International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 3 | Issue 8 | ISSN : 2456-3307
DOI : https://doi.org/10.32628/CSEIT183891

Intrusion Detection Using Back Propagation Neural Network and Quick Reduct Algorithms

[1]S. Vijaya Rani, [2]Dr. G. N. K Suresh Babu
[1]Research Scholar, Bharathiar University, Tamil Nadu, India
[2]Associate Professor, Department of MCA, Acharya Institute of Technology, Bangalore, Karnataka, India

ABSTRACT

It is a big challenge to safeguard a network and data due to various network threats and attacks in a network system. Intrusion detection system is an effective technique to negotiate the issues of network security by utilizing various network classifiers. It detects malicious attacks. The data sets available in the study of intrusion detection system were DARPA, KDD 1999 cup, NSL_KDD, DEFCON, ISCX-UNB, KDD 1999 cup data set is the best and old data set for research purpose on intrusion detection. The data is preprocessed, normalized and trained by BPN algorithm. Further the normalized data is discretized using Entropy discretization and feature selection carried out by quick reduct methods. After feature selection, the concerned feature from normalized data is processed through BPN for better accuracy and efficiency of the system.

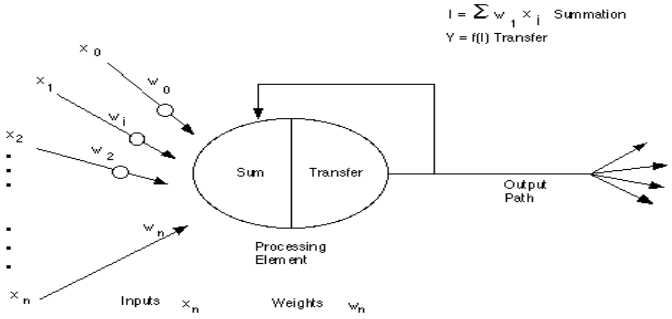
Keywords : IDS, Normalization, Neural Network, Threats, BPN, QR, ED

I. INTRODUCTION

The hackers, bad heads and fraudulent are generating new threats and attacks day by day which are beyond the capacity of the conventional IDS. Normally IDS examines every inbound and outbound traffic in a network and recognizes the patterns, protocols, signatures, anomalies and network policies from its stored database to detect malicious activities. MAC, IP spoofing are the smart activities being carried out by the hackers to intrude into the target network.. It is very difficult to identify threats and attacks whose signatures are not in its database. By implementing proper learning process, architecture and data mining algorithm, new attacks and threats can be identified meticulously.[12]

Neural Network:- A neural network[1] is a system that can simulate with the fundamental operation of human brain. Brains store information as patterns. Some of these patterns are very complicated and allow us the ability to recognize individual faces from many different angles.[5] This process of storing information as patterns, utilizing those patterns, and then solving problems encompasses a new field in computing. This field, as mentioned before, does not utilize traditional programming but involves the creation of massively parallel networks and the training of those networks to solve specific problems. This field also utilizes words very different from traditional computing, words like behave, react, self-organize, learn, generalize, and forget.. The schematic representation of neuron is as under:

The neuron consists of a number of inputs added by an adder module and further activated by activation function as desired by the user to give the desired output. The architecture selected in this concern is BPN.[13]



II. RELATED WORK

Ajay Shiv Sharma et al [17] proposed a techniques for Gene selection for tumour classification using resilient back propagation Neural Network. Classification of tumour into its subtypes is a basic problem nowadays in tumour diagnosis and treatment. The right classification of tumour into its subtypes prompts to the suitable treatment. Conventional strategies classify cancers depend on morphological appearance that prompts to misclassification on account of the comparative appearance of sub-sorts of cancer. To conquer this issue of misclassification, gene expression profiles of genes are utilized. But microarray data contains thousands of genes so the technique is need to diminish the data dimensionality. T-test scoring scheme is utilized for selecting important genes and Artificial NN is connected for the classification process. Resilient back propagation is utilized as training algorithm. The developed method classifies the information with more accuracy.

Alexandre et al [18] proposes machine learning algorithm for pattern classification refer high-dimensional vectors (perceptions) to classes in light of speculation from illustrations. Manufactured Neural Networks as of now accomplish state-of-the-art solution in this process. Although such networks are ordinarily utilized as black-boxes, they are additionally accepted to learn (high-dimensional) more elevated amount representations of the first perception. Envisioning the connections between scholarly representations of perceptions, and imagining the connections between artificial neurons. Through investigations led in three conventional image classification benchmark datasets, they demonstrate how representation can give profoundly important feedback for network designers. For example, our revelations in one of these datasets (SVHN) incorporate the nearness of interpretable clusters of learned representations, and the dividing of artificial neurons into gatherings with obviously related discriminative parts.

Ajit Danti et al [19] proposed Neural Network based Multiple Hierarchical Decision to predict Human Age and Gender. Faces assumes a vital section in the estimation/forecast of the age and sexual orientation of people, just by taking a gender at their face. Seeing human faces and displaying the unmistakable elements of human faces that contribute most towards face recognition are a portion of the difficulties faced by computer vision and psychophysics researchers.

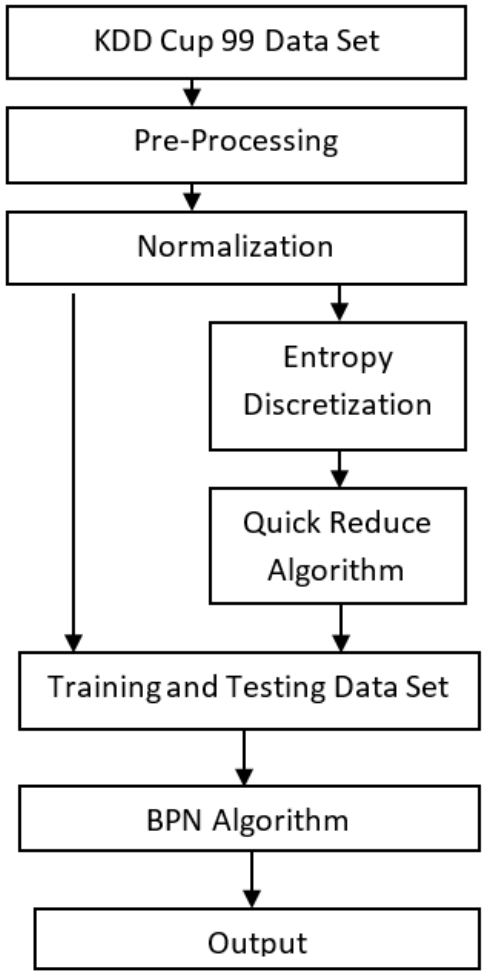
Dema Zaidan Andraws Swidan et al.[20] proposed the framework that examines the keystroke dynamics and utilizations it as a second authentication figure. The audit proposes a model for a reassure application made for gathering timing and non-timing data from keystroke flow. In addition to other mentioned in literature studies, they proposed complex password combination, which consists of text, numbers, and special characters. Also artificial neural networking utilized in the Strengthening access control. NN model based on multilayer perceptron classifier which uses back propagation algorithm is proposed. Several experiments have been done based on specific machine learning for data mining and classification toolkit named WEKA. The acquired solutions demonstrate that keystroke dynamics gives adequate level of execution measures as a moment confirmation figure. The discernible part for non-timing features close to the timing features is illustrated. These features have a critical part to

improve the execution measures of keystroke element behavioral authentication. The developed structure attain lower error rate of false acceptance of 2.2%, false rejection of 8.67%, and equal error rate of 5.43% which are better than most of references provided in the literature.

Vinod Kumar Giri et al. [21] proposed A NN approach and Wavelet analysis for ECG classification. ECG is essentially the graphical representation of the electrical action of cardiac muscles amid constriction and discharge stages. It helps in assurance of the cardiac arrhythmias in a well way. Because of this early recognition of arrhythmias should be possible legitimately. As it were can state that the bio-potentials produced by the cardiac muscles brings about an electrical flag called Electro-cardiogram (ECG). It goes about as an essential physiological parameter, which is being utilized only to know the condition of the cardiac patients. Include extraction of ECG assumes an imperative part in the manual and programmed examination of ECG for the utilization in uniquely outlined instruments. The examination ECG signal falls under the utilization of pattern recognition. The ECG signal created waveform gives all data about action of the heart. The ECG signal element extraction parameters, for example, spectral entropy, Poincare plot and Lyapunov type are utilized for study in this paper. This paper also includes artificial neural network as a classifier for identifying the abnormalities of heart disease.

LIMAM Selma et al. [22] proposed the integration and inflation of Three-Dimensional Periodic Phased Array Antenna using ANN Method. There frequent logical yields are not accessible for complex genuine frameworks, so that the computational cost of a single investigation can be restrictive. Consequently the plan technique must be extremely successful and adaptable. Optimization algorithms have given a difficult set as dependable methods for electromagnetic designs. At that point, this work concentrates on utilizing a proficient ANN scheme for the demonstrating and blending of the consistently separated linear phased array antenna. So, assuming the network's training database includes a finite set of samples of targets at certain angles are available. Neural Networks are multi-layered perception (MLP) with a back-propagation training algorithm. The given synthesis approach assured considerable improvements in terms of performances, computational speed (convergence's time) and software chosen displayed and examined by neural networks. However ANN is utilized generalization with early stopping method for produced the rapid solutions of synthesis.

III. PROPOSED WORK



In this paper an effective method of IDS is designed for detecting and classifying the kind of attacks[2] in an effective manner. In the suggested work, the pre processed KDD Cup99 data set is divided into training and testing segment. The training segment is fed to BPN algorithm for training and further tested for efficacy in detection of intrusions. Similarly the same training data set is discritized using entropy discritization[3] algorithm and further the prominent features are selected through quick reduct algorithm. Then the selected features are trained and tested accordingly for better detection. The schematic diagram is represented as below:

Training a Neural Network

Once a network has been structured for a particular application, that network is ready to be trained. To start this process the initial weights are chosen randomly. Then, the training, or learning, begins.[6]

There are two approaches to training - supervised and unsupervised. Supervised training involves a mechanism of providing the network with the desired output either by manually "grading" the network's performance or by providing the desired outputs with the inputs. Unsupervised training is where the network has to make sense of the inputs without outside help.

A. Input data for attack detection:-

The input data set is KDD 1999 Cup which has four categories of attacks and 22 attack types. One is customary and considered as normal. The Categories were as under:

DoS – The legitimate users are prevented from using services by an attacker. Generally six out of total 22 attacks fall into this group.[4] [7]

Probe:- The attacker trace to collect information about the target task. Generally four out of total 22 attacks fall into this group.

U2R:- The attacker as a local account on the victim and trace to gain root privileges. Generally four out of total 22 attacks fall into this group.

R2L:- The attacker does not have local account on a host and trace to obtain it. Generally eight out of total 22 attacks fall into this group.

The full KDD Cup 1999 data set consists of 5 million registers and very difficult to evaluate all registers. Hence 10% training set was utilized.60741 records have been selected for evaluation. Out of which 80% record were used for training and 20% record for testing. [8]Here the data set contains 41 conditional attributes(features) and one decision attribute.[14]

B. Preprocessing of input data.

The classified attack types are preceded for the methodology of preprocessing. The input data for the semantic network should be in the range [0 1] or [-1 1]. Therefore preprocessing and standardization data is needed. Preprocessing of KDD Cup 1999 format data is done.[9] Every record in KDD Cup 1999 format contain 41 features, each of which has been in one of the constant, discrete plus symbolic form, having differing ranges. For instance, the KDD CUP 99 dataset comprises numerical and symbolic types. Such symbolic features contain the protocol type (for instance, TCP, UDP plus ICMP), service type (e.g., HTTP, FTP, Telnet so on) plus TCP status flag (for instance, SF, REJ and so on).[10]

The technique simply exchanges the standards of the categorical characteristics having numeric values.[11] An integer code is allotted to every symbol for altering symbols into numerical form. For example, in the protocol type feature case, 0 is allotted to TCP, 1 to UDP, and 2 to the ICMP symbol.

Attack names were initially designed to one of five classes, 0 for Normal, 1 for Probe, 2 for DoS, 3 for U2R, and 4 for R2L.Moreover, two types spanned over a large integer range, called src_bytes [0, 1.3 billion] and also dst_bytes [0, 1.3 billion]. Logarithmic scaling (having base 10) was utilized to such features for deducting the range to [0.0, 9.14]. All another features were Boolean, in range [0.0, 1.0]. Therefore scaling was not required for these attributes.[11]

Normalization

Normalization is the initial process in the preprocessing procedure. A salient procedure of data preprocessing next to transmitting all symbolic types into numerical standards is normalization. Also, Data normalization is a method of ascending the worth of every feature into a well-proportioned collection; hence the favoritism favorable of attributes with greater standards is precluded from dataset. Each feature amidst each record is generalized by the corresponding optimum value then access to similar range of [0-1].

If (f > Max F) Nf= 1;
Otherwise Nf = (f/*MaxF);*
F: Feature;
F:Feature value;
Max F: Maximum acceptable value for F
Nf: Normalized or scaled value of

The transmitting and regulation technique is espoused for testing the data. Based on the optimum values and the succeeding easy formula, regularization of attributed values in the range [0,1] is computed. The normalized feature values are preceded for the training plus testing data set.[15]

C. Training and Testing

1) Training

The normalized 41 features from 48,593 records are chosen for training dataset. The networks are generally trained for performing tasks of configuration recognition plus decision-making.[16]

2) Testing

12148 records from the normalized 41 features are chosen for testing data set. Here the accuracy of every neural networks are calculated .[23]

D. Entropy Discretization Algorithm.

Entropy discretization method is working based on the heuristic of entropy minimization. This method is advantageous in that it is able to not only to remove noisy features but also effectively explore the remaining most representative features with critical cutting points. Multi-value discretization is same as binary discretization realized by recursively applying the same criterion to the subsets of the first partition until the stopping criterion met. The criterion is the minimum description length principle. The entropy-based discretization method is discussed in-detail.

The cutting operation first partitions the example sets S into the subsets S1 and S2. Assume that there are k classes D1, D2, …, D*k*. Let P(D*i*, S*j*) be the proportion of examples in S*j* that belong to class D*i*. The class entropy of a subset S*j*, *j* = 1, 2 is defined as:

$$E(S) = -\sum_{i=1}^{k} P(Di, S)\log(P(Di, Sj))$$

Suppose the subsets S1 and S2 are induced by partitioning a feature *A*

which is an attribute in this study. The information entropy of the partition, denoted as E(A, T; S), is given as.

$$E(A,T;S)=|S1|/|S| \text{ Ent}(S1)+ |S2|/|S| \text{ Ent}(S2)$$

The best cutting point is one, which gives the minimal class information entropy among all the candidate-cutting points. One can generalize the algorithm to be applicable to multiple intervals. From the binary discretization to multi-splitting discretization, recursive call of the binary discretization has performed until a stopping condition according to Minimal Description Length Principle (MDLP) is met.

Algorithm -Entropy Discretization

Input: Decision system *DS = (U, A U {D}), D* – decision attribute

Output: A best cut point and discrete value.

Step 1: Choose the best cut point according to the entropy criteria (Ta)

$E(S) = - \Sigma^{k}_{i=1} P(Di, S)\log(P(Di, Sj))$

Step 2: Evaluate, if the cut point is significant according to the

Minimum Description Length principle

if it is not significant, return,

else recursively call the discretization algorithm for each of

the split by the cut point Ta.

The Entropy based Discretization algorithm gives a qualitative data.

Quick Reduct Algorithm

QuickReduct(C, D)

C, the set of all conditional features;

D, the set of decision features.

Step 1: R ← {}

Step 2: **Do**

Step 3: T ← R

Step 4: ∀ X ∈ (C-R)

Step 5: **if** γR∪{x} (D) > γT (D)

where $\gamma R\ D = \frac{|POSR(D)|}{|U|}$

Step 6: T ← R∪{x}

Step 7: R ← T

Step 8: **until** γR (D) = = γC (D)

Step 9: **return** R

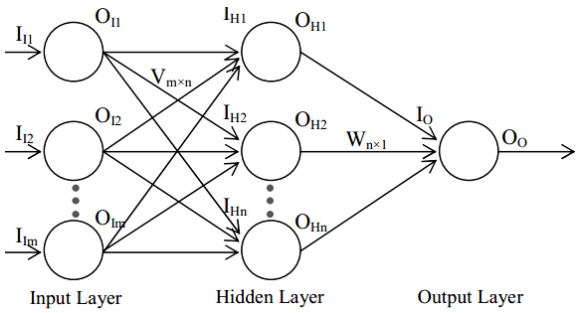
Quick Reduct, uses a forward selection, non exhaustive hill-climbing search prone to local optima. The evaluation is attribute subset measuring the rough set dependency value. The goal state is reached when the search finds the maximum possible dependency value for the dataset. It searches for a minimal subset without exhaustively generating all possible subsets. The search begins with an empty subset; attributes which result in the greatest increase in the rough set dependency value that is added iteratively.This process continues until the search produces its maximum possible dependency value for that data set.

Here the feature selected are 3,5,6 and 40 for 5001 rows with an elapsed time of 3800.147792 seconds.

BACK PROPAGATION NEURAL NETWORK

Neural Networks (NN) are important data mining tool used for classification and clustering. It is an attempt to build machine that will mimic brain activities and be able to learn by examples. If NN is supplied with enough examples, it should be able to perform classification and even discover new trends or patterns in data. The NNs have been efficaciously used for classification purposes in medical domains, including the classification of cancer samples in gene expression data. In this section, a three-layer Back Propagation Neural Network (BPN) is considered as a classifier. The three layers are input, hidden and output layers. Each layer can have number of nodes and nodes from input layer are connected to the nodes from hidden layer. Nodes from hidden layer

are connected to the nodes from output layer. Those connections represent weights between nodes.



Initially the inputs are normalized between [0, 1]. Two weight matrices are used: the weight matrix V represents the weights of synapses connecting input neurons and hidden neurons and the weight matrix W represents weights of synapses connecting hidden neurons and output neurons. The weights between neurons are initialized between [-0.5, 0.5].

[V]0 = [random weights]
[W]0 = [random weights]

Initially a training sample is presented to the input layer II as inputs to the input layer. By using linear activation function, the output of the input layer may be evaluated as: OI = II. Then the inputs to the hidden layer are computed by using the output of the input layer and weight
V as:
IH = V^TOI
Let the hidden layer units evaluate the output using the purelin transfer function as: OH = IH.
Then in the next step, the inputs to the output layer are obtained by multiplying corresponding weights of synapses as:
IO = WT OH
And the output layer units evaluate the output using linear activation
function as: OO = IO
From the output of the network, the error is calculated between the network output and the desired output as for the ith training set as:

E = |Tk– OOk|, if the error is negligible (≤ 0.0001), then the learning is continued with the next training sample, otherwise, the weights are updated based on the error and the learning is continued with the same training sample until the error becomes low.
The weights between hidden to output layer are back-propagated as:
D = (Tk-OOk)OOk(1-OOk)
[Y] = OH.d
[ΔW] = α[W]tη[Y]
where α and η are momentum and learning rate parameters to control the weight updation process in BPN architectures. The weights between inputs to hidden layer are backpropagated as:
E = [W] {d} and d* = ei(OHi)(1-OHi) and
[X] = OId* = IId*
[ΔV] = α[V]tη[X]
Update the weight matrices as:
[V]t+1 = [V]t + [ΔV] and [W]t+1= [W]t + [ΔW]
This learning process is then repeated for all the training samples.

IV. RESULTS AND DISCUSSION

In this portion, the investigational results attained for the suggested technique are proffered. The KDD Cup 99 dataset espoused for assaying the recommended intrusion detection methodology is taken from the publicly possible sources. The MATLAB software is espoused for fostering and testing the dataset plus the proficiency is computed.

A. Performance Analysis

The assessment metrics of sensitivity, specificity plus accurateness can be articulated regarding TP, FP, FN plus TN.

Sensitivity = TP/(TP+FN)

Specificity is the fraction of the true negatives to the sum of true negative and false positives.

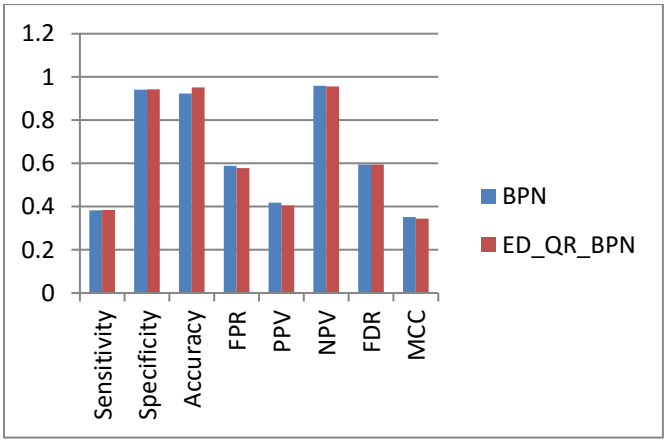
Specificity= TN/(TN+FP)

Accuracy has been the proportion of accurate outcomes, either true positive or else true negative, in a population. It calculates the veracity degree of a analytical test on a condition.

B. COMPARITIVE ANALYSIS

Measures	BPN	ED_QR_BPN
Sensitivity	0.3825	0.3838
Specificity	0.9412	0.9422
Accuracy	0.9230	0.9518
FPR	0.588	0.578
PPV	0.4184	0.4052
NPV	0.9584	0.9563
FDR	0.5948	0.5948
MCC	0.3509	0.3435

The graphical representation of the above received output is depicted as below:



C. DISCUSSION:

It is observed from the above results that the sensitivity of proposed method is slightly better than the existing BPN. Also the accuracy is improved after discretization and reducing the features through quick reduct. The accuracy is improved by the time factor and processing speed.

V. CONCLUSION

In this paper, it is seen that the accuracy in intrusion with Entropy Discretization +Quick Reduct +Back Propagation Neural Network is better than BPN. The accuracy in detection may be increased further by properly modifying the weights and biases of BPNN or through deep learning algorithms.

VI. REFERENCES

[1]. P. Ramasubramanian and A.Kannan, "Intelligent Multi Agent Based Multivariate Statistical Framework for Database intrusion prevention system". School of Computer Science and Engineering, Anna University, India.

[2]. PrashantDewan, ParthsDasgopts, Vijay Karamcheti,"Defending Against Denial of Service Attacks using Secure Name Resolution".

[3]. "An Analysis of current computer network attack procedures, their Mitigation measures and the development of an improved Denial of Service attack Model", IhekWeabaOgechi, Inyiama H.C, IhekWeabaChukwugoziem.

[4]. "Mitigation of Denial of Service attack", Payal Jain, Juhi Jain, Zatin Gupta.

[5]. "Using Artificial Intelligence in Intrusion Detection Systems", MattiManninen

[6]. "Artificial Intelligence Techniques Applied to Intrusion Detection", BharanidharanShunmugam.

[7]. "Study and performance evaluation on recent DDoS trends of Attack and Defense", Muhammad Aamir, Muhammad Arif

[8]. "Adaption of the neural network-based IDS to new attacks detection", PrzemyslawKukielka, ZbigniewKotulski

[9]. Data Security based on neural networks-Khaled M G Noaman and Hamid Abdullah Jalab.

[10]. Detecting and preventing attacks using network intrusion detection systems- Meera Gandhi and S K Srivatsa
[11]. Network and computer security tutorial version 0.4.0
[12]. Cisco certified network professional guide.
[13]. "Artificial Intelligence techniques applied to Intrusion detection" by Bharanidharan Shanmugam.
[14]. "Using Artificial Intelligence in Intrusion detection Systems", Matti Manninen
[15]. "Image based Authentication techniques to prevent Intrusion" by Kavya. S, Manasa.D
[16]. "Detecting and preventing attacks using network intrusion detection systems", by MeeraGhandhi, S.K. Srivasta
[17]. Kaur, Sukhdeep, Sharma AS, Kaur H, and Singh K (2016) Gene selection for tumor classification using resilient backpropagation neural network. In Advances in Computing, Communication, & Automation (ICACCA)(Fall), International Conference on, 1-5, IEEE.
[18]. Rauber, Paulo E., Fadel SG, Falcao AX, and Telea AC (2017). Visualizing the Hidden Activity of Artificial Neural Networks. IEEE Transactions on Visualization and Computer Graphics 23(1): 101-110.
[19]. Dileep, MR and Danti A (2016) Multiple hierarchical decision on neural network to predict human age and gender. In Emerging Trends in Engineering, Technology and Science (ICETETS), International Conference on, 1-6. IEEE.
[20]. Salem, Asma, Zaidan D, Swidan A, and Saifan R (2016) Analysis of Strong Password Using Keystroke Dynamics Authentication in Touch Screen Devices. In Cyber security and Cyber forensics Conference (CCC), 15-21. IEEE.
[21]. Gautam, Kumar M and Giri VK (2016) A Neural Network approach and Wavelet analysis for ECG classification." In Engineering and Technology (ICETECH), 2016 IEEE International Conference on, 1136-1141. IEEE .
[22]. Bilel, Hamdi, Selma L, and Taoufik A (2016) Artificial neural network (ANN) approach for synthesis and optimization of (3D) three-dimensional periodic phased array antenna." In Antenna Technology and Applied Electromagnetics (ANTEM), 2016 17th International Symposium on IEEE, 1-6.
[23]. "Neural Networks" by Simon Haykin

ABOUT AUTHORS

1. S.Vijaya Rani, MCA,M.Phil,(Ph.D). Research Scholar, Bharathiar University,Tamilnadu

She is a Research Scholar in Computer Science of Bharathiar University, Tamilnadu and having 10 years of teaching experience. The research area is Intrusion detection system using Artificial Neural Network. She has published 10 articles in various Journals and conference proceedings. She also attended various International ,National conferences and workshops.

2. Dr.G.N.K.Suresh Babu, M.E.,M.C.A.,M.Phil., Ph.D, Associate Professor,MCA Dept, Acharya Institute of Technology, Bangalore.

He is having 25 years of teaching and guiding experience. The research area is data mining and Artificial Neural Networks. He has published many articles in Journals and attended International and national conferences.