

# Advanced Stochastic Optimization Algorithm for Deep Learning Artificial Neural Networks in Banking and Finance Industries

Jamilu Auwalu Adamu<sup>1</sup>

<sup>1</sup> Mathematics Programme, 118 National Mathematical Centre, FCT-Abuja, Nigeria

Correspondence: Jamilu Auwalu Adamu, Mathematics Programme Building, 118 National Mathematical Centre, Small Sheda, Kwali, FCT-Abuja, Nigeria. Tel: 234-803-867-9094. E-mail: whitehorseconsult@yahoo.com

Received: October 22, 2019 Accepted: November 20, 2019 Online Published: November 26, 2019

## Abstract

One of the objectives of this paper is to incorporate fat-tail effects into, for instance, Sigmoid in order to introduce Transparency and Stability into the existing stochastic Activation Functions. Secondly, according to the available literature reviewed, the existing set of Activation Functions were introduced into the Deep learning Artificial Neural Network through the “Window” not properly through the “Legitimate Door” since they are “Trial and Error” and “Arbitrary Assumptions”, thus, the Author proposed a “Scientific Facts”, “Definite Rules: Jameel’s Stochastic ANNAF Criterion”, and a “Lemma” to substitute not necessarily replace the existing set of stochastic Activation Functions, for instance, the Sigmoid among others. This research is expected to open the “Black-Box” of Deep Learning Artificial Neural networks. The author proposed a new set of advanced optimized fat-tailed Stochastic Activation Functions EMANATED from the AI-ML-Purified Stocks Data namely; the Log – Logistic (3P) Probability Distribution (1st), Cauchy Probability Distribution (2nd), Pearson 5 (3P) Probability Distribution (3rd), Burr (4P) Probability Distribution (4th), Fatigue Life (3P) Probability Distribution (5th), Inv. Gaussian (3P) Probability Distribution (6th), Dagum (4P) Probability Distribution (7th), and Lognormal (3P) Probability Distribution (8th) for the successful conduct of both Forward and Backward Propagations of Deep Learning Artificial Neural Network. However, this paper did not check the Monotone Differentiability of the proposed distributions. Appendix A, B, and C presented and tested the performances of the stressed Sigmoid and the Optimized Activation Functions using Stocks Data (2014-1991) of Microsoft Corporation (MSFT), Exxon Mobil (XOM), Chevron Corporation (CVX), Honda Motor Corporation (HMC), General Electric (GE), and U.S. Fundamental Macroeconomic Parameters, the results were found fascinating. Thus, guarantee, the first three distributions are excellent Activation Functions to successfully conduct any Stock Deep Learning Artificial Neural Network. Distributions Number 4 to 8 are also good Advanced Optimized Activation Functions. Generally, this research revealed that the Advanced Optimized Activation Functions satisfied Jameel’s ANNAF Stochastic Criterion depends on the Referenced Purified AI Data Set, Time Change and Area of Application which is against the existing “Trial and Error” and “Arbitrary Assumptions” of Sigmoid, Tanh, Softmax, ReLu, and Leaky ReLu.

**Keywords:** activation functions, probability distribution, Fat-tail, Jameel’s ANNAF Stochastic Criterion, stocks, Referenced AI-Data Set

## 1. Introduction

Casper Hansen (2019) says “*Better optimized neural network; choose the right activation function, and your neural network can perform vastly better*”.

Artist Hans Hoffman wrote, “*The ability to simplify means to eliminate the unnecessary so that the necessary may speak.*”

On Tuesday, June 25, 2019, the U.S. Subcommittee on Communications, Technology, Innovation, and the Internet headed by Sen. John Thune, R-S.D. convened a hearing entitled, “Optimizing for Engagement: Understanding the Use of Persuasive Technology on Internet Platforms”, to find appropriate policy on algorithms’ transparency or explanation as regards to its decision-making and machine learning on internet platforms might be influencing the public. Some of the issues discussed includes; the AI Algorithms’ transparency, regulations, supervision, secrecy, the Black Box, difficulty in terms of explainability, human bias, and final ranking, the algorithms bad recommendations, AI was accused will be used to determine the next US President and exposes children to unsolicited sexual videos contents.

Thanks to the idea of Artificial Intelligence and Machine Learning, it makes approval of Loans and opening accounts automated using MyBucks, OnDeck, Kabbage, Lend up, Knab and Knab Finance, however, this has significantly reduced customers' time wait for the processing of loan request to just a few seconds. Presently, Artificial Intelligence has become an integral part of the Banking and Finance applications. The path to model deployment in banking and finance is traditionally cumbersome. Deploying a model is hard and takes time, resources and coordination across many teams throughout the bank.

In Bank and Finance Industries, one of giant and bravery first steps was taken in 2015 by Ally Bank (USA) thereby introduced Ally Assist: "a chatbot that could respond to voice and text, make payments on behalf of the customer, give an account summary, monitor savings, spending patterns, and use natural language processing to understand and address customer queries". Banks around the globe launched best versions of chatbots: Erica to iPAL, Eva and the most famous one, SBI's SIA. The banking and finance sectors have witnessed momentum development and continue to increase their spending on artificial intelligence and ML. The years 2016 and 2017 marked a significant milestone in the development of AI due to AlphaGo remarkable games successes. PwC (2017) stated that global spending in artificial intelligence in Banking has reached \$5.1 billion. The IHS Markit's stated that the report claims it will grow up to \$41.1 billion in 2018 and expected to reach \$300 billion by 2030.

The basic challenge is how to explore an Optimized ways where artificial intelligence and machine learning algorithms can be harnessed to be able to capture fat-tailed effects for accurate future prediction so as to improve Credit Decisions, Risk Assessment and Management, Fraud Prevention, Customer Experiences and apply to the other fields.

The key question you will be hearing in the Banking and Finance Industries is "What is Next? What comes Next? Where is ROI". Artificial intelligence and machine learning are said to revolutionize the financial world, changing the banking and finance experience for the better. However, there is an inherent risk of relying on AI and ML models. With a company's reputation, regulatory requirements, and finances at stake, it's critical to understand how to continuously monitor and validate the AI and ML models to ensure proper performance and sound business decisions. However, due to an increased reliance on models for everyday business processes and decisions, model risk must be effectively managed. If left unchecked, the consequences of model risk can be severe; where model risk is defined as the risk of financial or reputation loss due to errors in the development, implementation or use of models. Therefore, AI and machine learning models require constant monitoring and effective validation. This is not only a regulatory requirement, but it is also sound business practice.

Recent technological advancements have accelerated the integration of AI and machine learning models into more and more banking processes. In today's banking industry, institutions not using AI and machine learning risk losing their competitive edge, as competitors are increasingly enhancing their strategic decisions with the powerful analytical capabilities of AI and machine learning.

The contemporary Extremistan World (the statistical properties rapidly change over time), non-Gaussian (often skewed and exhibiting fat tails, making extreme events far more likely than they normally would be) where some things are inherently unpredictable even with Artificial Intelligence (Nassim N. Taleb (2013)). Couple with an inappropriate choice of activation functions.

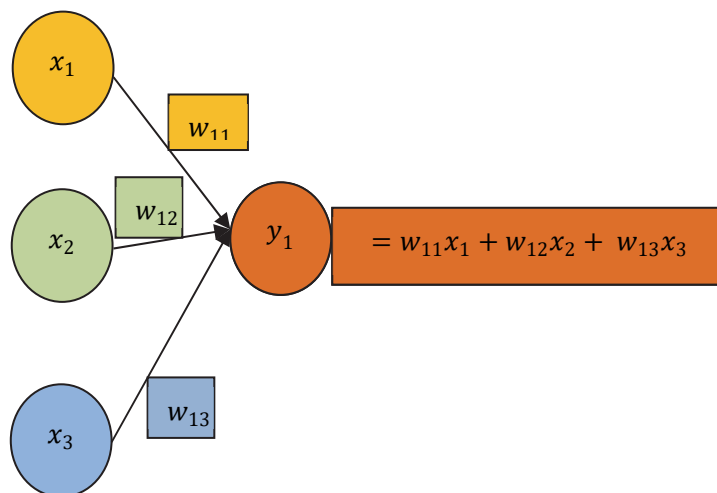
A model may be reasonable, but the world itself may be unstable. "A good model today may be inappropriate tomorrow" (Emanuel Derman, 1996). 'Change' is in the timeframe of the beholder. Ben Steiner (2019) stated that the challenges of Deep Learning are basically, the Non-Stationarity challenge which is one problem with many names, namely; Concept Drift, Covariate Shift or DATASET SHIFT. He defined Dataset Drift as when the statistical properties of the target variable, which the model is trying to predict, change over time in unforeseen ways. The unforeseen substitution of one data source  $S_1$  (with underlying probability distribution  $\Pi S_1$ ), with another source  $S_2$  (with distribution  $\Pi S_2$ ) imposes the Risk model. The second challenge was that we are learning what we already knew, Learning, but nothing new is incorporated into the learning process to capture the effects of unforeseen challenges, basically, the LOW-PROBABILITY, HIGH-IMPACT (fat-tailed effects).

According to the available literature review, the existing set of Activation Functions were introduced into the Deep learning Artificial Neural Network through the "WINDOW" not properly through the "LEGITIMATE DOOR" since they are "TRIAL and ERROR" and "ARBITRARY ASSUMPTIONS". The objectives of this paper are to incorporate fat-tail effects into the existing stochastic Activation Functions, for instance, Sigmoid and proposed a Stochastic "DEFINITE RULES" and "LEMMA" EMANATED from the AI-ML-Data Set to substitute not necessarily replace the existing set of "Trial and Error" stochastic Activation Functions. This research is expected to open the "Black-Box" of Deep Learning Artificial Neural networks.

The paper started with the Introduction, Literature Review, Material and Methods, Results; the paper crown up with Conclusion remarks.

**2. Literature Review**

Activation functions add a non-linearity to neural networks, thus gives it ability to learn complex functional mappings from data. A neural network without an activation function: (1) Can only represent a linear relationship between variables and, (2) Does not hold Universal approximation theorem. This can be seen below:



That means

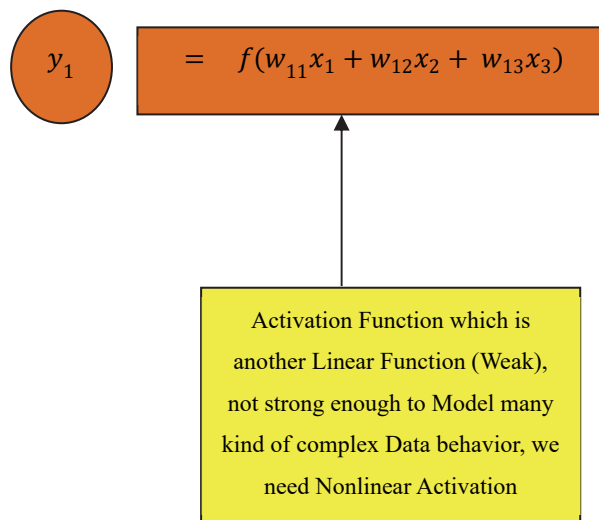


Figure 1. Linearity of Activation Functions

Source: Author (2019)

The choice of a Neural Network Activation Functions does not follow DEFINITE RULES; however, this makes the Deep Learning Artificial Neural Network a “Black Box” System and exposes its predictions to a “Model Risk”.

Sebastian Urban (2017) stated that the choice of activation functions was not seriously challenged by researchers (except for special purpose applications), until recently when Nair et al. (2010) introduced the rectified linear unit (ReLU), a neuron with an activation function that is linear for positive inputs and zero for negative inputs.

This achievement led to a wave of follow-up research in activation functions specifically tailored to deep networks. The year 2015, introduced stochasticity into the activation function by sampling the value for the slope with each training iteration from a fixed uniform distribution. Clevert et al. (2015) and Klambauer et al. (2017) replaced the negative part of ReLUs with a scaled exponential function and showed that, under certain conditions, this leads to automatic renormalization of the inputs to the following layer and thereby simplifies the training of the neural networks, leading to accuracy improvements of deep feed-forward networks on tasks from the UCI Machine

Learning repository (Lichman, 2013) amongst others. Nearly fully adaptable activation functions have been proposed by Eisenach et al. (2017). Similarly, Scardapane et al. (2017) also use a basis expansion, but with a set of Gaussian kernels that are equally distributed over a preset input range.

According to Asman Dureja and Payal Pahwa (2019), the choice of activation functions affects the neural network in term of OPTIMIZATION and to retrieve the better results. Several activation functions have been introduced in machine learning for many practical applications. But which activation function should use at hidden layer of deep neural networks was not identified. According to Chigozie Enyinna Nwankpa et al. (2018), LReLU, PReLU and RReLU perform better than the ReLU but some of these functions lack theoretical justifications to support their state-of-the-art results. Soufiane Hayou et al. (2019) stated that an inappropriate selection of Activation Function can lead to the loss of information of the input during forward propagation and the exponential vanishing/exploding of gradients during back-propagation. Schoenholz et al. (2017) applied Gaussian process approximation to Deep Neural Networks.

Joonho Lee et al. (2019), in their work, proposed a probabilistic activation function, called ProbAct. Unfortunately, the output value of ProbAct is sampled from a normal distribution; with the mean value same as the output of ReLU and with a fixed or trainable variance for each element.

Professionally, the opinions of the Members of StackExchange (2018) sponsored by IBM stated that *“I choose the activation function for the output layer depending on the output that I need and the properties of the activation function that I know. In hidden layers, I use a leaky ReLU to avoid dead neurons instead of the ReLU and the tanh instead of the sigmoid. Of course, I don't use a linear function in hidden units. However, the choice for them in the hidden layer is mostly due to trial and error. However, the more activation functions I discover the more I'm confused about the choice of the function to use in hidden layers. And I don't think that flipping a coin is a good way of choosing an activation function. It would be incredibly difficult to recommend an activation function that works for all use cases (although I believe that SELU was designed so that it would do the right thing with pretty much any input). At the end of the day, you are probably going to get as many opinions as there are people about the right choice of activation function. The bottom line is that there is no universal rule for choosing an activation function for hidden layers”*.

### **3. Materials and Methods**

#### *3.1 Materials*

##### 3.1.1 Basic Definitions

###### 3.1.1.1 Artificial intelligence (AI)

Artificial intelligence (AI) was created as a subfield of computer science targeted to solve tasks that humans are good at or the goal of AI is to mimic human intelligence. Artificial intelligence and machine learning technology in banking industry can be used for Credit decisions (It can allow for a fast, accurate assessment of a potential borrower, for far less cost than traditional methods, it eliminates bias, as machines have more objectivity than human employees), Risk assessment and management (banks mitigate risk because they receive accurate reporting, not prone to human error, AI does even more to reduce risks for banks and customers). By viewing the history of risk cases, AI can help banks forecast issues and take early steps to avoid problems), Fraud prevention (By analyzing spending patterns, location, and client behavior, machine learning can detect anomalies in spending and alert the cardholder, dramatically reducing credit card fraud) and Personalized approach (notifies members if their card was charged twice for an expense or if they tipped an exorbitant amount at a restaurant).

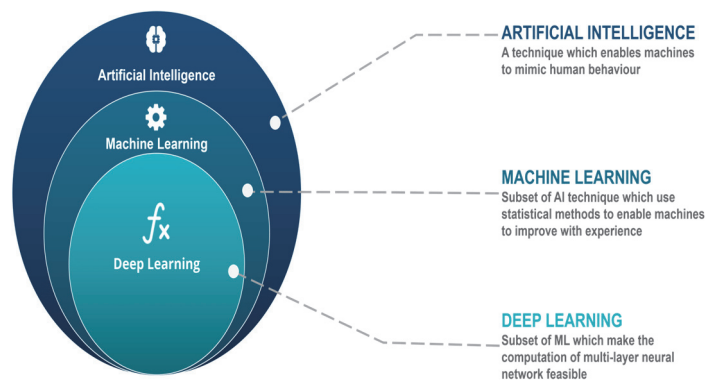


Figure 2. Artificial Intelligence

Source: Google Images (2019)

### 3.1.1.2 Machine Learning

Machine Learning uses algorithms to learn from data without relying on rules-based programming. Machine Learning originally developed as a subfield of Artificial Intelligence (AI), one of the goals behind machine learning was to replace the need for developing computer programs “manually.” Considering that programs are being developed to automate processes, we can think of machine learning as the process of “automating automation.” In other words, machine learning lets computers “create” programs (often, the intent for developing these programs is making predictions) themselves. In other words, machine learning is the process of turning data into programs. Machine learning (and deep learning) definitely helps to develop “AI,” however, AI doesn’t necessarily have to be developed using machine learning – although, machine learning makes “AI” much more convenient.

Machine learning models can also predict which banking tools individual members might use and recommend them so customers can make better financial decisions. Process automation (Using robotic process automation (RPA), banks can remove human error and restructure the workforce to focus on more pressing tasks. JPMorgan Chase & Co launched COIN or Contract Intelligence, which automated the processing of legal documents, extraction of data, and review of certain types of legal contracts. Machine learning algorithms could use image recognition to identify patterns in the agreements. What normally would take roughly 360,000 labor hours per year, took the model a few hours. Another example of automating tasks is the increased use of chatbots that provide quick and reliable answers to consumers. Using AI-powered mobile and web chatbots, banks can speed up the time it takes for consumers to receive answers and decrease the need for human assistants to answer questions. Other applications are improved customer services, Investment Predictions Marketing, Network Security and Algorithm Trading.

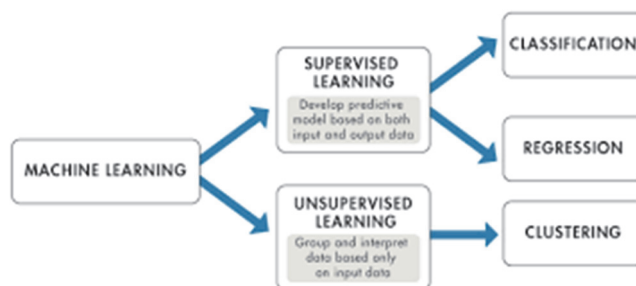


Figure 3. Machine Learning

Source: Google Images (2019)

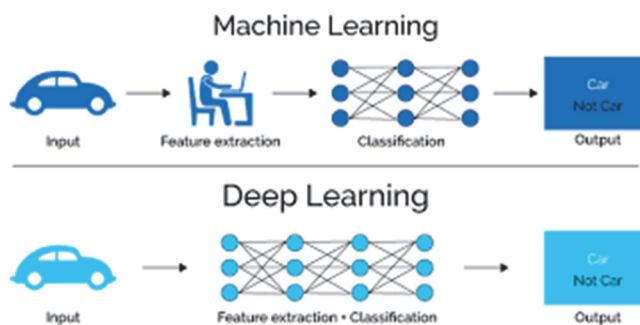


Figure 4. Difference between Machine Learning and Deep Learning

Source: Google Images (2019)

### 3.1.1.3 Deep Learning

Deep learning is a subfield of machine learning, referring to a particular subset of models that are particularly good at certain tasks. Deep Learning algorithms run data through several “layers” of neural network algorithms, each of which passes a simplified representation of the data to the next layer. The ability to process large numbers of features makes Deep Learning algorithms very powerful when dealing with unstructured data (such as images, audio and video). However, Deep Learning algorithms can be overkill for less complex problems because they require access to a vast amount of data to be effective and may be outperformed by Tree-Based Algorithms in these cases. Instead of organizing data to run through predefined equations, Deep Learning algorithms characterize basic parameters about the data and train the computer to learn on its own by recognizing patterns using many layers of processing. For instance, Deep Learning algorithms include: Recurrent Neural Networks (RNN), Deep Boltzmann Machines (DBM), Deep Belief Networks (DBN), Convolutional Neural Networks (CNN), and Stacked Auto-Encoders.

Deep Learning maps inputs to outputs using multiple layers of nonlinear processing units. It is capable of creating models for previously manual procedures. The first is creating risk management models for lending and credit risk management; the second is in fraud prevention, where AI systems identify, track and flag potential threats.

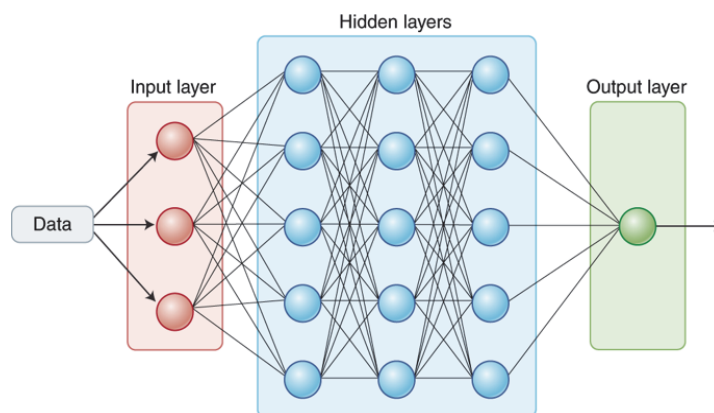


Figure 5. Deep Learning

Source: Google Images (2019)

### 3.1.1.4 Neuron

A neuron is just a **MATHEMATICAL FUNCTION** that takes inputs (the outputs of the neurons pointing to it) and **returns outputs**. These outputs serve as inputs for the next layer, and so on until we get to the final, output layer, which is the actual value we return. For each set of inputs, the Neural Network’s goal is to make each of its outputs **as close as possible to the actual expected values**. A neuron will take an input vector, and basically do three things to it:

- Multiply it by a weights vector.
- Add a bias value to that product.

- Apply an **activation function** to that value.

We get a **new vector as each layer's output**; feed it to the next layer as inputs, and so on. Bias is a constant which helps the model in a way that it can fit best for the given data. **Each layer in a Neural Network has an activation function.**

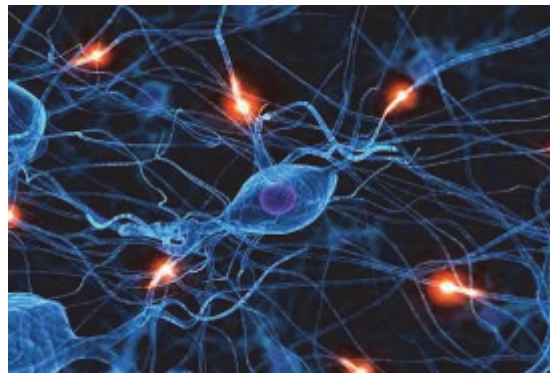
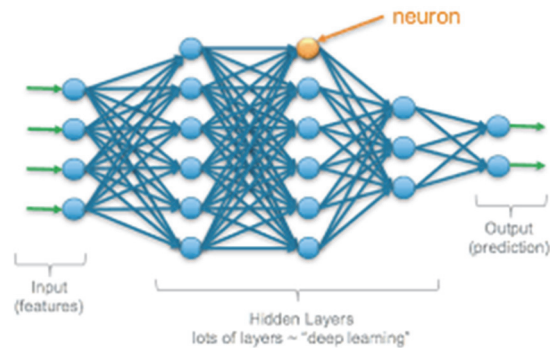


Figure 6. Neuron

Source: Google Images (2019)

### 3.1.1.5 Neural Network

A Neural Network is a Machine Learning model that, given certain input and output vectors, will try to “fit” the outputs to the inputs. What this means is, given a set of observed instances with certain values we wish to predict, and some data we have on each instance, it will try to generalize those data so that it can predict the values correctly for new instances of the problem. And we finally got to the core of our business: that’s what activation functions do. We’ll typically use **non-linear functions as activation functions**. This is because the linear part is already handled by the previously applied product and addition.



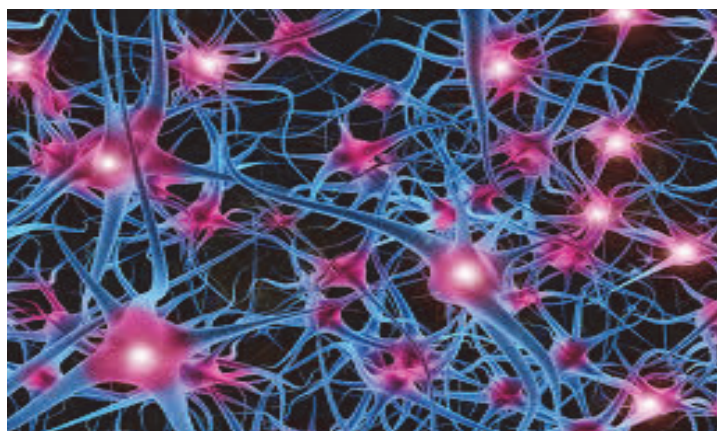
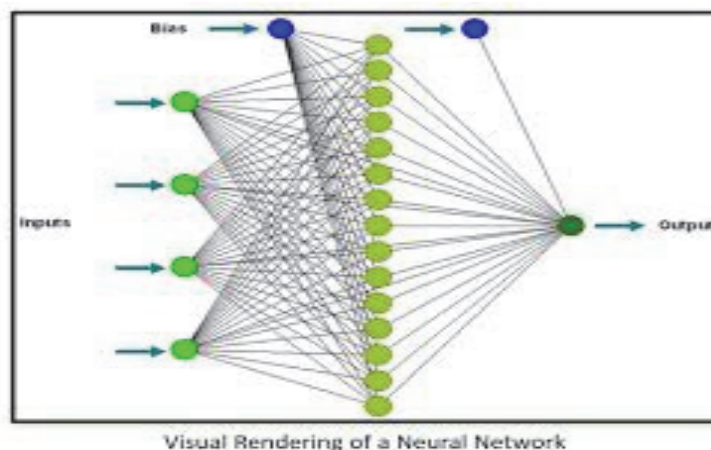


Figure 7. Neural Network

Source: Google Images (2019)

### 3.2 Methods

Model complexity has increased; models are being used in many areas of banking and finance applications. This scrutiny was kick-started by the aftermath of the global financial crisis, in which it became apparent that at least some incumbent models at the time were unfit for purpose, which was exacerbated by a lack of appropriate governance around models and their use.

The supervisory reaction has drastically increased regulatory scrutiny for how models are built, approved and maintained. In 2011, the Federal Reserve and the Office of the Comptroller of the Currency published the supervisory guidance on Model Risk Management which mandated banks to submit documentation on model risk management policies and practices. This is a regulatory standard for model risk management and the regulators around the globe picked up the challenge. In 2017, the European Central Bank (ECB) published guidelines which expect financial institutions to have a well-developed model risk management framework in place, and to apply it to all models used for business decision-making; especially those pertaining to capital requirements. In the United Kingdom (UK), the Prudential Regulation Authority (PRA) is also heavily focused on model risk; with stress testing models getting a lot of attention. In 2018, the PRA published a supervisory statement on model risk management for stress testing (SS3/18), covering firms' development, implementation and use of stress testing models. Algorithmic trading has also seen increased regulatory focus in the UK. The PRA published a supervisory statement on Algorithmic Trading (SS5/18) outlining expectations on the risk management and governance of Algorithmic trading, where machine learning models are increasingly being used to facilitate automated trading decisions. These developments are mirrored in many other jurisdictions around the world.



Most banks today are quite familiar with machine learning algorithms that learn by example from historical data and are used to predict outcomes and uncover patterns that are not easily spotted by humans. However, not all classes of machine learning algorithms are recent developments; many have been used in the banking and financial services industry for decades. These include basic statistical methods and techniques like linear and logistic regression and rule-based modeling. These traditional methods are relatively simple, well-understood by practitioners and easily explainable and interpretable. However, recently, the U.S. Congress has intensively criticized AI and ML as the Technological or Scientific advancement that lack transparency subject to Data Manipulations with unregulated outcomes that does not gives Companies possibilities to make well informed, AI and ML Algorithms are running in secrecy, the Black Box and beyond the control of the User and the Builder, Technology without having Human Independent supervisor, oversight or regulation, that the Technology cannot trusted anymore, AI & ML algorithms gives bad recommendations, care should be taken about Final Ranking of the content among others. To increase TRANSPARENCY and STABILITY in the Activation Functions output (final ranking of the content) so as to help policy makers, Companies and decision makers with optimum the right choice among the final activation functions to enable them make well informed decision, the paper propose the following Advanced Stochastic Algorithms. Also, this propose algorithm will logically make an attempt to answer the questions: "*why is that one works better than the other?*", "*how do we know which one to use?*", "*is it hardcore maths?*" and so on.

### 3.2.1 The Basic Idea

The basic IDEA was due to U.S. Congress meeting entitled, "Optimizing for Engagement: Understanding the Use of Persuasive Technology on Internet Platforms" held on Tuesday, June 25, 2019 by the Subcommittee on Communications, Technology, Innovation, and the Internet chaired by Sen. John Thune, R-S.D, with aim to find appropriate policy on algorithms' transparency or explanation as regards to its decision-making and machine learning on internet platforms might be influencing the public. The Author attempted to INCORPORATE TRANSPARENCY and STABILITY in the existing SET of Activation Functions thereby opening the Deep Learning Artificial Neural Network's "BLACK-BOX" in the Areas of Banking and Finance.

The Author observed that the NON-LINEAR Activation Functions were introduced into Deep Learning Artificial Neural Networks to introduce NON-LINEARITY into the system. Thus, this implies Activation Functions are the BRAIN behind the successful functioning of any Deep Learning Artificial Neural Network.

To INCORPORATE STABILITY in the existing the Deep Learning Artificial Neural Network System REFERENCE to out AI-DATA SET, the Author considered the set of existing Activation Functions thereby picking SIGMOID. Sigmoid is a fat-tailed Probability Distribution, however, the Author thinks on how to incorporate LOW-PROBABILITY, HIGH-IMPACT effects (additional fat-tailed effects) in it.

Jamilu (2015) proposed Jameel's Contractual-Expansional Stress Method that depends on JAMEEL'S CRITERION to incorporate fat-tail effects in BANKING and ECONOMIC Stochastic Models, particularly the Logistic Regression (LOGIT) and PROBIT Models as shown in the figure below:

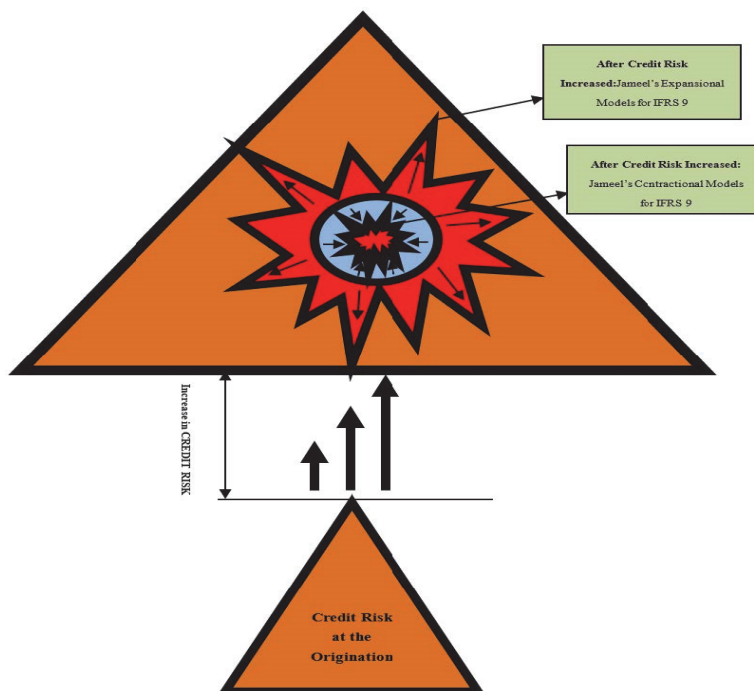


Figure 8. The Basic Idea

Source: The Author (2018)

The Author applied the above Idea to Sigmoid Activation Function:  
*Sigmoid/Logistic Function (famous function in the Neural Network):*

$$y = \frac{1}{1+e^{-x}} ; -\infty < x < +\infty$$

Thus, we have the following Stressed Sigmoid Types:

*Stressed Sigmoid Type A:*

$$\frac{1}{1 + \exp \mu_A(-x) \pm \sigma_A f(x)}$$

*Stressed Sigmoid Type B:*

$$\frac{1}{1 + \exp \mu_A(-x) \pm f(x)}$$

*Stressed Sigmoid Type C:*

$$\frac{1}{1 + \exp(-x) \pm \sigma_A f(x)}$$

*Stressed Sigmoid Type D:*

$$\frac{1}{1 + \exp(-x) \pm f(x)}$$

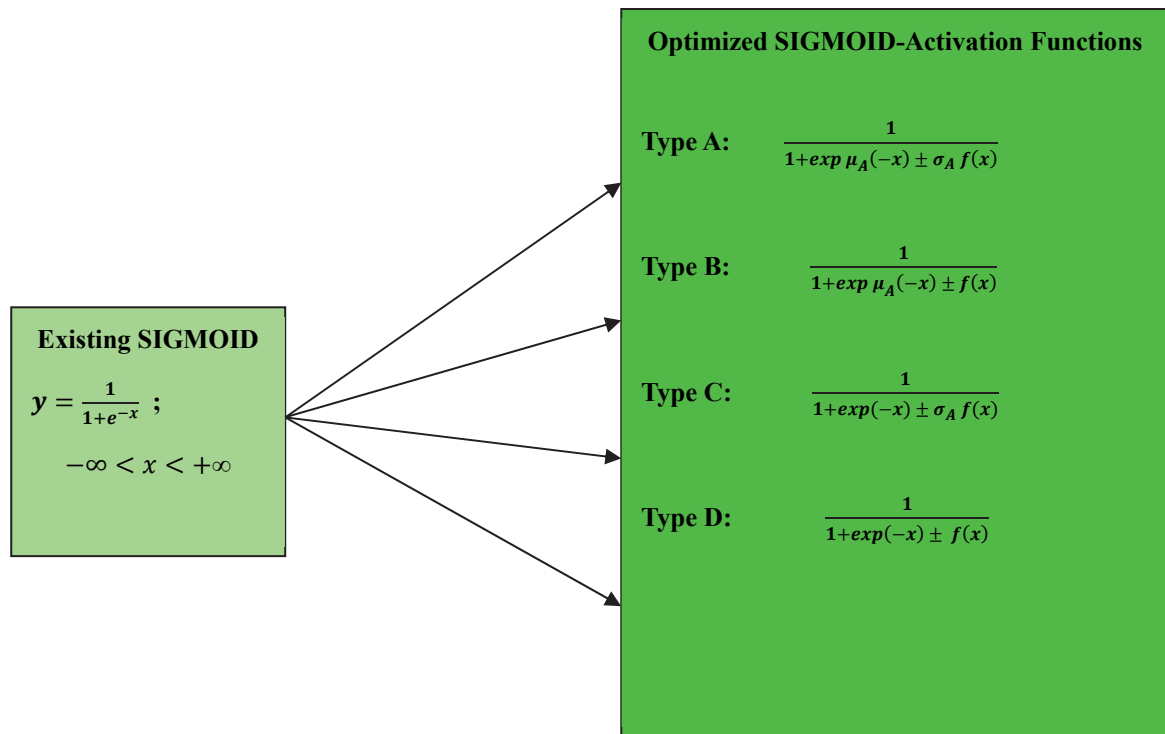


Figure 9. Stressed Sigmoid-Activation Functions

Source: The Author (2019)

Where, the Geometric Mean ( $\mu_A$ ) and Volatility ( $\sigma_A$ ) of the Macroeconomic Indicators and Task.  $f(x)$ , the Fat-tailed Stochastic Function satisfied Jameel’s Criterion.



Figure 10. Expected Stressed Sigmoid Output

Source: Google Images (2019)

**For instance, the Ultimate Goal is to drastically minimize the NOISES in the Right Photo of figure 10 above.**

After the Sigmoid was Stressed, the Author continued to brainstorm with available Literature concluded that the existing set of Activation Functions were introduced into the Deep Learning Artificial Neural Network System through the “WINDOW” not properly through the “LEGITIMATE DOOR”.

### 3.2.2 Existing Activation Functions and their Relationships with AI-ML-Purified Data Set

The existing Activation Functions has no correlation(s) in whatsoever with the AI-ML-Purified Data Set under consideration, in fact, their choices has no definite Rule of Thumb; it is just Trial and error as shown in the figure below:

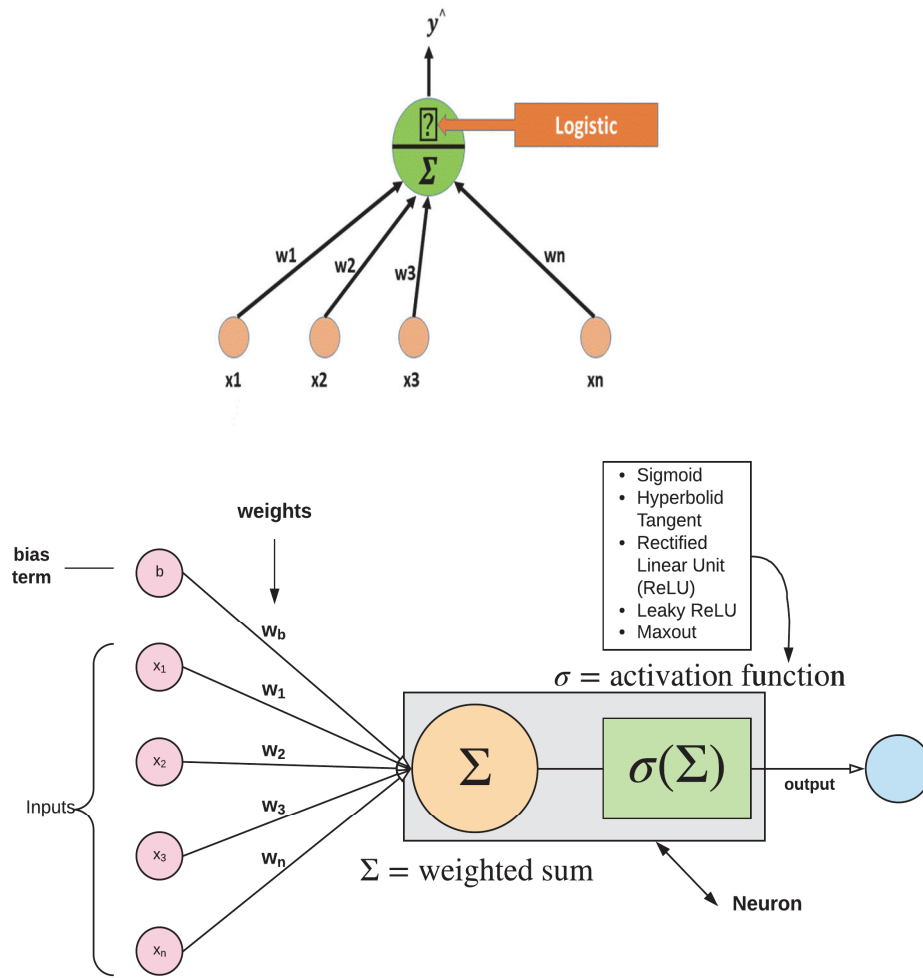


Figure 11. Activation Functions in ANN

Source: Google Images (2019)

### 3.2.3 Who are the Competent and Eligible Activation Functions for a successful Neural Network

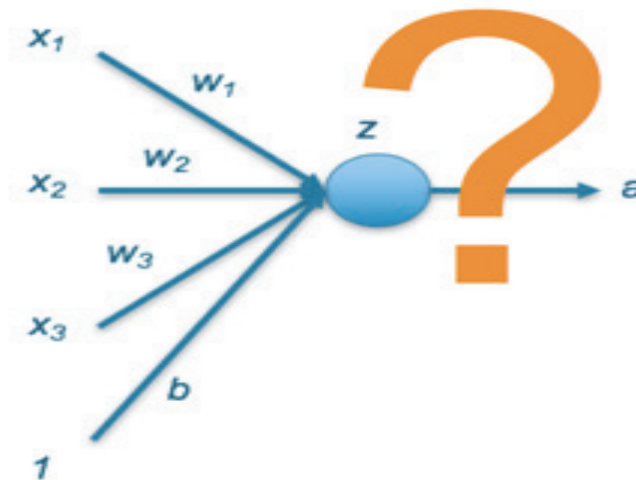


Figure 12. Question of Competent and Eligible Activation Functions

Source: Google Images (2019)

ACTIVATION FUNCTIONS Allows Artificial Neural Network to learn complex functional mappings from DATA and make sense of something complicated and Non-linear to produce meaningful output signal. Thus, should not be a Trial and error or a Black-Box Assumptions.

Biologically, Neuron performs three basic functions, namely: Receive signals (or information); Integrate incoming signals (to determine whether or not the information should be passed along); and, Communicate signals to target cells (other neurons or muscles or glands). However, this cannot be done successfully without action of a Non-linear Function (Brain of Neuron) residing in a Neuron of Human Brain EMANATED from incoming signals (information). This Non-linear Function (Brain of Neuron) residing in a Neuron of Human Brain EMANATED from incoming signals is what we called Activation Function of a Deep Learning Artificial Neural Network.

The Author PROPOSED that the Competent and Eligible Activation Functions for the successful conduct of Artificial Neural Networks are the Activation Functions EMANATED from the AI-ML-Purified Data Set under consideration satisfied AI-ML-Jameel's Stochastic or Deterministic Criterion because of the following SCIENTIFIC FACTS:

- (1) They EMANATED from the referenced AI-ML-Purified Data Set and satisfied AI-ML-Jameel's Stochastic and or Deterministic Criterion;
- (2) They have a very strong (if not perfect) CORRELATION with the referenced AI-ML-Purified Data Set. A link between the Data Set and Activation Functions MUST be strongly established since Artificial Neural network uses past historical data to predict the future of a given task with the aid of machines;
- (3) They relate better to the referenced AI-ML-Purified Data Set than the existing Assumed and Trial and error Activation Functions;
- (4) They indeed describe the distribution of our referenced AI-ML-Purified Data Set, which is a listing or function showing all the possible values (or intervals) of the data and how often they occur;
- (5) They represent real, virtual and un-virtual information about our referenced AI-ML-Purified Data Set;
- (6) They indeed captured the Symmetric, Left Skewed, Right Skewed, Mesokurtic, Leptokurtic, and Platykurtic properties of our referenced AI-ML-Purified Data Set;
- (7) They contain real, virtual and un-virtual information related to Measures of variability (the range, inter-quartile range, and standard deviation) and Measures of Central Tendency (Mean, Mode and Median, Minimum and Maximum) of our referenced AI-ML-Purified Data Set;

**(8)** They capture real, virtual and un-virtual information about the Correlation (autocorrelation) among the elements in our referenced AI-ML-Purified Data Set;

**(9)** In the case of Bivariate AI-ML-Data Set, they captured real, virtual and un-virtual information about Measures of Association (Covariance and Correlation) of our referenced AI-ML-Purified Data Set;

**(10)** They capture real, virtual and un-virtual information whether or not the parameters of our referenced AI-ML-Purified Data Set are constant over time; and,

**(11)** They also captured the presence of *outliers in our* referenced AI-ML-Purified Data Set under consideration.

Referenced AI-ML-Purified Data Set means our referenced Artificial Neural Network Data Set shall possess the following QUALITIES:

- a. Accuracy and Precision
- b. Legitimacy and Validity
- c. Reliability and Consistency
- d. Timeliness and Relevance
- e. Completeness and Comprehensiveness
- f. Availability and Accessibility
- g. Granularity and Uniqueness

Therefore, the author proposed that the practice of “TRIAL AND ERROR” choice of Assumed Activation Functions should be abandoned, however, the choice of Activation Functions should follow a “DEFINITE RULES”, and thus, the Author enhanced “JAMEEL’S CRITERION (2015)” to proposed “JAMEEL’S ANNAF STOCHASTIC CRITERION (2019)” as follows:

### 3.2.4 Proposed Jameel’s ANNAF Stochastic Criterion

ANNAF means Artificial Neural Network Activation Functions.

Under this criterion, we run the goodness of fits test on our referenced PURIFIED AL-ML-DATA SET such that:

- (i) We accept if the Average of the ranks of Kolmogorov Smirnor, Anderson Darling and Chi-squared is less than or equal to Three (3);
- (ii) We must choose the fat-tailed Probability Distribution  $f(x)$  follows by our referenced PURIFIED AL-ML-DATA SET ITSELF regardless of its Rankings;
- (iii) If there is tie, we include both the fat-tailed Probability Distributions in the selection;
- (iv) At least Two (2) fat-tailed Probability Distributions must be included in the selection;
- (v) We select the most occur Probability Distribution as the qualify candidate in each case of test of goodness of fit on our referenced PURIFIED AL-ML-DATA SET;
- (vi) **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 (i) to (vi) until we have a MONOTONE DIFFERENTIABLE FUNCTION. Note the Approach is from TOP-BOTTOM process;

- (vii) **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 ( $\chi^2$  Godness of Fit) Test, Lilliefors Corrected K-S Test, D’Agostino Skewness Test, Anscombe-Glynn Kurtosis Test, D’Agostino-Pearson and Omnibus Test. Let  $\{T_1, T_2, \dots, T_n\}$  be the set of such Best Goodness of Fit Tests,  $\{x_1, x_2, \dots, x_n\}$  be their **RANKS**



respectively then the generality of (i) can be re-expressed (or enhanced) if  $\frac{(x_1 + x_2 + \dots + x_n)}{n} \leq a$ , where

$0 < a \leq n, n \in N$  or equivalently,  $x_1 + x_2 + \dots + x_n \leq an$ ; and,

(viii) **Unit Axiom:** let  $f_{final}(x)$  be such that it satisfied axioms (i) to (iv) and or (vii). Let  $\{r_1, r_2, \dots, r_n\}$  be

the ranks of fitness test of  $f_{final}(x)$  obtained from the tests  $\{T_1, T_2, \dots, T_n\}$  respectively then if

$\forall i \in \{1, 2, \dots, n\}, r_i = 1$  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.**

3.2.5 Proposed Jameel’s Stochastic Lemma

All the TOP-RANKED Fat-tailed Monotone Continuously Differentiable Stochastic Functions **EMANATED** from referenced AI-ML-Purified Data satisfied Proposed Jameel’s Stochastic ANNAF Criterion are EXCELLENT STOCHASTIC ACTIVATION FUNCTIONS to perform well-informed Forward and Backward Propagations of an Artificial Neural Network.

3.2.6 Stocks Advanced Optimized Activation Functions

The Ranking of the Advanced Optimized Activation Functions was obtained using Jameel’s Criterion (second version). The author considered Eleven (11) out of Fifty (50) World’s Biggest Public Companies by FORBES as of 2015 Ranking Regardless of the platform in which they are listed, Time Series (Short or Long Terms), Old or Recently Listed Companies. The Ranking can be seen as follows:

**Log – Logistic (3P) Probability Distribution (1<sup>st</sup>):**

$$f(x; \mu, \sigma, \xi) = \frac{\left(1 + \frac{\xi(x - \mu)}{\sigma}\right)^{-\left(\frac{1}{\xi} + 1\right)}}{\left[1 + \left(1 + \frac{\xi(x - \mu)^{-\frac{1}{\xi}}}{\sigma}\right)\right]^2}; \quad x \geq \mu$$

Cauchy Probability Distribution (2<sup>nd</sup>):

$$f(x; \mu, \sigma, \pi) = \left(\pi \sigma \left(1 + \left(\frac{x - \mu}{\sigma}\right)^2\right)\right)^{-1}; \quad -\infty < x < +\infty$$

Pearson 5 (3P) Probability Distribution (3<sup>rd</sup>):

$$f(x; \alpha, \beta, \gamma) = \frac{\exp\left(-\frac{\beta}{(x - \gamma)}\right)}{\beta \Gamma(\alpha) \left(\frac{(x - \gamma)}{\beta}\right)^{\alpha + 1}}; \quad \alpha, \beta > 0$$

Burr (4P) Probability Distribution (4<sup>th</sup>):

$$f(x; \alpha, \beta, \gamma) = \frac{ak \left( \frac{x-\gamma}{\beta} \right)^{\alpha-1}}{\beta \left( 1 + \left( \frac{x-\gamma}{\beta} \right)^{\alpha} \right)^{k+1}}; \quad \alpha, \beta, k > 0$$

Fatigue Life (3P) Probability Distribution (5<sup>th</sup>):

$$f(x; \alpha, \beta, \gamma) = \frac{\sqrt{\frac{(x-\gamma)}{\beta}} + \sqrt{\frac{\beta}{(x-\gamma)}}}{2\alpha(x-\gamma)} \Phi \left( \frac{1}{\alpha} \left( \frac{\sqrt{x-\gamma}}{\beta} - \sqrt{\frac{\beta}{x-\gamma}} \right) \right); \quad \alpha, \beta, \gamma > 0$$

Inv.Gaussian (3P) Probability Distribution (6<sup>th</sup>):

$$f(x; \mu, \lambda, \gamma) = \sqrt{\frac{\lambda}{2\pi(x-\gamma)^3}} \exp \left( -\frac{\lambda(x-\gamma-\mu)^2}{2\mu^2(x-\gamma)} \right); \quad \lambda, \mu, \gamma > 0$$

Dagum (4P) Probability Distribution (7<sup>th</sup>):

$$f(x; \gamma, \beta, \alpha) = \frac{ak \left( \frac{x-\gamma}{\beta} \right)^{ak-1}}{\beta \left( 1 + \left( \frac{x-\gamma}{\beta} \right)^{\alpha} \right)^{k+1}}; \quad k, \alpha, \beta > 0$$

Lognormal (3P) Probability Distribution (8<sup>th</sup>):

$$f(x; \mu, \sigma, \gamma) = \frac{\exp \left( -\frac{1}{2} \left( \frac{\ln(x-\gamma) - \mu}{\sigma} \right)^2 \right)}{(x-\gamma)\sigma\sqrt{2\pi}}; \quad \gamma < x < +\infty$$

Note that the paper did not check the Monotone Differentiability of the functions presented above. However, the market trends are very volatile, many things had happened from 2019 – 2015. The distributions time series Data was from 2014 – 1990, Jamilu (2015) showed this. Different results may be obtained when conducted in the year 2019. Anyway, the paper adopted the fat-tailed probability distributions obtained as per as 2015.

## 5. Conclusion

The U.S. Congress, professionals, and the general public strongly believed that AI and ML have shortcomings despite the overwhelming advancement in Banking and Finance Industries, electronic trading platforms, medical diagnosis, robot control, entertainment, education, health, and commerce.

To address the FEARS raised by the U.S. Congress, the paper attempted to increase TRANSPARENCY and STABILITY in terms of Activation Functions output to help the policymakers, companies, and decision-makers with right Activation Functions using Definite Rules to enable them to make a well-informed decision.

To open the “Black Box” of Deep Learning Artificial Neural Network Systems, we need MONOTONE DIFFERENTIABLE FAT-TAILED PROBABILITY DISTRIBUTIONS **EMANATED** from our **REFERENCED** AI-ML-Purified Data Set satisfied Jameel’s Stochastic ANNAF Criterion.

This will strongly establish a very high (almost perfect) correlation between the AI-ML-Purified Data Set and the set of Stochastic Activation Functions of Deep Learning Artificial Neural Network Systems. The bottom line is that **“the Deep Learning Artificial Neural Network Stochastic Activation Functions Satisfied Jameel’s Criterion**

would depend on *REFERENCED AI-ML-Purified Data Set*” to successfully perform well-informed Deep Learning Artificial Neural Network Learning processes.

Furthermore, this research REVEALED that the Advanced Activation Functions satisfied Jameel’s ANNAF Stochastic or Deterministic Criterion depends on the REFERENCED PURIFIED AI DATA SET, TIME CHANGE and AREA OF APPLICATION (acronym DTA) as shown in the figure below:

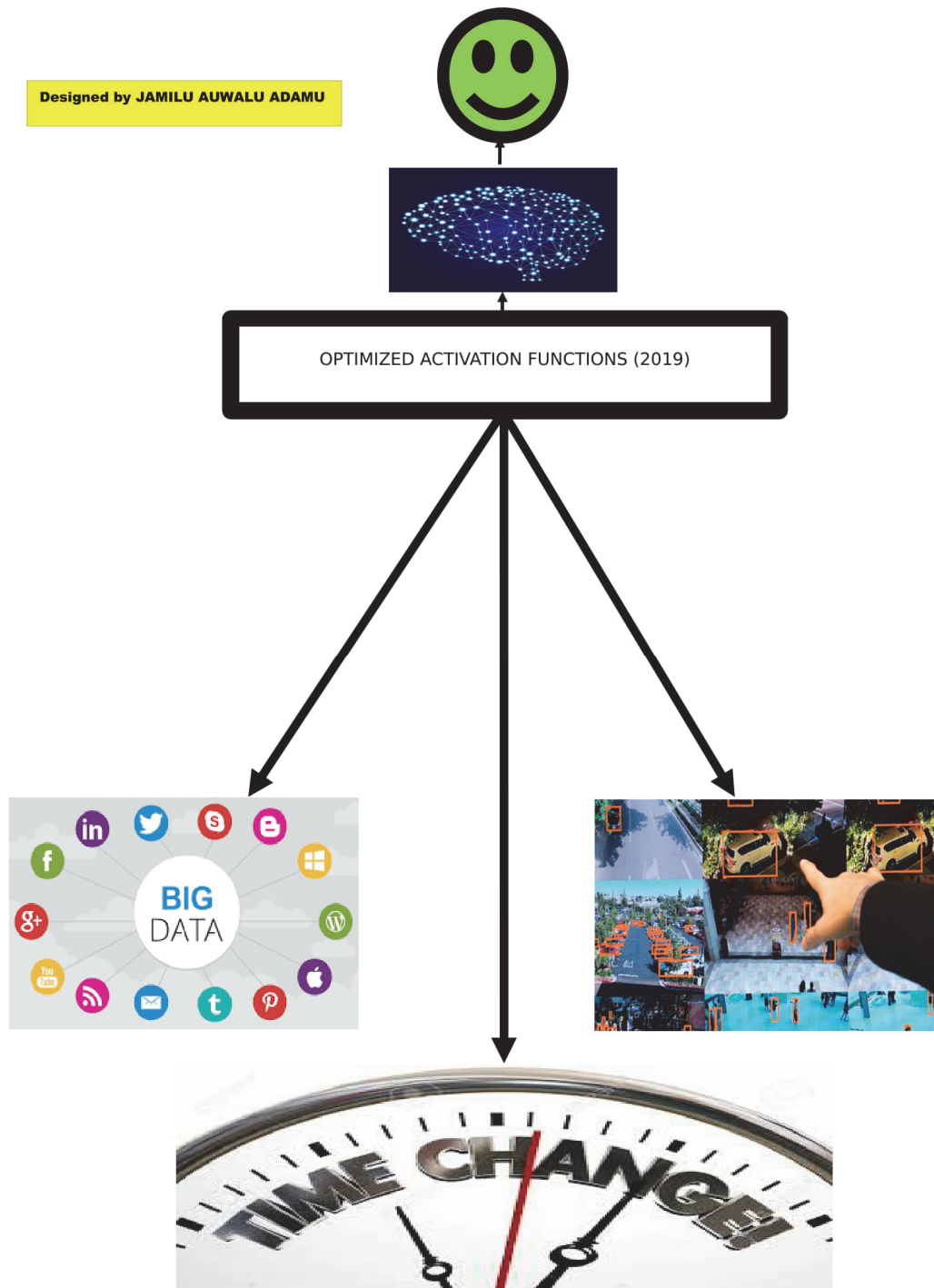


Figure 13. Optimized Activation Functions depends on AI DATA, TIME CHANGE & AREA OF APPLICATION

Source: The Author (2019)

This is against the traditional Trial and Error set of assumed Activation Functions, INDEPENDENT of **DTA**.

Research for better activation functions is still ongoing, however, the FUTURE of AI and ML lies on the emphasis given to the **DEFINITE RULES** of choosing Right Activation Functions considering the impact of LOW-PROBABILITY, HIGH-IMPACT EFFECTS, basically, attention given to FAT-TAILED EFFECTS.

This paper was **SUMMARIZED** in the following FIVE (5) Youtube Videos:

- (1) <https://www.youtube.com/watch?v=nth3cJggFts&t=5s>
- (2) <https://www.youtube.com/watch?v=lcyR4TCOBFw>
- (3) <https://www.youtube.com/watch?v=15NgJh71KRQ&t=3s>
- (4) <https://www.youtube.com/watch?v=6emMNluHMZg>
- (5) <https://www.youtube.com/watch?v=IIDTNWc7C-8>

**Declaration of Interest:** The Author reports no conflict of Interest. The views expressed in this paper are those of the Author and not his current employer.

### Acknowledgments

This research paper was a sequel to my Ph.D research work research findings extension with Ahmadu Bello University, Zaria, Nigeria. Firstly, I would like to thank the Federal Government of Nigeria through the National Mathematical Centre for releasing me for my Ph.D study. Second, I appreciate the efforts of Lan Smith, Assistant Editor, Risk, and Financial Management Journal, USA.

Also, I appreciate the Anonymous Reviewer(s) for his/her constructive criticism to improve the quality of this manuscript, however, if fact, He/she suggested to include Appendix A, B, and C for wider readability of the manuscript especially readers without strong Mathematical background.

Finally, I thank my mum (Hajiya Hauwa Ahmad), my lovely wife Halima and children Islam and Salman, they are the sources of my Creativity, Energy, and Aspiration.

### References

- Aman, D., & Payal, P. (2019). Analysis of Non-Linear Activation Functions for Classification Tasks Using Convolutional Neural Networks. *Recent Patents on Computer Science Journal*, 12(3). <https://doi.org/10.2174/2213275911666181025143029>
- Aman, D., & Payal, P. (2019). Analysis of Non-Linear Activation Functions for Classification Tasks Using Convolutional Neural Networks. *Recent Patents on Computer Science*, 12(3). <https://doi.org/10.2174/2213275911666181025143029>
- Artist Hans Hoffman wrote, “*The ability to simplify means to eliminate the unnecessary so that the necessary may speak.*” Retrieved from [https://www.brainyquote.com/quotes/hans\\_hofmann\\_107805](https://www.brainyquote.com/quotes/hans_hofmann_107805)
- Barnaby, B. et. at (2016). Complying with IFRS 9 Impairment Calculations for Retail Portfolios, Moody’s Analytics Risk Perspectives, the convergence of Risk. *Finance, and Accounting*, VII.
- Bellotti, T., & Crook, J. (2012). Loss Given Default Models Incorporating Macroeconomic Variables for Credit Cards. *International Journal of Forecasting*, 28(1), 171-182. <https://doi.org/10.1016/j.ijforecast.2010.08.005>
- Ben, S. (2019). Model Risk Management for Deep Learning and Alpha Strategies, BNP Paribas Asset Management, *Quant Summit 2019*.
- Burton, G. M. (2009). The Clustering of Extreme Movements: Stock prices and the Weather, *Princeton University*, AtanuSaha, Alixpartners, Alex Grecu, Huron Consulting Group, CEPS working paper No. 186 February, 2009.
- Casper, H. (2019). says “*Better optimized neural network; choose the right activation function, and your neural network can perform vastly better*”. Retrieved from <https://mlfromscratch.com/neural-networks-explainerd/#/>
- Chigozie, E. N. et al. (2018). Activation Functions: Comparison of Trends in Practice and Research for Deep Learning. Retrieved from <https://arxiv.org/pdf/1811.03378.pdf>
- Chigozie, E. N. et al. (2018). Activation Functions: Comparison of Trends in Practice and Research for Deep Learning, Preprint.
- Daniel, P. (2006). Estimating Probabilities of Default for German Savings Banks and Credit Cooperatives, *University of Applied Sciences, Mainz*, Ander Bruchspitze 50, D – 55122 Mainz

- David, M. R. (2012). Simulating Default Probabilities in Stress Scenarios, Presented to the *PRMIA Global Risk Conference*, New York, NY, May 14, 2012.
- David, R. (2019). Responding to the AI Challenge Learning from Physical Industries, @2019 *The Mathworks, Inc.*
- Djork-Arne Clevert, Thomas Unterthiner & Sepp Hochreiter (2016). FAST AND ACCURATE DEEP NETWORK LEARNING BY EXPONENTIAL LINEAR UNITS (ELUS), Published as a conference paper at ICLR 2016
- Jamilu, A. A. (2015). Banking and Economic Advanced Stressed Probability of Default Models. *Asian Journal of Management Sciences*, 3(08), 10-18.
- Jamilu, A. A. (2015). Estimation of Probability of Default using Advanced Stressed Probability of Default Models, Ongoing Ph.D Thesis, Ahmadu Bello University (ABU), Zaria, Nigeria.
- Jamilu, A. A. (2016). Reliable and Sophisticated Advanced Stressed Crises Compound Options Pricing Models. *Management and Organizational Studies*, 3(1), 39-55. <https://doi.org/https://doi.org/10.5430/mos.v3n1p39>.
- Jamilu, A. A. (2017). An Introduction of Jameel's Advanced Stressed Economic and Financial Crises Models and to Dramatically Increasing Markets Confidence and Drastically Decreasing Markets Risks. *International Journal of Social Science Studies*, 4(3), 39-71. <https://doi.org/https://doi.org/10.11114/ijsss.v4i3.1326>
- Jamilu, A. A. (2017). Jameel's Criterion and Jameel's Advanced Stressed Models: An Ideas that Lead to Non-Normal Stocks Brownian Motion Models. *Noble International Journal of Business and Management Research*, 1(10), 136-154. URL: <http://napublisher.org/?ic=journals&id=2>.
- Jamilu, A. A. (2017). Jameel's Criterion and Jameel's Advanced Stressed Models: An Ideas that Lead to Non-Normal Stocks Brownian Motion Models. *Noble International Journal of Business and Management Research*, 1(10), 136-154. URL: URL: <http://napublisher.org/?ic=journals&id=2>
- Jamilu, A. A. (2018). Jameel's Dimensional Stressed Default Probability Models are Indeed IFRS 9 Complaint Models. *Journal of Economics and Management Sciences*, 1(2), 104-114. <https://doi.org/10.30560/jems.v1n2p102>.
- Joonho, L. et al. (2019), ProbAct: A Probabilistic Activation Function for Deep Neural Networks, Preprint. Under review.
- Klambauer, et, al. (2017). Self-Normalizing Neural Networks, Institute of Bioinformatics, Johannes Kepler University Linz, Austria.
- Lichman. (2013). UCI machine learning repository. Retrieved from <http://archive.ics.uci.edu/ml/901>
- M, & Van, N. (1968). Fractional Brownian Motions, Fractional Noises and Applications (M & Van Ness (1968)). *SIAM Review*, 10, 422-437
- Mohit, G. et al. (2019). Learning Activation Functions: A new paradigm for understanding Neural Networks. *Proceedings of Machine Learning Research*, 101, 1–18.
- Nair, et al. (2010). Rectified linear units improve restricted boltzmann machines, ICML'10 Proceedings of the 27th International Conference on International Conference on Machine Learning Pages 807-814, Haifa, Israel — June 21 - 24, 2010.
- Nassim, N. T. (2007). Black Swans and the Domains of Statistics. *American Statistician*, 61(3).
- Nassim, N. T. (2009). Errors, Robustness, and Fourth Quadrant, New York University Polytechnic Institute and Universa Investment, United States. *International Journal of Forecasting*, 25, 744-759.
- Nassim, N. T. (2010). Convexity, Robustness, and Model Error inside the “Black Swan Domain”, Draft Version, September, 2010.
- Nassim, N. T. (2010). *Why Did the Crisis of 2008 Happen, Draft* (3rd ed.), August, 2010.
- Nassim, N. T. (2011). *A Map and Simple Heuristic to Detect Fragility, Antifragility, and Model Error* (1st ed.).
- Nassim, N. T. (2011). The Future has Thicker Tails Past: Model Error as Branching Counterfactuals, presented in Honor of Benoit Mandelbrot's at his Scientific Memorial, Yale University, April, 2011.
- Nassim, N. T. (2012). The Illusion of Thin – Tails under Aggregation, NYU – Poly, January, 2012
- Nassim, N. T. et al. (2009). Risk Externalities and Too bid to Fail, New York University Polytechnic Institute, 11201, New York, United States.
- Onali, E., & Ginesti, G. (2014). Pre-adoption Market Reaction to IFRS 9: A Cross-country Event-study. *Journal*

*of Accounting and Public Policy*, 33(6), 628-637.

- Peter Martey Addo et al. (2018). Credit Risk Analysis using Machine and Deep Learning Models. *Risks Journal, Risks*, 6(38). <https://doi.org/10.3390/risks6020038>
- Ram, Ananth, et al. (2019). Opening the “Black Box”. The Path to Deployment of AI Models in Banking, White Paper, DataRobot and REPLY AVANTAGE.
- Reney, D. E., & Michael, R. M. (2016). Forecasting of Stock Prices using Brownian Motion - Monte Carlo Simulation. Proceedings of the 2016 *International Conference on Industrial Engineering and Operations Management*. Kuala Lumpur, Malaysia, March 8-10, 2016.
- Schoenholz, et al. (2017). DEEP NEURAL NETWORKS AS GAUSSIAN PROCESSES, Published as a conference paper at ICLR 2018. Retrieved from [file:///C:/Users/pc/Downloads/Deep\\_Neural\\_Networks\\_as\\_Gaussian\\_Processes.pdf](file:///C:/Users/pc/Downloads/Deep_Neural_Networks_as_Gaussian_Processes.pdf)
- Sebastian, R. (2018), STAT 479: Machine Learning Lecture Notes. Retrieved from <http://stat.wisc.edu/~sraschka/teaching/stat479fs2018/>.
- Sebastian, U. (2017). Neural Network Architectures and Activation Functions: A Gaussian Process Approach, Technical University Munich, 2017.
- Soufiane, H. et al. (2019). On the Impact of the Activation Function on Deep Neural Networks Training. Retrieved from <https://arxiv.org/pdf/1902.06853.pdf>
- Soufiane, H. et al. (2019). On the Impact of the Activation Function on Deep Neural Networks Training, Proceedings of the 36 th International Conference on Machine Learning, Long Beach, California, PMLR 97, 2019.
- Spreadhar, T. B. et al. (2004). Forecasting Default with the KMV – Merton Model, University of Michigan, Ann Arbor MI 48109.
- Steven, R. D. (). Stochastic Processes and Advanced Mathematical Finance, The Definition of Brownian Motion and the Wiener process, Department of Mathematics, 203 Avery Hall, University of Nebraska-Lincoln, Lincoln, NE 68588-0130.
- Sven-Patrik, H. (). Machine Learning, Deep Learning, Experimental Particle Physics, University of Glasgow.
- Tidaruk, A. (2014). Mathematical Model of Stock Prices via a Fractional Brownian Motion Model with Adaptive Parameters.
- Ton, D. (2004). Simulation of Fractional Brownian Motion, Thesis, University of Twente, Department of Mathematical Sciences, P.O. BOX 217, 7500 AE Enschede, Netherlands.
- Wenyu, Zh. (2015). Introduction to Ito’s Lemma, Lecture Note, Cornell University, Department of Statistical Sciences, May 6, 2015.

## Web of References

- <https://www.stoodnt.com/blog/scopes-of-machine-learning-and-artificial-intelligence-in-banking-financial-services-ml-ai-the-future-of-fintechs/>
- <https://medium.com/datadriveninvestor/neural-networks-activation-functions-e371202b56ff>
- <https://missinglink.ai/guides/neural-network-concepts/7-types-neural-network-activation-functions-right/>
- <http://www.datastuff.tech/machine-learning/why-do-neural-networks-need-an-activation-function/>
- <https://medium.com/the-theory-of-everything/understanding-activation-functions-in-neural-networks-9491262884e0>
- <https://www.youthkiwaaz.com/2019/07/future-of-artificial-intelligence-in-banks/>
- <https://news.efinancialcareers.com/uk-en/328299/ai-in-trading-buy-side>
- <https://ai.stackexchange.com/questions/7609/is-nassim-taleb-right-about-ai-not-being-able-to-accurately-predict-certain-type/7610>
- <https://towardsdatascience.com/adam-latest-trends-in-deep-learning-optimization-6be9a291375c>
- The U.S. Hearing Youtube (Technology Companies & Algorithms): <https://www.youtube.com/watch?v=vtw4e68CkwU>



<https://www.commerce.senate.gov/2019/6/optimizing-for-engagement-understanding-the-use-of-persuasive-technology-on-internet-platforms>

<https://ai.stackexchange.com/questions/7088/how-to-choose-an-activation-function>

<https://mlfromscratch.com/activation-functions-explained/#/>

<https://github.com/nadavo/mood>.

### **Backward propagation and Activation Functions**

<https://www.youtube.com/watch?v=q555kfIFUCM>

<https://www.youtube.com/watch?v=-7scQpJT7uo>

<https://towardsdatascience.com/analyzing-different-types-of-activation-functions-in-neural-networks-which-one-to-prefer-e11649256209>

<http://vision.stanford.edu/teaching/cs231n-demos/linear-classify/>

<http://vision.stanford.edu/teaching/cs231n-demos/knn/>

<https://playground.tensorflow.org/#activation=tanh&batchSize=10&dataset=circle&regDataset=reg-plane&learningRate=0.03&regularizationRate=0&noise=0&networkShape=4,2&seed=0.41357&showTestData=false&discretize=false&percTrainData=50&x=true&y=true&xTimesY=false&xSquared=false&ySquared=false&cosX=false&sinX=false&cosY=false&sinY=false&collectStats=false&problem=classification&initZero=false&hideText=false>

## **Appendix A**

### **Research Companies and Data Sources**

This research paper was the extension of the Author's (Jamilu Auwalu Adamu (2015)) research findings published by *the Asian Journal of Management Sciences*, 03(08), 2015, 10-18 entitled "Banking and Economic Advanced Stressed Probability of Default Models". To statistically prove and explain to the general readers how the results of different types of Stressed SIGMOID and proposed Advanced Optimized Stochastic Activation Functions presented in this paper using financial industries **DATA SET**, the author uses the above research findings for illustration as repeating it by same author in this paper is an academic Tautology.

The author tested and implemented the models by considering **FINANCIAL DATA** of five (5) U.S. based companies listed on the platform of **NEW YORK STOCK EXCHANGE (NYSE)** from 2014 – 1991 (25 years) inclusive and extracted from Yahoo Finance as follows:

- 1) Microsoft Corporation (MSFT)
- 2) Exxon Mobil (XOM)
- 3) Chevron Corporation (CVX)
- 4) Honda Motor Corporation (HMC)
- 5) General Electric (GE)

### **Other Selected Data Sources**

- Yahoo Finance
- Google Finance
- Federal Reserve Bank
- Economic Research

### **Companies and Fundamental Macroeconomic Indicators used in the Research Work:**

In this research work, we consider:

- a) Five (5) companies listed on the platform of New York Stock Exchange (NYSE) namely; Chevron Corporation, Honda Motor Corporation, Microsoft Corporation, Exxon Mobil Corporation, and General Electric Corporation for the period of Twenty Five (25) years (1991 – 2014) data
- b) The stock returns of the five (5) companies under consideration
- c) The U.S. GDP
- d) The U.S. Inflation Rate

- e) The U.S. Prime Rate
- f) The U.S. unemployment Rate
- g) The U.S. USD/GBP Exchange Rate
- h) The U.S. House Price
- i) The U.S. Oil Price
- j) The U.S. Gold Price

**Simple Logistic Regression Model (SIGMOID or LOGIT):**

$$PD = \frac{1}{1 + \exp\left(\sum_{i=0}^K \beta_i X_i\right)}$$

PD is the probability of default.  $X = (X_1, X_2, \dots, X_k)$  is a vector of explanatory variables (Macro-economic Indicators).

The author **disturbs** it by introducing:

- (i) Geometric Mean ( $\mu_A$ ) and Volatility ( $\sigma_A$ ) of the Macroeconomic Indicators used in the research work;
- (ii) Geometric Mean ( $\mu_{Company}$ ) and Volatility ( $\sigma_{Company}$ ) of the Stock Return used in the research work; and
- (iii) Probability Distribution Function  $f(x; \mu_{company}, \sigma_{company}, \xi)$ .

As shown below:

Type A:

$$PD_{Stress} = \frac{1}{1 + \exp\left(\mu_A \left(\sum_{i=0}^K \beta_i X_i\right) \mp \sigma_A f(x; \mu_{company}, \sigma_{company}, \xi)\right)}$$

Type B:

$$PD_{Stress} = \frac{1}{1 + \exp\left(\mu_A \left(\sum_{i=0}^K \beta_i X_i\right) \mp f(x; \mu_{company}, \sigma_{company}, \xi)\right)}$$

Type C:

$$PD_{Stress} = \frac{1}{1 + \exp\left(\sum_{i=0}^K \beta_i X_i\right) \mp \sigma_A f(x; \mu_{company}, \sigma_{company}, \xi)}$$

Type D:

$$PD_{Stress} = \frac{1}{1 + \exp\left(\sum_{i=0}^K \beta_i X_i\right) \mp f(x; \mu_{company}, \sigma_{company}, \xi)}$$

**Stressed PROBIT:**

Type A:

$$PD_{Stress} = \Phi \left( \beta_0 + \mu_A \sum_{j=1}^J \beta_j X_j \right) \pm \sigma_A f(x; \mu_{company}, \sigma_{company}, \xi)$$

**Type B:**

$$PD_{Stress} = \Phi \left( \beta_0 + \mu_A \sum_{j=1}^J \beta_j X_j \right) \pm f(x; \mu_{company}, \sigma_{company}, \xi)$$

**Type C:**

$$PD_{Stress} = \Phi \left( \beta_0 + \sum_{j=1}^J \beta_j X_j \right) \pm \sigma_A f(x; \mu_{company}, \sigma_{company}, \xi)$$

**Type D:**

$$PD_{Stress} = \Phi \left( \beta_0 + \sum_{j=1}^J \beta_j X_j \right) \pm f(x; \mu_{company}, \sigma_{company}, \xi)$$

Using Multiple Regression Analysis on the above Macroeconomic Indicators and **QI Macros 2015 Software**, we obtained the following components for determining the proposed Banking and Economic Advanced Stressed Probability of Default models for the five (5) companies under considerations.

**Multiple Regression Model Component of CHEVRON Corporation (CVX) for calculating Probability of Default:**

$$Y_{CHEVRON} = 0.004 + 0.004 \times \Delta P(CHEVRON) - 0.199 \times \Delta P(GDP) + 0.009 \times \Delta P(OIL) + 0.009 \times \Delta P(INF) \\ - 0.018 \times \Delta P(UER) + 0.002 \times \Delta P(GOLD) + 0 \times \Delta P(INTEREST) + 0 \times \Delta P(USD / GBP)$$

**Multiple Regression Model Component of GENERAL ELECTRIC(GE) for calculating Probability of Default:**

$$Y_{GE} = 0.004 - 0.001 \times \Delta P(GE) - 0.207 \times \Delta P(GDP) + 0.009 \times \Delta P(OIL) + 0.016 \times \Delta P(INF) \\ - 0.017 \times \Delta P(UER) - 0.001 \times \Delta P(GOLD) + 0 \times \Delta P(INTEREST) + 0 \times \Delta P(USD / GBP)$$

**Multiple Regression Model Component of MICROSOFT (MSFT) Corporation for calculating Probability of Default:**

$$Y_{MSFT} = 0.004 - 0.006 \times \Delta P(MSFT) - 0.189 \times \Delta P(GDP) + 0.009 \times \Delta P(OIL) + 0.011 \times \Delta P(INF) \\ - 0.017 \times \Delta P(UER) + 0.001 \times \Delta P(GOLD) + 0 \times \Delta P(INTEREST) + 0 \times \Delta P(USD / GBP)$$

**Multiple Regression Model Component of EXXON MOBIL (XOM) Corporation for calculating Probability of Default:**

$$Y_{XOM} = 0.004 + 0.002 \times \Delta P(XOM) - 0.2 \times \Delta P(GDP) + 0.009 \times \Delta P(OIL) + 0.01 \times \Delta P(INF) \\ - 0.018 \times \Delta P(UER) + 0.001 \times \Delta P(GOLD) + 0 \times \Delta P(INTEREST) + 0 \times \Delta P(USD / GBP)$$

**Multiple Regression Model Component of HONDA MOTOR CO., Ltd for calculating Probability of Default:**

$$Y_{HONDA} = 0.004 - 0.004 \times \Delta P(HMC) - 0.204 \times \Delta P(GDP) + 0.009 \times \Delta P(OIL) + 0.01 \times \Delta P(INF) \\ - 0.018 \times \Delta P(UER) + 0.001 \times \Delta P(GOLD) + 0 \times \Delta P(INTEREST) + 0 \times \Delta P(USD / GBP)$$

**JAMEEL'S CRITERION (INITIAL VERSION)**

In this test of Goodness of fit, we consider Three (3) criteria:

- (i) We accept if the Average of the ranks of Kolmogorov Smirnor, Anderson Darling and Chi-squared is less than or equal to Three (3)
- (ii) At least Two (2) probability distributions must included in the selection
- (iii) We select the most occur probability distribution as the qualify candidate in each case of test of goodness of fit of the stock returns as follows

**LOG – LOGISTIC (3P) PROBABILITY DISTRIBUTION**

$$f(x; \mu, \sigma, \xi) = \begin{cases} \frac{\left(1 + \frac{\xi(x - \mu)}{\sigma}\right)^{-\left(\frac{1}{\xi} + 1\right)}}{\left[1 + \left(1 + \frac{\xi(x - \mu)}{\sigma}\right)^{-\frac{1}{\xi}}\right]^2} & ; x \geq \mu \\ 0 & ; \textit{elsewhere} (x \leq \mu) \end{cases}$$

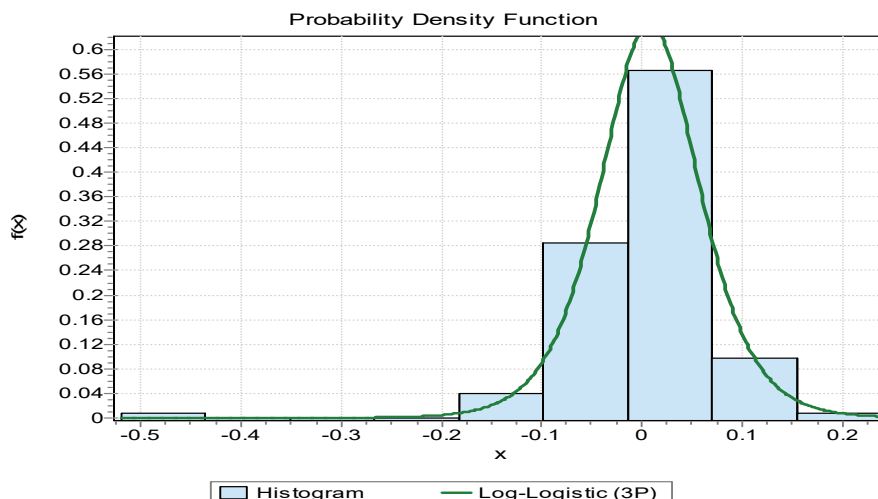
is called Generalized Log-Logistic or Log-Logistic (3P) Probability Distribution. Where,  $\mu \in \mathbb{R}$  is the location parameter,  $\sigma > 0$  the scale parameter and  $\xi \in \mathbb{R}$  the shape parameter. The shape parameter  $\xi$  is often restricted to lie in  $[-1, 1]$ , when the probability density function is bounded.

However, in this research work we will restrict  $\xi$  to three cases only  $\xi = 1$ . However, one can test, for  $\xi = 0$  or  $\xi = -1$  in the subsequent researchers.

**Appendix B**

**Statistical Data Analysis**

**SUMMARY RESULT OF TEST OF GOODNESS OF FIT OF CHEVRON CORPORATION STOCK RETURN**



Name of Probability Distribution	Kolmogorov Smirnov Rank	Anderson Darling Rank	Chi- Squared Rank	Average
Burr	3	3	2	2.07
Dagum	2	1	1	1.33
Log - Logistic(3P)	1	2	3	2.0

**Chevron Calculated Parameters:**

Using our **data sources, Microsoft EXCEL and QI Macros 2015**, we obtained the Chevron calculated parameters as follows:

Geometric Mean and Volatility (Standard Deviation) of the Macroeconomic indicators used in the case of **Chevron** are given by:

$$\mu_{GEO}(Chevron) = 0.030383975$$

$$\sigma_{GEO}(Chevron) = 0.111414539$$

Geometric Mean and Volatility (Standard Deviation) of **Chevron Stock Return** are given by:

$$\mu_{STOCK}(Chevron) = 0.004402791$$

$$\sigma_{STOCK}(Chevron) = 0.06909299$$

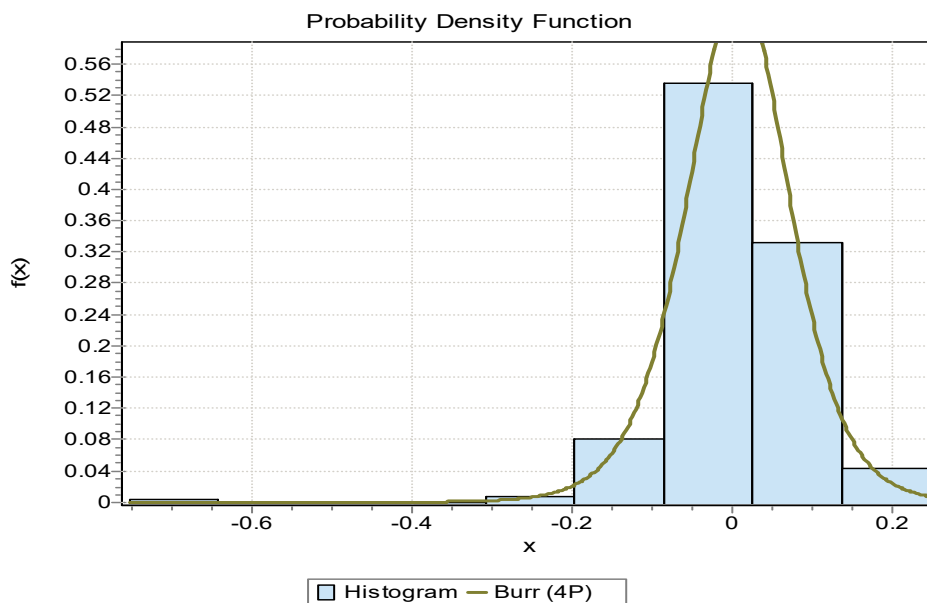
**Results of Stressed SIGMOID and Probit MODELS for Chevron Corporation:**

Under Chevron Corporation, on the month of June, 2014, the probability of default using the existing Sigmoid is 0.499097747% and that of probit is 0.501439786%, whereas, using Stressed Sigmoid and Probit models are: 0.499976914%, 0.499968258%, 0.500011436%, 0.499933742%, 0.49910206%, 0.499093434%, 0.499136461%, 0.499059039%, 0.501573711%, 0.50160834%, 0.501435622%, 0.501746429%, 0.501422471%, 0.5014571%, 0.501284382%, and 0.501595189% respectively. Generally, this can be seen in the table below:

Date	M1								LOGIT	T	M2							
	TYPE	TYPE	TYPE	TYPE	TYPE	TYPE	TYPE	TYPE			PROBI	TYPE	TYPE	TYPE	TYPE	TYPE	TYPE	TYPE
	A-	A+	B-	B+	C-	C+	D-	D+			A-	A+	B-	B+	C-	C+	D-	D+
01/12/20	0.4999	0.49996	0.4999	0.49995	0.4989	0.49896	0.4989	0.49895	0.4989	0.5016	0.5015	0.50160	0.5015	0.50167	0.5016	0.50165	0.5015	0.50175
14	70779	6686	87104	0363	72963	8885	89222	2627	70924	42165	88987	5363	23683	0667	33977	0353	68673	657
03/11/20	0.4999	0.49998	0.4999	0.49998	0.4996	0.49960	0.4996	0.49960	0.4996	0.5006	0.5015	0.50156	0.5015	0.50156	0.5006	0.50062	0.5006	0.500629
14	88014	8014	88014	8014	055	55	055	55	055	29531	66407	6407	66407	6407	29531	9531	29531	531
01/10/20	0.4999	0.49999	0.4999	0.49999	0.4999	0.49992	0.4999	0.49992	0.4999	0.5001	0.5015	0.50155	0.5015	0.50155	0.5001	0.50012	0.5001	0.500126
14	97599	7596	9761	7584	20916	0913	20928	0902	20915	26202	51108	112	51061	1167	26196	6208	26149	254
02/09/20	0.4999	0.49998	0.4999	0.49998	0.4994	0.49941	0.4994	0.49941	0.4994	0.5009	0.5015	0.50157	0.5015	0.50157	0.5009	0.50092	0.5009	0.500929
14	82294	2294	82294	2294	17245	7245	17245	7245	17245	29942	75535	5535	75535	5535	29942	9942	29942	942
01/08/20	0.4999	0.49998	0.4999	0.49998	0.4993	0.49939	0.4993	0.49939	0.4993	0.5009	0.5015	0.50157	0.5015	0.50157	0.5009	0.50096	0.5009	0.500970
14	81556	1525	81677	1404	92475	2445	92596	2323	9246	69493	76675	6797	76189	7284	69432	9554	68946	041
01/07/20	0.4999	0.49999	0.4999	0.49999	0.4998	0.49985	0.4998	0.49984	0.4998	0.5002	0.5015	0.50155	0.5015	0.50157	0.5002	0.50023	0.5002	0.500252
14	9598	5049	9969	1338	52831	1901	56539	8192	52366	3559	52576	6299	37734	1141	33729	7451	18886	294
02/06/20	0.4999	0.49996	0.5000	0.49993	0.4991	0.49909	0.4991	0.49905	0.4990	0.5014	0.5015	0.50160	0.5014	0.50174	0.5014	0.50145	0.5012	0.501595
14	76914	8258	11436	3742	0206	3434	36461	9039	97747	39786	73711	834	35622	6429	22471	71	84382	189
01/05/20	0.4999	0.49997	0.4999	0.49996	0.4992	0.49927	0.4992	0.49925	0.4992	0.5011	0.5015	0.50158	0.5015	0.50164	0.5011	0.50116	0.5010	0.501221
14	7971	6176	93802	2085	75826	2302	89878	825	74064	58426	75409	9545	19036	5917	51358	5494	94985	867
01/04/20	0.4999	0.49996	0.4999	0.49993	0.4989	0.49889	0.4989	0.49886	0.4988	0.5017	0.5015	0.50161	0.5014	0.50172	0.5017	0.50177	0.5016	0.501881
14	69966	3096	97362	5704	01885	5045	29164	7769	98465	57793	86947	4429	77356	402	44052	1534	34462	125
03/03/20																		







Name of Probability Distribution	Kolmogorov Smirnov Rank	Anderson Darling Rank	Chi-Squared Rank	Average
Burr	1	2	1	1.33
Dagum	2	3	2	2.33
Log - Logistic(3P)	3	1	3	2.33

Geometric Mean and Volatility (Standard Deviation) of the Macroeconomic indicators used in the case of **Honda Motor** are given by:

$$\mu_{GEO}(Honda) = 0.031352397$$

$$\sigma_{GEO}(Honda) = 0.114001187$$

Geometric Mean and Volatility (Standard Deviation) of **Honda Motor Stock Return** are given by:

$$\mu_{STOCK}(Honda) = 0.005839335$$

$$\sigma_{STOCK}(Honda) = 0.084945727$$

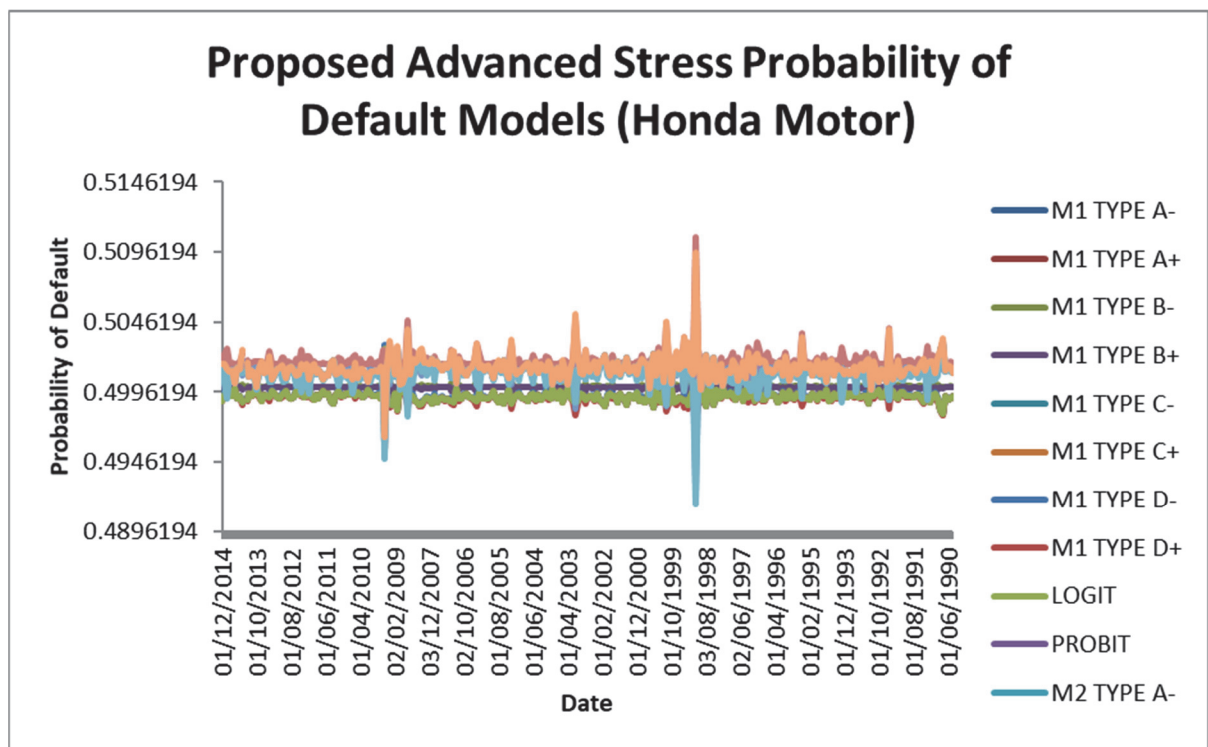
**Results of Stressed SIGMOID and Probit MODELS for Honda Motor Corporation:**

Under Honda Motor, on the month of December, 2014, the probability of default using the existing Sigmoid is 0.49898654% and that of probit is 0.501617235%, whereas, using Stressed Sigmoid and Probit models are: 0.49996827%, 0.499968182%, 0.49996784%, 0.49896591%, 0.498986503%, 0.498986932%, 0.498986163%, 0.501594751%, 0.501595103%, 0.501593383%, 0.50159647%, 0.501617059%, 0.50161741%, 0.501615691%, and 0.501618778% respectively..This can be seen below:

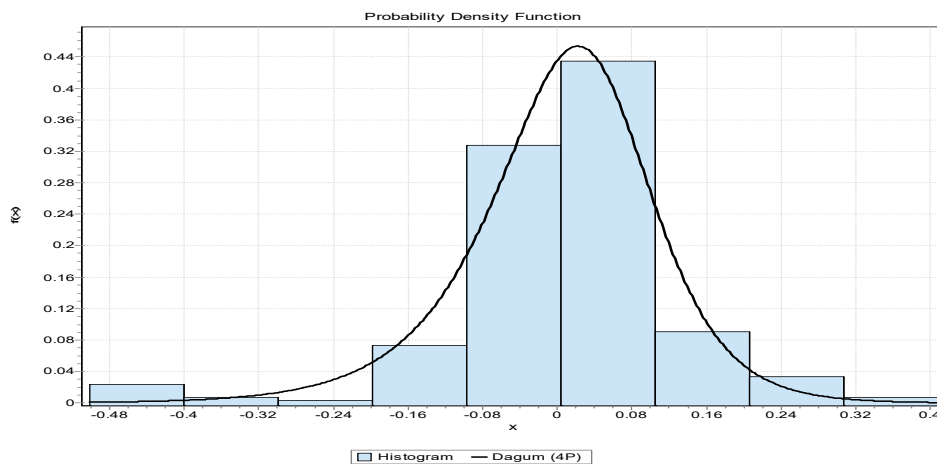
Date	M1								LOGIT	T	M2							
	A-	A+	B-	B+	C-	C+	D-	D+			A-	A+	B-	B+	C-	C+	D-	D+
01/12/2014	0.4999	0.49996	0.4999	0.49996	0.4989	0.49898	0.4989	0.49898	0.4989	0.5016	0.5015	0.50159	0.5015	0.50159	0.5016	0.50161	0.5016	0.50161
03/11/2014	0.5000	0.49996	0.5001	0.49980	0.4994	0.49942	0.4996	0.49926	0.4994	0.5008	0.5015	0.50170	0.5009	0.50234	0.5008	0.50097	0.5001	0.50160
01/10/2014	0.5000	0.49996	0.5002	0.49972	0.4998	0.49983	0.5001	0.49959	0.4998	0.5002	0.5015	0.50176	0.5005	0.50272	0.5000	0.50033	0.4991	0.50129
2014	26642	5058	66096	5893	98359	6807	3769	7764	67581	1131	22306	8645	65052	5899	8814	4479	30886	1733

02/09/	0.4999	0.49997	0.4999	0.49997	0.4993	0.49933	0.4993	0.49933	0.4993	0.5010	0.5016	0.50161	0.5016	0.50161	0.5010	0.50106	0.5010	0.50106
2014	79001	8997	79016	8981	30153	0149	30169	0134	30151	68923	11599	1615	11536	1677	68915	8931	68852	8994
01/08/	0.4999	0.49997	0.5000	0.49995	0.4993	0.49936	0.4993	0.49933	0.4993	0.5010	0.5016	0.50162	0.5015	0.50171	0.5010	0.50102	0.5009	0.50111
2014	82846	722	04707	5362	65952	034	87759	8535	63146	16271	05063	7566	17616	5013	0502	7523	17573	497
01/07/	0.4999	0.49999	0.4999	0.49999	0.4998	0.49985	0.4998	0.49985	0.4998	0.5002	0.5016	0.50163	0.5016	0.50164	0.5002	0.50023	0.5002	0.50023
2014	957	5279	97338	3642	5635	5929	57986	4292	56139	29569	3807	9755	31521	6305	28726	0411	22176	6961
02/06/	0.4999	0.49997	0.4999	0.49997	0.4991	0.49914	0.4991	0.49914	0.4991	0.5013	0.5016	0.50160	0.5015	0.50161	0.5013	0.50136	0.5013	0.50137
2014	7364	2967	76255	0352	4883	8159	51436	5552	48494	58805	02388	5081	91927	5542	57459	0151	46998	0613
01/05/	0.4999	0.49997	0.5000	0.49993	0.4993	0.49929	0.4993	0.49926	0.4992	0.5011	0.5015	0.50162	0.5014	0.50175	0.5011	0.50113	0.5009	0.50127
2014	82452	3628	16742	9343	03967	5167	38164	0975	99567	17729	87455	2752	50293	9914	0008	5378	62919	254
01/04/	0.4999	0.49994	0.5001	0.49976	0.4988	0.49885	0.4990	0.49868	0.4988	0.5017	0.5015	0.50168	0.5008	0.50238	0.5017	0.50188	0.5010	0.50258
2014	87283	2305	62143	7598	99443	466	73543	0713	77051	91965	0542	5359	06189	459	01995	1935	02765	1165
03/03/																		
2014																		
03/02/	0.4999	0.49996	0.5000	0.49990	0.4990	0.49906	0.4991	0.49900	0.4990	0.5014	0.5015	0.50163	0.5013	0.50187	0.5014	0.50150	0.5011	0.50174
2014	78844	3263	39398	2728	84505	898	44843	8661	76742	73304	72334	4664	30127	6871	4214	4469	99932	6677
02/01/	0.4999	0.49998	0.4999	0.49998	0.4995	0.49955	0.4995	0.49955	0.4995	0.5007	0.5016	0.50163	0.5016	0.50163	0.5007	0.50070	0.5007	0.50070
2014	86114	6114	86114	6114	57102	7102	57102	7102	57102	06763	32929	2929	32929	2929	06763	6763	06763	6763
02/12/	0.4999	0.49997	0.5000	0.49995	0.4993	0.49935	0.4993	0.49933	0.4993	0.5010	0.5016	0.50162	0.5015	0.50171	0.5010	0.50103	0.5009	0.50112
2013	8262	6974	04563	5033	58429	2796	80317	0911	55612	28293	04648	7236	16875	5009	16999	9587	29226	7361
01/11/	0.4999	0.49997	0.5000	0.49993	0.4994	0.49946	0.4995	0.49942	0.4994	0.5008	0.5015	0.50163	0.5014	0.50178	0.5008	0.50086	0.5006	0.50101
2013	88346	8427	26895	9885	7504	5141	1351	6679	7009	45613	93434	3114	39243	7305	25773	5453	71582	9644
01/10/	0.5000	0.49999	0.5000	0.49997	0.5000	0.50008	0.5001	0.50006	0.5000	0.4998	0.5016	0.50165	0.5015	0.50175	0.4998	0.49987	0.4997	0.49997
2013	05942	9583	30655	4873	91295	4934	16016	0215	88114	5939	32711	8147	33867	6992	46672	2108	47827	0953
03/09/	0.4999	0.49997	0.5000	0.49993	0.4993	0.49929	0.4993	0.49925	0.4992	0.5011	0.5015	0.50162	0.5014	0.50178	0.5011	0.50114	0.5009	0.50130
2013	83084	2806	2303	2869	01673	1423	41509	1594	96548	22547	83918	5035	24139	4814	01988	3105	42209	2884

Source: Jamilu Auwalu Adamu (2015), Asian Journal of Management Sciences, 03(08), 2015, 10-18



**SUMMARY RESULT OF TEST OF GOODNESS OF FIT OF MICROSOFT CORPORATION STOCK RETURN**



Name of Probability Distribution	Kolmogorov Smirnov Rank	Anderson Rank	Darling	Chi- Squared Rank	Average
Dagum	1	1		1	1.0
Log - Logistic(3P)	2	4		4	3.33

Geometric Mean and Volatility (Standard Deviation) of the Macroeconomic indicators used in the case of **Microsoft Corporation** are given by:

$$\mu_{GEO}(MSFT) = 0.031352397$$

$$\sigma_{GEO}(MSFT) = 0.117906073$$

Geometric Mean and Volatility (Standard Deviation) of **Microsoft Corporation Stock Return** are given by:

$$\mu_{STOCK}(MSFT) = 0.006798657$$

$$\sigma_{STOCK}(MSFT) = 0.115022493$$

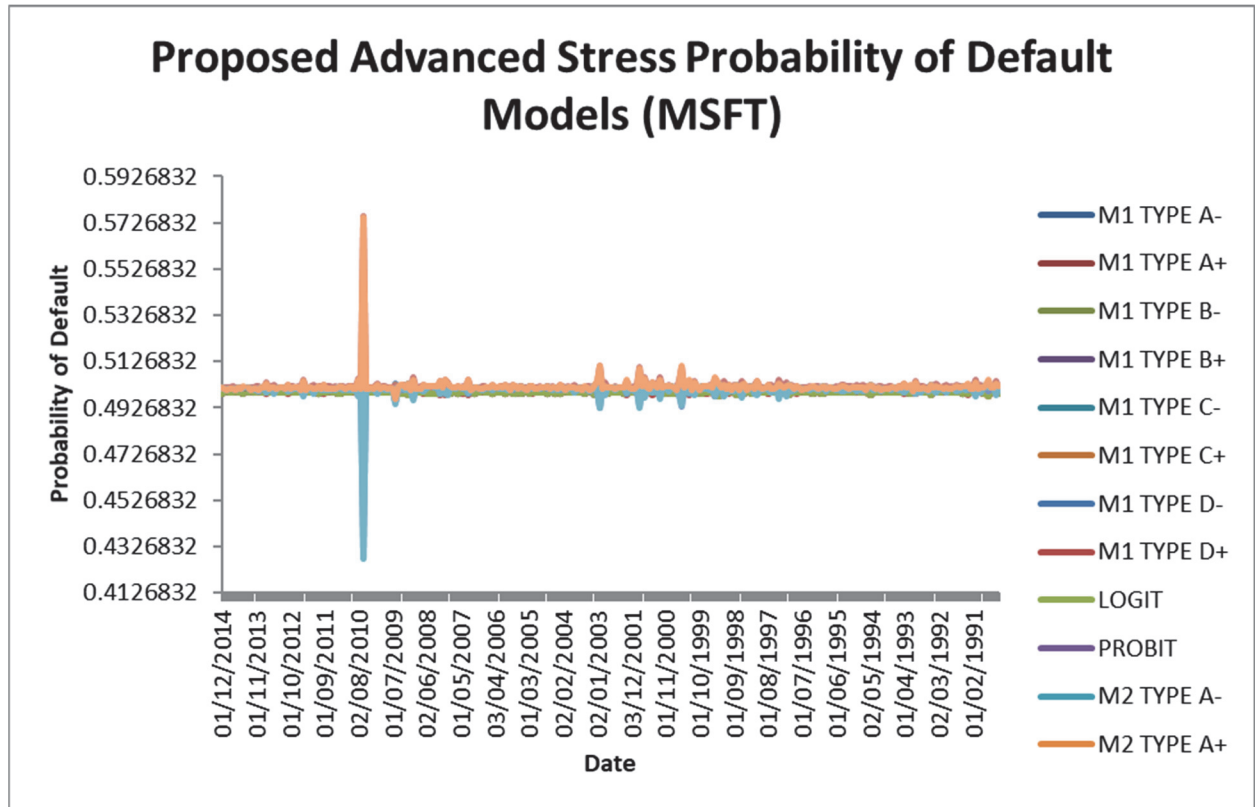
**Results of Stressed SIGMOID and Probit MODELS for Microsoft Corporation:**

Under Microsoft Corporation, on the month of December, 2014, the probability of default using the existing Simoid is 0.499506722% and that of probit is 0.500787158%, whereas, the Stressed Sigmoid and Probit models are: 0.499984689%, 0.499983853%, 0.499987817%, 0.499980725%, 0.499507139%, 0.499506304%, 0.49951026%, 0.499503183%, 0.50161708%, 0.501620425%, 0.501604569%, 0.501632936%, 0.500785486%, 0.50078831%, and 0.500772975%, 0.500801342% respectively. This can be seen as follows:

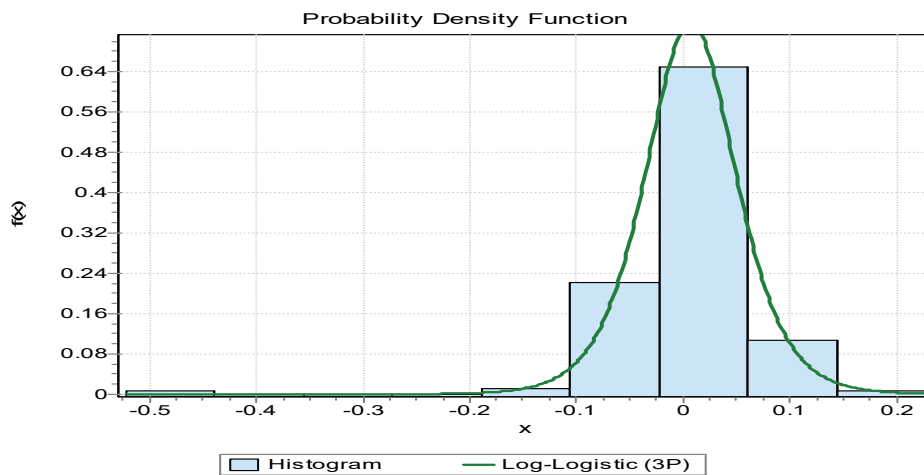
Date	M1								LOGIT	T	M2							
	A-	A+	B-	B+	C-	C+	D-	D+			PROBI	A-	A+	B-	B+	C-	C+	D-
01/12/2014	0.4999	0.49996	0.4999	0.49996	0.4989	0.49894	0.4989	0.49894	0.4989	0.5016	0.5015	0.50159	0.5015	0.50159	0.5016	0.50167	0.5016	0.50167
03/11/2014	0.4999	0.49998	0.4999	0.49998	0.4995	0.49950	0.4995	0.49950	0.4995	0.5007	0.5016	0.50162	0.5016	0.50163	0.5007	0.50078	0.5007	0.50080
01/10/2014	0.4999	0.49999	0.4999	0.49999	0.4998	0.49989	0.4998	0.49989	0.4998	0.5001	0.5016	0.50163	0.5016	0.50164	0.5001	0.50016	0.5001	0.50017
02/09/2014	0.4999	0.49997	0.4999	0.49997	0.4993	0.49933	0.4993	0.49932	0.4993	0.5010	0.5016	0.50161	0.5015	0.50162	0.5010	0.50106	0.5010	0.50108
01/08/2014	0.4999	0.49997	0.5000	0.49993	0.4994	0.49944	0.4994	0.49940	0.4994	0.5008	0.5015	0.50163	0.5014	0.50179	0.5008	0.50090	0.5006	0.50106
01/07/2014	0.4999	0.49999	0.5000	0.49997	0.4998	0.49984	0.4998	0.49982	0.4998	0.5002	0.5016	0.50164	0.5015	0.50171	0.5002	0.50025	0.5001	0.50031
02/06/2014	0.4999	0.49997	0.4999	0.49996	0.4991	0.49916	0.4991	0.49915	0.4991	0.5013	0.5015	0.50160	0.5015	0.50161	0.5013	0.50134	0.5013	0.50135
01/05/2014	0.4999	0.49997	0.4999	0.49997	0.4992	0.49924	0.4992	0.49924	0.4992	0.5012	0.5016	0.50160	0.5016	0.50161	0.5012	0.50121	0.5012	0.50121
01/04/2014	0.4999	0.49996	0.4999	0.49995	0.4989	0.49891	0.4989	0.49889	0.4989	0.5017	0.5015	0.50160	0.5015	0.50165	0.5017	0.50174	0.5016	0.50179
03/03/2014	0.4999	0.49997	0.5000	0.49993	0.4991	0.49911	0.4991	0.49911	0.4991	0.5014	0.5015	0.50160	0.5015	0.50160	0.5014	0.50141	0.5014	0.50141
02/01/2014	0.4999	0.49998	0.4999	0.49998	0.4996	0.49961	0.4996	0.49961	0.4996	0.5006	0.5016	0.50162	0.5016	0.50162	0.5006	0.50061	0.5006	0.50061
02/12/2013	0.4999	0.49997	0.5000	0.49995	0.4993	0.49932	0.4993	0.49930	0.4993	0.5010	0.5016	0.50162	0.5015	0.50170	0.5010	0.50108	0.5009	0.50116
01/11/2013	0.4999	0.49997	0.5000	0.49988	0.4995	0.49949	0.4995	0.49940	0.4995	0.5007	0.5015	0.50165	0.5012	0.50199	0.5007	0.50083	0.5004	0.50117
01/10/2013	0.5000	0.49999	0.5000	0.49993	0.5000	0.50006	0.5001	0.50000	0.5000	0.4998	0.5016	0.50167	0.5013	0.50191	0.4998	0.49991	0.4996	0.50015
03/09/2013	0.4999	0.49997	0.4999	0.49997	0.4992	0.49920	0.4992	0.49920	0.4992	0.5012	0.5016	0.50160	0.5015	0.50162	0.5012	0.50126	0.5012	0.50127
01/08/2013	0.4999	0.49996	0.5000	0.49993	0.4990	0.49908	0.4991	0.49905	0.4990	0.5014	0.5015	0.50161	0.5014	0.50175	0.5014	0.50146	0.5012	0.50160

2013	75756	6337	10991	1108	96682	7297	31794	2191	9199	48974	74127	1806	33181	2752	30134	7813	89188	8759
01/07/	0.5000	0.49991	0.5004	0.49948	0.4993	0.49919	0.4997	0.49876	0.4992	0.5011	0.5013	0.50185	0.4996	0.50360	0.5009	0.50142	0.4992	0.50317
2013	34658	7798	72277	1148	12548	6026	48904	0635	5428	89996	86894	4379	3819	3082	56253	3738	0755	2442

Source: Jamilu Auwalu Adamu (2015), Asian Journal of Management Sciences, 03(08), 2015, 10-18



### SUMMARY RESULT OF TEST OF GOODNESS OF FIT OF EXXON MOBIL CORPORATION STOCK RETURN



Name of Probability Distribution	Kolmogorov Smirnov Rank	Anderson Rank	Darling Rank	Chi-Squared Rank	Average
----------------------------------	-------------------------	---------------	--------------	------------------	---------

Burr	3	3	3	3.00
Dagum	2	2	4	2.67
Log - Logistic(3P)	1	1	1	1.00

Geometric Mean and Volatility (Standard Deviation) of the Macroeconomic indicators used in the case of **Exxon Mobil Corporation** are given by:

$$\mu_{GEO}(XOM) = 0.030729517$$

$$\sigma_{GEO}(XOM) = 0.110236167$$

Geometric Mean and Volatility (Standard Deviation) of **Exxon Mobil Corporation Stock Return** are given by:

$$\mu_{STOCK}(XOM) = 0.00487448$$

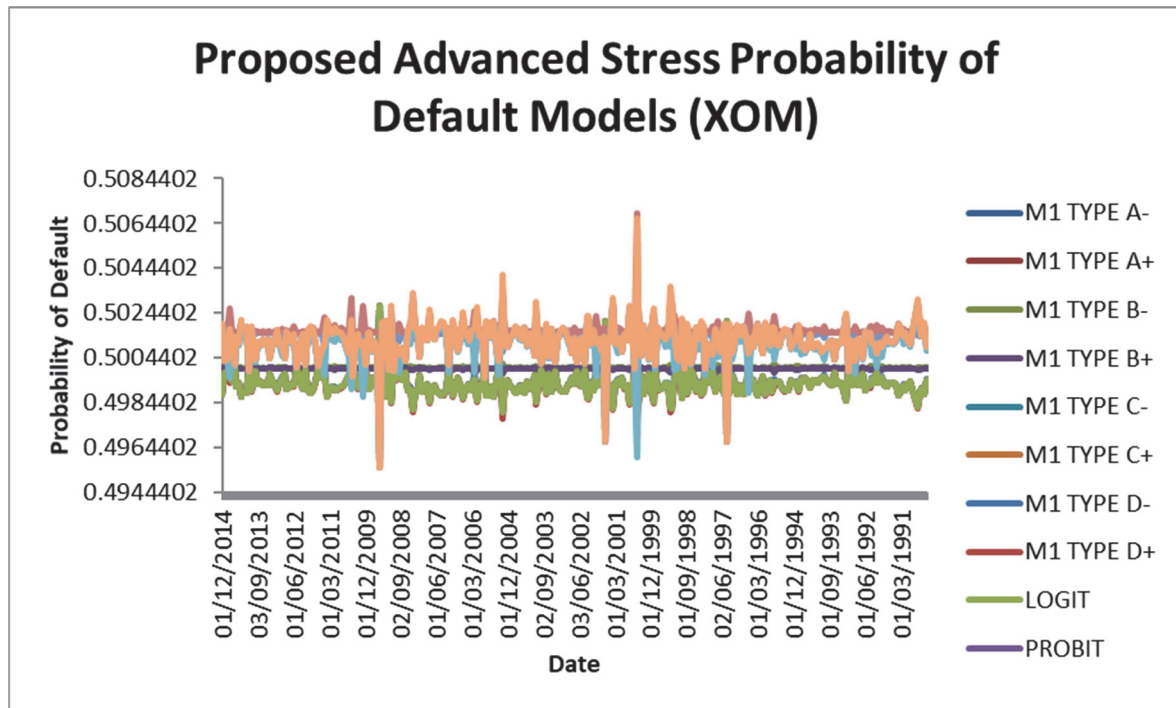
**Results of Stressed SIGMOID and Probit MODELS for Exxon Mobil Corporation:**

Under Exxon Mobil, on the month of October, 2014, the probability of default using the existing Sigmoid is 0.499779809% and that of probit is 0.500351375%, whereas, using the Stressed Sigmoid and Probit models are: 0.499996206%, 0.499994665%, 0.500002424%, 0.4999988447%, 0.499780578%, 0.499779039%, 0.499786791%, 0.499772826%, 0.5015668886%, 0.501573051%, 0.501542014%, 0.501597925%, 0.500348293%, 0.500354456%, 0.500323419%, and 0.500379339% respectively. This can be seen below:

Date	M1								LOGIT	T	M2							
	A-	A+	B-	B+	C-	C+	D-	D+			A-	A+	B-	B+	C-	C+	D-	D+
01/12/2014	0.4999	0.49997	0.4999	0.49996	0.4988	0.49880	0.4988	0.49879	0.4988	0.5019	0.5015	0.50160	0.5015	0.50164	0.5019	0.50191	0.5018	0.50195
03/11/2014	0.4999	0.49999	0.4999	0.49999	0.4998	0.49981	0.4998	0.49981	0.4998	0.5002	0.5015	0.50156	0.5015	0.50156	0.5002	0.50029	0.5002	0.50029
01/10/2014	0.4999	0.49999	0.5000	0.49998	0.4997	0.49977	0.4997	0.49977	0.4997	0.5003	0.5015	0.50157	0.5015	0.50159	0.5003	0.50035	0.5003	0.50037
02/09/2014	0.5000	0.49996	0.5002	0.49972	0.4996	0.49956	0.4998	0.49933	0.4995	0.5006	0.5014	0.50169	0.5005	0.50263	0.5005	0.50076	0.4995	0.50170
01/08/2014	0.4999	0.49998	0.4999	0.49998	0.4993	0.49936	0.4993	0.49936	0.4993	0.5010	0.5015	0.50158	0.5015	0.50158	0.5010	0.50102	0.5010	0.50102
01/07/2014	0.4999	0.49999	0.5000	0.49998	0.4999	0.49993	0.4999	0.49992	0.4999	0.5001	0.5015	0.50156	0.5015	0.50160	0.5001	0.50010	0.5000	0.50014
02/06/2014	0.4999	0.49998	0.4999	0.49998	0.4991	0.49914	0.4991	0.49914	0.4991	0.5013	0.5015	0.50159	0.5015	0.50159	0.5013	0.50136	0.5013	0.50136
01/05/2014	0.4999	0.49998	0.4999	0.49997	0.4993	0.49933	0.4993	0.49932	0.4993	0.5010	0.5015	0.50158	0.5015	0.50162	0.5010	0.50106	0.5010	0.50110
01/04/2014	0.4999	0.49997	0.4999	0.49995	0.4987	0.49869	0.4987	0.49867	0.4986	0.5020	0.5015	0.50161	0.5015	0.50168	0.5020	0.50208	0.5019	0.50216
03/03/2014	0.4999	0.49997	0.4999	0.49995	0.4988	0.49889	0.4988	0.49887	0.4988	0.5017	0.5015	0.50160	0.5015	0.50167	0.5017	0.50177	0.5016	0.50183
02/01/2014	0.5000	0.50000	0.5000	0.50000	0.5000	0.50008	0.5000	0.50008	0.5000	0.4998	0.5015	0.50155	0.5015	0.50155	0.4998	0.49985	0.4998	0.49985
02/12/2013	0.4999	0.49997	0.5000	0.49992	0.4989	0.49896	0.4990	0.49891	0.4989	0.5016	0.5015	0.50162	0.5013	0.50181	0.5016	0.50167	0.5014	0.50186
2013	84537	2617	32646	4519	7247	0599	20385	2695	66534	4917	73031	0713	80598	3145	25329	3012	32897	5444

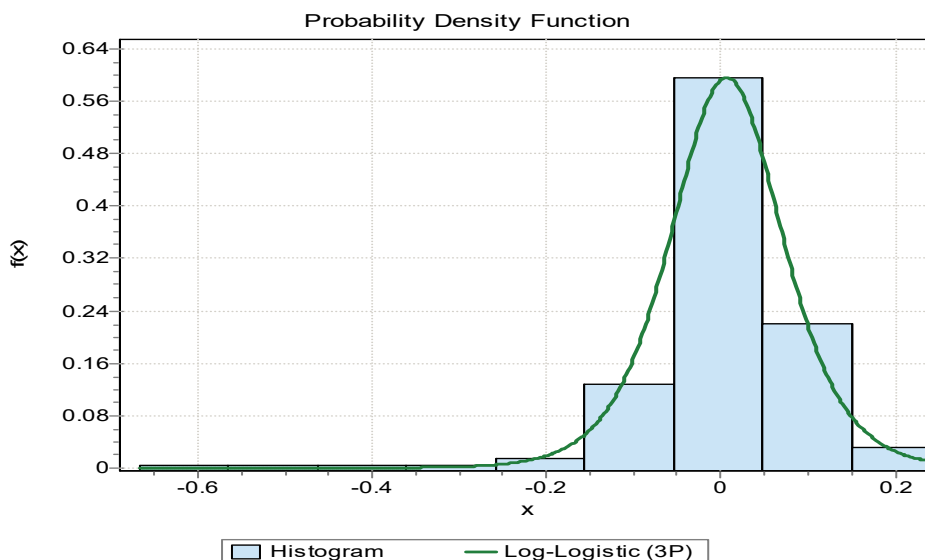
01/11/2013	0.4999	0.49998	0.4999	0.49996	0.4991	0.49919	0.4992	0.49917	0.4991	0.5012	0.5015	0.50159	0.5015	0.50165	0.5012	0.50129	0.5012	0.50135
01/10/2013	85144	1468	99982	6631	96509	2845	113	8055	94677	85108	81972	6679	22619	6031	77755	2462	18403	1814
03/09/2013	0.4999	0.49999	0.5000	0.49998	0.4998	0.49981	0.4998	0.49980	0.4998	0.5002	0.5015	0.50157	0.5015	0.50163	0.5002	0.50029	0.5002	0.50034
01/08/2013	98017	4581	11883	0716	23179	9745	37035	589	21462	84905	6172	5463	06257	0926	78034	1777	22571	7239
01/08/2013	0.4999	0.49998	0.4999	0.49997	0.4993	0.49929	0.4993	0.49929	0.4992	0.5011	0.5015	0.50158	0.5015	0.50161	0.5011	0.50112	0.5010	0.50114
01/08/2013	86147	4814	91527	9435	00249	892	05614	3555	99584	17701	83189	8521	61669	0041	15035	0367	93515	1887
01/08/2013	0.4999	0.49998	0.4999	0.49998	0.4993	0.49939	0.4993	0.49939	0.4993	0.5009	0.5015	0.50158	0.5015	0.50158	0.5009	0.50097	0.5009	0.50097
01/08/2013	87397	7397	87397	7397	92013	2013	92013	2013	92013	70206	82797	2797	82797	2797	70206	0206	70206	0206

Source: Jamilu Auwalu Adamu (2015), Asian Journal of Management Sciences, 03(08), 2015, 10-18



**SUMMARY RESULT OF TEST OF GOODNESS OF FIT OF GENERAL ELECTRIC CORPORATION STOCK RETURN**





Name of Probability Distribution	Kolmogorov Smirnov Rank	Anderson Darling Rank	Chi-Squared Rank	Average
Error	3	3	2	2.67
Laplace	2	2	3	2.33
<b>Log – Logistic (3P)</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1.0</b>

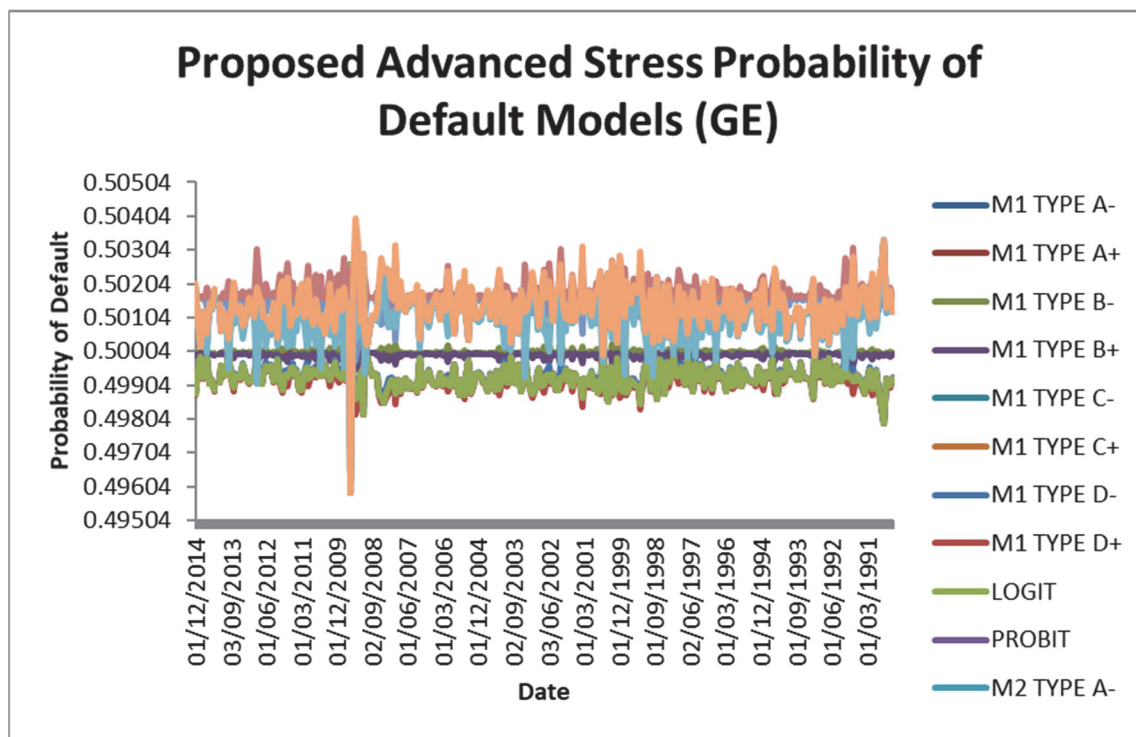
**Results of Stressed SIGMOID and Probit MODELS for General Electric:**

Under General Electric, on the month of September, 2014, the probability of default using the existing Sigmoid is 0.499256894% and that of probit is 0.501185825%, whereas, using the Stressed Sigmoid and Probit models are: 0.499973123%, 0.499971787%, 0.499979127%, 0.499965783%, 0.49925736%, 0.499256228%, 0.499263546%, 0.499250241%, 0.501608698%, 0.501614042%, 0.50158468%, 0.50163806%, 0.501183154%, 0.501188497%, 0.501159135%, and 0.501212515% respectively. All the SIXTEEN (16) proposed banking and economic advanced stressed probability of default models provides good numerical approximations at the time of financial crisis and of course incorporate crisis components as shown in the table below:

Date	M1								LOGIT	T	M2							
	A-	A+	B-	B+	C-	C+	D-	D+			A-	A+	B-	B+	C-	C+	D-	D+
01/12/2014	0.4999	0.49995	0.4999	0.49993	0.4987	0.49877	0.4987	0.49875	0.4987	0.5019	0.5015	0.50159	0.5014	0.50166	0.5019	0.50195	0.5018	0.50203
03/11/2014	0.4999	0.49997	0.4999	0.49996	0.4994	0.49944	0.4994	0.49943	0.4994	0.5008	0.5016	0.50162	0.5015	0.50166	0.5008	0.50089	0.5008	0.50093
01/10/2014	0.4999	0.49999	0.4999	0.49999	0.4998	0.49989	0.4998	0.49989	0.4998	0.5001	0.5016	0.50164	0.5016	0.50165	0.5001	0.50017	0.5001	0.50017
02/09/2014	0.4999	0.49997	0.4999	0.49996	0.4992	0.49925	0.4992	0.49925	0.4992	0.5011	0.5016	0.50161	0.5015	0.50163	0.5011	0.50118	0.5011	0.50121
01/08/2014	0.4999	0.49997	0.4999	0.49995	0.4993	0.49934	0.4993	0.49933	0.4993	0.5010	0.5016	0.50162	0.5015	0.50168	0.5010	0.50104	0.5009	0.50110
01/07/2014	0.5000	0.49998	0.5000	0.49992	0.4998	0.49982	0.4998	0.49975	0.4998	0.5002	0.5016	0.50167	0.5013	0.50193	0.5002	0.50030	0.4999	0.50056
02/06/2014	0.4999	0.49996	0.4999	0.49995	0.4990	0.49909	0.4991	0.49908	0.4990	0.5014	0.5015	0.50160	0.5015	0.50165	0.5014	0.50144	0.5013	0.50148

2014	67813	5361	78837	4338	99809	7366	10794	6381	98587	38445	97254	7064	53156	1162	3354	335	89442	7448
01/05/	0.4999	0.49996	0.4999	0.49996	0.4991	0.49918	0.4991	0.49918	0.4991	0.5012	0.5016	0.50160	0.5016	0.50161	0.5012	0.50129	0.5012	0.50130
2014	69937	9771	7068	9029	86806	6641	87546	59	86723	97801	06589	725	03617	0222	9747	8131	94498	1104
01/04/	0.4999	0.49995	0.4999	0.49993	0.4988	0.49887	0.4989	0.49885	0.4988	0.5017	0.5015	0.50159	0.5014	0.50168	0.5017	0.50179	0.5016	0.50188
2014	60837	6097	82143	4793	8188	7161	03094	5949	7952	88024	78015	6977	92781	2211	78543	7505	93309	2739
03/03/																		
2014																		
03/02/	0.4999	0.49996	0.4999	0.49996	0.4991	0.49910	0.4991	0.49910	0.4991	0.5014	0.5016	0.50160	0.5015	0.50161	0.5014	0.50143	0.5014	0.50144
2014	67051	6509	69488	4073	04064	3524	06492	1095	03794	30137	0042	2588	90672	2336	29052	1221	19305	0969
02/01/	0.4999	0.49998	0.4999	0.49998	0.4996	0.49961	0.4996	0.49961	0.4996	0.5006	0.5016	0.50163	0.5016	0.50163	0.5006	0.50062	0.5006	0.50062
2014	85569	5569	85569	5569	10685	0685	10685	0685	10685	21257	34947	4947	34947	4947	21257	1257	21257	1257
02/12/	0.4999	0.49997	0.5000	0.49993	0.4993	0.49933	0.4993	0.49929	0.4993	0.5010	0.5015	0.50163	0.5014	0.50177	0.5010	0.50107	0.5008	0.50121
2013	79433	1412	15489	5362	4094	2939	76904	6981	3694	58091	98132	0218	53905	4445	42048	4134	97821	8361
01/11/	0.4999	0.49997	0.4999	0.49996	0.4993	0.49934	0.4993	0.49933	0.4993	0.5010	0.5016	0.50161	0.5015	0.50164	0.5010	0.50104	0.5010	0.50106
2013	76374	5112	82049	9437	46222	4963	51883	9302	45592	44283	13093	8143	90392	0845	41758	6808	19057	9509
01/10/	0.5000	0.49998	0.5001	0.49988	0.5000	0.50000	0.5001	0.49990	0.5000	0.4999	0.5016	0.50169	0.5012	0.50210	0.4999	0.50001	0.4995	0.50042
2013	11844	9343	1301	8227	27252	475	28424	3627	16001	74467	08075	8077	0351	2642	29466	9468	24901	4033
03/09/	0.4999	0.49997	0.4999	0.49995	0.4992	0.49923	0.4992	0.49922	0.4992	0.5012	0.5016	0.50161	0.5015	0.50167	0.5012	0.50122	0.5011	0.50128
2013	73497	0109	8873	4878	40994	7616	56182	2429	39305	13893	02183	5739	4125	6673	07115	0671	46181	1604
01/08/	0.4999	0.49995	0.5000	0.49985	0.4989	0.49896	0.4990	0.49887	0.4989	0.5016	0.5015	0.50163	0.5011	0.50203	0.5015	0.50167	0.5011	0.50206
2013	73131	125	71513	2916	9087	9074	88865	1126	79972	27727	52304	9842	58814	3332	83958	1497	90468	4987

Source: Jamilu Auwalu Adamu (2015), Asian Journal of Management Sciences, 03(08), 2015, 10-18



## Appendix C

### Advanced Optimized Activation Functions

The Ranking of the Advanced Optimized Activation Functions was obtained using Jameel's Criterion (second version). The author considered Eleven (11) out of Fifty (50) World's Biggest Public Companies by FORBES as of 2015 Ranking Regardless of the platform in which they are listed, Time Series (Short or Long Terms), Old or Recently Listed Companies. This can be seen below:

Advanced Optimized Activation Functions	Frequency regardless of the Time Series	Order of Hierarchy
Log – Logistic (3P)	50	1 <sup>st</sup>
Cauchy	33	2 <sup>nd</sup>
Pearson 5 (3P)	25	3 <sup>rd</sup>
Burr (4P)	19	4 <sup>th</sup>
Fatigue Life (3P)	12	5 <sup>th</sup>
Inv. Gaussian (3P)	10	6 <sup>th</sup>
Dagum (4P)	9	7 <sup>th</sup>
Lognormal (3P)	8	8 <sup>th</sup>
Normal, Pearson 6 (4P), Gen. Gamma (4P), Error, Laplace, Hypersecant, Gamma, Johnson SU, Beta, and Uniform	Frequency less than or equal to 8	9 <sup>th</sup> and below it

Source: Jamilu Auwalu Adamu (2015), Asian Journal of Management Sciences, 03(12), 2015, 01-07

### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).