Adaptive Risk Analysis and Management (ARAM) for the Lightning Strike on Power Station Systems based on Machine Learning Modeling

Muhammad A. Sulaiman
Faculty of Computer Science & Information Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia

Ja'afar Zangina Sulaiman
Electrical and Electronics Engineering Faculty of Engineering Mevlana University Konya Yeni Istanbul Cad. 42003 Selcuklu Konya, Turkey

ABSTRACT
The effects of lightning strike to transmission and distribution systems are numerous and unfavorable. Just imagine the cost and havoc if a giant telecommunications company is shut down for an hour or day as a result of devices damage or a petrochemical plant catches fires due to lightning strike. Hence the needs to protect power apparatus from overvoltage surge are imperative. In this study an adaptive risk analysis & management (ARAM) based on artificial neural networks is proposed to analyze and proactively control the overvoltage at power substation due to lightning strike.

General Terms
Machine learning classification problem and System design.

Keywords
Machine learning modeling, Lightning strike, Risk management, Power & Control Systems.

1. INTRODUCTION
Lightning occurs randomly in nature and has adverse effects on overhead line faults, it has been estimated that 5% to 10% of overhead line faults course by lightening resulted to permanent damage to power system equipment [1]. Lightning surge voltages that arrive at a substation, traveling along a transmission system are caused by a lightning stroke terminating either on a tower, shield wire, or a phase conductor. The shape and magnitude of these voltages are functions of many parameters including polarity, shape, magnitude of the lightning stroke current, the footing impedance, line surge impedance and the tower surge impedance. For shielded transmission system, the induced surge voltages due to shielding failure are typically to a greater extent severe than those stimulated by a backflash. That is, they have greater crest voltage and greater steepness [2]

Having insight on the level of stroke (low, medium or high) and localizing the stroke to specific location of station and/or substation will help to control the risk of damaging power system equipment. This paper proposed ARAM based on artificial neural networks to accurately predict the level of stroke and mapped the stroke to a specific substation. This will make it possible to proactively isolate the power

equipment from damage. That is, by serving as an additional protection level which aimed at enhances lightning shielding performance and plays a crucial role in the lightning protection of transmission systems.

The rest of the paper is organized as follows: section 2 is the background of the study and it discusses the basic studies of lightning, effects of the lightening and lightning location system. Basic Artificial neural networks hypothesis is

explained in section 3, while section 4 presents the proposed adaptive risk analysis & management (ARAM) design and description. Section 5 is the experimental studies and it explains adopted experimental procedures, results and discussion. Finally, conclusion is presented in section 6.

2. BACKGROUND OF THE STUDY
Lightning is a physical phenomenon of random nature which generated fast transients in power systems. It is the major cause of power system outage or equipment damage by indirect strokes [3] and direct strokes [4]. Overhead transmission and distribution lines are often exposed to lightning overvoltage, whose waveforms vary widely. The transient overvoltage due to lightning are electromagnetically induced in power lines after a stroke within the vicinity of the lines.

Lightning surge voltages that arrive at a substation, traveling along a transmission system, are caused by a lightning stroke terminating either on a tower, shield wire, or a phase conductor. The shape and magnitude of these voltages are functions of the polarity, shape, and magnitude of the lightning stroke current, the footing impedance, line surge impedance, and the tower surge impedance. For lines that are effectively shielded, the induced surge voltages due to shielding failure are typically to a greater extent severe than those stimulated by a backflash. In other words, they are faced with fast front surge [2]. The voltage along the shielding wire varies significantly as a function of front time, peak current and position along the wire. Vast numbers of lightning shielding failures in a UHV and EHV transmission system result in fault [5].

Lightning strikes and associated effects are real serious threat. They are the cause of fires, injury or loss of life, damage & destruction of property, significant downtown and outage-related revenue loose due to equipment damage [6].

2.1 Basic study of lightning
Cloud flashes are the most frequent flashes in their classes. This group of flashes occurs whenever there is a severe storm. However, it shrinks spatial differences in charge among clouds. Generally cloud flashes within the region of primary negative charge at the height of 4-8 km in the cloud, propagates toward a region of predominately positive charge at a height of 8-12 km [7]. Cloud to Ground flash is the other category of flashes which is initiated by electrical breakdown in the cloud, and a negative charge from the electrified cloud spread-out to link earth as shown in Figure 1.

Subsequent flashes may result from lightning flash if cloud to cloud (CC) discharges originated from the depletion of charge in the main cloud cluster linking the adjacent charge cluster gather sufficient charges. Some of these charges are

channelled in a sequence of individual return strokes that feature a nominal duration of tens of microseconds, peak current ranging from few kilo amperes to more than 200 kA, and are spaced by 20–100 ms [7]. Since the discharge route from the main flash is quiet ionized, it assumes little charge to initiate a flash. In comparison with the primary stroke, the pre-ionized channel effects in subsequent strokes with about one-third the current and shorter rise time [8]

Both classes of flashes are considered to radiate electromagnetic energy in terms of high-frequency over a wide range of frequency spectrum. This makes it possible for the lightning sensing device to process electromagnetic signal in traditional radio-frequency ranges that relate to common signal-processing bands.
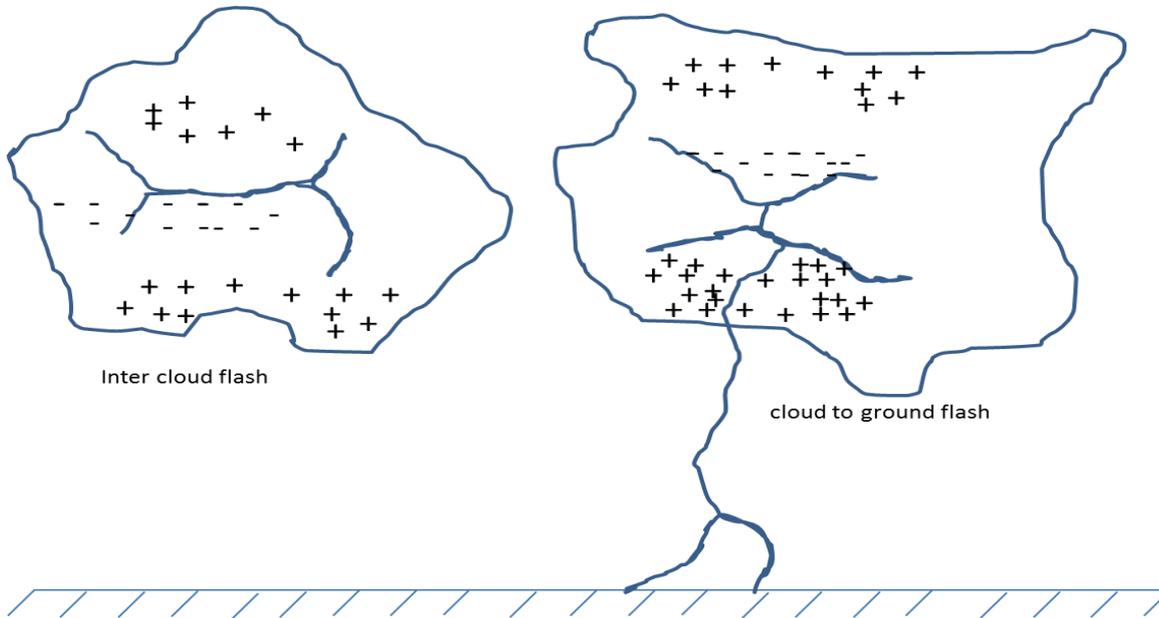

Fig 1: Classes of flashes

2.1 Mechanism of Lightning Location System

Emerging microelectronic technologies are offering advanced techniques to observe, examine and mitigate the effect of lightning on power systems. Cloud-to-Ground lightning (CG) is one of the principal causes of outages and transient overvoltage faults in electric power transmission and distribution systems in lightning-prone areas. The electromagnetic field in the VLF, LF and VHF frequency ranges radiated by lightning is measured by Lightning Location systems (LLS). These systems employed sensor which are usually separated by 50-400 km to measure lightning radiated magnetic field [9]. The CG discharges are located by using various techniques like Time-of-Arrival (TOA), Magnetic Direction Finding (MDF), and combination of both (MDF+TOA). Locating lightning using TOA technique requires precise synchronization of the sensors, which is available in GPS satellite. This made it possible to detect differences in the arrival time at various distances from the striking point as the radiated field from Lightning propagates in all directions with the speed of light.

The MDF systems determine the direction with the aid of crossed loop antennas once the upward propagating return stroke has attained a height of a few hundred meters [10]. So, the resulting direction vector indicates closely as possible to the position where the CG stroke linked to ground. The combined MDF+TOA technique provides azimuth data while absolute arrival time renders range information. This technique utilized location algorithm in conjunction with timing and angle information provided by the sensor to grants an optimized estimate of three variable parameters; discharge time, longitude and latitude [7].

2.2 Ground base Lightning Location System

Correctly identification and classification of each emitted lightning electromagnetic wave is not an easy task. This task is accomplished either by central processor (CP) in which acquired sensor data is examines base on a set of information used in locating the stroke or at the sensor level through "waveform criteria" established on a set of stroke features of CG flashes [11].

Uman et al. (1975) relates peak current *Ip* to the peak field *Ep* based on transmission line model (TL) as shown in Equation 1.

$$I_p = \frac{2\pi \cdot \varepsilon_0 \cdot c^2 \cdot D}{v} \cdot E_p \qquad (1)$$

Where;

D: horizontal distance between the lightning channel and the observation point,

V: return stroke speed

C: speed of light

Assuming a reference distance D of 100 km and 1.108 m/s return stroke velocity, Equation 1 becomes the linear relation used to infer lightning peak currents from measured peak fields as shown in equation 2.

$$I_p(kA) = 5.E_p(V/m) \qquad (2)$$

Lightning flashes over a very large area covered by the transmission and distribution Electrical Power System infrastructure can be monitored by a few numbered of installed ground base lightning detection systems (GLLS). In Figure 2, both the transmission system and the GLLS are on the same ground.

The lightning peak current, strike locations and discharge time information is determine by the GLLS. Due to the tremendous amount of data that can be gathered by this system, it remains a promising source of standard and customized lightning data format outputs to implement this work.
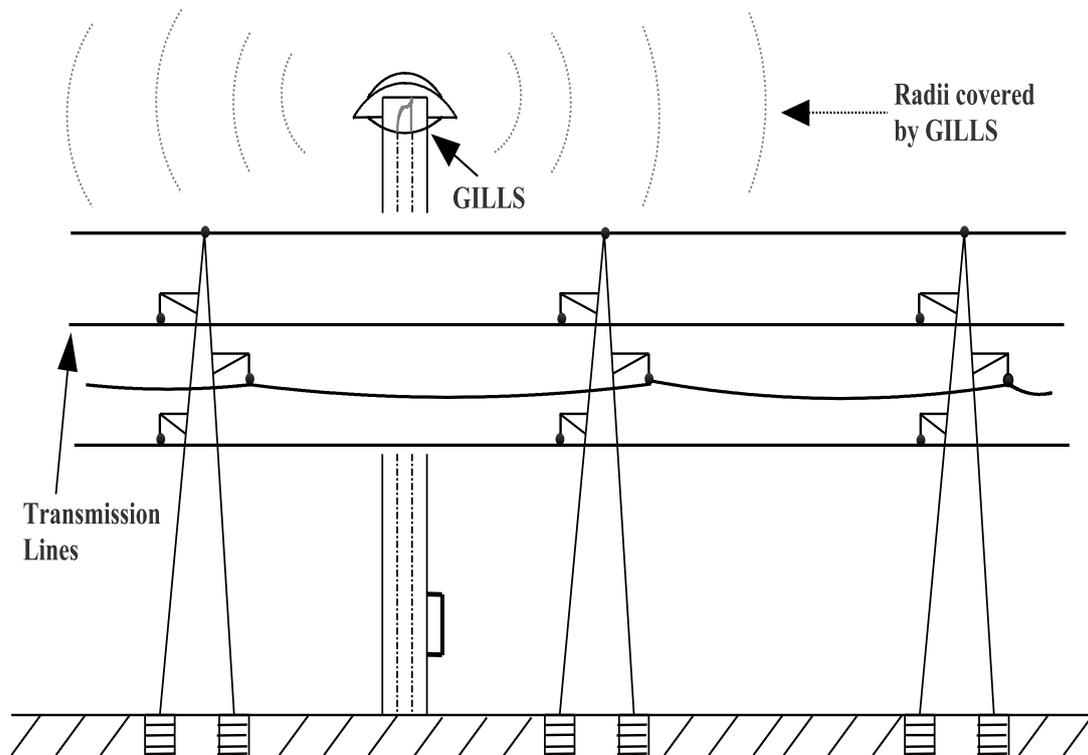
Fig 2: Installed ground based lightning detection systems

3. HYPOTHESIS OF THE ARTIFICIAL NEURAL NETWORKS (ANNs)

The artificial neural network (ANN) derives its computing ability through its massively parallel distributed structured and through its power to learn and therefore generalize. These two capabilities make it possible for ANN to find good approximation solution for intractable complex problems [12]. Among the features of ANN that made it attractive classification or prediction model in many fields of research are:

(1) Nonlinearity: It has the ability to solve complex nonlinear problems.

(2) Input-Output Mapping: The network learns from examples based on popular paradigm known as supervised learning by constructing an input-output mapping of the problem at hand.

(3) Adaptively: this is a built-in power of neural networks to adapt their weights to changes in the surrounding environment. ANN trained to operate in a specific environment can be easily retrained to deal with minor changes in the operating environmental conditions [12].

ANNs were developed by simulating human neurons or network of neurons in the brain. Neurons are cells with cell body and dendrites which are the input wires connected to the cell body and receive signals from body receptors. In addition to these two, neurons have output wire known as axon which sends signals or informative messages to other neurons. Thus, at primitive level a neuron is a computational unit that receives a number of inputs (electric pulses) through its input wires (dendrites), does some computation in the cell body and sends output through its axon to other neurons in the brain.
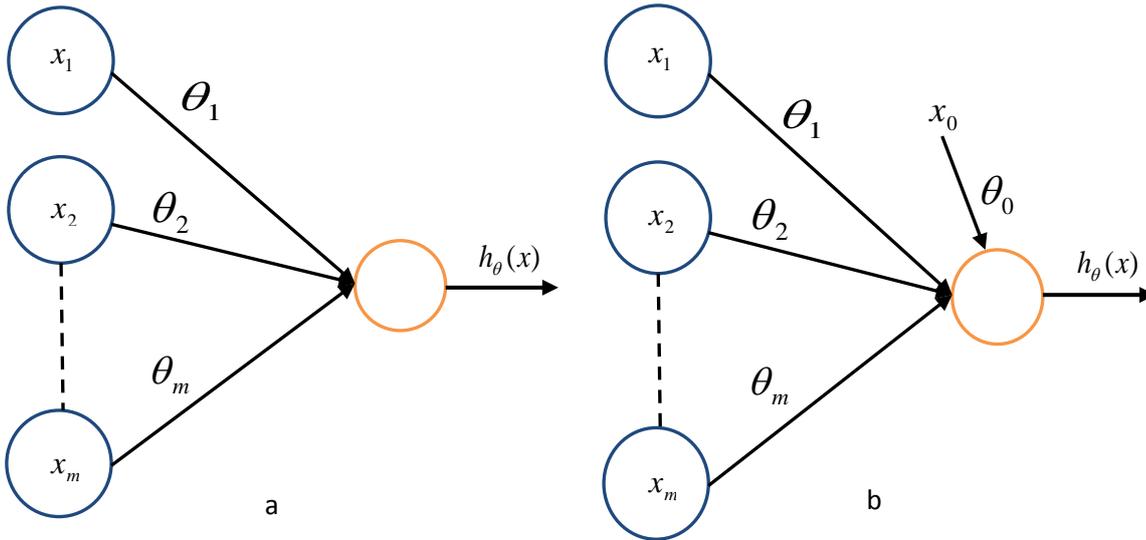
Fig 3: Neuron model: (a) without bias, (b) with bias input x_0 and weight θ_0 explicitly

Figure 3 depicted ANN implementation of a simple model of a neuron; the empty circle (in orange) plays a role analogous to the body of neuron, each arrow linking an input x_i to the neuron, has attached a parameter or weight θ_i, the diagram represent the computation of a sigmoid fed with the dot product of x and θ in the form of:

$$h_\theta(x) \frac{1}{1+e^{-\theta^T x}} \qquad (3)$$

Where x and θ are parameter vectors $x = \begin{bmatrix} x_0 \\ x_1 \\ . \\ . \\ x_m \end{bmatrix}$ and

$\theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ . \\ . \\ \theta_m \end{bmatrix}$. Mathematically, ANN implementation of a

simple model of a neuron figure3 can be represented as:

$$U_k = \sum_{j=1}^{m} \theta_{kj} x_j \qquad (4)$$

Equation (4) is refers to weighted sum of input signals without bias.

$$V_k = (U_k + b_k) \qquad (5)$$

and $\quad y_k = h(U_k + b_k) \qquad (6)$

where V_k is a weighted sum of input signals plus bias expressed as:

$$V_k = \sum_{j=0}^{m} \theta_{kj} x_j \qquad (7)$$

$$b_k = x_0 \theta_0 \qquad (8)$$

From equation (5) V_k can be seen clearly as a linear function, equivalent to equation of line $(ax + c)$ which is expressed in eqn. (7). And y_k from equation (6) is the output of classification based on activation function which says when

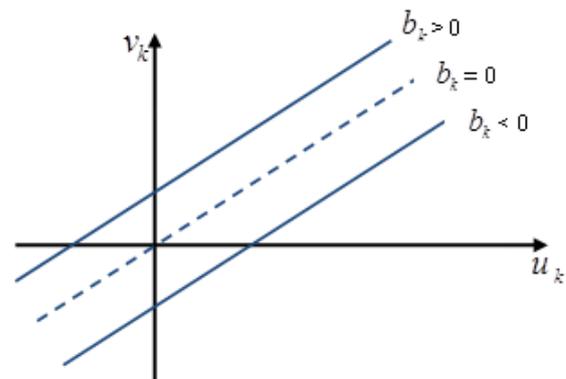$V_k \geq 0$ ' $y_k = 1$ or when $V_k < 0$ $\quad y_k = 0$

Fig 4: Affine transformation resulted by the presence of a bias

Note that in figure 4: V_k is a local induced field and b_k is an intercept at V_k. From equation 5, $V_k = b_k$ at $U_k = 0$ Thus $b_k > 0, V_k$ is intercepted at positive value, $b_k < 0$, V_k is intercepted at negative value and $b_k = 0, V_k$ is intercepted at origin.

Having shown both graphical and mathematical models of neuron, it is easy to see that Neural Network is a collection of different artificial neurons working together to form a predictor or classifier. Perhaps no other better way to prove the modeling capability of ANNs than to describe and prove the perceptron convergence theorem which was the first algorithmically description of ANNs as invented by Rosenblatt in 1958. The theorem stated that:

"If there is a set of weights that correctly classify the (linearly separable) training patterns, then the learning algorithm will find one such weight set, w* in a finite number of iterations."

Suppose the input vectors for perceptron (built around nonlinear neuron) originated from linearly separable classes. Let X_1 be the training subspace of training vectors $x_1(1), x_1(2), ...,$ and belongs to class (class 1 ε_1). And let X_2 be the training subspace of training vector $x_2(1), x_s(2), ...,$ and belongs to class (class 2). ε_2

So, given a set of training vectors X_1 & X_2, the training process involve adjusting the weight θ vector so that ε_1 and ε_2 are linearly separable stated as follows:

Let $\exists$ a weight vector θ such that:

$$\theta^T x > 0 \text{, for } x \in \varepsilon_1$$
$$\theta^T x \le 0 \text{, for } x \in \varepsilon_2 \qquad (9)$$

In case of adapting the weight vector in perceptron during training:

Case 1: $\theta(n+1) = \theta(n)$ if $\theta^T x(n) > 0$ & $x(n) \in \varepsilon_1$. And $\theta(n+1) = \theta(n)$; if $\theta^T x(n) \le 0$ & $x(n) \in \varepsilon_2$ are the cases of correct classification, in which there are no need for weight update.

Case 2: $\theta(n+1) = \theta(n) - \delta$, if $\theta^T x(n) > 0$ & $x(n) \in \varepsilon_2$. And $\theta(n+1) = \theta(n) + \delta$, if $\theta^T x(n) \le 0$ & $x(n) \in \varepsilon_1$ are the cases of misclassification, in which the weight need to be updated. On these bases, perceptron convergence theorem is proved as follows:

Assuming that $\delta = \eta \vec{x}(n)$; for n =1, 2... n, $\theta(0) = 0$ and $\eta = 1$ is the learning rate. Also, the input vector belongs to two linearly separable.

Using the 2nd line of case 2 and update the weight of misclassification in which $\theta^T x(n) \le 0$, and $x(n) \in \varepsilon_1$.

$$\theta(n+1) = \theta(n) + x(n) \qquad (10)$$

From equation (10) it can be deduced that:

$$\theta(1) = \theta(0) + x(0) = x(0); \quad n = 0.$$

$$\theta(2) = \theta(1) + x(1) = x(1); \quad n = 1.$$

$$\theta(3) = \theta(2) + x(2) = x(1) + x(2); \quad n = 2.$$

$$\theta(n+1) = x(1) + x(2) + ... + x(n) \qquad (11)$$

For n = n+1. The objective is to find weight θ_s such that: $\theta^T x(n) > 0$ & $x(n) \in \varepsilon_1$. For a fixed solution, θ_s let:

$$\alpha = \min_{x(n) \in \varepsilon_1} \theta_s^T x(n) \qquad (12)$$

Where α is positive. Multiplying both sides of equation (11) with θ_s^T to arrive at equation (13):

$$\theta_s^T \theta(n+1) = \theta_s^T x(1) + \theta_s^T x(2) + ... + \theta_s^T x(n) \qquad (13)$$

It is easy to see that equation (14) holds with reference to equation (12).

$$\theta_s^T \theta(n+1) \ge n\alpha \qquad (14)$$

According to Cauchy-Schwarz inequality, given two vectors θ_s and $\theta(n+1)$:

$$\|\theta_s\|^2 * \|\theta(n+1)\| \ge \left[\theta_s^T \theta(n+1)\right]^2 \qquad (15)$$

Where $\|.\|$ is Euclidean norms and inner product $\theta_s^T \theta(n+1)$ is scalar qualities. Equation (16) relates equation (14) and equation (15):

$$\|\theta_s\|^2 * \|\theta(n+1)\| \ge n^2 \alpha^2 \qquad (16)$$

It implies that:

$$\|\theta(n+1)\|^2 \ge \frac{n^2 \alpha^2}{\|\theta_s\|^2} \qquad (17)$$

Following another development route by rewriting equation (10) in terms of k as:

$$\theta(k+1) = \theta_{(k)} + x(k) \qquad (18)$$

For k = 1, 2... n. and $x(k) \in \varepsilon_1$. Taking squared Euclidian norms of both sides of equation (18):

$$\|\theta(k+1)\|^2 = \|\theta(k)\|^2 + \|x(k)\|^2 + 2\theta(k)^T x(k) \qquad (19)$$

But $\theta(k)^T x(k) \le 0$ based on condition of the second line of equation (9).Thus, equation (19) is deduced to:

$$\|\theta(k+1)\|^2 = \|\theta(k)\|^2 + \|x(k)\|^2 \text{, or equivalently}$$

$$\|\theta(k+1)\|^2 - \|\theta(k)\|^2 = \|x(k)\|^2 \qquad (20)$$

Given the initial condition $\theta(0) = 0$ and recall that $\theta(n+1) = \sum_{k=1}^{n} x(k)$ the following inequality is obtained:

$$\|\theta(n+1)\|^2 \le \sum_{k=1}^{n} \|x(k)\|^2 \qquad (21)$$

Now, let there be $\beta = \max_{x(k) \in \varepsilon_1} \|x(k)\|^2 \qquad (22)$

For sufficiently large value of n:

$$\|\theta(n+1)\|^2 \le n\beta \qquad (23)$$

However, equation (23) is in conflict with the eqn. (17) for a large value of n. So, to reconcile the two equations, n_{max} is considered, under which:

$$\frac{n_{max}^2 \alpha^2}{\|\theta_s\|^2} \le \|\theta(n+1)\|^2 \le n_{max}\beta \tag{24}$$

From equation (24), it is shown that the solution vector θ_s exists and it is guaranteed to be obtained after n_{max} iterations:

$$n_{max} = \frac{\|\theta_s\|^2 \beta}{\alpha^2}$$

3.1 ANN Modeling For Localizing Lightning Strike

The basic idea of the proposed method is to have an adaptive risk analysis & management system that will proactively protect power apparatus from overvoltage surge due to lightning strike. And for this reason ANN is used to model or predict: (1) the intensity of lightening (that is high, medium or low) and (2) to map the lightning strike to a specific region/area of station, so that appropriate action could be taken (e.g. to break the circuit).

However, in order to achieve this aim an archive of data for regions with lightning density and for a long period of time was sought for. For this reason lightening data were acquired from the national Lightning Detection Network (NLDN) which covers the United State and has been on operation for 30 years. The NLDN remained the most reliable source of lightning data. Data are collected at very high spatial accuracy and high detection efficiency.

This study used lightning flash dataset collected for 5 years and covered a region of approximately 25km by 25 km square area within the latitude/longitude bounds 37.15N – 37.375N and 95.95W – 95.667W. The data consist of the following fields (1) Date (in UTC), (2) Time (in UTC), (3) Latitude (in decimal degrees), (4) longitude (in decimal degrees), (5) Polarity and peak current (KA). Based on the description of this region, the satellite view of the area covered is captured and divided into smaller regional units based on the latitude and longitude bounds as in figure5. And it is assumed that these smaller regional unit areas consist of different type of power system equipment.

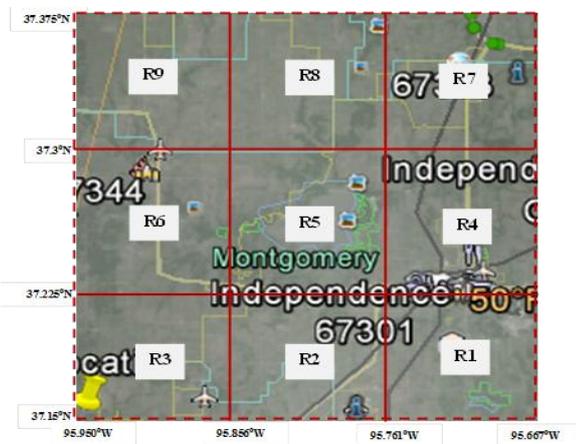**Fig. 5: Division of the study area into smaller regional units**

R1 to R9 refer to sub-region 1 to sub-region 9, each of which is within decimal values of latitude/longitude bounds. For example R1 is between 37.15N – 37.225N, 95.761W – 95.667W. Through this predefined division, localized labels are obtained for the case study locations. Also, using domain expert knowledge the peak current range of values are labeled into three classes namely, low, medium and high. Both the location labels and the lightning strike intensity labels combined to serves as targeted classes.

4. DESCRIPTION OF THE DESIGN OF ARAM

The proposed study considered that there is various power system installations at the case study substation location and these installed equipment have different resistance to voltage surge and could be found in different regional units. That is, it was assumed that there are documented inventory of all the installed equipment in the location handled by control system. This control system is linked to ARAM in one hand and to Circuit breaker system in the other hands. The ARAM receives relevant input from device for capturing the strike information (like GLLS) and produces strike level and the targeted location (sub-region), based on the output produced by ARAM a trigger will be initiated to alert control system for action, to remove fault. Table 1 is the outline of the steps involved, and the figure 6 provides the design of the proposed system.

Table 1. The ARAM steps for power substation system

1. Lightning strike occurrence
2. Device for capturing the strike information

   (like GLLs) captured the event
3. ARAM predicts and triggers control system.
4. Control system takes action to remove the fault.
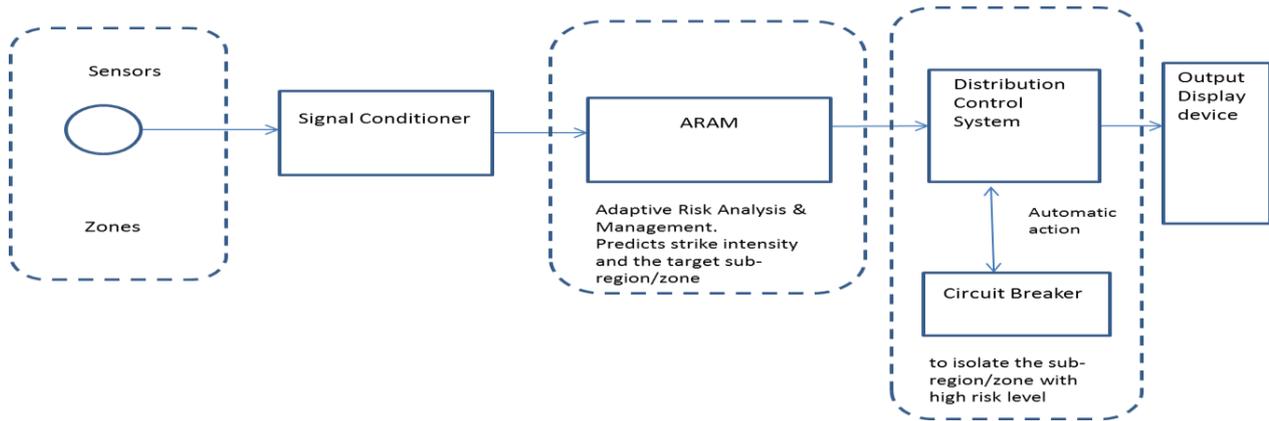5. Fault removal (such as circuit breaker)

Fig. 6: Design of the proposed method

5. EXPERIMENTAL STUDIES

This section presents various experiment carried out to ensure that the predicting model work well and corresponding results achieved indicates that the proposed method is promising.

The dataset for this study was acquired from NLDN as stated earlier on, the dataset consist of 5 features and 31,703 data points created from dataset for 2008 through 2012. However, for this experiment, only 3 features were considered as an input for the experiment due to their direct correlation with expected target attributes. This dataset is divided into training set, validation set and test set in the ratio of 3:1:1 respectively.

5.1 Experimental Setup

Based on the input dataset two attributes were generated, the location attribute was generated from location labels as explained in section 3.1. Also, strike intensity attribute for the data points from the lightening intensity label from section 3.1

were generated. Finally, the two attributes were combined into 27 multiclass labels for the neural network architecture.

The Neural Network implementation was evaluated to ensure that the learning process during training stage descent as expected. Figure 7 is generated to ensure that the learning curve algorithm is working fine. The cross validation errors decreases for different subset of the training set sizes. High variance problem can be observed in figure 7 as cross validation sets errors remains relatively high as the number of examples increase, while the training errors are relatively very small, as such is likely that the model will not generalized well. To resolve this over fitting problem, regularization parameters were included to the learning algorithm and the optimal regularization parameter (using cross-validation method) suitable for both training and cross validation sets is obtained.
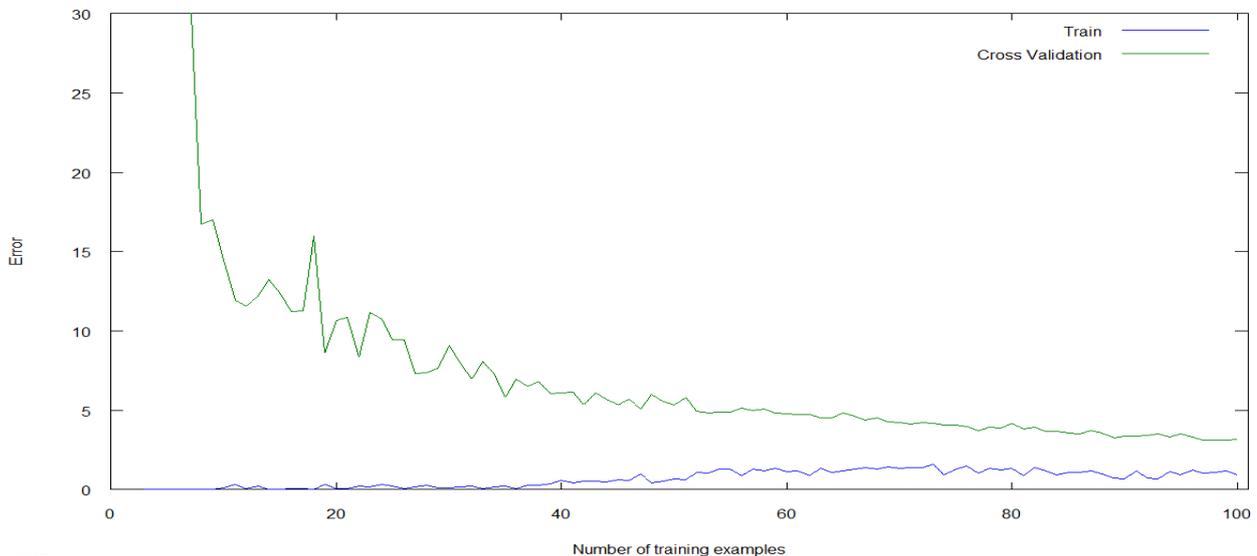

Fig. 7: Neural networks learning curve with zero regularization parameter
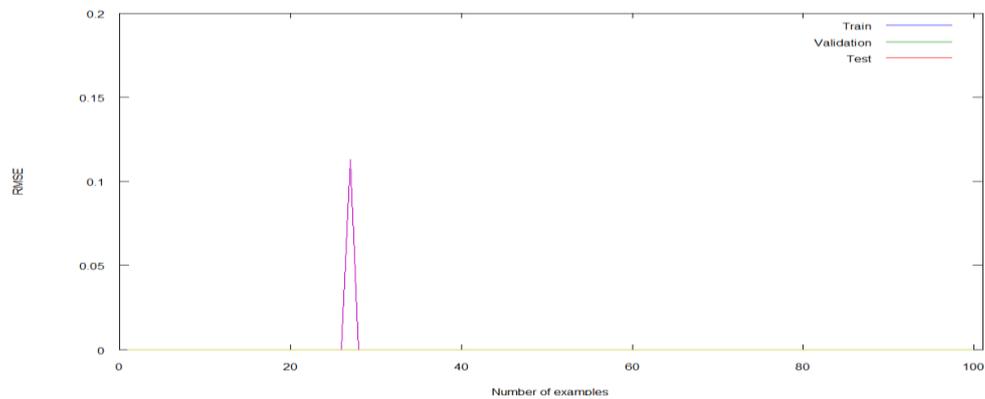
Fig. 8: Neural networks performance evaluation

5.2 Experimental Results and Discussion

After training the model, the performance of the resulted classifier is evaluated on the test set. Prediction accuracy of 99.668 is recorded with 0.302 RMSE. Also, the RMSE of each example in training, validation and test set samples of first 100 examples is presented in figure 8. Out of 100 first examples samples for each of the three subsets of the dataset, only one test example (unknown sample) was wrongly predicted, the remaining 99 unknown samples from the test set were predicted correctly.

6. CONCLUSION

In this study an adaptive risk analysis & management (ARAM) for the lightning strike on power substation system is proposed. The proposed design was illustrated hypothetically. Based on the real life lightning strike dataset obtained from NLDN the adaptive unit of the ARAM was implemented and the result obtained has shown that the choice of neural network is suitable for this study.

The key contribution of this study is to proactively reduce the effects of the lightning strike on power substation by isolating targeted substation equipment from damage due to strike. In addition, this study will serves as an additional protection level that aimed at enhance lightning shielding performance and will play a crucial role in the lightning protection of transmission systems.

Future work will focus on the simulation of the remaining part of the proposed ARAM as presented in this study, in particular the control system and fault removal sections.

7. ACKNOWLEDGMENTS

8. REFERENCES

[1] Kamran Hafeez, Sheheryar Khan. "Risk Management Analysis with the help of Lightning Strike mapping around 500 k-v Grid Station using Artificial Intelligence Technique.", *Robotics and Artificial Intelligence (ICRAI), 2012 International Conference on IEEE.*

[2] Technical Council of the IEEE Power Engineering Society, ", IEEE Guide for the Application of Insulation Coordination", *IEEE Std 1313.2-1999.*

[3] A. Borghetti, A. Morched, F. Napolitano, C. A. Nucci, and M. Paolone, "Lightning-Induced Overvoltages Transferred Through Distribution Power Transformers*" IEEE Transactions on Power Delivery, vol. 24, no. 1, pp. 360-372, Jan. 2009.*

[4]] S. Sekioke, K. Aiba, T. Miyazaki,nand S. Okabe, "Lightning Overvoltages in Low-Voltage Circuit for Various Lightning Striking Points*", IEEE Transactions on Power Delivery, vol. 25, no. 4, pp. 3095-3104, Oct. 2010.*

[5] J. Li, Q. Yang, W. Sima, C. Sun, T. Yuan, "A New Estimation Model of the Lightning Shielding Performance of Transmission Lines Using a Fractal Approach*", IEEE Transactions on Dielectrics and Electrical Insulation Vol. 18, No. 5; October 2011.*

[6] Roy B. Carpenter, Jr., Mark N. Drabkin, Ph.D. "Lightning Strike Protection", *Lightning Eliminators & Consultants, Inc. 6687 Arapahoe Road, Boulder, Colorado 80303 USA, 800-521-6101 or 303-447-2828.*

[7] K. L. Cummins, M. J. Murphy, "An Overview of Lightning Locating Systems: History, Techniques, and Data Uses, With an In-Depth Look at the U.S. NLDN", *IEEE Trans. on Electromagnet. Comp., Vol. 51, no. 3, August 2009.*

[8] J. C. Das "Transient in Electrical Systems" ISBN: 978-0-07-162603-3; C. 2010.

[9] K. L. Cummins, M. J. Murphy, "Overview of Lightning Detection in the VLF, LF, and VHF Frequency Ranges", *ILDC, Tucson, Arizona 2000.*

[10] E. P. Krider, R. C. Noggle, and M. A. Uman, "A gated wideband magnetic direction-finder for lightning return strokes", *J. Appl. Meteor. 1976, (15), pp. 301–306.*

[11] E. P. Krider, R. C. Noggle, A. E. Pifer and D. L. Vance, "Lightning Direction-Finding System for Forest Fire Detection" vol. 61, No. 9 AMS 1980.

[12] Simon Haykin, McMaster University Hamilton, Ontario, Canada, "Neural Networks and Learning Machines." Pearson International Edition, Third Edition.