

FAULT DETECTION AND DIAGNOSIS OF POWER SYSTEMS USING ARTIFICIAL NEURAL NETWORKS

K.S.Swarup

H.S.Chandrasekharaiah

DEPARTMENT OF HIGH VOLTAGE ENGINEERING
INDIAN INSTITUTE OF SCIENCE
BANGALORE 560012, INDIA

Abstract

Real time Fault Detection and Diagnosis (FDD) is an important area of research interest in Knowledge Based Expert Systems. Neurocomputing is one of fastest growing areas of research in the fields of Artificial Intelligence and Pattern Recognition. This paper explores the suitability of pattern classification approach of neural networks for fault detection and diagnosis. Suitability of using neural network as pattern classifiers for power system fault diagnosis is described in detail. A neural network design and simulation environment for real-time FDD is presented. An Analysis of the learning, recall and generalization characteristic of the neural network diagnostic system is presented and discussed in detail.

KEY WORDS: Pattern classification, Fault detection and diagnosis, neural network, power systems.

1 Introduction

An important application of artificial intelligence (AI) is the diagnosis of faults of mechanisms and systems in general. Traditional approaches to the problem of diagnosis is to construct a heuristic, rule-based system which embodies a portion of the compiled experience of a human expert. These systems perform diagnosis by mapping fault symptoms to generated hypothesis to arrive at diagnostic conclusions. The knowledge acquisition and search process involved in expert systems is exhaustive and hence time consuming. In addition, the simulation of models are usually too slow to be effectively applied in a real-time environment.

Artificial Neural Networks (ANN) are found to be suitable for the above requirements [7]. They are massively parallel interconnected networks of simple adaptive elements and their hierarchical organizations, which are intended to interact with the objects of the real world in the same way as the biological counterparts. Neural networks find wide applications in parallel distributed processing and in real-time environments [9,14,15]. Neural Networks have considerable advantages over expert systems in terms of knowledge acquisition, addition of new knowledge, performance and speed[8]. Recently, interest in the application of associative memories and neural networks to problems encountered in diagnostic expert systems development has increased. Neural Networks appear to offer features which coincide well with the requirements of pattern-based diagnosis [10,19]. An important feature of fault diagnosis using neural networks is that they can interpolate the training data to give an appropriate response for cases described by neighbouring or noisy input data. This paper describes the design and simulation of a neural network for fault detection and diagnosis of power systems. In this paper fault diagnosis is conceptualized as a pattern classification problem which involves the association of patterns of input data representing the behaviour of the power system to one or more fault conditions. The neural net-

work is trained off-line with different fault situations and used on-line. The diagnostic system was able to detect and diagnose the faulted component corresponding to the input pattern consisting of switching status of relays and circuit breakers.

1.1 Neuro Computing : A Pattern Recognition Perspective

Neural computing is one of the rapidly expanding areas of current research, attracting people from a wide variety of disciplines. Pattern recognition in one form or the other is currently the dominating area for the application of neural networks. Pattern recognition is generally defined as an abstract formulation of the categorization tasks in pattern classification. Given an input of some form, it is possible to analyse the input to provide a meaningful categorization of its data content. The input to the pattern recognition system is a pattern vector and the output is the decision as to the category or the class to which the input pattern belongs. A pattern recognition system is generally considered as a two stage device, The first stage is **feature extraction** and the second stage is **classification**. Feature extraction corresponds to selection of a definite characteristic of the input pattern. The classifier is supplied with the list of measured features. The task is to map these input patterns onto a classification state, that is, given the input features, the classifier must decide which type class category they must match closely. The classifiers typically rely on distance metrics and probability theories to perform the above task.

Once the feature extractor is fixed, the feature vector X is obtained. The next stage is to design an '*optimal*' decision rule so that observing X, the classifier can make a correct decision as to its class membership. Often it is not possible to derive such rules analytically. The classifier is then designed to learn the proper decision rule using a training set. The training set consists of feature vectors of known classification. During the training phase, the system is given the feature vectors one by one and is told what the classification should be. The system uses this information in a learning algorithm to learn the decision rules needed [1,2,6]. Pattern classification serves as a general framework in the learning phenomena of neural networks. In general terms, template matching may be interpreted as a special case of the '**feature extraction**' approach, where the templates are stored in terms of measurements, and a special classification criterion (matching) is used for the classifier

1.2 Pattern Classification

Pattern classification techniques fall into two broad categories of numeric and non-numeric methods [5]. Numeric techniques include deterministic and statistical measures which can be considered as measures made on the geometric space. Non-numeric techniques are those which deal with the domain of symbolic processing. Numeric techniques have large influence on the concept of pattern recognition approach to neural networks. The concept of pattern classification may be expressed in terms of

feature space or a mapping from feature space \mathfrak{R} to decision space Ψ . The input pattern consists of N measured patterns, where each set of N features can be considered as a vector X , called the feature or measurement vector, or a point in the N -dimensional feature space R . The problem of classification is to assign each possible vector or point in the feature space \mathfrak{R} to a proper pattern class. This corresponds to the partitioning of the feature space into mutually exclusive regions and each region will correspond to particular class.

1.3 Mathematical formulation of the pattern classification problem

Classification usually involves measurement of various parameters that are able to adequately distinguish inputs that belong to different categories or 'classes', as they are normally called. The input pattern, consisting of N measurements, each of which having a unique feature, is called a 'feature vector'. Mathematically, the problem of classification can be formulated in terms of 'discriminant functions'.

Let C_1, C_2, \dots, C_m be the m possible pattern classes to be recognized and let

$$X = (x_1, x_2, \dots, x_n)^t \quad \dots(1)$$

be the feature vector, where x_i represents the i th feature measurement. Then, the discriminant function $D_j(X)$ associated with class C_j , $j=1, 2, \dots, m$ i.e., such that, if the input pattern represented by the feature vector X is in class C_i , denoted as $X \sim C_i$, the value of $D_i(X)$ must be the largest, i.e., for all $X \sim C_i$,

$$D_i(X) \geq D_j(X) ; C_{ij} = 1, 2, \dots, m, i = j \quad \dots(2)$$

Thus in the feature space \mathfrak{R} , the boundary of partitioning called the 'DECISION BOUNDARY', between regions associated with class C_i and class C_j , respectively is expressed by the relation,

$$D_i(X) - D_j(X) = 0 \quad \dots(3)$$

Many different forms satisfying the condition (2) can be selected for $D_i(X)$. Discriminant functions are the basis for the majority of pattern recognition techniques. The basic idea behind discriminant function is a line or a decision boundary separating features belonging to a particular class. The mathematical definition of such a decision boundary is a discriminating function. It is a function that maps input features onto a classification space. Few of the important discriminant functions employed for pattern classification are 1. **Linear discriminant function**. 2. **Minimum distance classifier**. 3. **Piecewise linear Discriminant function**, (Nearest neighbour Classification) 4. **Polynomial discriminant function**. The linear discriminant function forms the basis of pattern classification using neural networks.

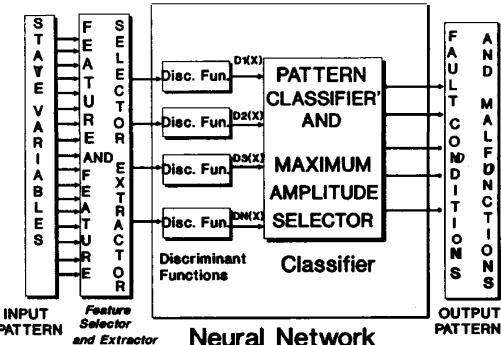
2 Fault Detection and Diagnosis using Neural Networks

Current techniques for fault detection and diagnosis rely on experts and expert systems modeling using classical techniques in the time or frequency domain. Neural network classifiers can learn and adapt themselves to different statistical distribution and non-linear mappings. The parallel structure of neural networks permits 'INCIPIENT FAULT DETECTION' which is an indication of an increase in the lead time for detecting faults. Neural Networks provide a greater degree of robustness or **fault tolerance**

than competitive fault detection methods because there are many more processing nodes, each with primarily local connections. Damage to a few nodes or links does not impair overall performance. In addition, neural networks are non-parametric and can make weaker assumptions about the character of their probability distributions of the sensor measurements. Artificial neural networks have the capacity to learn and store information about fault occurrence via associative memory and thus have an associative diagnostic ability with respect to the faults that occur in a power system [16,17].

2.1 Diagnosis as a Pattern Classification problem

Real-time diagnosis of faults occurring in complex systems is an active area of research in the field of knowledge-based expert systems. Neural networks offer features which coincide with the requirements of '**pattern-based diagnosis**'. In the present approach, fault diagnosis is conceptualized as the association or



Diagnosis as a Pattern Classification Problem

Figure 1: Fault Diagnosis as a Pattern Classification Problem

mapping of patterns of input data representing the behaviour of the physical system to one or more fault conditions [19]. The behaviour of a physical system is often described in terms of the temporal behaviour of various physical parameters relating to the system. Figure 1 shows the basic principle of the diagnostic process conceptualized as a pattern classification problem.

3 Learning

3.1 Neural Network Training

The artificial neural network is trained so that application of a set of input produces the desired set of outputs. Training is accomplished by sequentially applying input vectors, while adjusting network weights according to pre-determined procedures. During the training process, the network weights gradually converge to values such that each input vector produces the desired output vector. Training algorithms are categorized as supervised and unsupervised.

Supervised training requires the pairing of each input vector with a target vector representing the desired output. The input vector is applied, the output of the network is computed and compared with the corresponding target vector, and the difference is fed back through the network and the weights are changed according to an algorithm that tends to minimize the error. Unsupervised learning is an important concept in neural networks that led to the development of various configurations of learning phenomena. The training set consists solely of input

vectors. The training algorithm modifies network weights to produce output vectors that are consistent. The training process extracts the statistical properties of the training set and groups similar vectors into classes. Kohonen's self organising network is an example of the unsupervised training algorithm.

Artificial neural networks learn patterns of activations and hence, learning can be equated to determining the proper values of the connection strengths that allow all the nodes to achieve the correct state of activation for a given pattern of inputs. Once the pattern of activation is established, the resulting outputs let the network classify an input pattern. The adaptive nature of the neural network allows the weights to be learned by experience, thus producing a '**self-organizing system**'. Much of the recent interest in neural networks can be attributed to new and more effective learning methodologies. A robust learning heuristic for multi-layered feed forward neural network called the **GENERALIZED DELTA RULE (GDR)** or '**BACK PROPAGATION LEARNING RULE**'[14] has been implemented in this paper.

4 Case Study: Power System Fault Diagnosis

Figure 2 shows the model power system employed for fault diagnosis. The power system consists of three buses, two transformers and two transmission lines. The protective elements of the power system consists of 11 circuit breakers and 16 relays. The diagnostic system functions based on the protective co-ordination of the relays and circuit breakers. Figure 3 shows the topology of the neural network employed for fault detection and diagnosis. The input to the neural network consists of the status (on/off) of the relays and circuit breakers of the power system. The neural network is designed so that it captures the behaviour of the power system represented in the form of signals from relays and circuit breakers, resulting in a neural network consisting of 35 input nodes and 10 output nodes. Seventy four input nodes are also used to represent each device behaviour detail. The input pattern of the neural network is a feature vector consisting of 35 features. Eleven features correspond to the status of the circuit breaker relays, three buses (differential relays), three features each for the transformer protective relays and six features each for transmission line protection. The nodes in the hidden layer were varied from 4 to 50 for every input pattern and the performance of the network in determining the optimum hidden nodes was carried out. It was found that 18 nodes in the hidden layer gave the optimum and satisfactory performance in terms of generalization and recall performance

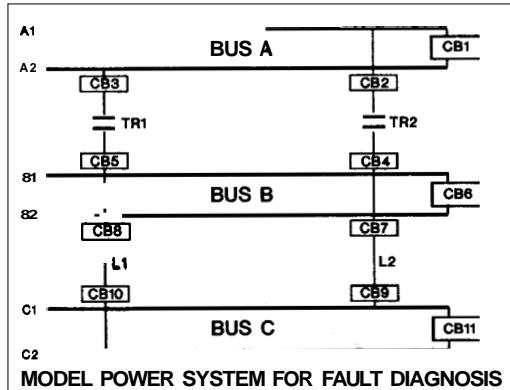


Figure 2: Model Power system for fault diagnosis

of the neural network. The output layer consists of ten nodes, corresponding to the various components of the power system

Fifty three major faults on the power system are considered. An input pattern corresponding to a particular fault condition is fed to the neural network and the desired output pattern corresponding to the fault condition is impressed at the output. The neural network is trained using the back propagation algorithm till the error in the weights between successive iterations is less than a specified minimum value. The neural network is trained off-line for 53 different types of faults and used on-line. When an unknown output pattern corresponding to the fault condition is given at the output layer of the neural network, the diagnostic system classifies the type duration and severity of the fault. The neural network is trained with the help of the data obtained through the simulation of a model-power system for different normal and abnormal operating conditions.

4.1 Behavioral representation

To train and use neural networks, behavioral data must be presented to the inputs of the networks. Two methods of behavioral data representation are widely used. In the first method, simulation data is presented to the input layer of the neural network as a vector of continuous real variables. In the second method, parameter behaviour is represented in a binary form. The performance of the power system under a fault is represented as a field of 0's and 1's corresponding to the 'off' and 'on' states of the circuit breakers and relays of power system. The field is represented as a single binary vector which is presented to

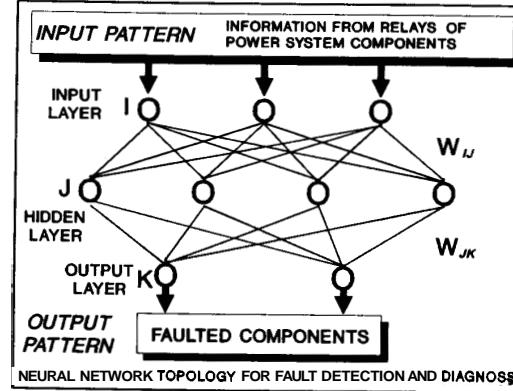


Figure 3: Neural Network Topology

the input of the neural network. The performance of the neural network in terms of discriminating power was found to be more suitable in the case of second method as the behaviour of the power system as a whole is exploited in training the network.

5 Neural Network Performance

Artificial neural networks are associated with the inherent properties of Learning, Generalization, Abstraction and Applicability and hence its performance is measured in terms of the above features. ANNs modify their response to the environment. Given a set of inputs along with desired outputs, they self organise to produce consistent properties. Once trained, the networks response can be insensitive to minor variations in its input. ANN generalises automatically as a result of its structure. ANNs are capable of abstracting the essence of a set of inputs. They can be trained to produce something it has-not seen before. They are

best suited for large class of pattern recognition tasks through the Connectionist property.

5.1 Results of Power System Fault Diagnosis

The performance of the neural network for fault diagnosis of power systems is carried out in terms of the following 1. Generalization of Faults.

2. Discrimination between single and multiple faults.
3. Recall and learning performance.
4. Learning and Momentum rate improvement.
5. Optimising the number of hidden nodes and layers.
6. Change in weights in the hidden layer (s).

Figures 4.1 to 4.8 show the Generalization and Recall performances of the neural network. Figure 4.1 shows the performance of the neural network with variation in the hidden layers. The learning performance of the neural network increases significantly with a double layer neural network. Figure 4.2 shows the Performance of the neural network with single hidden layer, the number of nodes varying from 4 to 50. Figure 4.3 shows the performance of a double hidden layer neural network with hidden nodes varying from 4 to 50. In both the cases, the learning performance has improved from 4 to 31 hidden nodes and decreased with further increase in hidden nodes. The optimum hidden nodes in single and double hidden layers was found to be 18 and 25 respectively. Figure 4.4 shows the generalization and recall performance of the neural network for different faults on the power system. Four different types of faults are simulated on the power system. Figure 4.5 shows the performance of the neural network with variation in the learning rate. The learning performance was most satisfactory at $\mu = 0.15$. Figure 4.6 shows the performance of the neural network with different momentum rates. The momentum rate of 0.075 was found to be suitable. The performance of the neural network for variations in the input is shown in figure 5. The magnitude of the input pattern is varied from 0.9 to 0.5. Figure 5 shows the performance of the single hidden network for $I=0.9$ and $I=0.5$. The performance of the double hidden layer network for $I=0.9$ and 0.5 is also shown in figure 5.

It can be observed from the figures that initially a local minima occurs in the second iteration and later with the increase in training period, a global minima is achieved. The network satisfactorily discriminates between the faults as shown in the above figures. It can be observed that the neural network is able to classify and discriminate between faults on the power system even at $I=0.5$. The generalization and recall performance for single and multiple faults is shown in figure 6.

6 Conclusions

Suitability of Artificial Neural Networks for fault detection and diagnosis of power systems is described in this paper. Various network topologies (number of hidden nodes and layers) have been tested and compared. The results show that accurate recall and generalization behaviours are obtained during the diagnosis of single faults. Performance during recall improves at first with an increase in the number of hidden nodes (units) and with the amount of training, and finally attains convergence. In general, performance during generalization improves with the duration of the training period. The neural network diagnostic system is also able to diagnose correctly even in the presence of faulty operation of the relays of the power system and under disturbances. The neural network diagnostic system trained for single

faults was found to be able to accurately diagnose abnormal behaviour resulting from simultaneous multiple faults. Graceful degradation of the diagnostic system was observed in situations where faults were not accurately diagnosed or under damage to a few nodes. Research is under progress in the implementation of advanced training algorithms, like Counter Propagation and Simulated Annealing using Boltzmann Machines, for efficient diagnostic problem Solving.

7 References

1. Nilsson, N. J 'LEARNING MACHINES ', Mc Graw-Hill, NY, 1965.
2. Fu, K. S 'SEQUENTIAL METHODS IN PATTERN RECOGNITION AND MACHINE LEARNING', Academic Press, NY 1968
3. Rosenblatt, 'PRINCIPLES OF NEURODYNAMICS', Spartan, 1962.
4. Minsky.M and Pappert, 'PERCEPTRONS', MIT Press, Cambridge, 1969.
5. Duda, R. O and P.E.Hart, 'PATTERN CLASSIFICATION AND SCENE ANALYSIS', Wiley , NY, 1973.
6. Michalski. R. S ,J.S Carbonell and T.M.Mitchel, 'MACHINE LEARNING :AN ARTIFICIAL INTELLIGENCE APPROACH'. Tioga Publications 1983.
7. Denker.S 'NEURAL NETWORKS FOR COMPUTING'. AIP Conference Proceedings, No. 151.NY, 1986.
8. Anderson, J.A and Rosenfield E 'NEURO-COMPUTING: FOUNDATIONS OF RESEARCH', MIT Press, Cambridge, Mass, 1988.
9. Wasserman, P. D 'NEURAL COMPUTING : THEORY AND PRACTICE', Van Nostrand, Reinhold, NY, 1989.
10. Pao, Y. H 'ADAPTIVE PATTERN RECOGNITION AND NEURAL NETWORKS' , Addison Wessly, NY, 1989.
11. Hopfield J, J. and Tank D, W. 'Computing with neural networks: A model', Science Vol. 233, August 1986, pp625-633.
12. Hopfield J, J 'Neural networks and physical systems with emergent collective computational abilities', proceedings of the National Academics of Science, USA, Biophysics, Vol. 79. April 1982, pp 2554-2558.
13. Hopfield J, J and Tank D, W ' Neural computation of decision in optimization problems', Biological cybernetics, Vol.52, Springer Verlag, 1985,pp141-152.
14. Rummelhart D, and McClelland J, 'PARALLEL DISTRIBUTED PROCESSING: EXPLORATIONS OF THE MICROSTRUCTURE OF COGNITION', Volumes I and II, MIT Press, Cambridge, Mass. 1987.
15. Lippmann. R ' An introduction to computing with neural networks',IEEE ASSP Magazine, August 1987,pp4-22.
16. Sobajic D and Pao Y, H 'Artificial neural network based dynamic security assessment for electric power systems', Case Western Reserve University, Technical report, TR87-135, December 1987.
17. Ebron S 'A NEURAL NETWORK PROCESSING STRATEGY FOR THE DETECTION OF HIGH IMPEDANCE FAULTS', Master's thesis, Electrical and Computer engineering department, North Carolina State University, NCSU, 1988.
18. Swarup, K S and Chandrasekhariah H, S 'Pattern directed inference system for fault classification and analysis', Sixth National Power Systems Conference, Bombay. June 1990., pp 145-150.

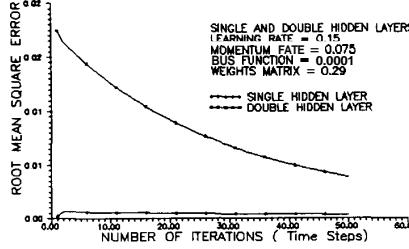


FIGURE 4.1 : Performance of the neural network with various hidden layers

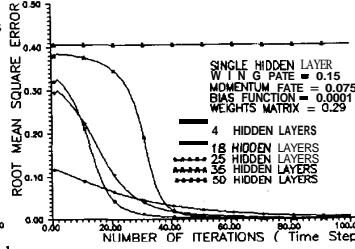


FIGURE 4.2 : Performance of the neural network with single hidden layer

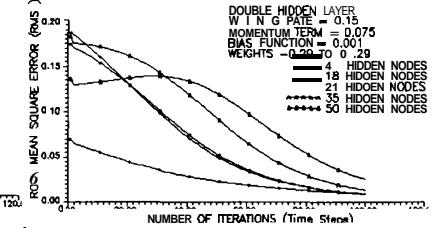


FIGURE 4.3 : Performance of the neural network with double hidden layer

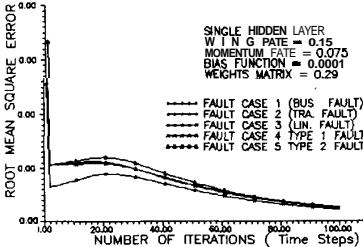


FIGURE 4.4 : Fault Generalization of the neural network

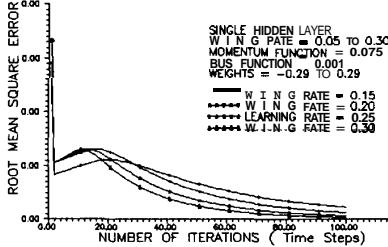


FIGURE 4.5 : Learning rate improvement of the neural network

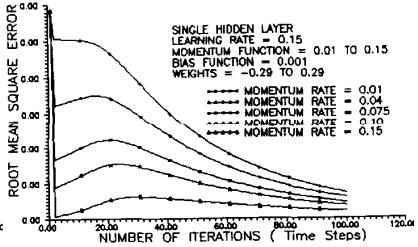


FIGURE 4.6 : Momentum rate improvement of the neural network

	I	2	3	4	5	6	7	8	9	10
A1	A .00000	10000	10000	10000	10000	10000	10000	10000	10000	10000
A1	P .73386	10247	10250	10273	10286	10288	10293	10294	10295	10295
A2	A .10000	00000	10000	10000	10000	10000	10000	10000	10000	10000
A2	P .72086	10254	10257	10280	10293	10299	10305	10315	10316	10316
B1	A .10000	72589	00000	10000	10000	10000	10000	10000	10000	10000
B1	P .10205	10345	.73591	10371	10383	10396	10398	10403	10414	10416
B2	A .10000	10000	10000	00000	10000	10000	10000	10000	10000	10000
B2	P .03370	03497	03508	.75157	03532	03544	05015	05121	04933	05110
C1	A .10000	10000	10000	00000	10000	10000	10000	10000	10000	10000
C1	P .10032	10172	10185	10198	.73524	10223	11504	11550	11550	11641
A2	A .10000	10000	10000	10000	00000	10000	10000	10000	10000	10000
A2	P .09784	09937	09951	09965	.09979	.74213	11498	11582	11543	11632
T1	A .10000	10000	10000	10000	10000	00000	10000	10000	10000	10000
T1	P .10260	10394	10406	10418	10430	10442	.74754	11526	11556	11647
T2	A .10000	10000	10000	10000	10000	10000	00000	10000	10000	10000
T2	P .10265	10399	10410	10422	10434	10447	11435	.74851	11478	11569
L1	A .10000	10000	10000	10000	10000	10000	10000	00000	10000	10000
L1	P .10319	10453	10465	10477	10489	10501	10571	11657	.75160	11529
L2	A .10000	10000	10000	10000	10000	10000	10000	10000	00000	10000
L2	P .03451	03588	03601	03613	03626	03639	03534	03547	04896	.51114

Single Hidden Layer

Performance of the Neural Network Diagnostic System

COMPARISON OF THE ACTUAL AND PREDICTED RESPONSE OF THE NEURAL NETWORK

FIGURE 5 Comparision of the actual and desired responses of the neural network

	1	2	3	4	5	6	7	8	9	10
1	A .90000	10000	10000	10000	10000	10000	10000	10000	10000	10000
A1	P .72086	11381	11381	11381	11381	11381	11381	11381	11381	11381
2	A .10000	.00000	10000	10000	10000	10000	10000	10000	10000	10000
A2	P .11381	.72086	11381	11381	11381	11381	11381	11381	11381	11381
3	A .10000	10000	.00000	10000	10000	10000	10000	10000	10000	10000
A2	P .11242	11242	.72086	11242	11242	11242	11242	11242	11242	11242
4	A .10000	10000	10000	00000	10000	10000	10000	10000	10000	10000
A2	P .11242	11242	11242	.72086	11242	11242	11242	11242	11242	11242
5	A .10000	10000	10000	10000	00000	10000	10000	10000	10000	10000
A2	P .11381	11381	11381	11381	.72086	11381	11381	11381	11381	11381
C1	P .11381	11381	11381	11381	11381	.72086	11381	11381	11381	11381
6	A .10000	10000	10000	10000	10000	00000	10000	10000	10000	10000
C2	P .11451	11451	11451	11451	11451	11451	.72095	11451	11451	11451
7	A .10000	10000	10000	10000	10000	10000	00000	10000	10000	10000
T1	P .10636	10636	10636	10636	10636	10636	10636	.76354	10636	10636
8	A .10000	10000	10000	10000	10000	10000	10000	00000	10000	10000
T2	P .10636	10636	10636	10636	10636	10636	10636	10636	.76354	10636
9	A .10000	10000	10000	10000	10000	10000	10000	10000	00000	10000
L1	P .10817	10817	10817	10817	10817	10817	10817	10817	10817	10817
10	A .10000	10000	10000	10000	10000	10000	10000	10000	10000	10000
L2	P .10925	10925	10925	10925	10925	10925	10925	10925	10925	.76816

Double Hidden Layer

A : Actual Output

P : Predicted Response

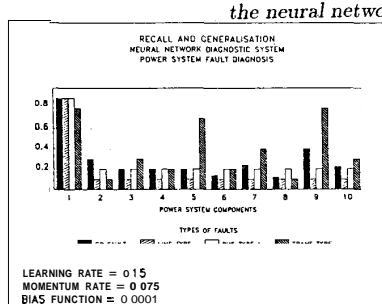


Figure 6: Generalization and Recall performance of the neural network diagnostic system