neural network;
- Introduces skillful and lovely approach and journey into deep learning artificial neural network;
- It has opened one of the deep learning artificial neural networks' 'Black boxes';
- It can help NAS experts to frequently obtain optimum architectures; and,
- It can help to discover more quantum computing applications;
- It opened up a new research area of study in non-linear activation functions of deep learning neural networks;
- It introduced a new teaching approach for the teaching staff of the Universities, Colleges of Education, Polytechnics, High Schools for the skillful and automatic selection of neural network activation functions;
- Researchers and Students can play around hundreds, thousands, millions or billions of neural network activation functions against the few traditional Sigmoid, ReLu among others. This can tremendously create job opportunities, rapid scientific and technological development and discoveries, innovations and creativities;
- The current study first published in Jamilu Adamu (2019) originated from finance conducted to increase non-linearity in neural network to accurately predict stock prices. The study introduced a new skillful approach to predict assets and derivatives prices as well as other financial series time series application;
- As scientifically claimed by Biologists, human brain consists of more than 100 billion neurons. The current E-brain, Micro Chip, and Autonomous Vehicle neural network architectures uses traditional activation functions, this study can introduce a new approach to Neuroscience computing and Autonomous Transportation, further can address pressing issues in the global Health and Transportation sectors;
- It can accelerate innovations and discoveries on Mathematical Modelling and Dynamical Systems to mention few;
- To crown it up, the humanity can benefit from the results of this study through Environment, Finance, Education, Health & Pharmaceuticals, Transportation & Security, Science & Technology, Innovation & Creativity.
- Strongly believing optimization of the modified neural network architectures obtained from 86 activation functions (models) generated in this paper can outperform the traditional trial and error architectures obtained from ReLu, Sigmoid and Tanh. The future research of the proposed study will focus on the automation of few hundred or thousands of activation functions emanate from input data that depend on time change, specific area of application and particular rules of thumb.

REFERENCES

[1]. Arute, F., Arya, K., Babbush, R. et al. Quantum supremacy using programmable superconducting processor, Nature 574, 505-510, https://doi.org/10.1038/s41586-019-1666-5 (2019).

[2]. Tamlin Magee, Quantum computing predictions for 2020: The quantum computing developments to watch out for in the coming year, Resellernews from IDG retrieved from www.reseller.co.nz/article/679935/microsoft-reveals-c-9-0-plans/ (2020).

[3]. Daniel Smith, What is AI Training Data? Lionbridge article available on https://lionbridge.ai/articles/what-is-ai-training-data/(2019).

[4]. Ambika Choudhury , 10 Open-Source Datasets to Learn Robotics available on https://analyticsindiamag.com/10-open-source-datasets-to-learn-robotics/ (2019).

[5]. Tarang Shah, About Train, Validation and Test Sets in Machine Learning available on https://towardsdatascience.com/train-validation-and-test-sets-72cb40cba9e7 (2017).

[6]. Jamilu Auwalu Adamu, Superintelligent Deep Learning Artificial Neural Networks, accepted in the International Journal of Applied Science, IDEAS SPREAD. INC (2019).

[7]. Jamilu Auwalu Adamu, Advanced Deterministic Optimization Algorithm for Deep Learning Artificial Neural Networks, accepted in the International Journal of Applied Science, IDEAS SPREAD. INC (2019).

[8]. Jamilu Auwalu Adamu, Advanced Stochastic Optimization Algorithm for Deep Learning Artificial Neural Networks in Banking and Finance Industries,Risk and Financial Management Journal, Vol 1 No1 (2019) available online: https://j.ideasspread.org/index.php/rfm/article/view/387 (2019).

[9]. TOPNEUNET Software (2023)

[10]. SUPERNEUNET Software Prototype (2023)

[11]. QI MACROS (2023)