VOL. 10, NO. 11, JUNE 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

A SURVEY ON NEURAL NETWORK MODELS FOR DATA ANALYSIS

Praveen Joe I. R. and Varalakshmi P.
Department of Computer Science & Engineering, College of Engineering, Anna University, Chennai, India
E-Mail: praveenjoeir@yahoo.com

ABSTRACT

Artificial Neural Network (ANN) is an information-processing archetype that draws inspiration from biological nervous systems, like the brain, in order to process information. The key unit of this information processing system is the large number of highly interconnected processing elements called neurons. Research activities started as early as in 1943 to simulate neural behavior upon building mathematical models. A neural network model is based on the nature of the problem, characteristics of the application domain and the learning procedure of the selected model. Currently, many neural network models have been built, each with distinct performance features. An attempt is made to study the various ANN suitable for clustering with an accuracy similar to the best statistical methods and which are characterized by parallel processing, a distributed architecture, and a large number of nodes, at the same time capable of proposing an optimal number of groups into which the patterns may be clustered .This paper gives an overview of such neural network models and their applications. This survey is a supplementary process in choosing a suitable ANN algorithm for web services clustering.

Keywords: ANN, neural network models, clustering.

1. INTRODUCTION

a) Neural network

A neural network is a very effective tool for data modeling which captures and represents complex input/output relationships. The inspiration for the development of neural network technology that branched out from the desire to develop an artificial system that could perform "intelligent" tasks similar to the activities of the human brain. According to Simon Haykin, neural networks are similar to the human brain in the following two ways:

i. A neural network gain knowledge through learning.
ii. A neural networks knowledge is stored within inter-neural connection strengths known as synaptic weights.

The intelligence of a neural network has become known from the collective behaviour of neurons, each of which performs only very partial operation. Even though each neuron works unhurriedly, they can arrive at solution by working in parallel. For this reason we can clearly trigger out why humans can recognize a visual scene faster than a digital computer, while an individual brain cell responds much more slow than a digital cell in a VLSI circuit. The following characteristics of neural networks have a significant role in a wide variety of applications: [1]

i. Adaptiveness: Powerful learning algorithms and self organizing rules permit self-adaptation as per the necessities in a repeatedly changing environment.
ii. Nonlinear processing: Ability to perform tasks involving nonlinear relationships and noise immunity makes it suitable for classification and prediction.
iii. Parallel processing: Architecture with a large number of processing units enhanced by extensive inter-connectivity provided for concurrent processing as well as parallel distributed information storage.

Neural Network models can be categorized in to classification models, Association models, Optimization models, Self Organization models and Hybrid models. The basic neural computational models are explored in this paper

b) ANN learning rules

The learning approaches in a neural network can be divided in to supervised, unsupervised and reinforcement learning. A teacher or supervisor who classifies the training set into classes and utilizes the information on the class membership of each training instance is available. In Unsupervised learning model there is a heuristic approach and no supervisor ie class labels are available. Reinforcement learning follows trial and error interactions with its environment (reward/penalty assignment). Learning depends on the space of interconnection neurons basically inspite of these classifications. That is, supervised learning learns by adjusting its inter connection weight combinations with the help of error signals where as unsupervised learning uses information associated with a group of neurons and reinforcement learning uses reinforcement function to modify local weight parameters. For learning to occur, ANN adjusts the free parameters of the network that are adapted where the ANN is embedded.

To know if the model is supervised or unsupervised or any other, the parameter adjustments play a vital role. Also, these learning algorithms are facilitated by various learning rules [8] as shown in the Figure-1.

VOL. 10, NO. 11, JUNE 2015 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

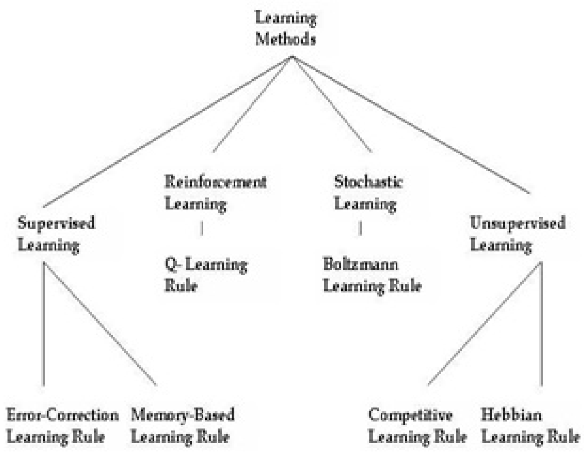

Figure-1. Learning rules of ANN.

2. CLASSIFICATION MODELS

A neural network classifies a given object presented to it according to the output activation. For binary outputs, 1 corresponds to one class and 0 corresponds to the other. For continuous outputs between 0 and 1, 0.5 can be used as the threshold to make the decision. In the case of multiple outputs, the object is assigned to the class corresponding to the output node with the maximum activation.

a) Single layer perceptrons

A single layer perceptron network is the oldest one, which consists of a single layer outputs via a series of weights. And therefore it is considered the simplest kind of feed-forward network. The sum of the products of the weights and the inputs is calculated in each node, and the resultant is a value above the threshold (typically 0) the neuron fires and takes the activated value (typically 1); otherwise it assumes the deactived value (typically -1). Neurons of the previous kind are also called McCulloch-Pitts neurons or threshold neurons. In the literature the term perceptron often refers to networks consisting of just one of these units. Warren McCulloach and Walter Pitts have already explained them in the 1940s. a simple learning algorithm that is usually called the delta rule can train Perceptron. It calculates the errors between calculated output and sample output data, and uses this to create an adjustment to the weights, thus implementing a form of gradient descent. Single layer perceptrons are only capable of learning linearly separable patterns; i.e., it can classify the input as to belonging to class A or class B, [5] unfortunately, many classification problems are not linearly separable. Multilayer Perceptron is a solution to problems that are nonlinearly separable. Fabio Roli and G N Marcialis used a single layer perceptron with a class separation loss function for identifying individuals based on their fingerprints [6].

b) Multilayer perceptron

Multi-layer Perceptron (MLPs) employs a number of learning techniques, especially back-propagation. Here the output values are compared with the correct answer to compute the value of some predefined error-function. The error is then fed back through the network by the means of certain techniques and accordingly the algorithm adjusts the weights of each connection so as to reduce the value of the error function to minimize the amount of error. Once a repeatable number of training cycles are over, the network will usually converge to some state where the error of the calculation is small. In this case the network seems to have learnt a certain target function. For training the network, generalized delta rule is used. Delta rule is also called as Adaline Rule, Widrow-Hoff Rule, or least Mean Square s (LMS) Rule. The training is continued until the total squared error reaches the minimum. Real-world task like the recognition of hand-written numerals may be efficiently and economically accomplished by means of a general-purpose MLP [3].

c) Adaptive Resonance Theory (ART) networks

Carpenter and Grossberg developed ARTI.ARTI nets are designed to solve the stability-plasticity dilemma. I.e. the network should stable as well as be able to learn new patterns. The basic architecture involves three groups of neurons: an input processing field (called the F1 layer), the cluster units (called the F2 layer), and a mechanism to control the degree of similarity of patterns placed on the same cluster (a reset mechanism). The degree of similarity required for patterns to be assigned to the same cluster unit is controlled by a user specified parameter known as the vigilance parameter. The net is a dynamic system, but the process can be simplified because the activations are assumed to change much more rapidly than the weights. Once an acceptable cluster unit has been selected for learning, the bottom-up and top-down signals are maintained for an extend period, during which time the weight changes occur. This is the 'resonance' that gives the net its name. ART1 networks are widely applied in clustering, character recognition, and pattern classification. Engineering design is a knowledge intensive process. Chin-Bin Wang et al. developed a novel scheme for functional feature-based reference design retrival using ART1 neural network to provide engineering designers with easy access to revelant design and other knowledge. [4] ART2 is designed to perform for continuous valued input vectors the same type of tasks as ART1 does for binary valued input vectors.

3. ASSOCIATION MODELS

The brain exhibits a particular type of memory which is of 'associative kind'. Associative memory could remember patterns. For example, a person could recognize another person just by hearing his / her name. Also, if somebody watches a scene from a movie on the television he or she could associate it to the movie from which it is taken even if he had watched the movie years ago. The method of recognizing is as follows: there will definitely be'underlying' information in a picture which lies in the

ARPN Journal of Engineering and Applied Sciences

ways that picture elements are ordered. This information gets loaded in the memory, and when a part of the picture is submitted to the network, it is possible to 'associate' and restore the entire picture. (Auto association memory). Yet another example is searching a book in the library; it is enough to know a part of the title to finally get the right book.

a) Hopfield networks

Hopfield Networks are recurrent neural networks, which means in an undirected way the output is send to the input as a feedback. The network learns through the act of modifying the strength of the connections between the neurons and the thresholds i.e., parameter of the activation function. The outputs of each neuron are connected to the input of every other neuron. Learning takes place as follows: feed a desired state to the network; if this state isn't stable, the values of the thresholds and inter-neuron connection strengths (weights) are changed until the desired stability is reached. Hebb's learning rule and the Widrow-Hoff or Delta learning rule are the rules used for the purpose. The Hopfield network can be performed as robust content-addressable memory, resistant to connection alteration if hebbian rule is applied. Hopfield Network is used in pattern recognition primarily [1].

b) Bidirectional Associative Memory (BAM)

The architecture of BAM consists of two laters of neurons, connected by directional weighted connection paths. The net iterates, sending back and forth between two layers until all neuron's reach equilibrium. i.e., until each neurons activation remains constant for several steps. The weights are bidirectional and the algorithm alternates between updating the activations for each layer, the X-layer and the Y-layer. Signals are sent only from one layer to the other at any step of the process, not simultaneously in both directions. BAM is used for pattern association [7, 9].

4. OPTIMIZATION MODELS

The optimization model offers solution to various combinatorial optimization problems that often lack efficient solution on a digital computer.

a) Hopfield networks

Hopfield Networks has been discussed in detailed in the last section. It has been shown that Hopfield Networks gives a better solution for optimization problems.

b) Boltzmann machines

The main problem with the Hopfield network is that it settles in to local minima by constant energy minimization. This phenomenon is desirable for association but undesirable for optimization or constraint satisfaction. The Boltzmann machine is a stochastic version of the Hopfield model; whose network dynamics incorporate a random component in correspondence with a given finite temperature. Starting with a high temperature and gradually cooling down, allowing the network to reach equilibrium at any step, chances are good, that the network will settle in a global minimum of the corresponding energy function. This process is called simulated annealing. Hinton and Sejnowski (1986) combined Hopfield networks and simulated annealing to result in networks known as Boltzmann machines. The network is used to solve a well-known optimization problem: The weight matrix is chosen such that the global minimum of the energy function corresponds to a solution of a particular instance of the traveling salesman problem [10]. Boltzman machine can be used in character recognition [7].

5. SELF ORGANIZATION MODELS

The term self-organization refers to the ability to learn and organize information without being given correct answers for input patterns.

a) Linear Vector Quantization (LVQ) networks

An LVQ network firstly has a competitive layer and then a second linear layer. The competitive layer learns to classify input vectors. The linear layer transforms the competitive layer's classes learned by the competitive layer are referred to as subclasses and the classes of the linear layer as target classes. The working of the network can be explained as follows: The network has two layers namely the input neurons layer and output neurons layer. The network is given by prototypes W=(w(i),…,w(n)). It changes the weights of the network accordingly aiming to classify the data correctly. For each data point, the prototype (neuron) that is closest to it is determined (called the winner neuron). The weights of the connections to the identified neuron are adapted, i.e. made closer if it correctly classifies the data point or made less similar if it incorrectly classifies it. LVQ networks have been applied to the problem of character recognition. Baykal and Yalabik [1992] have utilized a Kohonen LVQ net in conjunction with a feed forward net in the recognition of multifont characters. The recognition rate was 87% even with distorted, shifted and rotated characters.

b) Kohonen networks

Kohonen's Self Organizing Maps (SOM) is a type of unsupervised learning. The goal is to discover some underlying structure of the data. It is called a topology-preserving map because there is a topological structure imposed on the nodes in the network. A topological map is simply a mapping that preserves neighborhood relations. SOMs take into consideration the physical arrangement of the nodes. Nodes that are "close" together are going to interact differently than nodes that are "far" apart. The goal is to train the net so that nearby outputs correspond to nearby inputs.

E.g. if x1 and x2 are two input vectors and t1 and t2 are the locations of the corresponding winning output nodes, then t1 and t2 should be close if x1 and x2 are

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

similar. A network that performs this kind of mapping is called a feature map.

SOM is based on competition networks. The most extreme form of competition among a group of neurons is called Winner Take All i.e. the neuron with the maximum net value is fired. In the brain, neurons tend to cluster in groups. The connections within the group are much greater than the connections with the neurons outside of the group. Kohonen's network tries to mimic this in a simple way. It is applied in many clustering problems. Angeniol et al. has applied Kohonen SOM to the solution of the well-known traveling salesman problem [2].

c) Hebb net

Donald Hebb created a learning rule in the year 1949 [16] that states that the association between two neurons is made stronger if the neurons fire simultaneously or within a time interval. The Hebbian learning rule specifies the level to which the weight between two neurons has to be increased or decreased in accordance to their activation. In this learning system, learning is a local phenomenon without global feedback from the environment. i.e., the learning occurs without a teacher. Hebbian learning is a category of unsupervised learning where the result of the output need not be considered i.e. it is a correlation-based type of learning. The single layer feed forward neural net trained using the Hebb rule is referred to as Gebb net. The Hebbian form of learning seems to be present in the brain, e.g. experiments have showed that LTP in certain hippo campal synapses is hebbian [7] [9]. The Hebb rule is used for pattern association. [10]

6. HYBRID MODELS

In hybrid network, each layer implements a certain network type and performs its function. The network type concerned determines the connection pattern in each layer.

a) Counter Propagation Networks (CPN)

Counter propagation networks [6, 7] are multilayer networks based on a combination of input, clustering, and output layers. Counter propagation nets can be used to compress data, to approximated function, or to associate patterns. Counter propagation net approximates its training input vector pairs by adaptively constructing a look-up table. In this manner, a large number of training data points can be compressed to a more manageable number of look-up table entries. Training a counter propagation network occurs in two phases. During the first phase the winning neuron is allowed to learn. This is standard kohonen learning. During the second phase, the weights from the winning cluster unit to the output units are adjusted using delta rule. This is known as Grossberg learning. Pater and Rodney have used CPN for star identification. [4] This network takes a feature vector input and gives an output of what it as and the probability of the classification being correct.

b) Hamming networks

A Hamming network is a type of two layer feed-forward network. It has the ability to categorize noise corrupted patterns. The first layer of neurons, quantifier subnet, performs in parallel the hamming distance of a m-bit digital input vector with n previously stored exemplar patterns. The second layer is committed to the selection of the winner neuron. The winner neuron is the one with smallest Hamming distance to the input vector-this is the discriminator subnet. Therefore the network performs efficient classification for moderately low complexity, and always converges to one of the previously stored combinations. The number of independent neurons (n) corresponds to the number of patterns to be sorted out, and the number of synapses (m) association with each neuron corresponds to the number of input vector components. Hamming Networks finds its application in pattern matching.

Table-1. Supervised vs unsupervised learning and application domains.

Problem domain	**Supervised learning network**
Classification, Identification, Diagnosis, Expert Systems and Decision Systems	Adaline, Perceptron, Multi Layer Perceptron(MLP), Time Delay Neural Network(TDNN), Learning Vector
Feature Extraction	Quantization (LVQ), Radial Basis Function (RBF) Multi Layer Perceptron (MLP), Learning Vector Quantization (LVQ)
Function approximation, Forecasting	Multi Layer Perceptron (MLP), Radial Basis Function (RBF), Simultaneous Recurrent Network, Recurrent Network with back propagation
Problem domain	**Unsupervised learning network**
Character recognition and pattern classification	Adaptive Resonance Theory(ART)
Classification and Optimization Problems	Self Organizing Maps(SOM)
Pattern Association	Hebb Net
Classification Problems	CPN (both supervised and unsupervised)
Pattern Recognition	Hopfield, BAM
Character Recognition	Boltzmann
Optimization	Hopfield

The above table summarizes the various networks and the application domain. It is understood from the table that the classification, Association and Self-organization models are predominantly used in the field of Pattern Recognition and the optimization model play a significant role in character recognition and in optimization and in optimization problems.

7. CONCLUSIONS

The neural computational models have been discussed with regard to their application in this paper. In classification model, the given input data is assigned to one of the finite number of categories whereas in association model the given input pattern is associated model the given input pattern is associated with pattern is associated with the pattern already available. The objective of self-organization model is to cluster the given input based on the available features of the input patterns. Hybrid models try to emulate the behavior of two or more networks, specific to the application domain.

REFERENCES

[1] Abhijit Pandya and Robert Macy. 1996. “Pattern recognition with neurwl networks in C++”, CRC press Inc. p. 43.

[2] Awodele and O. Jegede. 2009. “Neural Networks and Its Application in Engineering”, Proceedings of Informing Science & IT Education Conference (InSITE). pp. 83-95.

[3] Jin-Song Pei, E. Mai and K. Piyawat. 2006. “Multilayer Feedforward Neural Network Initialization Methodology for Modeling Nonlinear Restoring Forces and Beyond”, 4th World Conference on Structural Control and Monitoring. pp. 1-8.

[4] Laurene Fausett. 1994. “Fundamentals of Neural Network Architectures, Algorithms and Applications”, Prentice Hal, New Jersey. USA. pp. 48-49.

[5] L. Fu. 2003. Neural Networks in Computer Intelligence, Tata McGraw-Hill.

[6] Marcialis G. and Roli F. 2005. “Fusion of multiple finger print matchers by singlelayer perceptron with class seperation loss function”, Pattern Recognition, Letters. 26: 1830-1839.

[7] Moghadassi, F. Parvizian and S. Hosseini. 2009. “A New Approach Based on Artificial Neural Networks for Prediction of High Pressure Vapor-liquid Equilibrium”, Australian Journal of Basic and Applied Sciences. Vol. 3, No. 3, pp. 1851-1862.

[8] R. Rojas. 1996. The Backpropagation Algorithm, Chapter 7: Neural Networks, Springer-Verlag, Berlin. pp. 151-184.

[9] R Sathya and Annaamma Abraham. 2013. “Comparision of Suervised and Unsupervised Learning Algorithms for Pattern Classification”, International Journal of Advanced Research in Artificial Intelligence. Vol. 2, No. 2.

[10] S. Haykin. 2005. Neural Networks- A Comprehensive Foundation, 2nd ed., Pearson Prentice Hall.

[11] T. Kohonen and O. Simula. 1996. “Engineering Applications of the Self-Organizing Map”, Proceeding of the IEEE. Vol. 84, No. 10, pp.1354 – 1384.